Intelligent Deep Learning Based Monkeypox Detection

Tadiwanashe Mandina
Department of Computer Science and Engineering
Aditya University
Surampalem, Andhra Pradesh, India
tadiwanashemandina@gmail.com

Pampana Murali
Department of Computer Science and Engineering.
Aditya University
Surampalem, Andhra Pradesh, India
murali.pampana@adityauniversity.in

***Abstract*— The recent rise of Monkeypox (MPox) has raised new concerns regarding its effective and precise diagnostic measures. Outlined in this paper is the construction of an automatic MPox detection system leveraging medical imaging through deep learning techniques. It proposes recognizing the characteristic skin lesions of the virus present in clinical and dermal images using the Residual Attention U-Net architecture and transfer learning. The model is trained on a dataset which is diverse and that has been augmented to improve robustness and fine-tuned to improve diagnostic accuracy. The framework is equipped with sophisticated pre-processing methods, real-time classification, and an intuitive interface that encourages prompt diagnosis and rapid clinical intervention. The developed system can help mitigate the responsibilities of healthcare providers in MPox detection and monitoring, thereby improving management and outcomes during an outbreak. The project demonstrates the system's performance based on the metrics of accuracy, precision, recall, and F1-score, proving its usefulness in medical practice. Traditional methods of MPox detection are often time-consuming and may lack real-time applicability. This project addresses the need for a technologically advanced solution by proposing a deep learning-based system for Monkey Pox detection. The project's main objective was to develop and test a robust model capable of accurately classifying skin lesions based on image analysis.**

Keywords—Monkeypox, deep learning, Residual U-Net, medical image classification, real-time detection

I. INTRODUCTION

This project introduces a MPox detection system that utilizes deep learning techniques. The development of artificial intelligence (AI) technologies has emerged with new healthcare opportunities, especially in disease diagnosis and management. Skin diseases like measles, chickenpox, and Mpox require early and precise diagnosis to mitigate complications and the potential spread of the disease. Most practitioners rely on clinical diagnosis, which painstakingly depends on clinical expertise that varies a lot in different regions [1]. This project is located under medical engineering technology and concerns deep learning algorithms for disease detection, with objectives focused on automating the classification of skin lesions through image capture. The application of powerful neural network architectures such as Residual U-Net aims to enable accurate lesion detection using low-tech equipment like laptop webcams.

MPox is a viral zoonotic illness that has symptoms resembling small pox, though it is less severe clinically. Over the last few years, the Mpox virus has raised public health concern due to its recurring outbreaks in different parts of the world [2]. Skin lesions are one of the most important signs of Mpox and timely detection is very important for effective mitigation. Unfortunately, distinguishing Mpox from other contagious skin pathologies like measles or chicken pox is difficult without some clinical tests. The use of computer vision and deep learning techniques for the early and accurate detection of Mpox lesions constitutes a remarkable achievement in the field of healthcare technology.

Deep learning is a subset of machine learning, which, employs the use of 'deep' neural networks to learn representations of data. In other words, image recognition using deep learning models transforms and simplifies images into pixels, eliminating the need for manual feature extraction. One embodiment of deep learning architectures, Convolutional Neural Networks, have been shown to be exceptionally accurate in performing automated analysis of medical images. The classic U-Net model has been improved variously, one of them is introducing residual learning, which enhances performance in pixel-wise image classification and segmentation, to produce the Residual U-Net model.

This work introduces an **Intelligent Deep Learning-Based MPox Detection** model of the system that:

1. Makes use of a **Residual Attention U-Net** to jointly segment and classify lesions in images.
2. Integrates a **lightweight real-time application** using Python and OpenCV on standard laptop hardware.
3. Achieves **>93 % validation accuracy**, **>82 % live-test accuracy**, processing frames in **<0.9 s** on CPU.

II. RELATED WORK

A. Skin Disease Using Effective Deep Learning

The use of deep learning has offered great advancement economically to the medical imaging sector. This is dominantly noticeable in pattern recognition and classification tasks. For instance, in dermatology, many Convolutional Neural Networks (CNNs) have been developed for detection of many skin diseases. Esteva et al. proved that the use of deep neural networks can classify skin cancer images with the same accuracy a dermatologist would using dermoscopy images [3]. Furthermore, Asuntha and Srinivasan demonstrated the use of deep learning models for the detecting diseases such as eczema, psoriasis, and Melanoma, proving that end-to-end learning works without the need for manual feature extraction [4]. These models not only automate the entire detection process but also add great value by providing faster support for diagnostics relative to manual examination.

