PoxDetect: Advancing Monkeypox Diagnosis with Machine Learning for Skin Lesion Classification

Md.Saiful Islam, Mahede Hasan and Sheikh Fazle Rabbi

October 25, 2024

PoxDetect: Advancing Monkeypox Diagnosis with Machine Learning for Skin Lesion Classification

***Abstract*— The monkeypox virus can infect nonhuman primates as well as humans. This dissertation examines the use of machine learning (ML) for early detection and classification of monkeypox disease caused by the varicella-zoster virus infects both humans and nonhuman primates. Emphasizes the importance of diagnosing monkeypox skin lesions in their early stages for effective treatment and disease prevention. Utilized images of monkeypox lesions obtained from Kaggle, then augmented to develop and test various test our own ML models. A basic mobile app was created to allow users to capture and send images for analysis by the models.Primarily focused on evaluating the effectiveness of different ML models,including ResNet50, InceptionV3, Xception, DenseNet121, and MobileNet. Results showed that the MobileNet and Xception models performed best with MobileNet achieving a mean accuracy 0.97 and an F-1 score of 0.968, while Xception achieved an accuracy of 0.986 and F-1 score of 0.98. Potential impact of ML in healthcare particularly in disease classification and identification.**

Keywords—Monkeypox, Machine learning, Skin illness, Resnet, Xception, DenseNet, MobileNet,Inception.

I. Introduction

The world recently emerged from the COVID-19 pandemic, which, despite being under control, has left significant lasting impacts. COVID-19, caused by the SARS-CoV-2 virus, first identified in Wuhan, China, in December 2019, rapidly spread globally, leading to over 609 million cases and 6.51 million deaths by september 2022. Although the COVID-19 pandemic persists, the focus has viral threat An infectious disease, monkeypox disease caused zoonotic monkeypox virus, a component of the Ortho poxvirus species, which also includes the smallpox-causing variola virus. Symptoms include fever, swollen lymph nodes, and a rash with blisters. The virus can spread from animals to humans through infected meat, bites, or scratches, and from person to person via contact with bodily fluids, contaminated surfaces, droplets, or airborne transmission. The ongoing monkeypox outbreak, which began in May 2022, has spread beyond Central and West Africa to over 105 countries, leading the WHO to declare it an international public health emergency (PHEIC). There were 57,669 confirmed cases as of september 2022. With the threat of monkeypox continuing into 2024, early detection is critical to preventing it from becoming a pandemic. Given its similarity to chickenpox and measles, computer-assisted detection may aid in early identification when PCR assays are unavailable [1].

II. RELATED WORKS

C. Sitaula and T. B. Shahi (2022) compared 13 pre-trained deep learning models for monkeypox detection, fine-tuning them with custom layers and evaluating them using Precision, Recall, F1-score, and Accuracy. Their ensemble approach achieved average metrics of 85.44% Precision, 85.47% Recall, 85.40% F1-score, and 87.13% Accuracy, outperforming existing methods and offering promising results for mass screening[1].

M. Altun, H. G¨ur¨uler (2023), Our study aimed to swiftly and accurately detect monkeypox through skin lesions using deep learning. We optimized transfer learning and hyperparameters in CNN models like MobileNetV3-s and Xception. The hybrid MobileNetV3-s model achieved top metrics: an AUC of 0.99, accuracy of 0.96, recall of 0.97, and F1-score of 0.98 demonstrating effective and rapid classification[2].

In GangHu's (2023), the proposed approach combines feature selection and a decision tree classifier, resulting in a sensitivity of 0.95, a specificity of 0.61, a p-value of 0.89, an N-value of 0.79, and a F1 score of 0.92. After the decision tree classifier settings were optimized, the methodology yielded an overall accuracy of 94.35%. Also, the Wilcoxon sign-rank test and an assessment of variance (ANOVA) were applied to assess statistical importance[3].

In M. M. Ahsan (2023), We evaluated our method using ten CNN models across three studies. Combining our approach with Xception achieved 77%-88% accuracy, while ResNet-101 excelled with 84%-99% accuracy. Our method was more computationally efficient than existing approaches and used LIME for explaining predictions and identifying key Monkeypox features[4].

