Monkeypox Disease Detection using Deep Learning Techniques

MSc Research Project
Data Analytics

Arpita Mitra
Student ID: x21116211

School of Computing
National College of Ireland

Supervisor: Bharat Agarwal

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Arpita Mitra
Student ID:	x21116211
Programme:	Data Analytics
Year:	2022-2023
Module:	MSc Research Project
Supervisor:	Bharat Agarwal
Submission Due Date:	15/12/2022
Project Title:	Monkeypox Disease Detection using Deep Learning Techniques
Word Count:	5980
Page Count:	20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	31st January 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Monkeypox Disease Detection using Deep Learning Techniques

Arpita Mitra
x21116211

Abstract

Following a pandemic like Covid-19, from which the world is still trying to recover, another unexpected epidemic of a new contagious infection called Monkeypox has captivated everybody's attention. It is the first example of a more wide population spreading of Monkeypox disease. The number of infected cases is still increasing even though it is less contagious and dangerous unlike Covid-19. Hence, developing a precautionary strategy is crucial for a brighter future. This resembles with other skin diseases in certain ways, but due to a lack of proper identification of the disease, it can be fatal in some situations. AI-based identification would be beneficial in this scenario as spotting Monkeypox at a very early stage will become easier with its help. Convolutional Neural Network (CNN) technology and an approach to use VGG-19 models with model's bottleneck features have been introduced in this study with fine-tuning the model by changing its custom layers to check and improve the performance and accuracy by lowering the complexities of the image classification procedure. A publicly accessible dataset has been employed for the research. In accordance with the model's acquired Accuracy, F1 Score, Precision, and Recall Value, the findings that were acquired using the recommended approach were confirmed and evaluated on various images of Monkeypox and non-Monkeypox to determine their correctness and efficacy. The achieved accuracy of 91.87% from VGG-19 model clearly states that this model will help in detecting Monkeypox at a wider range.

1 Introduction

While the entire world was rocked by the emergence of Covid-19 in 2020, cases of the advent of Monkeypox in 2022 revealed that yet another widespread virus is on the rise. This contagious illness is brought on by the Monkeypox virus, which belongs to the family of Zoonotic Orthopoxviruses. Humans may be a vector for the propagation of this virus. The number of reported cases of Monkeypox continues to rise, irrespective of the fact that the disease is now less contagious and has a lower risk of fatalities than Covid-19 has Ahsan et al. (2022). The Monkeypox epidemic which took place in the year 2022 affected a lot more people than the one which occurred in the USA in the year 2003. It's been almost 20 years since it happened for the first time in 2003, despite this, there is currently no medication available for it. Thus, in order to create effective medication, it is necessary to correctly identify and differentiate this virus from comparable but less dangerous skin illnesses. Furthermore, the use of image processing and evaluation supported by Artificial Intelligence (AI) might aid in the detection of the infection.

1.1 Research Motivation

Traditionally, it was believed that machine learning (ML) techniques would serve as the backbone for data mining and classification automation. In the field of medical science, several AI based models have been constructed using image analysis for the identification of different viruses, reflecting the large increase in the use of AI models in a wide range of fields. The enhanced learning potential of deep learning (DL) applications, especially variants of Convolutional Neural Networks (CNNs), has recently transformed several fields of medical research. With its capacity to build reliable prediction models from massive data, transfer learning has become more popular. Prior to Ahsan et al. (2022) breakthrough, very little study had been conducted on Monkeypox viral detection and their efforts and findings in this field of study were likewise quite motivating. However, other studies on other medical problems, such as brain tumor identification, Covid-19, and skin cancer detection, have already been conducted utilizing deep learning approaches and the positive outcomes in these areas serve as inspiration for further research and development in the research of Monkeypox viral detection.

1.2 Research Question

"How efficiently can Monkeypox disease be identified accurately by Image Classification using Deep Learning technique?"

1.3 Research Objectives

- The research work aims at providing a novel approach to detect and classify the images of Monkeypox from the other similar skin diseases using pre-trained model VGG-19 along with base model CNN. When compared to training from scratch, adopting a pre-trained model can save significant amounts of time and effort.
- Examine the related works and find current approaches that may be used as criteria for the suggested design and evaluation measures.
- Analyze the outcomes of the related works and the recommended methodology.
- Evaluate if the suggested approach yields better outcomes than prior research.

2 Related Work

Researchers have begun to pay attention to a number of illnesses in recent years, including Covid-19, other skin conditions and tumor-related diseases. Image processing is gaining popularity in medical and other scientific areas in recent years. Various algorithms and methodologies have been used in numerous research to categorize medical imaging data. Numerous research in the field of medicine have demonstrated that diseases like Covid-19, brain tumors, skin cancer, and many others can be recognized and diagnosed by Image segmentation. Therefore, in order to differentiate Monkeypox sickness from other skin diseases of a similar nature, such as chicken pox and measles, it must also be correctly identified.

The most recent and significant study is conducted by Ahsan et al. (2022). The lack of patient images that have been contaminated by the Monkeypox virus was the impetus for this research, and its primary objective was to find a solution to the data shortfall that has persisted for some time. This research paper had already incorporated an algorithm that was based on deep learning in order to diagnose the condition. Web mining techniques were used to construct the dataset, which consists of several photos of chickenpox, Monkeypox, measles, and normal skin. The usage of these techniques was done in the hopes of achieving more accurate results. They then implemented a transfer learning method using the VGG-16 model while taking into account two strategies. Monkeypox and chicken pox were the two virus groups that were taken into consideration for the initial approach while the second approach involved enhancing the images by data augmentation technique. They mentioned in their research that their accuracy was 97% when they categorized the images without using the data augmentation technique. On the other hand, when the data augmentation technique was introduced to the dataset, the accuracy level of the model was reduced to 78%.

For the purpose of detecting the Monkeypox virus, Sitaula and Shahi (2022) discussed and contrasted a total of 13 pre-trained deep learning models. The goal of this research was to find a model with the highest accuracy which will be able to detect Monkeypox virus correctly. They analyzed the outcomes of these 13 pre-trained deep learning models by incorporating common customized layers to all of them. They utilized the dataset from the study publication Ahsan et al. (2022), nevertheless, they chose the augmented photos to conduct their study and trained their 13 models. It included the ensemble method of these pre-trained DL models, including ResNet, Inception, Xception, MobileNet, VGG and DenseNet. They conducted their investigation using a 5-fold cross-validation method. In their research, they were able to achieve the highest of 87.13% accuracy from the ensemble approach among all models. A relatively smaller dataset had been used in their study which was one of the limitations as the pre-trained models work well with large amount of data that actually helps in models' performance. Additionally, using compact deep learning models would be preferable if the system has memory limitations.

