bioRxiv preprint doi: https://doi.org/10.1101/2022.08.08.503193; this version posted August 9, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Can Artificial Intelligence Detect Monkeypox from Digital Skin Images?

Towhidul Islam, Mohammad Arafat Hussain, Forhad Uddin Hasan Chowdhury, and B.M. Riazul Islam,

***Abstract*—An outbreak of Monkeypox has been reported in 75 countries so far, and it is spreading in fast pace around the world. The clinical attributes of Monkeypox resemble those of Smallpox, while skin lesions and rashes of Monkeypox often resemble those of other poxes, for example, Chickenpox and Cowpox. These similarities make Monkeypox detection challenging for healthcare professionals by examining the visual appearance of lesions and rashes. Additionally, there is a knowledge gap among healthcare professionals due to the rarity of Monkeypox before the current outbreak. Motivated by the success of artificial intelligence (AI) in COVID-19 detection, the scientific community has shown an increasing interest in using AI in Monkeypox detection from digital skin images. However, the lack of Monkeypox skin image data has been the bottleneck of using AI in Monkeypox detection. Therefore, recently, we introduced the Monkeypox Skin Image Dataset 2022, the largest of its kind so far. In addition, in this paper, we utilize this dataset to study the feasibility of using state-of-the-art AI deep models on skin images for Monkeypox detection. Our study found that deep AI models have great potential in the detection of Monkeypox from digital skin images (precision of 85%). However, achieving a more robust detection power requires larger training samples to train those deep models.**

***Index Terms*—Monkeypox, skin image data, deep learning, CNN, artificial intelligence.**

I. Introduction

MONKEYPOX virus has been spreading throughout the world at a rapid rate, while the world is still recovering from the aftermath of Coronavirus disease (COVID-19). Monkeypox is an infectious disease and its recent outbreak has been reported in at least 75 countries to date. The first human infection by the Monkeypox virus was reported in the Democratic Republic of Congo (formerly Zaire) in 1970 [1]. The Monkeypox virus belongs to the genus *Orthopoxvirus* of the family *Poxviridae* [2], which was first transmitted from animals to humans. Typically, Monkeypox infection shows symptoms similar to those of Smallpox infection [1]. Smallpox was largely eradicated in 1970, leading to the cessation of Smallpox vaccination. Since the 1970s, Monkeypox has been considered the most dangerous *Orthopoxvirus* for human health. In the past, Monkeypox was most frequent in the African continent, however, it has often been reported in urban areas outside the African continent now a days [1]. Scientists believe that the current Monkeypox outbreak in humans on a global scale is due to changes in biological attributes of the Monkeypox virus, changes in human lifestyle, or both [3].

The clinical attributes of Monkeypox are similar to that of Smallpox, however less severe [4]. On the other hand, skin lesions and rashes, caused by Monkeypox infection, often resemble those of Chickenpox and Cowpox. This clinical and visual similarity among different pox infections makes the early diagnosis of Monkeypox challenging for the healthcare professional. In addition, the rarity of Monkeypox infection in humans before the current outbreak [5] created a knowledge gap among healthcare professionals around the world. For the diagnosis of Monkeypox infection, the polymerase chain reaction (PCR) test is generally considered the most accurate tool [6], although healthcare professionals are often used to diagnosing pox infections by visual observation of skin rashes and lesions. Monkeypox infection has a low mortality rate (i.e., 1%-10%) [7], however, early detection of Monkeypox may help in patient isolation and contact tracing for effective containment of community spread of Monkeypox.

Different artificial intelligence (AI) tools, especially deep learning approaches, have been widely used in different medical image analysis tasks (e.g., organ localization [8], [9], organ abnormality detection [9], [10], gene mutation detection [11], cancer grading [12], [13] and staging [14]) in the last decade. Notably, AI methods have recently played a significant role in COVID-19 diagnosis and severity ranking from multimodal medical images (e.g., computed tomography (CT), chest X-ray, and chest ultrasound) [15]–[17]. This success motivates the scientific community in utilizing AI approaches for Monkeypox diagnosis from the digital skin images of patients. It is well-known that supervised or semi-supervised AI approaches are data-driven, which require a large number of data for developing AI methods effectively. However, there is no publicly available and reliable digital image database of Monkeypox skin lesions or rashes. To urgently tackle this situation, we recently used the web scraping (i.e., extracting data from websites) [18] approach to collect digital skin lesion images of Monkeypox, Chickenpox, Smallpox, Cowpox, and Measles [19] to facilitate the development of AI-based Monkeypox infection detection algorithms. We made this dataset (namely, Monkeypox Skin Image Dataset 2022) public and freely accessible in Kaggle.[1] However, the immediate question follows if we can utilize state-of-the-art AI techniques on digital skin images of rashes and lesions to accurately identify

T Islam is with the Department of Computer Science and Engineering, Northern University Bangladesh, Dhaka, 1215, Bangladesh. Email: towhidul.islam@nub.ac.bd

FUH Chowdhury is with the Department of Medicine, Dhaka Medical College Hospital, Dhaka, 1000, Bangladesh.

BMR Islam is with the Health Information Unit, Directorate General of Health Services, Dhaka, 1212, Bangladesh.

[1]https://www.kaggle.com/datasets/arafathussain/monkeypox-skin-image-dataset-2022

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.08.503193; this version posted August 9, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Monkeypox and distinguish it from infections by other types of poxes and diseases (e.g., Chickenpox, Smallpox, Cowpox, Measles, etc.). To our knowledge, only two AI-based Monkeypox detection studies have been conducted to date that appear as preprints [20], [21]. However, these studies have several limitations. First, these studies included only three cases of disease (i.e., Monkeypox, Chickenpox, and Measles). Second, these studies are conducted on very small datasets. Third, these studies tested one [20] or a few [21] AI deep models in pox classification tasks.

