

Hand Foot and Mouth Rash Detection Using Deep Convolution Neural Network

Naqibullah Vakili ¹, Jonathan H. Chan ¹, Worarat Krathu ¹, Nipat Phattarakijtham ¹, and Kazuya Hirata ¹

¹Affiliation not available

October 30, 2023

Abstract

This work was presented at the 9th Joint Symposium on Computational Intelligence (JSCI9), organized by the IEEE-CIS Thailand Chapter, that aims to support research students and young researchers, to create a place enabling participants to share and discuss on their research prior to publishing their works. The event was open to all researchers who want to broaden their knowledge in the field of computational intelligence.

Hand Foot and Mouth Rash Detection Using Deep Convolution Neural Network

Naqibullah Vakili
School of Information Technology
King Mongkut's University of Technology
Thonburi
Bangkok, Thailand
naqibullah.vakili@mail.kmutt.ac.th

Jonathan H. Chan
School of Information Technology
King Mongkut's University of Technology
Thonburi
Bangkok, Thailand
jonathan@sit.kmutt.ac.th

Worarat Krathu
School of Information Technology
King Mongkut's University of Technology
Thonburi
Bangkok, Thailand
worarat.kra@sit.kmutt.ac.th

Nipat Phattarakijtham
School of Information Technology
King Mongkut's University of Technology
Thonburi
Bangkok, Thailand
nipat.phattarakijtham@mail.kmutt.ac.th

Kazuya Hirata
Electronic Information Systems
Shibaura Institute of Technology
Tokyo, Japan
bp18095@sic.shibaura-it.ac.jp

Abstract— Hand-Foot-Mouth Disease (HFMD) is an extremely contagious viral infection common in infants that may quickly develop into a severe problem. The disease spreads easily through person-to-person contact with an infected person. To assist in diagnosing HFMD, we built a dataset of over 2000 images of various skin conditions. A state-of-the-art deep convolution neural network was applied to effectively distinguish Hand Foot and Mouth Disease rash from other skin conditions. Experimental results show 0.954 accuracy on our curated dataset. Our proposed model can potentially facilitate a proper and early detection of HFMD rash to assist in containing HFMD outbreaks in the Asia Pacific region.

Keywords—HFMD, HFMD rash, Asia Pacific Region, classification, Deep learning, Convolution Neural Network, CNN.

I. INTRODUCTION

HFMD is a common contagious disease that immediately develops into a severe problem, mainly affect infants, preschool children, and adults, too, sometimes [1]. Children younger than five years old are at greater risk of disease infection because they are less likely to have the necessary antibodies to protect them against viruses causing HFMD infection [2]. The epidemic of HFMD is mainly caused by human Enterovirus 71. The number of reported cases has recently been increasing as it has reached Over a million cases in Asia. Due to its threat to infants and its immense burdens on nations, mainly in the Asia Pacific region, including Thailand, Singapore, Malaysia, Japan, Australia, and across China, it has become a serious public health concern [3,4].

The primary defining symptoms of HFMD are the skin rash on the hands, feet, mouth, buttocks, ulcer in the oral mucosa, and fever that it causes, as the other signs closely mimic other illnesses. The disease's distinctive pattern and the method in which the rash progresses across the body, usually hand, foot, and mouth, are critical signs that healthcare providers visually diagnose the disease. Without immediate medical attention, the condition may be fatal for infants in the

short term. Consequently, With the virus continuing to spread globally, proper and efficient diagnosis of HFMD will be crucial in mitigating the virus rate. Moreover, detecting HFMD at the initial stage of the illness will increase the chance for treatment and mitigate fatal consequences.

Several works using data mining and machine learning have been accomplished in automatic HFMD disease diagnostics, devoted to identifying the Hand-Foot-Mouth virus from clinical data [2, 4]. To the best of our knowledge, the only published paper in detecting HFMD from digital images has been done by Aryan et al. [5]. The authors proposed a solution to classify HFMD rash from normal skin, which means skin without any inflammation, and rash using image processing and image recognition techniques. They tackled the challenge by applying handcrafted features and a classifier on top—the proposed algorithm trained and tested on 72 images collected from different sources.