In addition to other pre-trained CNN architectures like ResNet 50, Inception V3, and VGG-16, Ali et al. (2022) also employed the ensemble technique like Sitaula and Shahi (2022). Because of transfer learning, these models have shown good categorization results in a variety of AI based applications, especially the analysis of medical pictures. To maintain consistency and improved adaptation, the experiment was done by retaining various numbers of easily trained bottom layers. Threefold cross-validation was used, with the training, validation, and testing sets divided in a ratio of 70:10:20. Following significant testing of the pre-trained models, the lowest eight layers of the models were unfreeze and for binary classification of the virus, a softmax activation mechanism was applied with 2 nodes of fully connected layer. This study used a relatively smaller dataset, similar to other studies in the field. Nevertheless, with a smaller dataset, the ResNet50 model surpassed the ensemble model, achieving the greatest accuracy of 82.96%.

Arbane et al. (2021) proposed a technique for automatically detecting brain tumors using MRI scan images. A transfer learning based approach was implemented to differentiate between MRI images of tumorous and non-tumorous tissue using three Convolutional Neural Network (CNN) frameworks such as Xception, ResNet-50 and MobileNet V2. With a total of 1516 tumor and non-tumor images, the final outcome demonstrated that a classification model had been created using pre-trained algorithms that could successfully recognize the tumor from MRI images. The MobileNet V2 model was superior than that of the other two models in terms of F1-score, precision and accuracy.

The paper Seetha and Raja (2018) developed a deep CNN-based automated method for the diagnosis of brain tumors. For the initial step of brain image segmentation, the Fuzzy C-Means (FCM) method was used. This was done in conjunction with SVM and DNN-based categorization, which were also performed. In order to achieve a high level of accuracy using their suggested model, a loss function based on gradient descent was implemented. The accuracy of detection was then brought up to 97.5% by the utilization of these integrated models.

An innovative method for detecting and classifying skin cancer was described by Ansari and Sarode (2017). They used GLCM (Gray-Level Co-Occurrence Matrix) for feature extraction from data and then SVM (Support Vector Machine) was used for classification. This study focused on the preprocessing of images in 3 tiers, starting with the elimination of noise, continuing with image augmentation, and finally transitioning to the grayscale. In this study, necessary characteristics for further categorization were retrieved. These characteristics are Mean, Consistency, Vitality, and Contrast. Maximum thresholding is a technique that is used to classify images. This technique was utilized in order to derive features from an image, which required the foreground and background to be separated from one another. The SVM model achieved an accuracy of approximately 95% to classify photos of malignant lesions. Despite the fact that the impacts of a number of kernels, including Gaussian and polynomial within SVM was not considered prior to the writing of this study Ansari and Sarode (2017).

Sitaula and Hossain (2021) had completed their research on the chest x-ray picture classification of patients infected with Covid-19 virus. They proposed a deep learning model by combining VGG-16 and the attention module. They were able to fine-tune the classification method by designing a model that made use of the 4th pooling layer of VGG-16. Their investigation was carried out using three different datasets, each of which had chest x-ray pictures from Covid-19 patients. They compared the findings of VGG-16 with those of a number of other pre-trained models, such as VGG-19, EfficientNet, and so on; the results showed that VGG-16 had outperformed each of the other models with an accuracy of 87.49%. In a similar manner, another research effort was carried out on the classification of chest x-ray images of Covid-19 patients from different infections. For the purpose of detecting the Covid-19 virus, Madhavan et al. (2021) created a DL model called Res-CovNet that was based on the method of transfer learning. Their investigation was conducted on a database that had around 5856 images in total. They used ResNet-50 He et al. (2016) to extricate the characteristics from several X-ray images, and then they added a classification layer to the network to make it more comprehensive. Their approach was successful in achieving a remarkable accuracy of 98.4% in detecting Covid-19 versus the normal cases and 96.2% in detecting all other cases versus Covid-19 cases on X-ray pictures.

The research Sarumi (2020) also included the proposal of an ensemble learning strategy as a method for applying big data to the detection of diseases caused by the Ebola virus. In order to derive insights from the massive amounts of data utilizing Kafka and Apache Spark as a foundation, they used a combination of genetic algorithms and ANN (Artificial Neural Networks).

In this study Gouda et al. (2022), the deep learning technique CNN was utilized to distinguish between malignant and benign tumors. A total number of 3533 images were used. The images were categorized using a CNN approach, with the classification being based on a composite of the findings received after a large number of repetitions followed by fine-tuning the model using Resnet50, InceptionV3, and Inception Resnet. The suggested research achieved the highest total accuracy of 85.7% with Inception model though it had two drawbacks. First, it was done on a relatively smaller dataset and second is that it was possible to use a number of additional advanced DL models in order to gain improved performance and accuracy. In contrast to this, Yadav et al. (2022) had built a facial skin illness diagnosis system that was based on an HSV model segmentation and used deep learning. The purpose of this study (Nadeem et al.; 2020) was to investigate significant deep learning concerns pertaining to brain tumor studies by making use of the extensive toolkit offered by deep learning. This paper offered a straightforward overview of a variety of scholarly accomplishments that had been made in the area.

Another deep learning approach was implemented by Nasr-Esfahani et al. (2016) with medical images. In order to improve the precision of their method, they first divided the images as a part of the pre-processing stage, then a correlation was employed for the illumination of the images. The photos that had been improved and segmented were then transmitted to CNN to have their features extracted and classified which provided an accuracy of 81%.

It has been observed, based on all of the previous studies, that there have been a very little amount of effort done on the subject of Monkeypox detection however it has also been noticed that neural networks have become much more significant in the diagnosis of different disorders in the medical field. In Ahsan et al. (2022) and Ali et al. (2022) the majority of the study effort focused on binary classification, but their approach of using the pre-trained models to identify the disease did not perform well in terms of accuracy and precision. Moreover, the models did not provide a meaningful analysis. As a result, the primary focus of this research is on developing a DL model that is capable of diagnosing the condition with a higher degree of accuracy.