In this paper, we test the feasibility of using state-of-the-art AI techniques to classify different types of pox from digital skin images of pox lesions and rashes. The novelties of this work are the following.

1) We utilize a database that contains skin lesion/rash images of 5 different diseases (compared to 3 diseases in [20], [21], i.e., Monkeypox, Chickenpox, Smallpox, Cowpox, and Measles, as well as contains healthy skin images.
2) Our database contains more data for pox, measles and healthy images scraped on the Web (i.e., 804 images), before augmentation, compared to other similar databases used in [20], [21] (204 and 228 images, respectively).
3) We tested the disease classification power of seven state-of-the-art deep models (compared to one and three deep models in [20] and [21], respectively) from digital skin images. We tested the disease classification performance of ResNet50 [22], DenseNet121 [23], Inception-V3 [24], SqueezeNet [25], MnasNet-A1 [26], MobileNet-V2 [27], and ShuffleNet-V2 [28].
4) We performed 5-fold cross-validation tests for each of the AI deep models to more comprehensively analyze our findings, which is not done in previous studies [20], [21], although their dataset is smaller than ours.

II. Methodology

In this section, we describe our approach to data collection, data augmentation, and experimental setup.

A. Data Collection

1) Web-scraping for Image Collection: We use web scraping to collect skin infected with Monkeypox, Chickenpox, Smallpox, Cowpox and Measles, as well as healthy skin images from various sources, such as websites, news portals, blogs, and image portals using the Google search engine. We show the pipeline of our database development in Fig. 1. We searched to collect images that fall under "Creative Commons licenses." However, for several pox classes, we hardly find images. Therefore, we also collect images that fall under "Commercial & other licenses." Therefore, we include supplementary material that includes the uniform resource locator (URL) of the source, the access date, and the photo credit (if any) for all our collected images. In Fig. 2, we show some example images from our database.

Fig. 1. The pipeline of our database development.

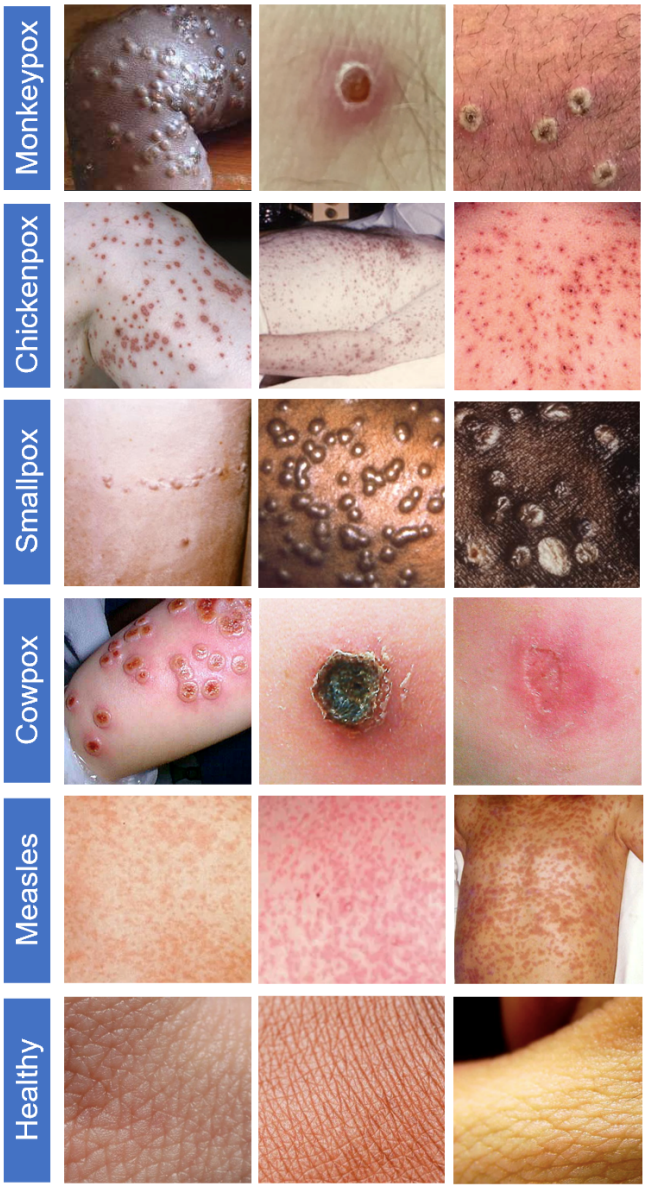

Fig. 2. Example skin images of Monkeypox, Chickenpox, Smallpox, Cowpox, Measles, and healthy cases (first to sixth rows, respectively) from our database.

2) Expert Screening: Two expert physicians, who were experts in infectious diseases, screened all images collected to validate the supposed infection. In Fig. 3, we show a pie chart of the percentage of original web-scraped images per class in our dataset (after expert screening).

