Detecting POX Disease in Skin Images Using Explainable Artificial Intelligence

Kalyan Kumar Jena[1], Sourav Kumar Bhoi[1], Sanjaya Kumar Panda[2], ***Senior Member*, *IEEE***
[1,2,3]Department of Computer Science and Engineering
[1,2]Parala Maharaja Engineering College (Govt.), Berhampur-761003, Odisha, India
[3]National Institute of Technology Warangal, Warangal-506004, Telangana, India
kalyan.cse@pmec.ac.in, sourav.cse@pmec.ac.in, sanjaya@nitw.ac.in

Abstract- The intricate domain of pox diseases, including chickenpox, cowpox, monkeypox, hand, foot, and mouth disease (HFMD), and measles, and their profound influence on human health is the focal point of extensive research. In response to the compelling demand for precise disease classification, an exploration is undertaken in this study, delving into the intersection of machine learning (ML) methodologies and the pursuit of interpretability through the application of explainable artificial intelligence (XAI). In this study, we use ML methods for pre-processing, then train the data and apply the XAI approach to those trained models. First, ML methods scikit-image were used to segregate the 15,000 images into train (70%), test (20%) and valid (10%). Then, we used 8 CNN models, namely AlexNet, LeNet, SeNet, GoogleNet, SpinalNet, MobileNetV1, VGG, and ZFNet, to train the model. The accuracy of GoogleNet is 82%, which is much better than that of other CNN models.

***Keywords*: Convolution Neural Network (CNN), Pox Disease, Chicken Pox, Cow Pox, Monkey Pox, Measles, Disease Classification.**

1. INTRODUCTION

Pox disease encompasses a range of highly contagious viral infections [1-3]. The earliest evidence of skin lesions resembling smallpox being one of them is found on the faces of Egyptian mummies [4-8]. It originated in 1350 BC, with cases found in Egyptian mummies. It is also known as variola or "la variole". Pox disease is a widespread health concern, but vaccination protects humans today. Pox is a complex of viral diseases in humans and other animals, marked primarily by skin and mucous membrane eruptions [9]. Pox diseases are prevalent globally and stem from various virus species. Such ailments include sheep pox, horse pox, fowl pox, cow pox, goat pox, and swine pox [10-15]. The mode of transmission varies according to the specific virus involved. It mainly influences individuals and children, manifesting signs and symptoms of regular itchy skin rashes, low-grade fever, and a feeling of preferred soreness. It causes an itchy rash, moderate fever and soreness. It spreads without difficulty through the air and by touching the rash. Although reasonable, it could cause problems for a few people. Recovering from chickenpox gives immunity. However, the virus can return later in the shape of shingles. Cases have reduced due to vaccination [16-20]. Studying chickenpox allows us to recognize and manipulate the ailment and improve vaccines.

In this work, we have used the Asus Tuf A17, which has 8 GB RAM and 512 GB SSD. It has an AMD Ryzen 4600H processor, an Nvidia GeForce GTX 1650 graphic processor, and a 64-bit operating system. We used Python 3.11 in the Jupyter Notebook integrated development environment (IDE) to embed our work. We have collected the data from Kaggle [21] and saved it in our local directory. We classified the images into three sub domains: train, test and valid. We used 8 CNN architectures to train the model. Then, we used the 5 XAI approach to get the result. We choose the ROC curve, F1 score, and accuracy as the performance matrices. The remainder of this paper is listed as follows. Section 2 shows the related work.

Section 3 presents the methodology. Section 4 shows the environmental setup and results. Section 5 concludes the work.

2. RELATED WORK

Saleh et al. [1] presented an approach for identifying individuals with monkey pox through image data analysis, marking the debut of the initial public monkey pox image classification. The procedural stages of its implementation were thoroughly discussed. Thieme et al. [6] proposed one skin lesion. They observed a high detection method to demonstrate the accuracy and receiver operating characteristic (ROC) curve of predicting an instance of several strategies in image processing, model selection and XAI technique. Chadaga et al. [10] proposed chicken pox and monkey pox detection using skin lesion images. Deep learning models were shown to be highly effective.