3 Methodology

The goal of this research is to develop an automated method that is capable of assisting medical professionals in distinguishing the skin illness known as Monkeypox from other skin disorders. Despite the fact that it has certain symptoms similar to chicken pox and other skin illnesses, it is far more dangerous and has a higher mortality rate than other skin conditions. As a result of its novelty and the lack of effective treatments, the number of reported deaths from Monkeypox continues to rise. Implementing an automation process might be a transformation since it is necessary to have a suitable system to identify the disease in order to avoid this from happening. This would allow for an initial assessment and treatment to be initiated. According to research that is relevant to this topic, deep learning is a highly successful module for detecting photos of a variety of medical conditions, including the identification of Covid-19 using X-ray images, the detection of brain tumors, the detection of skin cancer, and so on.

A well-designed procedure is required to execute the strategy that has been proposed for this research. The following figure 1 highlights all the steps that were taken while conducting this research:

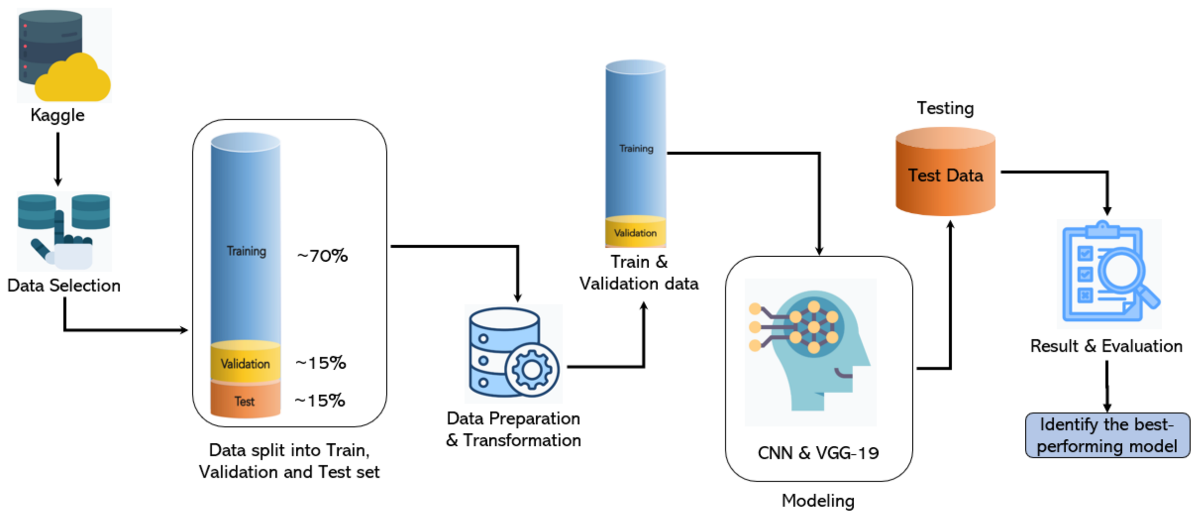

Figure 1: Process Flow of the Research

3.1 Data Collection & Description

The objective of this study is to classify photographs that include evidence of the Monkeypox virus's presence. In order to conduct this research, a dataset containing various images of Monkeypox as well as non-Monkeypox diseases (such as chicken pox and measles) have been considered. The authors Ali et al. (2022) have made the dataset [1] readily accessible to the general public for the purpose of scientific future investigations. The dataset is organized into three folders: Original Images, Augmented Images, and Fold 1. There are a total of 228 photographs in the original Images folder; out of which 102 are classified as 'Monkeypox' while the rest 126 are classified as 'Others' which includes instances of non-Monkeypox diseases like chickenpox and measles. Numerous data augmentations had already been done to the original dataset and the augmented images were saved in 'Augmented Image' folder because there were less number of images to work with and to aid the work of categorization. 'Monkeypox' and 'Others' are the two groups that make up the dataset's structural divisions. In order to complete this study, this 'Augmented Images' folder will be used as Deep learning models can better comprehend the image's context when augmented images are used to train the model. Some of the data samples for Monkeypox and Non-Monkeypox diseases have been displayed in figure 2.

3.1.1 Ethical Concerns

The use of the dataset for the purpose of this investigation does not raise any ethical difficulties as it has already been made available to the public.

3.2 Data Preprocessing

After deciding on a suitable dataset, it is necessary to clean and modify the data so that it is fully consistent with the network model that is constructed, hence maximizing

[1] `https://www.kaggle.com/datasets/nafin59/monkeypox-skin-lesion-dataset?datasetId=2308447`

Figure 2: Images of Monkeypox & Non-Monkeypox skin lesions

its performance. A lot of information is needed for deep learning models to provide reliable results. On the other hand, there could not be adequate data in other instances. The process of acquiring and interpreting data, especially when it pertains to medical concerns, is one that is both time-taking and expensive. One of them is the enhancement of data, which may help deep learning techniques achieve higher levels of precision Ayan and Ünver (2019). Despite the fact that enhanced photos are used in this research, further augmentation has been performed on such images in order to boost the model's ability to interpret data and prevent overfitting of the data. Every image is annotated and separated into a training set, validation set, and test set. In order to increase the performance of the model, certain challenges are included, such as image flipping, and including fuzzy and unsharpened versions of those images.

3.3 Model Selection

In contrast to Artificial Neural Networks(ANNs), CNNs have been shown to achieve superior results in the aforementioned literature study and past research studies. This has been used extensively in the field of medical image processing. The development of a model that may assist medical research in the more accurate diagnosis of illnesses via the use of AI has been attempted on several occasions by a large number of experts. In this study, an effort was made to construct a model that is capable of dependable detection of the Monkeypox virus. In this study, the CNN base model in conjunction with an advanced CNN model such as VGG-19 Xiao et al. (2020) have been used. With its pre-trained layers, VGG-19 is an advanced Convolutional neural network. This model is quite intelligent since it was pre-trained to classify millions of photos of varying types and complexity.

3.4 Evaluation

Accuracy, Precision, Recall and F1 Score are the evaluation metrics that will be considered to evaluate the proposed models. Also, the losses from each model will be taken into account while studying the model's performance throughout the evaluation process. The metrics will be formed by the use of the confusion matrix convention. Based on the information contained in the aforesaid matrix, specific metrics criteria will be used to evaluate the effectiveness of these models. Numbering the True Positive (TP), True Negative (TN), False Negative (FN), and False Positive (FP) results of a classifier that yields the confusion matrix. If the Monkeypox virus can be substantially categorized as TP and TN, then an appropriate diagnosis of the disease will be made and will receive an incorrect diagnosis if it is largely categorized as FP or FN.

