Multi-stage Classification of Monkeypox Disease Using Deep Learning Techniques

Orawan Chunhapran* and Maleerat Maliyeam

Department of Information Technology, Faculty of Information Technology and Digital Innovation, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand
Email: orawan_ch@rmutto.ac.th (O.C.); maleerat.m@itd.kmutnb.ac.th (M.M.)
*Corresponding author

***Abstract*—The ongoing monkeypox outbreak poses a significant global health challenge, underscoring the need for rapid and accurate identification of symptoms and rash stages. Deep learning algorithms have become increasingly valuable for disease diagnosis using medical imaging. This study addresses the classification of monkeypox skin lesions and rash stages using deep learning techniques, employing a combined and augmented dataset from "Monkeypox2022" and "Mpox Skin Lesion Dataset Version 2.0 (MSLD v2.0)", and a dataset of "RashStageMpox", which covers various rash stages. Three convolutional neural network architectures were evaluated—MobileNetV2, EfficientNetB0, and EfficientNetB2. EfficientNetB2 demonstrated strong performance, achieving 94.61% accuracy for skin lesion classification (Adam optimizer, batch size 32, learning rate 0.0001) and 90.90% accuracy for rash stage classification (Adam optimizer, batch size 32, learning rate 0.001). Building on this, the study further enhanced EfficientNetB2 by integrating the Squeeze-and-Excitation (SE) module. The SE-augmented EfficientNetB2 achieved even higher accuracy: 96.15% for skin lesions and 96.96% for rash stage classification. The SE module's dynamic channel-wise feature recalibration improved the model's focus on critical diagnostic features, resulting in more accurate classification, particularly in correctly identifying scabs and reducing misclassification of vesicles as pustules, as reflected in the confusion matrices. Despite these promising results, limitations remain regarding dataset diversity, image quality, and generalizability to broader patient populations. Future work should include validation with multi-institutional datasets, optimization for mobile deployment, and real-world clinical evaluation, focusing on Explainable Artificial Intelligence (XAI), robust data security, and user-friendly interfaces to ensure ethical and practical implementation.**

***Keywords*—monkeypox, deep learning, skin lesion, rash stages, EfficientNetB2, Squeeze-and-Excitation (SE) module**

I. Introduction

Monkeypox (MPox), a zoonotic disease caused by the Orthopoxvirus, presents a growing global health challenge. While historically concentrated in West and Central Africa, the virus has rapidly spread worldwide, prompting the World Health Organization (WHO) to declare a Public Health Emergency of International Concern (PHEIC) on July 23, 2022. Globally, Monkeypox has affected over 87,000 individuals across 110 countries, resulting in over 100 fatalities. Transmission occurs through respiratory secretions and direct contact with skin lesions. Accurate Monkeypox diagnosis is complicated by its clinical similarities to other infections, such as chickenpox, measles, herpes, and syphilis, making early differentiation challenging. Although Polymerase Chain Reaction (PCR) testing provides definitive confirmation, resource limitations, particularly in remote areas, necessitate alternative diagnostic approaches [1, 2].

Artificial Intelligence (AI), particularly Deep Learning (DL) and Convolutional Neural Networks (CNNs), has demonstrated significant potential in medical image analysis, aiding in the diagnosis of pulmonary diseases such as pneumonia, tuberculosis, COVID-19, and lung cancer [3], and Alzheimer's disease detection [4, 5]. In dermatology, DL techniques are increasingly used for skin lesion detection, dermatological diagnosis, and medical image classification. Studies have applied DL to classify monkeypox images using architectures like VGG16, ResNet-18, MobileNetV2, and EfficientNet-B0/B3, achieving high accuracy in distinguishing monkeypox from other conditions [6–8]. However, much research focuses on binary or multi-class classification, with limited attention to the specific stages of monkeypox rash, a critical factor for predicting disease progression and tailoring treatment strategies. Further research is needed to improve the accurate differentiation of these stages, particularly in early stages where lesions can appear visually similar, such as distinguishing Macules from Papules and improving the identification of Vesicles.

Addressing this gap, our study introduces a novel approach focused on classifying both Monkeypox lesions and the distinct rash stages using DL techniques. This stage-specific classification provides valuable insights into disease progression, enabling clinicians to make more informed decisions. Our datasets include Monkeypox2022 and MSLD v2.0 for Monkeypox lesion classification, and a dataset comprising five rash stages of Monkeypox: Macules (1–2 days), Papules (1–2 days), Vesicles (3–5 days), Pustules (5–7 days), and Scabs (7–14 days). To enhance classification performance, the study emphasizes hyperparameter optimization, including a selection of

Manuscript received January 28, 2025; revised April 23, 2025; accepted May 14, 2025; published August 18, 2025.

optimizers (e.g., Stochastic Gradient Descent (SGD) or Adam), batch size, and learning rate, to achieve optimal accuracy, precision, recall, and F1-Score. We also explore integrating Squeeze-and-Excitation (SE) modules into our architecture to improve feature discrimination. Our models enable improved initial Monkeypox diagnoses and efficient treatments.

The study's structure includes: Section II, a literature review of Monkeypox characteristics and DL-based classification, Section III, a description of the datasets and preprocessing, the methodology, DL architectures, and hyperparameter optimization, Section IV, a discussion of the experimental results, and Section V, a conclusion summarizing the findings and suggesting future research directions. This novel research approach provides significant advancements in improving Monkeypox diagnostics.

II. Literature Review

In the literature review, three topics as discussed: Monkeypox, Monkeypox Rash Stage, and Deep Learning for Monkeypox Classification.

A. Monkeypox

Since 2022, multiple outbreaks of Monkeypox (MPox) have been reported globally, including in countries where the disease was not previously endemic. As of July 2022, over 5000 confirmed cases had been reported in 66 countries across the Americas, Europe, Africa, the Eastern Mediterranean, and the Western Pacific, prompting the World Health Organization (WHO) to classify MPox as a moderate global health risk. Subsequently, on July 23, 2022, the WHO declared the MPox outbreak a Public Health Emergency of International Concern (PHEIC) due to the rapid global spread. Monkeypox is caused by the monkeypox virus, a member of the Orthopoxvirus genus, closely related to the smallpox virus [9, 10]. Despite its name, rodents are considered the primary reservoirs rather than monkeys. Genetic studies have identified two distinct clades of the virus—Clade I (Central African) and Clade II (West African)—with differing virulence in humans [11].

The virus was first identified in humans in 1970 in the Democratic Republic of Congo during smallpox surveillance. Early outbreaks were mostly limited, with a low secondary infection rate (~3%), significantly lower than the ~80% seen with smallpox in unvaccinated individuals. However, the decline in smallpox vaccination over time has increased population susceptibility. During the 1990s, outbreaks in Central Africa reported secondary attack rates exceeding 75%. In 2003, monkeypox was reported outside Africa for the first time during an outbreak in the United States. The virus was believed to have been imported via rodents from Ghana, which infected prairie dogs later sold as pets, resulting in over 70 human cases. Transmission typically occurs through direct contact with infected animals via bites, scratches, or exposure to bodily fluids. Although human-to-human transmission is less efficient, it played a significant role during the U.S. outbreak and has become more relevant in recent global cases, particularly through close skin-to-skin contact [11].

B. Monkeypox Rash Stage

Monkeypox is a re-emerging zoonotic disease caused by the monkeypox virus, a double-stranded DNA virus belonging to the Orthopoxvirus genus. It is closely related to the variola virus, which causes smallpox, but typically results in a milder clinical course. The disease can be transmitted from animals to humans, as well as from person to person, through direct contact with infected skin lesions, bodily fluids, contaminated objects (such as clothing or bedding), or respiratory secretions during prolonged face-to-face contact. After infection, the incubation period typically ranges from 7 to 14 days, though it may extend from 3 to 17 days. During this stage, individuals do not exhibit any symptoms, and most feel healthy. The first noticeable symptoms (prodromal phase) include fever, headache, muscle aches, fatigue, swollen lymph nodes, and, in some cases, back pain. These early symptoms may resemble other viral infections, making it difficult to differentiate clinically.

A key feature of monkeypox is the development of a distinctive skin rash, which generally appears 1–3 days after the onset of fever. However, in some cases, the rash may present as the first symptom. The rash can manifest on various parts of the body, such as the face, oral mucosa, palms of the hands, soles of the feet, chest, genital and perianal areas, and even the conjunctivae of the eyes. The number of lesions varies widely across individuals, from a few to several thousand, and may correlate with the severity of the infection.

The skin lesions evolve through five well-defined clinical stages:

- Macules—flat, discoloured spots on the skin.
- Papules—raised, firm bumps.
- Vesicles lesions are filled with clear fluid, resembling blisters.
- Pustules—vesicles that turn opaque and fill with yellowish pus.
- Scabs—the pustules dry out, forming crusts that eventually fall off.