Nayak et al. [11] offered an image classification modality. They have used the four-way classification of skin lesion images, including monkey pox. The XAI techniques can also aid medical image classification. They proposed pre-trained deep-learning networks to diagnose monkey pox. XAI techniques local interpretable model-agnostic explanations (LIME) and GradCAM enable the visual interpretation. Campana et al. [20] proposed a deep learning-based m-health solution to detect monkey poxes from the skin lesion field of CNN for medical image analysis and, in particular, into XAI techniques to make AI more predictable.

3. METHODOLOGY

3.1 DATA COLLECTION AND PRE-PROCESSING

We have collected 15000 images from Kaggle source [21]. These images belong to different pox sub domains like chicken, monkey, cow, measles, HFD, and healthy as shown in Figure 1. All the images have other characteristics and properties like height, width and quality. We use Scikit-image for the classification of the images. Scikit-image is an open-source image processing library designed for the Python programming language. The library encompasses a range of algorithms tailored for tasks such as segmentation, geometric transformations, color space manipulation, analysis, filtering, morphology, feature detection, and beyond.

We set the height and width in all images as 224 x 224. We segregate the images into train, test, and valid subdirectories. After the successful segregation, it is seen that the train contains 10000 images, the test contains 3000 images, and the valid one contains 2000 images as listed in Table 1. They are saved in our local directory with train, test, and valid names.

Table 1: Tabulation data model for image classification

Input Data	Image Classification Algorithm using ML	Output Data
POX Images (15000)	scikit-image, scikit-learn (224 x 224)	Train: 10000 Test: 3000 Valid: 2000

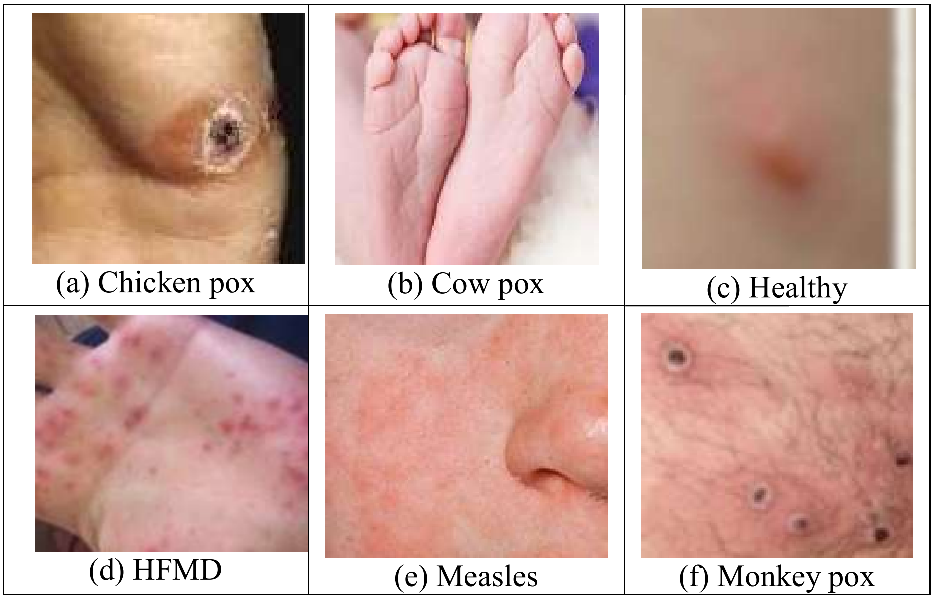

Figure 1: Image classification of data pre-processing.

3.3 MODEL TRAINING

In the model training phase, we used 8 CNN architectures described as follows. Le Net is a pioneering CNN architecture designed for handwritten digit recognition. It extracts features and pooling layers to reduce dimensionality, enabling efficient pattern recognition in images using convolution layers. It is used for handwritten digit recognition, image classification, and pattern recognition. Google Net is a deep-learning model for image recognition. It utilizes multiple layers to enhance accuracy, making it a powerful tool in computer vision applications. It is used in image recognition and optimizing accuracy through innovative convolution layer configurations. Mobile Net is a deep learning architecture that provides efficient and accurate image recognition, making it ideal for various real-time applications on smart phones and other portable gadgets. It is used in mobile apps for object detection and image recognition tasks.

