



Transfer learning with deep convolutional neural network for constitution classification with face image

Er-Yang Huan¹ · Gui-Hua Wen¹ 

Received: 4 December 2018 / Revised: 27 August 2019 / Accepted: 9 October 2019 /
Published online: 10 January 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Constitution classification is the basis and core content of constitution research in Traditional Chinese medicine. The convolutional neural networks have successfully established many models for image classification, but it requires a lot of training data. In the field of Traditional Chinese medicine, the available clinical data is very limited. To solve this problem, we propose a method for constitution classification through transfer learning. Firstly, the DenseNet-169 model trained in ImageNet is applied. Secondly, we carefully modify the DenseNet-169 structure according to the constitution characteristics, and then the modified model is trained in the clinical data to obtain the constitution identification network called ConstitutionNet. In order to further improve the accuracy of classification, we integrate the ConstitutionNet with Vgg-16, Inception v3 and DenseNet-121 to test according to the integrated learning idea, and judge the input face image to its constitution type. The experimental results show that transfer learning can achieve better results in small clinical dataset, and the final accuracy of constitution recognition is 66.79%.

Keywords Constitution classification · Convolutional neural networks · Transfer learning · DenseNet-169 · ConstitutionNet · Integrated test

1 Introduction

The constitution in Traditional Chinese medicine (TCM) refers to the relatively stable body traits of the individual to the internal and external environment of the body. It is a morphological structural psychological state and physiological function formed on the basis of congenital inheritance, which is a system concept formed by combining the Chinese medical discourse on human physique phenomena and the understanding of physique in many

✉ Gui-Hua Wen
crghwen@scut.edu.cn

¹ School of Computer Science and Engineering, China Guangdong Artificial Intelligence Engineering Research Center for Traditional Chinese Medicine, South China University of Technology, Guangzhou, China

disciplines and the purpose of medical research [10]. Constitution phenomenon is an important manifestation of human life phenomenon. It has the characteristics of individual differences, relative stability, dynamic variability, etc. [53, 54].

Constitution classification is the basis and core content of constitution research in TCM. The purpose is to standardize human constitution categories, and then to give different personalized conditioning solution for different constitution types. So it is especially important for specific people to accurately identify their constitution type. The existing constitution identification method is judged by questionnaire survey. The individual fills in the scale according to the national standard “Classification and Judgment of TCM Constitution” (ZYYXH/T157.2009) [5], and then evaluates the individual’s completed scale according to the scoring rules provided by the standard, and then determines the constitution type of the individual. This kind of scale-based approach has some shortcomings in clinic.

- 1) Subjective factors of the individual are greatly affected. Individuals are not very familiar with some issues and it is difficult to accurately choose an answer. Second, individuals have concerns about some issues and are reluctant to choose real answers.
- 2) The number of questions on the scale is relatively large. It takes too long to answer these questions, which makes it easy for individuals to lose patience in the process of filling out the body mass table. The latter problems on the scale are often randomly selected, which will inevitably affect the correct judgment of constitution.
- 3) The calculation formula of the scoring rules is complicated, and it is impossible to accurately calculate the physical types of many individuals [33, 60]. The original score and conversion score need to be calculated through the table, and then the score is compared to the score interval of each constitution.

In order to solve this problem, machine learning algorithms have been applied to constitution recognition, including the convolutional neural network algorithms [18, 29, 35]. In particular, the convolutional neural networks have led a series of breakthroughs in the field of image classification [12, 17, 19, 59]. However, training a convolutional neural network model from scratch is not easy, and requires a long time. It even requires some patience and expertise about training neural networks [46], especially it requires a large amount of labeled training data. However, the hospital currently only has medical records and prescription data in the field of Traditional Chinese medicine, no image data, and then need to be re-prepared. Moreover, it is very expensive for experts to label large amounts of data in TCM, which is difficult to meet in a short time. Secondly, if the convolutional neural network model is trained directly with small clinical dataset, the accuracy is not good and even over-fitting problems may be encountered. Therefore, we propose a constitution recognition method based on transfer learning, called transfer constitution recognition (TCR). This paper first uses the DenseNet-169 [19] model trained in ImageNet [9] and modifies its network. And then the modified network is retrained with a small sample of clinical face dataset. Finally, it integrates multiple models to achieve individual constitution type. The main contribution of the paper is:

- 1) This paper constructs a Constitution Identification Network called ConstitutionNet. Firstly, the DenseNet-169 model trained in ImageNet. Secondly, the DenseNet-169 structure is modified according to the physical characteristics, and then the modified model is trained with the clinical dataset to obtain the constitution type. The ConstitutionNet has obtained a better accuracy of constitution recognition.

- 2) In order to further improve the accuracy of classification, our paper uses the integrated learning idea to integrate the ConstitutionNet with Vgg-16 [40], Inception v3 [44] and DenseNet-121 [19], and determine the physique type of the input image. Finally, the accuracy of constitution recognition is improved.

The rest of this paper is organized as follows. In Section 2, we briefly present the related work. Section 3 details the proposed method. Experimental results as well as the discussion are given in Section 4, and Section 5 concludes this paper.

