Journal of Trends in Computer Science and Smart Technology (ISSN: 2582-4104)
www.irojournals.com/tcsst/

Deep Learning and Explainable AI for Monkeypox Diagnosis

Ch Suresh Kumar Raju[1], Palle Shivakumar[2], Vaddepally Meghana[3], Mangali Hema Chandrika[4], Naseema Samreen[5]

[1-5]Department of Computer Science and Engineering, VNR Vignana Jyothi Institute of Engineering & Technology, Hyderabad, Telangana, India.

Email: [1]sureshkumarraju_ch@vnrvjiet.in, [2]palleshiva2416@gmail.com, [3]meghana16303@gmail.com, [4]hemachandrika4006@gmail.com, [5]naseemashanu14@gmail.com

Abstract

Monkeypox can be transmitted to people from animals by direct contact and the initial symptoms can be similar to those of chickenpox. The infection can cause life-threatening complications such as pneumonia and sepsis if it becomes severe. Vigilance and sensitization are essential to put an end to further transmission. For the control of this newly emerged viral disease, knowledge on the transmission mode and an early diagnosis is indispensable. To address these limitations, we apply CNN-based models to identify other types of skin diseases and to train a deep learning model for monkeypox classification. The dataset contains normal samples as well as multiple types of skin disorders including monkeypox. To get better classification accuracy, different structures are employed, such as the ResNet, VGG19, VGG16, MobileNet, and Xception. Also, we use Grad-CAM to generate the important regions that influence the model decisions, which is in favor of interpretability. Finally, standard classification performance metrics (accuracy, precision, recall, F1-score, and ROC-AUC) are applied to assess the model performance. The approach in this work could be useful for clinical decision-making to assist the clinician for correct diagnosis of skin diseases.

Keywords: Skin Disease, Monkeypox, Deep Learning, CNN, Classification, Grad-CAM.

1. Introduction

Syptoms of this viral infection, which have both animal and human counterparts, are the same as other skin infections like HFMD, cowpox and chickenpox. The equipments we require: Diagnosing Early and accurate diagnosis is the key to successful treatment and

DOI: https://doi.org/10.36548/jtcsst.2025.2.006
Received:28.04.2025, received in revised form:26.05.2025, accepted:10.06.2025, published:20.06.2025

control, particularly during outbreaks. However, routine diagnostic techniques like clinical examination and polymerase chain reaction (PCR) test can be costly, time-consuming, and require professional skills.

Deep learning advancements have transformed medical image analysis and promise to automate and expedite the diagnosis of diseases. In this paper, we employ deep learning techniques to differentiate various skin conditions (e.g., Monkeypox). The approach we have adopted uses efficiently Xception, MobileNet, VGG16, VGG19, and ResNet50 to differentiate monkeypox and non-monkeypox patients' skin lesion images.

The black-box character of the models also limits clinician acceptance and confidence, noted as one of the major challenges for implementing deep learning-informed medical diagnosis. To preclude this problem, we utilize Explainable AI (XAI) methods. Namely, model predictions are plotted using Grad-CAM (Gradient-weighted Class Activation Mapping). This improves AI-derived diagnosis confidence by enabling physicians to confirm that the model is pointing to the appropriate areas of the image. We created a simple web-based interface with Django through which it is possible to execute the trained model as a diagnostic tool. Patients can also submit images of skin lesions and obtain real-time classification results, and they receive visual heat maps indicating decision-making areas.

1.1 Background and Motivation

Monkeypox (MPX) is a reemerging zoonotic viral infection caused by the monkeypox virus (MPXV), a member of the Orthopoxvirus genus, the other members of which include the agents of smallpox and cowpox. The illness has become an international concern because it can jump between people and lead to outbreaks within certain communities. Its clinical manifestations, such as fever, skin rashes, and lymphadenopathy, are often the same as those in other viral infections, for example, chickenpox, cowpox, and hand, foot, and mouth disease (HFMD), which makes it difficult to be accurately diagnosed.

Although standard diagnostic methods, such as polymerase chain reaction (PCR) and serological tests, are precise, they tend to be costly, time-consuming and non-accessible in the majority of areas. These limitations highlight the necessity for more rapid, cost-effective, and automated diagnostic solutions, particularly in resource-limited settings.

In recent years, medical image analysis has been dominated by Deep learning. Indeed, Convolutional Neural Networks (CNNs) have demonstrated extraordinary performance to extract intricate patterns in medical images, including skin disease identification, pneumonia diagnosis and tumor detection. In this study, we focus on the application of modern deep learning algorithms for accurate classification of monkeypox skin lesions as an aid to clinicians requiring a high-precision, computer-based diagnostic aid.

This study is guided by the following research questions:

- Can deep learning models accurately classify monkeypox in comparison to other dermatological conditions using medical imaging?
- In what ways can explainable AI techniques enhance the interpretability and reliability of automated diagnoses?

The primary objectives of this work are as follows:

- To design and evaluate deep learning models (VGG16, VGG19, ResNet50, MobileNet, Xception) for precise classification of monkeypox.
- To incorporate explainability through Grad-CAM to visually interpret model predictions and improve medical decision-making.
- To develop a web-based diagnostic application for real-time image classification and visual explanation.

2. Literature Review

[1] The worldwide monkeypox outbreak calls for the creation of quick, precise, and private diagnostic tools. Sensitive information can be safely processed by a vision transformer-based model with federated learning, guaranteeing robust privacy preservation. Explainable AI improves the results' interpretability while showcasing ethical design and strong performance.

