Scalable Skin Lesion Multi-Classification Recognition System

Fan Liu¹, Jianwei Yan², Wantao Wang², Jian Liu^{2, *}, Junying Li³ and Alan Yang⁴

Abstract: Skin lesion recognition is an important challenge in the medical field. In this paper, we have implemented an intelligent classification system based on convolutional neural network. First of all, this system can classify whether the input image is a dermascopic image with an accuracy of 99%. And then diagnose the dermoscopic image and the non-skin mirror image separately. Due to the limitation of the data, we can only realize the recognition of vitiligo by non-skin mirror. We propose a vitiligo recognition based on the probability average of three structurally identical CNN models. The method is more efficient and robust than the traditional RGB color space-based image recognition method. For the dermoscopic classification model, we were able to classify 7 skin lesions, use weighted optimization to overcome the unbalanced data set, and greatly improve the sensitivity of the model by means of model fusion. The optimization and expansion of the system depend on the increase of database.

Keywords: Skin disease, neural networks, ensemble, classification.

1 Introduction

Human skin is one of the most important organs of the human body. From the external environment, the skin is spread over the surface of the human body, accounting for about 15% of the adult's weight [Kanitakis (2002)]. Because the human body has the natural defense barrier of the skin, preventing most adverse effects such as bacteria and viruses, but skin diseases can threaten the safety of human skin. More than 5 million cases occur each year. A large part of these 5 million cases are caused by psoriasis, acne, and eczema. Caused by vitiligo, melanoma, etc., some of these skin diseases can be particularly fatal, such as melanoma [Pathan, Prabhu and Siddalingaswamy (2018)], and some may have a great impact on human appearance, such as vitiligo. Skin diseases also have differences in size. Some skin diseases such as vitiligo, psoriasis, etc. are directly visible to the human eye, convenient for observation and diagnosis, and some skin patients are difficult to observe, which requires physical means to enlarge, the appearance of dermatoscope makes this type Examination of lesions has become possible [Tschandl and Wiesner (2018)]. In short, the appearance of skin diseases can have a great impact on people's

¹ The First Clinical Medical College, Nanchang University, Nanchang, 330031, China.

² School of Computer & Communication Engineering, University of Science and Technology Beijing, Beijing, 100083, China.

³ Gu'an County People's Hospital, Langfang, 065000, China.

⁴ Amphenol AssembleTech, Houston, TX 77070, US.

^{*}Corresponding Author: Jian Liu. Email: liujian@ustb.edu.cn.

appearance or health. Therefore, these skin diseases should be taken seriously, and the spread of skin diseases should be detected and prevented as soon as possible.

1.1 Related works

In order to assist doctors to better detect and diagnose skin diseases, traditional image classification methods in medical image processing identify these diseases through preprocessing, edge segmentation, feature extraction and classifiers. Shrivastava et al. [Shrivastava, Londhe, Sonawane et al. (2017)] builds a segmentation based on segmentation. The Bayesian model extracts the features of the lesion image, and finally uses the SVM and FDR combination classification to achieve 99.84% of the psoriasis recognition rate. George et al. [George, Aldeen and Garnavi (2017)] performed five color space conversions and selected the voting results as the final diagnosis, which resulted in a higher precision psoriasis recognition rate, Alam et al. [Alam, Munia, Tavakolian et al. (2016)] based on the color-segmented eczema image, using the k-means and morphology. The image processing method was obtained, and the result was 90% accurate. Nezhadian et al. [Nezhadian and Rashidi (2017)] attempted to extract the texture features of melanoma lesion images and combined with support vector machines as classifiers for benign and malignant melanoma classification. Manerkar et al. [Manerkar, Snekhalatha, Harsh et al. (2016)] uses C-means and watershed algorithm to segment skin lesion images, and then uses gray level co-occurrence matrix algorithm to extract data features, finally he selects Support Vector Machine as the classifier. To achieve a good effect, Albay et al. [Albay, Kama and Sak (2015)] extracts the Fourier properties of the lesion boundary after segmentation of the skin lesion and serves as a classification feature.