In Gozde Yolcu Oztel's (2024), titled "Vision Transformer and CNN-Based Skin Lesion Analysis for Monkeypox," the proposed ensemble-based system achieved an accuracy of 81.91%, a Jaccard A 65.94% index, an 87.16% precision, a 74.12% recall, and a 78.16% F1-score were obtained. When measured against a range of criterion metrics, the system delivered competitive or superior results compared to existing literature[5].

III. PROBLEM STATEMENT

The main goal of this study is to develop a rapid method for identifying monkeypox through machine learning (ML), which can enhance virus surveillance and containment. By leveraging ML algorithms, the study aims to improve the accuracy, efficiency, and effectiveness of diagnosing skin disorders. ML can train computer algorithms to recognize patterns in skin lesion images, potentially automating diagnoses and saving time and costs compared to traditional methods like biopsies. ML's ability to detect subtle changes in lesions can lead to more accurate diagnoses, crucial for conditions like skin cancer where early detection is vital. Additionally, ML can assist in specific regimens for each patient, increasing results and lowering side effects. The study utilized models like DenseNet121, Inception V3, Xception, ResNet50, and MobileNet, all of which yielded moderate results and contributed to the research. Overall, ML has the potential to revolutionize dermatology by enhancing diagnostic precision and personalized treatment.

IV. OBJECTIVE

This Study Evaluates Utilizing ML to classify and diagnose A skin ailment called monkeypox is brought on by the varicella-zoster virus. Early and accurate identification of monkeypox skin lesions is crucial for effective treatment and reducing disease transmission. We collected monkeypox lesion images from Kaggle and examined the potential of ML classification models to categorize these images. Machine learning models are valuable for automating data categorization, which can enhance diagnostic accuracy for various diseases and streamline processes in fields like insurance fraud detection and sentiment analysis. These models also aid in image recognition and medical imaging by identifying specific objects and patterns. The study aims to evaluate the accuracy and effectiveness of ML models for monkeypox diagnosis and their potential impact on healthcare systems. Additionally, a mobile application will be developed to allow users to capture and upload images for analysis by the ML models, providing insights into practical applications of ML in the diagnosis and potting of illness classification.

V. PROPOSED STRATEGY

We used a Kaggle dataset for our research, which we divided into three categories: one for monkeypox images, one for measles and chickenpox images, and one for normal skin images.

Fig. 1. MonkeyPox and Others images [2]

To enhance the dataset and improve model training, we applied augmentation techniques such as rotation, shearing, and scaling. The dataset now includes 1,262 images of various conditions and 1,705 images of monkeypox[6].

TABLE 1. Sets for Training and Testing

Label	Train Set	Test Set
Monkey Pox	837	279
ChickenPox and Measles	594	198
Normal Skin	879	293

We split this dataset into training and testing sets, with about 80% allocated for training and the remaining 20% for testing.

A. DATA PREPROCESSING

The approach begins with preparing the data and constructing the embedding layer for ML models. Data cleaning is crucial, especially for image datasets where diseases like Monkeypox and Chickenpox can look similar. Preprocessing enhances accuracy and reliability by removing errors and inconsistencies, ensuring data consistency. While preprocessing includes geometric transformations such as rotation, scaling, and translation, its primary goal is to improve image quality by eliminating distortions and highlighting important features for further analysis.

B. DATA AUGMENTATION

Data augmentation involves modifying use previously collected information to increase its diversity without gather fresh data. This technique assists in limiting the formation of non-essential properties in neural networks. enhancing model performance[7].For monkeypox, chickenpox, measles, and normal skin datasets, augmentation included techniques such as flipping, rotating, cropping, resizing, adjusting brightness and contrast, adding noise, and jittering color channels.

TABLE 2. Configurations for augmenting data online

Rescale	1/255
Width shift	0.1
Height shift	0.1
Horizontal flip	False
Vertical flip	True
Optimizer	Adam
Batch size	16
Learning rate	0.1

These methods created a more varied dataset by altering object orientation, size, lighting, and position, which improved the model's ability to generalize and accurately classify new, unseen images. This resulted in a more robust model for disease detection.

C. WORKFLOW

We begin with Data Ingestion, which involves collecting the necessary data for model training. The raw data is then prepared for machine learning use through several preprocessing steps: Scaling adjusts the information to a conventional range, usually between 0 and 1; Flipping modifies an image's orientation if necessary; Cropping minimizes the size of photographs, while rotating them adjusts their angle as needed.The dataset is split after preprocessing, into two subsets: 20% for testing and 80% for training. The test set is used to evaluate the model's performance on data it has never seen before, while the training set is used to train the model to make prediction or classifies. Once training is complete.