The aim of this study is to use deep convolution neural networks (CNNs) to assist the medical practitioner and patient alike in the detection and possible diagnosis of the disease. Currently, no research specifically has been done for the detecting of HFMD rash from other skin conditions. Also, a system capable of HFMD virus characterization can be implemented in many fields. It could also be deployed as an application to be utilized in the schools, airport security, telemedicine, etc., to prevent transmission of the disease, or potentially as a phone application in developing nations, where healthcare assistants are rare, to facilitate the recognition of HFMD.

The rest of the paper is organized as follows. In the first section, our dataset and CNN model are explained. Next, present our experimental results. Followed by results in Section V. lastly provides the discussion and our conclusion.

TABLE I. SUMMARY TRANSFER LEARNING TECHNIQUES USED FOR THE CLASSIFICATION OF SKIN CONDITION

Work	Skin condition	Data-set	DNN	Performance
[7]	Seven common skin lesions.	HAM10000	DenseNet-169	Acc: 83.59%
[11]	Erythema migrans rash	1834 images	ResNet50	Acc: 95%
[12]	Nine most common skin conditions.	7259 images	ResNet-152	Acc: 85%.
[13]	Nine most common skin conditions.	41643 curated images	ResNet152	Acc: 80%
[14]	Melanoma Vs other	ISBI-16 ISBI-17 HAM10000	ResNet50 and ResNet101	Sensitivity:89.4% Specificity:64.4% ROC: 0.769
[15]	Melanocytic nevus, SK BCC, psoriasis.	1589 images	Inception v3	Dataset A: Accuracy: 87.25 $\pm$ 2.24% 522 Dataset B: Accuracy:86.63% 0.917 precision 0.925 recall
[16]	Eleven skin conditions	587 images	ResNet34	
[17]	Melanoma and other skin lesions	MED-NODE, Derm (IS & Quest) and ISIC	PNASNet-5	Acc: 0.76

II. RELATED WORK

This section emphasizes using various skin rash detection methods and diagnosis, focusing on the recent research that has used transfer learning for the same purpose. Deep neural networks achieve the state of the art performances in solving many computer vision problems. CNN's have recently gained enormous attention and have been increasingly implemented in medical study and radiology[6-8]. For instance, NasrEsfahani et al. developed a convolution neural network classifier to recognize melanoma lesions with a predictive accuracy of 81%. Recently, to discuss the most common skin diseases seen in health care, a CNN-based model has been developed by Google with accuracy on par with U.S board-certified dermatologist [8-9]. Likewise, CNN's have also been applied for other diseases as well, such as diabetic retinopathy, lymph node metastasis, and breast cancer [10-12]. CNN needs adequate data and computational resources for model training. If only a small image dataset is possible, superior strategies, e.g., data augmentation or transfer learning, are regularly applied to improve classification performance. Transfer learning uses a pre-trained classifier to work on a different but related problem [10]. Compared to the approach of creating a deep CNN classifier from scratch, it needs less training images and reduced runtime as a result.

Recently, comparative research presented on a small dataset, transfer learning is more helpful than other methods reviewed [7]. Another study that used transfer learning to implement a CNN classifier, by Kermany et al., proved performance similar to that of human experts in distinguishing age-related macular degeneration and diabetic macular edema [8]. Moreover, several CNN-based methods were developed by the crucial need for classifying skin images for timely treatment and diagnosis of diseases. TABLE I presents 8 CNN-based techniques developed lately for skin conditions classification. They all used transfer-learning to achieve final models—most of the studies in TABLE I employed residual networks (ResNet) to construct CNNs. Therefore, with the

lack of tools for visual identification of HFMD and increasing cases of HFMD in the Asia Pacific region countries, which was associated with contact-related transmissions with subsequent spread through populations. So, the models that can accurately recognize HFMD rash are urgently required. Improved HFMD rash identification ability would help primary care providers efficiently address new and various challenges in Asia Pacific region posed by the potential reemergence of HFMD into populations. The literature review found that such an expert method is needed to distinguish HFMD rash from other skin conditions.