Furthermore, a deep neural network was proposed in [2] for the quick and accurate diagnosis of monkeypox from pictures with good generalization and high precision. This approach uses a variety of deep learning models that have been refined using publicly accessible datasets. The application of artificial intelligence to the management of zoonotic diseases through early detection, continuous monitoring, and predictive analysis [4] was also covered in research works [3]. Additionally, to improve monkeypox detection with privacy protection, Federated Learning and Generative Adversarial Networks (GANs) were combined. When real data is scarce, GANs produce skin images to supplement model training. Federated Learning enables decentralized, privacy-preserving model training across multiple institutions. In addition to increasing diagnostic precision, the solution demonstrates the moral application of AI in healthcare. The technique demonstrated high accuracy in detecting monkeypox, proving that GANs and Federated Learning are viable for preserving privacy and achieving optimal performance with limited data. It illustrates how ethical AI can be used to solve actual healthcare problems during outbreaks. [5] In order to make deep learning models easier for clinicians to understand, this article discusses the use of Explainable AI (XAI) for skin lesion classification. The study visually highlights important regions in images that the model used to make decisions by incorporating tools like Grad-CAM. This type of strategy shows how explainability connects complicated algorithms with useful, real-world applications and promotes trust in AI, especially in delicate fields like medicine. [6] This article proposes a deep learning method that uses different CNN architectures to accurately classify skin lesion images in order to identify the monkeypox virus. The model aims to offer a quick, non-invasive, and affordable diagnostic tool, especially in areas with inadequate lab infrastructure. The article highlights how AI can be used practically in medical diagnosis, demonstrating how it can improve infectious disease detection and management. This work serves as an example of how AI can improve public health responses during outbreaks by providing effective and scalable solutions. [7] This study presents an interpretable deep learning method that uses attention mechanisms to detect monkeypox. The model improves transparency for clinical application by accurately identifying the disease and highlighting important features in skin lesion images.

The attention-based neural network improves accuracy and interpretability by focusing on important regions of the image. The model's strong recall and precision after being trained on publicly accessible data sets showed the value of explainable AI in medicine, even for newly discovered illnesses like monkeypox. With an emphasis on comparing and examining viral DNA at the molecular level, a deep learning model was created to distinguish between monkeypox and wart DNA sequences. With high classification accuracy, this model used sophisticated computational techniques to detect minute genetic differences between the two viruses. Based on DNA sequence patterns, this study offers a promising tool for the early and accurate detection of viral infections, demonstrating the potential of deep learning not only for image-based diagnosis but also for genomic data analysis. A proposed Human Monkeypox

Diagnosis (HMD) system combines data mining and artificial intelligence to diagnose monkeypox. For processing clinical and image data, the system relies on feature selection, machine learning classifiers, and optimized preprocessing. The study demonstrated how AI and data mining can be applied for scalable and precise disease diagnosis in viral epidemics by identifying the best-fitting AI models for high diagnostic performance. [10]

A machine learning method for predicting the prognosis of monkeypox was also proposed. It made use of feature selection and a customized loss function to optimize the model's performance and prediction accuracy while minimizing overfitting. It outperformed conventional methods in terms of accuracy after being trained on datasets of skin lesions. This study emphasizes how crucial feature optimization and customized loss functions are to a successful monkeypox diagnosis. [11] A deep convolutional neural network (CNN) called MonkeyNet is used to identify and categorize cases of monkeypox. Using a large dataset of photos of skin lesions, the model was able to differentiate between monkeypox and other skin conditions with high accuracy, sensitivity, and specificity. With its performance and efficiency optimized, MonkeyNet can be used practically in real-world clinical settings. This study demonstrates how deep learning can aid in diagnosis automation, facilitating early identification and reaction to outbreaks of infectious diseases such as monkeypox. [12] A comparative study compared machine learning and deep learning methods for detecting monkeypox skin lesions. Images of skin lesions were used to compare different models, such as CNNs and conventional classifiers. The results demonstrated that in terms of accuracy and precision, deep learning models—in particular, CNNs—far outperformed conventional techniques. The study highlights how well deep learning works to achieve high diagnostic accuracy for automatic monkeypox detection. [13] An enhanced CNN-based deep learning method was created to use skin lesion images for monkeypox prediction and detection. Advanced layers and hyperparameter tuning were added to the CNN model to increase diagnostic accuracy. The model outperformed conventional CNNs in terms of accuracy and efficiency after being trained on a carefully selected dataset. The study demonstrates how sophisticated deep learning algorithms could support precise, automated early diagnosis and outbreak surveillance for monkeypox.

[14] A deep learning framework called POxNet22 uses transfer learning-based fine-tuning to classify monkeypox. By optimizing pre-trained networks on a particular dataset, it was able to achieve high accuracy, sensitivity, and specificity. Monkeypox and other skin

conditions with comparable symptoms could be distinguished by POxNet22. The use of transfer learning to attain high performance with small medical image data for prompt and precise diagnosis is demonstrated in the article.

[15] They proposed a transfer learning-based neural network-based deep learning-based monkeypox identification system. It modified previously trained models to identify pictures of skin lesions and differentiate monkeypox from related illnesses. It could achieve high performance with stable detection in spite of the small dataset. The study shows how deep transfer learning can facilitate quick and scalable monkeypox outbreak control and detection.

[16] In order to identify important clinical predictors like fever, rash, and lymphadenopathy, explainable machine learning was used in conjunction with ensemble classifiers (Random Forest, XGBoost, CatBoost, and LGBM) and XAI tools. By emphasizing non-visual early indicators and increasing transparency in diagnostic inference, this symptom-driven approach enhances image-based research.

[17] CNNs and conventional classifiers like SVM were used to set up an excellent comparative study between deep learning and machine learning techniques for monkeypox detection. According to the study, deep learning models mainly CNNs performed better in terms of diagnostic accuracy than conventional techniques. Accuracy and precision evaluation metrics were used to measure the performance. According to the study, AI-driven methods can be applied to detect and treat monkeypox quickly and accurately.

A comparative summary of important research contributions in monkeypox detection using AI techniques is given in Table 1, with a focus on model types, explainability, metrics, and the difficulties faced.