2 Related work

The commonly used constitution type criteria is determined by the constitution questionnaire which is developed by wang [52] in the mainland, su [30–32, 41, 42] in Taiwan and wang [57] in Hong Kong. Wang et al. [52] divided the constitution into nine types, namely, gentleness, qi-deficiency, qi-depression, dampness-heat, phlegm-dampness, blood-stasis, special-diatheasis, yang-deficiency and yin-deficiency. Su et al. [43] studied the acoustic characteristics of eight different constitutions and applied them to constitution recognition. Wang et al. [55] classified the constitution through pulse which applied the BP neural network, and demonstrated the rationality and superiority of this method. Convolutional neural network is a specific type of neural network, which is a feed-forward neural network, including convolutional layer, pooled layer and fully connected layer. Due to its outstanding performance, the convolutional neural network is widely used in many fields, such as image classification [56, 62], target detection [27, 63], image segmentation [13, 34], visual tracking [24–26], etc. Hu et al. [16] applied the convolution neural network to the pulse diagnosis. In the case of feature ambiguity, the proposed method was superior to other well-known methods. Li et al. [28] used the convolution neural network to extract the features of the pulse, and then classified the body constitution type. The experimental results show that this method can obtain high accuracy. Huan et al. [18] proposed a constitution recognition algorithm based on convolutional neural network, which trained a convolutional neural network model for constitution recognition on face data. Li et al. [29] proposed a constitution recognition algorithm based on deep neural network, which first detected the tongue image and then determined the body constitution type. Ma et al. [35] proposed a complex perception-based algorithm for constitution recognition, whose dataset is tongue picture.

Facing the problem of collecting enough training data to train the model, the purpose of transfer learning is to transfer knowledge learned from source domains in big data to target domains in smaller data. The transfer learning based CNN has been used in many fields [14, 20, 49]. Burdick et al. [3] applied transfer learning to segment skin lesions and led to good classification results. Kermany et al. [22] used transfer learning to construct a diagnostic tool for screening patients with common treatments for blinding retinal diseases. Rajpurkar et al. [37] proposed the CheXnet network for pneumonia detection through chest X-rays images. The algorithm used transfer learning technology and was trained through the DenseNet-121 model.

3 Method

The algorithm proposed in this paper is divided into four main parts, (1) Data Acquisition (2) Data Preprocessing (3) Data Augmentation (4) Constitution Recognition through transfer

learning. The flow chart of the whole algorithm is shown in Fig. 1. First, the clinical face dataset is collected and preprocessed. Then, the pre-processed image dataset is subjected to data augmentation to obtain training data. Finally, the obtained training data is subjected to constitution recognition by transfer learning technology. The following sections provide a detailed description of the modules included in the architecture.

3.1 Data acquisition and preprocessing

The clinical face training dataset used in this paper has 12,730 images, with a type of constitution judged by clinical TCM experts. The constitution type is based on professor wang 's judgment criteria [52]. Before collecting data, the standard is discussed by nearly ten medical experts. Some agreed with this standard. Some professors were partially in favor of the standard. Some professors have a negative attitude on this standard. We chose three professors who were in favor of this standard. This means that they reached the consensus (agreement of standard) to determine the type of body constitution. Subsequently, they were in different hospitals to judge the patient's body constitution according to the standard. In this way, the impact of experience can be reduced as much as possible. Besides, these professors are well known and their ages are close, and the personal experience is not greatly different. Finally, the body constitution type of the patient in the same hospital is determined by the same medical professor. The entire dataset is determined by three Chinese medicine professors from three different hospitals according to the above-mentioned standard.

Therefore, all face images are taken by the same type of digital device and the patient's constitution type is specified by the doctor. The indoor environment is no sunshine, and lighting conditions are normal fluorescent lamps. In the face database, there are 8 kinds of constitution types, that is, gentleness, qi-deficiency, qi-depression, dampness-heat, phlegm-dampness, blood-stasis, yang-deficiency, and yin- deficiency. The number of each constitution type is shown in Table 1. Each constitution example is shown in Fig. 2. In the preprocessing

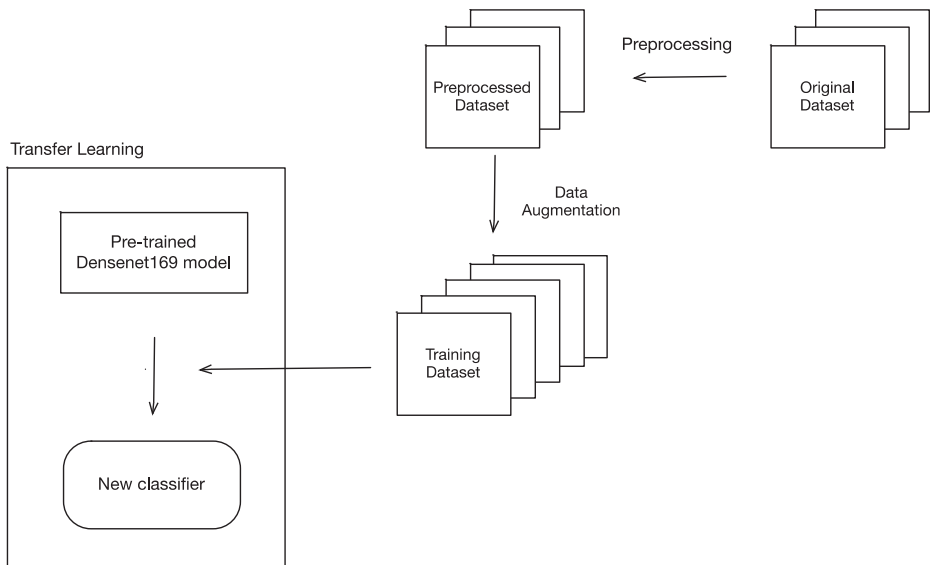


Fig. 1 The flow chart of the whole algorithm

Table 1 The number of each constitutional type

	Gentleness	Qi-deficiency	Yang-deficiency	Yin-deficiency	Phlegm-dampness	Dampness-heat	Blood-stasis	Qi-depression	Sum
Training Dataset	890	2153	800	2029	2021	2084	727	2026	12,730
Testing Dataset	57	75	60	75	75	75	41	75	533

