Distribution Agreement

Dermatologic Diseases Prediction Using Deep Learning Method on Facial Images

By

Xisha Weng

Master of Science in Public Health

Department of Biostatistics and Bioinformatics

Hao Wu, Ph.D.
Thesis Advisor

Tianwei Yu, Ph.D.
Reader

Dermatologic Diseases Prediction Using Deep Learning Method on Facial Images

By

Xisha Weng

Master of Science

University of Georgia

2017

Thesis Committee Chair: Hao Wu, Ph.D.

Reader: Tianwei Yu, Ph.D.

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health,

in Biostatistics and Bioinformatics

2019

Abstract

Dermatologic Diseases Prediction Using Deep Learning Method on Facial Images

By Xisha Weng

Background: Computer vision is a research field where algorithms are developed so that computers can gain high-level understanding from digital images or videos. With the rapid development and application of deep learning method, it has become one of the hottest area in the field of artificial intelligence. In clinical research, medical images are commonly used for disease diagnosis, which is a perfect application area for computer vision methods. In recent years, deep learning methods are widely applied in different types of medical images, which greatly improves the image-based disease diagnosis. In this thesis work, we focus on one of the dermatologic diseases, rosacea, to identify the subtypes based on facial images. With previous works utilizing deep neural networks on other medical image data achieving great performance, such as skin cancer and retinopathy screening, it is reasonable to apply deep neural networks on the facial images for rosacea disease prediction.

Methods and Materials: Facial images from rosacea patients with subtypes: ETR, PPR, PhR were collected as raw data. The raw image data were preprocessed to crop (solely facial region cropped from raw image) and mask (decoloring unnecessary region from cropped image) data. Each dataset were used under all the proposed models to evaluate the effect of image preprocessing. A simple 5-layer convolutional neural networks (CNN) was constructed as baseline model for disease prediction. Transfer learning from existing deep neural networks including ResNet, Inception, Inception-ResNet model were used to evaluate the prediction performance. To train and evaluate the model performance, 80% of each dataset were used as training set, 10% as validation set and 10% as testing set for final performance evaluation.

Results: Baseline CNN does not perform well on the current dataset with slightly higher than 50% of validation accuracy. Using transfer learning on all the deep neural network models has good performance on all three datasets, with worst performance occurs when using raw data, indicating the necessity of image preprocessing. ResNet152 and Inception-ResNet V2 were selected for disease prediction with highest validation accuracy of 93.6% and 93.95%, respectively, on the crop data. The final performance of ResNet152 and Inception-ResNet V2 on crop testing set had 85.84% and 93.24% testing accuracy, respectively.

Conclusion: Using transfer learning based on Inception-ResNet V2 model has achieved the best prediction performance on rosacea disease prediction with a 93.24% testing accuracy. Deep neural network architectures including ResNet or Inception can also be considered for dermatologic disease prediction with moderately good performance. The application of convolutional neural network on medical image analysis and disease diagnosis is promising and can be considered to extend to other medical area with image data analysis.

Dermatologic Diseases Prediction Using Deep Learning Method on Facial Images

By

Xisha Weng

Master of Science

University of Georgia

2017

Thesis Committee Chair: Hao Wu, Ph.D.

Reader: Tianwei Yu, Ph.D.

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health,

in Biostatistics and Bioinformatics

2019

ACKNOWLEDGEMENT

I would like to thank my thesis advisor Dr. Hao Wu at the Department of Biostatistics and Bioinformatics, Emory University. I sincerely express my gratitude to Dr. Wu for his guidance, motivation and support during my program of study and his great advises for this thesis. I also extend my deep sense of gratitude to my thesis reader Dr. Tianwei Yu, for his kind support, advises and contributions.

I would like to thank all of my fellow graduate students at Emory University for their friendship and support during my progress of study.

Finally, special thanks to my family and friends.

CONTENTS

Introduction ········ **1**

Computer Vision ········ *1*

Application of CNN ········ *2*

Rosacea ········ *3*

Convolutional Neural Networks ········ *3*

Building Blocks of CNN Architecture ········ *4*

Transfer Learning ········ *5*

Study Goal ········ *6*

Methods ········ **6**

Data collection and preprocessing ········ *6*

CNN Models ········ *9*

Model Training and Evaluation ········ *11*

Results ········ **12**

Discussion ········ **16**

References ········ **19**

Introduction

Computer Vision

Computer vision is one of the hottest research field within Deep Learning for various of applications in image classification, object detection, object tracking, etc. Earlier, the application of computer vision was focused on mimicking human visual system for digital image processing. It was first introduced in the area of artificial intelligence in the late 1960s in the Summer Vision Project endowing robots with ability to describe what it saw from a camera (Papert, 2004). However, with the desire shifted to fully understanding image features from three-dimensional image structure, studies in the 1970s formed the foundation of many computer vision algorithms that exist today (Szeliski, 2010). The algorithms include extraction of edges from images, labeling of lines, non-polyhedral and polyhedral modeling and motion estimation, etc (Szeliski, 2010). Recent work on image analysis or computer vision has been focusing more on feature-based methods using the conjunction of machine learning techniques and complex optimization frameworks (Sebe et al., 2005; Freeman et al., 2008). It has been shown that deep neural networks (DNN) have greater capabilities for image feature recognition and are widely used in Computer Vision algorithms in modern works, with convolutional neural network as a most common method in visual imagery analysis (Sebe et al., 2005). Different from other machine learning approaches which rely on explicit features to be addressed for the model training and testing process (Abramoff et al., 2010; Yousefi et al., 2014), the algorithm of CNN allows it to learn features from training data and maximize the network's ability to distinguish from different categories. In addition, due to the annual ImageNet Large Scale Visual Recognition Competition (He et al., 2016; Krizhevsky et al., 2017), well established DNN architectures has shown great performance in image pattern recognition and classification. The development of DNN architectures resulted from winners of ImageNet Competition for each year: LeNet (Lecun et al., 1998), AlexNet (Krizhevsky et al., 2012), ZFNet (Zeiler and Fergus, 2014), GoogleNet/Inception (Szegedy et al., 2015; Szegedy et al., 2016), VGGNet (Simonyan and Zisserman, 2014), ResNet (He et al., 2016). Among those, AlexNet has been shown to outperformed any of previous computer vision algorithms in image recognition, while Inception, VGGNet and ResNet had further improvement in the overall performance (Krizhevsky et al., 2012; Simonyan and Zisserman, 2014; Szegedy et al., 2015; Szegedy et al., 2016; He et al., 2016).