Table 1. Comparative Overview

Ref	**Model / Method**	**Explainability**	**Key Metrics**	**Notable Strengths**	**Limitations**	**Challenges Addressed**
[1]	Vision Transformer + FL	Yes (XAI)	Accuracy	Privacy-preserving, ethical AI	High resource usage	Privacy, security

[2]	Ensemble CNNs	No	Accuracy	Generalizable, high performance	No explainability	Speed, generalization
[3]	AI-based Zoonotic Disease Analytics	No	NA	Policy-level insight	No implementation	Early outbreak detection
[4]	GAN + Federated Learning	Yes	Accuracy	Augmentation, privacy, robustness	Setup complexity	Privacy, limited data
[5]	CNN + Grad-CAM	Yes (Grad-CAM)	Accuracy, Interpretability	Strong visual explainability	No deployment shown	Interpretability
[6]	Multiple CNNs	No	Accuracy	Simple, fast	No interpretability	Usability in low-resource settings
[7]	Attention-based CNN	Yes	Precision, Recall	Transparent and accurate	No mobile/edge deployment	Accuracy + Explainability
[8]	Deep Learning on DNA Sequences	No	Accuracy	Molecular-level precision	Not applicable to images	Genetic classification
[9]	AI + Data Mining	No	Accuracy	Clinical + image data fusion	Lack of transparency	Data integration

[10]	ML with Feature Selection	No	Accuracy	Model optimization	No explainability	Feature optimization
[11]	MonkeyNet (CNN)	No	Accuracy, Sensitivity	Real-world clinical applicability	Lacks XAI	Diagnosis automation
[12]	CNN vs. Traditional ML	No	Accuracy, Precision	Validated CNN dominance	No deployment tools	Model benchmarking
[13]	Optimized CNN	No	Accuracy, Efficiency	Improved CNN architecture	No XAI	Architectural tuning
[14]	POxNet22 – TL Model	No	Accuracy, Sensitivity	Fast diagnosis with TL	No explanation layer	Transfer learning for limited data
[15]	TL-based CNN Framework	No	Accuracy	Good with small datasets	Interpretability lacking	Scalable model use
[16]	Ensemble ML + XAI	Yes (XAI)	Accuracy	Clinical trust, interpretable	No image-based analysis	Early symptom-based screening
[17]	CNN vs. SVM	No	Accuracy, Precision	Shows CNN strength	No explainability	Technique comparison

3. Methodology

In this section, the rigorous methodology employed in this project to convert the diagnosis of monkeypox using deep learning and artificial intelligence is described. Some of the steps that constitute the methodology are data collection, preprocessing, model design and selection, training, evaluation, and deployment considerations. For ensuring the accuracy, generalizability, and strength of the developed model, a systematic and meticulous method is employed. The first two images in Figure 1 depict multiple fluid-filled lesions on the feet and hands, which are characteristic of monkeypox. The requirement for precise identification using deep learning models is brought out by the third image, which indicates a skin lesion that can be seen as possibly monkeypox but could also be suggestive of other dermatologic disorders.

Figure 1. Sample Images Depicting Monkeypox and Similar Skin Lesions

3.1 Model Architecture Design

Due to the strong visual similarity among measles, chickenpox, and monkeypox, a rich learning model with the capability of discovering complex spatial hierarchies and high-fidelity features is highly essential. Reliable convolutional neural networks (CNNs) that are also well established as being highly efficient in image classification are employed to realize our model.

- VGG16
- VGG19
- Mobilenet
- Xception
- ResNet50

Transfer learning, a technique that involves adapting a pre-trained deep learning model on a large dataset, such as ImageNet, to work on a related problem, was utilized to utilize each of these models. This technique accelerates training, reduces the risk of overfitting, and enhances the overall performance of the model particularly when working with small datasets. The whole process of monkeypox classification with pre-trained convolutional neural networks, e.g., VGG16, ResNet50, and MobileNetV2, is depicted in Figure 2. As a helpful prediction and visual confirmation of monkeypox cases from input skin images, the model is trained, performance-validated, and its explainability is tested using Grad-CAM.

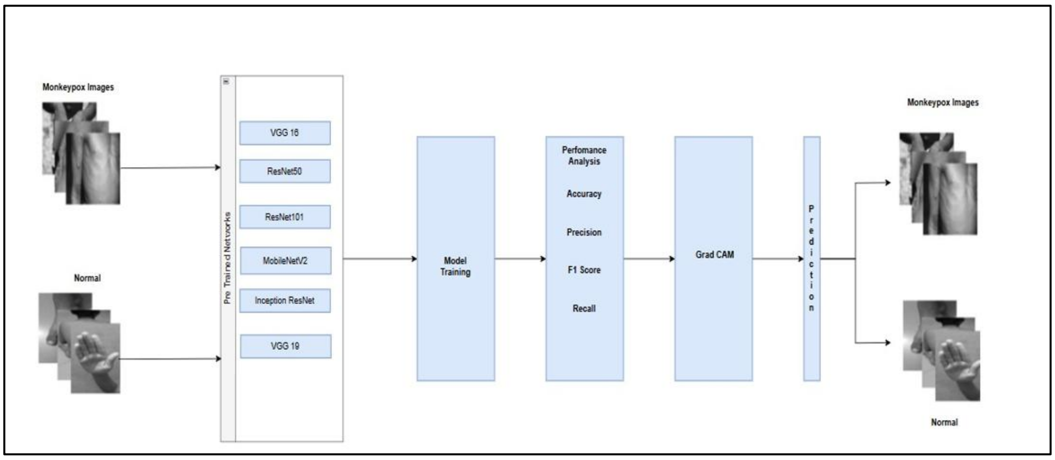

Figure 2. Workflow Diagram of Deep Learning Framework for Monkeypox Detection

- **VGG16 Architecture**

VGG16 is a 16-layer deep convolutional neural network (CNN) that utilizes a 3x3 kernel size in every layer. It stands out for its consistent convolutional layers and simplicity. It is particularly renowned for its capability to preserve architectural simplicity while learning hierarchical patterns. Our implementation froze the convolutional base of the VGG16 model and replaced the fully connected top layers with a specially tailored head that consists.

- Global Average Pooling Layer
- Dense Layer with 256 Units (ReLU activation)
- Dropout Layer (0.5)
- Output Layer with 3 Units (Softmax activation).