Once the scabs have completely detached and healthy skin has regenerated underneath, the individual is no longer considered contagious.

The total duration of illness is approximately 2–4 weeks, but this can vary depending on the patient's immune status and the severity of infection. Immunocompromised individuals, children, and pregnant people are more likely to experience severe complications, which may include secondary bacterial infections, bronchopneumonia, encephalitis, or even death in rare cases. Permanent scarring may result, particularly in visible areas such as the face and hands.

There are two genetic clades of the monkeypox virus: the Congo Basin clade, which is more virulent and associated with higher mortality rates, and the West African clade, responsible for the 2022–2023 global outbreak. The West African clade typically causes milder

symptoms and lower mortality, though complications can still arise.

Accurate diagnosis requires clinical observation of rash progression and laboratory confirmation, primarily via PCR testing from lesion swabs. The progression of lesions through defined stages is essential for distinguishing monkeypox from other rash-causing illnesses like chickenpox, measles, or syphilis [12, 13].

Understanding the rash evolution is crucial not only for early detection and isolation protocols but also for training effective machine learning models for image-based classification. This knowledge enables the development of robust models that distinguish between different stages of monkeypox rashes, thereby supporting clinical decision-making and enhancing diagnostic accuracy in digital health research.

C. Deep Learning for Monkeypox Classification

The detection of monkeypox has become an increasingly prominent area of focus in medical image analysis. In recent years, deep learning algorithms have shown significant potential in disease diagnosis, including the detection of monkeypox. Numerous studies have investigated the application of deep learning techniques to enhance both the accuracy and efficiency of monkeypox identification.

According to Ref. [6], a dataset consisting of 117 images, comprising 45 images of monkeypox cases and 74 images of other conditions such as normal skin, scarlet fever, and roseola, was used to classify monkeypox and non-monkeypox images. The images were manually collected by two medical professionals from various online sources and were made publicly available on Kaggle. Four pre-trained models were evaluated: VGG16, ResNet50, MobileNetV2, and EfficientNetB3. Among them, MobileNetV2 achieved the highest accuracy at 98.16%. However, the study noted limitations due to the small dataset size, which may not adequately represent the full spectrum of monkeypox cases.

Taruno *et al.* [7] utilized the "Mpox Skin Lesion Dataset," which was categorized into two classes: monkeypox and non-monkeypox. The dataset was trained using three pre-trained deep learning models—EfficientNet-B0, MobileNet, and InceptionV3. Among these, EfficientNet-B0 demonstrated the best performance with an accuracy of 85.12%. It also offered significant computational efficiency by utilizing approximately four times fewer parameters compared to other models, highlighting its suitability for resource-constrained environments.

In a related study [8], the "Monkeypox Skin Lesion Dataset" was employed for binary classification using five pre-trained convolutional neural networks: GoogLeNet, Places365-GoogLeNet, SqueezeNet, AlexNet, and ResNet-18. Among these, ResNet-18 achieved the highest classification accuracy of 99.49%, with 80% of the dataset used for training and 20% for validation. This study also integrated Explainable Artificial Intelligence (XAI) techniques such as Local Interpretable Model-agnostic Explanations (LIME) and Gradient-weighted Class Activation Mapping (Grad-CAM) to provide visual explanations of the model's predictions, offering interpretability to support clinical decision-making.

Harikiran *et al*. [14] using the Monkeypox Skin Lesion Dataset (MSLD) was created by collecting and analyzing images from various websites and publicly available case reports. This dataset was used for both binary and multi-class classification tasks. For binary classification (Monkeypox (MPX) vs. non-MPX), a modified ResNet50 architecture achieved an accuracy of 94.74%. For three-class classification (MPX vs. Chickenpox vs. Normal), an accuracy of 76.49% was reported. This study emphasizes the use of limited devices, such as smartphones, and incorporates explainable AI techniques to enhance interpretability for healthcare professionals. Meena *et al*. [15] developed a deep learning model based on transfer learning to assist healthcare professionals and the general public in identifying Monkeypox infections. The InceptionV3 model was trained using the publicly available Monkeypox Skin Lesion Dataset from GitHub, which originally contained 228 images—102 labeled as "Monkeypox" and 126 as "Other" (including chickenpox and measles). Due to the limited size of the dataset, extensive data augmentation techniques were applied, such as rotation, translation, reflection, jitter, noise, scaling, shear, and adjustments to hue, saturation, contrast, and brightness. This augmentation process expanded the dataset by nearly 14 times, resulting in 1428 images for the Monkeypox class and 1764 for the Other class. The InceptionV3 model achieved a high accuracy of 98%, demonstrating strong potential for use in real-world diagnostic applications. The proposed method is capable of handling large volumes of images, reducing the diagnostic burden on medical personnel and lowering the cost of laboratory resources and imported equipment. Moreover, the model can be updated with new data to improve its robustness over time. Future work may also explore transfer learning-based automatic hyperparameter tuning to further enhance model performance.

Sharma *et al*. [16] focused on the classification of skin lesions into four categories: Monkeypox, Measles, Chickenpox, and Normal, using the Monkeypox Skin Lesion Dataset (MSLD). A custom ResNet-18-based model was developed and compared against several established deep learning models, including the Visual Geometry Group 16-layer (VGG16), ResNet-50, and InceptionV3. Among all tested models, the custom ResNet-18-based model achieved the highest performance with an accuracy of 84.59%, using a 70:30 train-test split. However, the study acknowledges the limited availability of monkeypox data, as the disease remains relatively new in medical research. This limitation highlights the need for robust methodologies that can reduce the risk of overfitting and improve the generalizability of the classifier. The researchers propose that future work should focus on developing a more optimized model and training it on larger and more diverse lesion image datasets to improve performance and reliability in real-world clinical applications. Rajalaxmi *et al*. [17] employed the Monkeypox Skin Image Dataset to evaluate the effectiveness of Convolutional Neural Network (CNN)

classification models for distinguishing monkeypox from other visually similar skin conditions, such as smallpox, chickenpox, and measles. The detection of monkeypox was explored using three pre-trained deep learning models through transfer learning techniques. The model was initially developed using a custom CNN architecture via the Keras Application Programming Interface (API), and its performance was then compared with that of VGG16 and ResNet-50. Among the tested models, ResNet-50 demonstrated the highest potential, achieving an accuracy of 87% when trained with a batch size of 64. The research successfully established a multi-class image classification framework for differentiating skin diseases with similar visual features. Future directions include the development of a user-friendly interface for broader application. However, the study's major limitation lies in the relatively small dataset size, which may constrain the model's generalization capability. Improved performance is expected with the inclusion of larger and more diverse datasets in future work. In Ref. [18], various deep learning algorithms were employed to predict monkeypox with high accuracy. The models were trained on images of monkeypox as well as on images of other visually similar diseases such as chickenpox and measles. Among the tested models, InceptionV3 demonstrated a notable performance, achieving an accuracy of 89% when trained on the Monkeypox Images dataset, using a batch size of 32 and 50 epochs. The evaluation included precision, recall, and accuracy metrics to thoroughly assess the model's performance. Despite the promising results, the visual similarity between monkeypox, chickenpox, and measles presents a significant challenge in accurate differentiation, particularly in regions with limited access to infectious disease specialists. This research highlights the effectiveness of deep learning architectures in accurately classifying skin lesions commonly associated with monkeypox and related diseases.

Upon reviewing the literature, it is evident that several studies have successfully utilized the MobileNetV2 and EfficientNet architectures for image classification tasks, highlighting their effectiveness [6, 7, 19]. This success has led researchers to adopt MobileNetV2 in their studies due to its lightweight design, Suitable for portable devices and embedded systems, providing high computational efficiency. MobileNetV2 employs Depthwise Separable Convolutions to reduce computational complexity without compromising accuracy. Additionally, it offers Real-Time Performance with rapid inference times, making it ideal for real-time applications in healthcare settings and situations requiring quick diagnosis. In terms of EfficientNetB0, this architecture utilizes a Compound Scaling Method that simultaneously adjusts network depth, width, and resolution, achieving a balance between computational efficiency and high accuracy. EfficientNetB0 uses fewer parameters (5.3 million) than other models while maintaining high performance. Its design allows for effective feature extraction across diverse datasets, making it suitable for medical imaging tasks that require high-resolution differentiation. EfficientNetB2 builds upon EfficientNetB0 by enhancing input resolution and network depth, enabling it to handle complex data effectively. This robustness allows EfficientNetB2 to manage intricate patterns in medical images efficiently.

The advantages of MobileNetV2, EfficientNetB0, and EfficientNetB2 enable them to be applied to tasks with speed and resource constraints, such as medical image analysis requiring high accuracy and adaptability to diverse, refined datasets. This approach holds the potential for developing more effective systems for detecting and managing monkeypox, playing a crucial role in controlling outbreaks and enhancing public health safety.