Application of CNN

Modern biomedical research increasingly relies on image data as a primary source of information to identify the associations or differences between the image features. The complexity of the image data required state-of-art computational methods to fully explore the information. In addition, use of computers and modern software improve the accuracy and sensitivity of the biomedical image analysis to detect unnoticeable features (Meijering et al., 2016). With the development of computer vision algorithms deploying deep learning techniques, the application of CNN or transfer learning based on DNNs in medical research area has also achieved great performances. Pronounced work included the diabetic retinopathy screening study (Gulshan et al., 2016), dermatologist-level skin cancer classification (Esteva et al., 2017), lymph node metastasis detection (Bejnordi et al., 2017), and optic disease identification (Christopher et al., 2018), etc.

Rosacea

Rosacea is a chronic, inflammatory skin disease that affecting approximately 10% of the population (Tan et al., 2016). Symptoms present in various combinations and severity, often fluctuating between periods of exacerbation and remission (Tan et al., 2016; Rainer et al., 2017) . Rosacea is most often characterized by transient or persistent central facial erythema, visible blood vessels and often papules and pustules (Crawford et al., 2004). It can be classified into 4 broad subtypes based on patterns of physical findings: erythematotelangiectatic rosacea (Freeman et al., 2008), papulopustular rosacea (PPR), phymatous rosacea (PhR), and ocular rosacea (Wilkin et al., 2002). Among the four subtypes, ETR, PPR and PhR are characterized with different facial syptoms. Patients with ocular rosacea are commonly found to have blepharitis and conjunctivitis and usually found to be accompanies by other subtypes of rosacea (Crawford et al., 2004).

Convolutional Neural Networks

The classification on subtypes of rosacea are mainly done by medical staffs based on the standard classification criteria (Gallo et al., 2018). However, the process can be slow and inaccurate, especially when classifying the three main subtypes with facial symptoms (ETR, PPR and PhR). Previous applications using CNN in computer vision tasks and medical image classification indicated the potential capability of CNN in rosacea subtype classification task (Gulshan et al., 2016; Bejnordi et al., 2017; Esteva et al., 2017). Thus, it is reasonable to experiment the CNN algorithm with established DNN architectures on the rosacea disease prediction and classification.

Building Blocks of CNN Architecture

The main building blocks of a CNN model are convolution layers, pooling layers, and fully connected layers. A typical architecture consists of repetitions of a stack of one or several convolution layers and a pooling layer, followed by one or more fully connected layers (Figure 1).

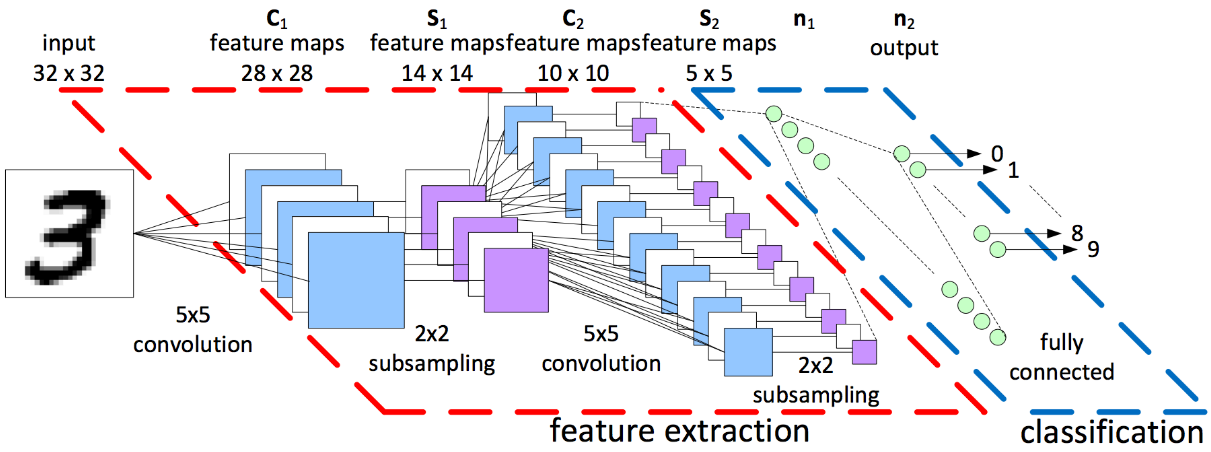


Figure 1. Typical structure of CNN model with 2 convolutional layers and 2 pooling (subsampling) layers followed by fully connected layers. Adapted from https://github.com/tavgreen/cnn-and-dnn.

