Monkeypox Detection using CSA based K-Means Clustering with Swin Transformer Model

[1]Prabhu M, [2]Sathishkumar A, [3]Sasi G, [4]Lau Chee Yong, [5]Shanker M C and [6]Selvakumarasamy K
[1,2,5,6]Department of Electronics and Communication Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India.
[3]Department of Biomedical engineering, Veltech Multitech Dr.Rangarajan Dr.Sakunthala Engineering College, Chennai, Tamil Nadu, India.
[4]School of Engineering, Asia Pacific University of Technology and Innovation, Kuala Lumpur, Malaysia.
[1]prabhu.ece05@gmail.com, [2]sat090579@gmail.com, [3]gsasikumar75@gmail.com, [4]laucheeyong@apu.edu.my, [5]rekshanmc@gmail.com, [6]selvakumarasamyk@gmail.com

Correspondence should be addressed to Selvakumarasamy K : selvakumarasamyk@gmail.com.

Article Info
Journal of Machine and Computing (http://anapub.co.ke/journals/jmc/jmc.html)
Doi: https://doi.org/10.53759/7669/jmc202404038
Received 14 July 2023; Revised from 22 January 2024; Accepted 25 February 2024.
Available online 05 April 2024.

Abstract – Despite the global COVID-19 pandemic, public health professionals are also concerned about a possible new monkeypox epidemic. Similar to vaccinia, cowpox, and variola, the ortho poxvirus that causes monkeypox has two strands that are double-stranded. Many people have propagated the current pandemic through sexual means, particularly those who identify as bisexual or gay. The speed with which monkeypox was detected is the most important element here. In order to catch monkeypox before it infects more people, machine learning could be a huge help in making a quick and accurate diagnosis. Finding a solution is the driving force behind this project, which aims to develop a model for detecting monkeypox using deep learning and image processing. For optimal cluster selection during photo segmentation, the Chameleon Swarm Algorithm (CSA) employs K-means clustering. Examining the accuracy with which the Swin Transformer model identified instances of monkeypox was the driving force for this study. The proposed techniques are evaluated on two datasets: Kaggle Monkeypox Skin Lesion Dataset (MSLD) besides the Monkeypox Skin Image Dataset (MSID). We assessed the outcomes of various deep learning models using sensitivity, specificity, and balanced accuracy. Positive results from the projected process raise the possibility of its widespread application in monkeypox detection. This ingenious and cheap method can be put to good use in economically deprived communities that may not have access to proper laboratory facilities.

Keywords: Chameleon Swarm Algorithm, Monkeypox Disease, Swin Transformer, K-Means Clustering, Image Segmentation.

I. INTRODUCTION

The eponymous virus, which is medically known as monkeypox, mostly infects monkeys [1]. Globally, monkeypox is most commonly found in Asia, Africa, and Central and West Africa [2]. Although the virus can infect any species of mammal, the most common way for humans to get it is through biting an infected animal, most commonly a bat or monkey [3]. The early symptoms of monkeypox include fever, headache, drowsiness, and muscle pain. Smallpox, chickenpox, and measles-like symptoms were present. You might find swollen glands in your neck, groyne, behind your ears, or under your chin before the rash shows up [4]. Viruses do not cause death on their own, although severe infections can cause blindness, pneumonia, and sepsis [5]. Although people rarely contract monkeypox, those living in areas where the disease is more prevalent should exercise caution around monkeys and rats due to the minimal possibility of infection. The Centres predict that millions of people worldwide will be infected with a new strain of monkeypox in the next years [6].

Although monkeypox is more closely related to smallpox in terms of clinical manifestations, the rash and skin lesions can resemble those of cowpox and chickenpox [7]. Its early stages are also difficult to detect due to its striking resemblance to chickenpox and measles. These shared features make it difficult for some medical professionals to distinguish monkeypox from other skin conditions only by looking at the lesions and rashes. A "substantial danger" to public health around the world, the pandemic has been warned by the World Health Organisation, which has not yet declared an emergency [8]. Also, healthcare groups like the World Health Network (WHN) were very worried and wanted the world to do something about the epidemic quickly.