B. Residual U-Net Based Classification of Skin Disease

Although U-Net began as a tool for segmenting medical scans, researchers now increasingly apply it to the finer, pixel-level study of skin lesions. [5]. Residual U-Net, which adds residual learning on top of the U-Net structure, rehabilitates issues like the vanishing gradients of deep networks. Zhou,

Siddiquee, Tajbakhsh and Liang showed that the incorporation of deeper layers and enhanced feature propagation in Residual U-Nets led to improved performance over classical U-Nets [6]. With regard to skin disease classification tasks, Residual U-Net facilitates precise lesion segmentation and enhances classification accuracy following the analysis of the segmented regions. This effectiveness stems from its construction, which is able to retain fine spatial resolution from the lesions while using deeper feature maps.

C. *Methods of Machine Learning Used in Detection Systems of Mpox*

The counted cases of Mpox drove the search towards automated computational methods of quicker detection. Time-tested machine-learning techniques- support vector machines, random forests, and k-nearest neighbors- have already proved reliable on many classification tasks linked to viral illness.. For the case of Mpox detection, initial work utilized image skin feature extraction and subsequent image classification using ML algorithms [7]. Nevertheless, these methods tend to be unscalable and inaccurate with complex, high-variance datasets. In addition, the need for manual domain expertise along with intensive feature engineering rend traditional machine learning systems inflexible in dynamic outbreak situations.

D. *Efficient Residual UNet MPox Detection*

Active research Mpox lesions detection by deep learning models, especially Residual U-Nets, demonstrates promise. Residual U-Net architectures are capable of isolating and classifying lesion regions simultaneously due to their ability to perform segmentation and classification in one step. As Zhang et al. illustrates, the integration of residual blocks into U-Net structures enables robust learning of fine details and patterns of Mpox lesions, as well as background noise and clutter, which makes the model forensic-grade robust. Coupled with webcam real-time streaming, outbreak control is further facilitated by enabling screening to occur outside the confines of laboratory-based diagnostics. Nonetheless, there is a great deal of unexplored potential in applying Residual U-Net for Mpox detection due to the lack of existing research [8].

The application of deep learning techniques to skin disease recognition has been widely researched, but Mpox has not received specific attention. Existing systems mostly cater to static high-resolution clinical photographs and do not adapt well for real-time webcam shots, which severely limits their deployment during outbreaks. Also, there seems to be a lack of literature on systems that incorporate both lesion segmentation and classification using Residual U-Net architecture designed for differentiating Mpox from measles and chickenpox. In addition, many other works disregard real imaging condition hurdles such as changes in light, motion blur, and background clutter, assuming that these imaging conditions are trivial, which is not the case for webcam-based detection.

The proposed system incorporates a Residual U-Net model that has been trained on a dataset that includes Mpox, measles, and chickenpox as well as healthy skin images. The system is intended to operate using live webcam feeds, which allows for instantaneous capture and classification of skin lesions. The architecture of the model applies residual connections to deepen feature learning, which improve classification strength to imaging and lesions order during scan capture.

III. Methodology

It was appropriate to use the DSRM, which is Design Science Research Methodology, because it gave the right order step by step systemically for the design, development, testing, and evaluation of the system toward achieving the goals.

A. *Problem Identification and Motivation*

- Problem: The stage of Mpox lesions identification and confirmation for potential patients to critically evaluate the possibility of performing a case management for example is challenging due to absence of fast and accurate machine/technological facilitated testing in place in many countries.
- Motivation: All over the world in context of the crux of health determining factors including the available bare minimum resources, and human capacity with deep learning techniques in place, provision of rapid CT scan and issue of delayed support system for health becomes possible.

B. *Defining Objectives of a Solution*

- Strategy: Successfully achieving automating Mpox lesion detection from images obtained through standard webcams required designing a webcam-based deep learning system with a low error rate and minimal operational costs.

C. *Design and Development*

- Product: A Deep Learning Framework for Skin Lesion Classification Using Webcam Stream: Application of a Residual U-Net architecture.
- Methodology: Built a dataset, constructed and trained a Residual U-Net model, evaluated its performance, and implement it using OpenCV for real-time webcam streaming.

