Heliyon 9 (2023) e20186

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

# Image classification and auxiliary diagnosis system for hyperpigmented skin diseases based on deep learning

Jianyun Lu [a,1], Xiaoliang Tong [a,1], Hongping Wu [b], Yaoxinchuan Liu [b], Huidan Ouyang [b,**], Qinghai Zeng [a,*]

[a] Department of Dermatology, Third Xiangya Hospital, Central South University, Changsha 410013, PR China
[b] Vocational Teachers College, Jiangxi Agricultural University, NanChang 330045, PR China

ARTICLE INFO

*Keywords:*
Hyperpigmented skin disease
Deep learning
Image classification
Diagnostic system

ABSTRACT

*Background and aim:* Melasma (ML), naevus fusco-caeruleus zygomaticus (NZ), freckles (FC), cafe-au-lait spots (CS), nevus of ota (NO), and lentigo simplex (LS), are common skin diseases causing hyperpigmentation. Deep learning algorithms learn the inherent laws and representation levels of sample data and can analyze the internal details of the image and classify it objectively to be used for image diagnosis. However, deep learning algorithms that can assist clinicians in diagnosing skin hyperpigmentation conditions are lacking.

*Methods:* The optimal deep-learning image recognition algorithm was explored for the auxiliary diagnosis of hyperpigmented skin disease. Pretrained models, such as VGG-19, GoogLeNet, InceptionV3, ResNet50V2, ResNet101V2, ResNet152V2, InceptionResNetV2, DesseNet201, MobileNet, and NASNetMobile were used to classify images of six common hyperpigmented skin diseases. The best deep learning algorithm for developing an online clinical diagnosis system was selected by using accuracy and area under curve (AUC) as evaluation indicators.

*Results:* In this research, the parameters of the above-mentioned ten deep learning algorithms were 18333510, 5979702, 21815078, 23577094, 42638854, 58343942, 54345958, 18333510, 3235014, and 4276058, respectively, and their training time was 380, 162, 199, 188, 315, 511, 471, 697, 101, and 144 min respectively. The respective accuracies of the training set were 85.94%, 99.72%, 99.61%, 99.52%, 99.52%, 98.84%, 99.61%, 99.13%, 99.52%, and 99.61%. The accuracy rates of the test set data were 73.28%, 57.40%, 70.04%, 71.48%, 68.23%, 71.11%, 71.84%, 73.28%, 70.39%, and 43.68%, respectively. Finally, the areas of AUC curves were 0.93, 0.86, 0.93, 0.91, 0.91, 0.92, 0.93, 0.92, 0.93, and 0.82, respectively.

*Conclusions:* The experimental parameters, training time, accuracy, and AUC of the above models suggest that MobileNet provides a good clinical application prospect in the auxiliary diagnosis of hyperpigmented skin.

* Corresponding author.
** Corresponding author.
*E-mail addresses:* ouyanghuidan@jxau.edu.cn (H. Ouyang), zengqinghai@csu.edu.cn (Q. Zeng).
[1] These authors contributed equally.

https://doi.org/10.1016/j.heliyon.2023.e20186
Received 10 December 2022; Received in revised form 11 September 2023; Accepted 13 September 2023
Available online 16 September 2023

## 1. Introduction

Melasma (ML) [1], naevus fusco-caeruleus zygomaticus (NZ) [2], freckles (FC) [3], cafe-au-lait spots (CS) [4], nevus of ota (NO) [5], and lentigo simplex (LS) [6] are common hyperpigmented skin diseases on the face that seriously affect the aesthetics and social activities of patients [7]. Medical doctors diagnose a disease mainly by integrating clinical symptoms, medical images, and laboratory test results [8]. However, as a result of incomplete extraction of critical information from medical data and the uncertainty of objective factors, doctors may misjudge a disease, thereby affecting the treatment of patients.

In the era of artificial intelligence and medical advancements, scholars are using cutting-edge deep learning algorithms to intelligently classify skin diseases through feature extraction, analysis, and training [9–13]. In 2017, Andre Esteva et al. used deep learning algorithms to classify skin diseases, achieving an accuracy of about 55.4% in nine classifications and 72.1% in three classifications, beyond the accuracy levels achieved by dermatologists [14]. Moreover, Philipp Tschandl used deep learning algorithms to classify different skin lesions, demonstrating once again the excellent performance of deep learning algorithms compared to expert doctors at diagnosing skin diseases [15]. Jianpeng Zhang et al. proposed a collaborative deep learning algorithm to classify medical images, and their algorithm innovation has dramatically improved the accuracy. The datasets, however, are limited to ImageCLEE and ISIC, other types of skin diseases have not been classified and explored, with the types of datasets being relatively few [16]. Yuan Liu et al. proposed a deep learning system to distinguish 26 skin diseases and assist in auxiliary diagnosis by the doctors [17]. Qianjin Lu et al., developed a mobile lupus erythematosus classification assistance platform based on a deep learning algorithm, enabling clinical dermatologists to diagnose patients online [18].

While several deep learning-based approaches have been applied for automated diagnosis of skin diseases, none has used the network architectures (ResNet50V2, InceptionResNetV2, VGG-19, DenseNet201, GoogLeNet, InceptionV3, MobileNet, NASNetMobile, ResNet101V2, and ResNet152V2) to classify the six hyperpigmented skin diseases (ML, NZ, FC, CS, NO, and LS) that have been handled in this work. We have thus explored the optimal deep network architecture for the auxiliary diagnosis of the diseases.