However, traditional image processing techniques also have some problems. In order to better feature extraction, a large amount of expert experience is needed to adjust the threshold in the segmentation stage. In addition, the classification system has poor scalability, and the system for detecting certain diseases cannot be simply used to detect other. Diseases, with the rapid development of deep learning, breakthroughs have been made in the field of image recognition. Convolutional neural networks (CNN) Krizhevsky et al. [Krizhevsky, Sutskever and Hinton (2012)] have strong feature extraction performance, and training image data using convolutional neural network models has become a lot of the first choice for research. Ge et al. [Ge, Demyanov, Bozorgtabar et al. (2017)] used skin data models with different depth neural networks, then used bilinear pooling techniques for some models, and finally used support vector machines as classifiers to get better results. Deep learning can be used for segmentation in addition to classification. Pal et al. [Pal, Garain, Chandra et al. (2018)] used U-shaped complete convolutional neural network (FCN) to train psoriasis segmentation images, and compared with traditional image segmentation techniques, 88.5% Jaccard values were obtained. Yu et al. [Yu, Chen, Dou et al. (2017)] constructed a Complete Convolutional Residual (FCRN) network to connect a residual network to identify skin cancer, and the previous network was used to separate skin lesions. The latter network layer is very deep, used for classification, and the final classification result is 94.9% accurate.

In this article, we want to implement a scalable skin lesion classification system based on convolutional neural networks. Because of the limited number and variety of data sets,

this means that we cannot achieve accurate classification of all skin diseases. However, this skin lesion classification system is scalable due to the mobility of the convolutional neural network.

Skin lesion class	Number per class		
Vitiligo	30000		
Normal	8677		
Melanoma	1113		
Melanocytic Nevus	6705		
Basal Cell Carcinoma	514		
Actinic Keratosis/Bowen's diease	327		
Benign keratosis	1099		
Dermatofibroma	115		
Vascular lesion	142		

Table 1: The number of skin lesion dataset



Figure 1: Representative non-dermoscopic (left) and dermo-scopic (right)

1.2 Organization

The structure of this paper is as follows. In Section 2.1, we discussed the skin lesion dataset. In Section 2.2, we introduced the conditions of the device, pre-trained the model,

and preprocessed the dataset. we achieve the division of dermoscopic lesion images and non-skin mirrors. In Section 3.1, we diagnose normal picture data and vitiligo lesion data for non-skin mirrors. In Section 3.2, we diagnose multiple types of dermoscopic lesion images. In Section 4, we show our system experiment results, in Section 5, we summarize the full text and put forward the outlook.

2 Prepared

2.1 Datasets

In this article, we use a database of vitiligo from the hospital, a total of 38,677, image size of 768×576 , including vitiligo and normal images, and dermatoscope dataset based on ISIC 2018: The Great Challenge Dataset for Skin Lesion Analysis to Melanoma Detection [Codella, Gutman, Celebi et al. (2018), Tschandl, Rosendahl and Kittler (2018)], a total of 10015, the image size is 450×600 , includes 7 types of skin lesions such as melanoma, as shown in Fig. 1. The distribution of each data set is shown in the Tab. 1.

For the dermatoscope dataset, it can be seen from the table that the amount of data between classes and classes is extremely uneven, and the imbalance of data volume may result in the model not learning the characteristics of the class with less data and the wrong classification. The problem we are going to try to solve.

2.2 Condition

First of all, our system needs to divide the skin mirror image and the non-skin mirror image. We will set the training set and test set according to 8:2. Because the image sizes of the two data sets are different, we will scale the size to 224×224 uniformly. The pre-trained model on ImageNet uses a batch size of 20 training models on 4 TITAN X. In the data preprocessing section, we performed random horizontal/vertical flipping of the image, normalization, luminance perturbation, and our learning rate was set to 0.0001, a total of 20 batches were used in the divided training.

Name	SEN	ACC	PPV	SPE	AUC
Xception	0.880	0.845	0.900	0.758	0.911
Iceptionv3	0.884	0.846	0.899	0.752	0.912
Vgg16	0.895	0.849	0.892	0.732	0.917
Resnet50	0.896	0.861	0.908	0.775	0.922

 Table 2: Percentage of correct classification (RGB image)

3 Expertiments

In our multi-dermatology classification system, we first use a two-classification model to determine whether the image is dermoscopic or non-dermoscopic. If it is a dermoscopic picture, it will be sent to the dermoscopic classification model of multiple diseases, otherwise it will be sent to the vitiligo classification model. In this paper, Resnet20 is

804

used to distinguish dermoscopic and non-dermoscopic pictures. The result is obvious, the accuracy is 99%, and the robustness is very good. Next, we introduce the implementation of the two models. The process of building the system is shown in Fig. 2.