In order to contain monkeypox, it is essential to identify cases early, track infected people through matched contacts, and isolate them as soon as possible. In this case, automated computer-aided methods based on AI could significantly restrict its global proliferation [9]. Once enough samples are available, it has been shown that DL methods can be used for automatic skin infection classification. Using training data from massive datasets, these deep networks can learn to choose the best approximations for certain tasks by analysing images in numerous layers [10]. The time and specialised computing resources needed to train DL-based frameworks restricts their practical application. Big datasets are also necessary. Another typical method when resources are restricted is transfer learning [11].

Deep learning techniques, such as forward and backward propagation, are used to build a model in convolutional neural network (CNN) image classification. The learned model is then used to classify new images [12]. Consequently, the most common approaches to picture classification include convolutional neural network (CNN) based classification, artificial neural network (ANN) based image classification, and support vector machine (SVM) based image classification. To evaluate how unique the clustering results are, the k-means clustering algorithm is suggested. Different clustering centroids occur when the central values do not change after using the clustering procedure. Unsupervised learning is exemplified by that model [13]. In this study, we hypothesise that finding the true number of clusters will lead to more reliable cluster outcomes. In data mining and image processing, clustering is a crucial technique for integrating numerical and visual data. Clustering is commonly employed in the medical and other research and development areas to efficiently group symptoms of illness and medicines, therefore saving time and improving efficiency [14]. Astronomy, data mining, and advertising are just a few of its numerous uses.

At the time this article was created, there were very few articles that showed how image processing and deep learning could be used to diagnose monkeypox. These articles provided the data used in this article. While this article was being written, these papers were leaked. The absence of a publicly available dataset for training and testing slowed down the development of a system for image-based diagnosis of monkeypox. The virus has recently infected huge numbers in several parts of the world, thus that's why it happened.

To tackle this problem, this study builds a monkeypox diagnostic model using deep learning and image processing. This includes the parts that follow: Introduction, Nature of the Experiment, Procedures, Data Collection and Preprocessing, Model Proposal, Results Analysis, Discussion, Literature Review, and Conclusion

II. RELATED WORKS

Using deep learning architectures, Taspinar et al. [15] want to diagnose monkeypox by predicting skin lesions in patients. We drew from a dataset that had images of individuals suffering from monkeypox and other skin illnesses. Separately, we evaluated how well the VGG16 and VGG19 CNN models classified images. It was also found that when these structures go through transfer learning and tuning processes, the success rate goes up. The VGG19 model got the highest score for classification accuracy (97.81%). The pictures were reviewed with a focus on refining the VGG19 model, which had the best results using the GradCAM method.

Savaş [16] introduces a method for visually differentiating diseases—in this instance, measles, chickenpox, and MPox—based on deep learning in response to the global pandemic of 2022. The research presented an optimisation strategy with two steps. This study examines seventy-one models of pre-trained ANNs with the goal of optimising accuracy using transfer learning, fine-tuning, and ensemble learning approaches. The first stage involved selecting models with a 97.5% accuracy rate, which included ConvNeXtBase, Large, and XLarge. Then, using the optimisation method's ensemble learning process, specific criteria were employed to choose the models from the first stage. In the second stage, the top ensemble model, EM3 (ResNetRS101, ResNet101, and RegNetX160), gets an AUC of 0.9971. Testing on unknown data improves the model's generalizability, validity, and reliability. The research has been fine-tuned in both design and execution to address the limitations literature. This technology enhances clinic efficiency by reducing the risk of human error and manual processes, while simultaneously delivering a fast and accurate decision support system for early MPox diagnosis. It directs the development of software for imaging instruments, aids in the early detection of MPox, and solves problems associated with other diseases. The study's far-reaching consequences lend credibility to global health programmes and demonstrate the potential of AI in medical informatics for illness identification and diagnosis.

Using vision transformer as its foundation, Aloraini [17] presented a novel approach to automatically detecting monkeypox virus in skin photographs. Tuning the vision transformer fixes the problems with human monkeypox classification. A picture is divided into small patches and fed sequentially into this fine-tuned vision transformer in order for it to recognise monkeypox virus. The experimental consequences show that the projected technique achieves outstanding detection accuracy using a publically available dataset.

