

Article

Classifying Conditions of Speckle and Wrinkle on the Human Face: A Deep Learning Approach

Tsai-Rong Chang *  and Ming-Yen Tsai

Department of Computer Science and Information Engineering, Southern Taiwan University of Science and Technology, No. 1, Nantai Street, Yongkang District, Tainan 71005, Taiwan

* Correspondence: trchang@stust.edu.tw

Abstract: Speckles and wrinkles are common skin conditions on the face, with occurrence ranging from mild to severe, affecting an individual in various ways. In this study, we aim to detect these conditions using an intelligent deep learning approach. First, we applied a face detection model and identified the face image using face positioning techniques. We then split the face into three polygonal areas (forehead, eyes, and cheeks) based on 81 position points. Skin conditions in the images were firstly judged by skin experts and subjectively classified into different categories, from good to bad. Wrinkles were classified into five categories, and speckles were classified into four categories. Next, data augmentation was performed using the following manipulations: changing the HSV hue, image rotation, and horizontal flipping of the original image, in order to facilitate deep learning using the Resnet models. We tested the training using these models each with a different number of layers: ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152. Finally, the K-fold (K = 10) cross-validation process was applied to obtain more rigorous results. Results of the classification are, in general, satisfactory. When compared across models and across skin features, we found that Resnet performance is generally better in terms of average classification accuracy when its architecture has more layers.

Keywords: image processing; skin condition detection; deep learning; ResNet



Citation: Chang, T.-R.; Tsai, M.-Y. Classifying Conditions of Speckle and Wrinkle on the Human Face: A Deep Learning Approach. *Electronics* **2022**, *11*, 3623. <https://doi.org/10.3390/electronics11213623>

Academic Editor: Chiman Kwan

Received: 30 August 2022

Accepted: 3 November 2022

Published: 6 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Speckles and wrinkles on the face are common skin conditions. They affect an individual's appearance to different degrees. Wrinkles typically increase with age. With skin care [1], the appearance of such skin features may be delayed. Skin conditions, when serious, are typically treated by clinicians after visual assessment, though they are often difficult to diagnose with the naked eye. Here, we aim to implement a diagnostic tool for clinicians to facilitate their professional assessments, focusing on classifying grades of skin conditions.

Skin is the largest organ in the human body [2]. It is divided into three parts: epidermis, dermis, and subcutaneous tissue (Figure 1).

Physicians routinely examine skin abnormality using the non-invasive dermatoscope with optical magnification. Some skin diseases are misdiagnosed, even by professional doctors [3]. To improve subjective judgment, images obtained by the dermatoscope have been subjected to deep learning models [4]. Lesions from other systemic diseases [5–8], and oral diseases have been similarly identified through deep learning [9].

In 2004, Mukaida et al. [10] proposed to extract wrinkles and spots based on local analysis of their shape characteristics. In 2015, Ng et al. [11] proposed the method of Hessian line tracking (HLT), based on the Hessian filter. This method strengthens connectivity of wrinkles, improving the accuracy of locating wrinkles. In 2017, Canak et al. [12] used local binary patterns to extract wrinkle features. In 2017, Zaghbani et al. [13] used the Gabor filter to extract wrinkles for facial emotion studies. These methods are used exclusively to detect wrinkles and spots, but do not distinguish the severity or grading of the skin conditions.

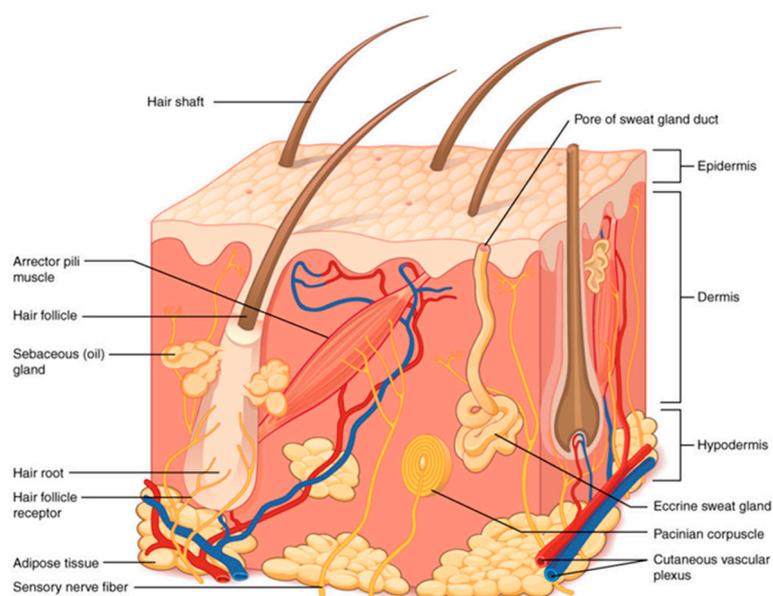


Figure 1. Diagram of skin anatomy [2].

In face detection, several methods have been proposed, including OpenCV's Haar Cascade Classifier, Dlib's Histogram of Gradient (HOG), and Dlib's convolutional neural network (CNN). The object-detection in OpenCV is Cascade Classifier, which is a boosting method serially connecting multiple weak classifiers. The early Cascade Classifier of OpenCV makes use of Haar-like features, later adding Local Binary Pattern (LBP) and Histogram of Gradient (HOG). In 2001, Viola and Jones proposed an algorithm of Viola-Jones' Object detection framework. Face feature-point detection is used to identify common features on the face [14], by comparing dark circles and upper cheeks, brighter area of the nose-bridge between two eyes, and the nose-bridge, and by detecting feature point positions of the eye, nose and mouth.