3.1 Vitiligo

According to incomplete statistics, the incidence of vitiligo in the world is about 0.5-2% [Krüger and Schallreuter (2012); Lotti and D'Erme (2014)]. The vitiligo is a common pigmented skin disease characterized by local or generalized loss of pigmentation and white spot formation. when vitiligo is still at an early stage, the number of vitiligo spots is small, the treatment effect is significant, the harm to human body is small, the treatment cost is low. Early diagnosis of vitiligo is very important, so automatic diagnosis of vitiligo is of great significance. In this paper, we propose a method base on three identical CNN models, trained with three different color-space images (RGB, HSV, and YCrCb) for the same vitiligo dataset. The result is better than a single model.

In the classification system of vitiligo, we use CNN model. The structure of CNN usually consists of two parts: automatic feature extraction and full connection layer, in which feature extraction includes convolution layer and pooling layer. We can get a good classification model without through to manually selecting features, this is the advantage of CNN. The Keras-Framework [Chollet (2018)] is used in this experiment. It has many pre-trained models, which can shorten the development time. This framework shows the powerful function of CNN and its simplicity. We choose four commonly (Resnet50, Vgg16, Xception, Iceptionv3) [Litjens, Kooi, Bejnordi et al. (2017)] used CNN models. After preliminary experiments, the Resnet50 has many indicators are better than other models, which is more suitable for classification of vitiligo data as shown in Tab. 2. This is because Resent model has residual module which can make the whole structure converge in the direction of identical mapping and ensure that the final error rate does not become worse and worse as the depth increases. Therefore, Resnet50 is used as the baseline model which is shown in Fig. 3. The formula of residual module as follows:

$$y = W_i \varphi(W_{i-1}x) + x \tag{1}$$

where φ represent the relu function, W_i represent the weight layer. The predictive value of vitiligo is averaged from the output proba-bilities of the three models, and the vitiligo recognition algorithm in detail as follows.

Step 1: Convert the RGB image into HSV and YCrCb image.

$$r = \frac{R}{255}, g = \frac{G}{255}, b = \frac{B}{255}, Max = max(r, g, b), Min = min(r, g, b)$$

YCrCb:

$$Y = 0.299*R + 0.587*G + 0.114*B$$

$$Cr = 0.5*R - 0.4187*G - 0.0813*B + 128$$

$$Cb = -0.1687*R - 0.3313*G + 0.5*B + 128$$

HSV:
(2)

$$S = \begin{cases} 0, \text{Max} = 0\\ \frac{Max - Min}{Max}, \text{Otherwise} \end{cases}$$
(3)

$$V = Max \tag{4}$$

$$H = \begin{cases} 0^{\circ}, Max = Min \\ 60^{\circ} * \frac{g - b}{Max - Min} + 0^{\circ}, Max = r \& g \ge b \\ 60^{\circ} * \frac{g - b}{Max - Min} + 360^{\circ}, Max = r \& g < b \\ 60^{\circ} * \frac{b - r}{Max - Min} + 120^{\circ}, Max = g \\ 60^{\circ} * \frac{r - g}{Max - Min} + 240^{\circ}, Max = b \end{cases}$$
(5)

Step 2: We will use the data preprocessing method mentioned in the previous section to normalize the image and enhance the data, and then send the processed image to the Keras-Resnet50 model for training.

Step 3: Averaging the probability values of the corresponding classes output (F1(x), F2(x), F3(x)) from the three models, where x is stand for the input data.

$$F(x_i) = \frac{F1(x_i) + F2(x_i) + F3(x_i)}{3}, i = 1, 2, 3, \dots N$$
(6)

Step 4: Diagnosis of vitiligo. We judge the image with a confidence greater than 0.5 as vitiligo.

Results =
$$\begin{cases} Vitiligo, F(x_i) \ge 0.5, i = 1, 2, 3, \dots N \\ Normoal, F(x_i) < 0.5, i = 1, 2, 3, \dots N \end{cases}$$
(7)

We can see that the images in three color spaces and the trained models have similar indexes, but through the method of mixing, the indexes have been improved.