The model is saved for future use[8].The final step involves integrating the saved model into a web application, which entails creating a user interface and creating backend code that loads the model to make use of its interface.

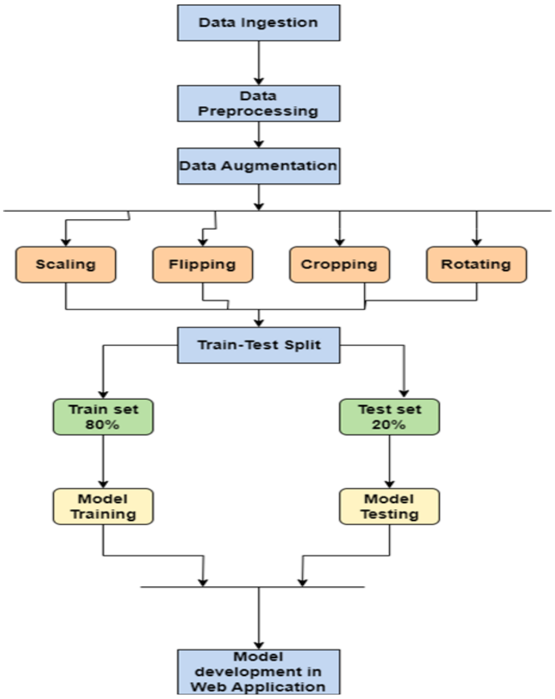


Fig. 2. Research Workflow

D. OVERVIEW OF THE MACHINE LEARNING MODEL

The proposed model utilizes image classification techniques on skin images collected through IoT devices to identify monkeypox. Machine learning aids physicians in more accurate diagnoses by providing quantitative analysis of lesions, potentially speeding up the clinical process. **Inception V3:** A convolutional neural network with several layers, including convolutions, max pooling, and fully connected layers. Inception V3 achieves 78.1% accuracy on the ImageNet dataset. Its loss function balances real loss with auxiliary losses to optimize both primary and secondary training objectives. The overall loss that the inception net experienced while being trained.

Real loss plus 0.3 × aux loss equals total loss. 1 plus 0.3 × aux loss 2.

DenseNet: Unlike typical CNNs, DenseNets require fewer parameters by eliminating redundant feature maps. Each layer in DenseNet has direct access to gradients based on the source image and loss function, making it efficient in training deep networks.**Xception Model:** An enhanced version of Google's Inception network, Xception uses depthwise separable convolutions to improve efficiency. The network consists of 71 layers and excels in extracting features from images.**MobileNet:** A lightweight a CNN created particularly for low-resource embedded and mobile gadgets. With MobileNet, depthwise separate convolutions, reducing computational costs while maintaining accuracy. Variants like MobileNetV2 and MobileNetV3 further optimize performance for mobile deployment.**ResNet-50:** A widely used ResNet architecture, ResNet-50 balances model complexity and performance, excelling in image classification tasks. Pre-trained on large datasets like ImageNet, it is particularly effective in transfer learning and representation learning, making it a top performer in classification challenges[9].

VI. IMPLEMENTATION RESULTS AND ANALYSIS

A. METRICS AND EVALUATION

This study evaluates model performance using metrics such as Accuracy, Precision, Recall, and F1-score, which are summarized in a classification report.

Accuracy measures the fraction of correct predictions out of all predictions, providing an overall assessment of the model's effectiveness.

Precision calculates the proportion of true positive predictions among all of the model's optimistic predictions. It indicates how well the model identifies positive instances while minimizing false positives.

Recall (or Sensitivity) is the proportion of true positive outcomes identified among all actual positive cases. It measures the model's ability to detect all positive instances while minimizing false negatives.

F1-Score is the harmonic mean of Precision and Recall, offering a balanced view of both metrics. It is particularly useful when class distributions are imbalanced, as it accounts for both false positives and false negatives, providing a more accurate evaluation of the model's performance.

$$\text{Accuracy} = \frac{Tp+Tn}{Tp+Tn+Fp+Fn}$$

$$\text{Precision} = \frac{Tp}{Tp+Fp}$$

$$\text{Recall} = \frac{Tp}{Tp+Fn}$$

$$\text{F1 Score} = 2 \times \frac{(Precision \times Recall)}{(\text{Precision} + Recall)}$$

In this case, a monkeypox patient was identified as Tp (True Positive).Monkeypox patient was appropriately identified (True Negative). False positive (Fp): A patient who was not infected with monkeypox was mistakenly recognized.

