AI-BASED LOCALIZATION AND CLASSIFICATION OF SKIN DISEASE WITH ERYTHEMA

ABSTRACT

Although computer-aided diagnosis (CAD) is used to improve the quality of diagnosis in various medical fields such as mammography and colonography, it is not used in dermatology, where noninvasive screening tests are performed only with the naked eye, and avoidable inaccuracies may exist. This study shows that CAD may also be a viable option in dermatology by presenting a novel method to sequentially combine accurate segmentation and classification models. Given an image of the skin, we decompose the image to normalize and extract high-level features. Using a neural network-based segmentation model to create a segmented map of the image, we then cluster sections of abnormal skin and pass this information to a classification model. We classify each cluster into different common skin diseases using another neural network model. Our segmentation model achieves better performance compared to previous studies, and also achieves a near-perfect sensitivity score in unfavorable conditions. Our classification model is more accurate than a baseline model trained without segmentation, while also being able to classify multiple diseases within a single image. This improved performance may be sufficient to use CAD in the field of dermatology.

INTRODUCTION

Computer-aided diagnosis (CAD) is a computer-based system that is used in the medical imaging field to aid healthcare workers in their diagnoses. CAD has become a mainstream tool in several medical fields such as mammography and colonography. However, in dermatology, although skin disease is a common disease, one in which early detection and classification is crucial for the successful treatment and recovery of patients, dermatologists perform most noninvasive screening tests only with the naked eye. This may result in avoidable diagnostic inaccuracies as a result of human error, as the detection of the disease can be easily overlooked. Furthermore, classification of a disease is difficult due to the strong similarities between common skin disease symptoms. Therefore, it would be beneficial to exploit the strengths of CAD using artificial intelligence techniques, in order to improve the accuracy of dermatology diagnosis. This paper shows that CAD may be a viable option in the field of dermatology using state-of-the-art deep learning models.

The segmentation and classification of skin diseases has been gaining attention in the field of artificial intelligence because of its promising results. Two of the more prominent approaches for skin disease segmentation and classification are clustering algorithms and support vector machines (SVMs). Clustering algorithms generally have the advantage of being flexible, easy to implement, with the ability to generalize features that have a similar statistical variance. Trabelsi et al. experimented with various clustering algorithms, such as fuzzy c-means, improved fuzzy cmeans, and K-means, achieving approximately 83% true positive rates in segmenting a skin disease. Rajab et al.implemented an ISODATA clustering algorithm to find the optimal threshold for the segmentation of skin lesions. An inherent disadvantage of clustering a skin disease is its lack of robustness against noise. Clustering algorithms rely on the identification of a centroid that can generalize a cluster of data. Noisy data, or the presence of outliers, can significantly degrade the performance of these algorithms. Therefore, with noisy datasets, caused by images with different types of lighting, non-clustering algorithms may be preferred; however, Keke et al. implemented an improved version of the fuzzy clustering algorithm using the RGB, HSV, and LAB color spaces to create a model that is more robust to noisy data. SVMs have gained attention for their effectiveness in high-dimensional data and their capability to decipher "…subtle patterns in noisy and complex datasets". Lu et al. segmented erythema in the skin using the radial basis kernel function that allows SVMs to separate nonlinear hyperplanes. Sumithra et al. combined a linear SVM with a k-NN classifier to segment and classify five different classes of skin lesions. Maglogiannis et al. implemented a threshold on the RGB value for segmentation and used an SVM for classification. Although more robust than clustering algorithms, SVMs are more reliant on the preprocessing of data for feature extraction. Without preprocessing that allows a clear definition of hyperplanes, SVMs may also underperform.

Owing to the disadvantages of these traditional approaches, convolution neural networks (CNNs) have gained popularity because of their ability to extract high-level features with minimal preprocessing. CNNs can expand the advantages of SVMs, such as robustness in noisy datasets without the need for optimal preprocessing, by capturing image context and extracting high-level features through down-sampling. CNNs can interpret the pixels of an image within its own image-level context, as opposed to viewing each pixel in a dataset-level context. However, although down-sampling allows CNNs to view an image in its own context, it degrades the resolution of the image. Although context is gained, the location of a target is lost through downsampling. This is not a problem for classification, but causes some difficulty for segmentation, as both the context and location of the target are essential for optimal performance. To solve this, up-sampling is needed, which works in a manner