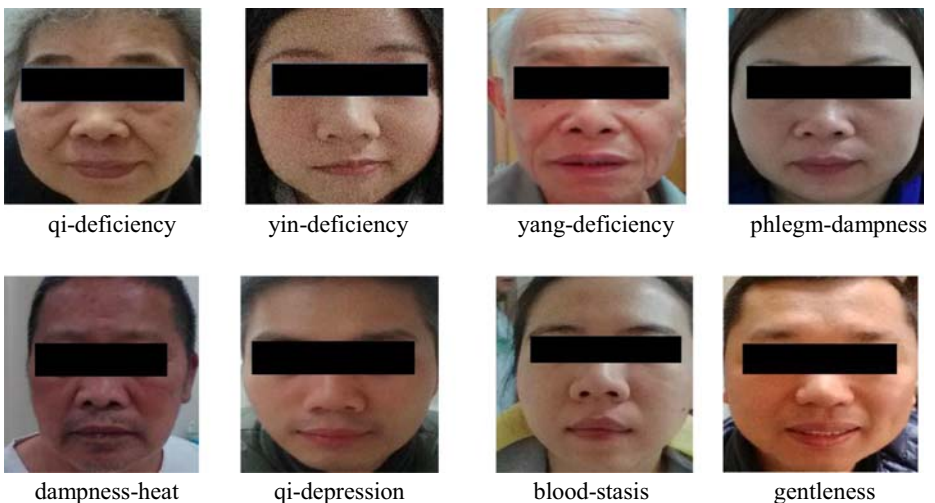
process, firstly, the face detection algorithm is used to detect the acquired picture, and the corresponding bounding box is obtained. Considering both time complexity and precision, this paper uses OpenCV tool to complete the face detection.

3.2 Data augmentation

The dataset collected in this paper is limited. At the same time, the data augmentation technology can not only increase the size of the dataset, but also avoid the over-fitting. Then the data augmentation is applied in the collected dataset. The original image is preprocessed in the training phase. Each image is 224×224 in size. In this paper, the width and height of the image are scaled proportionally and the image is zoomed in both length and width direction. This paper uses the Keras [7] tool to achieve data augmentation through the functions it contains, just through setting the values of `width_shift_range`, `height_shift_range` and `zoom_range` in the `ImageDataGenerator` function. After data augmentation, these pictures are trained through transfer learning.

3.3 Classifier architecture

The current clinical data is limited. We don't use this small dataset to train the entire CNN, but use the transfer learning method, which takes advantage of features previously learned from

**Fig. 2** Each constitution example

larger dataset. We propose a new constitution identification network (ConstitutionNet), as shown in Fig. 3. It first uses the DenseNet-169 model trained in ImageNet, and modifies the DenseNet-169 model according to the characteristics of the constitution.

- 1) The final fully connected output layer performs eight classifications (gentleness, qi-deficiency, qi-depression, dampness-heat, phlegm-dampness, blood-stasis, yang-deficiency, and yin-deficiency), not 1000 classes as previously designed for the ImageNet dataset.
- 2) In DenseNet-169, the Google’ Inception block [21] and the ResNet’ Residual block [12] are added before fully-connected layer. The Inception block can increase the width of the network. The receptive field of different branches is different, so there is multi-scale information contained in it, and the parameters are reduced. The Residual block enables the network to increase with depth without gradient degradation.

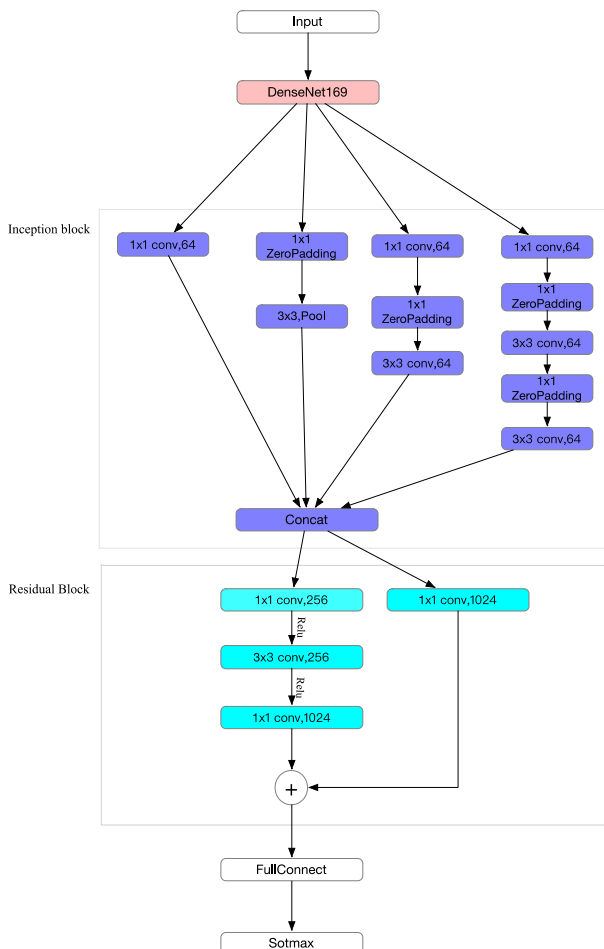


Fig. 3 The network structure of ConstitutionNet

3.4 Integrated constitution identification

To further improve the accuracy of classification, we adopt integrated learning ideas. Different classification models have different classification effects for different categories, and the classification effect is improved by the complementarity between model classifications. Through experimental comparison, this paper chooses to integrate the ConstitutionNet with Vgg-16, Inception v3 and DenseNet-121. First, the network models VGG16, Inception v3, and DenseNet-121 are trained separately through the transfer learning, which all models are implemented and trained through the Keras tool. Second, a test face image is inputted into the VGG16, Inception v3, DenseNet-121, and ConstitutionNet models to separately compute the probability of each constitution in the test process. Third, the average operation is performed to the four probabilities of each constitution from each model so as to obtain the final probability of each constitution. Finally, the constitution type corresponding to the index of the maximum probability is taken as the recognized constitution type for the input image. As shown in Fig. 4, the input face image is judged by integrated model, and then improves the accuracy of the constitution recognition.