3) Data Preprocessing: We cropped images to remove unwanted background regions and blocked the eye region with black boxes to make patients nonidentifiable from their corresponding images. We also did the same to hide revealed private parts. Since typical AI deep models take square-shaped images as inputs in terms of pixel counts (often 224×224×3

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.08.503193; this version posted August 9, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

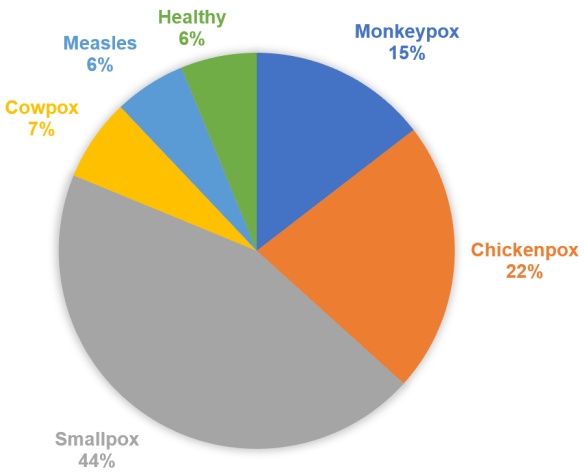

Fig. 3. A pie chart showing the percentage of original web-scraped images per class in our dataset.

pixels), we added extra blank pixels in the periphery of many images to avoid excessive stretching of the actual skin lesions during image resizing. Finally, we cropped and resized all the images to 224×224×3 pixels using bilinear interpolation.

4) Augmentation: We performed 19 augmentation operations on the web-scrapped images using Python Imaging Library (PIL) version 9.2.0, and scikit-image library version 0.19.3. to increase the number of images and introduce variability in the data. Our augmentation operations include (1) brightness modification with a randomly generated factor (range [0.5, 2]), (2) color modification with a randomly generated factor (range [0.5, 1.5]), (3) sharpness modification with a randomly generated factor (range [0.5, 2]), (4) image translation along height and width with a randomly generated distance between -25 and 25 pixels, (5) image shearing along height and width with randomly generated parameters, (6) adding Gaussian noise of zero mean and randomly generated variance (range [0.005, 0.02]) to images, (7) adding zero-mean speckle noise and randomly generated variance (range [0.005, 0.02]) to images, (8) adding salt noise to randomly generated number of image pixels (range [2%, 8%]), (9) adding pepper noise to the randomly generated number of image pixels (range [2%, 8%]), (10) adding salt & pepper noise to randomly generated number of image pixels (range [2%, 8%]), (11) modifying the values of the image pixels based on the local variance, (12) blurring an image with a Gaussian kernel with a randomly generated radius (range [1, 3] pixels), (13) contrast modification with a randomly generated factor (range [1, 1.5]), (14) rotating all images by 90°, (15) rotating images at random angle (range [-45°, 45°]), (16) zooming in an image by about 9%, (17) zooming in an image by about 18%, (18) flipping images along the height, and (19) flipping images along the width. In Table I, we show the number of original and augmented images per class in our database. We also show example images after augmentation in Fig. 4. We used these 19 augmentation operations in different combinations that increased the data by 49×. We made this dataset publicly available in Kaggle [19].[2]

TABLE I
DISTRIBUTION OF IMAGE CLASSES IN OUR MONKEYPOX SKIN IMAGE DATASET 2022.

Class	# Original Images	# Augmented Images
Monkeypox	117	5,733
Chickenpox	178	8,722
Smallpox	358	17,542
Cowpox	54	2,646
Measles	47	2,303
Healthy	50	2,450
Total	804	39,396

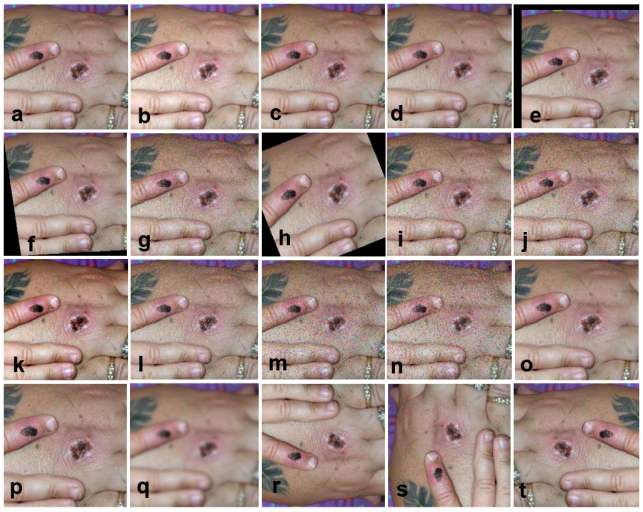

Fig. 4. Illustration of (a) an original image and corresponding augmented images by (b) brightness modification, (c) color modification, (d) sharpness modification, (e) translation (f) shear, (g) adding Gaussian noise, (h) random rotation, (i) adding speckle Noise, (j) adding local variance noise, (k) contrast modification (l) adding salt noise, (m) adding pepper noise, (n) adding salt & pepper noise, (o) 9% zooming in, (p) 18% zooming in, (q) Gaussian blurring, (r) flipping along the height, (s) rotation by 90°, and (t) flip along the width.

B. Experimental Setup

1) Deep Models: We implement seven deep AI models, namely ResNet50 [22], DenseNet121 [23], Inception-V3 [24], SqueezeNet [25], MnasNet-A1 [26], MobileNet-V2 [27], and ShuffleNet-V2-1× [28], to check their feasibility in pox classification task. These models varies in terms of the number of trainable parameters as shown in Table II. Since our dataset is small, we choose deep models of different size ranges (number of trainable parameters ranges 1-26 Millions) to check the effect of training sample size.