In recent years, applications of face images have been growing. For example, the COVID-19 epidemic has forced people to wear face masks. Other applications include popular software for special effects such as processing of faces, fast face synthesis sketching [15], and face swapping [16]. In 2017, Rosebrock proposed a 68-point face model for multi-point image capture for face alignment [17], as shown in Figure 2a, based on the Dlib library. Through this algorithm, facial features are first marked with numbers, and then the target block is intercepted according to these points. With this approach, face images can be used for training by deep learning. Subsequently, various applications have been achieved: e.g., face changing [16], head pose estimation [18], expression discrimination [19], and emotion analysis [20]. In 2019, other researchers added 13 to the 68 points to form an 81-point model [21], further improving the performance.

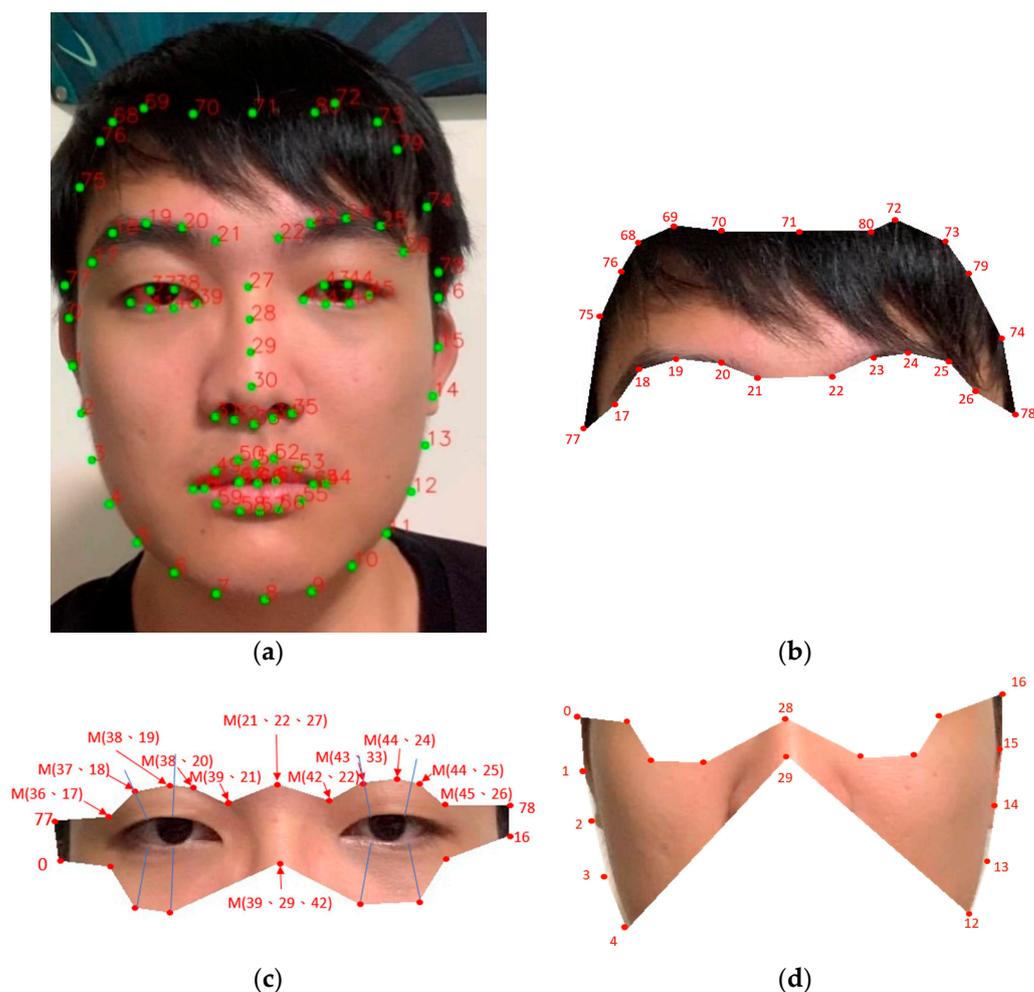


Figure 2. (a) A result plot produced by the 81-point model showing the point-locations in green; (b) Defined forehead region extracted from (a); (c) Defined eyes region; and (d) Defined cheeks region also extracted from (a).

2. Materials and Methods

Deep learning is gradually becoming integrated into everyday life. Through deep learning, computers can learn and may even surpass human performance. In this study, we applied a supervised system for feature detection. For relevant training, the system was given multiple pictures with answer labels. These answer labels were produced by human experts. To generate classified results consistent with the perspectives of both the public and experts, we invited professional beauticians and non-professional people to help label the skin conditions in the images, with slightly more weight given to the professionals.

In order to isolate facial regions, we used Dlib's 81-point model for detecting facial feature points and for their polygonal segmentation. To improve the image processing, some technical issues needed to be resolved regarding image capturing. These issues included sufficient lighting, and camera stability using a tripod.