D. *Demonstration*

Split testing the system's ability to recognize and classify four Measles, Chickenpox, Mpox, and Normal skin into live and static image feeds was done for the demonstration.

E. *Evaluation*

As a simple means of assessing constituent tasks of systems such as recognition and time critical operation, traditional classification checks such as accuracy, precision, recall, F1 score, alongside practical checks of time, responsiveness, and speed were employed for the analysis.

F. *Communication*

Discussion concerning results and findings, together with recommendations for future work, is presented in this paper. With all this accomplished, it is fair to conclude DSRM has successfully structured both the theoretical and practical framework of this research.

G. System Development Model.

A systematic, structured methodology called the Software Development Life Cycle (SDLC) was followed when developing the solution. SDCL is a clear, step by step framework that moves through design, development, integration, and testing of a proposed software solution to ensure the final program satisfies user needs. The process also lays out the roadmap for later maintenance, updates, and feature expansions. For this project, the classic Waterfall model was chosen, which treats each phase as a distinct and non-overlapping stage.

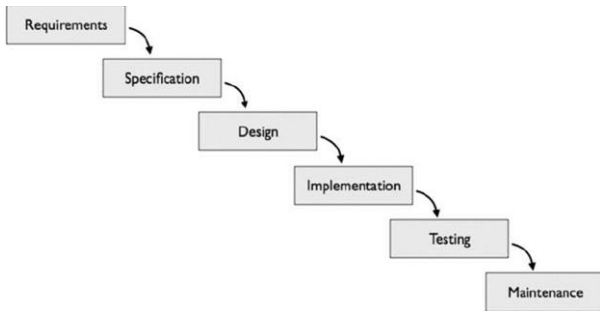


Figure 1 Waterfall model

H. Data Collection

The dataset for this particular project was composed of images grouped into the following four folders.

- Measles
- Chickenpox
- Mpox
- Normal

The sources included dermatology datasets available in the public domain as well as other image repositories online. Public domain as a consideration for ethics was respected. Each image has been checked manually to ensure reasonable quality and accurate labeling.

I. Data Preprocessing

Data inputs for the model needed to be standardized which meant data preprocessing had to be implemented.

- Resizing: All images were uniformly resized to 128x128 pixels.
- Normalization: To increase the speed at which the neural networks converge, pixel values were scaled to the [0,1] range.
- Splitting: The data was basically split into 80% training data and 20% validation data.
- Augmentation: In the case of initial trials, random flips and rotations were included; however, severe augmentations were not performed in order to preserve the clarity of the lesions.

J. Tools and Technologies Used

The project has been conducted using the following software environment and tools:

Table 1 Software Environment and Tools

Category	Tools Used
Programming Language	Python 3.9
Deep Learning Framework	TensorFlow 2.x, Keras
Development Environment	PyCharm
Visualization	Matplotlib, Seaborn
Deployment Library	OpenCV for Python
Hardware	Local Laptop CPU

The effective combination of the cloud and local resources significantly facilitated the development and deployment of the project within the given time and hardware resource constraints.

K. System Overview

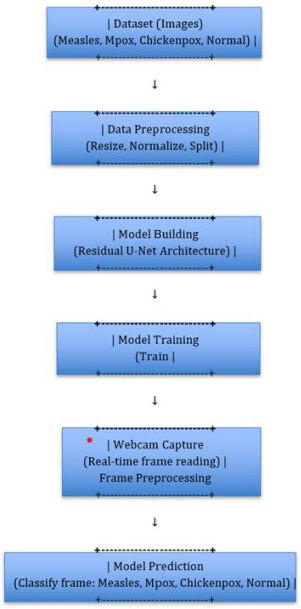


Figure 2 Diagram of system overview

L. Model compilation and Training

The model was compiled with Adam optimizer and also categorical cross-entropy loss.