The image is input as arrays of pixels for the CNN models to train and test. The convolution layer is a fundamental component that performs feature extraction. Each kernels of the convolutional layer learn a specific feature of the image. A pooling layer provides a typical down sampling operation which reduces the in-plane dimensionality of the feature maps and decrease the number of subsequent learnable parameters. The fully connected layers, also known as dense layers, transform the output feature map of the final convolution or pooling layer to a one-dimensional vector, with the length of the last fully connected layer equals to the categories of the classification task. The output of the last fully connected layer is applied with a softmax function to yield probabilistic values between 0 and 1 for each class.

Transfer Learning

The major techniques to successfully employ CNN in image recognition task included: 1). training CNN from scratch; 2). or using transfer learning to fine tune pre-train CNN model on current data. As a common problem to all deep learning networks, a much larger amount of data and computation time are required compared to more traditional machine learning techniques. However, in medical image analysis area, the scale of data is unavailable in most medical image classification tasks in order to train the complex deep learning neural networks. Small dataset may also result in model overfitting. Transfer learning is now a common technique to be used to overcome this situation (Pan and Yang, 2010). It allows us to utilize the pre-trained CNN model, usually pretrained on large general image dataset, such as ImageNet database (Russakovsky et al., 2015), as a starting point for an specific imaging task (Pan and Yang, 2010). Transfer learning has previously been employed to train models for other medical image recognition tasks, such as skin cancer cases (Esteva et al., 2017) and optic disease images (Christopher et al., 2018). In addition, both works illustrated reduced training time and computation cost as well as good testing performance. Thanks to the yearly ImageNet Challenge (Russakovsky et al., 2015), advanced and deep CNN models have been explored and applied to different areas. The most popular deep CNN models include VGGNet (Simonyan and Zisserman, 2014), ResNet (He et al., 2016) and Inception (Szegedy et al., 2015; Szegedy et al., 2016), etc.

Study Goal

The aim of this project is to develop and evaluate convolutional neural networks to identify the three rosacea subtypes with facial symptoms (ETR, PPR, PhR) in the patient skin images. The classification accuracy of different CNN architectures, including simple CNN model from scratch and transfer learning from deep neural networks (Szegedy et al., 2015; Szegedy et al., 2016) were evaluated.

Methods

Data collection and preprocessing

Study participants were selected from Xiangya Hospital Central South University. All participants were diagnosed as having one of the three facial subtypes of rosacea: ETR, PPR, PhR. Face images of each patients were taken and used in this project. For analysis, photographs were stored as high resolution (~1800 × ~2500 pixels) JPG images. Total 338 ETR, 756 PhR and 368 PPR images were collected as the raw data for analysis. All images were preprocessed to a new dataset evaluate the effect of image preprocessing on model performance.

1. *Face Cropping*

To avoid the unnecessary region effect on model computation and performance, the face regions of the images were detected and cropped using the Multitask Cascaded Convolutional Networks Model (MTCNN) (Zhang et al., 2016). There was a few data loss due to the unsuccessful detection in faces on some raw images. The final cropped dataset included 336 ETR, 694 PhR and 362 PPR images. The programming

2. *Color Mask*

To further eliminate the uninformative regions, additional color masking was processed based on the cropped dataset. Since all the symptoms are shown as redness in the human face region, region with color rather than red or pink, such as background or hair, are masked as black color based on the RGB and HSV skin color model (Anwar et al., 2019). There were 336 ETR, 694 PhR and 362 PPR images in the mask dataset.

The examples of raw, cropped and masked images are shown in Figure 2.

