A WeChat Application Combined with Convolutional Neural Network for Skin Cancer Recognition

Junlin Wu[1*]

[1]Sixth College, University of California San Diego, 9500 Gilman Dr, La Jolla, United States of America
*Corresponding author: juw029@ucsd.edu

Abstract
The skin cancer detection process is complex and time-consuming. The rise of machine learning led to the invention of skin detection models. However, the models are rarely used in the real world and the few implementations are slow and not lightweight. This paper sought to solve the issue by developing a lightweight mobile application that runs diagnosis on the cloud, allowing users to assess their conditions preliminarily. Convolutional Neural Network as a powerful type of deep learning methods was considered to recognize the skin cancer in this study. Datasets collected from a large skin cancer dataset called HAM10000 was employed to verify the proposed method. Based on convolutional neural network, an application using WeChat platform was able run a skin-cancer detection model with 0.81 test accuracy in the cloud to diagnose a user's image in under 20 seconds while maintaining the app's size under two megabytes, achieving ease of access, lightweight, and speed.

Keywords: *WeChat application, Convolutional Neural Network, skin cancer recognition, HAM10000*

1 INTRODUCTION

Skin cancer is one of the most common diseases in the world. It can be roughly divided into two groups: nonmelanoma and melanoma. The melanoma skin cancer is the more serious type. Out of all deaths related to skin cancer, 75% of them are caused by melanoma skin cancer. The number of incidences is increasing globally. In the U.S., statistics showed that nonmelanoma skin cancers have constituted one third of all cancers [1]. For melanoma skin cancer, study has shown that the rate has been steadily increasing by about 3%to 7% annually. The rates have become so high that about 1% of the U.S. population born in 1993 will develop malignant melanoma skin cancer [1].

To avoid skin cancer related death, it is critical to detect skin cancer early to assure patients receiving treatment as early as possible. In the past, diagnosis has typically constituted of the following steps. First, the doctors' queries about patients' personal information such as family medical history, gender, etc. to identify high risk individuals. Then, the examiner closely inspects every part of the skin of the patient through the naked eye or microscopy and identify potential skin cancer by appearance. Asymmetry, Border, Color, Diameter (ABCD) checklist is often used to assist detection [4]. For suspected areas, a skin biopsy is needed to further evaluate the condition of the skin.

These steps are often time consuming and detection accuracy could vary based on the doctor's experience. Thus, a faster approach to detect skin cancer is needed. Studies has been conducted to incorporate computers into skin cancer detection by having computers analyze photos of suspected spots. M and Karki proposed a solution that involves feature extraction using Asymmetry, Border, Color, Diameter (ABCD) rule, Gray Level Co-occurrence matrix (GLCM) and Histogram of Oriented Gradients (HOG), and machine learning models of Support Vector Machine, naïve bayes, and K-Nearest-Neighbor. The solution was able to achieve an accuracy of 97.8% after training [9]. Building on their works, others have tried predicting skin cancer using convolutional neural network. A group of scholars compared the performance of InceptionV3, ResNet, and VGG19 for the task of skin cancer detection and concluded the InceptionV3 to be the best model with accuracy of 86.90%, precision of 87.47%, sensitivity of 86.14%, and the specificity of 87.66% [6]. Zhang N, et al developed an algorithm that optimally selects weights and biases. Their proposed model achieved the highest specificity, accuracy, sensitivity, NPV, and PPV when comparing to models such as VGG-16, ResNet-101, AlexNet, InceptinoV3, etc [10]. Gouda and J took it

© The Author(s) 2023
B. Fox et al. (Eds.): IC-ICAIE 2022, AHCS 9, pp. 975-980, 2023.
https://doi.org/10.2991/978-94-6463-040-4_147

further by having their model directly classify a skin spot image into types of skin cancer using ResNet. Their model achieved an accuracy of 92% and recall of 0.8331 [3].

However, the application of those models has rarely been discussed. Models only become useful when people are using them in the real world. Karargyris et al. proposed a mobile app that processed the images users upload through colour transformation, binarization, etc. and used SVM to classify the image into benign or malignant [5]. This should run quite efficiently on modern phones, but result could be improved by using neural networks. Dai et al took it further by running a pre-trained Convolutional Neural Network (CNN) model on a mobile app and classify the user image into its predicted skin cancer class. However, the process is entirely run on a mobile phone and the authors failed to mention the running time of their application. The process could be very slow as the phones' computing power is extremely limited. In addition, updating the model is difficult as users need to redownload and install a new version of the app.