4 Experiments

In this section, we conducted a series of experiments to measure the effectiveness of transfer learning applied to the body constitution recognition algorithm. The details of these experiments are described below.

4.1 Experiment settings

The tools used in this experiment are based on Keras [7], TensorFlow [48], Scikit-learn [39] and Scikit-image [51]. The GPU is NVIDIA GTX Titan X, and its memory size is 12 GB. The operating system is Ubuntu 14.04. The face training dataset used in this paper has a total of 12,370 images, and the test dataset has 533 images. The whole network is trained by random

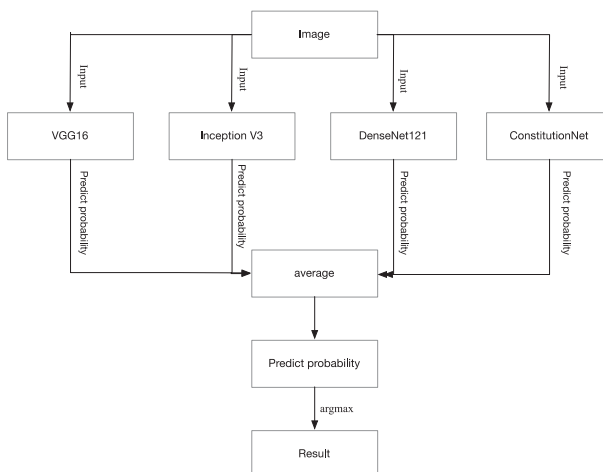


Fig. 4 The structure diagram for integrated constitution recognition

Table 2 Classification results of different classifiers under different feature extraction

	SVM	Random forest	KNN	Softmax	Decision Tree	Naive Bayes
Haar-like	20.82%	14.07%	13.69%	14.63%	15.01%	9.38%
LBP	62.10%	57.22%	24.20%	39.21%	56.09%	16.88%
HOG	14.07%	60.78%	22.14%	17.64%	56.84%	15.76%
ConstitutionNet	64.91%	64.16%	63.79%	65.67%	62.10%	63.41%

gradient method. The learning rate is 0.0002, the momentum is set to 0.9, and the batch size is set to 30. In data augmentation processing, the values of width_shift_range, height_shift_range, and zoom_range are set to 0.2.

4.2 Experiment results

The purpose of the experiment is to verify the effectiveness of the proposed method. First of all, this paper chooses to compare with the traditional feature extraction methods to verify the advantages of the deep learning method. Second, comparing with some representative deep learning models, we demonstrate the effectiveness of our proposed constitution recognition network. Finally, our proposed integrated constitution recognition method is compared with the method we proposed earlier, indicating that the method of this paper has made new progress.

4.2.1 Comparison of different feature extraction methods

There are many methods for traditional image feature extraction. The ConstitutionNet network is also used to extract features. We compare it with traditional feature extraction methods to prove the superiority of deep learning methods. Methods for extracting facial image features include histogram of oriented gradient(HOG) [8], local binary patterns (LBP) features [1], Haar-Like features [50], and based on the ConstitutionNet. Considering that the same classifier is used, the classifiers of different feature extraction methods have different effects, and the same feature extraction method has different effects of different classifiers. Therefore, the classifier chooses a different principle classifier, namely, Logistic Regression classifier (LR) [61], Naive Bayes classifier (NB) [2], Support Vector Machine classifier (SVM) [11], Random Forest Classifier(RF) [36], KNN classifier(KNN) [38], Decision Tree classifier(DC) [4]. Among them, the kernel function in the support vector machine is the RBF.

Table 3 the confusion matrix under the SVM algorithm based on LBP feature

	Qi-deficiency	Yin-deficiency	Yang-deficiency	Phlegm-dampness	Dampness-heat	Qi-depression	Blood-stasis	Gentleness
Qi-deficiency	75	0	0	0	0	0	0	0
Yin-deficiency	16	59	0	0	0	0	0	0
Yang-deficiency	25	0	35	0	0	0	0	0
Phlegm-dampness	17	0	0	58	0	0	0	0
Dampness-heat	26	0	0	0	49	0	0	0
Qi-depression	50	0	0	0	0	25	0	0
Blood-stasis	11	0	0	0	0	0	30	0
Gentleness	57	0	0	0	0	0	0	0

Table 4 the confusion matrix of the random forest algorithm based on HOG feature

	Qi-deficiency	Yin-deficiency	Yang-deficiency	Phlegm-dampness	Dampness-heat	Qi-depression	Blood-stasis	Gentleness
Qi-deficiency	37	11	0	9	12	5	0	1
Yin-deficiency	3	64	0	1	5	2	0	0
Yang-deficiency	5	4	35	6	3	7	0	0
Phlegm-dampness	2	4	0	65	2	2	0	0
Dampness-heat	5	3	0	7	56	4	0	0
Qi-depression	13	7	0	7	13	35	0	0
Blood-stasis	4	1	0	2	1	1	32	0
Gentleness	19	10	0	8	11	9	0	0