## 2. Methods

### 2.1. Collection of skin lesion pictures

Pictures of six common hyperpigmented skin diseases from the Department of Dermatology at the Third Xiangya Hospital of Central South University. Each disease was diagnosed by two specialized dermatologists. In some controversial cases, reflectance confocal microscopy (RCM) was used as an auxiliary diagnosis. Below are the characteristics of each skin disorder's lesions: In ML, brown or dark brown lesions occur symmetrically in the zygomatic and temporal regions, and are irregularly distributed or butterfly-shaped. Generally, symmetrically distributed, NZ affects the bilateral zygomatic, temporal, and lower eyelids. NZ is usually round or oval, gray-brown, or dark-gray, isolated macules that are not confluent. FC typically appears as multiple rice-grain-sized yellow-brown or brown, isolated, and unfused spots. CS are well-defined, uniformly pigmented, light brown macules of varying sizes, being either single or multiple. NO forms bluish-black or bluish-gray patches of uniform color; it is commonly found on the temporal, zygomatic, upper, and lower eyelids and cheeks. LS are mostly unilaterally distributed as flakes, ranging in size from the tip of a needle to that of a grain of rice, with uniform color. This study was approved by the Ethics Committee of the Third Xiangya Hospital of Central South University (NO.2022-S303). Patient information was kept strictly confidential.

### 2.2. Sample selection

In this study, 1366 images of skin lesions were included, with 233, 255, 346, 212, 214, and 112, respectively, for ML, NZ, FC, CS, NO, and LS. Data from dermatological images are categorized into training and test sets in a 4:1 ratio (Table 1).

### 2.3. Deep learning algorithms

The deep learning algorithm [19] model originates from the human neural network, comprising multiple units connected to each other to form a network. Each layer of the network is equivalent to the human cerebral cortex; the more cortex content, the deeper the network. The depth generally refers to more than ten layers. Since the introduction of neural network algorithms, there have been more and more deep learning algorithms, attributed to the number of nodes in each layer, the depth of each layer, and different design

**Table 1**
Distribution of training set and the test set.

| | Train Dataset num | Test Dataset num | Sum |
|---|---|---|---|
| CS | 169 | 43 | 212 |
| FC | 276 | 70 | 346 |
| LS | 89 | 23 | 112 |
| ML | 186 | 47 | 233 |
| NO | 171 | 43 | 214 |
| NZ | 204 | 51 | 255 |
| Sum | 1089 | 277 | 1366 |

of the layers, and so on. In the optimal situation for pretrained model training, the validation loss no longer improves or achieves the expected accuracy rate, and the training can be considered to converge.

In this study, 10 pretrained models were used, including VGG-19, GoogLeNet, InceptionV3, ResNet50V2, ResNet101V2, ResNet152V2, InceptionResNetV2, DesseNet201, MobileNet, and NASNetMobile. The structure diagram of the 10 deep learning networks is presented in Fig. 1.

Considering the VGG-19 model architecture [20], VGG-19 consists of 19 sequentially stacked layers, including a convolutional layer, maximum pooling layer, and fully connected layer. VGG-19 has a simple and unified architecture, which can be easily understood and implemented. The model performed well in various image classification tasks. The VGG-19 has many parameters and high computational and memory consumption, which may not be applicable to environments with limited resources.

GoogLeNet model architecture [21] contains multiple Inception modules, with a total of 22 layers. GoogLeNet uses the concept of Inception modules and overlays them with the maximum pooling layer and the average pooling layer to achieve high accuracy while reducing computational complexity. It can efficiently extract features at different scales. The main drawback of GoogLeNet is its complexity, making it difficult to understand and implement.

The InceptionV3 model architecture [22] consists of approximately 48 layers, including convolutional layers, Inception modules, and fully connected layers. InceptionV3, like GoogLeNet, overlays multiple Inception modules with pooling and fully connected layers. InceptionV3 further introduces techniques such as decomposition and dimensionality reduction according to GoogLeNet. It performs well in various image classification tasks. Nonetheless, Inception V3 is a relatively complex architecture, and may require high computational resources for its training.

The ResNet50v2, ResNet101, and ResNet152v2 models architecture [23] include 50 layers, 101 layers, and 152 layers, respectively. These layers include convolutional layers, residual blocks, and fully connected layers. ResNet uses residual blocks, and the output of one block is added to the input of subsequent blocks. The ResNet model solves the vanishing gradient problem, making training easier and building a very deep network. The model could achieve state-of-the-art performance in image classification tasks. Compared to other architectures, the ResNet model is deeper, having more parameters, slower training speed, and requires more computing resources.

The InceptionResNetv2 model architecture [24] consists of approximately 572 layers and the stacks Inception modules and uses residual connections within these modules. InceptionResNetV2 further improves on GoogLeNet and ResNet, while using residual connections and has excellent accuracy in various computer vision tasks. InceptionResNetV2 is a relatively complex architecture and requires high computational resource during the training process.

The DenseNet201 model architecture [25] has a total of 201 layers, that includes convolutional layers, densely connected blocks, and fully connected layers. DesseNet201 emphasizes dense connections, and each layer is directly connected to all subsequent layers. The dense connection of DesseNet201 helps better reuse of features and reduction in the number of parameters. It performs well in tasks such as image classification, has higher computational complexity, and may require longer training time compared to other models.