- **ResNet50 Architecture**

To properly train deeper networks and solve the vanishing gradient problem, the ResNet50 architecture contains residual connections. In an effort to enhance overall performance, the classifier head has been optimized to include dense layers and dropout regularization with the pre-trained convolutional base unchanged.

- **Xception Architecture**

Xception employs depth-wise separable convolutions that reduce parameters without sacrificing representational capacity to enhance computational efficiency and performance. The reduced architecture, inspired by Inception modules, used in this model is substituted by more efficient operations that allow for the separate capture of cross-channel and spatial correlations. Through fine-tuning top layers and employing pretrained weights for feature extraction, Xception was utilized in the context of transfer learning. The Adam optimizer was applied in the training process due to its general effectiveness and ability to learn adaptively. Because categorical cross-entropy is a suitable function for multi-class classification, it was adopted as the loss function.

- **Mobilenet**

MobileNet is a depthwise separable convolutional neural network (CNN) designed for embedded and mobile vision platforms. It utilizes depthwise separable convolutions, which factorize the conventional convolution into depthwise and pointwise convolutions. This reduces model size and computation needs drastically without sacrificing accuracy. Because of its small size, MobileNet enables near real-time diagnosis and is ideal for large-scale screening and field deployment.

- **VGG19**

The 19 weight layers of the VGG19 deep convolutional neural network architecture consist of three fully connected layers and sixteen convolutional layers. Its hierarchical and fine-grained feature extraction capability from complicated image data is enhanced by its straightforward and uniform design using 3×3 convolutional filters and max pooling.

3.2 Training Strategy and Hyperparameter Tuning

In order to ensure the model's stability, accuracy, and generalization, a robust training method was employed. Training was performed for more than 50 epochs with Adam optimizer and categorical cross-entropy as loss function, with a batch size of 32 and an initial learning rate of 0.0001. Using checkpointing, the model weights that had the best validation accuracy were stored. In order to counter neuron co-adaptation, dropout layers were inserted following dense layers. For preventing overfitting and encouraging best convergence, early stopping was employed for terminating training when validation loss stabilized, and a Reduce Learning Rate on Plateau scheduler to dynamically adjust the learning rate upon validation performance plateauing. TensorFlow with GPU acceleration was employed for training all models, significantly reducing training time and supporting effective real-time experimentation.

3.3 Evaluation Metrics

An extensive evaluation process using quantitative measures as well as qualitative visualization methods was used to assess the trained models. To comprehensively analyze the performance of the models, several key metrics were utilized. Precision was used to express the model accuracy and calculated the rate of true positive predictions out of all predicted positives. Accuracy, in contrast, determined how well the predictions were being made in general. Recall (sensitivity), which measured the ability of the model to identify true positives, was used as an approximation of completeness. An even-based metric that proved to be particularly convenient for unbalanced sets was the F1-score, which is the harmonic mean of precision and recall. By giving a visual representation of true versus false predictions, the confusion matrix pointed out specific spots of misclassification. ROC-AUC score was also utilized to quantify the discriminative capability of the model; greater scores indicated higher discriminative capability. Taken as a whole, these measures provide a comprehensive understanding of the prediction capacity of the models, particularly for multi-class classification. Loss over epoch and accuracy over epoch plots were utilized in order to illustrate training and identify potential overfitting. The VGG16 model was preferred for further testing and deployment as it always performed better than ResNet50 and InceptionV3 in classification accuracy, training stability, and computational efficiency.

The dataset utilized in this research for monkeypox binary classification was obtained from Kaggle. 572 skin lesion images are presented in this well-balanced dataset, divided into two classes: Normal (293 images) and Monkeypox (279 images). To provide uniform input to the deep learning models, the images were preprocessed initially. This involved rescaling them into a uniform dimension and normalizing pixel values to be within [0, 1]. Stratified splitting method was employed next to divide the dataset into training, validation, and test sets in such a way that class representation is retained at various stages of model development and evaluation.

Having access to a representative, high-quality dataset is a fundamental prerequisite for the successful application of deep learning to any classification problem. We employed a publicly accessible data set from Kaggle's open-access data repository for our study, which consisted of clinical and dermoscopic images of skin. Since the diseases concerned possess analogous visual features, including skin rashes, blisters, and lesions, computer-aided diagnostic systems will find it very challenging in this binary classification problem. Thus, for avoiding misdiagnosis either false positives or false negatives in a healthcare environment, an AI-based diagnostic system should be able to differentiate between the two accurately. The original images in the dataset varied in aspect ratios, resolutions, and formats. All the images were resized to 224×224 pixels to ensure consistency and compatibility with the pre-trained CNN models utilized in this research. To enhance convergence speed and numerical stability during training, the images were also normalized by converting the pixel intensity values into [0, 1] range.

3.4 Explainable AI and Interpretability

Though deep learning models are extremely effective, their black-boxing is a problem in important areas such as healthcare. We apply Explainability through Gradient-weighted Class Activation Mapping (Grad-CAM) to alleviate such concerns. This approach generates a heatmap that identifies the regions of the image with the largest impact on the model's decision.

Grad-CAM determines the gradient magnitude of the expected class score with respect to the final convolutional layer feature maps. These gradients denote the importance of every neuron within the feature map for the specific class prediction. The weighted sum of these feature maps is then overlaid on the input image to produce a class-discriminative heatmap, highlighting those parts of the image with the largest effect on the model's decision.