3.2 Dermoscopic

For relatively small skin lesions, it is difficult for doctors to detect by naked eye observation, and it is easy to misdiagnose. Because of the appearance of dermoscopy, doctors can diagnose as early as possible and prevent it in time. This paper designs a CNN model for assistant doctors, which can classify seven skin lesions under dermoscopy and reduce the misdiagnosis rate as much as possible.

806





Figure 3: The structure of Resnet50 model

In the dermatoscopy classification system, we divide the data set into 8:1:1, which is divided into training set, verification set and test set. A total of 150 rounds of batches are trained. The other data preprocessing methods are the same as before, due to the skin mirror data. The amount of data in the set of 7 types is very different, so we use the weighted loss optimizer of the unbalanced class, and the weight is inversely proportional to the proportion of the class, so as to achieve the model to learn the characteristics of all classes equally.

Firstly, we mainly choose the following five models ResNet101, DenseNet201, SENet154, SE-Resnext101, and DPN68b [Huang, Liu, Van Der Maaten et al. (2017); Hu, Shen and Sun (2017); Xie, Girshick, Dollár et al. (2017); Chen, Li, Xiao et al. (2017)]. These models

are selected to have a large depth and the parameter quantity is also large. Many of them help the model to learn better, and different models have different learning effects. The fusion model obtained later combines the advantages of various models, and the recognition accuracy will be higher. Taking Resnet101 as an example, the model is finetuned, and the zero-initialization method is applied to the BatchNorm layer in each residual module. The accuracy of the model is slightly improved. In addition, we turn the fullyconnected layer into a multi-layer perceptron structure, which leads to a significant increase in sensitivity.

Next, in the validation set and test set, we use the data expansion method, and we use multiple cropping to obtain N images of different positions for validation, and average N results, we use the following formula to cut N copies around the original image center.

$$\begin{cases} P_{[k,0]} = \frac{x (m-x)^* i}{2! \sqrt{N-1}}, i \in [0, \cdots, \sqrt{N}-1] \\ P_{[k,1]} = \frac{y (n-y)^* j}{2! \sqrt{N-1}}, j \in [0, \cdots, \sqrt{N}-1] \end{cases}$$
(8)

where $(P_{[k,0]}, P_{[k,1]})$ coordinate value represents the center coordinates of the crop, \$N\$ represents the number of crops, (x, y)\$ represents the size you want to crop, (m, n)represents the size of the original image, so the cropped image position for validation and test set is:

$$BOX_{k} = \left(\frac{2P_{[k,0]} - x}{2}, \frac{2P_{[k,1]} - y}{2}, \frac{2P_{[k,0]} + x}{2}, \frac{2P_{[k,1]} + y}{2}\right)$$
(9)

We average the cropped N pictures as the result, by using the following formula:

$$\overline{pred} = \frac{\sum_{k=0}^{N-1} pred[k]}{N}, \qquad k \in [0, \cdots, N-1]$$
(10)

We compare N to none, 16, 36 and 64 cases, where the value of 36 is the best. Using this method, the performance is improved a lot compared to the various indicators obtained by using a single random cut. By cutting the validation set and the test set around the center into multiple copies, 36 images of different positions are obtained. The randomness obtained by multiple cropping is reduced, and the lesion image recognition is more accurate for some image lesions.

According to this method, we obtain various metrics for different models on the validation set and test set, and calculate the mean. The results of DenseNet201, ResNet101, SENet154, SE-Resnext101, DPN68b and ResNet101MLP are obtained in turn, as Tab. 3 shows. From the results of these models, the best overall model is SENet154, the model with the best sensitivity is the resnet101 model with MLP structure which is shown in Fig. 4. However, we hope to get higher sensitivity results at the same time. we use model ensemble to combine the advantages of each model to get better results.

Finally, for the results of validation set, we adopt an approach of ensemble combined

with search strategy of Bayesian optimization. From the above model, we select the following five models for the ensemble, including DenseNet201, ResNet101, SENet154, SE-Resnext101, DPN68b and ResNet101MLP. The vectors are extracted before passing the softmax module, we connect all the vectors from the validation set through these five models, and train these vectors with RandomForest as a classifier. In the field of hyperparameter automatic search, there are common grid search, random search and beyesian optimization search, but the first two searches are not efficient, so we adopt bayesian optimization search, we will introduce the implementation process of Bayesian optimization Algorithm 1.