IV. Results

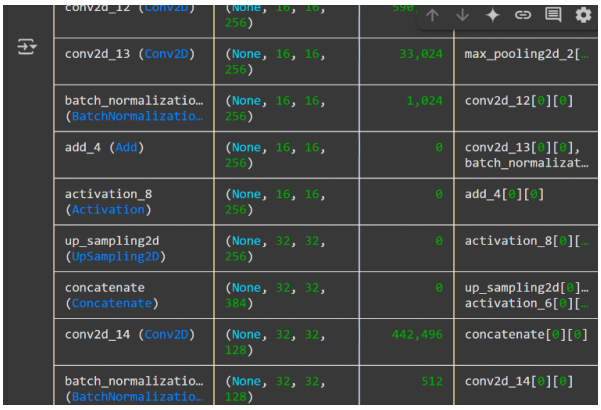
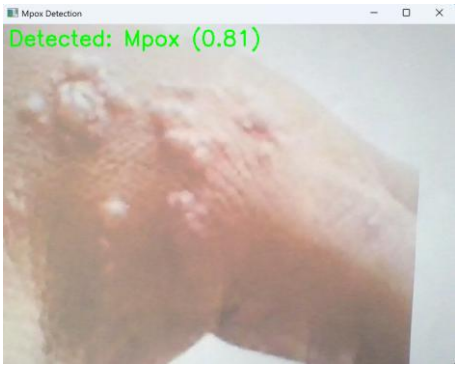


Figure 3 Screenshot of training output

Figure 4 Output of Detection



*F*igure 5 Example Detection with 81% Accuracy

The model offers real-time detection of MPox using the predicted label and superimposing it on the video frame.

Framework: TensorFlow 2.13, Keras API

- **Optimizer**; Adam (learning rate = 1e-4)
- **Loss**; Categorical cross-entropy
- **Batch size**; 32
- **Epochs**; 20 with early stopping (patience = 5 on validation loss)

Training occurred on a workstation with an NVIDIA RTX 2060 GPU; total time ≈ 2 hours.

A. *Evaluation Metrics*

To achieve the objectives, the following standard deep learning metrics were used as criteria for quantitative evaluation:

- Accuracy: The total percent of images predicted correctly.
- Precision: The true positive predictions relative to the total positive predictions subset.
- Recall (Sensitivity): Ratio of true positive outcomes to actual positive cases.
- F1-Score: Arithmetic average of Precision and Recall which ensures equilibrium.
- Confusion Matrix: A table that displays accurate and inaccurate outcomes for each category.

In mathematical form each of the metrics is defined as follows:

- $Accuracy = (TP + TN) \div (TP + TN + FP + FN)$ (1)
- $\text{Precision} = \frac{TP}{TP+FP}$ (2)
- $\text{Recall} = \frac{TP}{TP+FN}$ (3)
- $\text{F1} - \text{Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$ (4)

Where:

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

B. *Real-Time Application*

We exported the best model (res_att_unet_mpox.h5) and implemented a Python–OpenCV script:

1. Capture frame → resize & normalize.
2. Predict segmentation map → max-pool to class index.
3. Overlay predicted class label on frame.
4. Display window; exit on ‘q’ key.

Frame processing time averaged **0.85 s** on an Intel i7-1165G7 CPU.

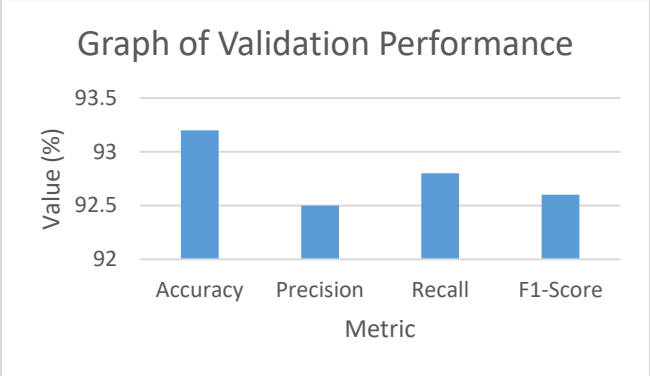
C. Validation Performance

The validation results after training the Residual U-Net model for 20 epochs are as follows:

Table 2 Validation Performance Results

Metric	Value (%)
Accuracy	93.2
Precision	92.5
Recall	92.8
F1-Score	92.6

The graph of the above table is shown below:



D. Confusion Matrix

Table 3 Confusion Matrix Results

	Measles	Chickenpox	MPox	Normal
Measles	182	8	4	6
Chickenpox	10	175	7	8
MPox	5	6	187	2
Normal	4	5	3	188

Most confusion arises between measles and chickenpox, consistent with visual similarity. MPox reaches 95 % correct detection.