It can be seen from Table 2 that under the premise of the same classifier, the classification effect based on the ConstitutionNet is better than that single LBP, HOG and Haar-like features. Moreover, the classification effect based on LBP features is significantly better than that based on single HOG and Haar-like in the traditional feature extraction method under the same classifier, namely, SVM, KNN, Softmax, Naïve Bayes. But the classification effect based on HOG features is significantly better than that based on single LBP and Haar-like under the same classifier, namely, Random Forest, Decision Tree. At the same time, the classification effects of different classifiers are compared under the same feature extraction method. Based on the single Haar-like feature and LBP feature, the SVM has the best classification accuracy. Based on the single HOG feature, the random forest classification works best. Based on the features extracted by the ConstitutionNet, the logistic regression has the best classification effect. Overall, the ConstitutionNet network is far better than other feature extraction methods.

Further, we use the confusion matrix to analyze the sensitivity of various methods to different constitution type. The confusion matrix under the SVM algorithm based on LBP feature is shown in Table 3. It can be seen from Table 3 that the algorithm has the best classification effect on qi-deficiency and the worst classification effect on gentleness. The gentleness is misclassified to qi-deficiency. The confusion matrix of the random forest algorithm based on HOG feature is shown in Table 4. It can be seen from Table 4 that the algorithm has the best classification effect on the phlegm-dampness and the classification effect on the gentleness is not good. And the confusion matrix of Softmax based on ConstitutionNet is shown in Table 5. It can be seen from Table 5 that the algorithm has the best classification of yin-deficiency and the worst classification effect on gentleness. From these confusion matrices, the combination of different feature extraction methods and classifiers has different characteristics, providing a basis for further combination classification.

In order to analyze the characteristics of the ConstitutionNet network from different aspects, we analyze the classification effect of the combination with multiple classifiers through the

Table 5 the confusion matrix of softmax based on ConstitutionNet

	Qi-deficiency	Yin-deficiency	Yang-deficiency	Phlegm-dampness	Dampness-heat	Qi-depression	Blood-stasis	Gentleness
Qi-deficiency	48	2	5	7	11	0	1	1
Yin-deficiency	5	65	1	0	2	2	0	0
Yang-deficiency	9	0	43	1	6	0	1	0
Phlegm-dampness	3	3	0	63	5	1	0	0
Dampness-heat	9	1	4	2	58	1	0	0
Qi-depression	9	1	1	2	6	37	1	18
Blood-stasis	4	1	1	1	1	0	33	0
Gentleness	11	6	9	6	10	4	8	3

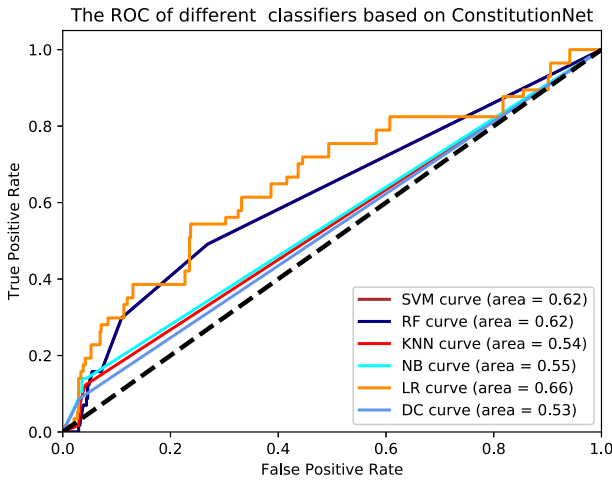


Fig. 5 the ROC curves of different classifiers based on the ConstitutionNet

receiver operating characteristic curve. Based on the ConstitutionNet feature extraction, the ROC curves of different classifiers are shown in Fig. 5. As can be seen from Fig. 5, the Logistic Regression classifier has the largest area, with a value of 0.66, indicating that the Logistic Regression classifier performs best. The Decision Tree classifier has the smallest area with a value of 0.53, which indicates that the Decision Tree classifier has the worst performance.

Since the combination of the ConstitutionNet feature extraction method and the softmax classifier works best, we list the ROC curves for each label similar to the confusion matrix. As shown in Fig. 6, it can be seen that the ROC area of the yin -deficiency is 0.96, and the area of the gentleness is 0.66, indicating that the classification of yin-deficiency is the best and the classification of gentleness is the worst. In addition to the ROC curve, we also use other indicators to evaluate the classifier effects, namely macro-averaging and precision-recall curves. The experimental results are shown in Figs. 7 and 8. The macro-averaging is used by classifiers to measure the validity of small class discrimination. From Fig. 7, it can be seen