One approach to classifying using CNNs is proposed by Eliwa et al. [18]. Better (AUC) were achieved by optimising the CNN model approach, as compared to the non-optimized model. The GWO optimisation technique can improve the performance of CNN representations when used to similar tasks. With an impressive 95.3% accuracy, the GWO optimizer boosted the model's ability to differentiate between classifications. There are a number of potential benefits to using the proposed technique for monkeypox diagnosis and surveillance. If monkeypox skin lesions could be diagnosed more quickly and accurately, it would allow for earlier finding and better patient outcomes. Furthermore, the approach might have major ramifications for public health in terms of monkeypox outbreaks. All things considered, this study's findings offer a novel and encouraging approach to monkeypox diagnosis, which might have significant real-world consequences.

The computer-vision-based technique of diagnosing MPX by analysing images of skin lesions put out by Almufareh et al., [19] is safer and more intelligent than traditional procedures of diagnosis. The proposed method services deep learning algorithms to ascertain whether skin lesions are positive for MPXV or not. The proposed techniques are evaluated on two datasets: the Kaggle Dataset (MSID). We assessed the outcomes of various sensitivity, accuracy. Positive results from the proposed method raise the possibility of its widespread application in monkeypox detection. This ingenious and low-cost strategy might be a lifesaver for low-income areas that lack the resources to establish their own laboratories.

The aim of Pal et al., [20] was to achieve a rapid and concise mpox diagnosis. Bisexuals and sexually active gay males are among the groups that have been identified as potentially susceptible to the spread of mpox. Therefore, isolating, vaccinating, and treating them selectively seems unfeasible due to the societal has great promise for image-based diagnosis and may even help with mass diagnosis that is free of errors. The proposed innovation, the system used, and the methodology and approaches are all explored in detail in the essay. In this research, we propose a method for diagnosis that makes use of DL models. The proposed methods were evaluated in this study using a dataset that is available to the public. Extensive details are provided regarding the data gathering process that the study employed, with the dual goals of testing and training. We pitted numerous algorithms against one another, including Convolutional Neural Networks (CNNs), VGG19, ResNet50, Inception v3, and Autoencoder. Inception v3, CNN, and VGG19 could all help diagnose mpox skin lesions early; however, Inception v3 achieved the accuracy at 96.56%.

III. PROPOSED METHODOLOGY

We use a fully mechanical, non-invasive deep learning technology to detect viruses. Using an input image of normal skin, the proposed method scans for monkeypox skin lesions. Once the pre-processing steps have been executed, the input is transferred learning-classified as monkeypox or another sort of skin condition. **Fig 1** shows the entire blueprint.

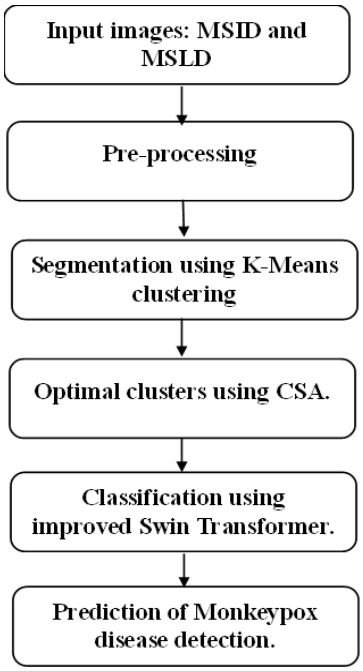

Fig 1. Workflow of Projected Model

Data Acquisition and Preparation

The present study employs two datasets for its experimental purposes: the Monkeypox Skin Image Dataset (MSID) and the Kaggle Monkeypox Skin Lesion Dataset (MSLD) [21]. **Table 1** displays the dataset details. For the sake of consistency, the four separate groups that make up the MSID dataset—combined into one. The picture classes were gathered from a variety of web resources. Creating the full dataset was the responsibility of the Department of Engineering at Islamic University in Kushtia-7003, Bangladesh. Images collected and processed from a variety of web-scraping sources, including websites, news portals, and publicly obtainable case reports, are used to create the MSLD. Differentiating monkeypox patients from other similar non-monkeypox cases was the primary goal of developing the MSLD. The first rash and pustules of monkeypox are similar to chickenpox and measles skin lesion pictures, which are part of the MSLD to help achieve this. A binary categorization system divides these pictures into two groups: "monkeypox" and "others."

Table 1. Data Delivery.