III. MATERIAL AND ARCHITECTURES

A. Data Collection

To analyze skin rash and accurately classify the HFMD rash from other skin conditions and normal skin, we implement CNN models in a binary classification context so that all the images are labeled as either “HFMD” or “Non-HFMD.” For the non-HFMD section, it is mainly necessary to have images, including classes similar to the HFMD disease in appearance in the same way classes that significantly differ to provide overall diversity to the dataset.

Since there are no public resources to obtain qualitative datasets from real patients due to its high cost and sensitive information, we decided to collect the images from different sources. We used the Google Images service to parse images from the web, and the following queries were performed: “HFMD rash,” “HFMD inflammation.” TABLE II presents the entire list of samples we collected. There are 1782 images of the Non-HFMD rash and 315 HFMD images presents in the dataset. It is hard to perform a high-quality method by training it from scratch due to the shortage of dataset. The quality of such models is almost invisible from a coin toss. To mitigate the shortage of proper HFMD images dataset for training and testing. Firstly, we applied image augmentation techniques such as minor adjustments, horizontal flips, and random rotations to create HFMD duplicates images. The evaluation showed that our final classifier, adapted from a

pre-trained model, presented better performance without augmentation.

TABLE II: DISTRIBUTION OF IMAGE SAMPLES

Disease	Number of images
Measles	112
Scabies	211
Eczema	238
Psoriasis	115
Erythema	103
Impetigo	86
Urticaria	77
Atopic Dermatitis	109
Lupus	68
Ringworm	125
Bowen	109
Keratosi	96
Dermatofibroma	65
HFMD	315
Normal	268
Total	2097

B. Proposed Method

The method presents all the detailed stages to distinguish HFMD from other skin conditions using deep convolutional neural network.

Since we had no control over the images and camera resectioning acquisition, pre-processing is performed as the images are collected from different sources and had varying dimensions. We consider the input size $224 \times 224 \times 3$ based on ResNet50 requirements.

C. CNN Model

In this study, we used transfer learning approach to develop our CNN model to identify HFMD rash through skin images. We experimented with three well-known CNN architecture of (ResNet), ResNet34, and ResNet50 and inception v3 with transfer learning scenario initialized on ImageNet. Inception v3 is regularly applied for image detection that has been shown an accuracy rate of over 78.1% on the ImageNet dataset. The classifier is the culmination of many ideas developed by multiple researchers over the years. The Inception v3 consists of a 42-layer deep neural network. The inception v3 model consists of symmetric and asymmetric structure blocks, including convolutions layer (Conv), max-pooling, average pooling, dropouts, and fully-connected layers [19].

ResNet-50 and some other deep ResNet classifiers were first proposed in 2015 by Microsoft Asia and demonstrated outstanding performances in many fields. ResNet models won ImageNet ILSVRC 2015 competitions (Large Scale Visual Recognition Challenge) in image detection, localization, and classification. Also, ResNet show state of art

on the tasks of ImageNet localization, detection, etc. ResNet stands for residual networks. ResNet34 and ResNet50 have already been pre-trained on the ImageNet dataset containing more than 14 million images [9]. Mainly, the ResNet model consists of 5 steps, where each step consists of Convolutional (Conv) layer, Rectified-Linear-Unit (ReLU), and Normalization (Batch Norm) layers. Step 1 consists of the convolution layer and max pooling. A convolutional layer applies a filter named kernel to pass over an image to build a feature map. Initially, kernel size is 7×7 with 64 output channels and a stride of 2. Max pooling reducing image dimensionality by retaining each maximum value in each patch of each feature map, which the kernel carries over. Steps 2-5 are residual blocks. Residual blocks especially look like main road networks without gates in their skip connections to permit data flow from the first layers to the final layers. Following step 5, there is an average-pooling and a fully-connected layer—average pooling reducing image dimensionality by utilizing each matrix square's average value. The fully connected layer takes the result of the convolution and pooling and outputs the image classification's final probabilities. [9, 18].