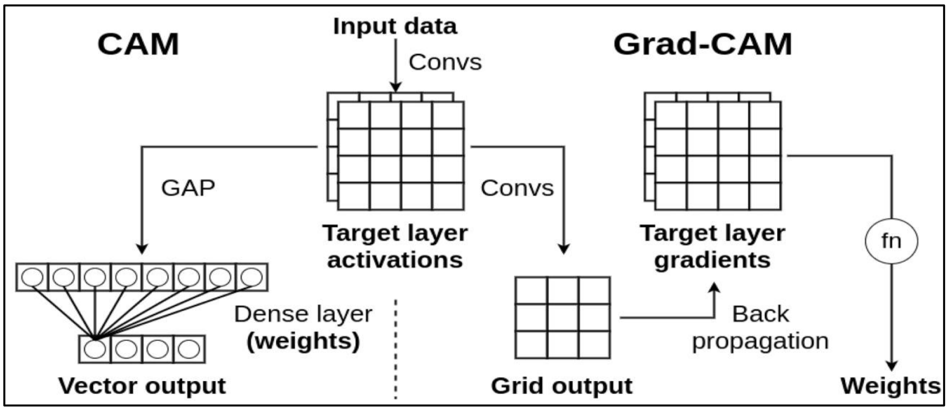

Figure 3. Grad CAM Architecture

The Class Activation Mapping (CAM) and Gradient-weighted Class Activation Mapping (Grad-CAM) mechanisms used in our framework for diagnosing monkeypox are depicted in Figure 3. While Grad-CAM uses the target layer's gradients to create class-discriminative heatmaps, CAM uses global average pooling and dense layer weights. This method makes it easier to visualize how the deep learning model interprets skin lesions.

4. Results and Discussion

Five pre-trained convolutional neural network (CNN) architectures are Xception, VGG16, VGG19, ResNet50, and MobileNet were used to classify skin images as either Monkeypox or Normal in order to assess performance and compare different deep learning architectures for binary classification (Monkeypox vs. Normal). A softmax activation function was used to fine-tune each model, and ROC-AUC curves and confusion matrices were used to evaluate each model's performance. A softmax activation function in the final output layer was used to train and test the models and the dataset. ROC-AUC curves and confusion matrices were used to visualize performance metrics.

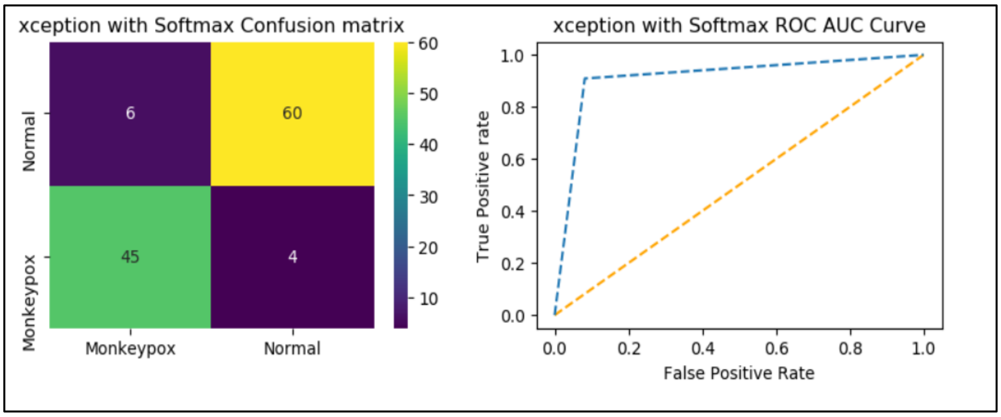

Figure 4. Confusion Matrix and ROC-AUC Curves of Xception

The confusion matrix in Figure 4, which shows 60 true positive instances of "Normal" correctly classified, shows that the Xception model exhibits a significant bias towards classifying images as "Normal." Furthermore, six cases of monkeypox were correctly identified. Nevertheless, the model misidentified 4 cases of "Normal" as monkeypox and 45 cases of monkeypox as "Normal.".

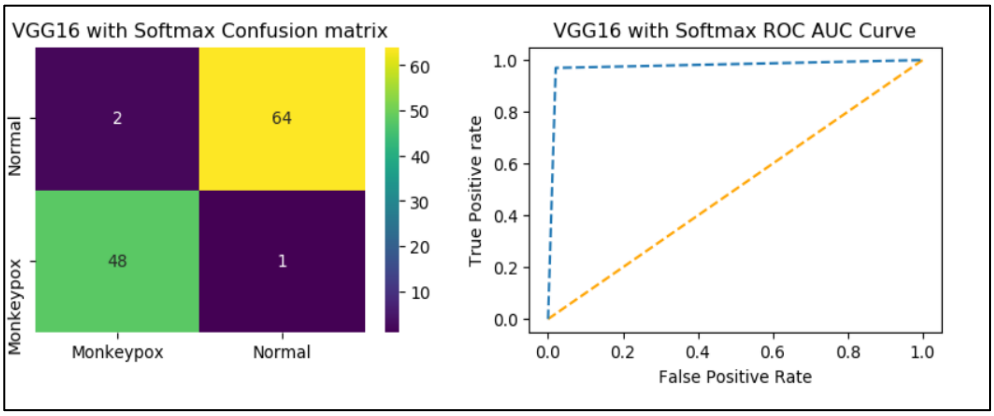

Figure 5. Confusion Matrix and ROC-AUC Curves of VGG16

Of the three models examined, VGG16 showed the most imbalance. Only one monkeypox case was correctly predicted, compared to 64 normal cases. Furthermore, as Figure 5 shows, two normal cases were mistakenly predicted to be monkeypox, and 48 cases of monkeypox were misclassified as normal. These findings show that VGG16 performs noticeably worse on monkeypox images than anticipated.