2) Fine-tuning of Deep Models in Cross-validation: Since our digital skin image dataset is small, we use ImageNet-based pretrained weights to initialize our deep models and fine-tune on our digital skin image data in 5-fold cross-validation settings. We split our preprocessed original images (shown in Table I second column) in 5 equal folds per class. Since there is a class imbalance among our original data, we use different

[2]https://www.kaggle.com/datasets/arafathussain/monkeypox-skin-image-dataset-2022

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.08.503193; this version posted August 9, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

TABLE II
APPROXIMATE NUMBER OF TRAINABLE PARAMETERS IN MILLIONS (M) OF THE SEVEN DEEP AI MODELS USED IN THIS STUDY.

Model	# Parameters (M)
ResNet50 [22]	25.6
Inception-V3 [24]	23.8
DenseNet121 [23]	7.2
MnasNet-A1 [26]	3.9
MobileNet-V2 [27]	3.4
ShuffleNet-V2-1× [28]	2.3
SqueezeNet [25]	1.2

number of augmented images per class during training to make the training data balanced, i.e., all the classes have about same number of training image samples (∼1,700 images per class). We also make sure that augmented images of an original image do not get split into different folds during cross-validation training. In Table III, we show the distribution of the number of images used in validation per class and the number of augmented images used in training per class in each fold.

TABLE III
NUMBER OF IMAGES PER CLASS IN EACH FOLD IN A 5-FOLD CROSS-VALIDATION SETUP.

Class	# Validation Images (Validated on)	# Training Images	# Augmented Images (Trained on)
Monkeypox	23	92	1,748 (=92×19)
Chickenpox	35	140	1,680 (=140×12)
Smallpox	71	284	1,704 (=284×6)
Cowpox	10	40	1,760 (=40×44)
Measles	9	36	1,764 (=36×49)
Healthy	10	40	1,760 (=40×44)

C. Computation Setup

We schedule fine-tuning each deep model for 100 epochs/fold, and the whole 5-fold cross-validation was supposed to take about 1 day/model. We picked the best validation performance in each fold. We chose Adam optimizer with a learning rate of 0.001. We also chose a batch size of 16 images. We implemented our models in PyTorch version 1.6.0 and Python version 3.8.10. Training was performed on a workstation with Intel E5-2650 v4 Broadwell 2.2 GHz processor, an Nvidia P100 Pascal GPU with 16 GB of VRAM, and 64 GB of RAM.

III. RESULTS

In this section, we present the comparative classification performance of our seven state-of-the-art deep AI models as summarized in Table II. In Figs. 5, 6, 7, 8, 9, 10, and 11, we show the confusion matrices for 5-fold cross-validation predictions by ResNet50, Inception-V3, DenseNet121, MnasNet-A1, MobileNet-V2, ShuffleNet-V2-1×, and SqueezeNet, respectively. We see many miss-classifications for disease classes in these figures for all the deep models, except for the healthy skin class. The least number of missclassifications for all 5 folds is seen for ShuffleNet-V2 (see Fig. 10).

We also present the quantitative comparison of mean precision, mean recall, mean F1 score and mean accuracy, estimated over the 5-fold cross-validation, for all classes by all deep AI models in Table IV. Here, we see that the best accuracy is achieved by the ShuffleNet-V2 (79%). ShuffleNet-V2 has fewer trainable parameters (i.e., 2.3 M) than ResNet50 (25.6 M), Inception-V3 (23.8 M), DenseNet121 (7.2 M), MnasNet-A1 (3.9 M) and MobileNet-V2 (3.4 M) (see Table II). From our observation on the prediction performance by different deep models, as summarized in Table IV, we hypothesize that models with a larger number of trainable parameters may be underfitted due to the small training sample size. On the other hand, although the SqueezeNet has fewer number of trainable parameters (i.e., 1.2 M) than the ShuffleNet-V2 (2.3 M), it is overfitted on the small number of training samples, thus resulted in worse prediction performance on the validation data.

TABLE IV
QUANTITATIVE COMPARISON OF MEAN PRECISION, MEAN RECALL, MEAN F1 SCORE AND MEAN ACCURACY OVER THE 5-FOLD CROSS-VALIDATION.

Methods	Class	Mean Precision	Mean Recall	Mean F1 Score	Mean Accuracy
ResNet50	Chickenpox	0.80	0.58	0.68	0.72
	Cowpox	0.58	0.70	0.64	
	Healthy	0.88	0.98	0.92	
	Measles	0.38	0.44	0.41	
	Monkeypox	0.59	0.51	0.55	
	Smallpox	0.77	0.86	0.81	
Inception-V3	Chickenpox	0.73	0.70	0.71	0.71
	Cowpox	0.70	0.70	0.70	
	Healthy	0.93	1.00	0.96	
	Measles	0.45	0.51	0.48	
	Monkeypox	0.71	0.53	0.61	
	Smallpox	0.77	0.83	0.80	
DenseNet121	Chickenpox	0.76	0.74	0.75	0.78
	Cowpox	0.64	0.72	0.68	
	Healthy	0.94	0.98	0.96	
	Measles	0.50	0.60	0.55	
	Monkeypox	0.71	0.53	0.61	
	Smallpox	0.83	0.87	0.85	
MnasNet-A1	Chickenpox	0.70	0.61	0.65	0.72
	Cowpox	0.58	0.72	0.64	
	Healthy	1.00	0.92	0.96	
	Measles	0.44	0.47	0.45	
	Monkeypox	0.63	0.41	0.50	
	Smallpox	0.77	0.88	0.82	
MobileNet-V2	Chickenpox	0.77	0.72	0.74	0.77
	Cowpox	0.57	0.78	0.66	
	Healthy	0.94	1.00	0.97	
	Measles	0.44	0.47	0.45	
	Monkeypox	0.73	0.62	0.67	
	Smallpox	0.85	0.86	0.85	
ShuffleNet-V2	Chickenpox	0.79	0.67	0.72	**0.79**
	Cowpox	0.58	0.78	0.67	
	Healthy	1.00	0.98	0.99	
	Measles	0.54	0.71	0.62	
	Monkeypox	0.79	0.58	0.67	
	Smallpox	0.83	0.89	0.86	
SqueezeNet	Chickenpox	0.67	0.54	0.60	0.65
	Cowpox	0.45	0.60	0.52	
	Healthy	0.94	0.92	0.93	
	Measles	0.20	0.40	0.27	
	Monkeypox	0.55	0.44	0.49	
	Smallpox	0.79	0.78	0.78	