Since the ResNet50 is already tuned for identifying features of different images, we only require to change the last section of ResNet50 to distinguish HFMD rashes. The version of ResNet50 used in this work is the latest implementation 2020 in a Python package fastai. In the fastai library, ResNet50 is layered on top of the Pytorch library, a Python environment in fastai [9].

IV. RESULTS

Applying various fine-tuning methods determined that the best approach is to freeze all Convolution layers (Conv) of the ResNet, i.e., the classifier's backbone, with their weights pre-trained on ImageNet dataset and trained only the last section of the classifier. The dataset was divided randomly into training and validation sets for each iteration of training with an 80 20 % split, sequentially. Normalization was applied to model training. The batch size set to 64 and the number of epochs 5 for the initial phase of model training.

Images in the dataset were divided into one of the two classes: HFMD (positive) and non-HFMD (negative). On each round, we measured three commonly standard metrics to assess our model: accuracy, specificity, and sensitivity, which are represented as follows: Sensitivity= $TP / (TP + FN)$, Specificity = $TN / (TN + FP)$, Accuracy = $(TP + TN) / (TP + FN + TN + FP)$. Like TP, TN, FN, and FP indicate the number of true positive, true negative, false negative, and false positive classifications, respectively. Moreover, Sensitivity: indicate the ability of a test to distinguish patients with a disease correctly. Specificity: the ability of a test to distinguish people without the disease correctly. Accuracy: Number of correct predictions/ Total number of predictions [9].

A. Model Refinement

After building the initial classifier, we fine-tuned the entire model by unfreezing the model's backbone layers for retraining. As learning significantly affects the algorithm performance [9]. We visualized the relation between learning-rate and loss-function on each iteration for 5 epochs to find suitable learning rates for the validation iterations.

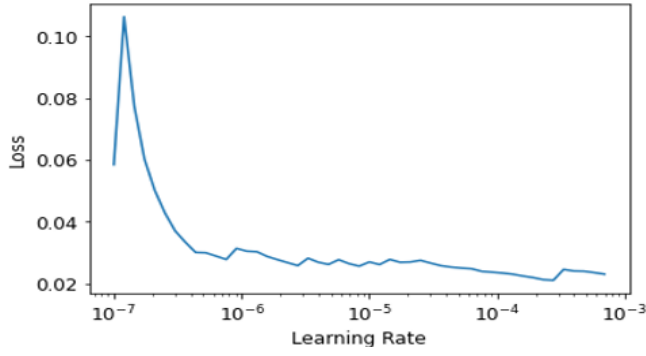


Fig.1. Presents learning rate vs values of the loss function.

It presents in Figure 1 that the recorded loss is directed to decrease with the increase of learning rate before it changes immediately after a point near to $1e-3$. Based on this view and three other similar plots, we set the range $[1e-6, 1e-4]$ as our differential learning rates for classifier refinement, with which ten more epochs were then applied to obtain our final model. Moreover, it's helpful to see where exactly our errors are occurring. To know whether they're due to a dataset issue (e.g., images that aren't HFMD at all, or are mislabeled, etc.), or a classifier problem (maybe it isn't handling images taken with unusual lighting, or from a various angle, etc.). We again sort our dataset based on reliable dermatologist websites curation. Therefore, our final model's performance on the validation sets shows that model refinement improved accuracy, sensitivity, and specificity from the original 89.6%, 0.619, and 91.7. TABLE III presents the validation performance measurement of the HFMD and Non-HFMD rash for each model. ResNet-50 presented the highest accuracy on our original image dataset. Also, we build a test dataset of about 245 images of HFMD and Non-HFMD. The images were used to evaluate the quality of classification from our fine-tuned model. A simple forward pass of image on a network initialized with the saved parameters was performed. Whereas testing performance measurement Sensitivity, Specificity, and Accuracy are presented in TABLE IV.

