Mpox Detection Using Convolutional Neural Networks

Mopuru Venkata Srikanth Reddy(mvsr26032005@gmail.com), Kothapally Kartikeya(kartikeyavk18@gmail.com) and Kanduri Rama Narasimha(kramanarasimha1@gmail.com).

Vellore Institute of Technology, Amaravathi

Abstract:

Mpox, formerly known as monkeypox, is a zoonotic disease within the Orthopoxvirus genus. The World Health Organization (WHO) reports that between January 2022 and September 2024, there were 109,699 confirmed Mpox cases globally, leading to 236 deaths across 123 countries. Symptoms often include fever, headaches, swollen lymph nodes, and fatigue, resembling illnesses such as chickenpox and measles, which complicates diagnosis based solely on clinical features. Early symptoms in these diseases overlap, making initial differentiation challenging. This research uses deep learning, specifically convolutional neural networks (CNNs), to enhance diagnostic accuracy in Mpox. We trained models like ResNet152, DenseNet169, EfficientNetB7, and VGG16 on the Mpox Skin Lesion Dataset Version 2.0 (MSLD v2.0) from Kaggle, containing images of Mpox and other skin lesions. Among these, DesneNet169 was our best performing model with 93% of accuracy, ResNet152 and DenseNet169 achieved the accuracy more than 90%, outperforming other models, which ranged between 78-85% accuracy. Our CNN-based model offers a valuable tool for self-assessment and preliminary diagnostics, especially where Polymerase Chain Reaction (PCR) tests are inaccessible, and aids healthcare professionals in remote consultation scenarios.

Keywords: Deep learning, Convolutional Neural Network, Residual Network, Densely Connected Convolutional Network, DenseNet, ResNet , Monkeypox

Introduction

Mpox, formerly known as monkeypox, is a zoonotic virus within the Orthopoxvirus genus, closely related to smallpox. Though previously endemic to Central and West African regions, the global spread of Mpox has intensified since 2022. The World Health Organization (WHO) reports that, as of September 2024, there were 109,699 confirmed cases and 236 deaths across 123 countries, marking a case fatality rate that varies widely depending on healthcare access and population vulnerability. Higher fatality rates are typically observed in immunocompromised individuals, children, and pregnant women, underscoring the importance of prompt and accurate diagnosis

Differentiating Mpox from similar infections, such as chickenpox and measles, poses clinical challenges, as these illnesses often share early symptoms like fever, muscle pain, and fatigue. Mpox, however, progresses with unique pustular lesions that are generally more firm, while chickenpox (caused by the varicella-zoster virus) manifests as itchy, fluid-filled blisters, and measles displays a red, flat rash originating on the face and spreading down the body. However, reliance solely on symptom presentation can lead to misdiagnosis, especially in settings without Polymerase Chain Reaction (PCR) testing. This difficulty has driven a growing interest in the application of artificial intelligence (AI) and deep learning (DL) to improve diagnostic reliability for such conditions.

Deep learning, and especially convolutional neural networks (CNNs), has become a cornerstone in modern medical diagnostics, transforming how complex imaging data is processed and analyzed. CNNs excel in pattern recognition tasks that involve large, complex datasets, making them especially suited for medical imaging, where the precise differentiation of diseases is critical. By learning nuanced patterns in imaging data, CNNs can assist in the differentiation of visually similar diseases, such as Mpox, chickenpox, and measles, which often share overlapping visual and symptomatic characteristics. This capability allows for more refined diagnostic accuracy, supporting clinicians in quickly identifying diseases that might otherwise require time-intensive and costly tests, such as Polymerase Chain Reaction (PCR) tests.

In our study, we explored several CNN architectures, including ResNet152, DenseNet169, EfficientNetB7, and VGG16, trained on the Mpox Skin Lesion Dataset (MSLD v2.0). This dataset, curated to provide a wide variety of skin lesion images, offers a substantial base for developing models that can discern between Mpox and other similar viral infections. Among these architectures, ResNet152 demonstrated a classification accuracy of 90%, significantly outperforming other models, which exhibited performance in the 80-85% range. ResNet152's layered residual connections likely contributed to this outcome, enabling deeper learning without the vanishing gradient issues often faced in deep networks. This architecture, known for its depth and ability to handle complex hierarchical features, proved highly effective for lesion classification tasks.