To solve all of the above problems, a WeChat mini-program app and a web-framework were developed to process users' uploaded images and run it through a pre-trained CNN model on the server. In this way, users have the ease of accessing the app on mobile as well as the speed of getting the result.

2 METHOD

2.1 Dataset

The HAM10000 dataset is used for training and testing the model. It is a dataset collected by Harvard containing 10, 015 dermatoscopic images [8]. The images are labeled with akiec for Actinic keratoses and intraepithelial carcinoma / Bowen's disease, bcc for basal cell carcinoma, bkl for benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses, df for dermatofibroma, mel for melanoma, melanocytic nevi for nv, and vasc for vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage. In addition, 340 images of healthy skin were added and labeled as nod. Figure 1 provides some examples for the collected dataset.

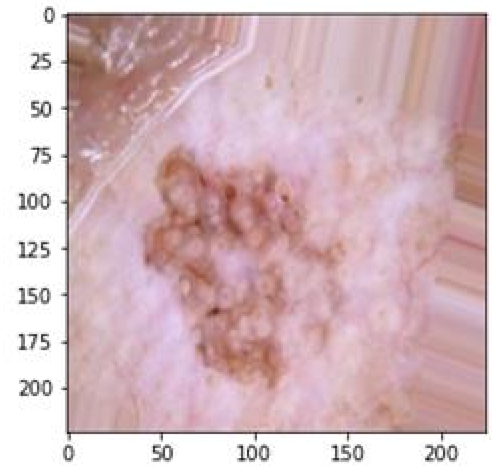

Figure 1: The sample image for the HAM10000 dataset.

2.2 Data preparation

The images are greyscaled and normalized using image data generator in Keras by setting rescale values to 1/255. Also using Keras, each image is adjusted to have the size of 224 by 224.

2.3 Model

Convolutional Neural Networks is a type of neural networks used for machine learning, which has been widely in various interesting domains e.g. medical analysis [2], rehabilitation training [7] and automatic driving [11]. It consists of at least one convolutional layer. A convolutional layer contains one or more kernels, which are matrices. Each kernel scans the input data from top left using a sliding window that selects a submatrix with dimension equal to the kernel's dimension and calculate the dot product as the output. The sliding window then slides to the right or bottom and repeat the process. The process has a great ability to reduce the dimensionality of data. For this reason, it is often used for image classification where data typically has large dimensions.

The model is built on top of Google's InceptionV3 architecture. It is a convolutional neural network architecture that achieved great result in ImageNet Large Scale Visual Recognition Challenge 2014. Since the ImageNet and skin detection are both image classification problems, InceptionV3 model can be selected for transfer learning to increase model training time. The appropriate input was set for the InceptionV3 model to fit the HAM10000 datasets. Then a global average pooling layer was added to massively reduce the amount of computations. Then a drop out layer is used to prevent overfitting. This is followed by a dense layer with Relu activation function, another dropout layer, and a final output layer using softmax as the activation function. Softmax is chosen for its ease of computing the gradient. Figure 2 and Table 1 represent the employed architecture of InceptionV3.

Table 1: The architecture of the employed incepctionV3.

Layer (type)	Output Shape	Param #
inception_v3 (Functional)	(None, 5, 5, 2048)	21802784
global_average_pooling2d (GlobalAveragePooling2D)	(None, 2048)	0
dropout (Dropout)	(None, 2048)	0
dense (Dense)	(None, 512)	1049088
dropout_1 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 8)	4104

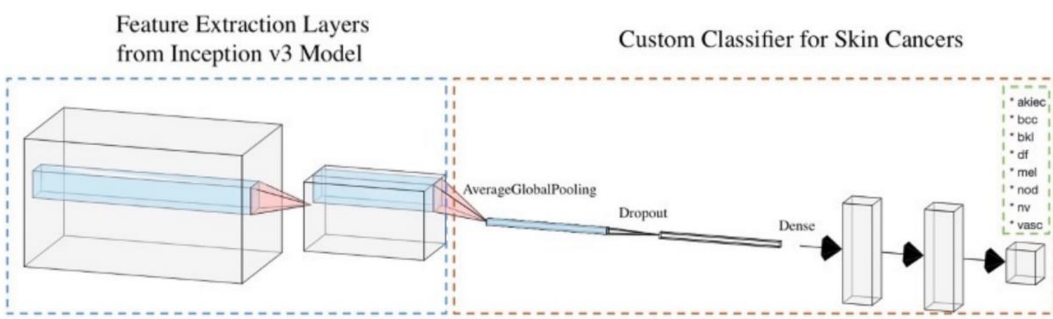

Figure 2: The workflow for skin cancer detection based on Inception V3.