TABLE III. VALIDATION PERFORMANCE MEASUREMENT.

Model	Accuracy	Sensitivity	Specificity
Inception v3	0.941	0.809	0.96
ResNet-34	0.94	0.796	0.961
ResNet-50	0.954	0.833	0.974

TABLE IV. TESTING PERFORMANCE MEASUREMENT FOR RESNET-50 MODEL

Images classified	Accuracy	Sensitivity	Specificity
	0.951	0.813	0.971

We manually analyzed some misclassified images to verify our results. Figure 2 presents an example of a False-positive prediction. The right photos in Figure 2 are a mouth measles rash compared to HFMD rash images on the left. This example presents the difficulty of identifying the HFMD rash from other skin diseases. As most of the false-positive identification was made on images so comparable in

appearance to the HFMD rash, like measles, scabies, etc. To decrease such a problem, more training images of both the HFMD infection and these similar subgroups are needed.

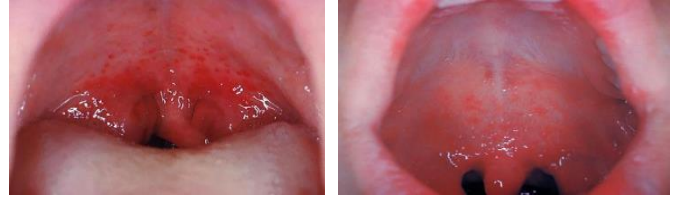


Fig. 2. Sample image of false-positive prediction (right), measles mouth classified as HFMD disease.

V. CONCLUSIONS

This study proposed a novel automatic deep learning-based classifiers (CNN) to distinguish the distinctive HFMD rash from various other skin conditions. With the spread of the HFMD disease continuing to increase in Asia Pacific countries. An adequately trained model capable of recognizing HFMD rash is vital in combatting the outbreak. Our research showed that deep learning solutions could perform on par and even better method on the task than classifiers to pick the color, texture manually, and shape features [5], even provided a minimal number of labeled images. Moreover, an advantage of our proposed method is that it does not need manual pre-processing of images. We have tested three pre-trained CNN models, ResNet (34, 50) and Inception v3, on 2097 clinical images. In the experiment stage, we analyzed the performance of ResNet 50, 34, and Inception v3, which ResNet50 shows better performances than others. Our proposed model achieved the classification average accuracy of 0.954 and an average sensitivity of 0.833 on validation datasets. Also, we can improve performance with the addition of more images. Moreover, Our Classifier only focuses on the rash's appearance and structure and does not consider the distribution of the rash or part of the body.

VI. FUTURE WORK

Further research includes information about different simultaneous symptoms, like, fever, poor appetite, feeling unwell and painful sores due to the lack of these data. When available in the future, such information can be easily combined into our model to further develop diagnosis efficiency. We have various ages, gender, and body parts across images but do not have a wide variety of skin colors. Another possible future development of this research is collecting more ethnically diverse images to apply for model retraining. By combining more images from various racial groups, our prototype model can be further improved. More importantly, it can be more readily implemented in telemedicine to assist health professionals and primary care providers in diagnosis.