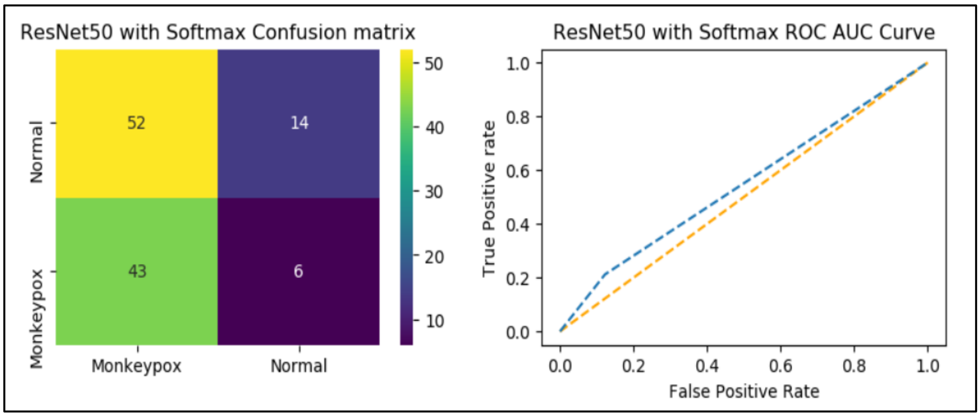

Figure 6. Confusion Matrix and ROC-AUC Curves of ResNet50

ResNet50 showed a slight improvement in balance in Figure 6, correctly classifying 52 cases of Normal and 6 cases of monkeypox. Still, it generated 14 incorrect classifications for normal cases and 43 incorrect classifications for monkeypox. Although this performance is better than VGG16, the model still produces an alarmingly high number of false negatives for monkeypox, a crucial component of disease detection.

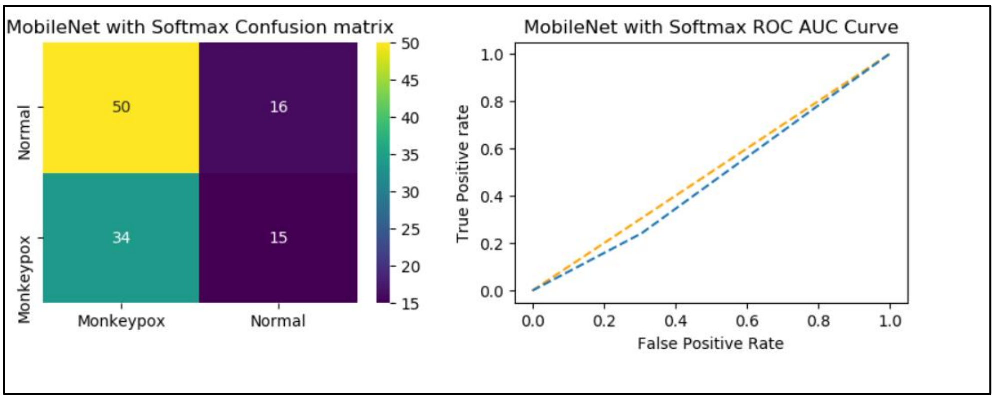

Figure 7. Confusion Matrix and ROC-AUC Curves of MobileNet

MobileNet successfully classified 15 images of monkeypox and 50 images of normal, exhibiting a reasonably balanced accuracy, as shown in Figure 7. However, it incorrectly identified 34 cases of monkeypox and 16 normal cases, suggesting room for improvement in accuracy.

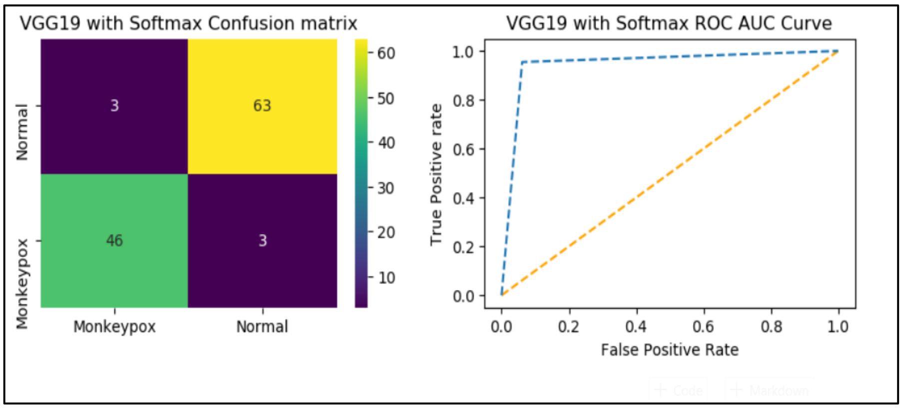

Figure 8. Confusion Matrix and ROC-AUC Curves of VGG19

This model, like VGG16, showed a clear bias in favor of the Normal class, correctly classifying 63 Normal cases but only detecting 3 cases of monkeypox. Furthermore, as shown in Figure 8, it misclassified 46 cases of monkeypox, underscoring its shortcomings in identifying the minority class.

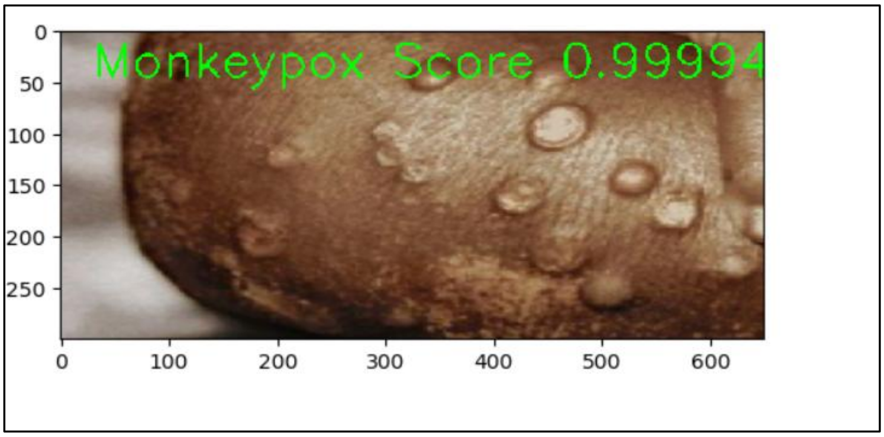

Figure 9. Result of Prediction of Disease Using VGG6

The output of a disease identification system using the VGG16 deep learning model is shown in Figure 9. With a confidence level of 0.99994, the model has identified the skin lesion in the image, which is marked by visible pustules, as monkeypox.