ZF Net is a deep learning architecture that introduced deeper neural networks, improving image recognition accuracy and paving the way for advanced computer vision applications. Its advanced convolution layers enhance accuracy, making it valuable in computer vision applications like object detection and facial recognition. Alex Net is a pioneering convolution neural network analyzed for large-scale image recognition tasks. It showcases deep learning's potential to revolutionize computer vision applications. It is widely used for tasks like object recognition; its deep layers enhance accuracy, enabling advanced applications in computer vision. Spinal Net operates as a learning system that undergoes a series of steps to formulate predictions. It begins with input data intake, proceeds to generate a prediction, and concludes by comparing the prediction with the desired output. Visual geometry group (VGG) is a renowned deep learning model characterized by its simplicity and depth. It excels in image recognition tasks, utilizing deep convolution layers to capture intricate patterns, enhancing accuracy. It accurately identifies objects, making it indispensable in healthcare, robotics, and autonomous vehicles. Se Net enhances deep learning models by focusing on essential features. Its adaptive recalibration technique improves accuracy in various computer analysis tasks, ensuring efficient image recognition. It emphasizes crucial features, boosting accuracy in object detection and image classification tasks. We use five XAI approaches. In the CNN architecture, we got different results. However, we predict the ROC curve and accuracy as the performance matrix.

4. ENVIRONMENTAL SETUP AND RESULTS

In this experiment, we used 15000 pox images downloaded from Kaggle. These images belong to different pox sub domains like chicken, monkey pox, cow, measles, HFMD, and healthy. All the images have different characteristics and properties, like height, width, and quality. We use scikit-image to classify the images into three categories, with the names train, test, and valid and set the image size as 224 x 224. Then we got 10000 images for training, 3000 for testing, and 20000 for validation. Then, we used the 8 CNN architectures, such as Spinal Net, Alex Net, Google Net V1, SE Net, ZF Net, VGG, Le Net, and Mobile Net, to train our model. We moved to the XAI approach, called LIME, for analysis. We got results by measuring the CNN and XAI, i.e., explanations for model predictions, which helps understand why a model makes a specific prediction for a particular data instance. LIME is a technique used in ML to make a specific decision for a particular input, such as classifying an image. It provides local and interpretable explanations for model predictions. The ROC curve of LIME is shown in Figure 2.

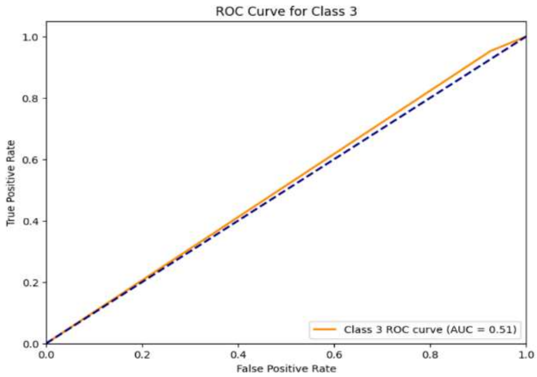

Figure 2: ROC curve of LIME.

In the Lime XAI model, we took eight models, i.e., Le Net, Alex Net, Google Net, Mobile Net, ZF Net, Spinal Net, VGG, and Se Net. The values of all eight models are defined based on the following parameters: test accuracy, test loss, XAI accuracy, precision, recall, and F1 score. Le Net achieves test accuracy of 35.54%, test loss of 2.6808, XAI accuracy of 37.36%, precision of 0.9316, recall of 0.1865, and F1 score of 0.3401. Alex Net achieves a test accuracy of 34.06%, test loss of 2.6808, XAI accuracy of 31.92%, precision of 0.9368, recall of 0.2676, and F1 score of 0.4103. Google Net achieves a test accuracy of 81.69%, test loss of 1.5867, XAI accuracy of 83.27%, precision of 0.9809, recall of 0.8617 and F1 score of 0.6532. Mobile net achieves test accuracy of 46.04%, test loss of 2.7584, XAI accuracy of 50.92%, precision of 0.9434, recall of 0.0758 and F1 score of 0.1256.