REFERENCES

- [1] S. Aswathyraj, et al. (July 2016). "Hand foot and mouth disease (HFMD): emerging epidemiology and the need for a vaccine strategy." *Medical Microbiology and Immunology*, 205(5), 397-407.
- [2] J.F. Wang, et al. (Apr 2011). "Hand, foot and mouth disease: spatiotemporal transmission and climate." *International journal of health geographics*, 10(1),25.
- [3] J. Zhao, et al. (2019, Apr). "The transmission of Hand, Foot, and Mouth Disease in East and Southeast Asia." *BioRxiv*, 612580.

- [4] G. Liu, et al. (Nov 2017). "Developing a machine learning system for identification of severe hand, foot, and mouth disease from electronic medical record data," *Nature Scientific Reports*, vol. 7, no. 1, p. 16341.
- [5] F. Aryan, et al. (Jun 2020). "Image Recognition for Detecting Hand Foot and Mouth Disease." *IAIT2020: The 11th International Conference on Advances in Information Technology*. No.: 35 P. 1–11.
- [6] E. Nasr-Esfahani, et al. (Aug 2016). "Melanoma-detection-by analysis-of-clinical-images-using-convolutional-neural-network." 2016-38th-Annual International-Conference-of-the-IEEE-Engineering-in-Medicine-and Biology-Society (EMBC) . (pp. 1373-1376). IEEE.
- [7] S.S. Yadav, et al. (Dec 2019). "Deep convolutional neural network-based medical image classification for disease diagnosis." *J Big Data* 6, 113 (2019). available at: <https://doi.org/10.1186/s40537-019-0276-2>.
- [8] D. S.Kermany, et al. (Feb 2018). "Identifying medical diagnoses and treatable diseases by image-based deep learning." 2018;172(5):1122–31.
- [9] K. Glock, et al. (Aug 2020). "Measles Rash Image Detection Using Deep Convolutional Neural Network." [eess.IV]. available at: <https://arxiv.org/ftp/arxiv/papers/2005/2005.09112.pdf>
- [10] A. Mobiny, et al. (Aug 2019). "Risk Aware Machine Learning Classifier for Skin Lesion Diagnosis." *J Clin Med*. 8(8) 1241.
- [11] P.M. Burlina, et al. (Nov 2019) "Automated detection of erythema migrans and other confounding skin lesions via deep learning." *Comput Biol Med.*, 105:151-156.
- [12] S. Mishra, et al. (Mar 2020). "Assessing Robustness of Deep Learning Methods in Dermatological Workflow," *ACM CHIL 2020 Workshop*, arXiv: 2001.05878.
- [13] S. Mishara, et al. (Oct 2018). "Supervised classification of Dermatological diseases by Deep learning," no. DI, 2018.
- [14] M.A. Khan, et al. (Apr 2019), "Multimodel deep neural network-based features extraction and optimal selection approach for skin lesion classification," In 2019 international conference on computer and information sciences (ICCIS) 2019 Apr 3 (pp. 1-7) IEEE.
- [15] X. Zhang et al. (Jul 2018). "Towards improving diagnosis of skin diseases by combining deep neural network and human knowledge," *BMC medical informatics and decision making*, 2018;18(2):59.
- [16] S. Sharma, et al. (2019 Aug). "Derma Detect: a computer vision and deep learning approach for an accurate diagnosis of skin conditions and rashes" Available at <https://www.researchgate.net/publication/335083461>.
- [17] M.D. Milton, et al. (Jan 2019). "Automated Skin Lesion Classification Using Ensemble of Deep Neural Networks in ISIC2018: Skin Lesion Analysis Towards Melanoma Detection Challenge". arXiv:1901.10802. available at <https://arxiv.org/ftp/arxiv/papers/1901/1901.10802.pdf>.
- [18] N. N. Sultana, et al. (Apr 2018). "Deep residual network with regularized fisher framework for detection of melanoma." *IET Comput Vis*.2018;12(8):1096–104.
- [19] A. Demir, et al. (Oct 2019). "Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3". 2019 Medical Technologies Congress (TIPTEKNO), DOI: 10.1109/TIPTEKNO47231.2019.8972045