Algorithm 1 Sequential Model-Based Optimization

Input: f, X, S, M1: $D \leftarrow InitSamples(f, X)$ 2: for $i \leftarrow |D|$ to T do 3: $p(y|x, D) \leftarrow FitModel(M, D)$ 4: $x_i \leftarrow argmax_{x \in X}S(x, p(y|x, D))$ 5: $y_i \leftarrow f(x_i)$ 6. $D \leftarrow D \cup (x_i, y_i)$ 7: end for where f is the unknown function relationship

where f is the unknown function relationship, X is the input data, S is acquisition Function, M is model based on input data hypothesis, firstly we get the initialized data set based on the input data, then make a loop to select T times parameters. The model we chose is based on Gaussian distribution, the mean μ and covariance $K(x, x^*)$ of a Gaussian function are fixed, formulated as follows: $f \sim GP(\mu, K)$, when the Gaussian process is used as a priori for Bayesian inference, the posterior function can be used to predict new data, we suppose y is a function value known by training data, y_* is the function value of the test set input x_* , $\overline{\mu}$ is the mean of training set, $\overline{\mu}_*$ is the mean of test set, \sum_* is the covariance of the training set, \sum_{**} is the covariance of the test set. Eq. (5) shows the relationship.

$$\begin{bmatrix} f \\ f_* \end{bmatrix} \sim N(\begin{bmatrix} \overline{\mu} \\ \overline{\mu}_* \end{bmatrix}, \begin{bmatrix} \Sigma & \Sigma_* \\ \sum_{*}^T & \sum_{**} \end{bmatrix})$$
(11)

Gaussian process extends multivariate Gaussian distribution to infinite dimension, a training set y can be represented as a sample taken from a multivariate Gaussian distribution: $y = [y_1, y_2, \dots, y_n]^T$. We set the mean of the Gaussian process to 0 and the

most common choice for covariance is the squared exponential, See Eq. (6):

$$k(x, x') = \sigma_f^2 \exp(\frac{-(x - x')^2}{2l^2})$$
(12)

Due to the existence of noise, we express the formula as Eqs. (7), (8):

$$y = f(x) + N(0, \sigma_n^2) \tag{13}$$

$$k(x, x') = \sigma_f^2 \exp(\frac{-(x - x')^2}{2l^2}) + \sigma_n^2 \sigma(x, x')$$
(14)

where $\sigma(x, x')$ is the KroneckerDelta function, in addition to calculating the covariance of the training set K, see Eq. (9), we also need to calculate the covariance between the new independent variable and the training set independent variable K_* Eq. (10) and the covariance of the new independent variable K_{**} Eq. (11). Then Eq. (12) shows the relationship

$$K = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \cdots & k(x_n, x_n) \end{bmatrix}$$
(15)

$$K_* = [k(x_*, x_1), k(x_*, x_2), \cdots, k(x_*, x_n)]$$
(16)

$$K_{**} = k(x_*, x_*) \tag{17}$$

$$\begin{bmatrix} y \\ y_* \end{bmatrix} \sim N(0, \begin{bmatrix} K & K_*^T \\ K_* & K_{**} \end{bmatrix})$$
(18)

Table 3: Validation/Test set pass models' mean metrics

Name	SEN	ACC	F1	SPE	AUC
DenseNet201	0.828/0.817	0.868/0.862	0.862/0.861	0.972/0.969	0.976/0.974
ResNet101	0.825/0.819	0.862/0.860	0.859/0.860	0.971/0.968	0.967/0.971
SENet154	0.855/0.800	0.886/0.881	0.884/0.880	0.976/0.971	0.980 /0.970
SE-Resnext101	0.844/0.790	0.881/0.872	0.881/0.874	0.972/0.967	0.976/0.970
DPN68b	0.828/0.801	0.861/0.852	0.860/0.851	0.968/0.964	0.973/0.970
ResNet101MLP	0.859/0.831	0.814/0.805	0.808/0.791	0.968/0.962	0.972/0.975



Figure 4: Resnet50 model structure (add multilayer perceptron)