Currently, deep learning techniques, particularly those using MobileNetV2 and EfficientNet architectures, have shown great promise in medical applications, especially in analyzing medical images. These techniques not only enable rapid and accurate disease detection but also play a significant role in preventing disease spread. Moreover, identifying the stages of monkeypox is crucial as it aids healthcare professionals in diagnosis, patient isolation, treatment planning, and prevention. Rapid and accurate interventions based on disease stage can significantly reduce community transmission risks.

This study has compiled and evaluated the performance of datasets from multiple sources, aiming to enhance image classification accuracy and facilitate real-world applications. Accurate stage identification provides valuable insights for healthcare professionals, assisting in diagnosis, patient isolation, treatment planning, and disease prevention. Furthermore, timely and precise interventions based on disease stages can significantly mitigate the risk of outbreaks in communities.

III. Methodology

The methodology of multi-stage classification is divided into six main topics: 1) Proposed Framework. Describes the overall structure and approach used for the multi-stage classification process, including the steps and workflow from data input to output. 2) The dataset provides details about the dataset used, including the instances, types of images (such as skin lesions or rash stages), and any preprocessing or organization steps taken to prepare the data for training. 3) Data Augmentation explains the techniques used to artificially expand the dataset by applying transformations to the images to improve the robustness of the model and reduce overfitting. 4) Model Building Outlines selecting and building the deep learning models (e.g., MobileNetV2, EfficientNetB0, and EfficientNetB2), including the architecture, layers, and configurations used to classify the images in the different stages. 5) Experiment Setup: Describes the experimental environment, including the choice of hyperparameters, training settings, evaluation metrics, and hardware or software resources used for model training and testing. And 6) Performance Metrics discusses the key metrics used to assess model performance, including accuracy, precision, recall, and F1-Score. Accuracy reflects the proportion of correctly predicted instances (Monkeypox and rash stages) out of the total instances.

A. Proposed Framework

The concept of multi-stage classification is divided into three steps: 1) Skin Lesion Classification to identify whether the skin image shows monkeypox infection. 2) Filter of Infected Skin lesions on monkeypox lesions only. And 3) Rash Stage Classification categorizes the rash caused by monkeypox into different stages, such as Macules, Papules, Vesicles, Pustules, and Scabs. The proposed architecture model uses CNN architectures like MobileNetV2, EfficientNetB0, and EfficientNetB2, as shown in Fig. 1.

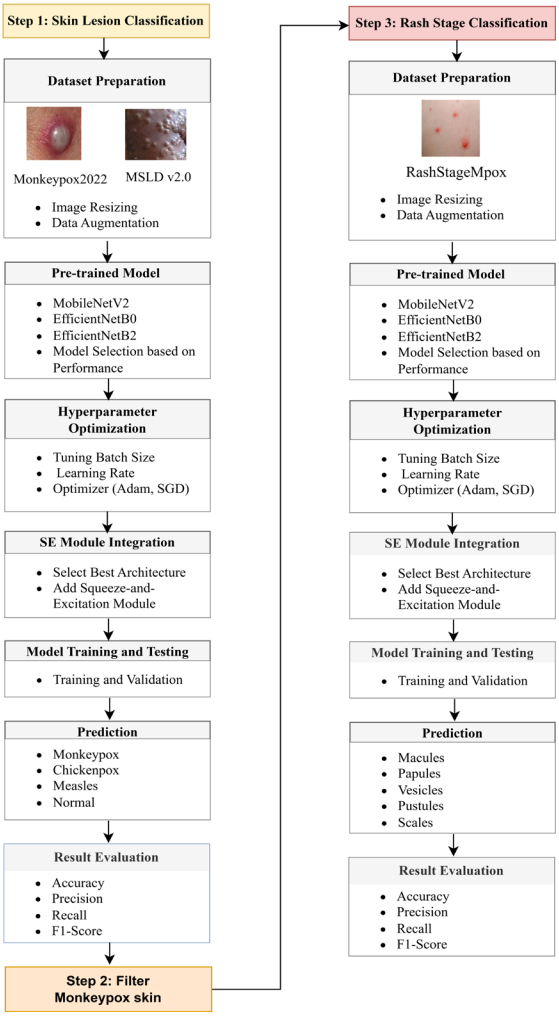

Fig. 1. The architecture for classifying skin lesions and rash stages of Monkeypox rash using deep learning.

B. Dataset

The image data is used in the classification of skin lesions and rash stages of monkeypox using deep learning techniques. The dataset is structured into two main components:

1) Dataset for classificatoion

The dataset comprises two parts: the Skin Lesion Image Dataset and the Rash Stage Image Dataset, which include images representing various stages of monkeypox rash.

a) Skin lesion image dataset

Classification extended to four distinct classes: Monkeypox, Chickenpox, Measles, and Healthy Skin. The image data used in this classification task is derived from two primary datasets:

- "Monkeypox2022", developed by the Department of Computer Science and Engineering, Islamic University, Kushtia-7003, Bangladesh, contains 770 images.
- "MSLD v2.0" (Mpox Skin Lesion Dataset Version 2.0) consists of images from monkeypox patients, created under the guidance of professional dermatologists and approved by regulatory authorities. Includes a total of 528 images.

These datasets serve as the foundation for building and evaluating models aimed at accurately classifying skin conditions and supporting timely diagnosis and treatment. Fig. 2 shows the sample of Monkeypox and healthy skin images from two different datasets. The skin image dataset consists of four classes: Monkeypox, Chickenpox, Measles, and Healthy Skin, as shown in Table I.

TABLE I. THE NUMBER OF MONKEYPOX AND NORMAL SKIN IMAGES FROM TWO DATASETS

Skin Class	Data Set		Total
	Monkeypox2022	MSLD v2.0	
Monkeypox	279	284	563
Chickenpox	107	75	182
Measles	91	55	146
Normal	293	114	407
Total	**770**	**528**	**1,298**

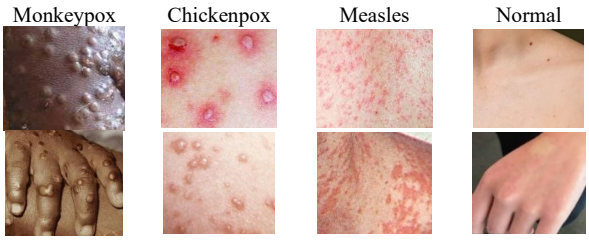

Fig. 2. Sample of Monkeypox and Healthy skin images from two different datasets.

b) Rash stages image dataset

"RashStageMpox" Dataset: This dataset contains 325 images of different rash stages of Monkeypox, provided by health organizations, including the U.S. Centers for Disease Control and Prevention (CDC) and the U.K. National Health Service (NHS).

The rash stage of Monkeypox is divided into five levels: Macules have 71 images, Papules have 52 images, Vesicles have 48 images, Pustules have 120 images, and Scales have 35 images, as shown in Table II and sample images from the dataset of each category. Monkeypox rash stages in Fig. 3.

TABLE II. NUMBER OF IMAGES FOR EACH RASH STAGE OF MONKEYPOX

StageID	Rash Stage	#Images
1	Macules (1–2 days)	71
2	Papules (1–2 days)	52
3	Vesicles (3–5 days)	48
4	Pustules (5–7 days)	120
5	Scabs (7–14 days)	35
	Total	**325**

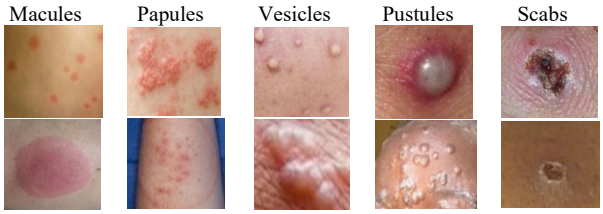

Fig. 3. Sample images from the dataset of each rash stage of Monkeypox.

2) Data split and model evaluation

The dataset was systematically divided into three main subsets to ensure rigorous evaluation of the model's performance: 1) Training + Cross-Validation (80%) split to the training (60%) to train subset was utilized the model, enabling it to learn patterns and features from the data and Cross-Validation (20%) used to a 5-fold cross-validation technique was employed on this subset to optimize the model's hyperparameters and assess its performance. This approach also mitigates the risk of overfitting during training by evaluating the model across multiple data splits. 2) Hold-Out Set (10%) This subset was reserved to monitor the model's performance during training and to detect potential overfitting. It served as an independent validation set, distinct from the training data, ensuring unbiased evaluation during model development. 3) Test Set (10%), which consisted of completely unseen data not utilized during training or validation. It was exclusively used for the final assessment of the model's generalizability and robustness in predicting outcomes on new data. Shown as detailed in Table III are the data split techniques on skin lesions, and Table IV shows the data split techniques on skin lesions.