In addition to the superior performance of ResNet152, CNN-based diagnostics like this one highlight the growing role of deep learning in healthcare. Beyond lesion classification, CNNs are increasingly being used for applications such as radiology, histopathology, and even genomics. They enable early detection and monitoring of various diseases, including cancer, diabetic retinopathy, and neurodegenerative disorders, where early intervention can significantly alter patient outcomes. The automation of these processes not only supports overburdened healthcare systems by reducing diagnostic delays but also aids in areas with limited access to specialist care. The growing integration of deep learning into medical imaging is thus reshaping the future of healthcare, offering high-throughput, cost-effective solutions that democratize access to diagnostic services globally, particularly in underserved regions.

The adoption of deep learning in healthcare is revolutionizing diagnostics beyond infectious disease, from early cancer detection to personalized medicine and robotic-assisted surgeries. This shift represents a move toward AI-enhanced healthcare, where DL algorithms not only improve diagnostic precision but also enable remote healthcare access, triage efficiency, and rapid clinical decision-making. The integration of CNNs and other AI tools in real-time diagnostic support is reshaping how healthcare professionals respond to public health challenges, making these technologies essential in both present and future healthcare frameworks.

Literature:

The study titled **“Monkeypox Skin Lesion Detection Using Deep Learning Models: A Feasibility Study”** [1] utilized deep learning models to automate the detection of monkeypox skin lesions. The methodology involved gathering images of monkeypox, chickenpox, and measles lesions from publicly available sources and employing data augmentation techniques to enhance the dataset, significantly increasing the sample size for training. The research evaluated several pre-trained deep learning models, including VGG-16, ResNet50, and InceptionV3, to assess their effectiveness in differentiating monkeypox from other diseases. The findings revealed that ResNet50 achieved the highest accuracy, indicating the potential for AI-assisted early diagnosis of monkeypox, which could facilitate timely intervention. However, the study acknowledged certain drawbacks, such as the need for a larger and more diverse dataset to improve the models' generalizability and robustness. Additionally, while a prototype web application was developed for preliminary screenings, further validation and refinement are necessary before clinical application.

“A Novel CNN Model for Monkeypox Classification and Detection” [2] highlights the significant advancements in deep learning applications for medical image analysis, particularly concerning infectious diseases. It references prior studies that utilized pre-trained models like ResNet50 and VGG-16, demonstrating their effectiveness in monkeypox classification through data augmentation and transfer learning. These studies emphasize the need for rapid and accurate diagnostic tools, addressing the limitations of traditional testing methods. By contextualizing its novel CNN model within this framework, the paper underscores its potential to enhance early detection and management of monkeypox, particularly in regions with limited healthcare resources.

Both the papers highlight use of ResNet50 in classification of Mpox and VGG-16 .These two models provided great outcome and thus we have used VGG-16 and we belive ResNet152 could be of better use compared with ResNet50 in image classification thus started our experimentation with those. The drawback is the number of images in dataset studied. There is no publicly available dataset which have high number of valid images to classify Mpox.

Contribution:

In our study, we advanced the existing body of knowledge on monkeypox classification by focusing on the effectiveness of various deep learning models, specifically ResNet152, DenseNet169, EfficientNet, and VGG-16. Building upon prior research that highlighted the utility of ResNet50 and VGG-16 for differentiating monkeypox from other infectious diseases, we aimed to assess whether a more robust model, ResNet152, could yield improved classification outcomes.

Our research employed data augmentation techniques to enhance the limited dataset available for monkeypox image classification, which is a common challenge in this domain. We systematically evaluated the performance of the aforementioned models and found that ResNet152 demonstrated superior accuracy, achieving a remarkable 90% in classification tasks. This outcome suggests that DenseNet169 and ResNet152,with proper hyperparameter tuning could significantly enhance the diagnostic capabilities for monkeypox, thereby facilitating earlier detection and timely medical intervention.

Furthermore, while our study recognized the limitations associated with the availability of a large and diverse dataset, our findings contribute to the ongoing discourse on the need for innovative diagnostic tools in regions where traditional healthcare resources may be lacking.