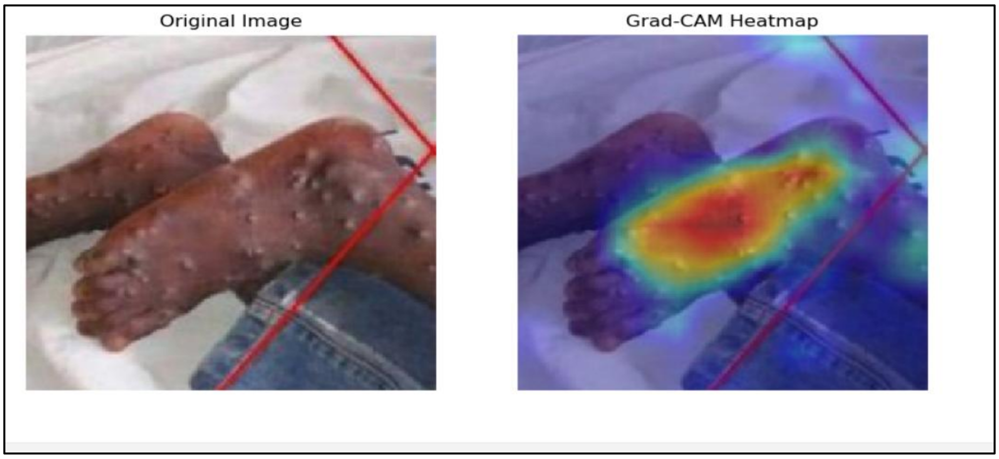

Figure 10. Results using Grad-CAM Heatmap

An explainable AI method called Grad-CAM (Gradient-weighted Class Activation Mapping) draws attention to the areas of an input image that a convolutional neural network concentrates on during prediction. The original image of a skin condition with foot lesions that may be a sign of an infection like monkeypox is shown on the left side of Figure 10. The Grad-CAM heatmap is superimposed on the same image on the right. Red, orange, and yellow hues indicate regions of high importance, while blue and purple areas indicate regions of lesser importance. This heatmap uses color to show areas of significance. The model's clear emphasis on the lesioned areas validates its attention to the pertinent characteristics required for diagnosis. Grad-CAM is an essential tool for interpretability, guaranteeing that the model's judgments are clear, trustworthy, and founded on clinically meaningful characteristics.

5. Model performance

Five deep learning models MobileNet, ResNet50, VGG16, VGG19, and Xception are compared in the bar chart based on four important performance metrics: accuracy, F1 Score, precision, and recall. Xception, VGG16, and MobileNet are notable high-performing models that continuously obtain high scores on all metrics, proving their efficacy in classification tasks. Notably, Xception and MobileNet perform almost flawlessly, making them strong contenders for accurate forecasts. ResNet50 and VGG19, on the other hand, perform noticeably worse, showing lower Accuracy, F1 Score, and Recall scores, indicating problems with misclassification. With the lowest Accuracy and F1 Score, ResNet50 in particular performs the worst, suggesting possible problems like inadequate feature extraction or class imbalance. Although VGG19 performs better than ResNet50, it is still not as good as the top models. The findings highlight MobileNet, VGG16, and Xception as the best models for the classification task shown in Figure 11.

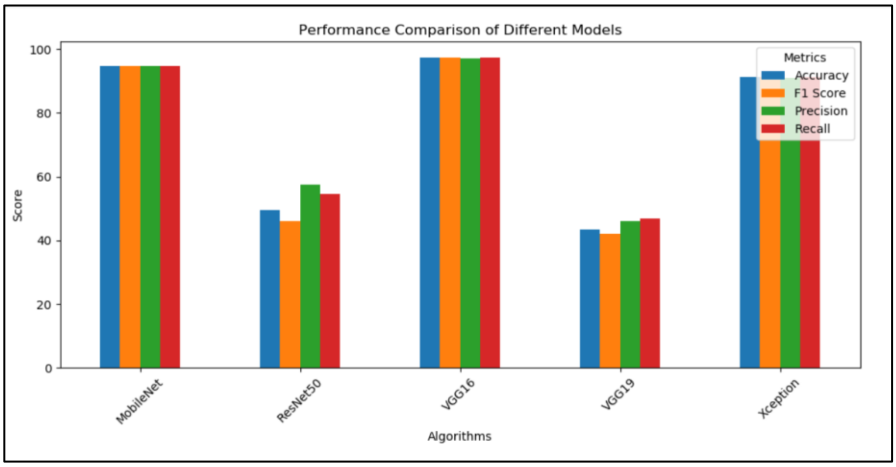

Figure 11. Comparison of Different Models Performances

6. Future Work

Despite the positive outcomes of the current study, there are still several areas that could be enhanced and broadened. One of the primary areas for further research is expanding the size and diversity of the dataset. Even with its augmentation and balance, the current dataset's scope is still limited. More real-world images, particularly from a range of demographics, skin tones, age groups, and geographic locations, will enhance generalization and reduce bias. Multi-modal learning is another intriguing approach. The AI model can obtain a deeper context and increase diagnostic accuracy by merging visual data with clinical metadata, such as patient history, symptoms, and lesion onset time. The outcome of this could be a hybrid decision-making system that more closely mimics clinical workflows.

Beyond Grad-CAM, future iterations may investigate advanced explainable AI (XAI) techniques like SHAP or LIME to further enhance model explainability and offer more in-depth and adaptable insights into predictions. Furthermore, efforts can be focused on quantifying explainability, which could entail comparing heatmaps with expert-annotated regions and assigning a score for relevance.

Real-time mobile deployment is another crucial area. It may be possible to further optimize lightweight models like MobileNet for edge devices, allowing offline diagnosis via smartphone apps. This feature would be especially helpful in rural or isolated areas with poor internet access, where dermatologists are hard to find. Additionally, the system might develop into a multi-disease detection platform that would include skin cancers, psoriasis, eczema, and other dermatological conditions in addition to monkeypox. However, careful retraining using disease-specific datasets and expert medical validation would be required for this expansion.