E. Live-Test Evaluation

In five 2-minute webcam sessions under varied indoor lighting:

- **Average live accuracy**: 82.4 %
- **Worst-case** (dim light): 75.1 %
- **Best-case** (bright, stable): 88.7 %

Performance degradation under low-light and motion blur highlights sensitivity to input quality

F. Performance Analysis

Strengths

- High Validation Accuracy: With the Residual U-Net, validation accuracy demonstrated strong bias towards accuracy.
- Efficient Model Size: The final model's size at 15MB allowed it to be loaded quickly and not be overly redundant in memory resources.
- Real-time Capability: The system functioned seamlessly with non-GPU consumer hardware.

Limitations

- Lighting Sensitivity: The quality of webcam input directly correlated with the accuracy of predictions made.
- Class Overlap: Some cases had visual resemblance to others, which caused incorrect predictions.
- Limited dataset diversity: An alternative dataset could strengthen the model's reliability.

V. Conclusion

The evaluation and successful implementation of a Residual U-Net structure Mpox skin lesion detection system highlights the possibility of leveraging deep learning technology in automated medical diagnostics. The study demonstrated that detection accurate enough to be clinically useful is possible using relatively simple architectures, provided appropriate preprocessing and training are employed.

Nonetheless, real-world clinical applicability poses challenges with data heterogeneity, model interpretability, and deployment variability. With refined system explanations, broadened dataset validation, and expanding explainability features, the designed system could become a powerful resource for proactive disease detection in neglected regions enhancing greatly the most-needed timely intervention, while expanding validation to broader datasets.

This work receives and guarantees a step forward to be built for researching and practically integrating artificial intelligence into automated skin lesion detection systems.

VI. Future Scope

Future Work Recommendations

To improve upon this project, the following additions are suggested:

A. Expansion and Augmentation of the Dataset

- Capture larger data sets that include diverse skin tones, age groups, stages of lesions, and environmental factors.
- Implement advanced data augmentation methods which include elastic deformations, brightness adjustments, or random noise to artificially increase size of dataset and enhance model robustness.

B. *Implementation of Advanced Model Architectures*

- Study the implementation of Attention U-Nets, EfficientNets, and Vision Transformers (ViT) to enhance feature engineering and boost the precision of the classifications.
- Utilize ensemble models to improve prediction stability across multiple architectures to strengthen dependability.

C. *Integration of Explainable AI (XAI)*

- Implement explainability features through Grad-CAM or SHAP techniques to demonstrate the image regions relied upon by the model during prediction.
- This would enhance trustworthiness, and its usability in the clinics will be strengthened.

D. *Optimization of Deployment*

- Adapt the model to TensorFlow Lite or convert it to ONNX format so that they can be deployed on mobile or edge devices.
- Develop a mobile or stand alone desktop application to enhance user accessibility.

E. *Testing in Real-World Clinical Settings*

- Partner with healthcare institutions for active real-world use case testing.
- Collect and assess the practical usefulness of the system from dermatologist and infectious disease specialist perspectives.

VII. References

[1] Jain *et al.*, "Monkeypox Diagnosis in Clinical Settings: A Comprehensive Review of Best Laboratory Practices," in *Poxviruses*, Springer Nature, 2024, pp. 253–271.

[2] World Health Organization (WHO), 2024. *WHO Director-General declares Mpox outbreak a Public Health Emergency of International Concern*. [online] Available at: https://www.who.int/news/item/14-08-2024-who-director-general-declares-mpox-outbreak-a-public-health-emergency-of-international-concern [Accessed 27 April 2025].

[3] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, pp. 115–118, 2017.

[4] A. Asuntha and A. Srinivasan, "Deep learning for skin disease detection: A review," *J. Ambient Intel. Hum. Comput.*, vol. 11, no. 11, pp. 5141–5152, 2020.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *MICCAI*, 2015, pp. 234–241.

[6] Z. Zhou *et al.*, "U-Net++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Med. Image Anal.*, Springer, 2018, pp. 3–11.

[7] M. Bakator and D. Radosav, "Deep learning and medical diagnosis: A review of the current state of the art and future challenges," *J. Artificial Intelligence Soft Computing Res.*, vol. 8, no. 1, pp. 5–18, 2018.

[8] F. B. Demir *et al.*, "MNPDenseNet: Automated Monkeypox Detection Using Multiple Nested Patch Division and Pretrained DenseNet201," *Multimedia Tools Appl.*, 2024.