2.4 Implementation Details

For model training parameters, batch size is chosen at 32. The model uses categorical cross entropy as the loss function and uses accuracy for the evaluation metrics. Adam optimizer is used with base learning rate set to 0.0001.

2.5 Development for Application

To apply the model into real-world application, a WeChat mini-program was developed to allow user interaction. The reason to use the WeChat platform is because it has already had a large user base so millions of users can have easy access to the app. The huge user base also proved the reliability of the platform as it needs to handle constant server stress from user activities. WeChat also constricts the mini-program to be under 2M so downloading and updating the app will be a nonissue since the size is so small. WeChat mini program has a structure similar to Vue, where the data are set in the logical page and call the variable in the layout page. A page for WeChat Mini-Program has four components– wxml, wxss, javascript, and json. The first two corresponds to html and css. Javascript is responsible for logical components and the json file is for WeChat's built-in function and import external components.

To handle the actual diagnosing part of the program, a web framework was built using Python's flask package. The mini-program contains three pages: welcome page, main page, and result page. The user can upload an image in the main page. After receiving an image, the program will upload the image to WeChat's cloud storage and get a fileId. The fileId will be stored in WeChat's cloud database and return a "_id" for identification. The reason to use WeChat's cloud service as opposed to sending the image directly to the server is that the server may run into storage issues and had to be reconfigured if user counts become too great. WeChat cloud service allows easy storage expansion so that the program can be scaled easily. Then, the mini-program will send requests to "/detect" route of the server with the image's _id. After getting an access token from WeChat's API, the server will query the cloud database to find fileId and used it to find the fileId of the image stored in cloud storage, which will then be fed into WeChat's API to get an HTTPS address of the image. The image is then downloaded and processed by running it through the pre-trained model and return a result.

To increase usability, an NLP chat-bot was developed with datasets containing common questions regarding skin cancer. The final detection result is showed in a chat-box manner. The user can type to send a message to the chat-bot. The program will send an HTTPS post request with the user-input message as the data to the server's "/chat" route. After feeding the message into the chat-bot model, the result is sent back to the mini-program. The page's layout consists of a for loop that dynamically links to a list variable in JavaScript. User and chat-bot's messages are added to that variable to allow the page to update without needing to refresh. Figure 3 indicates the framework for developed application. The numbers indicate the sequence of the app's logic.

3 RESULTS AND DISCUSSION

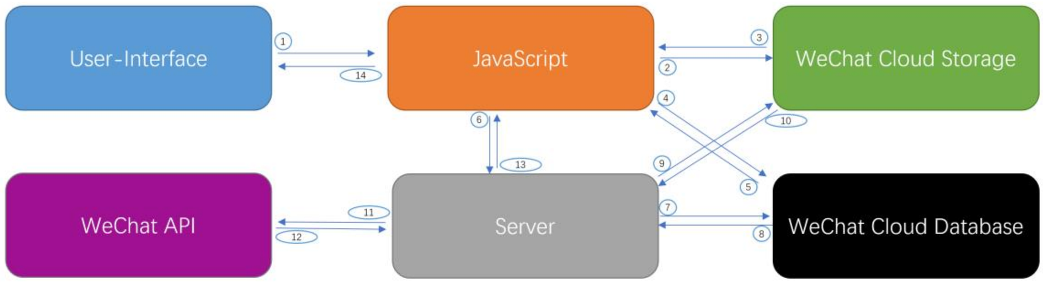

Figure 3: The framework for developed application.