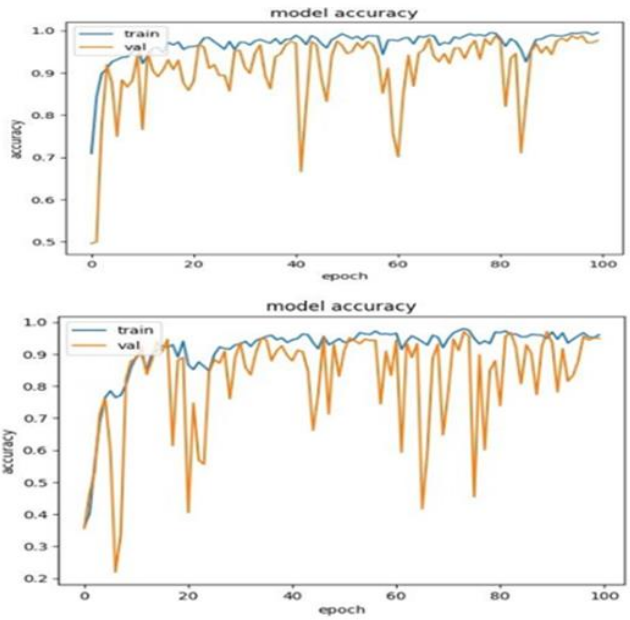


Fig. 3. Accuracy of MobileNet and InceptionV3 models for every era

MobileNet model, we achieved impressive results after normalizing pixel values during the data pretreatment phase and using the Adam optimizer[10].
With a precision of 96.9%, recall of 96.8%, and an F1 score of 96.8%, the model achieved an accuracy of over 97.6%.
InceptionV3 model, optimization techniques were employed to enhance model adaptation. This model achieved a precision of 94.9%, a recall of 94%, and an accuracy of above 94%.

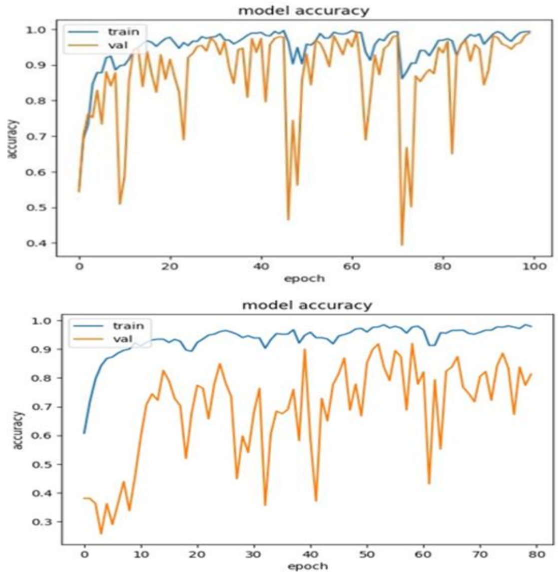


Fig. 4. Each epoch's model accuracy for Xception and Resnet50

Xception model, Our recommended approach revealed an excellent level of accuracy in differentiating between those who were and weren't affected by monkeypox if paired to the Ultra Inception structure of 98.6%. The model also achieved a precision, recall, with a 98.9% F1 score. A sophisticated depth-wise separable convolutional neural network called Xception was developed by Google as an advancement over traditional convolutional processes[11].
ResNet50 model, we began by normalizing pixel values and enhancing the dataset for better generalization. The model was trained using gradient descent optimization techniques, specifically the Adam optimizer, to achieve optimal results.

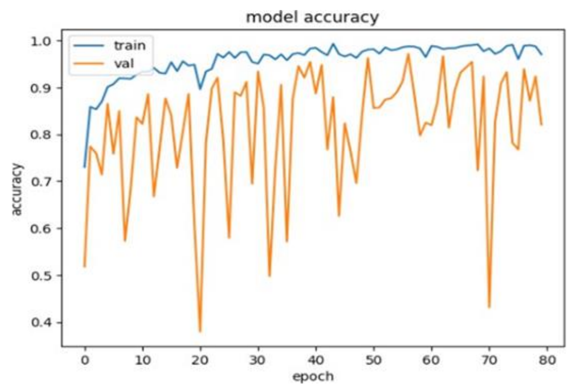


Fig.5. DenseNet121 model accuracy per epoch

DenseNet121 model, we utilized Adam optimizer during training. The model achieved a recall of 82%, precision of 84.3%, and accuracy of above 82%.