Dataset Name	Others	Monkeypox
MSLD	126	102
MSID	198	279

Preprocessing

A three-channel picture data set with a spatial dimension of MxN is the input ϰ(i,j). With abd (i,j) pointing to the pixel site in the image, the image ℂ has M rows besides N columns. Prior to moving on to the feature extraction module, the pixels are normalized. The input image is subjected to the mapping function ϰ. This is the definition of the mapping purpose ϰ: [0,255]→[0,1].

$$\mathcal{X}(\varkappa(i,j)^c) = \frac{\varkappa(i,j)}{max_{s,t}\varkappa(s,t)} \tag{1}$$

$$i \in [1,M], j \in [1,M] \; and \; C \in \; R,G,B.$$

Segmentation using Improved K-Means Clustering

The clustering method that relies on partitioning. It is an unsupervised method for clustering. After the centroid has been meticulously chosen and compared to the data points using their intensity and features to establish the distance, points that are similar to it are assigned to the cluster where the centroid is situated. In order to compute new 'k' centroids and create new k-clusters, it is necessary to identify the data points that are nearest to the clusters. The following are the main components of the k-means algorithm:

- Using the Chameleon Swarm Algorithm (CSA), select k locations and set them as the initial centers.
- If a data point from the collection finds a match (minimum distance) with each centroid, it is placed in the cluster for that centroid. We break ties (at equal distance) at random.
- After assigning each data point to a cluster, recalculate the values of the centroid for each of the k-point clusters.
- Keep going with the previous steps until there is no change in data points between the clusters.

CSA

The CSA, developed by Braik [22] in 2021, mimics the way chameleons find and catch prey. Species that can change their coloration and pass themselves off as another are very adaptable and can live in a wide variety of habitats.

Evaluation of Initialization and Function

Being a population- algorithm, CSA starts the search process by randomly creating a population of individuals. Each member of the n-chameleon population represents a distinct key to the optimization issue, and the search area has d dimensions. At every particular iteration, the location of each chameleon in the search region is described by Equation (2).

$$y_t^i = \left[y_{t,1}^1, y_{t,2}^i, \ldots, y_{t,d}^i\right] \tag{2}$$

In order to generate the initial population, the issue space are calculated, as shown in Equation (3):

$$y^i = l_j + r(u_j - l_j) \tag{3}$$

Every new position in every step has its solution quality determined by the evaluation of the fitness function.

Searching for a Target

Equation characterizes the approach for updating the chameleons' positions during the search (4):

$$y_{t+1}^{i,j} = \begin{cases} y_t^{i,j} + P_1\left(P_t^{i,j} - G_t^j\right)r_2 + P_2\left(G_t^j - y_t^{i,j}\right)r_2 \\ y_t^{i,j} + \mu(u^j - l^j)r_3 + l_b^j sgn(rand - 0.5) r_1 < P_p \end{cases} r_1 \geq P_p \tag{4}$$

Eyes Rotation of Chameleon

Through the use of their eye-rotating abilities, chameleons are able to detect the position of their prey in all directions. Here is the order in which the following steps take place:

- The chameleon's center of gravity or focus point is its original location or starting point.
- By computing matrix, the prey's location can be determined.
- The chameleon's focal point location is then rationalized using matrix.
- Lastly, they are returned to their initial position.

Hunt of Target

When a chameleon sees its prey getting too close, it will launch an attack. Being the most ideal option, it has the shortest distance to the prey. Chameleons use their tongues to assault their prey. Chameleons can improve their positioning by using

their tongues, which can expand up to double their length. This allows the chameleon to make good use of its hunting area and bring its victim to hand. To simulate the speed of a chameleons toward its victim, one can utilize Equation (5):

$$v_{t+1}^{i,j} = wv_t^{i,j} + c_1\left(G_t^j - y_t^{i,j}\right) + c_2\left(P_t^{i,j} - y_t^{i,j}\right) + c_2\left(P_t^{i,j} - y_t^{i,j}\right)r_2 \quad (5)$$

Classification using Deep Learning

The research used Swin Transformer to aid with monkeypox detection. Three parts make up the Swin Transformer architecture: processing images, blocking attention, and performing activities downstream.

Using a convolutional neural network means treating the picture as a matrix for the convolution process; on the other hand, the transformer is derived from natural language processing and is used to handle sequences of natural language. As a CNN, it isn't very user-friendly when it comes to extracting features from images. Because of this, the model started using patching techniques including masking, patch merging, and patch embedding.