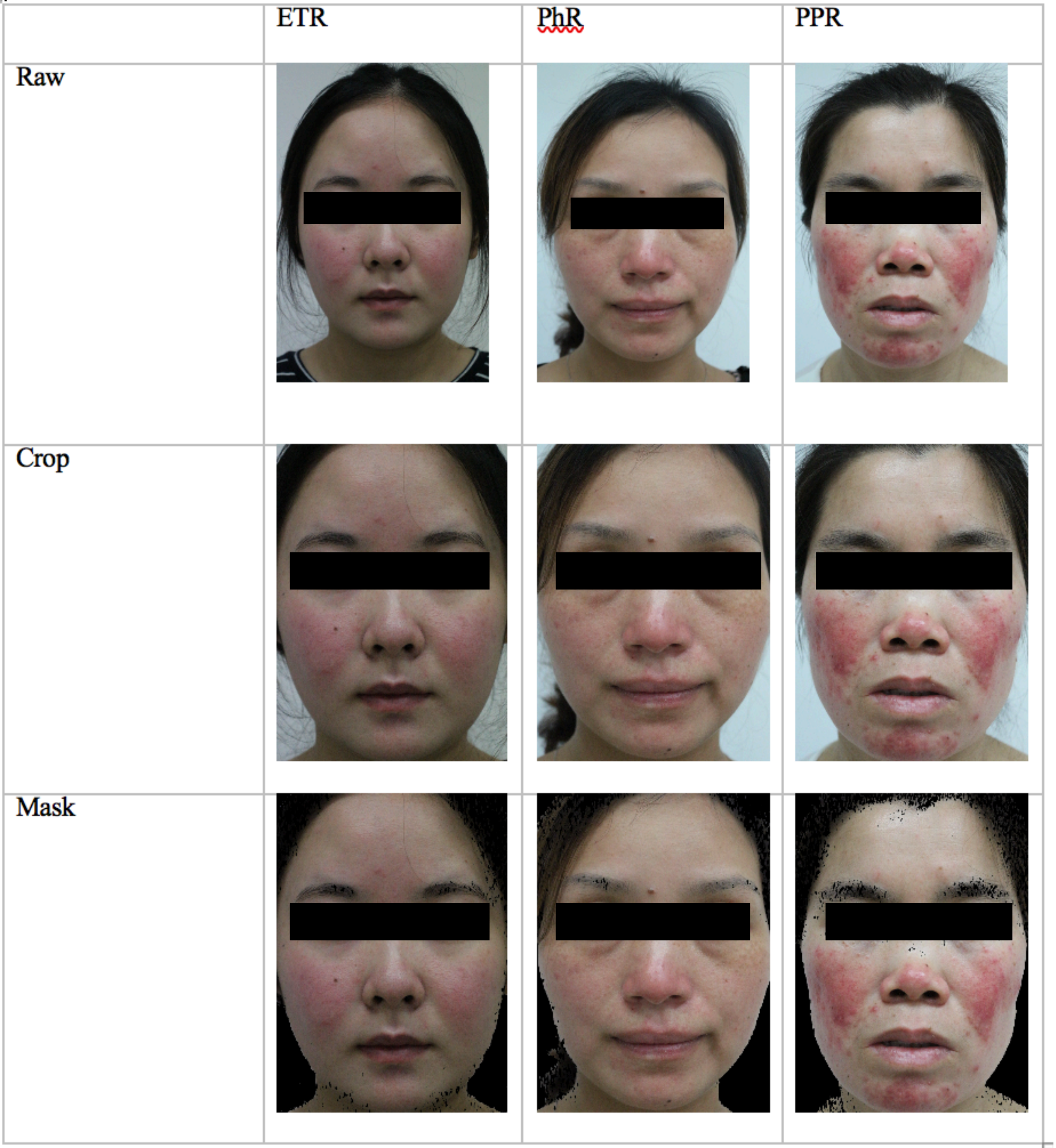


Figure 2. Examples of raw, face cropped, color masked images of ETR, PhR and PPR rosacea cases.

CNN Models

1. *Baseline CNN*

The baseline model is a simple convolutional neural networks with 5 convolutional layers and a max pooling layer followed by each convolutional layer. The final max pooling layer is followed by three fully connected layers and softmax function was applied for the final classification. The output of all convolutional layer and first two fully connected layers was applied with the ReLU activation function for nonlinearity. The model architecture is shown in Figure 3A.

2. *Deep CNN models*

Three main different deep learning architecture were evaluated using transfer learning technique: Inception V3 (Szegedy et al., 2016), ResNet (He et al., 2016) and Inception-ResNet V2(Szegedy et al., 2016). These architectures have been widely adopted for both general and medical image classification tasks and their performances are commonly used for comparison. The architectures are shown in Figure 3A and Figure 4. For the ResNet architecture, ResNet18, ResNet50, ResNet101 and ResNet152 models were evaluated in the current project. All ResNet models have similar architecture but differs in the numbers of residual blocks and layers.

All the deep CNN models were pretrained on ImageNet database (Russakovsky et al., 2015) so that model weights were initialized based on pretraining on a general image dataset, and transfer learning approach can be applied on current dataset. Additional training was performed on rosacea medical image data.

