IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY SECTION

Received April 17, 2021, accepted June 24, 2021, date of publication June 29, 2021, date of current version July 6, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3093358

Automatic Diabetic Foot Prediction Through Fundus Images by Radiomics Features

XIAOKANG LIANG[1,2,3], EBTESAM N. ALSHEMMARY[4], MING MA[5], SHENGWU LIAO[6], WEI ZHOU[6], AND ZHENTAI LU[1,2,3]

[1]School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China
[2]Guangdong Provincial Key Laboratory of Medical Image Processing, Southern Medical University, Guangzhou 510515, China
[3]Guangdong Province Engineering Laboratory for Medical Imaging and Diagnostic Technology, Southern Medical University, Guangzhou 510515, China
[4]IT Research and Development Center, University of Kufa, Najaf 540011, Iraq
[5]Department of Ophthalmology, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong 510515, China
[6]Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong 510515, China

Corresponding author: Zhentai Lu (13422322117@163.com)

This work was supported in part by the Natural Science Foundation of Guangdong Province under Grant 2016A030313574, in part by the Science and Technology Planning Project of Guangdong Province under Grant 2020A1414040021, and in part by the Science and Technology Planning Project of Guangzhou City under Grant 202103000037.
This work did not involve human subjects or animals in its research.

ABSTRACT Current clinical approaches to diabetic foot (DF) treatment mainly rely on clinician vigilance and laboratory test, which have significant limitations, such as the high cost involved in the diagnosis and the high demands for professional skills of clinicians. At present, the research on DF prediction has mainly focused on the regression analysis of clinical data and the recognition based on foot ulcers skin. In view of this situation, we examine the patients' fundus images to explore an efficient way for DF prediction. In this paper, we have proposed a DF prediction model through fundus images by radiomics features. Twelve kinds of radiomics features are extracted, including a variety of features first applied in the field of medical imaging, describing the information of image texture, direction, phase, and gradient. Subsequently, a two-step feature selection model is put forward for a large number of radiomics features we extracted for the purpose of searching for the best combination. Considering the simplicity and performance of the model, we chose 19 features to train the support vector machine model. The obtained model is evaluated with 5-fold cross validation on abundant clinical data, and the mean prediction performance: area under the curve: 0.9678; sensitivity: 0.9786; specificity: 0.9161; accuracy: 0.9247, showed the excellence of the model prediction. Here, we present a new, noninvasive, and efficient detection means for the automatic prediction of DF, which can help clinicians find potential diabetic foot patients earlier and is expected to be a novel auxiliary diagnostic tool.

INDEX TERMS Diabetic foot, feature selection, fundus image, radiomics, support vector machine.

I. INTRODUCTION

According to the *Global report on diabetes* [1], first published in 2016 by World Health Organization (WHO), the prevalence of diabetes has risen sharply over the past 30 years. The number of adults living with diabetes has tripled in the last 40 years. In 2014, the global prevalence of diabetes was 8.5%, and the number of adults with the disease has reached 422 million. In 2012, the number of deaths directly caused by diabetes was 1.5 million, ranking it in the top eight causes of death and the top five among women. Diabetes has become one of the major causes of death, disability, and shortened life. The real horror of diabetes lies not in itself but in a series of related complications, such as diabetic retinopathy, diabetic nephropathy, and diabetic foot (DF), which is the main cause of amputation in diabetic patients.

DF is a common chronic complication in type 2 diabetes, which is mainly caused by the destruction of vascular endings and neuropathy, resulting in infection or ulcer of the lower limbs, or even amputation if treatment is not timely. Bi Y's study [2] showed that the average prevalence of DF in the world was 6.3%. The prevalence rate in North America was 13%, ranking first, and that in Africa was 7.2%, ranking second. Asia and Europe were ranked third and fourth with

The associate editor coordinating the review of this manuscript and approving it for publication was Alessandra Bertoldo.

5.5% and 5.1%, respectively. Australia had the lowest incidence, only 3%. The current clinical examination methods for DF mainly include neurologic examination, vascular examination, and foot pressure examination, but these methods involve tedious steps and strict requirements for the examiners. A simple and efficient inspection method is urgently needed.

In recent years, radiomics, with its extraction and analysis of hidden information in medical images, has helped researchers make significant achievements in various fields of medical imaging, especially tumor research [3], [4]. Radiomics refers to the high flux to extract the image features of a region of interest (ROI) in great quantities, to help researchers in the data mining of medical images; this data can be recognized by machine learning algorithms, so as to help doctors diagnose patients, assess the stages of cancer, determine the best treatment more accurately, and reduce the burden on doctors.