ZF net produces test accuracy of 18.69%, test loss of 1.9637, XAI accuracy of 22.82%, precision of 0.9040, recall of 0.2792 and F1 score of 0.1745. Spinal Net achieves test accuracy of 42.84%, test loss of 1.6438, XAI accuracy of 39.45%, precision of 0.9630, recall of 0.4833 and F1 score of 0.4263. VGG produces test accuracy of 41.48%, test loss of 1.6384, XAI accuracy of 43.48%, precision of 0.9620, recall of 0.1834 and F1 score of 0.2463. Se Net produces a test accuracy of 23.74%, test loss of 1.8248, XAI accuracy of 22.76%, precision of 0.9268, recall of 0.2179 and F1 score of 0.1896. We got the best test accuracy in the Google Net model, i.e., 81.69%. The results are summarized in Table 2.

Table 2: Comparison of results in various models with XAI

XAI	Model	Test Accuracy	Test Loss	XAI Accuracy	Precision	Recall	F1 Score
LIME	Le Net	35.54%	2.6808	37.36%	0.9316	0.1865	0.3401
	Alex Net	34.06%	2.5957	31.92%	0.9368	0.2676	0.4103
	Google Net / Inception Net	**81.69%**	1.5867	83.27%	0.9809	0.8617	0.6532
	Mobile Net V1	46.04%	2.7584	50.92%	0.9434	0.0758	0.1256
	ZF Net	18.69%	1.9637	22.82%	0.9049	0.2792	0.1745
	Spinal Net	42.84%	1.6438	39.45%	0.9630	0.4833	0.4263
	VGG	41.48%	1.6384	43.48%	0.9620	0.1834	0.2463
	Se Net (Squeeze and Excitation Network)	23.74%	1.8248	22.76%	0.9268	0.2179	0.1896

5. CONCLUSION

In this work, we used Python 3.11 language to experiment with our work and used Jupyter Notebook IDE and different XAI approaches to classify the diseases. We collected the dataset from the Kaggle source, downloaded it, and saved it in our directory. We trained the model using 8 CNN architecture and also re-trained the model using the Tensor flow/Pytorch. We used an XAI approach, namely LIME. It is seen that Google Net shows a high accuracy of 82% concerning other CNN models. LIME provides local and interpretable explanations for model predictions.

REFERENCES

1. Saleh, A. I., & Rabie, A. H. (2023). Human monkeypox diagnose (HMD) strategy based on data mining and artificial intelligence techniques. Computers in Biology and Medicine, 152, 106383.
2. Farzipour, A., Elmi, R., & Nasiri, H. (2023). Detection of Monkeypox cases based on symptoms using XGBoost and Shapley additive explanations methods. Diagnostics, 13(14), 2391.
3. Lakshmi, M., & Das, R. (2023). Classification of Monkeypox Images Using LIME-Enabled Investigation of Deep Convolutional Neural Network. Diagnostics, 13(9), 1639.
4. Sharma, N., Mohanty, S. N., Mahato, S., & Pattanaik, C. R. A novel dataset and local interpretable model-agnostic explanations (LIME) for monkeypox prediction. Intelligent Decision Technologies, (Preprint), 1-12.
5. Yasmin, F., Hassan, M. M., Hasan, M., Zaman, S., Kaushal, C., El-Shafai, W., & Soliman, N. F. (2023). PoxNet22: A fine-tuned model for the classification of monkeypox disease using transfer learning. IEEE Access, 11, 24053-24076.
6. Thieme, A. H., Zheng, Y., Machiraju, G., Sadee, C., Mittermaier, M., Gertler, M., & Gevaert, O. (2023). A deep-learning algorithm to classify skin lesions from mpox virus infection. Nature medicine, 29(3), 738-747.
7. GhoshRoy, D., Alvi, P. A., & Santosh, K. C. (2022). Explainable AI to Predict Male Fertility Using Extreme Gradient Boosting Algorithm with SMOTE. Electronics, 12(1), 15.
8. Sizikova, E., Vendrow, J., Cao, X., Grotheer, R., Haddock, J., Kassab, L., & Needell, D. (2022). Automatic Infectious Disease Classification Analysis with Concept Discovery. arXiv preprint