Proposed Methodology:

The proposed methodology involves data collection, preprocessing, and the application of various deep learning techniques such as ResNet , DenseNet, EfficientNet, Visual Geometry Group. The models are evaluated based on their performance in detecting diseases. We have used four metrices namely accuracy, precision, F1 score and Recall

FORMULA

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

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

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

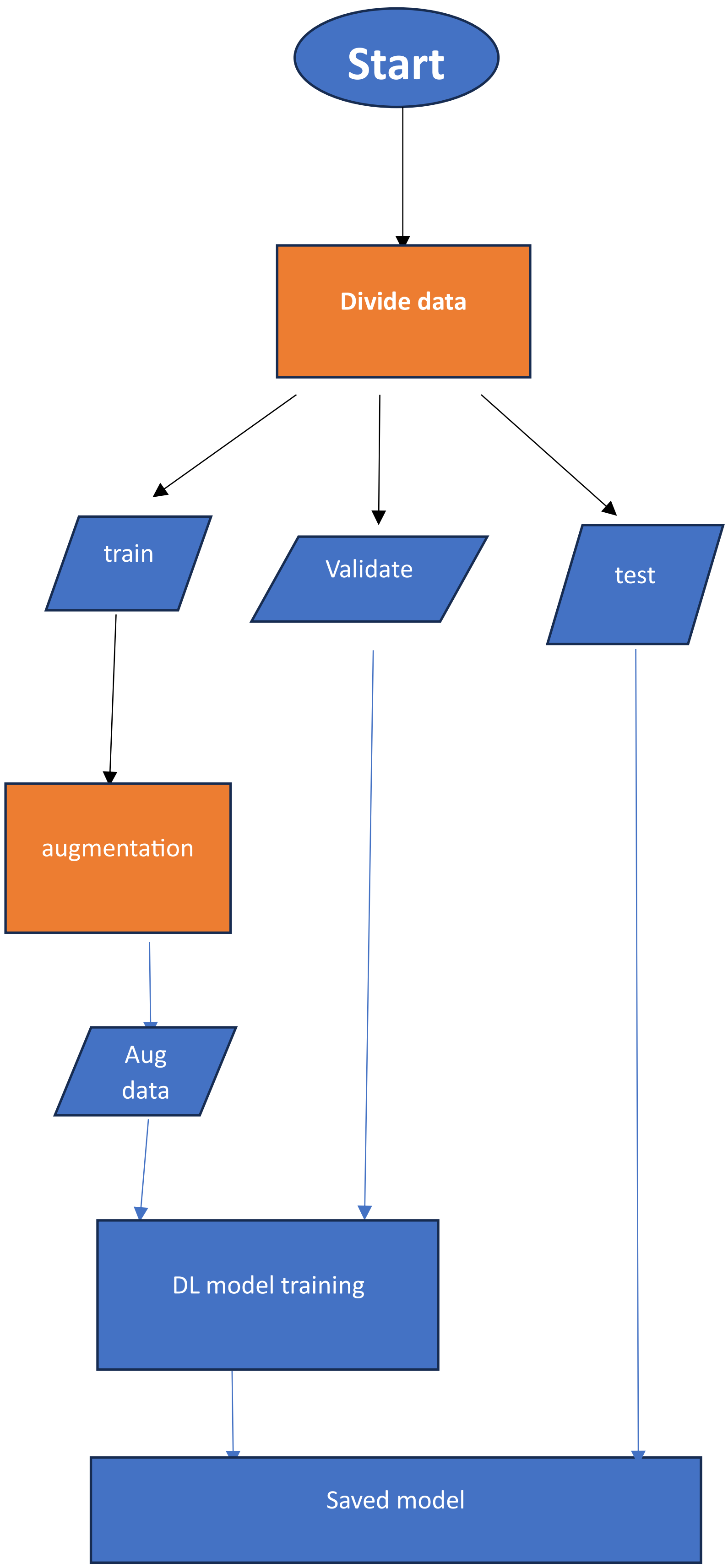
$$F1\ SCORE = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

FLOW Chart

The following flow chart depicts the experimental procedure



Experimental setup

The experiments are conducted using Python ,in Jupyter Notebook and VScode . Performance metrics such as Accuracy ,Precision,Recall, F1-score are used to evaluate the models

1)data collection:

We have used Mpox Skin Lesion Dataset Version 2.0 (MSLD v2.0) from kaggle for dataset of Mpox and non Mpox images

IMAGE CLASS	NUMBER OF IMAGES
MPOX IMAGE	284
NON MPOX IMAGES	130
TOTAL	414

MPOX	NON MPOX (Chickenpox)	NON MPOX (Measles)