The MobileNet model architecture [26] convolutional neural network (CNN) is designed for mobile and embedded devices, with different variants and usually comprises dozens of layers. MobileNet uses deep separable convolutions to reduce computational complexity, simultaneously maintaining high accuracy. The model typically includes convolutional layers, deep separable convolutional layers, and fully connected layers. The MobileNet architecture is efficient and suitable for limited computing resource environments. It is widely used in mobile applications. Compared to other models, MobileNet may sacrifice some accuracy. For some

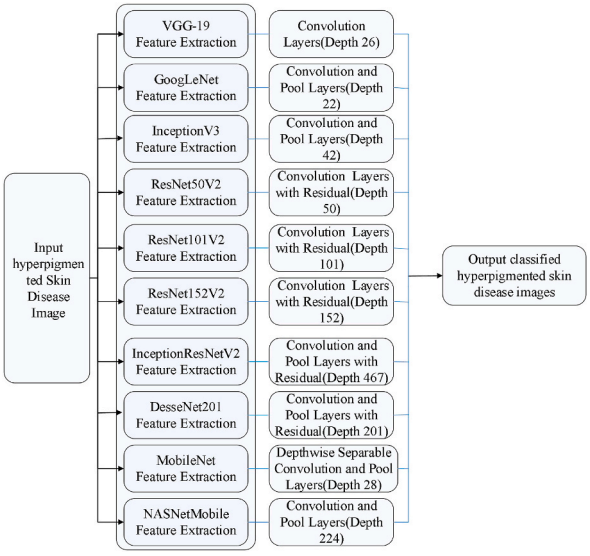


**Fig. 1.** Ten pretrained models.

complex tasks, performance may not be as good as other deeper models.

The NASNetMobile (Neural Architecture Search Network) model architecture [27] is a series of CNN designed using neural architecture search. NASNetMobile is a variant optimized for mobile devices, and has dozens of layers. NASNetMobile uses Reinforcement learning and other technologies to automatically discover the network architecture, which comprises multiple modules and connection modes. The design of NASNetMobile has been optimized to achieve competitive performance on mobile devices. It achieves efficient network architecture through automatic search. Compared to traditional predefined architectures, the network search and training of NASNetMobile may require more computing resources and time.

### 2.4. Statistical analysis

#### 2.4.1. Normalization

The images do not usually have homogenous intensities and affect the performance of the automated methods. Several normalization algorithms have been implemented by using different types of images for high performance [28,29]. However, they may lead to high computational costs. Therefore, in the proposed approach, an efficient algorithm has been provided for normalization. The normalization and standardization formulas are as follows:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{1}$$

$$x'' = \frac{x' - \bar{x}}{\sigma} \tag{2}$$

In formulas 1 and 2, where $X$, $\min(X)$, $\max(X)$, and $X'$, respectively, represent the characteristic value, the minimum value under the characteristic, the maximum value under the characteristic, and the normalized value [19]. They represent the mean value, the variance, and the standardized result under the feature.

#### 2.4.2. Accuracy

Accuracy is one of the evaluation indicators for deep learning algorithm models. Accuracy is the degree to which the predicted value matches the true value under certain experimental conditions, and is expressed in error [30]. A confusion matrix, also known as an error matrix, is a standard format to evaluate the accuracy and is expressed as a matrix with N rows and N columns. Accuracy= (TP + TN)/(TP + FN + FP + TN).

#### 2.4.3. Confusion matrix

The Confusion matrix is a graphical way to intuitively display the quality of the deep learning classification model. A confusion matrix, also known as a possibility table or error matrix in artificial intelligence and machine learning [31]. The confusion matrix is used to visually present the results of algorithm performance, especially in supervised learning. In unsupervised learning, matching matrices are generally used. In a confusion matrix, the accuracy of classification results can be displayed as the differences between classification results and actual values. The confusion matrix is calculated by classification of each measured pixel and comparing the position with the corresponding status and classification in the classified image.

#### 2.4.4. Sensitivity specificity curve

An AUC is a performance metric used to measure the quality of deep learning classification model. An area under the ROC curve (AUC) is a performance metric used to measure the quality of learners [32]. As ROC curves cannot clearly explain which classifier is better, AUC value is often used as the evaluation standard. The classifier with a larger AUC value is better. When AUC = 1, it represents a perfect classifier; when 0.5 < AUC <1, it is better than the random classifier; when 0 < AUC <0.5, it is worse than the random classifier. The abscissa of a ROC curve represents specificity (sensitivity = TP/(TP + FN)) and its ordinate represents sensitivity (specificity = TN/(TN + FP)).

Among them, true negative (TN), known as true negative rate, indicates the number of negative samples. The false positive rate (FP), or false positive rate, is the number of negative samples that are expected to be positive. False negative (FN), also called the false negative rate, shows the number of positive samples that are predicted to be negative samples. True positive (TP), or true positive rate, measures the number of positive samples that are expected to be positive.

#### 2.4.5. Experimental environment

Experiments were conducted using a supercomputer to build a deep learning network. The configurations of CPU, memory, gpu0, gpu1, and gpu2 of the first notebook computer were 2.90ghz, 160.0gb, 92.0gb, 92.0gb, and 92.0gb, respectively. Finally, the ten network models VGG-19, GoogLeNet, inception V3, ResNet50v2, ResNet101v2, ResNet152v2, inception ResNetv2, dessenet201, MobileNet, and NASNetMobile were trained on the supercomputer. The trained models were then tested on the test set. Lastly, the confusion matrix, AUC curve, and accuracy evaluation index were visualized.