Patch Embedding

Patch partition is responsible for transforming RGB maps into blocks of patches that do not overlap. The patch in this case is 48, which is obtained by multiplying 4 x 4 by the relevant RGB channels. By transforming the processed patches into a feature matrix according to the given dimensions,

Patch Merging

The feature matrix is alienated into 2×2 $size\ windows$, and then the consistent positions of each window are combined. Finally, after merging, the four feature mediums are connected.

Mask

The design of the mask ensures that the window will solely focus on the continuous portion following the relocation of the SW-MSA. With a little adjustment, you can find the initial window in the matrix in the upper left. The following is the formulation for the link between the size of the shift and the size of the window:

$$s = \left[\frac{w}{2}\right] \quad (6)$$

in which w is window and s is the size of the shift.

It is necessary to use the mask matrix to separate the region to the right and below, which is visible and horizontal slicing areas are equal, with the former consisting of [0, −window_size], the latter of [−window_size, −shift_size], and the former of [−shift_size]. Gulf the window size into blocks equally. Merge on behalf of the sum and the measurement of the batch size. This will produce the labeled mask matrix that follows the window partition (function window partition). Mask slicing requires partitioning the unique matrix mask into smaller windows, which is why this division is necessary.

Algorithm Design: First and Second Stage

Image processing was finished in the first step of the investigation. A common practice in transformer architecture is to first divide the RGB monkeypox image into many non-overlapping patches. As a result of every pixel in the Swin Transformer [23] arrangement having RGB three-channel values, every 4 × 4 × 3 patch is ultimately turned embedding layer.

The Swin Transformer block is the second step. Swin [23] imprisonments deep properties by stacking many blocks, similar to most CNN topologies. For the purpose of learning picture features, this research makes use of four repeated attention blocks. The space you specify is used to project the processed patches. We used linear embedding to partition the input feature into C dimensions before passing it on to the Swin Transformer Block. The Swin Transformer block is made up of two layers of MLP and a shift window-based MSA. A residual connection follows each MSA module and MLP, after which there is a layer specification (LN) layer. Afterwards, patches within the immediate 2 × 2 range are stitched first using the Patch Merging operation. The result is that the feature dimension is 4C and the sum of patch blocks is H/8 × W/8. Following the same linear embedding procedure as in stage 1, 4C is reduced to 2C before being inputted into the Swin block. The layered representation that results from combining these blocks is equivalent in feature mapping resolution to that of a standard convolutional network.

Self-Attention in Non-Overlapped Windows

Many vision applications necessitate large sets of dense forecast or representation of images, and global computation is not suited for these tasks due to its quadratic difficulty in terms of the sum of tokens. Alternatively, efficient modeling can be achieved by computing self-attention within a limited window. These windows ensure that the photographs are consistently split without overlapping. Under the assumption that every window has M × M patches, the MSA module and h × w patch pictures is high.

$$\Omega(MSA) = 4hwC^2 + 2(hw)^2C \quad (7)$$

$$\Omega(W-MSA) = 4hwC^2 + 2M^2hwC \tag{8}$$

as the latter shadows a linear path (with a default value of 7) when M is constant and the former is squared with the sum of hw. Global self-attentive computing is usually expensive due to the massive hardware requirement, in contrast to self-attentiveness, which is scalable.

Shifted Window Partitioning in Successive Blocks
They add cross-window connections since their modeling skills are severely limited due to the lack of info sharing among non-overlapping windows. This method switches among W-MSA and SW-MSA every two blocks of Swin Transformer. By increasing the perceptual range of opinion, the shifted window division method connects consecutive windows in the higher layer. Here are the calculations for the Swin Transformer blocks using the rearranged structure:

$$\begin{aligned} \hat{z}^l &= W-MSA\big(LN(z^{l-1})\big) + z^{l-1} \\ z^l &= MLP\big(LN(\hat{z}^l)\big) + \hat{z}^l \\ \hat{z}^{l+1} &= SW-MSA\big(LN(z^l)\big) + z^l \\ z^{l+1} &= MLP\big(LN(\hat{z}^{l+1})\big) + \hat{z}^{l+1} \end{aligned} \tag{9}$$

where $\hat{z}^l$ and z^l represents the MSA module for block and the MLP module for window-based multi-head algorithms, respectively; W-MSA and SW-MSA employ normal besides removed window partitioning configurations, singly.