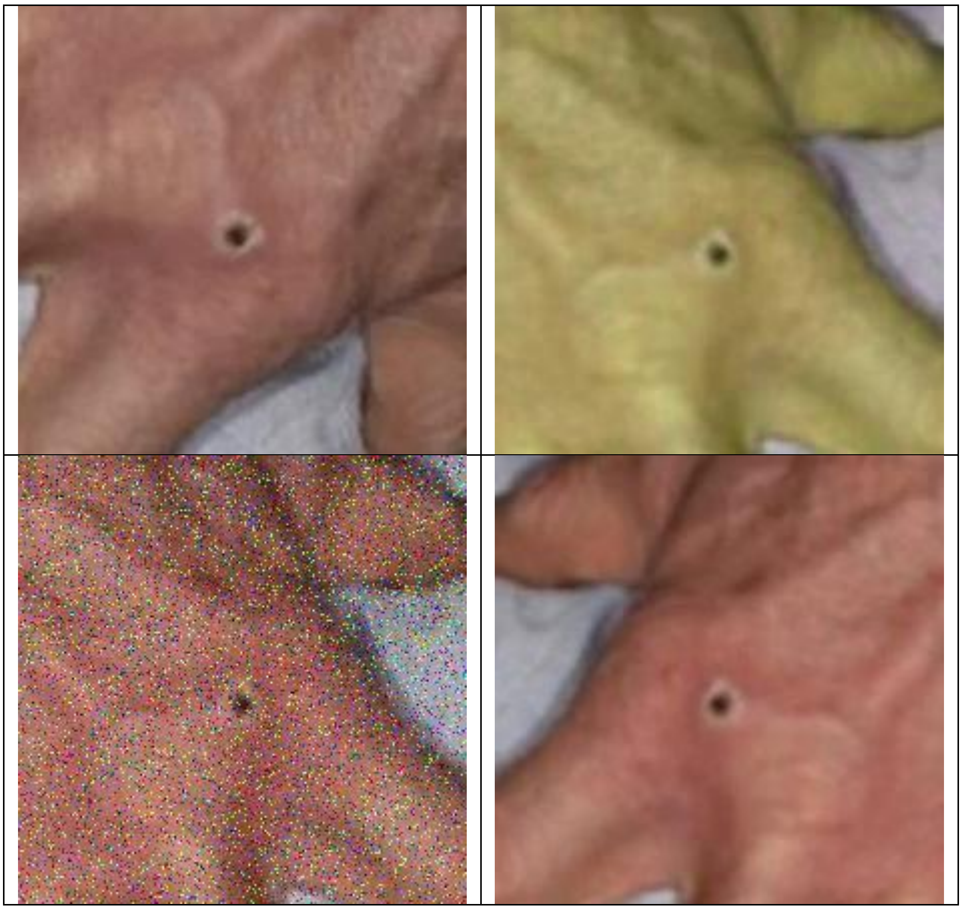
2) **Dataset Screening**

The collected skin images were processed through a 2- stage screening process. First, the out-of-focus, low-resolution, and low-quality images were discarded, and only the unique images that satisfy the quality requirements were selected. Next, the images were cropped to their region of interest and resized to 224 224 pixels while maintaining the aspect ratio

3)**Data split and Augmentation**

Data iis split into 70:20:10 ratio while to assist in the classification task and improve the generalizability of the learning algorithms, several data augmentation methods, including rotation, translation, reflection, shear, hue, saturation, contrast and brightness jitter, noise, and scaling, were applied to the training data and given to the model along with validation and test dataset

CLASS OF TRAINING IMAGE	NO OF TRAINING IMAGES AFTER AUGMENTATION
MPOX	2828
NON MPOX	1218



Models

1)Resnet152

ResNet152 is a deep convolutional neural network architecture that consists of 152 layers and is part of the ResNet family, which introduced the concept of residual learning. This architecture leverages skip connections, allowing the network to bypass certain layers, which helps mitigate the vanishing gradient problem commonly associated with training deep networks. One of the significant positives of ResNet152 is its ability to capture complex features in images, making it highly effective for various tasks such as image classification, object detection, and segmentation, often achieving state-of-the-art performance on benchmark datasets like ImageNet. Additionally, its design enables effective transfer learning, allowing pretrained models to be fine-tuned for specific applications. However, despite its strengths, ResNet152 has some drawbacks. The depth of the network makes it computationally intensive, requiring substantial memory and processing power for training and inference. Moreover, the complexity of the architecture can lead to overfitting, especially when applied to smaller datasets without sufficient regularization techniques. Thus, while ResNet152 is a powerful tool in computer vision, careful consideration of its computational demands and the nature of the dataset is essential for optimal performance.