B. MODEL LOSS

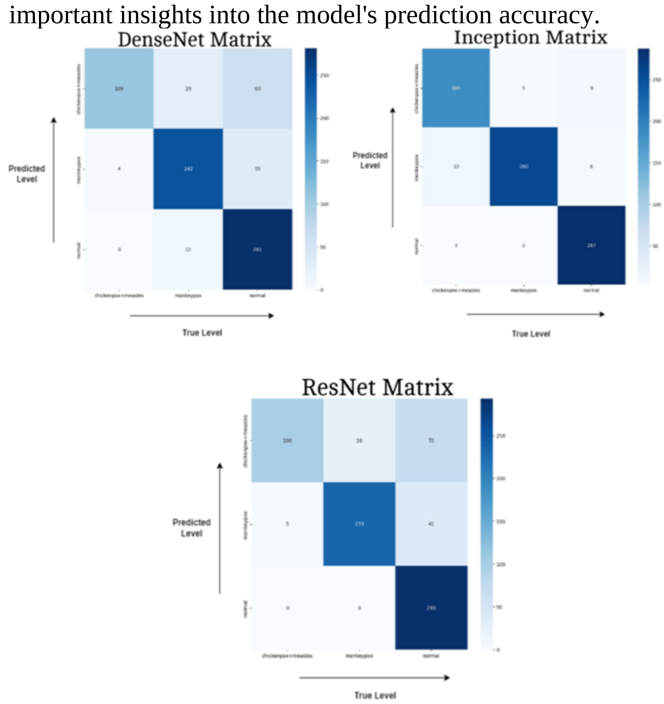
A loss function is crucial in evaluating how well a neural network can predict test data by comparing the predicted and actual output values. During the training process, the objective is to minimize this difference. Any modifications to the model's parameters are guided by the loss function, that calculates the difference between the model's predictions and the actual results.The average of all losses, known as the cost, guides the model to better represent the underlying data distribution. Machine learning models like neural networks often rely on maximum likelihood estimation to build accurate representations of data[12].
During both training and testing phases, monitoring model loss provides insights into how well the model is learning. The figure showing model loss helps to visualize this. For most models—except DenseNet121—the loss decreased over time, indicating effective learning. This reduction in loss is achieved through methods that modify the model's parameters iteratively, such as gradient descent to minimize loss and improve accuracy.

TABLE 3. Earlier computational outputs for the ML approaches employed research

Model	Precision	Recall	F_1 Score	Accuracy
Mobile net	0.96	0.96	0.96	0.97
Inception V3	0.94	0.94	0.94	0.94
Resnet 50	0.84	0.81	0.82	0.81
Densenet121	0.84	0.82	0.83	0.82
Xception	0.98	0.98	0.98	0.98

C. CONFUSION MATRIX

A confusion matrix is used to summarize predictions by showing the number of correct and incorrect forecasts for each class. It highlights which classes are often confused with others, aiding in the assessment model performance[13]. The matrix divides predictions separated into two groups: negative (normal) and positive (normal), with the positive class typically being underrepresented. Matrix helps evaluate the model's ability to minimize false positives (incorrectly predicting monkeypox) and false negatives (failing to predict monkeypox). True positives correctly identify monkeypox cases, while true negatives accurately classify non-monkeypox images. The confusion matrix provides important insights into the model's prediction accuracy.