Accuracy: The accuracy of a prediction system is measured as the percentage of correct predictions relative to the number of training instances. The formula, which is used to calculate Accuracy is :

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision: This term refers to the proportion of correct positive results relative to the sum of true positive and false positive results. The mathematical expression for precision is:

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

Recall: The proportion of items correctly classified as positive, or recall, is calculated as follows: recall = true positives / total number of items correctly classified as positive. The mathematical expression of Recall is :

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

F1-Score: Harmonic average of Precision and Recall yields the F1 score. The greater the F1 score, the higher it should be. A score on the F1 test that is much closer to one demonstrates both excellent recall and precision. The mathematical expression for F1 Score is:

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

4 Design Specification

4.1 CNN(Convolutional Neural Network)

CNN has been used extensively in the field of medical image processing. The development of a model that may assist medical research in the more accurate diagnosis of illnesses via the use of AI has been attempted on several occasions by a large number of experts. High-quality results in both video and image identification tasks have been achieved by Convolutional Networks. The accessibility of public-image archives is a contributing factor in the development of these characteristics.

CNN models are utilized for the identification of Monkeypox, and these models have features like max-pooling and dense layers. The CNN architecture that is employed in this research is presented in figure 3.

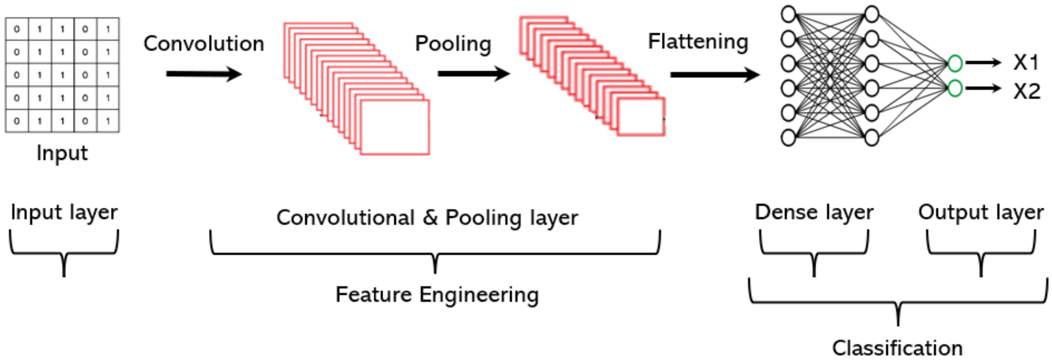

Figure 3: Architecture of CNN

The application of the pooling layer results in the generation of a max pooling layer. This research utilized a three-layer Conv2D architecture for its base model CNN, along with two dense layer configurations. In order to lessen the complexity of the picture, maxpooling was conducted in between the layers. When using pooling, the network's variables and processing load are decreased by compressing the spatial size. The flattening layers come in right after the pooling layers since they are responsible for transforming the incoming picture matrix into a matrix with one column. Therefore, the model makes use of a flattening layer in order to transform three-dimensional feature mappings into one-dimensional feature units. Two dense layers have been used. The dense technique is a useful tool for performing network analysis in Keras. In addition to this, one dropout layer is included in the last dense layers in order to include a classifier in the base of the Convolutional model. ReLu was implemented as an activation function because it makes the model more statistically efficient and has excellent performance. It is implemented in the dense layer. Because of the requirement to lessen the number of computing resources, the last dense layer's activation function is decided to be a Sigmoid function. This decision was made because of the necessity to conserve more time.

4.2 VGG-19

VGG-19 is a pre-trained model which has been used in this research. This model consists of 19 layers which include 3 fully connected layers along with 16 Convolutional layers. This model is quite intelligent since it was trained to classify millions of photos of varying types and complexity. To carry out the research, as it is a pre-trained model so, further training was not done on the model, however, to complete the categorization job, some layers of this model were frozen and a shallow 2 layer was constructed on this model. Here total 5 separate VGG-19 models have been constructed and compared by fine-tuning the models. Also, bottleneck Song et al. (2015) features have been implemented in this study. The presence of this layer enables the compression of extracted features inside the network. The characteristics known as bottlenecks are those that appear on the maps of activation that come prior to the fully-connected final layers of the model. Table 1 displays all the details of VGG-19 models that have been used while performing this experiment.

Model	Architecture
VGG19-Model 1	2 Dense Layer
VGG19-Model 2	3 Dense Layer+1 Dropout
VGG19-Model 3	3 Dense Layer+1 Dropout+1e-2 LR
VGG19-Model 4	3 Dense Layer+1 Dropout+1 Adam
VGG19-Model 5	2 Dense Layer+1 Dropout

Table 1: Types of VGG-19 models used in this study

5 Implementation

This section provides an explanation of how correctly the deep learning models are able to diagnose the Monkeypox illness. During the course of this research, one standard CNN model and five custom-built VGG-19 models were utilized to compare the results.

5.1 Technologies & Tools used

The below figure 4 depicts the tools, technologies and libraries used in this study.

Figure 4: Tools, Technologies and Libraries used in this study

Below Table 2 shows the specification of hardware that has been used in the study.

Hardware used	Specification
Processor	Intel(R) Core(TM) i7 8th generation
RAM	16 GB
Hard Drive	1 TB
Operating System	Windows-10 64 bit

Table 2: Hardware Specification

5.2 Data Processing & Transformation

The dataset is initially imported into Jupyter notebook from a local drive. It has an 'Augmented Image' folder that contains subfolders of augmented Monkeypox and non-Monkeypox images. The source of data is reliable and the acquired images are of the highest quality. Each of the images within its associated subfolders is shuffled, and then randomly split, before being saved within the training, validation, and testing folders of their respective subfolders. 70% of the total images have been stored as a training set and have been used to train the model and adjust the attributes of a classifier. Half of the remaining images have been used as a validation set in order to tweak the attributes, model's performance has been tested using the remaining half. Following that, additional blurry and unsharp images of Monkeypox and non-Monkeypox are incorporated into the training set in order to enhance the performance of the model. Python code written in Jupyter notebooks has been used to successfully carry out all of the necessary tasks of data cleansing, transformation, and handling. Then each image is scaled down to fit the input size of 228 by 228 pixels so that the model can be more generic, effective, and precise. Additionally, in order to increase the complexity of the base model CNN, several hyperparameters like zooming, rescaling, flipping the images horizontally, shear range have been used during data augmentation. Overfitting of the model has been prevented by using Early stopping. The number of epochs for the same has been set to 5 to optimize the model performance. The model's performance is sensitive to its setup, making it challenging to establish optimal settings for the model as well as the training procedure. The entire process of cleansing, transformation, and handling were performed using python in Jupyter notebook.