## 3. Result

### 3.1. The area under the ROC curve results

AUC area and training set accuracy for 10 different pretrained models are shown in Table 2. After 50 iterations of VGG-19, GoogLeNet, inception V3, ResNet50v2, ResNet101v2, ResNet152v2, inception ResNetv2, dessenet201, MobileNet, and NASNetMobile algorithms, the respective parameters obtained were 18333510, 5979702, 21815078, 23577094, 42638854, 58343942, 54345958, 18333510, 3235014, and 4276058. The training time was 380, 162, 199, 188, 315, 511, 471, 697, 101, and 144 min, respectively and the respective accuracy rates of the training set were 85.94%, 99.72%, 99.61%, 99.52%, 99.52%, 98.84%, 99.61%, 99.13%, 99.52%, and 99.61%. The accuracy rates of the test set data were 73.28%, 57.40%, 70.04%, 71.48%, 68.23%, 71.11%, 71.84%, 73.28%, 70.39%, and 43.68%, respectively. The respective areas of AUC curves were 0.93, 0.86, 0.93, 0.91, 0.91, 0.92, 0.93, 0.92, 0.93, and 0.82. The optimal convolution algorithm MobileNet was then analyzed in detail employing four indicators: the number of model parameters, training time, the accuracy of test set data, and the area of the AUC curve.

As shown in Fig. 2(a–j), a comparison of the AUC curves of the above ten classifiers revealed that the MobileNet classifier was the best performer. According to Fig. 2 (i), the MobileNet algorithm model performed the best on CS, followed by NO, FC, NZ, and LS, and the worst on ML. Most classification algorithms have shown good classification results for CS skin pigmentation disease. Experimental results indicate better performance of the MobileNet classical CNN algorithm used in this paper than the InceptionV3 algorithm used in the study of the team in Stanford, published in Nature in 2017. In conclusion, we observed that the MobileNet classical CNN algorithm exhibited higher accuracy and more vital generalization ability in classifying objects in this study.

### 3.2. The results of confusion matrices

In Fig. 3(a–j) are the confusion matrices of 6 classification tasks of 10 classical CNN, respectively. The figure shows that the elements (i and j) of each confusion matrix represent the probability of predicting class j when it is actually class i. Where, 0,1,2,3,4 and 5 represent CS, FC, LS, ML, NO, and NZ, respectively. The MobileNet algorithm for the recognition of pigmentary dermatosis was better selected only by the confusion matrix. VGG-19, GoogLeNet, InceptionV3, ResNet50V2, ResNet101V2, ResNet152V2, InceptionResNetV2, DesseNet201, MobileNet, and NASNetMobi were also compared using AUC in this paper. Ten classical CNN were analyzed in six common pigmented skin diseases as follows.

### 3.3. The classification system of hyperpigmented skin diseases

A deep learning algorithm is used to classify skin and venereal diseases by analyzing images of common skin diseases, whose data are entered into the system. The input images are cropped into 224x224 images after data preprocessing, and then the by the MobileNet algorithm was used for comprehensive prediction. Model parameters were stored in the Model layer of the MobileNet algorithm. Fig. 4 shows the classification result of medical image data of the patient. After adding 2%, 5%, 10%, 20%, 30%, and 50% noise to the images of hyperpigmented skin diseases in the test set, their sensitivity, specificity, and accuracy slightly decreased, but the change was not significant (Supplementary Fig. 1). Low-definition photos have little impact on the model, and the MobileNet algorithm is highly robust.

Fig. 5 shows the auxiliary diagnosis system of hyperpigmented skin diseases. First, the medical image images of six pigmented skin diseases were classified and sorted. After comparing ten deep learning algorithms, a model trained by the optimal MobileNet algorithm was deployed to the server. The patient uploaded a picture of their skin lesions by the client application (a mobile app, or a browser, or a desktop computer application). The data was fed into an algorithm using the MobileNet network model to predict the pigmentodermatosis category of the patient. The results were reviewed by dermatologist, and then forwarded patients with the diagnosis to the client application, to realize a computer-aided diagnosis and pigment dermatologist, thus once again significantly improving the efficiency of diagnosis.

**Table 2**
The accuracy table of each convolution neural network.

| | Train Accuracy(%) | Test Accuracy(%) | AUC | Params | Train Time(min) |
|---|---|---|---|---|---|
| VGG19 | 85.94 | 73.28 | 0.93 | 18333510 | 380 |
| GoogleNet | 99.72 | 57.40 | 0.86 | 5979702 | 162 |
| InceptionV3 | 99.61 | 70.04 | 0.93 | 21815078 | 199 |
| ResNet50V2 | 99.52 | 71.48 | 0.91 | 23577094 | 188 |
| ResNet101V2 | 99.52 | 68.23 | 0.91 | 42638854 | 315 |
| ResNet152V2 | 98.84 | 71.11 | 0.92 | 58343942 | 511 |
| InceptionResNetV2 | 99.61 | 71.84 | 0.93 | 54345958 | 471 |
| DenseNet201 | 99.13 | 73.28 | 0.92 | 18333510 | 697 |
| MobileNet | 99.52 | 70.39 | 0.93 | 3235014 | 101 |
| NASNetMobile | 99.61 | 43.68 | 0.82 | 4276058 | 144 |