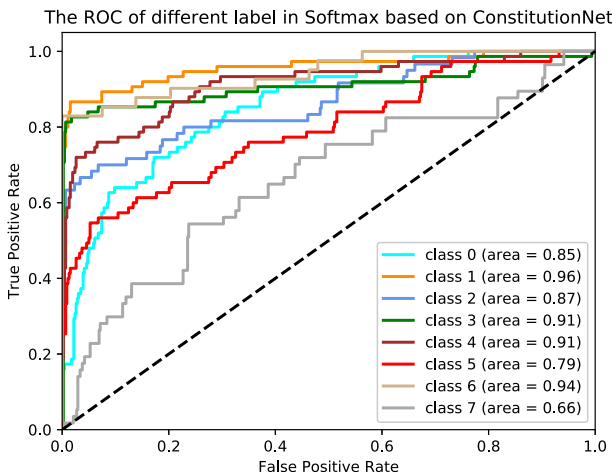


Fig. 6 the ROC curve of each label in softmax based on ConstitutionNet

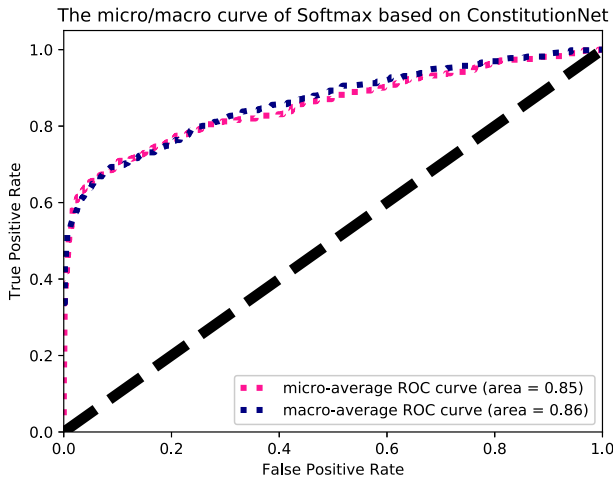


Fig. 7 the micro/macro-average of softmax based on ConstitutionNet

that the ROC area of macro-averaging is 0.86. In the training dataset collected in this paper, the number of physical types of gentleness, yang-deficiency and blood-stasis is small, indicating that the classifier has a good classification effect on the three type. From Fig. 8, it can be seen that the precision-recall curve of the support vector machine, the random forest and the decision tree fluctuates significantly, and the precision-recall curve of the Regression classifier is relatively smooth. It can also be seen that the area under the precision-recall is relatively large, indicating the Logistic Regression classification works well.

4.2.2 Comparison with other deep network models

There are some representative deep learning networks. Krizhevsky et al. [23] constructed an AlexNet network to classify in the ImageNet and achieved the best results. Later, different scholars proposed different network for classification in the ImageNet, such as VGG-16 [40],

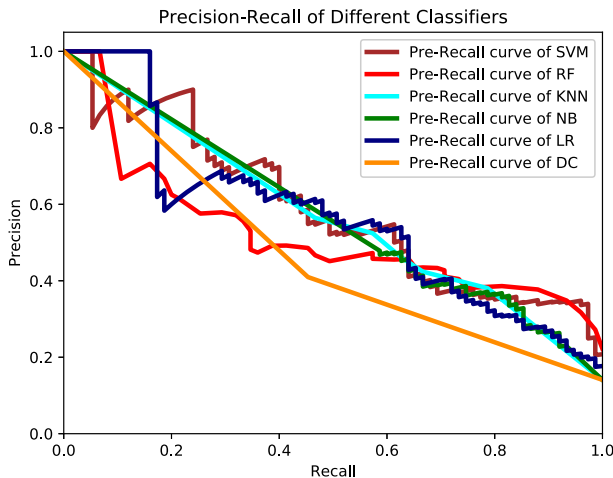


Fig. 8 the precision-recall curve of different classifiers based on ConstitutionNet

Table 6 The classification result of different model

Model	Result
VGG-16 [40]	62.66%
Inception V3 [44]	63.79%
ResNet-50[12]	62.24%
InceptionResnet [45]	61.72%
Mobilenet v1 [15]	62.10%
Xception [6]	62.60%
DenseNet-121 [19]	64.25%
SE-DenseNet121 [17]	62.10%
SE-InceptionV3 [17]	62.67%
SE-MobileNet [17]	61.91%
SE-ResNet50 [17]	64.16%
CBAM-ResNet50 [58]	62.48%
CBAM-MobileNet [58]	60.66%
CBAM-DenseNet [58]	59.29%
CBAM-InceptionV3 [58]	60.23%
EfficientNets [47]	62.66%
ConstitutionNet	65.67%

ResNet [12], Inception V4 [45], Xception [6], MobileNet v1 [15], SENet [17], CBAM [58] and EfficientNets [47]. These networks can be used for constitution identification, and the effects are different. In order to verify the effect of the constitution recognition network proposed in this paper, we also use transfer training to train these networks on the same training dataset. The experimental results are shown in Table 6. It is not difficult to see that the ConstitutionNet model works best, with an accuracy rate of 65.67%.

4.2.3 Comparison of integrated constitution recognition

In order to further improve the accuracy of face recognition, this paper uses integrated learning ideas and selects four deep learning models in Table 6 for integration. They are Vgg-16, Inception v3, DenseNet-121 and ConstitutionNet. These four models are all trained on the same training dataset using transfer learning. However, the test dataset uses the literature [18], the purpose is to compare with the previous method, indicating that the proposed method has been further improved. The test results are shown in Table 7. It can be seen from Table 7 that the test result of the convolutional neural network is 64.54%, and then the result by fusion of CNN and color feature is 65.29% in the literature [18]. In this paper, the result of ConstitutionNet is 65.67%, which is little higher than the result of the literature [18]. The accuracy of the integrated recognition is 66.79%, which is higher than that of all comparison models. The confusion matrix the integrated constitution recognition is shown in Table 8. It is not difficult to see that the integrated model is the best for yin-deficiency, but the classification of the gentleness is the worst, indicating that there is model with poor classification for

Table 7 Comparison of results from different literature methods

	Huan et al. [18]		Our paper	
	CNN	CNN + color	ConstitutionNet	Model integration test
accuracy	64.54%	65.29%	65.67%	66.79%

Table 8 Confusion matrix of model integration test

	Qi-deficiency	Yin-deficiency	Yang-deficiency	Phlegm-dampness	Dampness-heat	Qi-depression	Blood-stasis	Gentleness
Qi-deficiency	55	3	4	5	6	2	0	0
Yin-deficiency	6	65	0	1	1	2	0	0
Yang-deficiency	9	0	42	0	8	0	1	0
Phlegm-dampness	3	3	0	63	3	2	0	1
Dampness-heat	15	0	1	2	56	1	0	0
Qi-depression	4	0	1	4	8	39	0	19
Blood-stasis	6	1	0	1	1	0	32	0
Gentleness	16	8	5	8	8	6	2	4

gentleness in the model integration. If the model is well classified for gentleness, it can further improve the classification effect by model integration test.