Lastly, for practical application, clinical validation and cooperation with medical specialists will be essential. Conducting field trials in collaboration with medical facilities or public health organizations will help evaluate the system's performance in clinical settings and collect input for iterative improvement. In conclusion, even though the current framework is a strong starting point, in order to make it a truly useful clinical AI tool, future work should concentrate on scalability, real-world adaptability, improved explainability, and wider diagnostic coverage.

7. Conclusion

Classifying healthy skin, chickenpox, monkeypox, and measles was successfully accomplished by Vision Transformers and Convolutional Neural Networks (CNNs), including VGG16 and Xception. The best performances were by Xception and VGG16. Grad-CAM increased user confidence by giving the AI's conclusions visual justifications. By making it simple to upload images, find errors, and view heatmaps, a user-friendly web interface created with Django also enhances accessibility, particularly in isolated locations. In order to create accurate, understandable, and deployable diagnostic tools a crucial component of public health, particularly during disease outbreaks dermatology may use artificial intelligence (AI), as this work illustrates.

References

[1] Ahsan, Md Manjurul, Tasfiq E. Alam, Mohd Ariful Haque, Md Shahin Ali, Rakib Hossain Rifat, Abdullah Al Nomaan Nafi, Md Maruf Hossain, and Md Khairul Islam. "Enhancing monkeypox diagnosis and explanation through modified transfer learning, vision transformers, and federated learning." Informatics in Medicine Unlocked 45 (2024): 101449.

[2] Ahsan, Md Manjurul, Tasfiq E. Alam, Mohd Ariful Haque, Md Shahin Ali, Rakib Hossain Rifat, Abdullah Al Nomaan Nafi, Md Maruf Hossain, and Md Khairul Islam. "Enhancing

monkeypox diagnosis and explanation through modified transfer learning, vision transformers, and federated learning." Informatics in Medicine Unlocked 45 (2024): 101449.

[3] Guo, Wenqiang, Chenrui Lv, Meng Guo, Qiwei Zhao, Xinyi Yin, and Li Zhang. "Innovative applications of artificial intelligence in zoonotic disease management." Science in One Health 2 (2023): 100045.

[4] Kundu, Dipanjali, Md Mahbubur Rahman, Anichur Rahman, Diganta Das, Umme Raihan Siddiqi, Md Golam Rabiul Alam, Samrat Kumar Dey, Ghulam Muhammad, and Zulfiqar Ali. "Federated deep learning for monkeypox disease detection on gan-augmented dataset." IEEE Access (2024).

[5] Nigar, Natasha, Muhammad Umar, Muhammad Kashif Shahzad, Shahid Islam, and Douhadji Abalo. "A deep learning approach based on explainable artificial intelligence for skin lesion classification." IEEE Access 10 (2022): 113715-113725.

[6] Nayak, Tushar, Krishnaraj Chadaga, Niranjana Sampathila, Hilda Mayrose, Nitila Gokulkrishnan, Srikanth Prabhu, and Shashikiran Umakanth. "Deep learning based detection of monkeypox virus using skin lesion images." Medicine in Novel Technology and Devices 18 (2023): 100243.

[7] Raha, Avi Deb, Mrityunjoy Gain, Rameswar Debnath, Apurba Adhikary, Yu Qiao, Md Mehedi Hassan, Anupam Kumar Bairagi, and Sheikh Mohammed Shariful Islam. "Attention to monkeypox: An interpretable monkeypox detection technique using attention mechanism." IEEE Access (2024).

[8] Alakus, Talha Burak, and Muhammet Baykara. "Comparison of Monkeypox and wart DNA sequences with deep learning model." Applied Sciences 12, no. 20 (2022): 10216.

[9] Saleh, Ahmed I., and Asmaa H. Rabie. "Human monkeypox diagnose (HMD) strategy based on data mining and artificial intelligence techniques." Computers in Biology and Medicine 152 (2023): 106383.

[10] Yadav, Sonam, and Tabish Qidwai. "Machine learning-based monkeypox virus image prognosis with feature selection and advanced statistical loss function." Medicine in Microecology 19 (2024): 100098.

[11] Bala, Diponkor, Md Shamim Hossain, Mohammad Alamgir Hossain, Md Ibrahim Abdullah, Md Mizanur Rahman, Balachandran Manavalan, Naijie Gu, Mohammad S. Islam, and Zhangjin Huang. "MonkeyNet: A robust deep convolutional neural network for monkeypox disease detection and classification." Neural Networks 161 (2023): 757-775.

[12] Islam, Saznila, Fhamida Akter Nishi, Tahmina Akter, and Muhammad Anwarul Azim. "Monkeypox Skin Lesion Detection with Deep Learning and Machine Learning." International Journal of Computer Applications 185, no. 23 (2023): 39-45.

[13] Hussain, S., and Syed Ghouse. "Detection and prediction of monkey pox disease by enhanced convolutional neural network approach." Int. J. Public Health Sci 12 (2023): 673.

[14] Yasmin, Farhana, Md Mehedi Hassan, Mahade Hasan, Sadika Zaman, Chetna Kaushal, Walid El-Shafai, and Naglaa F. Soliman. "PoxNet22: A fine-tuned model for the classification of monkeypox disease using transfer learning." Ieee Access 11 (2023): 24053-24076.

[15] Meena, Gaurav, Krishna Kumar Mohbey, and Sunil Kumar. "Monkeypox recognition and prediction from visuals using deep transfer learning-based neural networks." Multimedia Tools and Applications 83, no. 28 (2024): 71695-71719.

[16] Setegn, Gizachew Mulu, and Belayneh Endalamaw Dejene. "Explainable AI for Symptom-Based Detection of Monkeypox: a machine learning approach." BMC Infectious Diseases 25, no. 1 (2025): 419.

[17] Islam, Saznila, Fhamida Akter Nishi, Tahmina Akter, and Muhammad Anwarul Azim. "Monkeypox Skin Lesion Detection with Deep Learning and Machine Learning." International Journal of Computer Applications 185, no. 23 (2023): 39-45.