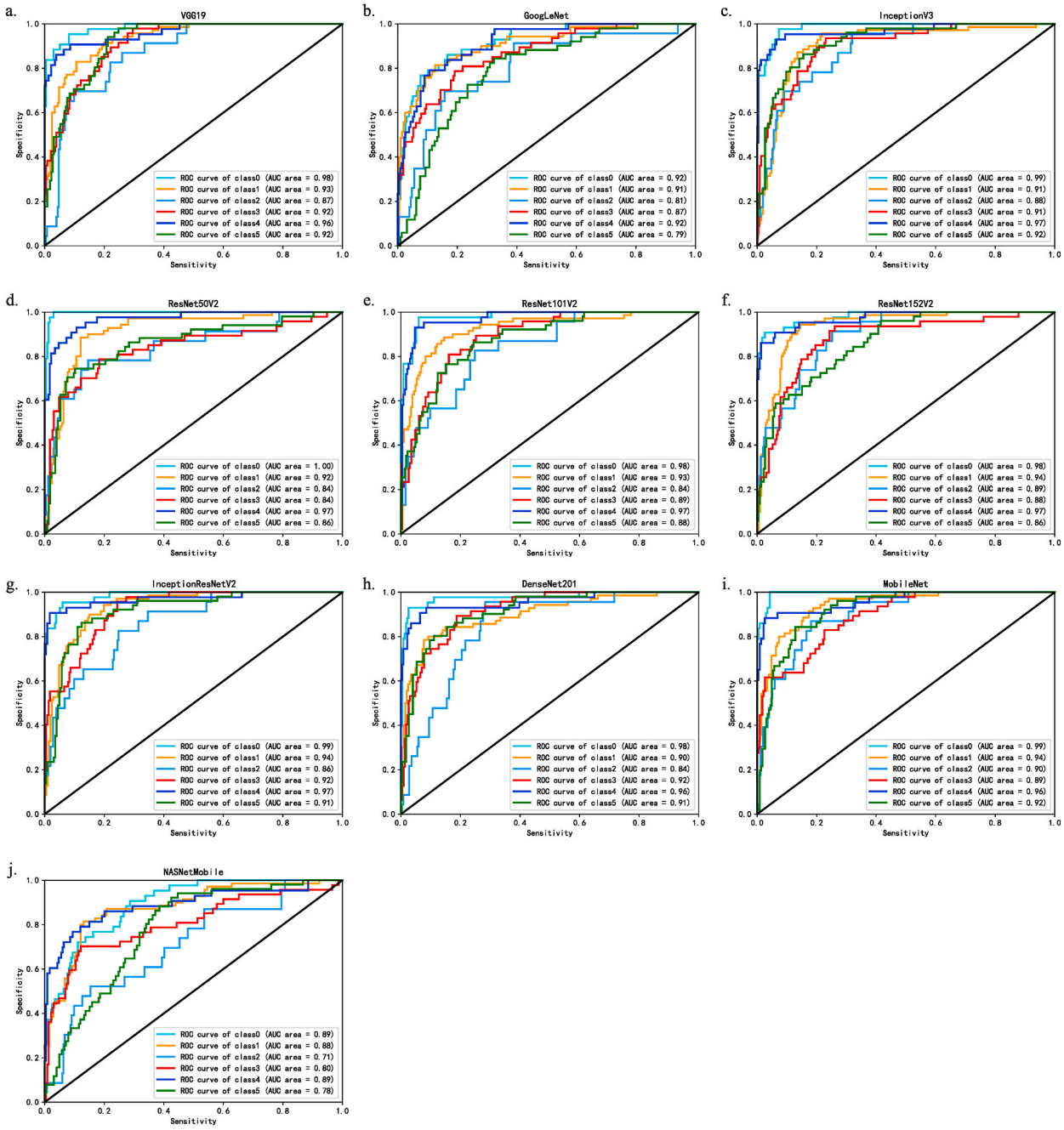


**Fig. 2.** Area Under the ROC Curve for 10 common convolutional neural network (CNN) algorithms for skin disease classification. The AUC of VGG-19 (a), GoogLeNet (b), InceptionV3 (c), ResNet50V2 (d), ResNet101V2 (e), ResNet152V2 (f), InceptionResNetV2 (g), DesseNet201 (h), MobileNet (i), and NASNetMobil (j) for predicting skin diseases.

## 4. Discussion

In this study, multiple pretrained models were used to classify the images of the most common hyperpigmented skin diseases. The experimental results were visualized using sensitivity-specificity curves and confusion matrices. To evaluate the 10 algorithms, VGG-19, GoogLeNet, InceptionV3, ResNet50V2, ResNet101V2, ResNet152V2, InceptionResNetV2, DesseNet201, MobileNet, and NAS-NetMobile, we used the model parameter, training time, test set accuracy, and AUC as evaluation indicators. According to the research findings, the classification effect of the MobileNet algorithm model was the best. Lastly, a system for online diagnosis of hyperpigmented skin diseases was developed. Doctors will now be able to improve their diagnosis efficiency and reduce their misdiagnosis rate