2)DenseNet121

DenseNet121 is a deep convolutional neural network (CNN) architecture that stands out for its dense connectivity pattern, where each layer receives input from all preceding layers, creating a highly connected architecture. This design improves information and gradient flow throughout the network, which enhances training efficiency, reduces the risk of vanishing gradients, and allows for a more compact model with fewer parameters compared to traditional CNNs of similar depth. DenseNet121 is especially useful in applications like medical imaging, object detection, and feature extraction, as it captures intricate patterns efficiently with a smaller model size.

However, the dense connectivity also has its downsides. It requires more memory during training due to the additional feature maps being passed to subsequent layers, which can lead to slower training times on memory-constrained hardware. Additionally, for very large datasets, the network may become more prone to overfitting because of the increased model complexity.

2)Densenet169

DenseNet-169 is a deeper version of the DenseNet architecture compared to DenseNet-121, with 169 layers as opposed to 121. Like DenseNet-121, it uses dense connectivity, where each layer is connected to all previous layers within the same dense block. This design enhances gradient flow and promotes feature reuse, making it highly efficient in parameter usage. With its increased depth, DenseNet-169 can learn more complex features and is often preferred in applications that demand high accuracy, such as medical imaging, detailed object detection, and fine-grained classification tasks.

The primary advantage of DenseNet-169 over DenseNet-121 is its greater capacity for modeling intricate patterns, thanks to its deeper structure. However, this increased depth also comes with higher memory requirements and longer training times, as the additional layers produce more feature maps and require more computations. This can be a disadvantage on memory-constrained hardware or in situations where real-time inference is necessary. Furthermore, DenseNet-169, with its larger model size, may risk overfitting when applied to smaller datasets without sufficient regularization techniques. Despite these trade-offs, DenseNet-169 is an excellent choice when accuracy is prioritized over model size and computation speed.

3) EfficenetB0

EfficientNetB0 is the foundational model in the EfficientNet family, designed for image classification tasks with an emphasis on optimizing accuracy while minimizing computational costs. It employs a unique scaling method that uniformly adjusts the depth, width, and resolution of the network, resulting in state-of-the-art performance with significantly fewer parameters and FLOPS compared to traditional convolutional neural networks. The model's efficiency makes it ideal for deployment on resource-constrained devices, and its scalability allows for easy adaptation to larger variants, enhancing its applicability across various tasks. However, the architecture's complexity can pose challenges in implementation and tuning, and while it is efficient, inference speed may be slower than simpler models, particularly on limited hardware.

4)**VGG-16**

VGG-16 is a deep convolutional neural network known for its effective image classification capabilities. Comprising 16 weight layers that primarily utilize small 3××3 convolutional filters, it excels at capturing intricate features in images. Its architecture is straightforward, allowing for ease of implementation and fine-tuning. The key advantages of VGG-16 include its strong performance on benchmarks like ImageNet and its ability to learn hierarchical features. However, it has drawbacks, such as high computational and memory requirements, making it less suitable for resource-constrained environments. Additionally, its large number of parameters can lead to overfitting, necessitating careful regularization during training.

Hyperparameters of the models

Hyperparameters play a crucial role in machine learning models, as they determine the learning rate, batch size, number of epochs, and layer configurations, ultimately influencing the model's performance and ability to generalize to unseen data.

Hyperparameter	ResNet152	DenseNet121	DenseNet169	EfficenetB0	VGG-16
Batch size	32	32	32	32	32
Input Shape	(224,224,3)	(224,224,3)	(224,224,3)	(224,224,3)	(224,224,3)
Epochs	20	20	20	15	15
Initial Learning rate	1e-4	1e-4	1e-4	1e-4	1e-4
Optimiser	Adam	Adam	Adam	Adam	Adam
Unfreeze layers	Top 100	Top 20	Top 20	None	Top 5
Early stop paitence	5	5	5	5	5