3.1 Performance for the Inception V3

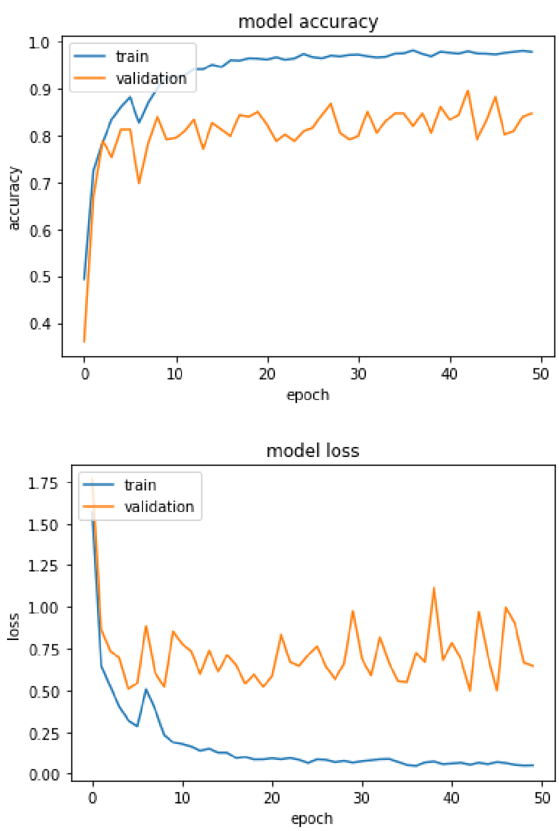

Figure 4: Training curve.

Table 2: The performance of Inception V3.

Epochs	Validation Accuracy	Validati on Loss	Test Accura cy	Test Loss
50	0.7891	1.0219	0.8090	0.0783

Figure 4 and Table 2 indicate the training curve and detailed performance. Loss in Table 2 is mean categorical crossentropy.

After 50 epochs, the validation accuracy and loss are more volatile. However, the overall trend is still a flattening out process with validation accuracy and loss achieving 0.7891 and 1.0219 after 50 epochs. The test accuracy yields 0.81 and test loss of 0.0783.

The results are quite impressive. After only 50 epochs, the model can have such an impressive result on the test set. An accuracy of 0.81 is much better than randomly guessing the result, suggesting that the model has truly picked up some of the key essential features from the datasets and found some crucial patterns.

3.2 Application

After saving the model, a WeChat mini-program was developed along with a web framework using Flask. The users are able to upload a photo of their skin to the app and the server will receive the image, process it, and return the result in under 20 seconds using server-level hardware. The user first encounters the welcome page which will direct them to the next page where they can upload a photo. After receiving the diagnosis from the server, the result will be displayed in a chat-box, resembling the interface for texting. In addition, an NLP model was trained to implement a chat-bot function. The user can type in some common questions regarding skin cancer like “what are the preventive measures of dermatofibroma?”. The message will be sent to the server comes up with replies that provide information regarding the disease. This application will provide patients with the ability to have a preliminary assessment of their skin condition and educate them about their symptoms. In addition, since the installation size of the app is under 2M, it is very easy for the developer to push updates to users. Moreover, WeChat being a giant platform allows ease of access for millions of potential users and its cloud functionality allows scalability down the road if the demand rises. Finally, the actual photo processing taking place on a self-built server provided data security for the users as their diagnosis will not go through WeChat’s sever thus decreasing the chance of the platform leaking the user’s medical condition. Figure 5, Figure 6, and Figure 7 show the specific interfaces for the developed application.

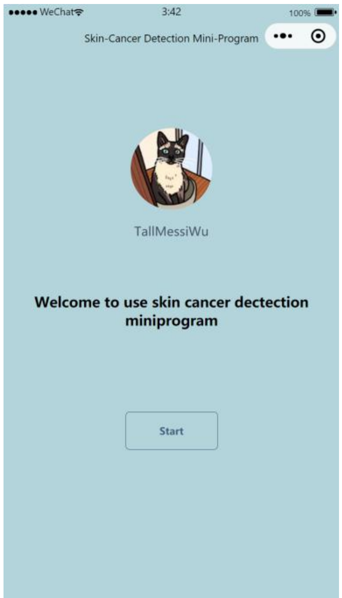

Figure 5: Homepage in the application.