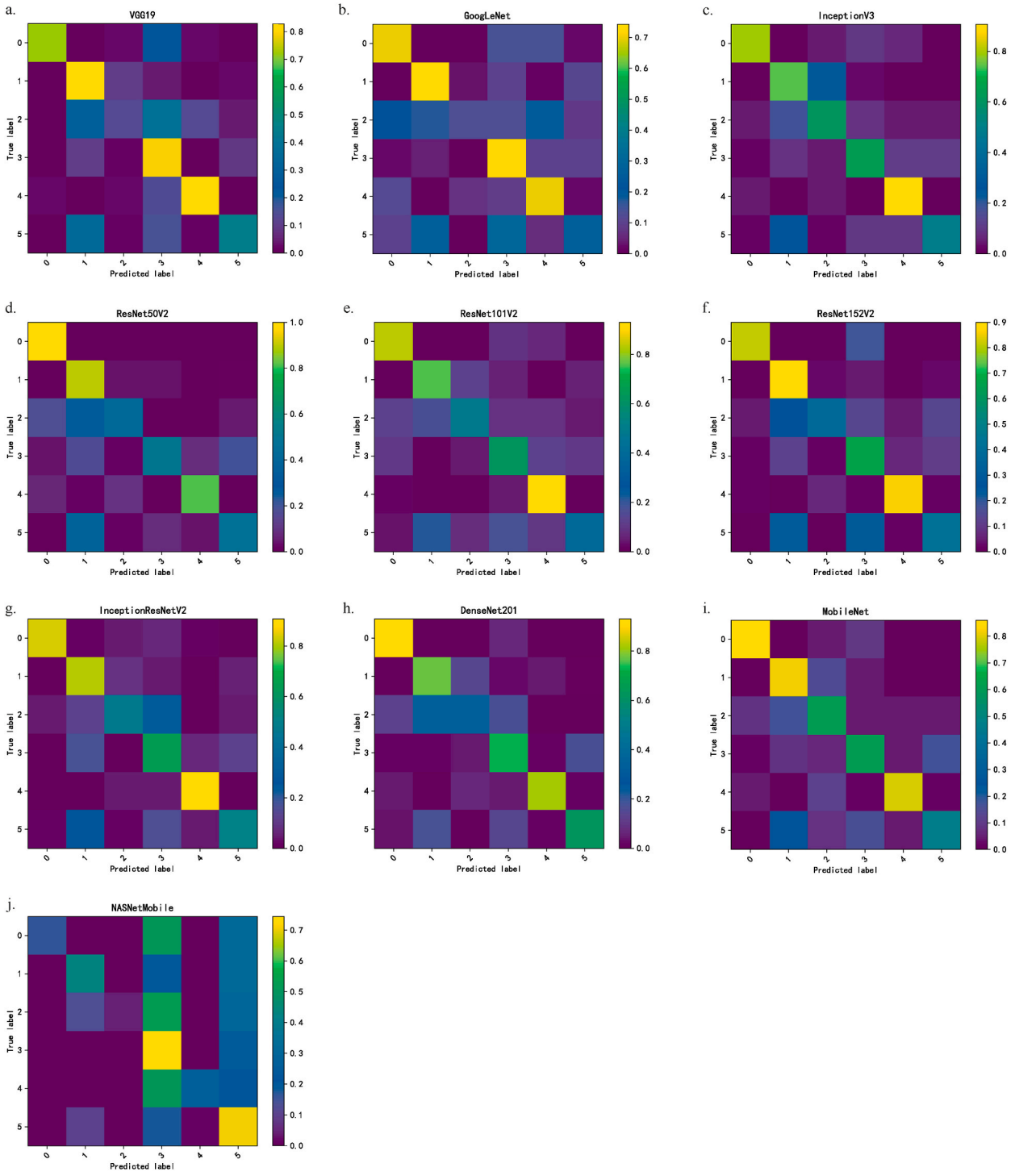


**Fig. 3.** Confusion matrix of 10 convolutional neural network (CNN) algorithms for predicting skin diseases, The confusion matrix of VGG-19 (a), GoogLeNet (b), InceptionV3 (c), ResNet50V2 (d), ResNet101V2 (e), ResNet152V2 (f), InceptionResNetV2 (g), DesseNet201 (h), MobileNet (i), and NASNetMobil (j) for predicting skin diseases, respectively.

using the optimal MobileNet algorithm for comprehensive diagnosis. Furthermore, it can be used to solve problems such as cross-regional diagnosis of patients and on-time diagnosis. Further advancements in deep learning will lead to better CNN network models for diagnosing pigmented skin diseases in the future. Online clinical diagnosis can be improved by combining the results of dermoscopy, RCM, skin pathology, etc. More accurate diagnosis results can be obtained using the comprehensive multimodal data as