5 Conclusion

Face consultation is an important diagnosis method in the Traditional Chinese medicine. This paper applies face diagnosis to constitution recognition. Because the clinical dataset of physique type is very limited and we want to take the great advantages of deep learning network, this paper proposes the ConstitutionNet for constitution recognition which is obtained through the transfer learning. The DenseNet-169 model trained in ImageNet is used for physique recognition through transfer learning technology, and then the model is modified to further suit the TCM constitution recognition. In order to further improve the accuracy of constitution recognition, an integrated physique recognition method is proposed. The basic classifier includes ConstitutionNet and the other three most representative deep networks, with accuracy of 66.79%. Experiments show that transfer learning and integrated learning are effective for constitution recognition with limited clinical data. The future work is to explore more mechanisms of transfer learning methods and to be used for constitution recognition.

References

1. Ahonen T, Hadid A, Pietikäinen M (2004) Face recognition with local binary patterns. Proceedings of Springer European Conference on Computer Vision, In, pp 469–481
2. Bermejo P, Gámez JA, Puerta JM (2014) Speeding up incremental wrapper feature subset selection with Naive Bayes classifier. Knowledge-Based Systems 55:140–147
3. Burdick J, Marques O, Weinthal J, Furht B (2018) Rethinking skin lesion segmentation in a convolutional classifier. J Digit Imaging 31(4):435–440
4. Chen KH, Wang KJ, Wang KM et al (2014) Applying particle swarm optimization- based decision tree classifier for cancer classification on gene expression data. Appl Soft Comput 24:773–780
5. China Association of Chinese Medicine (2009) Classification and identification of constitution theory of TCM (ZYYXH/T157-2009). World Journal of Traditional Chinese Medicine 4(4):303–304
6. Chollet F (2017) Xception: deep learning with depthwise separable convolutions. arXiv preprint: 1610.02357
7. Chollet F.: Keras (2015). <https://github.com/fchollet/keras>
8. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, In, pp 886–893
9. Deng J, Dong W, Socher R et al (2009) Imagenet: a large-scale hierarchical image database. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, In, pp 248–255

10. Ding YW (2010) Development of constitution theory in TCM. *Yunnan Journal of Traditional Chinese Medicine and Materia Medica* 2:71–75
11. Geng Y, Chen J, Fu R et al (2016) Enlighten wearable physiological monitoring systems: on-body rf characteristics based human motion classification using a support vector machine. *IEEE Trans Mob Comput* 15(3):656–671
12. He KM, Zhang X, Sun S et al (2016) Deep residual learning for image recognition. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, In, pp 770–778
13. He KM, Gkioxari G, Dollár P et al (2017) Mask r-cnn. *Proceedings of IEEE International Conference on Computer Vision*, In, pp 2980–2988
14. Hoo-Chang S, Roth HR, Gao M et al (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging* 35(5):1285
15. Howard AG, Zhu M, Chen B et al (2017) Mobilenets: efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*
16. Hu X, Zhu H, Xu J, Xu D et al (2014) Wrist pulse signals analysis based on deep convolutional neural networks. *Proceedings of IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology*, In, pp 1–7
17. Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, In, pp 7132–7141
18. Huan EY, Wen GH, Zhang SJ et al (2017) Deep convolutional neural networks for classifying body constitution based on face image. *Computational and Mathematical Methods in Medicine*
19. Huang G, Liu Z, Maaten VD et al (2017) Densely connected convolutional networks. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, In, pp 4700–4708
20. Huynh BQ, Li H, Giger ML (2016) Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. *Journal of Medical Imaging* 3(3):034501
21. Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*
22. Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, Dong J, Prasadha MK, Pei J, Ting MYL, Zhu J, Li C, Hewett S, Dong J, Ziyar I, Shi A, Zhang R, Zheng L, Hou R, Shi W, Fu X, Duan Y, Huu VAN, Wen C, Zhang ED, Zhang CL, Li O, Wang X, Singer MA, Sun X, Xu J, Tafreshi A, Lewis MA, Xia H, Zhang K (2018) Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 172(5):1122–1131
23. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* pp:1097–1105
24. Lan X, Ye M (2019) Shao R et al. A Robust RGB-Infrared Tracking System. *IEEE Transactions on Industrial Electronics, Learning Modality-Consistency Feature Templates*
25. Lan X, Ma AJ, Yuen PC, Chellappa R (2015) Joint sparse representation and robust feature-level fusion for multi-cue visual tracking. *IEEE Trans Image Process* 24(12):5826–5841
26. Lan X, Zhang S, Yuen PC et al (2017) Learning common and feature-specific patterns: a novel multiple-sparse-representation-based tracker. *IEEE Trans Image Process* 27(4):2022–2037
27. Law H, Deng J (2018) Cornernet: detecting objects as paired keypoints. *Proceedings of the Springer European Conference on Computer Vision*, In, pp 734–750
28. Li H, Xu B, Wang N et al (2016) Deep convolutional neural networks for classifying body constitution. *Proceedings of Springer International Conference on Artificial Neural Networks*, In, pp 128–135S
29. Li HH, Wen GH, Zeng HB (2018) Natural tongue physique identification using hybrid deep learning methods. *Multimedia Tools and Applications*, pp1–22
30. Lin JD, Lin JS, Chen LL et al (2012) BCQs: a body constitution questionnaire to assess stasis in traditional Chinese medicine. *European Journal of Integrative Medicine* 4(4):e379–e391
31. Lin JS, Chen LL, Lin JD (2012) BCQ-: a body constitution questionnaire to assess Yin-Xu. Part II: Evaluation of reliability and validity. *Forschende Komplementarmed* 19(6):285–292
32. Lin JD, Chen LL, Lin JS et al (2012) BCQ-: a body constitution questionnaire to assess Yin-Xu. Part I: establishment of a provisional version through a Delphi process. *Forschende Komplementarmedizin* 19(5): 234–241
33. Liu X, Wang Q (2013) Suggestion and analysis on revise of standard of classification and determination of constitution in TCM. *Beijing University of Chinese Medicine* 5:005
34. Liu S, Qi L, Qin H et al (2018) Path aggregation network for instance segmentation. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, In, pp 8759–8768
35. Ma JJ, Wen GH, Hu Y et al (2018) Tongue image constitution recognition based on complexity perception method. *arXiv preprint arXiv:1803.00219*
36. Masetic Z, Subasi A (2016) Congestive heart failure detection using random forest classifier. *Comput Methods Prog Biomed* 130:54–64