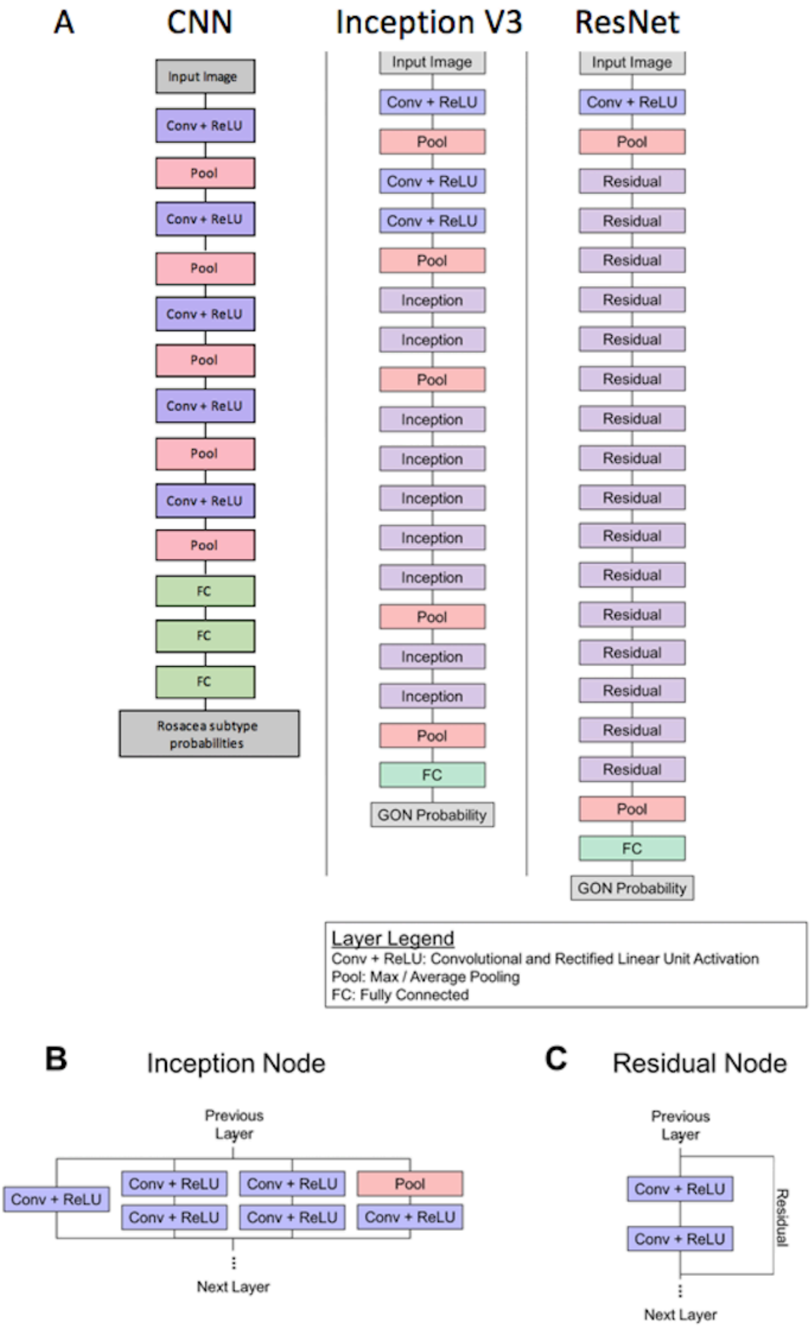


Figure 3. (A) Schematic diagrams of the baseline CNN, Inception V3, and ResNet. The Inception (B) and Residual (C) are used as building blocks for the Inception and ResNet architectures, respectively. Adapted from Christopher et al., 2018.

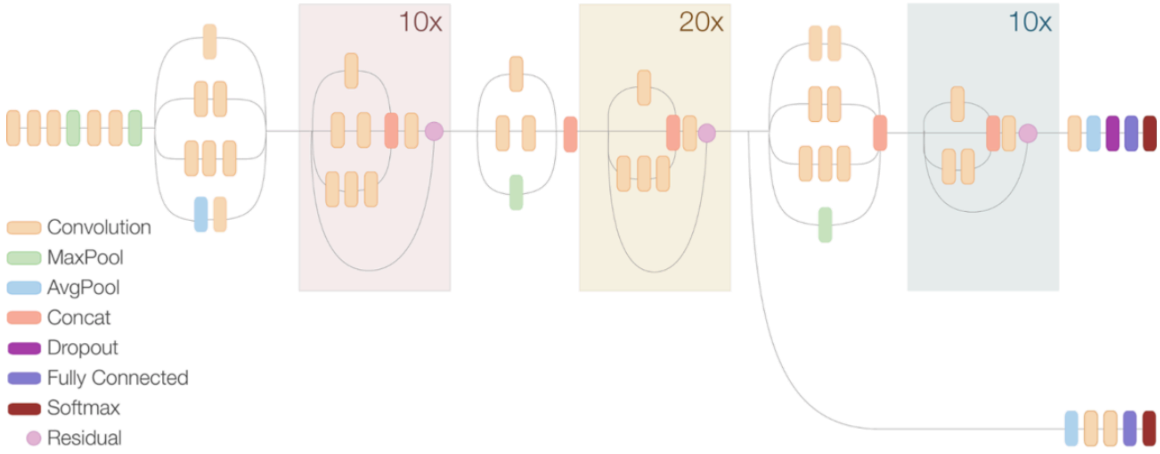


Figure 4. Schematic diagram of Inception-ResNet V2. Adapted from https://ai.googleblog.com/2016/08/improving-inception-and-image.html

Model Training and Evaluation

The image datasets (raw, crop, mask set) were randomly divided into independent training, validation and testing sets using an 80-10-10 percentage split. For training, a total of 100 epochs with batch size of 4 were performed. The fine-tuning procedures included adding dropout layer, learning rate schedule modification and random image augmentation. The model was evaluated on the validation set at every epoch. This process was repeated for each CNN model, including baseline CNN, Inception V3, ResNet and Inception-ResNet V2. The model training for baseline CNN, Inception V3 and ResNet were performed using PyTorch (Paszke et al., 2017), and for Inception-ResNet V2 was performed using Keras (François, 2015). The model that achieved the highest performance on the validation set was selected for evaluation on the testing set. The highest training accuracy, validation accuracy and final best model testing accuracy were reported.

Results

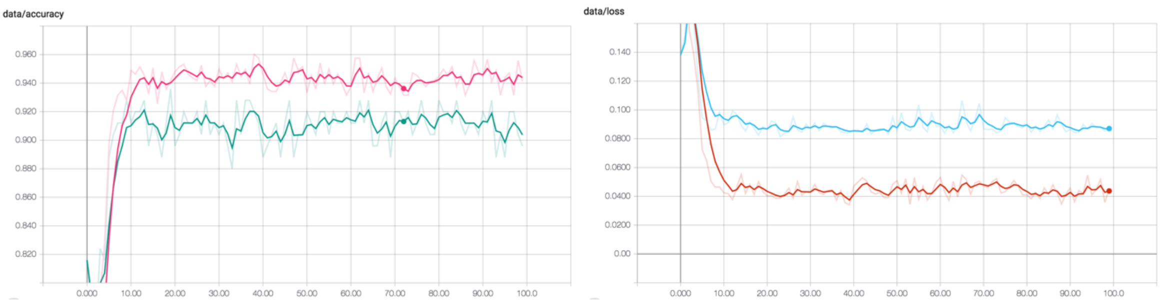
Model performance was evaluated on the validation datasets for each epoch. Table 1 shows the best validation accuracy of the baseline CNN, ResNet and Inception models. The baseline CNN did not perform very well on any of the dataset, with only slightly higher than 50% of validation accuracy. The transfer learning on the deep neural network models performed quite well on raw, crop and mask data set. The models have an overall worst performance on the raw dataset. The Inception V3 model with 0.5 dropout on the last layer had higher performance compared without dropout. It performed the best on the mask data with a 91.47% validation accuracy. For the ResNet models, the ResNet152 performed the best on the crop data with a validation accuracy of 93.6%. For the Inception-ResNet V2 model performed the best on the crop data with a validation accuracy of 93.95%, which is the highest among all the experiments. In addition, the ResNet18, 50 and 101 had similar performance on all three dataset, while the deeper neural network ResNet152 and Inception-ResNet V2 performed better on the crop data set.