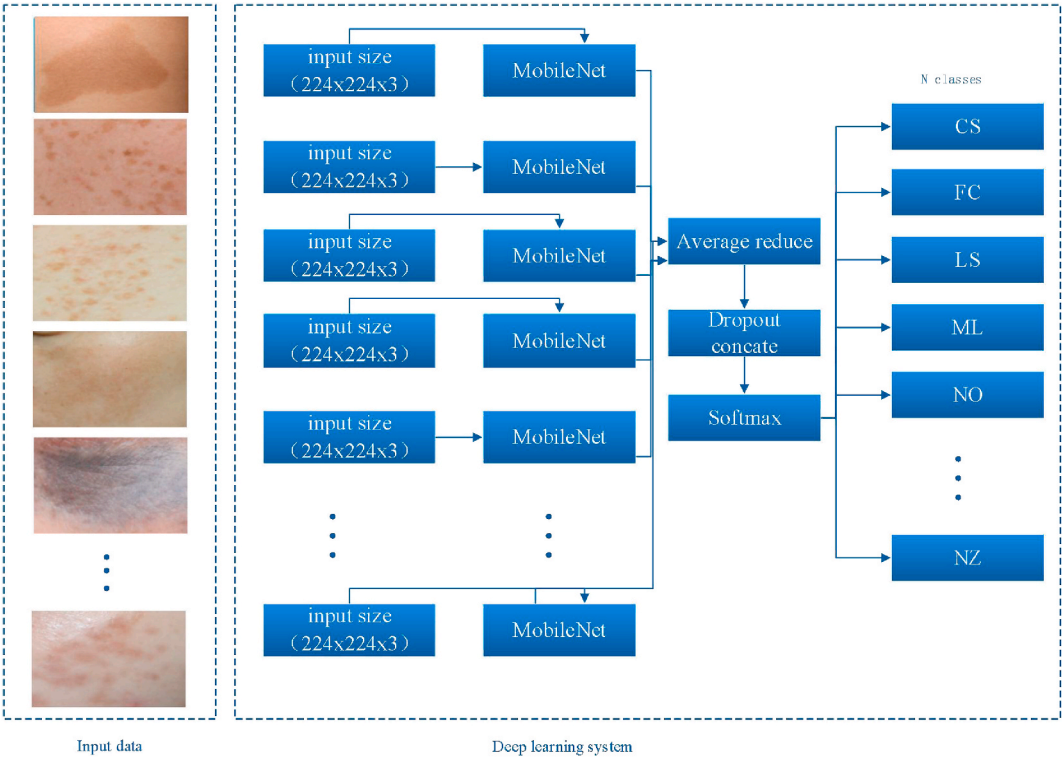


**Fig. 4.** The classification system of hyperpigmented skin diseases based on MobileNet.

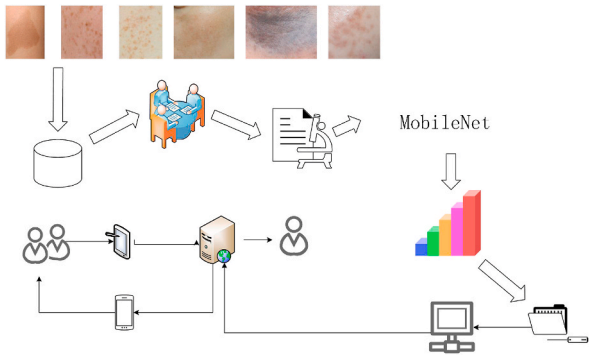


**Fig. 5.** The auxiliary diagnosis system of hyperpigmented skin diseases.

input to the deep learning algorithm.

The limitation of this study is that we only worked on 10 deep learning algorithm models. In the following research work, more algorithm models can be attempted for comparison. Second, according to the basic theory of deep learning, the more data trained, the more accurate is the resulting model. Therefore, the next step can be to consider using more image data of pigmented skin diseases for training.

Noise is another important factor affecting image processing and there are different kinds of noise. Different denoising techniques have been compared and evaluated recently to assess their performances in the classification of skin lesions [33]. In the proposed approach, we did not use any denoising technique. As an extension of this work, a denoising method in a pre-processing stage can be used to evaluate the performance of the proposed approach.

In the future, the performance of the proposed approach can be compared with that of a capsule network because capsule networks can preserve spatial relationships of learned features and have been used recently for image classification [34–36].

Deep network architectures are data-hungry; several augmentation methods in the literature have been applied to increase the reliability and robustness of the automated lesion classifications from skin images [37–40]. Therefore, an appropriate data augmentation can be employed and the performance of the proposed approach can be evaluated after training with expanded image

datasets as an extension of this study.

## 5. Conclusion

MobileNet provides a good prospect of clinical application in the auxiliary diagnosis of hyperpigmented skin, with high accuracy and vital generalization ability.

## Funding

This work was supported by the Hunan Provincial Health Commission (No. 202104121924).

## Author contribution statement

Jianyun Lu; Xiaoliang Tong: Contributed reagents, materials, analysis tools or data; Wrote the paper.
Hongping Wu; Yaoxinchuan Liu: Analyzed and interpreted the data.
Huidan Ouyang: Conceived and designed the experiments; Performed the experiments; Wrote the paper.
Qinghai Zeng: Conceived and designed the experiments; Wrote the paper.

## Data availability statement

Data associated with this study has been deposited at an open-source website (https://github.com/AIdental/six-classifications-skin-disease).

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e20186.