Multihead Self-Attention
To transition from Transformer to vision, a multihead attention technique is employed. Here is the exact formula:

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d}} + B\right)V\left(Q,K,V \in R^{M^{2+d}}\right) \tag{10}$$

where B stands for identical to how Transformer is introduced. d is the size corresponding dimension, which helps to maintain a balance between QK^T and B. Q, K, V computation: following a linear layer, the derived for the incoming window information.

Third Stage
We utilized a Swin Transformer to accomplish the organization besides segmentation missions, correspondingly, after introducing how to use one to extract features in the previous section. Both the classification and segmentation missions are located downstream in the third stage. In our studies, there were two categories: with and without nodules. The number of classifications is used as the output dimension for the classification mission. To acquire the final classification probability, the output is passed via softmax.

Architecture Variants
With a perfect size and computational difficulty comparable to ViT-B/DeiT-B, the technique utilized Swin Transformer's basic perfect, Swin-B. Training with all of the Swin-T besides Swin-S variants was not possible due to limitations in experimental equipment. A model with 0.25 times the computational complexity is Swin-S, while another model with 0.5 times the complexity is Swin-T. In terms of complexity, Swin-T is on par with ResNet-50 (DeiT-S) and Swin-S is on par with ResNet-101. M= 7 is the default value for the window size. Each MLP has an expansion layer of 4, and the dimension for all experiments is d = 32 for each head. The hyperparameters for the model's architecture in these iterations are:

- Swin-T- -C = 96, layer statistics = $\{2, 2, 6, 2\}$;
- Swin-S- -C = 96, layer statistics = $\{2, 2, 18, 2\}$;
- Swin-B- -C = 128, layer statistics = $\{2, 2, 18, 2\}$.
-

where C is sum of the hidden in the primary phase.

Loss Function
The mission is as shadows:

$$L = \frac{1}{N}\sum_i L_i = -\frac{1}{N}\sum_i \sum_{c=1}^{M} y_{ic} log(p_{ic}) \tag{11}$$

where M is the sum of groups and y_{ic} is the purpose (0 *or* 1). If group of the taster i is equivalent to 2, take 1; then, take 0. p_{ij} is the foretold likelihood experiential example i belongs to group 2.

IV. RESULTS AND DISCUSSION

To ensure that the model performs well during training, it is crucial to monitor its performance on a validation set. If the performance starts to drop, training should be stopped. In addition to enhancing generalization performance, it saves computational resources by preventing overfitting. The $NVIDIA\ GeForce\ GTX\ 1060\ GPU$ was used for all investigations. For every experiment, the mini batch was set at 16.

Validation Analysis of proposed model on First dataset – MSLD

Table 2. Analysis of Proposed Model

Methods	Accuracy	Specificity	Sensitivity
DBN	92.1	86.7	83.4
CNN	89.4	81.1	85.4
RNN	87.7	83.4	82.2
Swin Transformer	90.7	89.4	86.2
Proposed model	**95.9**	**91.2**	**88.1**

In above **Table 2** characterise that the Investigation of proposed perfect. In the scrutiny of DBN model accomplished the accuracy as 92.1 and specificity as 86.7 and then sensitivity as 83.4 congruently. Then the CNN flawless reached the accuracy as 89.4 and specificity as 81.1 and then sensitivity as 85.4 correspondingly. Then the and then sensitivity as RNN model attained the accuracy as 87.7 and specificity as 83.4 and then sensitivity as 82.2 similarly. Then the Swin Transformer 90.7 89.4 and then sensitivity as 86.2 correspondingly. Then the Proposed perfect accomplished the accuracy as 95.9 and specificity as 91.2 and then sensitivity as 88.1 correspondingly.