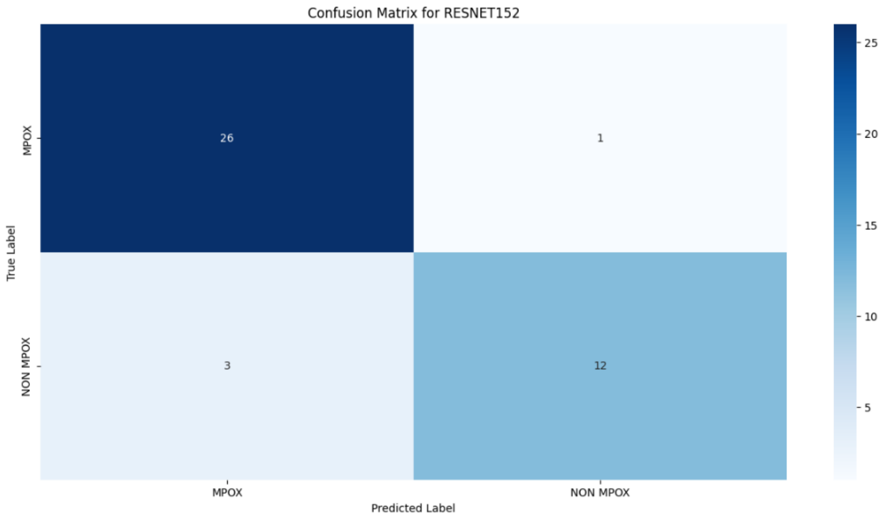
Activation function	Softmax	Softmax	Softmax	Softmax	Softmax

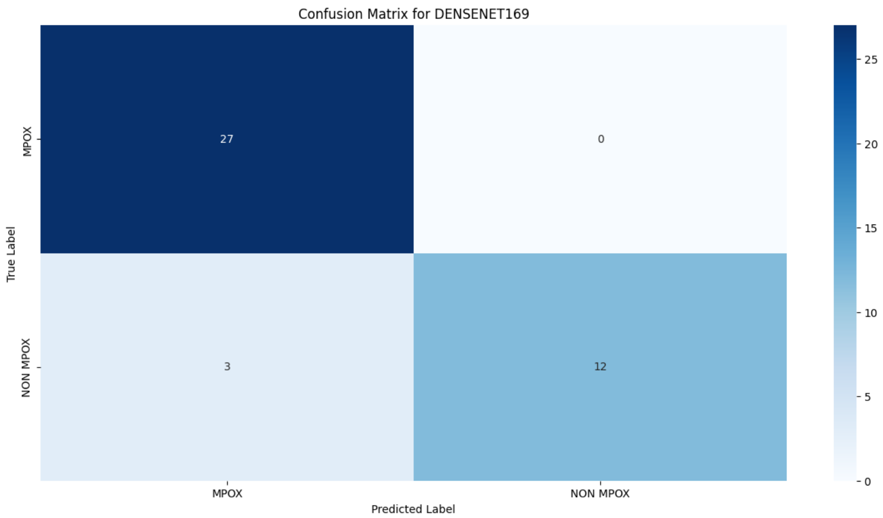
Results

Model	**Accuracy**	**Precision**	**Recall**	**F1 score**
RESNET152	**90.47**	**90.47**	**90.47**	**90.3**
DENSENET121	**85.71**	**88.31**	**85.71**	**84.6**
DENSENET169	**92.85**	**93.57**	**92.85**	**92.6**
EFFICENET	**80.95**	**81.06**	**80.95**	**80.14**
VGG16	**78.57**	**78.79**	**78.57**	**77.33**

Confusion Matrix

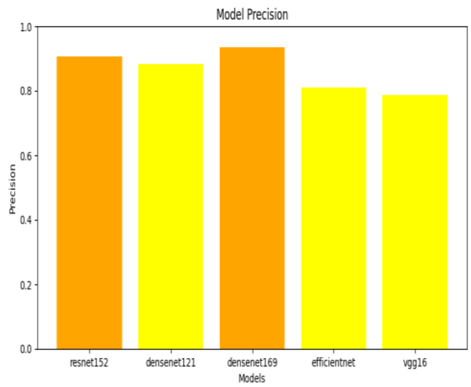
Top 2 models confusion matrix

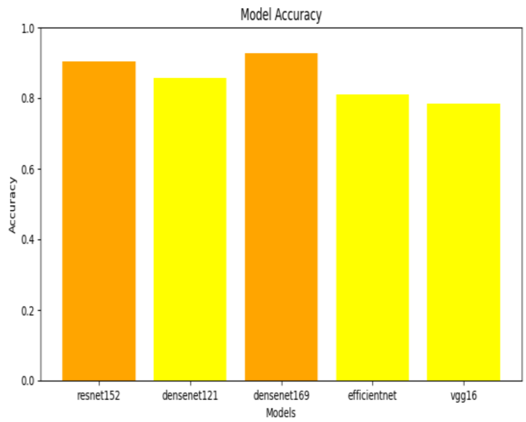

Confusion Matrix for RESNET152
MPOX
NON MPOX
True Label
26
1
3
12
MPOX
NON MPOX
Predicted Label
25
20
15
10
5