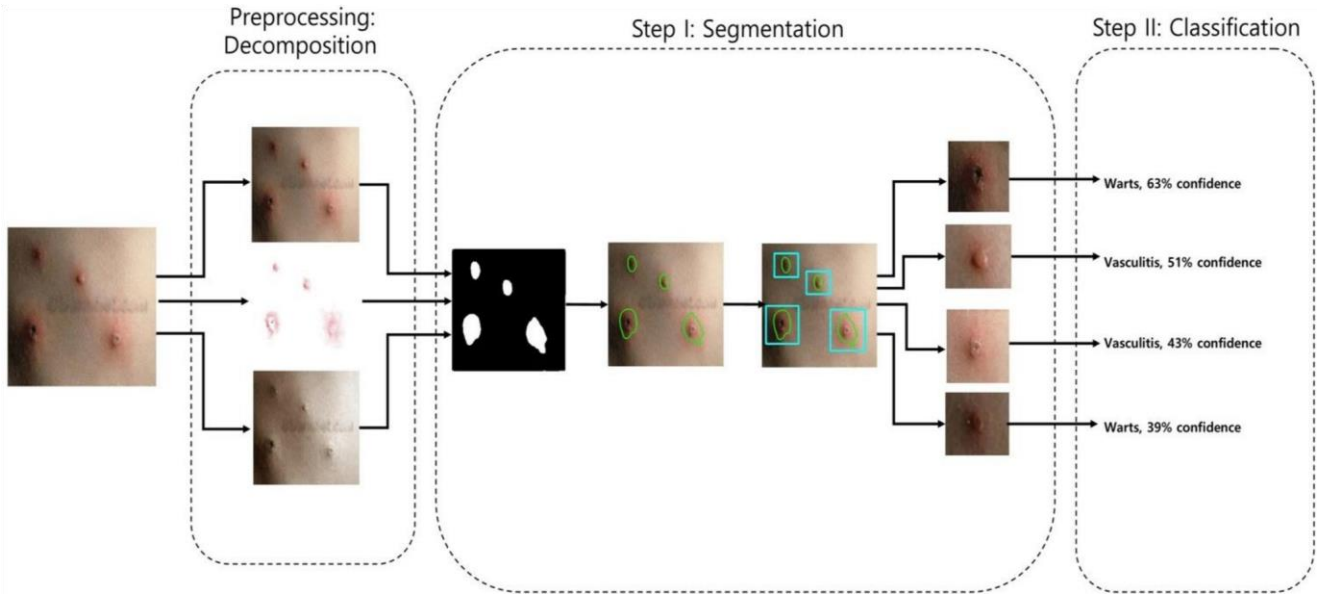
opposite to that of down-sampling, in the sense that it increases the resolution of the image. While down-sampling takes a matrix and decreases it to a smaller feature map, up-sampling takes a feature map and increases it to a larger matrix. By learning to accurately create a higher-resolution image, CNNs can determine the location of the targets to segment. Thus, for segmentation, we use a combination of down-sampling and upsampling, whereas for classification, we use only down-sampling. To further leverage the advantages of CNNs, skip-connections were introduced, which provided a solution to the degradation problem that occurs when CNN models become too large and complex. We implement skip-connections in both segmentation and classification models. In the segmentation model, blocks of equal feature numbers are connected between the down and up-sampling sections. In the classification model, these skip-connections exist in the form of inverted residual blocks. This allows our models to grow in complexity without any performance degradation.

In this paper, we present a method to sequentially combine two separate models to solve a larger problem. In the past, skin disease models have been applied to either segmentation or classification. In this study, we sequentially combine both models by using the output of a segmentation model as input to a classification model. In addition, although past studies of nonCNN segmentation models used innovative preprocessing methods, recent CNN developments have focused more on the architecture of the model than on the preprocessing of data. As such, we apply an innovative preprocessing method to the data of our CNN segmentation model. The methods described above lack the ability to localize and classify multiple diseases within one image; however, we have developed a method to address this problem. Our objective is two-fold. First, we show that CAD can be used in the field of dermatology. Second, we show that state-ofthe-art models can be used with current computing power to solve a wider range of complex problems than previously imagined. We begin by explaining the results of our experimentation, followed by a discussion of our findings, a more detailed description of our methodology, and finally, the conclusions that can be drawn from our study.