37. Rajpurkar P, Irvin J, Zhu K et al (2017) Chexnet: radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225
38. Samanthula BK, Elmehdwi Y, Jiang W (2015) K-nearest neighbor classification over semantically secure encrypted relational data. *IEEE Trans Knowl Data Eng* 27(5):1261–1273
39. scikit-learn (2018): machine learning in python. <http://scikit-learn.org/>
40. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
41. Su YC (2007) Establishment of traditional Chinese medical constitutional scale and classificatory index (2–1). *Yearbook of Chinese Medicine and Pharmacy* 25(5): 45–144
42. Su YC (2008) The creation of traditional Chinese medical constitutional scale and classification index (2–2). *Yearbook of Chinese Medicine and Pharmacy* 26(5):65–152
43. Su SY, Yang CH, Chiu CC, Wang Q (2013) Acoustic features for identifying constitutions in traditional Chinese medicine. *J Altern Complement Med* 19(6):569–576
44. Szegedy C, Vanhoucke V, Ioffe S et al (2016) Rethinking the inception architecture for computer vision. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, In, pp 2818–2826
45. Szegedy C, Ioffe S, Vanhoucke V et al (2017) Inception-v4, inception-resnet and the impact of residual connections on learning. *AAAI Conference on Artificial Intelligence* 4:12
46. Tajbakhsh N, Shin JY, Gurudu SR et al (2016) Convolutional neural networks for medical image analysis: full training or fine tuning. *IEEE transactions on medical imaging* 35(5):1299–1312
47. Tan M, Le QV (2019) EfficientNet: rethinking model scaling for convolutional neural networks. arXiv preprint arXiv:1905.11946
48. TensorFlow (2018). <https://www.tensorflow.org>
49. Van OA, Ikram MA, Vernooij MW et al (2015) Transfer learning improves supervised image segmentation across imaging protocols. *IEEE Trans Med Imaging* 34(5):1018–1030
50. Viola P, Jones MJ (2004) Robust real-time face detection. *Int J Comput Vis* 57(2):137–154
51. Walt SD, Schönberger JL, Nunez-Iglesias J et al (2014) Scikit-image: image processing in python. *PeerJ* 2: e453
52. Wang Q (2005) Classification and diagnosis basis of nine basic constitutions in Chinese medicine. *Journal-Beijing University of Traditional Chinese Medicine* 28(4):1
53. Wang Q (2006) Three key issues in the study of TCM constitution (Part I). *J Tradit Chin Med* 4:250–252
54. Wang Q (2006) Three key issues in the study of TCM constitution (Part II). *J Tradit Chin Med* 5:329–332
55. Wang YC, Bai LN (2014) Classification of body constitution of pulse signal in TCM based on BP neural network. *J Tradit Chin Med* 55(15)
56. Wang F, Jiang M, Qian C et al (2017) Residual attention network for image classification. arXiv preprint arXiv:1704.06904
57. Wong W, Lam CK, Su YC, Lin SJ, Ziea ET, Wong VT, Wai LK, Kwan AK (2014) Measuring body constitution: validation of the body constitution questionnaire (BCQ) in Hong Kong. *Complementary therapies in medicine* 22(4):670–682
58. Woo S, Park J, Lee JY et al (2018) Cbam: convolutional block attention module. In: *Proceedings of the European Conference on Computer Vision*, pp3–19.
59. Xie S, Girshick R, Dollár P et al (2017) Aggregated residual transformations for deep neural networks. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, In, pp 5987–5995
60. Yu RX, Wang Q, Wang J et al (2013) An analysis of the status quo of application of constitution identification. *Chinese Journal of Information on Traditional Chinese Medicine* 2:107–109
61. Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks. *Proceedings of Springer European Conference on Computer Vision*, In, pp 818–833
62. Zhang X, Li Z, Loy CC et al (2017) Polynet: a pursuit of structural diversity in very deep networks. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, In, pp 3900–3908
63. Zhang SF, Wen LY, Bian X et al (2017) Single-shot refinement neural network for object detection. arXiv preprint arXiv:1711.06897

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.