TABLE III. THE DATA SPLIT ON THE SKIN LESION DATASET

Skin Class	Total	Training + CV	Hold-Out	Test
Monkeypox	563	450	56	56
Chickenpox	182	146	18	18
Measles	146	117	15	15
Normal	407	326	41	41

TABLE IV. THE DATA SPLIT ON THE RASH STAGE DATASET

Rash Stage	Total	Training + CV	Hold-Out	Test
Macules (1–2 days)	71	57	7	7
Papules (1–2 days)	52	42	5	5
Vesicles (3–5 days)	48	38	5	5
Pustules (5–7 days)	120	96	12	12
Scabs (7–14 days)	35	28	4	4

C. Data Augmentation

Data augmentation is a technique used to increase the volume and diversity of training data while addressing class imbalance within the dataset. This technique generates new images or data by applying various transformations to existing data. It is particularly beneficial in machine learning and deep learning, especially for image classification tasks, where larger *and* more diverse datasets contribute to improved model performance and generalization. Examples of commonly used transformations include horizontal flip, vertical flip, random rotation, random resized cropping, random brightness and contrast adjustments, Gaussian noise, and random translation. These transformations are typically applied randomly to the original images with a specified probability to generate new variations. Such augmentation helps to prevent overfitting and enhances the model's ability to generalize effectively to the real world scenarios [20–22].

In this study, two datasets were used: MSLD v2.0 and Monkeypox2022. Although the default version of MSLD v2.0 includes built-in image augmentation, it was not used in this research due to the integration of data from both sources and the implementation of a custom data-cleaning process. This process involved removing duplicate images that may exist across sources and resizing and standardizing image dimensions and aspect ratios to ensure consistency across the combined dataset. The MSLD v2.0 dataset was merged with the Monkeypox2022 dataset to support a four-class classification task (Monkeypox, Chickenpox, Measles, and Healthy Skin). As a result, custom image augmentation techniques were applied specifically for this study, as shown in Table V.

TABLE V. THE DATA AUGMENTATION TECHNIQUES

Technique	Generate Augmented
1. HorizontalFlip	50%
2. VerticalFlip	50%
3. RandomRotation	45 degree
4. RandomResizedCrop	50%
5. RandomBrightnessContrast	20%

D. Model Building

To effectively classify monkeypox lesions and rash stages, three deep learning models—MobileNetV2, EfficientNetB0, and EfficientNetB2—were selected based on their proven capabilities in image classification and their varying computational complexities. These models were employed in the skin lesion and rash stage classification pipelines. MobileNetV2 was initially considered for its efficiency and reduced parameter count. EfficientNetB0 and B2, on the other hand, were chosen for their potential to achieve higher accuracy through compound scaling, systematically balancing network depth, width, and resolution. Transfer learning was employed, using ImageNet pre-trained weights to accelerate learning and improve generalization performance on the relatively limited monkeypox dataset. These models enable us to classify images at different stages of monkeypox, which can provide healthcare professionals with crucial insights for diagnosis, treatment, and disease management. The models are explained in two sections: 1) Model Architectures, and 2) Squeeze-and-Excitation Module. Performance in skin lesion and rash stage classification was improved.

1) Model architectures

This section provides an overview of the three Convolutional Neural Network (CNN) architectures utilized in this study: MobileNetV2, EfficientNetB0, and EfficientNetB2. These models were selected due to their balance between computational efficiency and classification accuracy, particularly in the context of limited medical image datasets.

a) MobileNetV2

The MobileNetV2 architecture, as described in [23], is a deep neural network known for its efficiency and suitability for resource-constrained environments. It achieves excellent performance with relatively fast processing and low power consumption compared to larger models, making it well-suited for various image processing tasks, including image classification. MobileNetV2 builds upon the foundation of MobileNetV1 with the addition of Inverted Residual Blocks, which enhance feature learning and improve backpropagation efficiency. These blocks contribute to a reduction in model parameters, improved processing efficiency, and the use of depthwise separable convolutions, which further reduce computational cost. In our implementation, the MobileNetV2 architecture was adapted for both the four-class skin lesion classification task (Monkeypox, Chickenpox, Measles, and Normal) and the five-class rash stage classification task (Macules, Papules, Vesicles, Pustules, and Scales). For both tasks, the final classification layer was modified to output probabilities for the respective classes. The input images were resized to 224×224 pixels, and the Inverted Residual Blocks were retained to maintain efficiency. The general architecture of MobileNetV2 is shown in Fig. 4.

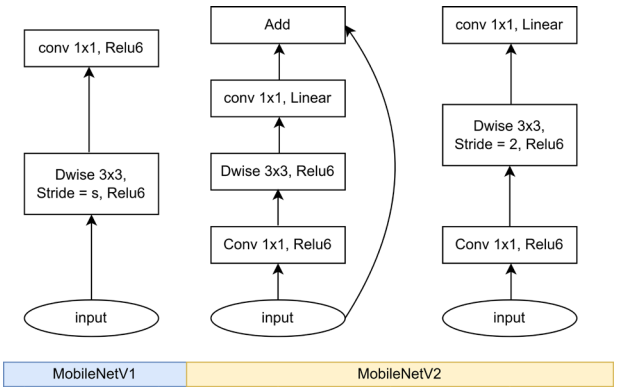

Fig. 4. The architecture for MobileNetV2.

b) EfficientNetB0

The EfficientNet model, which was developed by exploring the relationship between network depth, layer width, and image resolution, represents model scaling in three dimensions [24]. EfficientNetB0 is a baseline model designed for high computational efficiency and memory usage through an Automated Parameter Search Technique (AutoML). EfficientNetB0 has been pre-trained on ImageNet and can be used as a base model in Transfer Learning for fine-tuning to tasks requiring high accuracy and efficient resource utilization.

Compound scaling, unlike traditional methods of improving model accuracy that often focus on increasing only one dimension (e.g., depth, width, or resolution), systematically scales all three dimensions. These dimensions are scaled together using a constant ratio to ensure balanced and efficient model expansion without overburdening computational resources. This innovative approach makes EfficientNet highly accurate and resource-efficient, as shown in Fig. 5.

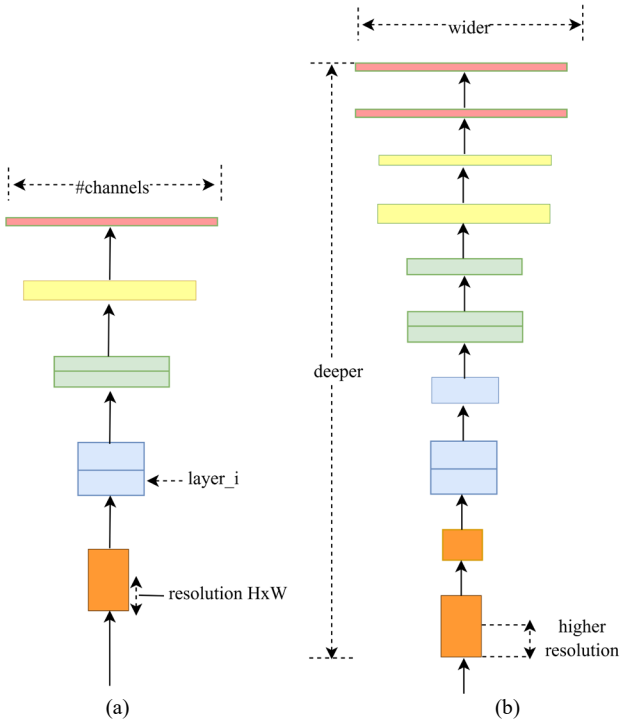

Fig. 5. Model Scaling of EfficientNet. (a) baseline network, and (b) A Compound scaling method that uniformly scales all three dimensions with a fixed ratio.

c) EfficientNetB2

EfficientNetB2 is one of the models in the EfficientNet family, designed using the Compound Scaling principle to enhance the accuracy and efficiency of the model. EfficientNetB2 is an extended version of EfficientNetB0, achieved by increasing the depth, width, and image resolution in appropriate proportions to balance computational performance and resource utilization. EfficientNetB2 builds upon EfficientNetB0, the base model designed with an Automated Search Mechanism (AutoML) using Neural Architecture Search (NAS) to discover highly efficient architectures. Its activation function employs the Swish function, which enhances non-linearity and enables the model to learn features more effectively. The advantage of EfficientNetB2 is high accuracy: it supports complex image classification tasks with improved precision and efficient Resource Utilization and optimally uses memory and computation, making it suitable for mobile and resource-constrained devices.