We first applied a facial positioning technique of 81-points to divide the face into three gross areas (forehead, eyes, and cheeks). We then increased the original number of images through data augmentation. Data augmentation involved the following manipulations: changing the HSV, and changing the rotational angle of the original image to produce new images. These data were then trained via the residual neural network (ResNet). Since there is only one result after training, it is necessary to use K-fold ($K = 10$) cross validation in order to obtain more rigorous results. In K-fold cross validation, 10% of data were used only for testing, not for training. The original data were hence segregated randomly into

ten equal parts, and training was repeated ten times at the ratio of 9:1. For each 1000 epochs of training, we also checked results with K-fold ($K = 10$) cross validation. Whenever the accuracy rate appeared to grow rapidly, training was terminated. This step was to prevent model overfitting.

2.1. Face Detection

Facial landmarks are useful in predicting shapes. In the input image, we judged and assigned areas of interest, based on the shape predictors and the locations of all points of interest along the shape outline. A two-step procedure was followed:

1. Locating the entire face in the image: this step can be achieved in many ways. Three built-in training detectors are available in OpenCV: Haar, LBP, and HOG. Haar feature detection is most popular. Here, we applied the mainstream deep learning algorithm for face positioning. Once the face in the image was detected and the image positioned, we proceeded to the next step.
2. Detecting areas of interest of the face: during the first step, the (x, y) coordinates of the face were obtained, and various facial marker detectors were then applied. Specifically, the following face areas were marked: right eyebrow, left eyebrow, right eye, left eye, nose, mouth, and jaw. By importing the image file in '.dat' format, and using Dlib as the predictor, we were able to display the desired points on the face and the cropped rectangular or polygonal areas from the image.

2.2. Polygon Images

To avoid irrelevant parts affecting training, the image was cut through those 81 points to form polygons each representing areas of forehead, eyes, or cheeks. Figure 2a shows an example of the 81-point model. We further defined intermediate points. If A and B are two of the 81 feature points, the intermediate point $M(A, B)$ is the mid-point between them as represented by:

$$M(A, B) = \text{round}\left(\frac{A_x + B_x}{2}, \frac{A_y + B_y}{2}\right) \quad (1)$$

Sometimes it is necessary to involve three instead of two points according to the following:

$$M(A, B, C) = \text{round}\left(\frac{A_x + B_x + C_x}{3}, \frac{A_y + B_y + C_y}{3}\right) \quad (2)$$

2.3. Residual Neural Network (ResNet)

Residual Neural Network (ResNet) was proposed by He et al. in 2016 [22]. Figure 3 shows the various architectures of the ResNet network including: ResNet-18, ResNet-34, ResNet-50, and ResNet-101. Their difference is in the number of layers. Taking the example of ResNet-152, there is a layer of input at the front, plus the middle building block, $3 + 8 + 36 + 3 = 50$, and then each block is divided into three layers, giving $50 \times 3 = 150$, plus the last classification layer. It therefore has a total of $1 + 150 + 1 = 152$ layers. In addition, the training parameters used in this study are: number of epochs: 1000, batch size: 128, learning rate: 0.001, momentum: 0.9, optimizer: stochastic gradient descent (SGD), loss function: cross entropy.

2.4. Data Augmentation

In the event of insufficient data, deep learning training will likely fail. It is necessary to augment the amount of data without changing the parameters that affect training. For example, in training for skin color, image volume can be increased by rotating the angle of an image (1 to 5° clockwise or anti-clockwise). This approach is not applicable in training for wrinkles, because the training will be affected. For wrinkles or speckles, we increased the data volume by changing the H and S in HSV, as this does not affect training.

Layer name (output size)	ResNet-18	ResNet-34	ResNet-50	ResNet-101	ResNet-152
Conv1 (112 × 112)	7 × 7, 64, stride 2				
	3 × 3 max pool, stride 2				
Conv2_x (56 × 56)	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv3_x (28 × 28)	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
Conv4_x (14 × 14)	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
Conv5_x (7 × 7)	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
FC (1 × 1)	Average pool, 1000-d Full Connected, Softmax				

Figure 3. Details of each layer in the ResNet architecture [22].

Through changing the HSV and rotation angle of the image, and with various combinations, we increased 200 raw images to >10,000 images for training. Specifically, the hue of the image is subjected to $H \pm 0$ to 5 for a total of 11 amplifications. The altered images did not appear very different to the naked eye. Regarding the change to saturation of the image, we applied five amplifications at $S \pm 0$ to 2, finally giving $11 \times 5 - 1 = 54$ permutations and combinations. In addition, there are another two types of amplifications with horizontal flips, and six types of amplifications with three rotational angles (left or right by 5, 10, and 15°). Finally, data augmentation increased the original 200 images to $200 \times (11 \times 5 - 1) \times 2 \times 6 = 129,600$ images.

3. Results

We analyzed 129,600 facial images obtained after data augmentation. Based on the 81-point model, the face is segmented into three regions: forehead, eyes, and cheeks. Images of the whole face and the segmented areas were first viewed by professional beauticians to score wrinkles and speckles in terms of quality grading. After completing a model's training, we used K-fold ($K = 10$) cross validation to show the model's performance (namely, accuracy in classifying skin conditions in terms of gradings). Note that in the K-fold cross validation, for each iteration, we used 90% original data plus their augmented data for training, with the remaining 10% (including original and their augmented data) being used for testing and not for training.

3.1. Dataset

The data used in this study is a dataset obtained by collaborating with a commercial company. There is a total of 200 face images. All subjects in the images provided written informed consent (in Chinese) after a clear explanation of the study protocol. As facial features of individual participants were extracted in fragments and randomly coded, facial data were hence not traceable to individuals. There was no privacy issue involved. At the time of taking these images, there was no lighting problems (too bright or too dim). For subsequent training and pre-treatment, we adopted the influence of elimination. Then, we solved the problem of over-reconciliation caused by insufficient data volume, that is, the number of images was increased to 129,600 through data augmentation as described above.