arXiv:2209.02415.
9. Kundu, D., Siddiqi, U. R., & Rahman, M. M. (2022, December). Vision transformer based deep learning model for monkeypox detection. In 2022 25th International Conference on Computer and Information Technology (ICCIT) (pp. 1021-1026). IEEE.
10. Chadaga, K., Prabhu, S., Sampathila, N., Chadaga, R., KS, S., & Sengupta, S. (2022). Predicting cervical cancer biopsy results using demographic and epidemiological parameters: A custom stacked ensemble machine learning approach. Cogent Engineering, 9(1), 2143040.
11. Nayak, T., Chadaga, K., Sampathila, N., Mayrose, H., Gokulkrishnan, N., Prabhu, S., & Umakanth, S. (2023). Deep learning based detection of monkeypox virus using skin lesion images. Medicine in Novel Technology and Devices, 100243.
12. Athina, F. H., Sara, S. A., Sarwar, Q. S., Tabassum, N., Era, M. T. J., Ashraf, F. B., & Hossain, M. I. (2022, November). Multi-classification Network for Detecting Skin Diseases using Deep Learning and XAI. In 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT) (pp. 648-655). IEEE.
13. Ahsan, M. M., Abdullah, T. A., Ali, M. S., Jahora, F., Islam, M. K., Alhashim, A. G., & Gupta, K. D. (2022). Transfer learning and Local interpretable model agnostic based visual approach in Monkeypox Disease Detection and Classification: A Deep Learning insights. arXiv preprint arXiv:2211.05633.
14. Ahsan, M. M., Abdullah, T. A., Ali, M. S., Jahora, F., Islam, M. K., Alhashim, A. G., & Gupta, K. D. (2022). Transfer learning and Local interpretable model agnostic based visual approach in Monkeypox Disease Detection and Classification: A Deep Learning insights. arXiv preprint arXiv:2211.05633.
15. Ahsan, M. M., Ali, M. S., Hassan, M. M., Abdullah, T. A., Gupta, K. D., Bagci, U., & Soliman, N. F. (2023). Monkeypox diagnosis with interpretable deep learning. IEEE Access.
16. Aggarwal, P., Mishra, N. K., Fatimah, B., Singh, P., Gupta, A., & Joshi, S. D. (2022). COVID-19 image classification using deep learning: Advances, challenges and opportunities. Computers in Biology and Medicine, 144, 105350.
17. Kawamleh, S. (2023). Against explainability requirements for ethical artificial intelligence in health care. AI and Ethics, 3(3), 901-916.
18. Khan, G. Z., & Ullah, I. (2023). Efficient technique for monkeypox skin disease classification with clinical data using pre-trained models. Journal of Innovative Image Processing, 5(2), 192-213.
19. Ahamed, B. S., Usha, R., &Sreenivasulu, G. (2022, October). A Deep Learning-based Methodology for Predicting Monkey Pox from Skin Sores. In 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon) (pp. 1-6). IEEE.
20. Campana, M. G., Colussi, M., Delmastro, F., Mascetti, S., & Pagani, E. (2023). A Transfer Learning and Explainable Solution to Detect mpox from Smartphones images. arXiv preprint arXiv:2305.18489.
21. Pox disease, accessed on Sep. 2023, https://www.kaggle.com/datasets/joydippaul/mpox-skin-lesion-dataset-version-20-msld-v20.