The training set obeys a multidimensional normal distribution, according to K ,we can know that the posterior probability of the test set y_* is $p(y_* | y \sim N(\overline{\mu}, \hat{\sigma}^2))$, the mean $\overline{\mu}$ and variance $\hat{\sigma}^2$ of y_* are expressed as follows Eqs. (19)-(20):

$$\overline{\mu} = K_* K^{-1} y \tag{19}$$

$$\hat{\sigma}^2 = K_{**} - K_* K^{-1} K_*^T \tag{20}$$

The hyperparameter to be determined is $\theta = [\sigma_f^2, l]$, since the training set obeys a multidimensional normal distribution, the likelihood function is Eq. (21):

$$L = \log p(y | x, \theta) = -\frac{1}{2} \log |K| - \frac{1}{2} (y - \mu)^T K^{-1} (y - \mu) - n \log(2\pi) / 2$$
(21)

Bayesian optimization maps x to the real space R through the Acquisition function, indicating the probability that the objective function value of the point can be larger than the current optimal value. The two main types of acquisition functions are commonly used, the first is probability of improvement, see Eq. (22).

$$POI(X) = P(f(X) \ge f(X^{+}) + \xi) = \Phi(\frac{\mu(x) - f(X^{+}) - \xi}{\sigma(x)})$$
(22)

where f(x) is the value of the X objective function, $f(x^{+})$ is the optimal X objective function value so far, $\mu(x)$, $\sigma(x)$ are the mean and variance of the objective function obtained by the Gaussian process, respectively, ξ is the trade-off factor which adjust to select the points around the X^{+} . In general, we use MonteCarlo simulation

method to find X so that POI(X) is the largest.

The second is expected improvement. The POI is a probability function, so only the probability that f(x) is larger than $f(x^{+})$ is considered, and expected improvement is a desired function, so it is considered how much f(x) is larger than $f(x^{+})$. We get x by the following Eq. (23).

$$x = \arg\max_{x} E(\max\{0, f_{t+1}(x) - f(X^{+})\} | D_{t})$$
(23)

where Dt is the first t samples, under the premise of normal distribution, we can get the following Eq. (24): expression $(Z = \frac{\mu(x) - f(x^{+})}{\sigma(x)})$: $\int (\mu(x) - f(x^{+}))\Phi(Z) + \sigma(x)\phi(Z) \text{ if } \sigma(x) > 0$

$$EI(x) = \begin{cases} (\mu(x) - f(x^{+}))\Phi(Z) + \sigma(x)\phi(Z), & \text{if } \sigma(x) > 0\\ 0, & \text{if } \sigma(x) = 0 \end{cases}$$
(24)

For the classifier model, we use SVM and RandomForest, SVM configuration is as follows, we use 10 fold cross-validation. For unbalanced data sets, we also use class balance weighted, we use sensitivity as the main evaluation indicator. Search parameters include C and kernel, the best performance parameters are thefollowing values: C=[0.1, 1000], kernel=['linear', 'poly','rbf', 'sigmoid']. The parameter control for the SVM classifier is mainly from the following formula:

According to validation set vectors $x_i \in \mathbb{R}^p$, $i = 1, \dots, n$ and a vector $y \in \{1, -1\}^n$, the SVM formula can be expressed as Eq. (25):

$$\begin{cases} tar = \min_{\omega,b,\zeta} \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n \zeta_i, \\ y_i(\omega^T \phi(x_i) + b) \ge 1 - \zeta_i, \end{cases}$$
(25)

Due to the duality, this formula is equivalent to the following Eq. (26):

$$\begin{cases} tar = \min_{\alpha} \frac{1}{2} \alpha^{T} Q \alpha - e^{T} \alpha, \\ subject to \quad y^{T} \alpha = 0, \end{cases} \quad a \le \alpha_{i} \le C, \ j = 1, \cdots, n$$
(26)

where e is the vector of all ones, Q is an n by n positive semidefinite matrix. $Q_{ij} = y_i y_j K(x_i x_j)$. Validation set vectors are implicitly mapped into a higher dimensional space by the function ϕ . The decision function is Eq. (27):

$$dec = sgn(\sum_{i=1}^{n} y_i \alpha_i K(x_i, x) + \rho)$$
(27)

where $K(x_i x_j) = \phi(x_i)^T \phi(x_i)$ is the kernel. C > 0 is the upper bound, we change the value to get best results.

We consider the classification model. The random forest model is a combination of the

bagging model plus decision trees which creates multiple subtrees by splitting features. The difference is that decision trees usually generate nodes and rules by calculating the information gain and the Gini index. In contrast, random forests are random. Deeper decision trees tend to have over-fitting problems, while random forests can prevent most situations by creating random subsets of features and using them to build smaller trees, which then form subtrees, this method can prevent overfitting in most cases.