3.2. Skin Grading

To avoid visual fatigue of viewers, images were rated in batches of 20 images each time (from the 200 original images), resulting in ten rating sessions for each viewer. Finally, these

ratings were averaged across viewers to become the rating of each image. The images were then augmented as described above. In rating skin conditions, wrinkles have five grades: Grade 1 is the best, and Grade 5 is the worst. For speckles, because they are less common than wrinkles, the sample size is smaller, and there are only four grades. Section 3.2.4 below shows an example of the image database.

3.2.1. Forehead Wrinkles

Figure 4 shows that Grade 1 has almost no wrinkles with obvious skin luster. This grade is found with most young people aged between 20 and 30 years old. With good skin care, some 40 years old also have this skin grade. Grade 2 shows a few fine lines. Grade 4 has obvious wrinkles, but these are not as deep and abundant as in Grade 5. Wrinkles of Grades 2 and 3 are age-related and are found in people aged between 40 to 60 years old. Grades 4 and 5 are common in the elderly.



Figure 4. Examples of the five grades of forehead wrinkles (regions are similar to Figure 2b).

3.2.2. Eye Wrinkles

Figure 5 shows that Grade 1 has no wrinkles around the eyes. Grade 2 has some faint wrinkles at the lower eyelid. Grade 3 has more wrinkles. In Grade 4 the wrinkles appear more obviously, and appear even on the nose. In Grade 5, wrinkles appear deeper. Grades 3 to 5 are age-related, and are found in subjects >40 years of age.

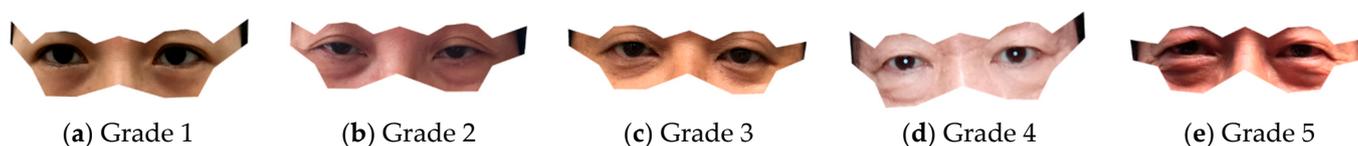


Figure 5. Examples of the five grades of eye wrinkles (regions are similar to Figure 2c).

3.2.3. Cheek Wrinkles

Although there are some red speckles on cheeks, this is also a classification of wrinkles. Thus, only wrinkles are classified as shown in Figure 6.

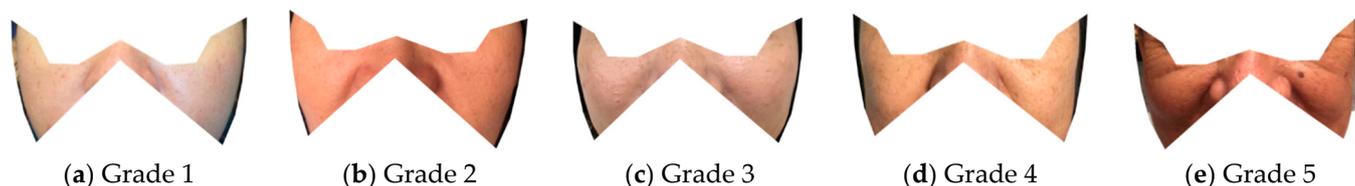


Figure 6. Examples of the five grades of cheek wrinkles (regions are similar to Figure 2d).

3.2.4. Cheek Speckles

Regarding speckles, only four categories are classified. Grade 1 has very few speckles. Grade 2 has one or two speckles. Grade 3 has a small number of speckles. Grade 4 has even more speckles, commonly in middle-age or older subjects (Figure 7).

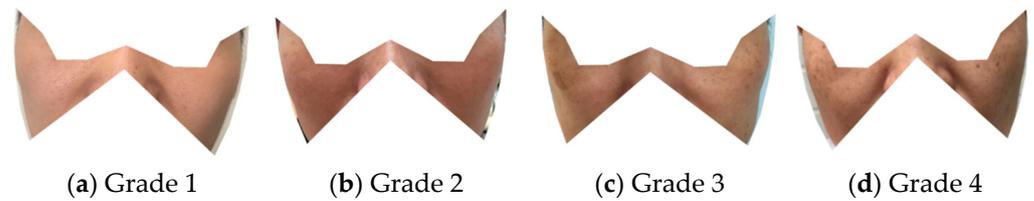


Figure 7. Examples of the four grades of cheek speckles (regions are similar to Figure 2d).

3.3. Training Results

3.3.1. Training Results with Polygon Images

Figure 8 shows an example of Grade 1 forehead wrinkles, and the training results using various models. This type of image was originally classified as Grade 1. Through training with ResNet-18, ResNet-34 and ResNet-50, it was classified as Grade 2, while with ResNet-101 and ResNet-152 it is correctly classified as Grade 1. Here, ResNet-152 shows the biggest improvement in accuracy. The poorer performances by models with fewer layers are likely due to presence of hair bangs in the images.