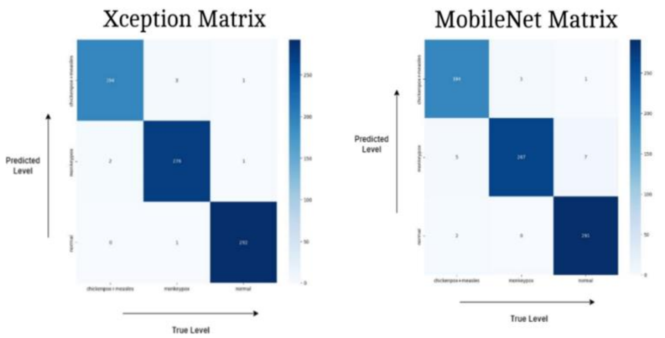


Fig. 6. Each model's confusion matrix

The evaluation utilized various models for the purpose prediction, with the Xception model being the most successful. The confusion matrix for this model highlights its effectiveness, particularly in never misclassifying a healthy person as having monkeypox[14].Evident in confusion matrix, where the row for healthy skin shows no instances of healthy skin being incorrectly identified as monkeypox.

D. RECEIVER OPERATING CHARACTERISTIC (ROC)

A ROC curve, or receiver operating characteristic, is a tool used to evaluate the performance of binary classification models across different thresholds. It plots the true positive rate (sensitivity) against the false positive rate (1-specificity) to show how changes in the classification threshold affect model performance. The curve helps in selecting an optimal threshold by illustrating the trade-off between true positives and false positives.

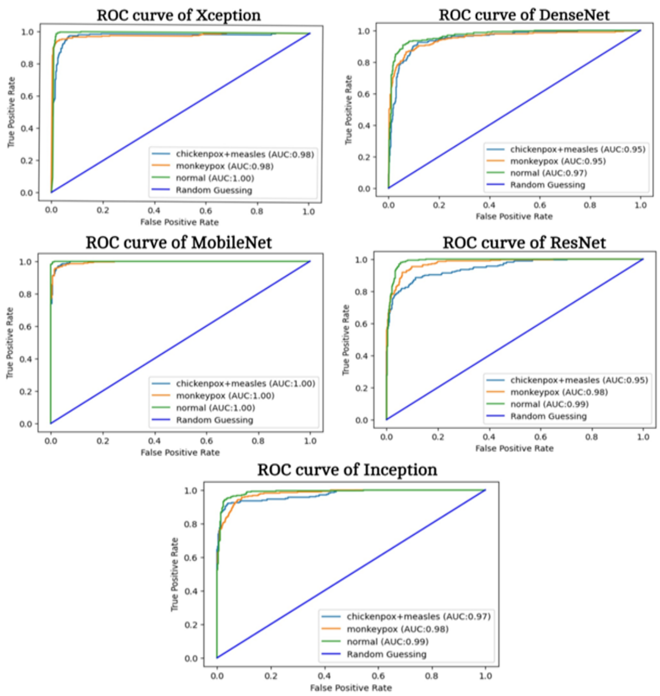


Fig.7. Every model's ROC curve

ROC curve region underneath the curve(AUC-ROC) quantifies the overall performance, with values ranging from 0 to 1, where higher values indicate better performance.In our analysis, we evaluated several CNN models using the ROC curve. MobileNet achieved an AUC of 1.00, indicating near-perfect classification[15]. Xception followed with an AUC of 0.98, demonstrating strong discriminative ability. InceptionV3 had an AUC of 0.97, showing effective feature capture. Both DenseNet121 and ResNet50 had AUCs of 0.95, reflecting competent performance in the classification task.

TABLE 4. ROC curve's AUC for the classes

Model/Class	ChickenPox+Measles	Monkey Pox	Normal Skin
Xception	0.98	0.98	1
Inception V3	0.97	0.98	0.99
MobileNet	1	1	1
DenseNet	0.95	0.98	0.99
ResNet	0.95	0.95	0.97

E. DISCUSSION

According to the analysis, the chosen algorithms effectively detect monkeypox in digital skin scans, with Xception performing the best. Xception achieved an F1-score of 0.989, surpassing other models in accuracy, precision, and recall. The MobileNet model also performed well, with an accuracy of 97.6% and an F1-score of 0.968, though Xception outperformed it in these metrics.Our approach successfully distinguishes monkeypox from other conditions like measles and chickenpox, highlighting its effectiveness in differentiating skin diseases.The challenge of acquiring accurate monkeypox images underscores the need for careful validation by experts before including them in training datasets.Compared to other studies using models such as custom CNNs, Inception-ResNet-v2, and VGG-16, our study achieved superior accuracy and F1-scores, demonstrating the robustness of our model selection and evaluation approach.