In this experiment, we select the hyperparameter range to include the following values: N(the number of trees in the forest)=[10, 500], MinSamplesSplit (the minimum number of samples required to split an internal node)=[2, 100], MaxFeatures (the number of features to consider when looking for the best split)=[0.1, 0.999], MaxDepth (the maximum depth of the tree)=[5, 80], to find the best sensitivity value by Bayesian optimization combined with random forest model.

4 Results

The performance of the classification is evaluated in terms of accuracy (ACC), precision (PPV), sensitivity (SEN), specificity (SPE) and the Area Under ROC Curve (AUC) [Powers (2011)]. The respective definitions of these common metrics adopting true positive (TP), true negative (TN), false positive (FP), and false negative (FN) are as follows: ACC= (TP+TN)/(TP+TN+FP+FN), SEN=TP/(TP+FN), SPE=TN/(TN+FP), PPV=TP/(TP+FP).

For vitiligo classification model, we can see that the specificity of these three models is very low, no more than 80%, however the combination of different color spaces reach SPE=0.802. The vitiligo is a pigmented disease, and lesion areas have high contrast under different color space transformation, the fusion method can effectively recognition effect of lesion areas, reduce the misdiagnosis rate. At the same time, the indicators of the classification model are ACC=0.878, PPV=0.919, SEN=0.909, AUC=0.942 (see Tab. 4).

Name	SEN	ACC	PPV	SPE	AUC
RGB	0.896	0.861	0.908	0.775	0.922
HSV	0.896	0.863	0.910	0.780	0.929
YCrCb	0.889	0.860	0.913	0.789	0.927
Average	0.909	0.878	0.919	0.802	0.942

Table 4: Performance of vitiligo classification

In conclusion, our method is simple effective and low complexity, that can be better applied in the assistant of doctor's diagnosis, also can provide a research idea for other pigmented diseases. For another model, we use SVM classifier to search for hyperparameters C and kernel, the best performing set of coefficients is C=10 and kernel='rbf', the result of the test set after passing the best set of SVM classifiers is SEN=0.834, ACC=0.842, F1=0.837, SPE=0.966, AUC=0.900. For random forest classifier, the best performing set of coefficients is N=255, MaxFeatures=0.1385, MaxDepth=75 and MinSamplesSplit=100, the result of the test set after passing the best set of random forest classifier is SEN=0.846, ACC=0.840, PPV=0.833, SPE=0.968, AUC=0.907. The correct rates of seven classifications (melanoma, nevus, basal cell carcinoma, actinic keratosis/bowen's diease, benign keratosis, dermatofibroma, vascular lesion) are 0.651, 0.884, 0.942, 0.914, 0.675, 0.922 and 0.993 respectively. The confusion matrix of the two classification models is shown in Fig. 5.



Figure 5: The Confusion Matrix of Vitiligo Model(a) and Dermoscopic Model(b)

5 Discussion

In this work, we implement a multi-dermatological classification system. We first classify dermoscopic and non-skin mirrors (vital) and then construct models of dermoscopic and non-skin mirrors, respectively. For the vitiligo model, we propose an approach to recognition of vitiligo based on probability-average value of three CNN models which are same structures, compared with traditional images recognition methods based on RGB color space, the propose method is more effective and robust. For the dermoscopic model, we use weighted loss optimization for unbalanced data. In the validation and test set, multiple clipping methods are used. Finally, the validation set is connected through vectors of multiple models, and Bayesian optimization is used to search for hyperparameters to train a performance-optimized sensitivity classifier. The research on the classification of skin lesions is worthy of more trials, there are still many shortcomings in our work. In the future, we will do further research and make a progress on classification methods, and we are committed to helping doctors reduce the fatigue caused by diagnosis.

Acknowledgement: This work is supported by 2018 Sugon Intelligent-Factory on Advanced Computing Devices (No. MIIT2018-265-137).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Alam, M. N.; Munia, T. T. K.; Tavakolian, K.; Vasefi, F.; MacKinnon, N. et al. (2016): Automatic detection and severity measurement of eczema using image processing.

38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1365-1368.

Albay, E.; Kamaşak, M. (2015): Skin lesion classification using fourier descriptors of lesion borders. *Medical Technologies National Conference*, pp. 1-4.