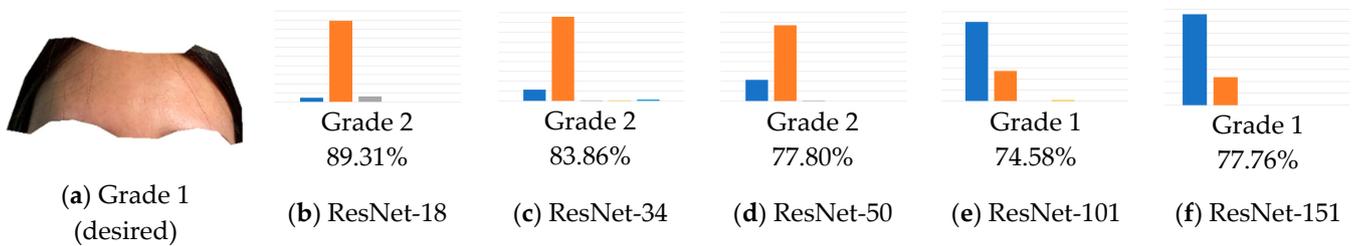


Figure 8. Example and training results of Grade 1 forehead wrinkles (wrongly classified grades are underlined).

Figure 9 similarly shows Grade 2 eye wrinkles. In ResNet, several architecture models show the correct classifications rather satisfactorily.

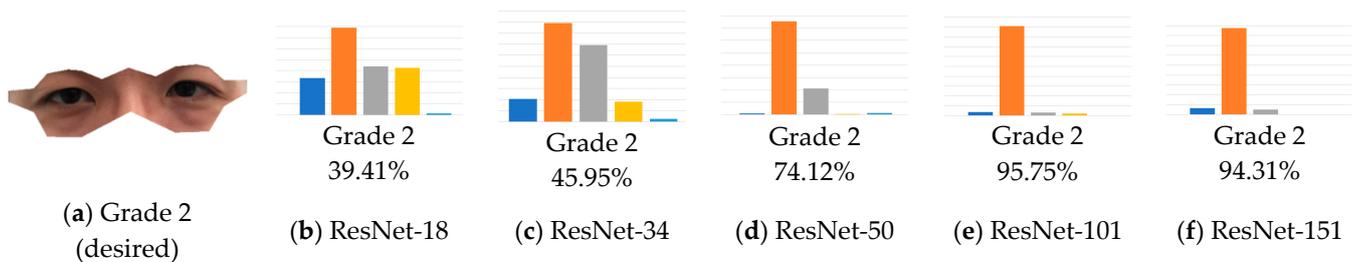


Figure 9. Example and training results of Grade 2 eye wrinkles.

Figure 10 shows Grade 3 cheek wrinkles. This type of images is not classified correctly even with ResNet-152. Shadows on both sides of the nose likely cause such poor performances.

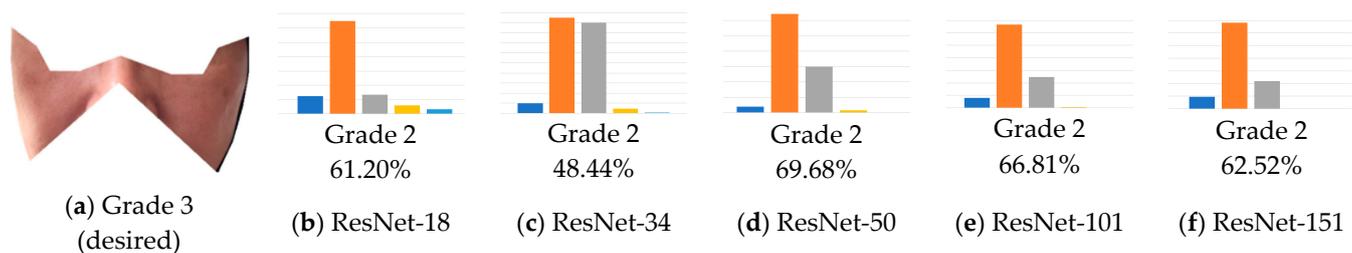


Figure 10. Example and training results of Grade 3 cheek wrinkles (wrongly classified grades are underlined).

Figure 11 shows Grade 4 cheek speckles. Speckles on the face appear more clearly in the images compared with other skin features. Model classification accuracy is not so good using models with fewer layers. With more layers, model accuracy greatly improves.

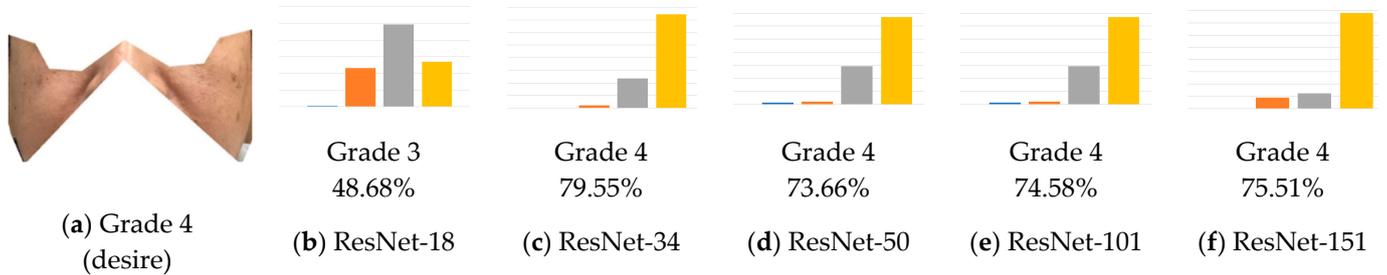


Figure 11. Example and training results of the Grade 4 cheek speckles (wrongly classified grades are underlined).