TABLE 5. Table of Summaries

Model	Precision	Recall	F1 Score	AUC	Accuracy
MobileNet	0.96	0.96	0.96	1.00	0.97
InceptionV3	0.94	0.94	0.94	0.97	0.94
ResNet 50	0.84	0.81	0.82	0.95	0.81
DenseNet121	0.84	0.82	0.83	0.95	0.82
Xception	0.98	0.98	0.98	0.98	0.98

VII. BUILDING A GRAPHICAL USER INTERFACE (GUI)

We created a web application to showcase the efficacy of our best-performing model. The application, built with a tech stack including Windows 11, Matplotlib, Seaborn, Pandas, a T4 GPU on Google Colab, TensorFlow, Python 3, and Streamlit, follows these steps: Load the Model: The app accesses the saved Keras model files (.h5) from the server.Load Labels: It loads label files indicating class indexes: 0 for normal, 1 for chickenpox and measles, and 2 for both. Image Input: Users can upload images to Streamlit. Resize Image: The app resizes the image to 244x244 pixels match the model's training resolution.Normalize Image: The image array is normalized to a range 0 to 1. Predict Class: Model predicts utilizing the picture class the label file to interpret the class index.Read Confidence Score: The confidence score is extracted and scaled to a percentage. Display Results: The app presents the user's class and ratings of confidence using Markdown formatting.

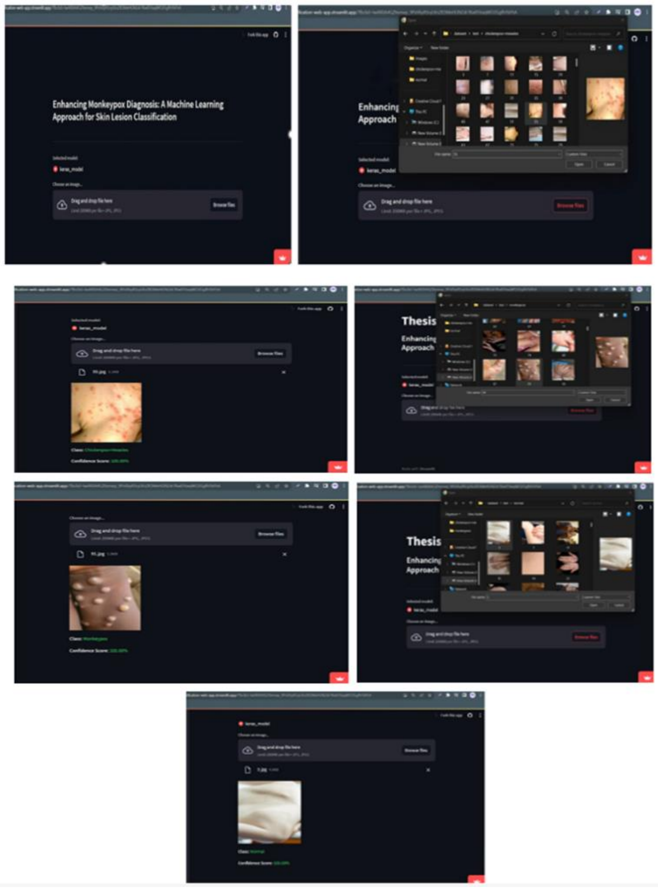


Fig. 8. Website homepage, image selection, and result showing

VIII. DECLARATION OF COMPETITIVE INTEREST

The authors confirm that they have no known financial conflicts of interest or personal relationships that could have influenced the work presented in this paper.

X. CONCLUSION

Recent research tested machine learning techniques to determine if digital skin photos are useful for diagnosing monkeypox..Models evaluated performed similarly, but Xception emerged as the most accurate with the best F1 scores. The models showed high accuracy in identifying healthy skin and differentiating between monkeypox and chickenpox, which is crucial for early diagnosis and prompt treatment. Accurate differentiation is vital, as symptoms of chickenpox and monkeypox can be similar, leading to potential misdiagnosis, especially in areas with few communicable disease experts. The study concludes that a deep learning framework can effectively classify skin lesions related to these diseases. Implementing such a framework, either independently or alongside experts, can enhance early diagnosis and help prevent future outbreaks of monkeypox and chickenpox. Future work includes upgrading our web application for better user experience and performance. We plan to expand the data set to enhance accuracy and ensure the software can manage computational and memory demands effectively. Offline functionality will be added to maintain app performance without internet connectivity. Security measures will be strengthened to protect user data and privacy. Additionally, we will implement automatic updates for the app to ensure it always has the most recent updates, including new features, bug fixes, and security improvements.