Chen, Y.; Li, J.; Xiao, H.; Jin, X.; Yan, S. et al. (2017): Dual path networks. *Advances in Neural Information Processing Systems*, pp. 4467-4475.

Chollet, F. (2018): *Keras: The Python Deep Learning Library*. Astrophysics Source Code Library.

Codella, N. C.; Gutman, D.; Celebi, M. E.; Helba, B.; Marchetti, M. A. et al. (2018): Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC). *IEEE 15th International Symposium on Biomedical Imaging*, pp. 168-172.

Ge, Z.; Demyanov, S.; Bozorgtabar, B.; Abedini, M.; Chakravorty, R. et al. (2017): Exploiting local and generic features for accurate skin lesions classification using clinical and dermoscopy imaging. *IEEE 14th International Symposium on Biomedical Imaging*, pp. 986-990.

George, Y.; Aldeen, M.; Garnavi, R. (2017): A pixel-based skin segmentation in psoriasis images using committee of machine learning classifiers. *Digital Image Computing: 2017 International Conference on Techniques and Applications*, pp. 1-8.

Hu, J.; Shen, L.; Sun, G. (2017): Squeeze-and-excitation networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 7, pp. 7132-7141.

Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K. Q. (2017): Densely connected convolutional networks. *CVPR*, vol. 1, pp. 4700-4708.

Kanitakis, J. (2002): Anatomy, histology and immunohistochemistry of normal human skin. *European Journal of Dermatology*, vol. 12, no. 4, pp. 390-401.

Krizhevsky, A.; Sutskever, I.; Hinton, G. E. (2012): Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, pp. 1097-1105.

Krüger, C.; Schallreuter, K. U. (2012): A review of the worldwide prevalence of vitiligo in children/adolescents and adults. *International Journal of Dermatology*, vol. 51, no. 10, pp. 1206-1212.

Litjens, G.; Kooi, T.; Bejnordi, B. E.; Setio, A. A. A.; Ciompi, F. et al. (2017): A survey on deep learning in medical image analysis. *Medical Image Analysis*, vol. 42, pp. 60-88.

Lotti, T.; D'Erme, A. M. (2014): Vitiligo as a systemic disease. *Clinics in Dermatology*, vol. 32, no. 3, pp. 430-434.

Manerkar, M. S.; Snekhalatha, U.; Harsh, S.; Saxena, J.; Sarma, S. P. et al. (2016): Automated skin disease segmentation and classification using multi-class SVM classifier. *Electrical, Electronics, Engineering Trends, Communication, Optimization and Sciences, 3rd International Conference*, pp. 1-6.

Nezhadian, F. K.; Rashidi, S. (2017): Melanoma skin cancer detection using color and

new texture features. Artificial Intelligence and Signal Processing Conference, pp. 1-5.

Pal, A.; Garain, U.; Chandra, A.; Chatterjee, R.; Senapati, S. (2018): Psoriasis skin biopsy image segmentation using deep convolutional neural network. *Computer Methods and Programs in Biomedicine*, vol. 159, pp. 59-69.

Pathan, S.; Prabhu, K. G.; Siddalingaswamy, P. (2018): Techniques and algorithms for computer aided diagnosis of pigmented skin lesions-a review. *Biomedical Signal Processing and Control*, vol. 39, pp. 237-262.

Powers, D. M. (2011): Evaluation: from precision, recall and f-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, vol. 2, pp. 37-63.

Shrivastava, V. K.; Londhe, N. D.; Sonawane, R. S.; Suri, J. S. (2017): A novel and robust bayesian approach for segmentation of psoriasis lesions and its risk stratification. *Computer Methods and Programs in Biomedicine*, vol. 150, pp. 9-22.

Tschandl, P.; Rosendahl, C.; Kittler, H. (2018): The ham10000 dataset: a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, vol. 5.

Tschandl, P.; Wiesner, T. (2018): Advances in the diagnosis of pigmented skin lesions. *British Journal of Dermatology*, vol. 178, no. 1, pp. 9-11.

Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. (2017): Aggregated residual transformations for deep neural networks. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5987-5995.

Yu, L.; Chen, H.; Dou, Q.; Qin, J.; Heng, P. A. (2017): Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Transactions on Medical Imaging*, vol. 36, no. 4, pp. 994-1004.