Figure 12 shows Grade 5 forehead wrinkles. Despite the clarity of these wrinkles, ResNet-18 gives only a fair accuracy in classification. With ResNet-34 or more layers, classification accuracy improves steadily.

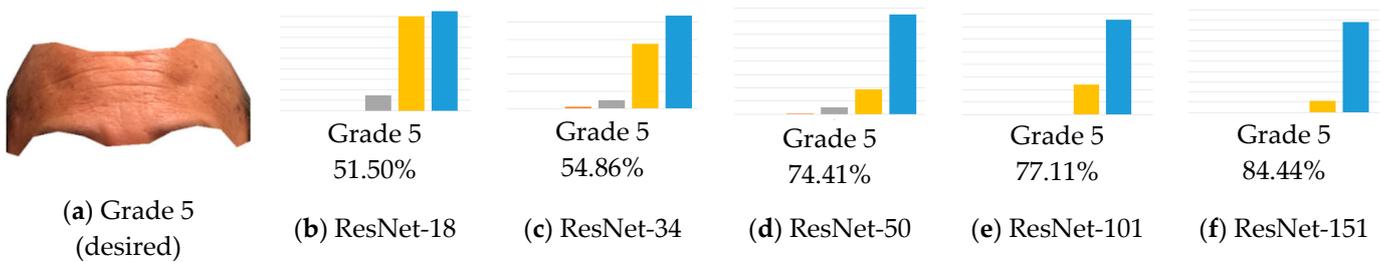


Figure 12. Example and training results of Grade 5 forehead wrinkles.

In brief, as the number of model layers increases progressively, the accuracy increases in parallel. As expected, model performance is closely related to its number of layers (Figure 13).

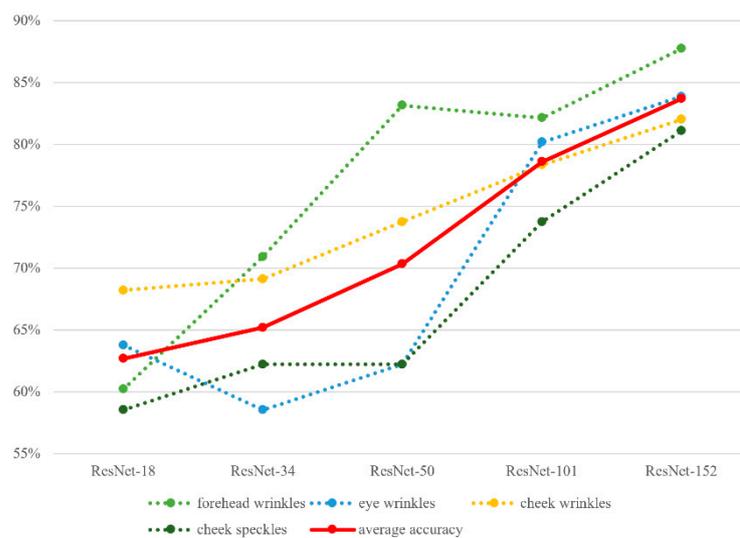


Figure 13. Model accuracy results (from K-fold cross validation) plotted against model layers and facial features.

3.3.2. Results of Model Classification Using Polygon Images Presented in Terms of Precision and Recall

Figure 14 shows results of precision, and Figure 15 shows results of recall. The overall pictures are basically similar to those of accuracy (Figure 13), despite some minor disparities. These disparities are likely related to characteristics of the dataset, such as sample size, hair or shadow noise, and relative abundance of skin features across facial regions. Again, as the number of model layers increases, the precision or recall rate increases.

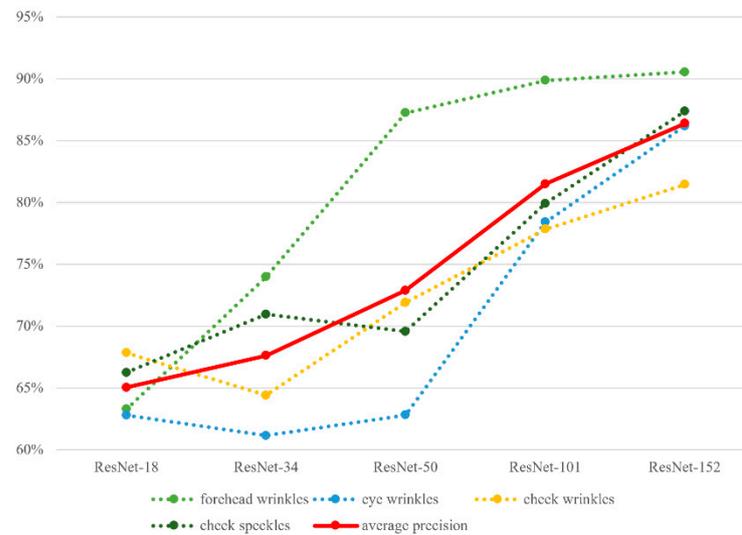


Figure 14. Results of classification precision (from K-fold cross validation) plotted against model layers and facial features.