RESULT AND DISCUSSION

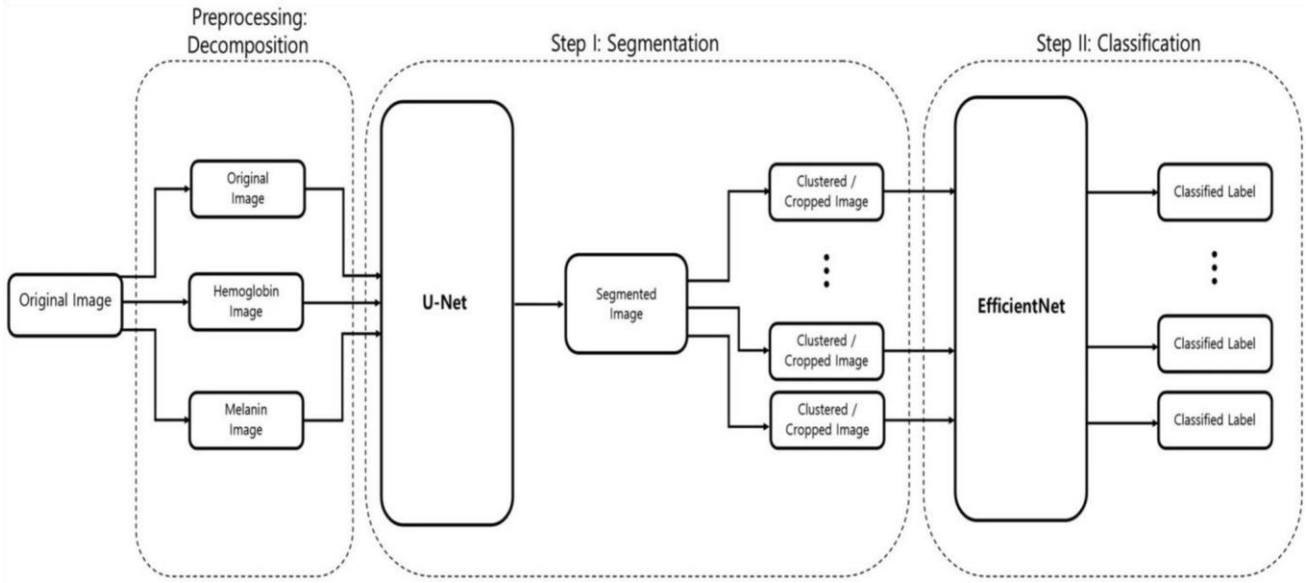
We started with the original image. We preprocessed this image by decomposing it into its hemoglobin and melanin constituents. These images were then input to the U-Net to generate the segmented output. We drew contours around each cluster and used a convex hull algorithm to draw rectangles around these clusters and crop them as individual images. These cropped images were used as input to the EfficientNet, which generated a prediction along with the confidence

rate. The K-means clustering algorithm showed sub-optimal performance, owing to its limitations with noisy data. The SVM method showed a significant improvement in performance, that was attributed to the advantages of using SVMs to extract information from decomposition, rather than clustering algorithms. Even without the extra information, the U-Net trained without decomposition outperformed the previous two methods in terms of sensitivity. The U-Net model was also trained with decomposition and showed the highest sensitivity rate. This was a result of an improved performance when there is a smaller area to search for the disease. Because we segmented only the abnormal areas of the skin, the EfficientNet model showed better performance compared to images with a larger ratio of normal skin. Thus, we can learn about the location of the disease that is present in an image and improve performance by training a CNN model to focus on particular subsections of the images. Figure 3 shows a visual representation of this claim using an implementation of the Grad-CAM method11. Activation, which is the intensity with which a model focuses on an area, is represented on a rainbow colormap. Red represents areas of highest activation, while violet represents areas of lowest activation. When trained with unsegmented data, our model focused on an area larger than that of abnormal skin. The area of activation was highest around the erythema, although there were other areas of high activation. In these cases, the model utilized the shapes of body parts for classification. This decreases performance because skin disease can appear in virtually any part of body and there is a lack of data required to form an association between the probability of a skin disease based on the body part. When trained with contextually segmented data, however, our model correctly focused only on erythema. The area of activation was highest around the erythema, while areas of low activation were demonstrated elsewhere. Not only does this add validity to our reported results, but this is also a justification for the inclusion of the segmentation phase before the classification phase because there were clear improvements in all metrics regarding the use of the U-Net before the EfficientNet. The main contribution of our study is researching the viability of CAD in the field of dermatology. This is achieved through the increase in the classification performance of skin disease images, owing to the increase in performance of segmentation. However, our model is most effective with camera images of skin diseases with erythema, which is a limitation of our study. We chose to focus on camera images and erythema because these images are very accessible, and erythema is one of the most common symptoms of skin disease. In addition, currently we only classify diseases into 18 categories due to the limitations of the data. In the future, we plan to create a more comprehensive skin disease classification model, and this seems to be viable if enough data can be obtained. In addition, we plan to work on a method to help dermatologists with time-series analysis of patients. This seems viable with the accumulation of data through CAD.



ANAIYSIS METHODOLOGY

Our 2-phase analysis model for localization and classification is shown via the pseudocode in Algorithm 1 and visually in . We decomposed the original image into its hemoglobin and melanin constituents using preprocessing, to help our model extract valuable information from data that would have been otherwise unavailable. We provide these images as input to our segmentation model, the U-Net, which generated a segmented image. This segmented image was then analyzed for clusters, which were subsequently cropped and input to our classification model, the EfficientNet, which then produced a classified label, thus completing our analysis model.