Models	Raw	Crop	Mask
Baseline CNN	51.77%	54.40%	51.16%
Inception V3	89.6%	87.23%	89.15%
Inception V3(with dropout)	87.94%	90.40%	91.47%
ResNet18	90.78%	90.40%	90.70%
ResNet50	89.36%	89.60%	91.47%
ResNet101	89.36%	88.80%	89.15%
ResNet152	90.78%	93.60%	87.15%
Inception-Resnet V2	84.92%	93.95%	88.15%

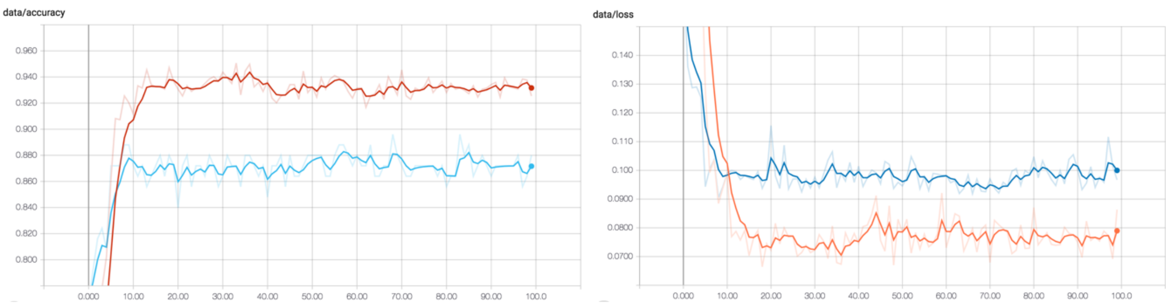
Table 1. The best validation accuracy of the models on raw, crop and mask validation set

The Figure 5 shows the model performance including the accuracy and loss for training and validation process. Only the plots for ResNet152, Inception V3 and Inception-ResNet V2 are shown to illustrate the model evaluation. All the models converge fast within about 10 epochs of training. The Inception V3 models shows the largest disparity between training and validation accuracy and loss, which indicates a higher possibility of overfitting. The plots of Inception-ResNet V2 model had smallest disparity indicating no apparent sign of overfitting.

a) ResNet152



b) Inception V3



c) Inception-ResNet V2

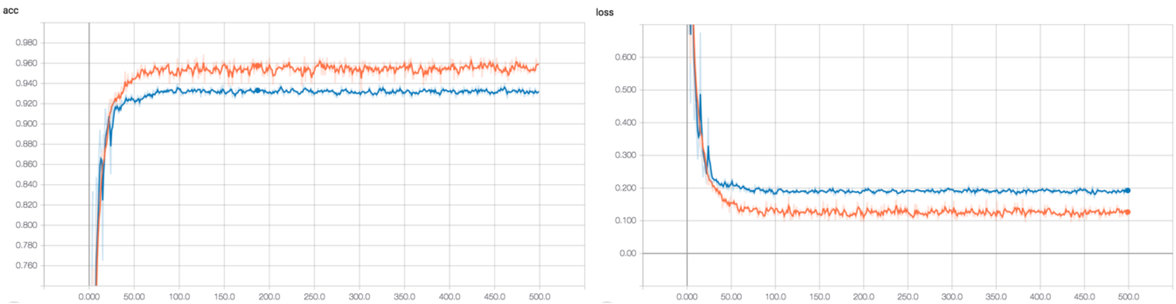
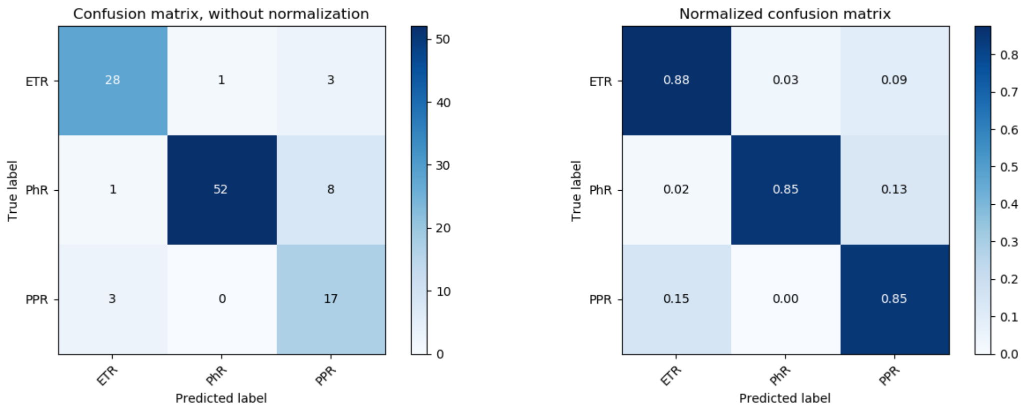


Figure 5. Model performance on the training and validation dataset by epochs for a). ResNet152, b). Inception V3, c). Inception-ResNet V2.