EfficientNetB2 is an excellent choice for tasks requiring high accuracy and efficient resource usage, such as disease classification from images or real-time image processing in mobile applications [25]. The architectures of all three models are highly efficient and well-suited for tasks requiring rapid processing on mobile devices. These capabilities align with the critical need for classifying monkeypox from images and identifying its stages based on rash progression. High-performance classification is essential for early disease detection, enabling measures to prevent its spread—a significant and challenging task.

Furthermore, these models support healthcare professionals in promptly identifying and isolating infected individuals, enabling accurate and timely treatment before the disease spreads or is transmitted within the community. Early detection is crucial for effective disease prevention, diagnosis, and management, significantly reducing the risk of outbreaks through timely intervention

2) Squeeze-and-Excitation (SE) module

The Squeeze-and-Excitation (SE) module is a deep learning mechanism, particularly effective in Convolutional Neural Networks (CNNs), designed to enhance a model's ability to focus on the most informative features within input data. This is especially important in medical imaging tasks, where subtle differences between disease stages, such as those in monkeypox skin lesions and rash stages, are critical for accurate classification. The SE module operates in two main steps:

Squeeze: Aggregates spatial information across each feature map using global average pooling, producing a compact channel descriptor that summarizes the global distribution of responses for each channel.

Excitation: Processes this descriptor through two fully connected layers and a sigmoid activation to generate channel-specific weights. These weights (ranging between 0 and 1) are then used to rescale the original feature maps, enhancing important features while suppressing less informative ones [6].

This makes the SE Module well-suited for image classification tasks, such as the classification of skin lesions and rash stages in monkeypox, where subtle visual cues are critical for accurate diagnosis.

E. Experiment Setup

The experiment setup is divided into two main steps: 1) Computation Setup, which explains the computational environment used for training and testing the model, and 2) Hyperparameters Optimization and Model Evaluation, which discusses the hyperparameters used for model training as well as the evaluation metrics employed to assess the model's performance, ensuring that the model training is appropriate and can function effectively.

1) Computation setup

In this experiment, training and testing were conducted on Google Collaboratory using input images of size 224×224 pixels, which aligns with the input format of the MobileNetV2 model, and with a batch size set to 32. Performance evaluation was done using K-Fold Cross-Validation, specifically 5-fold cross-validation, where the data was shuffled in each fold. Data augmentation techniques were employed to enhance the model's ability to handle new data and reduce overfitting. These techniques included horizontal flip, vertical flip, random rotation, a randomly resized crop, random brightness/ contrast, and Gaussian noise, which were fed into three pre-trained models: MobileNetV2, EfficientNetB0, and EfficientNetB2, pre-trained models available on TensorFlow Hub designed to classify images into 1,000 categories. This technique is highly beneficial for re-training models to classify categories in a custom dataset, such as distinguishing between custom datasets, including Monkeypox, Chickenpox, Measles, and Healthy Skin.

The model architecture was modified by removing the original classification layer and freezing the weights of the pre-trained model to prevent updates. The spatial dimensions of the data were reduced to a vector, followed by the addition of a fully connected layer. Dropout was applied to deactivate 50% of the neurons, reducing overfitting. The model was trained for 50 epochs per fold using the training data and evaluated on the validation data. This approach aims to minimize the loss using the Adam optimizer for parameter optimization with an initial learning rate of 0.001. The batch size is set to 32, and the loss function used was Categorical Crossentropy from the Keras API.

2) Hyperparameter optimization and model evaluation

To maximize the performance of the deep learning architectures, a comprehensive hyperparameter optimization strategy was implemented. This process involves tuning key hyperparameters to identify the optimal configuration for each model. The following hyperparameters were explored:

Optimizer: The Adaptive Moment Estimation (Adam) algorithm and Stochastic Gradient Descent (SGD) [26] were evaluated. Adam, known for its adaptive learning rates, and SGD, a foundational optimization technique, were chosen to explore different convergence characteristics and their impact on model generalization.

Batch Size: Batch sizes 32, 64, and 128 were investigated. These values were selected to examine the trade-off between computational efficiency and the accuracy of gradient estimation. Smaller batch sizes offer more frequent updates, potentially leading to faster initial convergence, while larger batch sizes provide smoother gradient estimates but may require more memory.

Learning Rate: The learning rate, a critical parameter governing the magnitude of weight updates, was fine-tuned across a logarithmic scale with values of 0.001, 0.0001, and 0.00001. This range was chosen based on empirical observations and established best practices in deep learning, as it encompasses typical values that yield stable and effective training.

A grid search methodology was adopted to efficiently navigate the hyperparameter space. Grid search systematically evaluates all combinations of the specified hyperparameters, leading to an exhaustive exploration of 18 distinct configurations (2 optimizers × 3 batch sizes × 3 learning rates). This comprehensive approach ensures that the optimal hyperparameter configuration is identified within the defined search space.

The evaluation of each hyperparameter combination was conducted using a stratified 5-fold cross-validation technique. Stratified cross-validation was employed to preserve the original class distribution within each fold, thereby mitigating potential biases and providing a more robust estimate of model performance. The model was trained on the data during each fold and validated. This process was repeated five times, with each fold using a different subset of the data for validation, to ensure comprehensive evaluation across the entire dataset.

The primary evaluation metric for model selection was validation accuracy, defined as the average accuracy across the five cross-validation folds. This metric provides a reliable measure of the model's ability to generalize to unseen data and serves as a key indicator of its overall performance.

The grid search process, encompassing training and validation for all hyperparameter combinations, necessitated approximately. This extensive computational effort. This extensive computational effort was conducted using Google Collaboratory, leveraging its cloud-based infrastructure with an NVIDIA T4 Tensor Core Graphics Processing Unit (GPU) and a high Random Access Memory (RAM) instance. The NVIDIA T4 GPU—developed by NVIDIA Corporation, a global leader in GPU and artificial intelligence (AI) hardware innovation—is specifically optimized for accelerated deep learning workloads. Access to these resources was critical to exhaustively explore the hyperparameter space and ensure the statistical robustness of the results. The optimal hyperparameters identified through this rigorous grid search were subsequently used to train the final models, as detailed in the Results section.

F. Performance Metrics

In classification tasks, commonly used performance metrics include accuracy, precision, recall, and F1-Score. Accuracy measures the proportion of correctly predicted instances (both monkeypox and rash stages) out of the total number of instances. Recall indicates how effectively the model identifies positive instances by calculating the ratio of correctly predicted positives to all actual positives. Precision evaluates the quality of positive predictions by measuring the ratio of correctly predicted positives to all predicted positives (i.e., monkeypox or a rash stage), and the F1-Score is the harmonic mean of precision and recall, providing a balance between the two. It is particularly useful when the dataset is imbalanced, as it considers both false positives and false negatives.

The performance of the image classifier is evaluated using accuracy, recall, precision, and F1-Score, as defined in Eqs. (1)–(4), which are determined using TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) observations.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{F1-Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

IV. Results and Discussion

This section systematically evaluates the proposed framework for Monkeypox (MPox) skin lesion and rash stage classification. Findings are organized into six subsections, each addressing a critical aspect of the experimental outcomes.

A. Model Performance

The classification performance of three deep learning architectures—MobileNetV2, EfficientNetB0, and EfficientNetB2—was evaluated for the tasks of monkeypox skin lesion classification and rash stage classification. All models were initially trained using the Adam optimizer with a learning rate of 0.001, a batch size of 32, and 50 epochs. The categorical cross-entropy loss function from the Keras API was employed during training. The experiments were conducted on Google Colaboratory, utilizing an NVIDIA T4 GPU and a high-RAM environment to ensure computational efficiency.

1) Skin lesion classification

The efficiency results from the initial trial indicated that the EfficientNetB2 model with data augmentation outperformed the approach without data augmentation in classifying monkeypox skin lesions. It delivered superior performance compared to other models evaluated in this study. The EfficientNetB2 model effectively distinguished between different types of lesions and achieved the best overall performance. It obtained an accuracy of 91.53%, a precision of 92.46%, a recall of 91.53%, and an F1 score of 91.20% for skin lesion classification. These results are presented in Table VI, which summarizes skin lesion classification.

TABLE VI. Classification Performance with Different Pre-Trained Skin Lesion Image Datasets

Model		Performance (unit%)			
		Accuracy	Precision	Recall	F1-Score
With data augmentation	MobileNetV2	61.53	67.93	61.53	58.82
	EfficientnetB0	89.99	89.65	90.00	89.65
	EfficientnetB2	91.53	92.46	91.53	91.20
Without-data augmentation	MobileNetV2	55.38	67.05	55.38	52.93
	EfficientnetB0	83.07	83.85	83.07	82.73
	EfficientnetB2	87.69	88.03	87.69	86.53

2) Rash stage classification

The performance results from the first experiment revealed that the EfficientNetB2 model with data augmentation outperformed MobileNetV2 and EfficientNetB0 in classifying the progressive stages of monkeypox rash. The model achieved the highest accuracy in identifying stages such as macules, papules, vesicles, pustules, and scabs. This suggests that the architecture's capacity to capture subtle temporal and morphological changes in rash progression is advantageous for stage classification.