The data for training and testing were obtained from Dermnet NZ, an archive of skin disease information launched and maintained by a group of dermatologists from New Zealand. The site provides open source images with labels. We selected 18 top-level categories each of which included enough data, besides including erythema as one of its common symptoms. Using a web crawler, we gathered a total of 15,851 images. Among the images obtained through Dermnet, the erythema of 100 images was masked by dermatologists, to be used as a ground truth. For segmentation, 60 images were used for training, and 40 images were used for testing. For classification, 13,473 images were used for training, and 2,378 images were used for testing. In addition, the test set for classification was split before segmentation cropping to prevent the subsections of one image from appearing in both the training and testing sets. Table shows the distribution of data in greater detail. We chose the 100 images for segmentation in a balanced manner from each class, to minimize any bias that could occur during the classification phase.

CONCLUSION

We have shown that even without a large dataset and high-quality images, it is possible to achieve sufficient accuracy rates. In addition, we have shown that current state-of-the-art CNN models can outperform models created by previous research, through proper data preprocessing, self-supervised learning, transfer learning, and special CNN architecture techniques. Furthermore, with accurate segmentation, we gain knowledge of the location of the disease, which is useful in the preprocessing of data used in classification, as it allows the CNN model to focus

on the area of interest. Lastly, unlike previous studies, our method provides a solution to classify multiple diseases within a single image. With higher quality and a larger quantity of data, it will be viable to use state-of-the-art models to enable the use of CAD in the field of dermatology.

REFERENCE

1. Rajab, M. I., Woolfson, M. S. & Morgan, S. P. Application of region-based segmentation and neural network edge detection to skin lesions. *Comput. Med. Imaging Graph.* **28**, 61– 68. (2004).

2. Keke, S., Peng, Z. & Guohui, L., Study on skin color image segmentation used by fuzzyc-means arithmetic. In *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, Yantai, 612–615. (2010).

3. Hongmao, S. *Quantitative Structure-Activity Relationships: Promise, Validations, and Pitfalls in A Practical Guide to Rational Drug Design* 163–192 (Woodhead Publishing, Sawston, 2016).

4. Lu, J., Manton, J. H., Kazmierczak E. & Sinclair, R., Erythema detection in digital skin images. In *2010 IEEE International Conference on Image Processing*, Hong Kong, 2545–2548.(2010).

5. Sumithra, R., Suhil, M. & Guru, D. S. Segmentation and classification of skin lesions for disease diagnosis. *Proced. Comput. Sci.* **45**, 76–85. (2015).

6. Maglogiannis, I., Zafiropoulos, E. & Kyranoudis, C. Intelligent segmentation and classification of pigmented skin lesions in dermatological images in Advances in Artificial Intelligence. SETN 2006. In *Lecture Notes in Computer Science* Vol. 3955 (eds Antoniou, G. *et al.*) 214–223 (Springer, Berlin, 2006).

7. Albawi, S., Mohammed, T. A. & Al-Zawi, S., Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)*, Antalya, 1–6. (2017).

8. Selvaraju, R. *et al*. Grad-CAM: Visual explanations from deep networks via gradientbased localization. *Int. J. Comput. Vis.* **128**, (2019).

9. Gutman, D., Codella, N., Celebi, E., Helba, B., Marchettic, M., Mishra, N., & Halpern, A., Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC).(2016).

10. Codella, N., Gutman, D., Celebi, ME., Helba, B., Marchetti, MA., Dusza, S., Kalloo, A.,
Liopyris, K., Mishra, N., Kittler, H., & Halpern, A., Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC). (2017).

11. Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci. Data* **5**,(2018).

12. Tsumura, N., Haneishi, H. & Miyake, Y. Independent-component analysis of skin color image. *J. Opt. Soc. Am. A* **16**, 2169–2176.(1999).

13. Hyvärinen, A. & Oja, E. Independent component analysis: Algorithms and applications. *Neural Netw.* **13**, 411–430. (2000).

14. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. Medical image computing and computer-assisted intervention— MICCAI 2015. MICCAI 2015. In *Lecture Notes in Computer Science* Vol. 9351 (eds Navab, N. *et al*.) 234–241 (Springer, Berlin, 2015).

15. Taha, A. & Hanbury, A. An efficient algorithm for calculating the exact hausdorff distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **37**(11),(2015).

16. Tan, M. & Le, Q., Efficientnet: Rethinking model scaling for convolutional neural networks, in *ICML*, 6105–6114. (2019).