We also present classification performance by an ensemble of our implemented and fine-tuned seven deep models, where we use majority voting to make the prediction decision. In Fig. 12, we show the confusion matrices for 5-fold cross-validation predictions by the ensemble of ResNet50, Inception-V3, DenseNet121, MnasNet-A1, MobileNet-V2, ShuffleNet-

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.08.503193; this version posted August 9, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

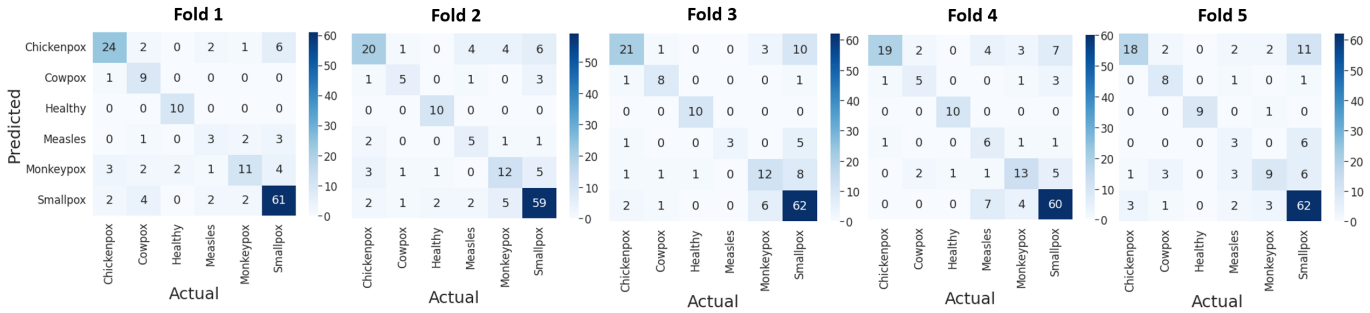

Fig. 5. Confusion matrices of prediction by ResNet50 in 5-fold cross-validation.

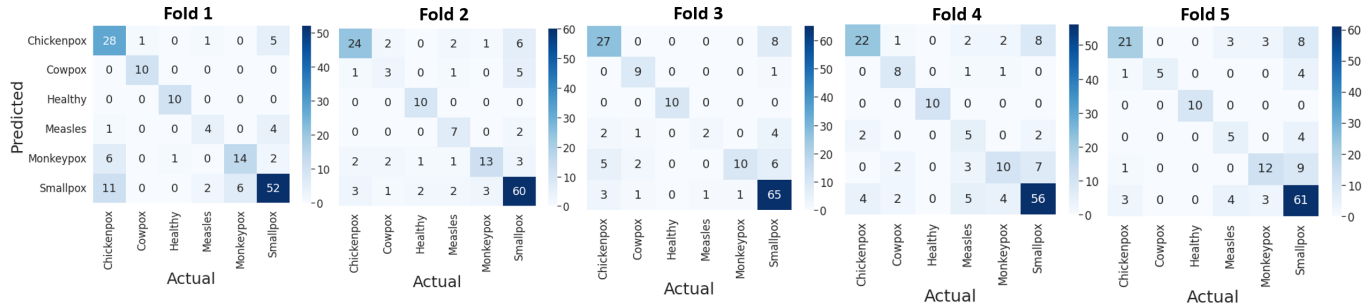

Fig. 6. Confusion matrices of prediction by Inception-V3 in 5-fold cross-validation.

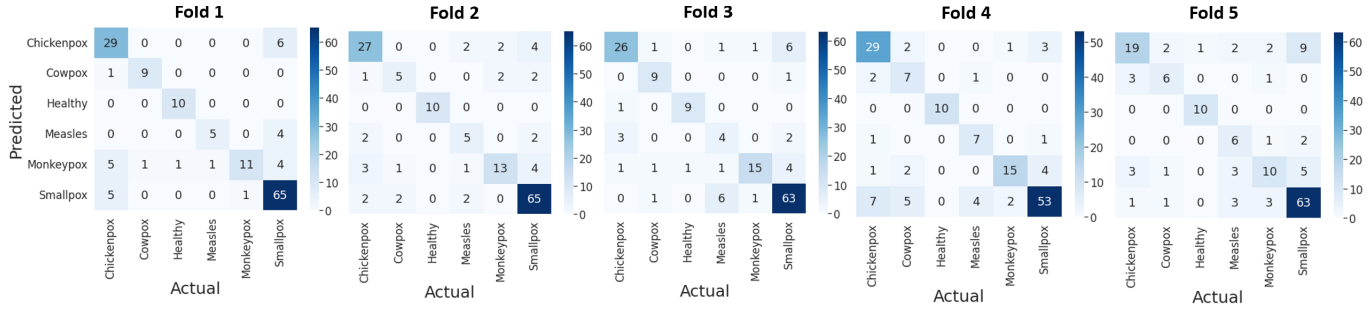

Fig. 7. Confusion matrices of prediction by DenseNet121 in 5-fold cross-validation.