## References

[1] S. Desai, et al., Optimizing Melasma management with topical tranexamic acid: an expert consensus, J. Drugs Dermatol. JDD 22 (4) (2023) 386–392.
[2] C.C. Sun, et al., Naevus fusco-caeruleus zygomaticus, Br. J. Dermatol. 117 (5) (1987) 545–553.
[3] M. Kukla-Bartoszek, et al., DNA-based predictive models for the presence of freckles, Forensic Sci Int Genet 42 (2019) 252–259.
[4] B.D. McLarney, J.J. Parker, S. Hsu, Large café-au-lait spots on a 5-year-old boy, JAAD Case Rep 28 (2022) 127–129.
[5] J. Patrocínio, D. de Sousa, J.V. Frade, Nevus of ota, J. Gen. Intern. Med. 38 (5) (2023) 1302.
[6] L. Samuelov, et al., Extensive lentigo simplex, linear epidermolytic naevus and epidermolytic naevus comedonicus caused by a somatic mutation in KRT10, Br. J. Dermatol. 173 (1) (2015) 293–296.
[7] N.C. Syder, C. Quarshie, N. Elbuluk, Disorders of facial hyperpigmentation, Dermatol. Clin. 41 (3) (2023) 393–405.
[8] D.C. Castro, I. Walker, B. Glocker, Causality matters in medical imaging, Nat. Commun. 11 (1) (2020) 3673.
[9] E. Goceri, Automated skin cancer detection: where we are and the way to the future, in: 44th International Conference on Telecommunications and Signal Processing, TSP 2021, 2021.
[10] E. Goceri, Convolutional neural network based desktop applications to classify dermatological diseases, in: 2020 IEEE 4th International Conference on Image Processing, Applications and Systems, IPAS), 2020.
[11] E. Goceri, Impact of deep learning and smartphone technologies in Dermatology: automated diagnosis, in: 2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA), 2020.
[12] E. Goceri, A.A. Karakas, Comparative evaluations of CNN based networks for skin lesion classification, in: The 14th International Conference on Computer Graphics, Visualization, Computer Vision and Image Processing, CVGCVIP 2020, 2020.
[13] E. Goceri, An application for automated diagnosis of facial dermatological diseases, İzmir Katip Çelebi Üniversitesi Sağlık Bilimleri Fakültesi Dergisi 6 (3) (2021) 91–99.
[14] A. Esteva, et al., Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115–118.
[15] P. Tschandl, et al., Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study, Lancet Oncol. 20 (7) (2019) 938–947.
[16] J. Zhang, et al., Medical image classification using synergic deep learning, Med. Image Anal. 54 (2019) 10–19.
[17] Y. Liu, et al., A deep learning system for differential diagnosis of skin diseases, Nat Med 26 (6) (2020) 900–908.
[18] H. Wu, et al., A deep learning-based smartphone platform for cutaneous lupus erythematosus classification assistance: simplifying the diagnosis of complicated diseases, J. Am. Acad. Dermatol. 85 (3) (2021) 792–793.
[19] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[20] M. Mateen, et al., Fundus image classification using VGG-19 architecture with PCA and SVD, Symmetry 11 (1) (2018).
[21] P. Tang, H. Wang, S. Kwong, G-MS2F: GoogLeNet Based Multi-Stage Feature Fusion of Deep CNN for Scene Recognition, Neurocomputing, 2017.
[22] C. Szegedy, et al., Rethinking the Inception Architecture for Computer Vision, 2015.
[23] K. He, et al., in: [IEEE 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) - Las Vegas, NV, USA (2016.6.27-2016.6.30)] 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) - Deep Residual Learning for Image Recognition, 2016.

[24] C. Szegedy, et al., Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, 2017.
[25] G. Huang, et al., Densely connected convolutional networks, in: IEEE Conference on Computer Vision and Pattern Recognition, 2017.
[26] A.G. Howard, et al., MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, 2017.
[27] F. Saxen, et al., Face attribute detection with MobileNetV2 and NasNet-mobile, in: 2019 11th International Symposium on Image and Signal Processing and Analysis, ISPA), 2019.
[28] E. Goceri, Intensity normalization in brain mr images using spatially varying distribution matching, in: Conferences Computer Graphics, Visualization, Computer Vision and Image Processing 2017, 2017.
[29] E. Goceri, Fully Automated and Adaptive Intensity Normalization Using Statistical Features for Brain MR Images, Celal Bayar Üniversitesi Fen Bilimleri Dergisi, 2018, pp. 125–134.
[30] B. Lyu, et al., Deep leaning based medicine packaging information recognition for medication use in the elderly, Procedia Computer Science 187 (2021) 194–199.
[31] T. Deimel, D. Aletaha, G. Langs, OP0059 AUTOSCORA: deep learning to automate scoring of radiographic progression in rheumatoid arthritis, Ann. Rheum. Dis. 79 (Suppl 1) (2020), 39.2-40.
[32] J. Huang, C.X. Ling, Using AUC and accuracy in evaluating learning algorithms, IEEE Trans. Knowl. Data Eng. 17 (3) (2005) 299–310.
[33] E. Goceri, Evaluation of denoising techniques to remove speckle and Gaussian noise from dermoscopy images, Comput. Biol. Med. 152 (2023), 106474.
[34] E. Goceri, Classification of skin cancer using adjustable and fully convolutional capsule layers, Biomed. Signal Process Control 85 (2023).
[35] E. Goceri, Analysis of capsule neural networks for image classification, in: 15th International Conference on Computer Graphics, Visualization, Computer Vision and Image Processing, CGVCVIP 2021, 2021.
[36] E. Goceri, Capsule neural networks in classification of skin lesions, in: 15th International Conference on Computer Graphics, Visualization, Computer Vision and Image Processing, CGVCVIP 2021, 2021.
[37] D. Shen, G. Wu, H.I. Suk, Deep learning in medical image analysis, Annu. Rev. Biomed. Eng. 19 (2017) 221–248.
[38] E. Goceri, Comparison of the impacts of dermoscopy image augmentation methods on skin cancer classification and a new augmentation method with wavelet packets, Int. J. Imag. Syst. Technol. 33 (5) (2023) 1727–1744.
[39] E. Goceri, Medical image data augmentation: techniques, comparisons and interpretations, Artif. Intell. Rev. (2023) 1–45.
[40] E. Goceri, Image augmentation for deep learning based lesion classification from skin images, in: 2020 IEEE 4th International Conference on Image Processing, Applications and Systems, IPAS), 2020.