It obtained an accuracy of 90.90%, a precision of 89.75%, a recall of 90.90%, and an F1 score of 90.13% for skin lesion classification. These results are presented in Table VII, which summarizes the classification of rash-stage datasets.

TABLE VII. CLASSIFICATION PERFORMANCE WITH DIFFERENT PRE-TRAINED MODELS USING RASH STAGE DATASETS

Model		Performance (unit%)			
		Accuracy	Precision	Recall	F1-Score
With augmentation	MobileNetV2	72.72	73.58	72.72	70.99
	EfficientnetB0	87.87	88.63	87.87	88.01
	EfficientnetB2	90.90	89.75	90.90	90.13
Without-data augmentation	MobileNetV2	66.67	67.53	66.67	63.98
	EfficientnetB0	84.84	86.33	84.84	85.11
	EfficientnetB2	84.84	85.45	84.84	84.90

B. *Effect of Hyperparameter Optimization*

Based on the comparison of the three models, the EfficientNetB2 architecture demonstrated superior performance in classifying both monkeypox skin lesions and rash stages. Consequently, hyperparameter optimization was conducted on EfficientNetB2 to enhance its classification capabilities further. The optimization process focused on tuning key parameters, including the choice of optimizer (Adam and SGD), learning rates (0.001, 0.0001, and 0.00001), and batch sizes (32, 64, and 128). These hyperparameters are critical for improving model generalization and preventing overfitting. This systematic optimization aims to maximize classification accuracy and robustness across diverse lesion and rash stage images.

1) Skin lesion hyperparameter optimization

Analysis of EfficientNetB2 Performance on Monkeypox Classification via Hyperparameter Tuning. Based on the experimental results presented in Table VIII, Adam optimizer consistently outperformed SGD, demonstrating higher stability and overall performance. Notably, the combination of batch size 32 with a learning rate of 0.0001 using Adam yielded the highest accuracy of 94.61% and an F1-Score of 94.63%. In contrast, SGD exhibited instability, particularly at learning rates $\leqslant$ 0.0001, where accuracy dropped to a range of 50–79.23 %.

Additionally, the table shows that Adam was more tolerant of changes in learning rate. For example, even at a low learning rate of 0.00001, the model still maintained a high accuracy of 93.07%. Impact of Batch Size: Adam's smaller batch size of 32 yielded the best performance with an accuracy of 94.61%. Meanwhile, SGD, with a batch size of 64, performed well, achieving 92.30% accuracy.

TABLE VIII. PERFORMANCE FROM PARAMETER TUNING FOR CLASSIFICATION USING SKIN LESION DATASETS

Parameter Configuration			Performance (unit%)			
Optimizer	Batch size	Learning rate	Accuracy	Precision	Recall	F1-Score
Adam	32	0.001	91.53	92.46	91.53	91.20
		0.0001	94.61	94.69	94.61	94.63
		0.00001	93.07	93.07	93.07	93.04
	64	0.001	87.69	87.47	87.69	87.47
		0.0001	88.46	88.39	88.46	87.80
		0.00001	89.99	90.11	90.00	89.56
	128	0.001	86.15	87.22	86.15	86.30
		0.0001	89.99	90.83	90.00	90.21
		0.00001	93.07	93.06	93.07	93.06
SGD	32	0.001	86.92	87.59	86.92	87.07
		0.0001	70.76	57.24	70.76	62.99
		0.00001	57.69	54.83	57.69	52.10
	64	0.001	92.30	93.10	92.30	92.42
		0.0001	79.23	66.63	79.23	72.31
		0.00001	55.38	56.69	55.38	49.48
	128	0.001	87.69	87.52	87.69	87.50
		0.0001	68.46	51.97	68.46	58.53
		0.00001	50.00	52.88	50.00	44.47

2) Rash stage hyperparameter optimization

In optimizing hyperparameters for rash stage classification, similar strategies for skin lesions were applied to ensure effective differentiation between stages, such as macules, papules, vesicles, pustules, and scabs. The highest rash stage classification performance was achieved using the Adam optimizer with a batch size of 32 and a Learning Rate (LR) of 0.001, reaching an accuracy of 90.90%. Additionally, configurations with batch size 64 and LR values of 0.0001 or 0.00001 maintained the same accuracy (90.90%), although they resulted in slightly lower precision and F1-Score, ranging from approximately 88% to 89.57%.

Adam optimizer demonstrated greater robustness to changes in batch size and learning rate due to its adaptive learning rate mechanism. In contrast, SGD showed higher sensitivity to learning rate variations. For example, with SGD at LR 0.00001 and batch size 32, the accuracy dropped significantly to 30.30%. However, the best configuration for SGD was with batch size 64 and LR 0.001, achieving an accuracy of 87.87% and an F1-Score of 88.50%. These results are presented in Table IX. Hyperparameter performance in the classification of rash-stage datasets.

Hyperparameter optimization proved critical in maximizing the performance of EfficientNetB2 for both skin lesion and rash stage classification tasks. By systematically tuning optimization, learning rates, batch size, and employing data augmentation, the model achieved superior generalization and robustness, consistent with state-of-the-art approaches in dermatological image analysis.

TABLE IX. PERFORMANCE FROM PARAMETER TUNING FOR CLASSIFICATION USING RASH STAGE DATASETS

Parameter Configuration			Performance (Unit (%))			
Optimizer	**Batch size**	**Learning rate**	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Adam	32	0.001	90.90	89.75	90.90	90.13
		0.0001	84.84	87.27	84.84	85.05
		0.00001	87.87	87.87	87.87	87.87
	64	0.001	87.87	86.32	87.87	86.84
		0.0001	90.90	88.34	90.90	89.57
		0.00001	90.90	88.34	90.90	89.57
	128	0.001	84.84	88.47	84.84	84.19
		0.0001	81.81	82.87	81.81	81.26
		0.00001	81.81	82.89	81.81	81.52
SGD	32	0.001	78.78	78.96	78.78	78.32
		0.0001	57.57	74.87	57.57	55.38
		0.00001	30.30	14.64	30.30	19.05
	64	0.001	87.87	91.22	87.87	88.50
		0.0001	63.63	55.05	63.63	56.07
		0.00001	39.39	37.46	39.39	31.87
	128	0.001	84.84	86.11	84.84	84.91
		0.0001	66.66	53.03	66.66	58.18
		0.00001	45.45	42.34	45.45	41.01

C. Experimental Results

This research utilized three image classification models and found that the model based on the EfficientNetB2 architecture delivered the best performance compared to the other models. Table X shows it achieved an accuracy of 94.61% in classifying skin lesions and an accuracy of 90.90% in classifying the rash stages of monkeypox. The hyperparameter set for skin lesion classification uses the Adam optimizer, the batch size is 32, and the learning rate is 0.0001. These parameters resulted in an accuracy of 94.61%, precision of 94.69%, recall of 94.61%, and an F1-Score of 94.63%. The hyperparameter set for classifying the rash stages of monkeypox is the Adam optimizer, the batch size is 32, and the learning rate is 0.001. These parameters yielded an accuracy of 90.90%, a precision of 89.75%, a recall of 90.90%, and an F1-Score of 90.13%.

D. Integrate Squeeze-and-Excitation Module

To optimize classification performance, the Squeeze-and-Excitation (SE) module was incorporated into the EfficientNetB2 architecture. The SE module dynamically recalibrates channel-wise feature responses by explicitly modeling inter-channel relationships, thereby enhancing feature discriminability. As demonstrated in Table XI, this table compares the classification performance for monkeypox skin lesions and rash stages before and after integrating the SE module.

TABLE X. HYPERPARAMETER SET FOR SKIN LESION AND RASH STAGE OF MONKEYPOX CLASSIFICATION BASED ON THE EFFICIENTNETB2 ARCHITECTURE

Hyperparameter/ performance	Skin Lesion	Rash Stage
Optimizer	Adam	Adam
Batch size	32	32
Learning rate	0.0001	0.001
Accuracy	94.61%	90.90%
Precision	96.69%	89.75%
Recall	94.61%	90.90%
F1-Score	94.63%	90.13%

TABLE XI. PERFORMANCE OF THE INTEGRATED SQUEEZE-AND-EXCITATION MODULE

Model	Performance (unit%)				
	Accuracy	**Precision**	**Recall**	**F1-Score**	**AUC-ROC**
Skin Lesion	94.61	94.69	94.61	94.63	98.88
+ SE Module	96.15	96.49	96.15	96.17	99.88
Rash Stage	90.90	89.75	90.90	90.13	97.71
+ SE Module	96.96	97.15	96.96	96.71	98.29

The integration of the Squeeze-and-Excitation (SE) module into the EfficientNetB2 architecture significantly enhanced its performance in classifying monkeypox-related skin conditions, as evidenced by:

1) Skin lesion classification

The base model achieved 94.61% accuracy on test data, while the SE Module improved accuracy to 96.15%, reflecting a 1.54% increase. The SE Module reduced False Positives (misclassifications as Monkeypox) from 5.31% to 3.51%, as evidenced by the precision metric. The 99.88% AUC-ROC demonstrates near-perfect discrimination between classes, attributable to the SE Module's dynamic recalibration of feature channel weights in CNNs, which enhances focus on lesion-specific regions. Notably, the +1.00% AUC-ROC improvement confirms superior class separation even for borderline cases, underscoring the SE Module's effectiveness in refining feature discrimination.