5.3 Transfer learning with Pre-Trained Model

Advanced CNN with pre-trained layer VGG-19 has been used in the second phase. In order to assess the outcome of the detection of the Monkeypox photos, five different sets of VGG-19 models have been implemented, fine-tuned, and evaluated to one another. In order to accomplish this goal, a method has been carried out that involves freezing the main layer of the model in order to obtain the features and labels of each image that comes from the training set, as well as the validation set, and the testing set. As a result of this, the model will be able to implement a transfer learning technique using its previously trained images. Additionally, the bottleneck feature of the model, which was obtained through the extraction, was saved and has been used as the model's last layer of classification. During the course of this research, VGG-19 model that had been pre-trained using the ImageNet dataset has been used. The architecture of VGG-19 models used in this study, is given in the table Table 2. These networks make use of the complete image that is sent as input and provide a probable outcome for each component of the image.

6 Evaluation

Efficiency measurements of each implemented model in this research have been discussed in the section. This section also conducted a comparison study of the outcomes obtained by each model architecture (CNN and pre-trained VGG19) on the basis of evaluation criteria such as accuracy, loss, precision, recall, and F1 score. During the training phase, each model's accuracy and loss have been determined which helps in calculating the accuracy and misclassification of images in the testing phase as a percentage of how much tested data is predicted accurately.

6.1 Case Study 1 : Experiment with CNN

The baseline CNN model consists of three Convolutional 2D layers, 3 maxpooling layers, one flattening layer, two dense layers and one dropout layer. The model alternates between maxpooling and Conv2D layers in order to optimize the performance and minimize the model's runtime. Also to increase the model efficiency the epoch limit has been set to 30 and the batch size has been set at 50 which determines the number of images that will be sent over the network. It has used 6699 images for training and 479 images for validation. Early stopping value has been set to 5 epochs in this model which indicates that the model will preserve the best value at the time and cease seeking for better ones if validation accuracy does not increase after 5 iterations of the training process. In order to make the model much more efficient in terms of computation, ReLu activation function has been used in this study. It has been observed that the training and validation accuracy increased every time the number of epochs was increased. The highest accuracy and lowest loss were observed at epoch 29 with 92.66% of training accuracy and 89.78% of validation accuracy respectively. The validation loss has been reduced from 0.64 to 0.28 whereas the training loss has been reduced from 0.7 to 0.18. Finally, the CNN model has acquired 90% of testing accuracy with 24.13% of loss. Figure 5 and figure 6 depict the model summary and the training curve of the model respectively.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 226, 226, 32)      896

 activation (Activation)     (None, 226, 226, 32)      0

 max_pooling2d (MaxPooling2D  (None, 113, 113, 32)     0
 )

 conv2d_1 (Conv2D)           (None, 111, 111, 32)      9248

 activation_1 (Activation)   (None, 111, 111, 32)      0

 max_pooling2d_1 (MaxPooling  (None, 55, 55, 32)       0
 2D)

 conv2d_2 (Conv2D)           (None, 53, 53, 64)        18496

 activation_2 (Activation)   (None, 53, 53, 64)        0

 max_pooling2d_2 (MaxPooling  (None, 26, 26, 64)       0
 2D)