V2-1× and SqueezeNet. We see in these confusion matrices that the ensemble approach reduces the missclassifications significantly, compared to those seen in Figs. 5, 6, 7, 8, 9, 10 and 11 for individual deep models. Furthermore, we present the quantitative comparison of mean precision, mean recall, mean F1 score, and mean accuracy, estimated over the 5-fold cross-validation, for all classes using the ensemble approach in Table V. We see in this table that the ensemble approach shows the best performance in terms of all the metrics, notably in terms of accuracy (83%), compared to those by individual deep models as summarized in Table IV.

TABLE V
QUANTITATIVE PERFORMANCE BY THE ENSEMBLE METHOD IN TERMS OF MEAN PRECISION, MEAN RECALL, MEAN F1 SCORE, AND MEAN ACCURACY OVER THE 5-FOLD CROSS-VALIDATION.

Methods	Class	Mean Precision	Mean Recall	Mean F1 Score	Mean Accuracy
Majority Voting	Chickenpox	0.81	0.75	0.78	0.83
	Cowpox	0.74	0.84	0.79	
	Healthy	0.98	1.00	0.99	
	Measles	0.58	0.58	0.58	
	Monkeypox	0.85	0.61	0.71	
	Smallpox	0.85	0.94	0.89	

IV. CONCLUSION

In this paper, we tested the feasibility of using seven state-of-the-art AI deep models in classifying different pox types and Measles from the digital skin images of lesions and rashes. We built and utilized a digital skin database containing skin lesion/rash images of five different diseases, namely, Monkeypox, Chickenpox, Smallpox, Cowpox, and Measles. We also chose seven state-of-the-art deep models that vary in terms of the number of trainable parameters, ranging from 1 to 26 million. Our 5-fold cross-validation experiments showed that AI deep models have the ability to distinguish among different pox types using digital skin images of pox/Measles lesions and rashes. We also observed that deep models tend to overfit or underfit, perhaps due to the trade-off between the training sample size and the number of trainable parameters in a model. Thus, to achieve better classification accuracy

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.08.503193; this version posted August 9, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

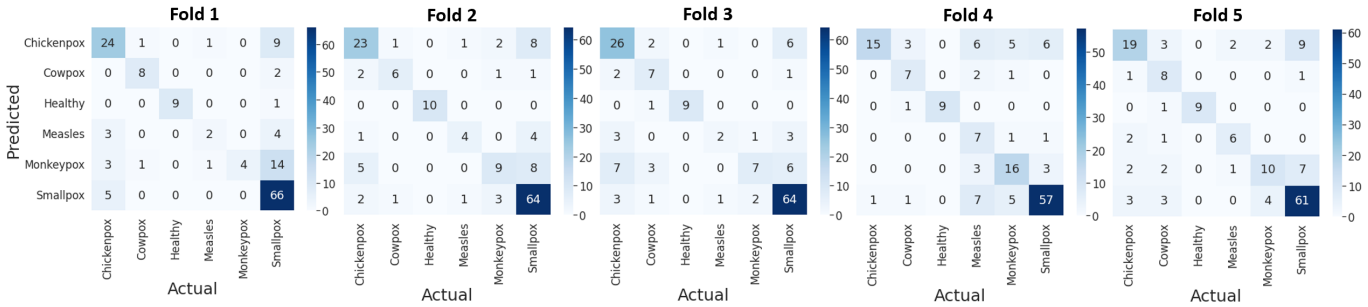

Fig. 8. Confusion matrices of prediction by MnasNet-A1 in 5-fold cross-validation.

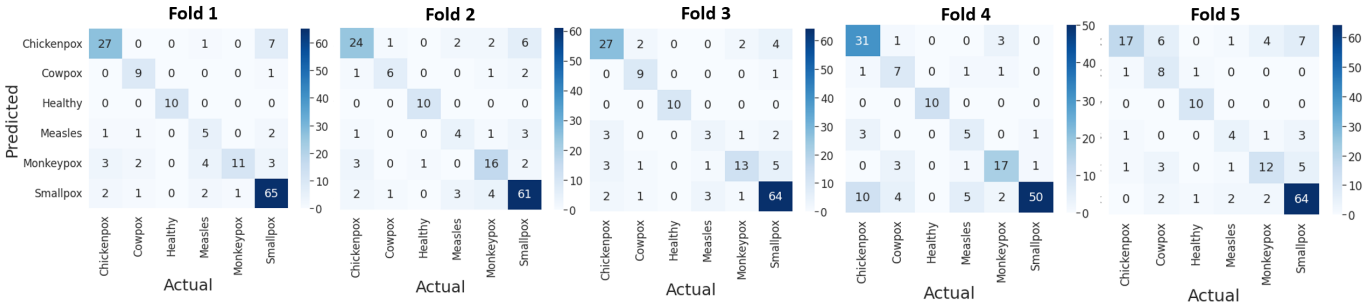

Fig. 9. Confusion matrices of prediction by MobileNet-V2 in 5-fold cross-validation.