2) Rash stage classification

The SE Module boosted accuracy from 90.90% to 96.96%, while the F1-Score rose from 90.13% to 96.71%, reflecting balanced gains in both precision and recall. Despite significant class imbalance (e.g., scarce Scabs images), the AUC-ROC marginally improved from 97.71% to 98.29%, underscoring the SE Module's ability to refine class separation, particularly in distinguishing subtle differences between rash stages like Macules and Papules. Additionally, improved recall minimized misdiagnoses of early-stage Macules, which are frequently mistaken for rashes caused by other conditions.

E. Confusion Matrix and Error Analysis

Fig. 6 presents the confusion matrices comparing the best-performing model under optimized hyperparameters for skin lesion classification with and without integrating the Squeeze-and-Excitation (SE) module into the EfficientNetB2 architecture.

Before SE Module Integration, the model misclassified some Monkeypox images as Chickenpox or Normal. Although it performed reasonably well, confusion in borderline cases (e.g., Monkeypox vs. diseases with overlapping symptoms) negatively impacted precision and recall.

After integrating the SE Module, the model improved class separation significantly. The confusion matrix shows fewer misclassifications, particularly in the Monkeypox category, with clearer predictions across all four classes: Chickenpox, Measles, Monkeypox, and Normal.

This analysis demonstrates that the SE Module enhances the model's ability to focus on lesion-specific features, leading to reduced errors, higher accuracy, and more reliable classification, especially for critical cases like Monkeypox.

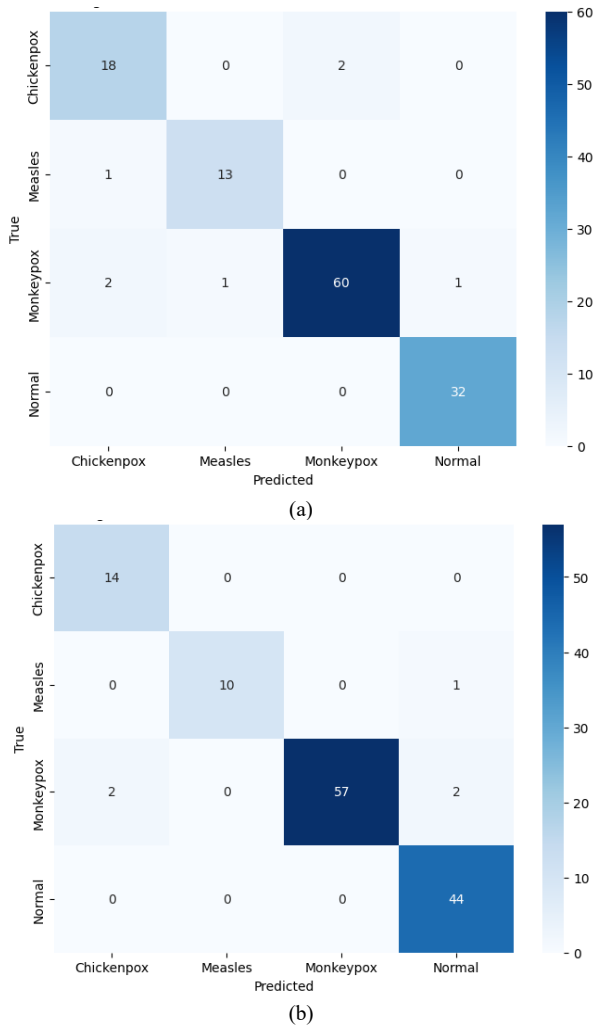

Fig. 6. Confusion matrix comparison before (a) and after (b) SE module integration into EfficientNetB2 for Monkeypox skin lesion classification.

Fig. 7 presents a detailed confusion matrix from the Monkeypox rash stage classification model, comparing the performance before and after the integration of the Squeeze-and-Excitation (SE) module into the EfficientNetB2 architecture.

Before integrating the SE module, the model achieved the highest accuracy in classifying Pustules but exhibited notable difficulty in distinguishing Vesicles and Scabs.

After integrating the SE module, the model's overall classification accuracy improved, particularly for the Vesicles and Scabs stages, Macules and Papules: All samples were correctly classified, Vesicles: Significant reduction in misclassifications, Pustules: Remained the most accurately predicted stage, Scabs: Accuracy improved considerably.

An analysis of the confusion matrix indicates that the SE module enhances the model's ability to classify Monkeypox rash stages, especially in cases where the original model exhibited weaknesses. The integration of the SE module leads to a noticeable improvement in overall classification performance.

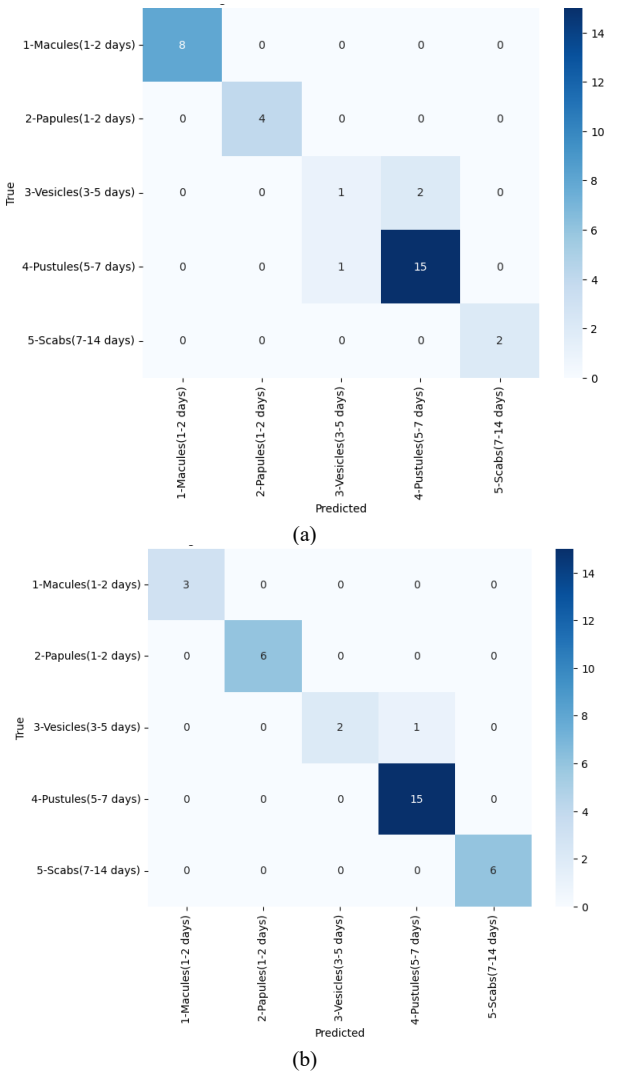

Fig. 7. Comparison of confusion matrices before (a) and after (b) SE module integration in EfficientNetB2 for Monkeypox rash stage classification.

F. Comparison with Previous Studies

Based on the comparative analysis presented in the relevant Table XII. The proposed method—EfficientNetB2 integrated with the Squeeze-and-Excitation (SE) module—demonstrated superior performance in classifying both skin lesions and rash stages associated with monkeypox. Specifically, it achieved an accuracy of 96.15% for skin lesion classification and 96.96% for rash stage classification, outperforming other methods listed.

Ref. [16–18] employed various architectures, including VGG16, ResNet50, and Inceptionv3 models on the same 4-class skin lesion classification task (monkeypox, chickenpox, measles, normal). However, none were able to reach the accuracy levels achieved by the proposed Integration SE Module into the EfficientNetB2 approach.

A key distinction of the proposed method is its utilization of a combined dataset comprised of both Monkeypox2022 and MSLD v2.0, whereas most prior research relied on the MSLD dataset alone. This integration likely contributed to the enhanced performance by providing more relevant and diverse data for the classification tasks.

The incorporation of the SE module significantly improved classification accuracy by 1.54% for skin lesions (from 94.61% to 96.15%) and 6.06% for rash stages (from 90.90% to 96.96%). These improvements highlight the efficacy of the SE module in enhancing the model's ability to discriminate subtle features, aligning with findings from other studies that have successfully applied SE modules to models like MobileNet and ResNet.