Radiomics mainly consists of four steps, namely dividing the ROIs, extracting the features of the ROIs, selecting and reducing the dimensions of the features extracted, and establishing a model to complete the prediction or classification of diseases. Compared with traditional clinical examination methods, computer-aided diagnosis technology is known as the doctor's "third eye" due to its noninvasive nature, high efficiency, and accuracy, which plays a crucial role in the diagnosis and recognition of many diseases [5]–[8].

How to use machine learning algorithms for the data mining of massive raw data remains a current research hot spot. As a comprehensive data mining method, ensemble learning has made great achievements in opinion mining [9], [10], text mining [11]–[13], web mining [14], and medical information mining [15]. The main idea is to construct a better and more comprehensive model by comprehensively considering the results of various feature selection algorithms or classifier algorithms. Inspired by the idea of ensemble learning, we separate the features into two steps and get better results than the single basic algorithm.

Fundus imaging is crucial to medical imaging research. Since the blood vessels in the fundus are the only blood vessels that can be observed directly from the body surface, doctors can check whether there are lesions in the optic nerve and retina of the fundus by analyzing the fundus images. For the fundus examination of diabetic patients, current studies mainly focused on the automatic recognition and classification of diabetic retinopathy [16], [17], while another complication of diabetes, DF, has not been reported. In view of this research status, we proposed an automatic prediction model for fundus images of DF patients based on radiomics features. The main contributions of this paper are as follows:

- To our best knowledge, this is the first time that DF has been determined by an automated analysis through fundus images, with encouraging results. It can be used as an effective way to find potential DF patients.

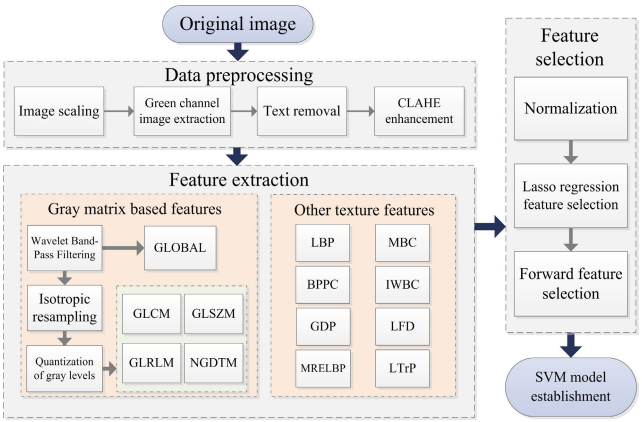

FIGURE 1. Main process of the proposed method.

- On the basis of the four commonly used gray-scale matrix features, we added a total of eight categories of feature descriptors, which greatly enriched the extracted information. At the same time, compared with the case that used only the four gray-scale matrix features, the newly added features can play a role in assisting the classification and improving the prediction performance.
- Compared with other researches on radiomics, the method we proposed does not require a professional physician to delineate the ROIs of microaneurysms or bleeders, but adopts a threshold segmentation method to generate a holistic mask, which saves a lot of manpower.
- A two-step feature selection model was proposed to reduce the complexity of the model and find the optimal combination of the features extracted, which produced a good effect.

II. MATERIALS AND METHODS

In this section, we mainly describe the clinical data set used in the experiment and all the details of our proposed method. The main work of this paper is shown in Fig. 1.

A. DATASET

This study retrospectively analyzed the general clinical data and color fundus images of 1198 patients with type 2 diabetes admitted to Nanfang Hospital, Guangdong, China from 2004 to 2018, including 1026 cases of common diabetes mellitus (DM) and 172 cases of DF, as shown in Table 1. The fundus images are all 1400 × 1400 pixels resolution. All patients were photographed binocularly by professional nurses using the mydriasis-free TRC-NW2000 fundus camera (Topcon, Japan) centered on the macular area and the mid-point of the optic disc. Among them, there were 125 cases of left eye image deletion and 87 cases of right eye image deletion. Finally, we obtained 2184 fundus images. Due to the long collection time span, the quality of the fundus images was not uniform, which was caused by various factors. Some of

TABLE 1. Characteristics of patient cohort.

Type	DM	DF
Male	560	83
Female	466	89
Male Age Range (y)	18-86	37-89
Female Age Range (y)	18-94	29-89

TABLE 2. Distribution of training and test set.

Type	Training	Test
DM	1232	521
DF	192	90
Total	1424	611

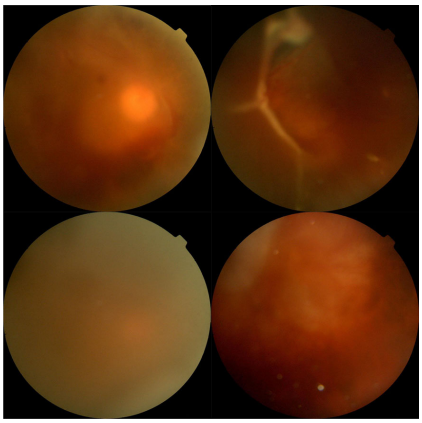

FIGURE 2. Fuzzy image example. The top line images are of DF patients, and the bottom line are of DM patients.