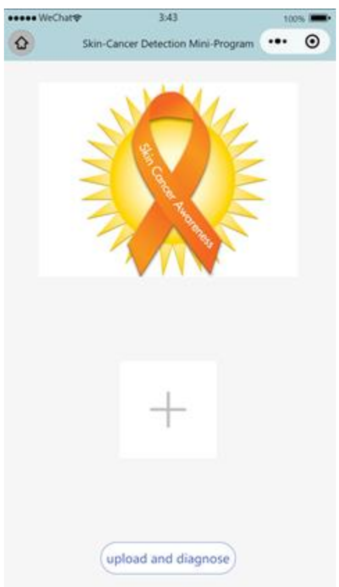

Figure 6: Interface for image uploading in the application.

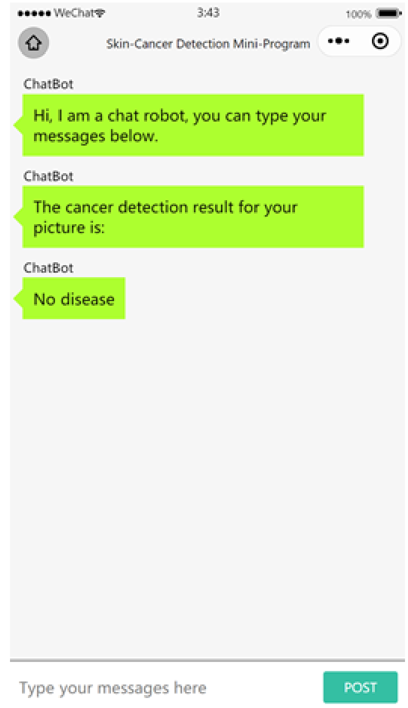

Figure 7: Chatbot in the application.

4 CONCLUSIONS

In this work, a real-world application was proposed to preliminarily diagnose skin cancer based on user's photos. The application needs to be fast and light-weight in order to be truly practical. In response, a skin-cancer-detecting CNN model, a WeChat application, and a web framework were developed. The model achieved great results and the application was tested to be able to return a diagnosis in under twenty seconds and only taking up less than two megabytes of space. This is much faster and lightweight than previous solutions. In the future, a more complex neural network architecture will be used to increase the accuracy of the model and a task distributing system will be constructed to allow multi-user use case. In addition, the app should allow inputting optional doctor measured data to further increase the reliability of the diagnosis.

REFERENCES

[1] Diepgen, T. L., & Mahler, V. (2002). The epidemiology of skin cancer. British Journal of Dermatology, 146, 1-6.

[2] Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018). GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing, 321, 321-331.

[3] Gouda, N., & Amudha, J. (2020, October). Skin cancer classification using ResNet. In 2020 IEEE 5th International Conference on Computing

Communication and Automation (ICCCA) (pp. 536-541). IEEE.

[4] Jerant, A. F., Johnson, J. T., Sheridan, C. D., & Caffrey, T. J. (2000). Early detection and treatment of skin cancer. American family physician, 62(2), 357-368.

[5] Karargyris, A., Karargyris, O., & Pantelopoulos, A. (2012, November). DERMA/care: An advanced image-processing mobile application for monitoring skin cancer. In 2012 IEEE 24th International Conference on Tools with Artificial Intelligence (Vol. 2, pp. 1-7). IEEE.

[6] Mijwil, Maad M. "Skin cancer disease images classification using deep learning solutions." Multimedia Tools and Applications 80.17 (2021): 26255-26271.

[7] Qiu, Y., Wang, J., Jin, Z., Chen, H., Zhang, M., & Guo, L. (2022). Pose-guided matching based on deep learning for assessing quality of action on rehabilitation training. Biomedical Signal Processing and Control, 72, 103323.

[8] Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific data, 5(1), 1-9.

[9] Vidya, M., & Karki, M. V. (2020, July). Skin cancer detection using machine learning techniques. In 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT) (pp. 1-5). IEEE.

[10] Zhang, N., Cai, Y. X., Wang, Y. Y., Tian, Y. T., Wang, X. L., & Badami, B. (2020). Skin cancer diagnosis based on optimized convolutional neural network. Artificial intelligence in medicine, 102, 101756.

[11] Zhang, J., Xie, Z., Sun, J., Zou, X., & Wang, J. (2020). A cascaded R-CNN with multiscale attention and imbalanced samples for traffic sign detection. IEEE Access, 8, 29742-29754.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.