Validation Analysis of proposed model on Second Dataset – MSID

Table 3. Experimental Analysis

Methods	Accuracy	Specificity	Sensitivity
DBN	87.1	84.3	84.1
CNN	86.2	83.7	82.4
RNN	85.9	82.1	83.7
Swin Transformer	79.9	87.7	86.9
Proposed model	89.8	90.6	93.4

In above **Table 3** characterise that the Trial Investigation. In the investigation of DBN perfect attained the correctness as 87.1 and specificity as 84.3 and then sensitivity as 84.1 similarly. Then the CNN faultless got the accuracy as 86.2 and specificity as 83.7 and then sensitivity as 82.4 consistently. Then the RNN perfect accomplished the accuracy as 85.9 and specificity as 82.1 and then sensitivity as 83.7 consistently. Then the Swin Transformer model attained the accuracy as 79.9 and then sensitivity as 86.9 similarly. Then the Proposed perfect accomplished accuracy as 89.8 and specificity as 90.6 and then sensitivity as 93.4 correspondingly.

V. CONCLUSIONS `

In areas where medical services are scarce, monkeypox is most commonly reported. Consequently, this approach is quite beneficial for these regions. The dataset including monkeypox skin lesions was pre-processed to improve the quality of the pictures for this investigation. Two publicly available datasets, MSID and MSLD, were used to test several deep networks. The transfer learning technique was used to convey domain knowledge. Utilizing K-means clustering, with clusters optimally chosen by CSA, for segmentation. Applying a novel classification approach based on an efficient transformer to the analysis of medical images is the focus of this research article. The computational cost of this novel visual converter is directly proportional to the input image size, and it produces hierarchical feature representations. For monkeypox detection using AI, there are a number of promising avenues for future study. Scientists can improve their computer vision algorithms to the point where they can distinguish between monkeypox stages in photos. This study is constrained by its dataset. Future studies can enhance it. Scientists can test the viability and efficacy of monkeypox detection systems powered by artificial intelligence in real-world scenarios. Collecting massive datasets and working together with healthcare providers and patients will be necessary for this. In addition, the aforementioned issue can also be investigated with alternative architectures. By including supplementary data sources like patient results, researchers might enhance the accuracy of AI-based monkeypox identification. A more complete picture of the illness and better diagnostic accuracy can be achieved in this way.

Data Availability
No data was used to support this study.

Conflicts of Interests
The author(s) declare(s) that they have no conflicts of interest.

Funding
No funding agency is associated with this research.

Competing Interests
There are no competing interests.