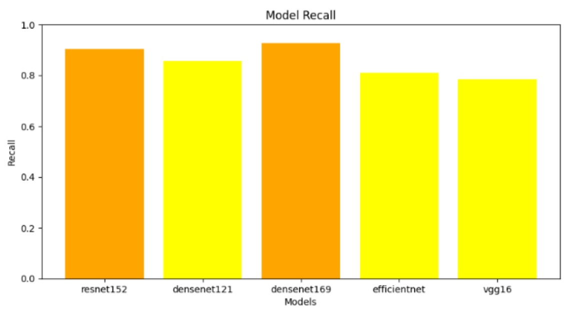

Confusion Matrix for DENSENET169
MPOX
NON MPOX
True Label
27
0
3
12
MPOX
NON MPOX
Predicted Label
25
20
15
10
5
0

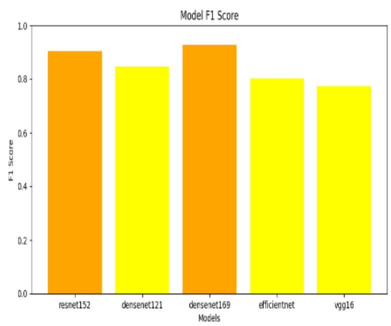
Evaluation

The comparision of Evaluation metrics for the 5 models are as follow









Future Scope

The future of AI in Mpox diagnostics and broader medical applications presents numerous opportunities for advancement. Potential directions for further research include:

1. Dataset Expansion: To improve the generalizability and robustness of our models, we recommend the development and curation of larger, more diverse datasets that include variations in skin tone, lesion presentation, and demographic factors.

2. Transfer Learning: Exploring transfer learning techniques by fine-tuning models pre-trained on large datasets (like ImageNet) can lead to enhanced performance, especially in scenarios with limited labeled data.

3. Real-time Deployment: Future work could focus on developing mobile or web applications that leverage these models for real-time diagnosis, enabling healthcare professionals to perform assessments in remote or resource-limited settings.
4. Integration with Other Modalities: Combining CNN-based image analysis with other diagnostic tools (such as clinical data and patient history) could create a comprehensive decision support system for healthcare providers.
5. Longitudinal Studies: Conducting longitudinal studies to monitor the performance of these models in clinical settings will be crucial in understanding their real-world efficacy and reliability.
6. Research on Other Skin Lesions: Expanding the research to classify and differentiate other similar viral infections and skin conditions could further enhance the utility of CNNs in dermatological diagnostics.

By addressing these areas, future research can contribute to a more effective and accessible healthcare system, ultimately leading to better disease management and improved patient care outcomes.

Conclusion

In this study, we explored the effectiveness of various convolutional neural network (CNN) architectures for the classification and detection of Mpox lesions using the Mpox Skin Lesion Dataset Version 2.0. Our findings demonstrate that deep learning models, particularly ResNet152 and DenseNet169, significantly enhance diagnostic accuracy for Mpox, achieving accuracies of 90.47% and 92.85%, respectively. These results underscore the potential of CNNs in distinguishing Mpox from visually similar diseases, such as chickenpox and measles, which can aid healthcare professionals, especially in regions where access to confirmatory PCR testing is limited.

The integration of data augmentation techniques also played a crucial role in addressing the challenges posed by the limited availability of high-quality imaging data. By improving the model's robustness and generalizability, we have provided a framework that could be adapted to other medical imaging tasks where dataset scarcity is a concern. Our work contributes to the growing body of knowledge advocating for AI-driven tools in clinical settings, emphasizing their role in early detection and improved patient outcomes.

References

[1] Ali, Shams Nafisa, et al. "A web-based mpox skin lesion detection system using state-of-the-art deep learning models considering racial diversity." *Biomedical Signal Processing and Control* 98 (2024): 106742.

[2] Singh, Rahul, et al. "A Novel CNN Model for Monkeypox Classification and Detection." *2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC)*. IEEE, 2024.