 flatten (Flatten)           (None, 43264)             0

 dense (Dense)               (None, 64)                2768960

 activation_3 (Activation)   (None, 64)                0

 dropout (Dropout)           (None, 64)                0

 dense_1 (Dense)             (None, 1)                 65

 activation_4 (Activation)   (None, 1)                 0

=================================================================
Total params: 2,797,665
Trainable params: 2,797,665
Non-trainable params: 0
_________________________________________________________________
None
```

Figure 5: Summary of CNN model

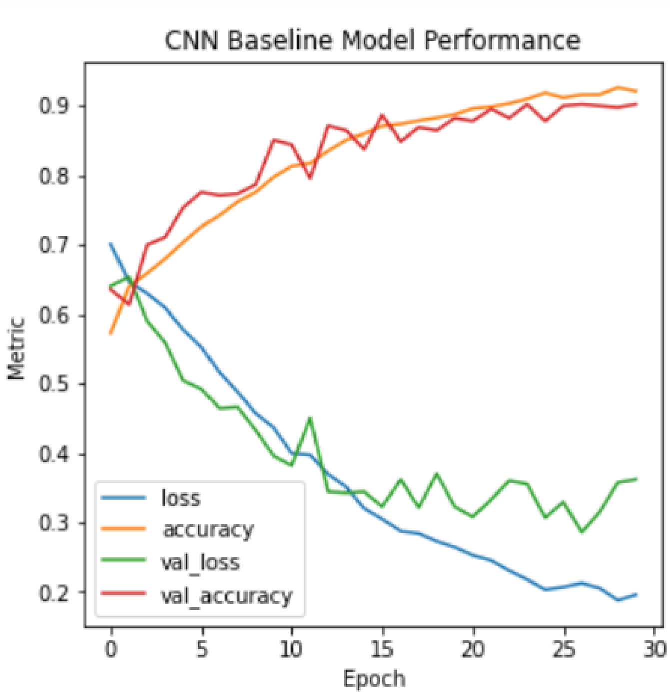

Figure 6: Learning curve of CNN Model

In order to check and improve the baseline CNN model, an advanced pre-trained model VGG-19 has been used. To carry out the research further, the bottleneck feature of the main VGG-19 model was extracted and appended with a classifier that includes the last dense layer. After that several others VGG-19 models were constructed by changing the layers of the main model.

6.2 Case Study 2 : VGG19-(Model 1)

In the first VGG-19 model, a classifier has been implemented on top of the VGG-19 model as well as two dense layers have been added. Due to the fact that not all of the model's neurons would be active at the same time, Relu activation function is employed in the model over others. Also, hyperparameters like learning rate have been adjusted to 0.0001 with a total of 30 epochs and 50 batch sizes. The following curve represents the model's training performance. With the highest 95.89% of training accuracy and 92.07% of validation accuracy at epoch 30, the model has encountered loss from 0.6 to 0.09 which has been highlighted in figure 7. In the end, a testing accuracy of 91.87% was achieved with a loss of 22.98%.

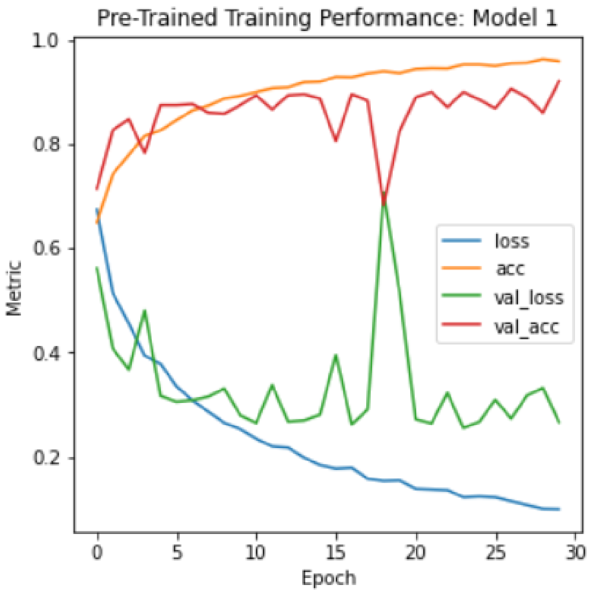

Figure 7: Learning curve of VGG19 - Model1

6.3 Case Study 3 : VGG19-(Model 2)

The second VGG-19 model is deeper than the previous one. It uses three dense layers, and one dropout layer has been added with the intention of getting better accuracy results. Extracted bottleneck features of the main model are loaded with training and validation set in this model. Like the previous model, ReLu function has been used with a learning rate of 0.0001, epoch limit 30 and batch size 50. This model acquired 95.70% of training accuracy with 91.65% of validation accuracy, where the loss has been reduced from 0.61 to 0.10. All of this information has been highlighted in figure 8. This model's performance is almost similar to the previous version of VGG-19 model, but in terms of accuracy, this model cannot exceed the previous model. Additionally, it has been observed that the accuracy of the model was not improved by the addition of a dropout layer. Post-testing, an accuracy of 90% was achieved with 30.39% loss.

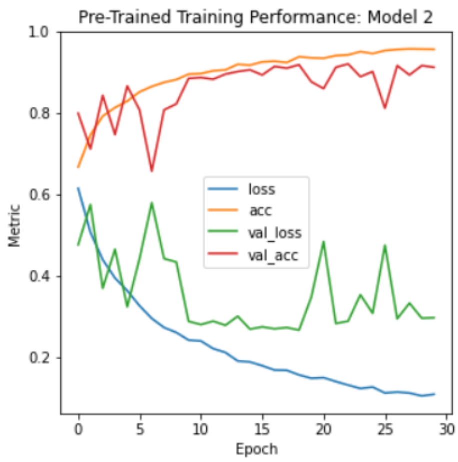

Figure 8: Learning curve of VGG19 - Model2

6.4 Case Study 4 : VGG19-(Model 3)

The next model consists of the same layers as the previous one, but with a different learning rate. In this case, learning rate has been increased to check the performance of this model and it has been observed that when the learning rate was increased to 0.01, the model's performance was really bad as it flattened down every single one of the learning curves as shown in figure 9. Following the test, an accuracy of 89.17% along with a loss of 32.60% was obtained.

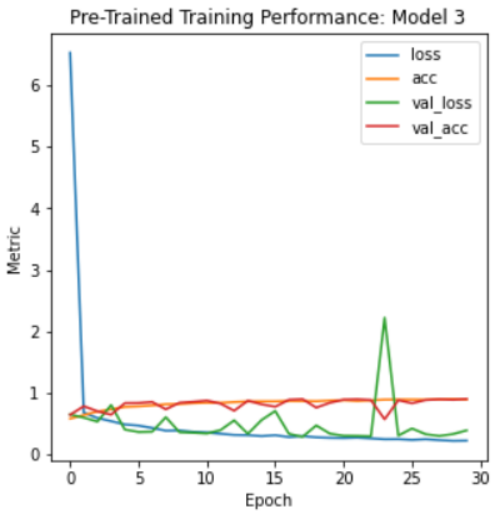

Figure 9: Learning curve of VGG19 - Model3 with increase learning rate

6.5 Case Study 5 : VGG19-(Model 4)

A much higher validation loss as compared to the training loss was observed when VGG-19 model was applied with Adam optimizer. A training accuracy of 94.82% was acquired along with a validation accuracy of 90.40%. The learning curve of this model has been laid out in below figure 10.