References

[1]. M. C. Irmak, T. Aydin, and M. Yaganoglu, "Monkeypox Skin Lesion Detection with MobileNetV2 and VGGNet Models," 2022 Medical Technologies Congress (TIPTEKNO), Oct. 2022, doi: 10.1109/tiptekno56568.2022.9960194.
[2]. V. H. Sahin, I. Oztel, and G. Yolcu Oztel, "Human Monkeypox Classification from Skin Lesion Images with Deep Pre-trained Network using Mobile Application," Journal of Medical Systems, vol. 46, no. 11, Oct. 2022, doi: 10.1007/s10916-022-01863-7.
[3]. T. Islam, M. A. Hussain, F. U. H. Chowdhury, and B. M. R. Islam, "Can Artificial Intelligence Detect Monkeypox from Digital Skin Images?," Aug. 2022, doi: 10.1101/2022.08.08.503193.
[4]. D. A. León-Figueroa et al., "Detection of Monkeypox Virus according to The Collection Site of Samples from Confirmed Cases: A Systematic Review," Tropical Medicine and Infectious Disease, vol. 8, no. 1, p. 4, Dec. 2022, doi: 10.3390/tropicalmed8010004.
[5]. M. Altindis, E. Puca, and L. Shapo, "Diagnosis of monkeypox virus – An overview," Travel Medicine and Infectious Disease, vol. 50, p. 102459, Nov. 2022, doi: 10.1016/j.tmaid.2022.102459.
[6]. C. Sitaula and T. B. Shahi, "Monkeypox Virus Detection Using Pre-trained Deep Learning-based Approaches," Journal of Medical Systems, vol. 46, no. 11, Oct. 2022, doi: 10.1007/s10916-022-01868-2.
[7]. K. D. Akin, C. Gurkan, A. Budak, And H. Karataş, "Açıklanabilir Yapay Zeka Destekli Evrişimsel Sinir Ağları Kullanılarak Maymun Çiçeği Deri Lezyonunun Sınıflandırılması," European Journal of Science and Technology, Sep. 2022, doi: 10.31590/ejosat.1171816.
[8]. S. Gürbüz and G. Aydin, "Monkeypox Skin Lesion Detection Using Deep Learning Models," 2022 International Conference on Computers and Artificial Intelligence Technologies (CAIT), Nov. 2022, doi: 10.1109/cait56099.2022.10072140.
[9]. D. de Sousa et al., "Monkeypox Diagnosis by Cutaneous and Mucosal Findings," Infectious Disease Reports, vol. 14, no. 5, pp. 759–764, Sep. 2022, doi: 10.3390/idr14050077.
[10]. Shah, "Monkeypox Skin Lesion Classification Using Transfer Learning Approach," 2022 IEEE Bombay Section Signature Conference (IBSSC), Dec. 2022, doi: 10.1109/ibssc56953.2022.10037374.
[11]. M. M. Ahsan et al., "Deep transfer learning approaches for Monkeypox disease diagnosis," Expert Systems with Applications, vol. 216, p. 119483, Apr. 2023, doi: 10.1016/j.eswa.2022.119483.
[12]. E. J. Tarín-Vicente et al., "Clinical presentation and virological assessment of confirmed human monkeypox virus cases in Spain: a prospective observational cohort study," The Lancet, vol. 400, no. 10353, pp. 661–669, Aug. 2022, doi: 10.1016/s0140-6736(22)01436-2.
[13]. J. Rodriguez-Morales, "Monkeypox and the importance of cutaneous manifestations for disease suspicion," Microbes, Infection and Chemotherapy, vol. 2, p. e1450, Jun. 2022, doi: 10.54034/mic.e1450.
[14]. V. Kumar, "Analysis of CNN features with multiple machine learning classifiers in diagnosis of monkeypox from digital skin images," Sep. 2022, doi: 10.1101/2022.09.11.22278797.
[15]. Taspinar, Yavuz & Cinar, Ilkay & Kursun, Ramazan & Koklu, Murat. " Monkeypox Skin Lesion Detection with Deep Learning Models and Development of Its Mobile Application". International Journal of Research in Engineering and Science, (2024), 12. 273-285.
[16]. S. Savaş, "Enhancing Disease Classification with Deep Learning: a Two-Stage Optimization Approach for Monkeypox and Similar Skin Lesion Diseases," Journal of Imaging Informatics in Medicine, Jan. 2024, doi: 10.1007/s10278-023-00941-7.
[17]. M. Aloraini, "An effective human monkeypox classification using vision transformer," International Journal of Imaging Systems and Technology, vol. 34, no. 1, Jul. 2023, doi: 10.1002/ima.22944.
[18]. E. H. I. Eliwa, A. M. El Koshiry, T. Abd El-Hafeez, and H. M. Farghaly, "Utilizing convolutional neural networks to classify monkeypox skin lesions," Scientific Reports, vol. 13, no. 1, Sep. 2023, doi: 10.1038/s41598-023-41545-z.
[19]. M. F. Almufareh, S. Tehsin, M. Humayun, and S. Kausar, "A Transfer Learning Approach for Clinical Detection Support of Monkeypox Skin Lesions," Diagnostics, vol. 13, no. 8, p. 1503, Apr. 2023, doi: 10.3390/diagnostics13081503.
[20]. M. Pal et al., "Deep and Transfer Learning Approaches for Automated Early Detection of Monkeypox (Mpox) Alongside Other Similar Skin Lesions and Their Classification," ACS Omega, vol. 8, no. 35, pp. 31747–31757, Aug. 2023, doi: 10.1021/acsomega.3c02784.
[21]. M. F. Almufareh, S. Tehsin, M. Humayun, and S. Kausar, "A Transfer Learning Approach for Clinical Detection Support of Monkeypox Skin Lesions," Diagnostics, vol. 13, no. 8, p. 1503, Apr. 2023, doi: 10.3390/diagnostics13081503.
[22]. M. S. Braik, "Chameleon Swarm Algorithm: A bio-inspired optimizer for solving engineering design problems," Expert Systems with Applications, vol. 174, p. 114685, Jul. 2021, doi: 10.1016/j.eswa.2021.114685.
[23]. B. Gunapriya, T. Rajesh, A. Thirumalraj, and M. B, "LW-CNN-based extraction with optimized encoder-decoder model for detection of diabetic retinopathy," Journal of Autonomous Intelligence, vol. 7, no. 3, Dec. 2023, doi: 10.32629/jai.v7i3.1095.