Based on the above results, the Inception-ResNet V2 and ResNet152 were chosen for evaluation on the test set of the crop data. The final testing accuracy for ResNet152 and Inception-ResNet V2 were 85.84%, and 93.24%, respectively. The confusion matrix for the testing set are shown in Figure 6. It is indicated that the subtypes ETR and PhR are more likely to be mistakenly classified as subtype PPR by both models. The Inception-ResNet V2 had an overall better performance than ResNet152 on rosacea subtype prediction based on current dataset results. However, it is hard to compare between models since the different testing set were selected through the randomization process.

a) ResNet152



b) Inception-ResNet V2

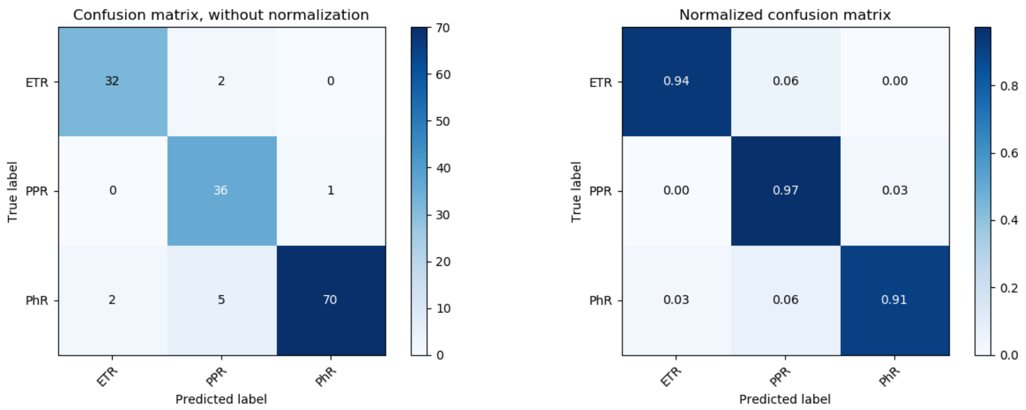


Figure 6. Confusion matrix on crop data testing set with best models for a). ResNet152, b). Inception-ResNet V2.

Discussion

The baseline CNN model did not perform well, possibly due to the complexity of image and small sample size for each category. Since most of the features are similar between the categories, the simple CNN model is not sufficient to acquire the difference between the categories. Simply increasing the number of layers in the CNN model won't help due to the small sample size. On the other hand, transfer learning methods show significantly improved results. They utilize the learnt features from larger scale of image set and fine-tune the existing model to capture more advanced difference between the categories of one specific domain (Pan and Yang, 2010; Feng et al., 2019). Such procedure borrows information from historical data, and is proved to be very effective in our data where the sample size is small.

The model fine-tune procedure for deep neural networks included learning rate selection, modifying the last few layers of the existing model, adding dropout layers, horizontal image flip, etc. It is shown that the dropout layers increased performance for Inception V3 model and decreased overfitting (Table 1). In addition, the horizontal image flipping increased sample size which may explain part of increased performance for deep neural networks compared with baseline CNN.

As regard to the deep neural network model's performances, all the models performed quite well on the validation set with an approximately ~90% validation accuracy. As expected, all the models have a similar to worse performance on the raw dataset compared with crop and mask dataset. This indicates that the background information on the image do have a noisy impact on the model performance. The Inception V3, ResNet18, 50 and 101 had slightly better performance on the mask dataset compared with crop dataset. Contradictory, the much deeper ResNet152 and Inception-ResNet V2 models performed best on the crop dataset with the validation accuracy of 93.60% and 93.95%, respectively. The underlying mechanism is not entirely clear, however, one assumption is that the mask dataset eliminated too much information on the image which resulted a worse performance with a deeper neural networks.

The plots from Figure 5 indicated the evidence of overfitting for the Inception V3 model, but less extent of overfitting for ResNet and Inception-ResNet V2 models. The residual blocks in the ResNet model are designed as a way to eliminating the possibility of vanishing gradient with deeper neural networks, and to give at least the same performance from the output before the residual block with the idea of identity mapping (He et al., 2016). With introducing the residual blocks into the Inception models, it significantly improved the model performance compared with ResNet or Inception V3, which is consistent with the results published previously (Szegedy et al., 2016).

The overall model selection gives priority to the ResNet152 and Inception-ResNet V2 to be applied on the crop data set (Table 1), with a final testing accuracy of 85.84% and 93.24%, respectively. It shows that Inception-ResNet V2 performed better when predicting the incoming unseen data, while the ResNet152 had a lower performance. One reason may be the proposed increased performance applying residual blocks with inception blocks connection (Szegedy et al., 2016). In the current project, the train test set, although randomly split, were different between the two models, due to the different framework utilization (PyTorch framework for ResNet modeling, and Keras for Inception-ResNet V2). The small sample size and different train test set may play a role in the disparity of model performance. For further investigation, a K-fold cross-validation procedure is strongly recommended to minimizing the differences in performance introduced by the different train test set (Schaffer, 1993; Erickson, 2017).

In conclusion, the transfer learning on existing deep neural networks had good prediction results (93.24% testing accuracy) on human skin medical images in the current project. Applying machine learning algorithm such as convolutional neural networks on image data is a promising way for the medical disease diagnosis. Future application for other image data not limited to human face may be considered.

References

Abramoff, M. D., Garvin, M. K. and Sonka, M. (2010). Retinal Imaging and Image Analysis. IEEE Reviews in Biomedical Engineering, 3, 169-208. doi:10.1109/RBME.2010.2084567

Anwar, N., Rahman, A., Chong, K. and See, J. (2019). RGB-H-CbCr Skin Colour Model for Human Face Detection.