REFERENCES

1. C. Sitaula and T. B. Shahi, "Monkeypox virus detection using pre-trained deep learning-based approaches", *en J. Med. Syst.*, vol. 46, no. 11, 2022.
2. M. Altun, H. G¨ur¨uler, “Discuss the detection of Monkeypox using convolutional neural networks (CNN) combined with transfer learning techniques”, in an article published in *Sensors (Basel)*, vol. 23, issue 4, p. 1783,2023.
3. Gang Hu, Weiguo Zhao,“Diagnosis of Monkeypox Disease Using Transfer Learning and Binary Advanced Dipper Throated Optimization Algorithm”,Medicine, Computer Science, Environmental Science, DOI:10.3390/biomimetics8030313, 1 July 2023.
4. M. M. Ahsan, M. R. Uddin, M. S. Ali, et al., "Deep transfer learning approaches for monkeypox disease diagnosis," Expert Systems with Applications, vol. 216, no. 119483, pp. 119-483, 2023.
5. Gozde Yolcu Oztel,“Vision transformer and CNN-based skin lesion analysis: classification of monkeypox”,Open access,Published: 09 July 2024, Volume 83, pages 71909–71923, (2024).
6. K. Chadaga,T. Nayak, "Deep learning based detection of monkeypox virus infection using skin lesion images," *Medical Novel Technologies and Devices*, vol. 18, pp. 100-243, 2023.
7. A. S. Azar, H. Pirnejad, and U. K. Wiil, "The infection known as monkey diagnosis using deep neural networks," BMC Infectious Diseases, vol. 23, no. 1, 2023.
8. A. Shawon, Y. Bao, and Y. Huang, "DenseNet convolutional neural networks use for predicting COVID-19 using CT imaging," SN Computer Science, vol. 2, no. 5, 2021.
9. D. Kmiec and F. Kirchhoff, "a form of monkey: an emerging threat?" International Journal of Molecular Sciences, vol. 23, no. 14, p. 7866, 2022.
10. L. Shapo,M. Altindis, "Monkeypox virus illness detection: a summary "Travel Health and Chronic Disease," Travel Medicine and Infectious Disease, vol. 50, no. 102459, p. 102459, 2022.
11. H. A. Jassmi, and M. Gocho, "Performance analysis of deep CNN-based identification of cracks and location methods in solid structures," Sensors (Basel), vol. 21, no. 5, p. 1688, 2021.
12. G. Muhammad, and M. S. Hossain, "Discriminative feature learning for skin disease classification using deep convolutional neural networks," IEEE Access, vol. 8, pp. 39,025–39,033, 2020.
13. J. Cheng *et al.*, “Retrieval of brain tumors by adaptive spatial pooling and fisher vector representation,” *PLoS One*, vol. 11, no.6, p. e0157112, 2016.
14. Shams, M.Y. Hybrid Neural Networks in Generic Biometric System: A Survey. J. Artif. Intell. Metaheuristics,1, 20–26,[CrossRef], 2022.
15. S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, “A survey of MRI-based medical image analysis for brain tumor studies,” *Phys. Med. Biol.*, vol. 58, no. 13, p. R97, 2013.
16. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv Prepr. arXiv1409.1556*, 2014.
17. T. Hillel, M. Bierlaire, M. Z. E. B. Elshafie, and Y. Jin, "A comparative review of artificial intelligence classification methodologies for modeling customer mode choice," *Journal of Choice Modelling*, vol. 38, article 100221, 2021. [CrossRef]
18. R. dos Santos Lobo, E. Trindade, L. A. Carneiro D'Albuquerque, and L. Haddad, "COVID-19-related healthcare cost-outcome evaluation: The effect of medical and societal factors," Brazilian Journal of Infectious Diseases, vol. 25, article 101609, 2021.
19. H. S. Kim, and R. A. Naqvi, "CardioNet: Automatic semantic segmentation to calculate the cardiothoracic ratio for cardiomegaly and other chest diseases," Journal of Personalized Medicine, vol. 12, article 988, 2022.
20. H. Pirnejad and U. K. Wiil, "Monkeypox detection using deep neural networks," BMC Infectious Diseases, vol. 23, no. 1, 2023. [CrossRef]