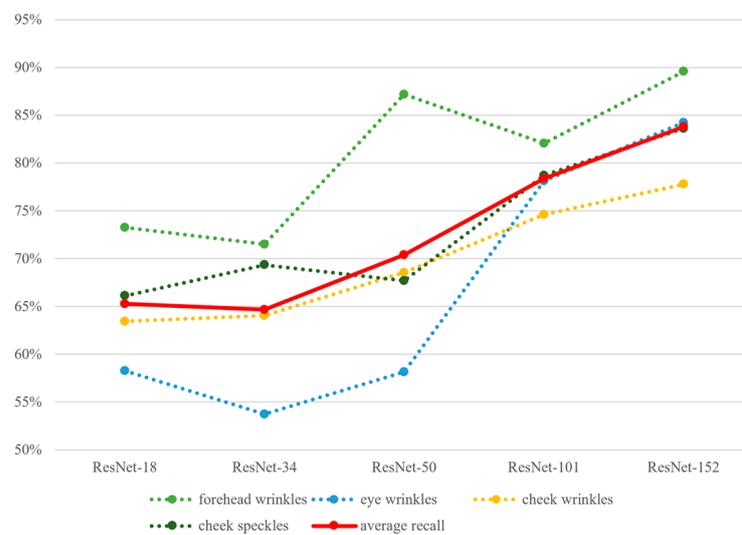


Figure 15. Results of classification recall (from K-fold cross validation) plotted against model layers and facial features.

3.3.3. Training Results Using Whole Face Images

Because the whole face has a larger picture size, training takes more time (at least twice as long) when compared with training using the smaller polygon images. Figure 16 clearly shows that the average accuracy using polygon images is much higher than that using whole face images. Such discrepancy is likely due to interference of the non-feature regions in the whole face images.

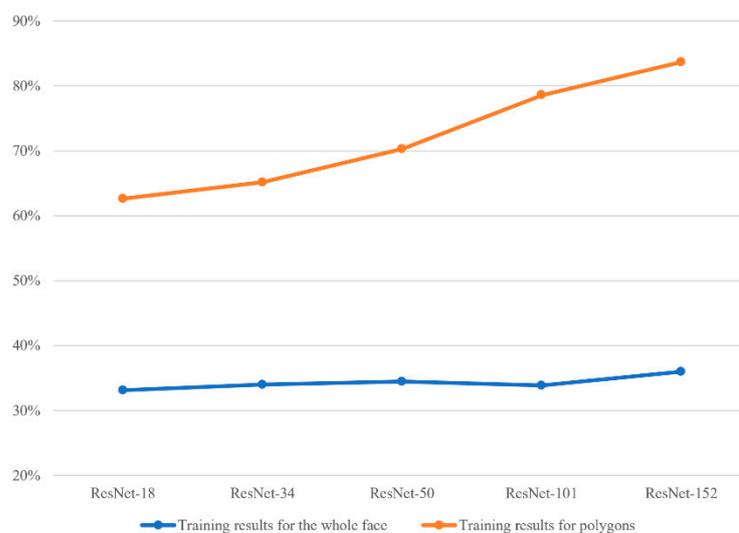


Figure 16. Average accuracy results (from K-fold cross validation), comparing training with the whole face and with polygon images.

3.3.4. Extended Results Using Other Models

After our study with Resnet, other models like AlexNet, and VGG were also tested. Figure 17 confirms that training results using Resnet152 are still the best. These other models give high correct rates comparable to those from Resnet50.

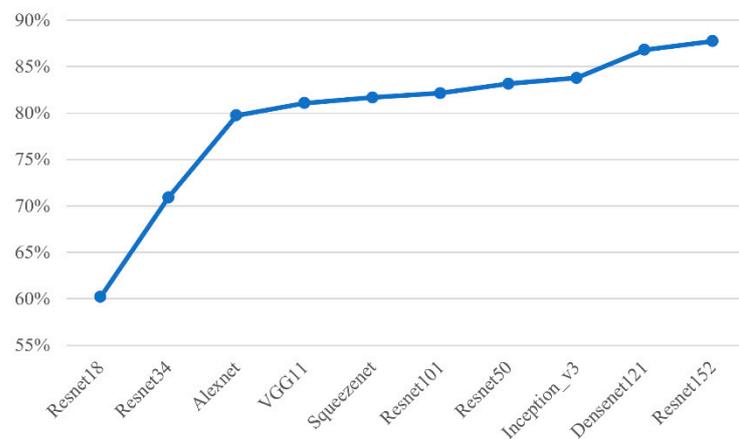


Figure 17. Accuracy results (from K-fold cross validation) on forehead wrinkles plotted across several models.

3.3.5. Limitations

Two main limitations of the study are as follows: (a) the data augmentation step undoubtedly introduced strong bias. If we had larger datasets (e.g., 20,000 original images) our conclusions would definitely be more convincing; (b) we have not compared other models exhaustively. For example, after segmenting the face into various regions, we did not continue with the face detection model for wrinkle and speckle detection and grading classification. The reason for this is technical, as we needed to manually mark and grade the features on the images, a process that is labor-intensive and has questionable accuracy. We therefore do not rule out the possibility of a better performance by other models that we have not fully tested.

4. Conclusions

This is the first deep learning modeling study on classifying grades of two common skin conditions (wrinkles and speckles) using polygon images. The overall test results

from classifying such conditions using Resnet models are satisfactory. Deep learning using polygon images produces better results than those using whole face images. A greater number of layers in ResNet produces better performance. ResNet-152 so far shows the best results compared with the other models we tested.

Author Contributions: T.-R.C.: Problem conceptualization, Methodology, Data analysis, Writing-Review and edit final draft, Results tabulate and graphic presentation. M.-Y.T.: Data collection, Software development and execution, Investigation. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported in part by the A²IBRC, STUST from the Higher Education Sprout Project of the Ministry of Education, Taiwan, and MOST 109-2221-E-218 -022—from the Ministry of Science and Technology, Taiwan.