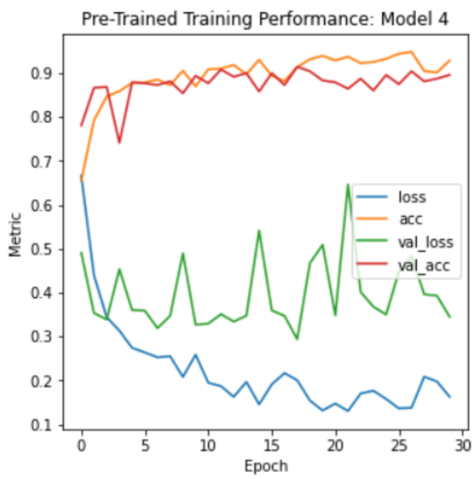

Figure 10: Learning curve of VGG19 - Model4 with Adam Optimizer

Adam was found to be a better alternative than Relu for model optimization as mentioned by Kingma and Ba (2014), however in this case accuracy is not as good as first VGG-19 model. After completing the test, it was determined that the accuracy was 90%, while the loss was also 36.45%.

6.6 Case Study 6 : VGG19-(Model 5)

The last VGG-19 model consists of two dense layers and one dropout layer with a learning rate set at 0.000005. This model was implemented with 30 epochs and 50 batch sizes. It was trained well, as shown by a reduction in loss from 0.64 to 0.26. It was observed that this model had acquired the highest 90.57% of training accuracy as shown in below figure 11, but it still was not able to beat the training accuracy of 95.89% which was achieved from the first VGG-19 model. Following the completion of the test, it was found that the accuracy was 89.79%, while the loss was 29%.

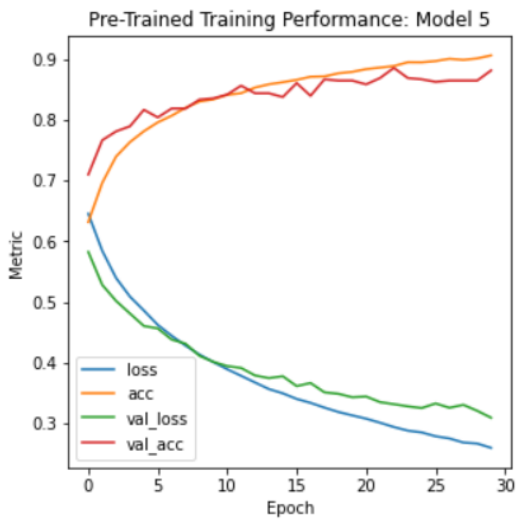

Figure 11: Learning curve of VGG19 - Model5 with shallow network

The table 3 that follows presents a comprehensive summary of the results of all models with regard to the accuracy and losses of test data.

Model Names	Accuracy(%)	Loss(%)
CNN	90.00	24.13
VGG19 Model 1	91.87	22.98
VGG19 Model 2	90.00	30.39
VGG19 Model 3	89.17	32.60
VGG19 Model 4	90.00	36.45
VGG19 Model 5	89.79	29.00

Table 3: Modelwise Accuracy & Loss

6.7 Discussion

Neural network has been used by deep learning technique in order to learn feature representation from the data. It can acquire proper accuracy and sometimes outperform human performance. Experiments were done using a variety of neural network models in order to construct an image classifier, where CNN turned out to be the top performer. The success of this research is shown by an improvement in the value of accuracy and precision when distinguishing the Monkeypox disease from other similar diseases. The base CNN model, along with all of the other pre-trained VGG-19 models that were used in this research, has reached a satisfactory level of accuracy. Every other models were outperformed in terms of accuracy, loss, precision, recall and F1 score by the first VGG-19 model. An accuracy of 91.87% along with a loss of 22.98% and 8% of misclassification of images were obtained by first VGG19 model. The confusion matrix of this model has been displayed in figure 12. This model has acquired higher accuracy than all other models because it has identified an appropriate combination of weights to enhance the performance of the model in identifying the target variable accurately, while the other models with custom layers were merely creating noise. Hence the first VGG-19 model can be preferred over every other model for Monkeypox detection.

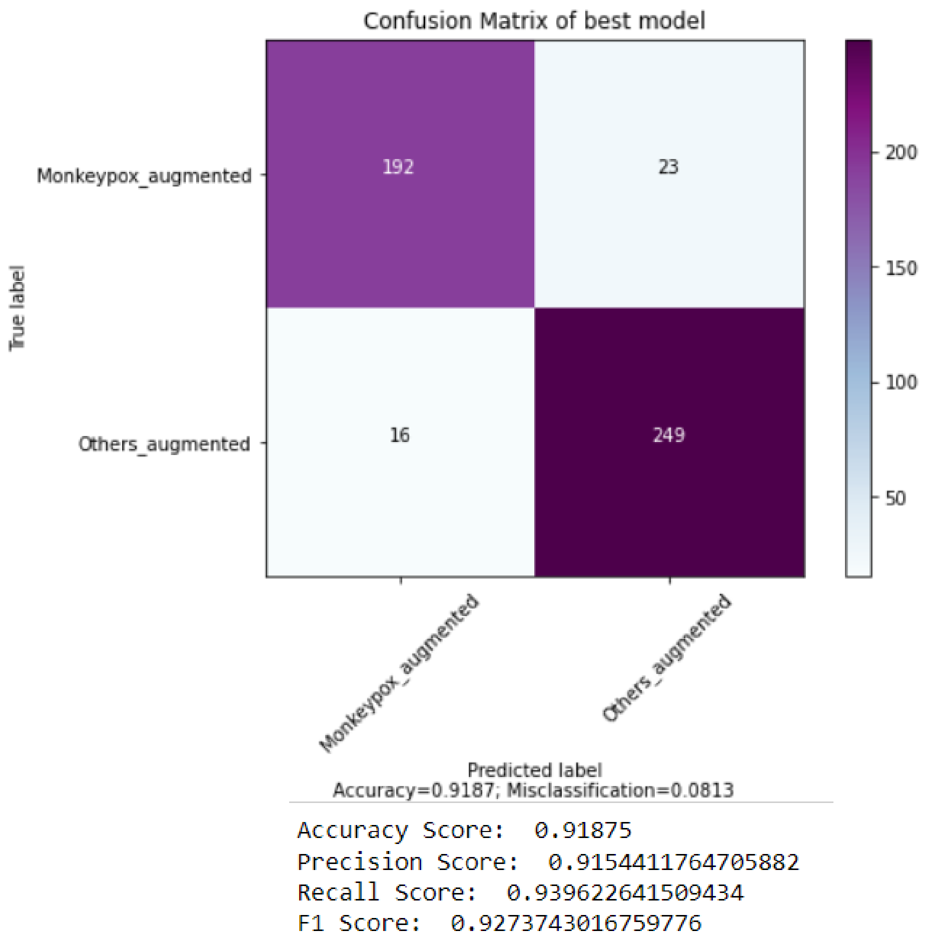

Figure 12: Confusion Matrix and Evaluation scores of VGG-19 Model1 (Best Model)

The comparative analysis outcome has been highlighted in the below figure 13:

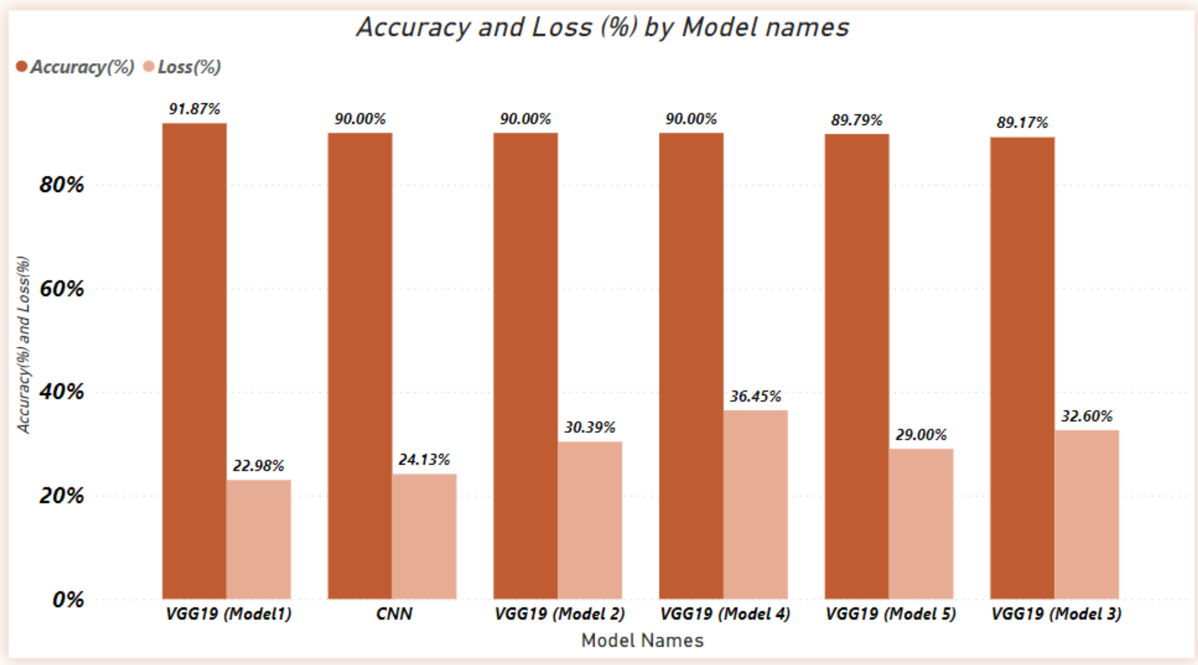

Figure 13: Modelwise comparative analysis

7 Conclusion and Future Work

In order to detect Monkeypox accurately in a short frame of time, multiple Convolutional Neural Network models have been used in this research, one of which is CNN base model and the other one is the advanced pre-trained version of CNN model known as VGG-19. Out of 6 models which have been used in this experiment, VGG-19 model with 2 dense layers provides an improved accuracy of 91.87% with only 8% of misclassification of the images in addition to faster processing. The accuracy of the model is exceptional and much more reliable when compared to earlier techniques of transfer learning that were used in other studies. The recommended strategy helps the model learn instances quickly compared to others and the accuracy also increases after a certain number of epochs. As a result, this model has the potential to be used in the future to properly diagnose a variety of additional skin diseases. The performance of the model may also be improved further with the addition of a complete and accurate data set derived from medical sources.