In summary, the proposed Integration SE Module into the EfficientNetB2 exhibits superior performance in classifying monkeypox-related skin conditions, attributable to both the strategic integration of relevant datasets and the effective enhancement provided by the SE module.

TABLE XII. COMPARISON OF THE PROPOSED METHOD AND PREVIOUS RESEARCH METHODS

Ref	Dataset	Model	Skin Lesion Accuracy	Rash Stage Accuracy
[16]	MSLD	VGG16 ResNet50 InceptionV3 ResNet18	81.48 82.59 74.07 84.59	N/A
[17]	MSLD	CNN VGG16 ResNet50 (batch size 64)	64.00 84.00 87.00	N/A
[18]	GitHub	InceptionV3	89.00	N/A
proposed method	Monkeypox2022 and MSLDv2.0	Integration SE Module into the EfficientNetB2	96.15	96.96

V. CONCLUSION

This study demonstrated the effectiveness of integrating the Squeeze-and-Excitation (SE) module into the EfficientNetB2 architecture for improved classification of monkeypox-related skin lesions and rash stages. Through comparative analysis against existing methodologies, the proposed Integration SE Module into the EfficientNetB2 approach achieved superior accuracy in classifying both skin lesions (96.15%) and rash stages (96.96%). The enhanced performance can be attributed to the SE module's ability to recalibrate channel-wise feature responses, enabling the model to focus on salient features critical for accurate diagnosis. As shown in the confusion matrices (Figs. 6 and 7), the SE module significantly reduced misclassifications, particularly for visually similar conditions, and improved the identification of distinct rash stages. Furthermore, the utilization of a combined dataset, incorporating both Monkeypox2022 and MSLD v2.0, contributed to improved model generalization.

Despite these promising findings, the study has some limitations. First, while the dataset was enhanced by the inclusion of both Monkeypox2022 and MSLD v2.0, and data augmentation was used, the size and diversity of the dataset may not fully represent the wide spectrum of real-world cases. The model's performance could be affected by variations in image quality, lighting conditions, and patient demographics. Second, while several architectures were explored, including MobileNetV2, EfficientNetB0, and EfficientNetB2, this study primarily focused on EfficientNetB2 due to its superior performance. Although hyperparameter tuning was conducted for EfficientNetB2, further optimization may be possible. Finally, the study primarily examined the impact of the Squeeze-and-Excitation (SE) module on the EfficientNetB2 architecture. While the SE module proved effective, other attention mechanisms and architectural modifications were not explored and might offer further improvements.

For future research directions, (1) validate the accuracy and robustness of the model on larger, more diverse, and multi-institutional datasets to ensure its broad applicability; (2) optimize the model for deployment on mobile devices to increase its accessibility in resource-constrained settings; (3) implement and evaluate the model in real-world clinical environments to assess its practical utility and impact on diagnostic accuracy and patient outcomes. Furthermore, emphasis should be placed on data security and patient privacy measures, the development of user-friendly interfaces for medical personnel, and the use of explainable AI (XAI) techniques to enhance transparency, trust, and acceptance among physicians and system users.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

OC had the conceptualization, and performed the methodology, model building process, writing—original drafting, reviewing and editing, and managing data for all authors, who had approved the final version. MM also had the conceptualization and performed the validation, resources, writing, review, editing, supervision, and project administration. All authors had approved the final version.

REFERENCES

[1] World Health Organization. (Aug. 26, 2024). Monkeypox. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/monkeypox

[2] L. Jarrell and K. Perryman, "Mpox (monkeypox): Diagnosis, prevention, and management in adults," *Nurse Pract.*, vol. 48, no. 4, pp. 13–20, 2023.

[3] S. Saleem and M. I. Sharif, "An integrated deep learning framework leveraging NASNet and vision transformer with mixprocessing for accurate and precise diagnosis of lung diseases," arXiv preprint, arXiv:2502.20570, 2025.

[4] R. Hafeez *et al.*, "Deep learning in early Alzheimer's disease detection: A comprehensive survey of classification, segmentation, and feature extraction methods," arXiv preprint, arXiv:2501.15293, 2025.

[5] S. Ucan, M. Ucan, and M. Kaya, "Deep learning based approach with EfficientNet and SE Block attention mechanism for multiclass Alzheimer's disease detection," in *Proc. 2023 4th International Conference on Data Analytics for Business and Industry (ICDABI)*, IEEE, 2023, pp. 285–289.

[6] A. S. Jaradat *et al.*, "Automated monkeypox skin lesion detection using deep learning and transfer learning techniques," *Int. J. Environ. Res. Public Health*, vol. 20, no. 5, 4422, 2023.

[7] P. E. N. Taruno, G. S. Nugraha, R. Dwiyansaputra, and F. Bimantoro, "Monkeypox classification based on skin images using CNN: EfficientNet-B0," in *Proc. E3S Web Conf.*, 2023, vol. 465, 2031.

[8] T. Nayak *et al.*, "Deep learning-based detection of monkeypox virus using skin lesion images," *Med. Nov. Technol. Devices*, vol. 18, 100243, 2023.

[9] A. Park. (May 27, 2022). There's already a monkeypox vaccine. But not everyone may need it. *Time.* [Online]. Available: https://time.com/6179429/ monkeypox-vaccine/

[10] Z. Ježek and F. Fenner, *Human Monkeypox*, Basel, Switzerland: Karger Medical and Scientific Publishers, 1988.

[11] A. M. Almars, "DeepGenMon: A novel framework for monkeypox classification integrating lightweight attention-based deep learning and a genetic algorithm," *Diagnostics*, vol. 15, no. 2, 130, 2025.

[12] M. Altindis, E. Puca, and L. Shapo, "Diagnosis of monkeypox virus—An overview," *Travel Med. Infect. Dis.*, vol. 50, 102459, 2022.

[13] CDC. (2024). Clinical features of mpox. *Mpox*. [Online]. Available: https://www.cdc.gov/mpox/hcp/clinical-signs/index.html

[14] J. Harikiran, S. Pavan, and C. B. Sai, "Detection of monkeypox using modified ResNet50," in *Proc. 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*, IEEE, 2023, pp. 317–322.

[15] G. Meena, K. K. Mohbey, and S. Kumar, "Monkeypox recognition and prediction from visuals using deep transfer learning-based neural networks," *Multimedia Tools and Applications*, vol. 83, no. 28, pp. 71695–71719, 2024.

[16] K. Sharma, V. Kumar, and M. Mittal, "MonkeyPox, measles and chickenpox detection through image-processing using residual neural network (ResNet)," in *Proc. 2023 6th International Conference on Information Systems and Computer Networks (ISCON)*, 2023, pp. 1–6.

[17] S. P. Rajalaxmi, P. S. Shetty, and S. Namita, "Automated classification of monkeypox disease based on deep learning," in *Proc. 2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, IEEE, 2024, pp. 1186–1190.

[18] R. Dhiman, U. K. Giri, A. Kumar, R. Tripathy, A. K. Sahoo, and S. P. Kumar, "Deep learning based monkeypox virus detection model using skin lesion images," in *Proc. 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS)*, IEEE, 2024, pp. 1362–1366.

[19] G. Z. Khan and I. Ullahx, "Efficient technique for monkeypox skin disease classification with clinical data using pre-trained models," *J. Innov. Image Process.*, vol. 5, pp. 192–213, 2023.

[20] Z. Wang *et al*., "A comprehensive survey on data augmentation," arXiv preprint, arXiv:2405.09591, 2024.

[21] M. Abbas, M. Arslan, R. A. Bhatty, F. Yousaf, A. A. Khan, and A. Rafay, "Enhanced skin disease diagnosis through convolutional neural networks and data augmentation techniques," *Journal of Computing & Biomedical Informatics*, vol. 7, no. 01, pp. 87–106, 2024.

[22] K. S. Krishna, B. D. Kumar, M. D. Reddy, C. Saketh, and G. Ramasamy, "A multiclass classification framework with smote based data augmentation technique for Alzheimer's disease progression," in *Proc. 2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE)*, IEEE, 2024, pp. 1–6.

[23] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520.

[24] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *Proc. International Conference on Machine Learning*, 2019, pp. 6105–6114.

[25] M. Tan and Q. Le, "Efficientnetv2: Smaller models and faster training," in *Proc. International Conference on Machine Learning*, 2021, pp. 10096–10106.

[26] D. Kurniadi, A. Mulyani, and S. Moch, "Eye disease detection model using convolutional neural network with rmsprop, SGD, and adam optimizers," in *Proc. 2024 International Conference on ICT for Smart Society (ICISS)*, IEEE, 2024, pp. 1–6.