Informed Consent Statement: Informed consent was obtained from all participants involved in the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Asakura, K.; Nishiwaki, Y.; Milojevic, A.; Michikawa, T.; Kikuchi, Y.; Nakano, M.; Iwasawa, S.; Hillebrand, G.; Miyamoto, K.; Ono, M.; et al. Lifestyle factors and visible skin aging in a population of Japanese elders. *J. Epidemiol.* **2009**, *19*, 251–259. [CrossRef] [PubMed]
2. 5.1 Layers of the Skin—Anatomy and Physiology 2e | OpenStax. Available online: <https://openstax.org/books/anatomy-and-physiology-2e/pages/5-1-layers-of-the-skin> (accessed on 2 November 2022).
3. Riaz, F.; Naeem, S.; Nawaz, R.; Coimbra, M. Active contours based segmentation and lesion periphery analysis for characterization of skin lesions in dermoscopy images. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 489–500. [CrossRef] [PubMed]
4. Albert, B.A. Deep learning from limited training data: Novel segmentation and ensemble algorithms applied to automatic melanoma diagnosis. *IEEE Access* **2020**, *8*, 31254–31269. [CrossRef]
5. Liao, H. A Deep Learning Approach to Universal Skin Disease Classification. University of Rochester Department of Computer Science, CSC, 2016. Available online: <https://studylib.net/doc/14025144/a-deep-learning-approach-to-universal-skin-disease-classi> (accessed on 2 November 2022).
6. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017**, *542*, 115–118. [CrossRef] [PubMed]
7. Lopez, A.R.; Giro-i-Nieto, X.; Burdick, J.; Marques, O. Skin lesion classification from dermoscopic images using deep learning techniques. In Proceedings of the 2017 13th IASTED International Conference on Biomedical Engineering (BioMed), Innsbruck, Austria, 20–21 February 2017; pp. 49–54.
8. Huang, H.; Kharazmi, P.; McLean, D.I.; Lui, H.; Wang, Z.J.; Lee, T.K. Automatic Detection of Translucency Using a Deep Learning Method from Patches of Clinical Basal Cell Carcinoma Images. In Proceedings of the 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Honolulu, HI, USA, 12–15 November 2018; pp. 685–688.
9. Anantharaman, R.; Velazquez, M.; Lee, Y. Utilizing Mask R-CNN for Detection and Segmentation of Oral Diseases. In Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Madrid, Spain, 6 December 2018; pp. 2197–2204.
10. Mukaida, S.; Ando, H. Extraction and Manipulation of Wrinkles and Spots for Facial Image Synthesis. In Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, Korea, 17–19 May 2004; pp. 749–754.
11. Ng, C.; Yap, M.H.; Costen, N.; Li, B. Wrinkle detection using hessian line tracking. *IEEE Access* **2015**, *3*, 1079–1088. [CrossRef]
12. Çanak, B.; Kamaşak, M.E. Automatic Scoring of Wrinkles on the Forehead. In Proceedings of the 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, Turkey, 15–18 May 2017; pp. 1–4.
13. Zaghbani, S.; Boujneh, N.; Bouhlel, M.S. Facial Emotion Recognition for Adaptive Interfaces Using Wrinkles and Wavelet Network. In Proceedings of the 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), Hammamet, Tunisia, 30 October–3 November 2017; pp. 342–349.
14. OpenCV: Cascade Classifier Training. Available online: https://docs.opencv.org/3.4/dc/d88/tutorial_traincascade.html (accessed on 25 August 2021).
15. Peng, C.; Gao, X.; Wang, N.; Li, J. Superpixel-based face sketch-photo synthesis. *IEEE Trans. Circuits Syst. Video Technol.* **2017**, *27*, 288–299. [CrossRef]
16. Nirkin, Y.; Masi, I.; Tuan, A.T.; Hassner, T.; Medioni, G. On Face Segmentation, Face Swapping, and Face Perception. In Proceedings of the 2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018), Xi'an, China, 15–19 May 2018; pp. 98–105.
17. Facial Landmarks with Dlib, OpenCV, and Python—PyImageSearch. Available online: <https://pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/> (accessed on 2 November 2022).

18. Khan, K.; Mauro, M.; Migliorati, P.; Leonardi, R. Head Pose Estimation through Multi-Class Face Segmentation. In Proceedings of the 2017 IEEE International Conference on Multimedia and Expo (ICME), Hong Kong, China, 10–14 July 2017; pp. 175–180.
19. Benini, S.; Khan, K.; Leonardi, R.; Mauro, M.; Migliorati, P. Face analysis through semantic face segmentation. *Signal Process. Image Commun.* **2019**, *74*, 21–31. [[CrossRef](#)]
20. Hassan, S.Z.; Ahmad, K.; Al-Fuqaha, A.; Conci, N. Sentiment Analysis from Images of Natural Disasters. In *Image Analysis and Processing—ICIAP 2019*; Ricci, E., Rota Bulò, S., Snoek, C., Lanz, O., Messelodi, S., Sebe, N., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 104–113.
21. Niko 81 Facial Landmarks Shape Predictor | Github. Available online: https://github.com/codeniko/shape_predictor_81_face_landmarks (accessed on 2 November 2022).
22. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. *arXiv* **2015**, arXiv:1512.03385.