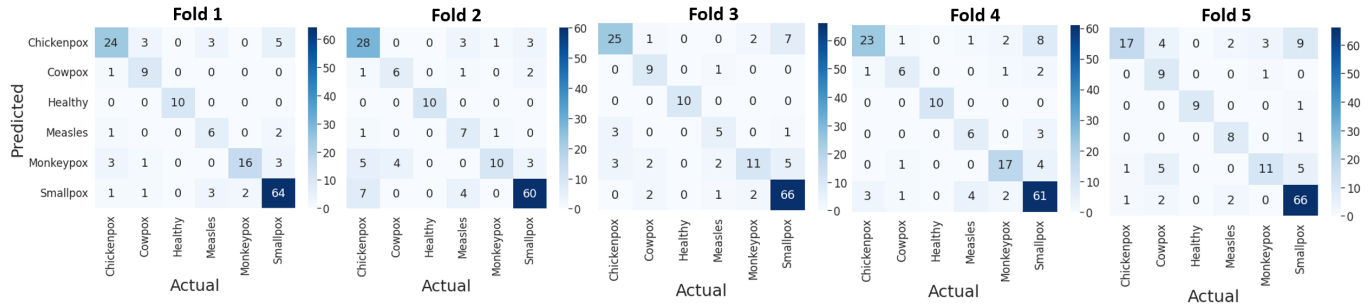

Fig. 10. Confusion matrices of prediction by ShuffleNet-V2-1× in 5-fold cross-validation.

and harness the full strength of state-of-the-art deep models, we must ensure a larger sample size for model training. We also observed that lighter deep models, having fewer trainable parameters, also have potential in pox classification, which could be used on smartphones for Monkeypox diagnosis in the latest outbreak. Furthermore, Monkeypox detection based on digital skin images can facilitate remote diagnosis by healthcare professionals, leading to early isolation of patients and effective containment of community spread.

V. Supplementary Material

We share a spread sheet that contains the list of sources of all images in our dataset. This spreadsheet can be found at this link.[3]

References

[1] J. P. Thornhill, S. Barkati, S. Walmsley, J. Rockstroh, A. Antinori, L. B. Harrison, R. Palich, A. Nori, I. Reeves, M. S. Habibi *et al.*, "Monkeypox virus infection in humans across 16 countries—april–june 2022," *New England Journal of Medicine*, 2022.

[2] S. Shchelkunov, A. Totmenin, P. Safronov, M. Mikheev, V. Gutorov, O. Ryazankina, N. Petrov, I. Babkin, E. Uvarova, L. Sandakhchiev *et al.*, "Analysis of the monkeypox virus genome," *Virology*, vol. 297, no. 2, pp. 172–194, 2002.

[3] E. M. Bunge, B. Hoet, L. Chen, F. Lienert, H. Weidenthaler, L. R. Baer, and R. Steffen, "The changing epidemiology of human monkeypox—a potential threat? a systematic review," *PLoS neglected tropical diseases*, vol. 16, no. 2, p. e0010141, 2022.

[4] J. G. Rizk, G. Lippi, B. M. Henry, D. N. Forthal, and Y. Rizk, "Prevention and treatment of monkeypox," *Drugs*, pp. 1–7, 2022.

[5] N. Sklenovska and M. Van Ranst, "Emergence of monkeypox as the most important orthopoxvirus infection in humans," *Frontiers in public health*, vol. 6, p. 241, 2018.

[6] N. Erez, H. Achdout, E. Milrot, Y. Schwartz, Y. Wiener-Well, N. Paran, B. Politi, H. Tamir, T. Israely, S. Weiss *et al.*, "Diagnosis of imported monkeypox, israel, 2018," *Emerging infectious diseases*, vol. 25, no. 5, p. 980, 2019.

[7] Q. Gong, C. Wang, X. Chuai, and S. Chiu, "Monkeypox virus: a re-emergent threat to humans," *Virologica Sinica*, 2022.

[8] M. A. Hussain, G. Hamarneh, and R. Garbi, "Cascaded regression neural nets for kidney localization and segmentation-free volume estimation," *IEEE Transactions on Medical Imaging*, vol. 40, no. 6, pp. 1555–1567, 2021.

[9] M. A. Hussain, A. Amir-Khalili, G. Hamarneh, and R. Abugharbieh,

[3] https://www.kaggle.com/datasets/arafathussain/monkeypox-skin-image-dataset-2022?select=image_list_with_sources_and_credits.xlsx

bioRxiv preprint doi: https://doi.org/10.1101/2022.08.08.503193; this version posted August 9, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

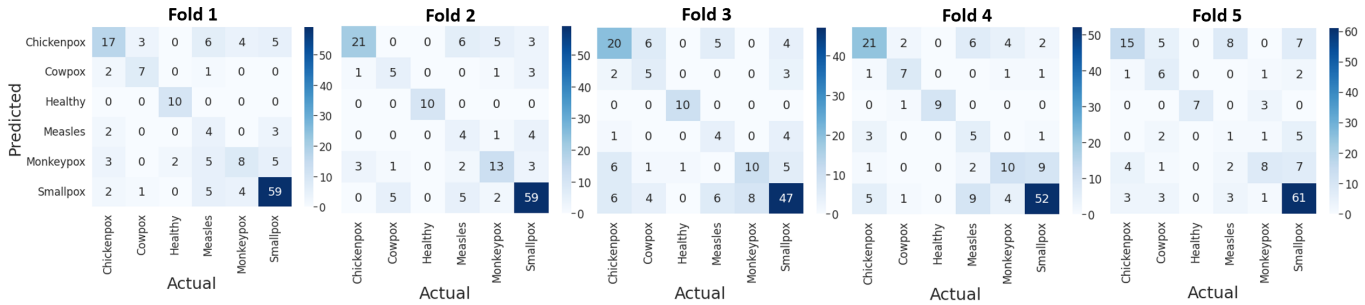

Fig. 11. Confusion matrices of prediction by SqueezeNet in 5-fold cross-validation.