Bejnordi B. E., Veta, M., Johannes V. D. P., van Ginneken B., Karssemeijer N., Litjens G., . . . Venancio R. (2017). Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. Jama, 318(22), 2199-2210. doi:10.1001/jama.2017.14585

Crawford, G. H., Pelle, M. T. and James, W. D. (2004). Rosacea: I. Etiology, pathogenesis, and subtype classification. Journal of the American Academy of Dermatology, 51(3), 327-341. doi:https://doi.org/10.1016/j.jaad.2004.03.030

Christopher, M., Belghith, A., Bowd, C., Proudfoot, J. A., Goldbaum, M. H., Weinreb, R. N., Christopher, A. G., Liebmann J. M. and Zangwill, L. M. (2018). Performance of Deep Learning Architectures and Transfer Learning for Detecting Glaucomatous Optic Neuropathy in Fundus Photographs. Scientific Reports, 8(1), 16685. doi:10.1038/s41598-018-35044-9

Erickson, B. J., Korfiatis P., Akkus Z. and Kline T. L. (2017). Machine Learning for Medical Imaging. RadioGraphics, 37(2), 505-515. doi:10.1148/rg.2017160130

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118. doi:10.1038/nature21056

François, C. (2015). Keras. In. https://github.com/fchollet/keras: Github.

Freeman, W., Perona, P. and Schölkopf, B. (2008). Guest Editorial. International Journal of Computer Vision, 77(1), 1-1. doi:10.1007/s11263-008-0127-7

François, C. (2015). Keras. In. https://github.com/fchollet/keras: Github.

Feng, S., Zhou, H. and Dong, H. (2019). Using deep neural network with small dataset to predict material defects. Materials & Design, 162, 300-310. doi:https://doi.org/10.1016/j.matdes.2018.11.060

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., . . . Webster, D. R. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. Jama, 316(22), 2402-2410. doi:10.1001/jama.2016.17216

Gallo, R. L., Granstein, R. D., Kang, S., Mannis, M., Steinhoff, M., Tan, J. and Thiboutot, D. (2018). Standard classification and pathophysiology of rosacea: The 2017 update by the National Rosacea Society Expert Committee. Journal of the American Academy of Dermatology, 78(1), 148-155. doi:10.1016/j.jaad.2017.08.037

He, K., Zhang, X., Ren, S. and Sun, J. (2016). Deep Residual Learning for Image Recognition. Paper presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Krizhevsky A., Sutskever, I. and Hinton G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. 1097--1105.

Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Commun. ACM, 60(6), 84-90. doi:10.1145/3065386

Lecun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324. doi:10.1109/5.726791

Meijering, E., Carpenter, A. E., Peng, H., Hamprecht, F. A. and Olivo-Marin, J.-C. (2016). Imagining the future of bioimage analysis. Nature Biotechnology, 34, 1250. doi:10.1038/nbt.3722

Papert, S. (2004). The Summer Vision Project.

Pan, S. J. and Yang, Q. (2010). A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345-1359. doi:10.1109/TKDE.2009.191

Paszke, A., Gross S., Chintala S., Chanan G., Yang E., DeVito Z., Lin Z., Desmaison A., Antiga L. and Lerer A. (2017). Automatic differentiation in PyTorch.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., . . . Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. Int. J. Comput. Vision, 115(3), 211-252. doi:10.1007/s11263-015-0816-y

Rainer, B. M., Kang, S. and Chien, A. L. (2017). Rosacea: Epidemiology, pathogenesis, and treatment. Dermato-endocrinology, 9(1), e1361574-e1361574. doi:10.1080/19381980.2017.1361574

Schaffer, C. (1993). Selecting a classification method by cross-validation. Machine Learning, 13(1), 135-143. doi:10.1007/BF00993106

Sebe, N., Cohen, I., Garg, A. and Huang, T. S. (2005). Machine Learning in Computer Vision (Computational Imaging and Vision): Springer-Verlag.

Szeliski, R. (2010). Computer Vision: Algorithms and Applications: Springer-Verlag.

Simonyan, K. and Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition.

Szegedy, C., Wei, L., Yangqing, J., Sermanet, P., Reed, S., Anguelov, D., . . . Rabinovich, A. (2015). Going deeper with convolutions. Paper presented at the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Szegedy, C., Ioffe, S. and Vanhoucke, V. (2016). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z. B. (2016). Rethinking the Inception Architecture for Computer Vision.

Tan, J., Schofer, H., Araviiskaia, E., Audibert, F., Kerrouche, N. and Berg, M. (2016). Prevalence of rosacea in the general population of Germany and Russia - The RISE study. J Eur Acad Dermatol Venereol, 30(3), 428-434. doi:10.1111/jdv.13556

Wilkin, J., Dahl, M., Detmar, M., Drake, L., Feinstein, A., Odom, R. and Powell, F. (2002). Standard classification of rosacea: Report of the National Rosacea Society Expert Committee on the Classification and Staging of Rosacea. J Am Acad Dermatol, 46(4), 584-587.

Yousefi, S., Goldbaum, M. H., Balasubramanian, M., Medeiros, F. A., Zangwill, L. M., Liebmann, J. M., . . . Bowd, C. (2014). Learning From Data: Recognizing Glaucomatous Defect Patterns and Detecting Progression From Visual Field Measurements. IEEE Transactions on Biomedical Engineering, 61(7), 2112-2124. doi:10.1109/TBME.2014.2314714

Zeiler, M. D. and Fergus, R. (2014). Visualizing and understanding convolutional networks. Paper presented at the European conference on computer vision.

Zhang, K., Zhang, Z., Li, Z. and Qiao, Y. (2016). Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. IEEE Signal Processing Letters, 23(10), 1499-1503. doi:10.1109/LSP.2016.2603342