7.1 Acknowledgement

I would want to give recognition to Professor Bharat Agarwal since he was a significant contributor to my academic achievement and because, without him, I would not have been able to achieve this level of accomplishment. I am really thankful of his assistance, direction, and insightful inputs which were provided by him during the research.

References

Ahsan, M. M., Uddin, M. R., Farjana, M., Sakib, A. N., Momin, K. A. and Luna, S. A. (2022). Image data collection and implementation of deep learning-based model in detecting monkeypox disease using modified vgg16, *arXiv preprint arXiv:2206.01862* .

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* .

Ansari, U. B. and Sarode, T. (2017). Skin cancer detection using image processing, *Int Res J Eng Technol* **4**(4): 2875–2881.

Arbane, M., Benlamri, R., Brik, Y. and Djerioui, M. (2021). Transfer learning for automatic brain tumor classification using mri images, *2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH)*, IEEE, pp. 210–214.

Ayan, E. and Ünver, H. M. (2019). Diagnosis of pneumonia from chest x-ray images using deep learning, *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, Ieee, pp. 1–5.

Gouda, W., Sama, N. U., Al-Waakid, G., Humayun, M. and Jhanjhi, N. Z. (2022). Detection of skin cancer based on skin lesion images using deep learning, *Healthcare*, Vol. 10, MDPI, p. 1183.

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

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980* .

Madhavan, M. V., Khamparia, A., Gupta, D., Pande, S., Tiwari, P. and Hossain, M. S. (2021). Res-covnet: An internet of medical health things driven covid-19 framework using transfer learning, *Neural Computing and Applications* pp. 1–14.

Nadeem, M. W., Ghamdi, M. A. A., Hussain, M., Khan, M. A., Khan, K. M., Almotiri, S. H. and Butt, S. A. (2020). Brain tumor analysis empowered with deep learning: A review, taxonomy, and future challenges, *Brain sciences* **10**(2): 118.

Nasr-Esfahani, E., Samavi, S., Karimi, N., Soroushmehr, S. M. R., Jafari, M. H., Ward, K. and Najarian, K. (2016). Melanoma detection by analysis of clinical images using convolutional neural network, *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, pp. 1373–1376.

Sarumi, O. A. (2020). Machine learning-based big data analytics framework for ebola outbreak surveillance, *International Conference on Intelligent Systems Design and Applications*, Springer, pp. 580–589.

Seetha, J. and Raja, S. S. (2018). Brain tumor classification using convolutional neural networks, *Biomedical & Pharmacology Journal* **11**(3): 1457.

Sitaula, C. and Hossain, M. B. (2021). Attention-based vgg-16 model for covid-19 chest x-ray image classification, *Applied Intelligence* **51**(5): 2850–2863.

Sitaula, C. and Shahi, T. B. (2022). Monkeypox virus detection using pre-trained deep learning-based approaches, *Journal of Medical Systems* **46**(11): 1–9.

Song, Y., McLoughLin, I. and Dai, L. (2015). Deep bottleneck feature for image classification, *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, pp. 491–494.

Xiao, J., Wang, J., Cao, S. and Li, B. (2020). Application of a novel and improved vgg-19 network in the detection of workers wearing masks, *Journal of Physics: Conference Series*, Vol. 1518, IOP Publishing, p. 012041.

Yadav, N., Alfayeed, S. M., Khamparia, A., Pandey, B., Thanh, D. N. and Pande, S. (2022). Hsv model-based segmentation driven facial acne detection using deep learning, *Expert Systems* **39**(3): e12760.