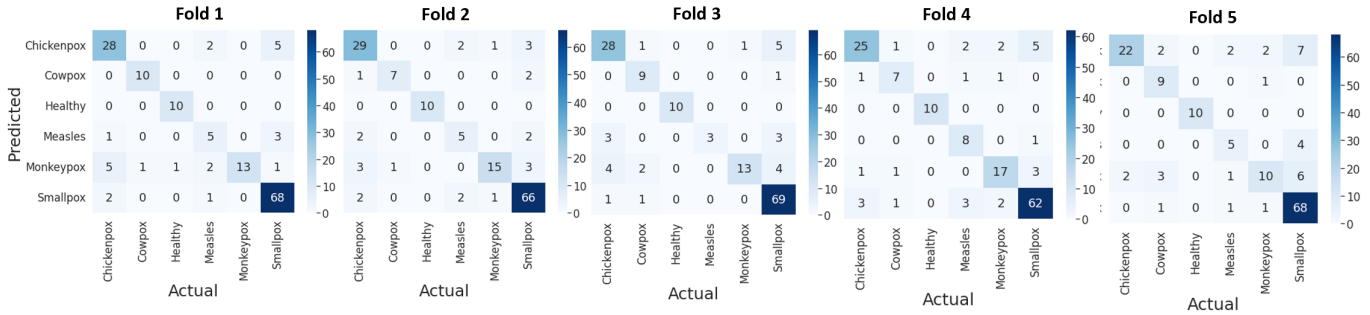

Fig. 12. Confusion matrices of prediction by majority voting in 5-fold cross-validation.

"Segmentation-free kidney localization and volume estimation using aggregated orthogonal decision cnns," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2017, pp. 612–620.

[10] M. A. Hussain, G. Hamarneh, T. W. O'Connell, M. F. Mohammed, and R. Abugharbieh, "Segmentation-free estimation of kidney volumes in ct with dual regression forests," in *International Workshop on Machine Learning in Medical Imaging*. Springer, 2016, pp. 156–163.

[11] M. A. Hussain, G. Hamarneh, and R. Garbi, "Noninvasive determination of gene mutations in clear cell renal cell carcinoma using multiple instance decisions aggregated cnn," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 657–665.

[12] M. A. Hussain, G. Hamarneh, and R. Garbi, "Learnable image histograms-based deep radiomics for renal cell carcinoma grading and staging," *Computerized Medical Imaging and Graphics*, vol. 90, p. 101924, 2021.

[13] M. A. Hussain, G. Hamarneh, and R. Garbi, "Imhistnet: Learnable image histogram based dnn with application to noninvasive determination of carcinoma grades in ct scans," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2019, pp. 130–138.

[14] M. A. Hussain, G. Hamarneh, and R. Garbi, "Renal cell carcinoma staging with learnable image histogram-based deep neural network," in *International Workshop on Machine Learning in Medical Imaging*. Springer, 2019, pp. 533–540.

[15] J. Sun, L. Peng, T. Li, D. Adila, Z. Zaiman, G. B. Melton-Meaux, N. E. Ingraham, E. Murray, D. Boley, S. Switzer *et al.*, "Performance of a chest radiograph ai diagnostic tool for covid-19: A prospective observational study," *Radiology: Artificial Intelligence*, vol. 4, no. 4, p. e210217, 2022.

[16] A. Akbarimajd, N. Hoertel, M. A. Hussain, A. A. Neshat, M. Marhamati, M. Bakhtoor, and M. Momeny, "Learning-to-augment incorporated noise-robust deep cnn for detection of covid-19 in noisy x-ray images," *Journal of Computational Science*, p. 101763, 2022.

[17] M. Momeny, A. A. Neshat, M. A. Hussain, S. Kia, M. Marhamati, A. Jahanbakhshi, and G. Hamarneh, "Learning-to-augment strategy using noisy and denoised data: Improving generalizability of deep cnn for the detection of covid-19 in x-ray images," *Computers in Biology and Medicine*, vol. 136, p. 104704, 2021.

[18] M. Dogucu and M. Çetinkaya-Rundel, "Web scraping in the statistics and data science curriculum: Challenges and opportunities," *Journal of Statistics and Data Science Education*, vol. 29, no. sup1, pp. S112–S122, 2021.

[19] T. Islam, M. A. Hussain, F. U. H. Chowdhury, and B. R. Islam, "A web-scraped skin image database of monkeypox, chickenpox, smallpox, cowpox, and measles," *bioRxiv*, 2022.

[20] M. M. Ahsan, M. R. Uddin, and S. A. Luna, "Monkeypox image data collection," *arXiv preprint arXiv:2206.01774*, 2022.

[21] S. N. Ali, M. Ahmed, J. Paul, T. Jahan, S. Sani, N. Noor, T. Hasan *et al.*, "Monkeypox skin lesion detection using deep learning models: A feasibility study," *arXiv preprint arXiv:2207.03342*, 2022.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[23] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[24] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[25] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size," *arXiv preprint arXiv:1602.07360*, 2016.

[26] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q. V. Le, "Mnasnet: Platform-aware neural architecture search for mobile," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2820–2828.

[27] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.

[28] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "Shufflenet v2: Practical guidelines for efficient cnn architecture design," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 116–131.