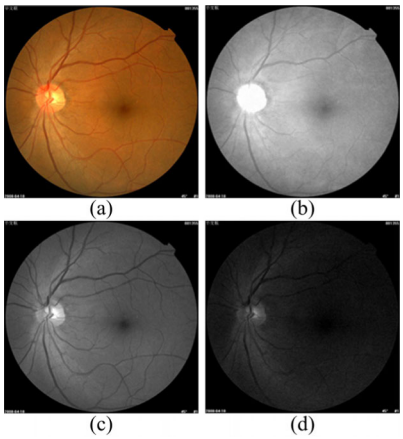

FIGURE 3. Image of different channels. (a) Original image (b) R channel image (c) G channel image (d) B channel image.

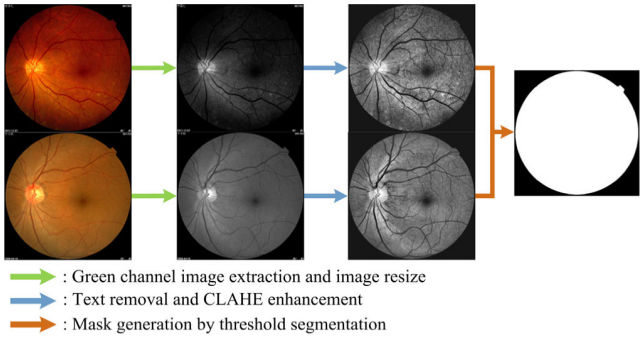

FIGURE 4. Image preprocessing steps on images of different quality.

the fundus images were very fuzzy, as shown in Fig. 2, which will adversely affect the follow-up studies. After removing the fuzzy images, the final dataset was randomly divided into a training set and test set according to the ratio of 7:3, and the specific distribution is shown in Table 2. DF patients and DM patients are labeled as 1 and 0, respectively, which will be used as the classification label of our model.

B. FEATURE EXTRACTION

In this section, we describe in detail how we got features from the original dataset, and briefly introduce the basic principles of the radiomics features we adopted.

1) DATA PREPROCESSING

Image size also has a certain impact on feature extraction. As we know, high-resolution images contain more information, which may have a beneficial impact on the prediction model. However, sizes that are too large will also increase the cost of calculation. In this paper, we discussed two different sizes of images, 350 × 350 pixels and 1400 × 1400 pixels, to explore the impact of image size on feature extraction. The results will be shown in section 3. At the same time, the original fundus images also contained some useless text information of patients and shots, which needed to be eliminated. As shown in Fig. 3, experiments show that the structure of the fundus images extracted from the green channel was more obvious, which was more conducive to feature extraction. Meanwhile, it was necessary to use CLAHE [18] on the green channel image to eliminate the impact of external factors as much as possible. Fig. 4 shows the preprocessing step in detail.

2) TEXTURE FEATURE EXTRACTION

Texture features reflect both the homogeneity phenomenon in the image and the surface structure arrangement attribute of the object with slow or periodic changes. In order to extract more abundant texture features of the ROIs, we tested many types of texture features, and the specific types are shown in Table 3. Different from previous radiomics studies, it was not necessary to manually sketch any ROIs, but a threshold segmentation algorithm was used to automatically generate a mask.

Global features are commonly used first-order statistical features, which can show the statistical characteristics of images. Gray-level co-occurrence matrix (GLCM [19]), gray-level run-length matrix (GLRLM [20], [21]), gray-level size zone matrix (GLSZM [22]), and neighborhood gray-tone difference matrix (NGTDM [23]) are four common matrix-based texture features in texture classification. GLCM

TABLE 3. Radiomics features list.

Descriptor Name	Dimension
GLOBAL	3
GLCM	9
GLRLM	13
GLSZM	13
NGDTM	5
LBP	256
MRELBP	800
BPPC	1062
IWBC	2048
GDP	256
LFD	512
MBC	3072
LTrP	256

describes the joint distribution of two pixels with a certain spatial position relationship and can be regarded as the joint histogram of two-pixel gray pairs, which is a kind of second-order statistics. GLRLM can reflect the comprehensive information of gray images about direction, adjacent interval, and amplitude variations. GLSZM is used to calculate the connected number of voxels in an image. If the adjacent voxels have the same gray level, then these voxels are connected. NGTDM represents the difference between the gray value of a point and the average gray value in its neighborhood, and stores the sum of the differences between all gray levels and their average gray values in the matrix. Before extracting the above features, we carried out three preprocessing operations on the images, namely wavelet band-pass filtering, isotropic resampling, and gray-scale transformation to obtain more abundant texture features with reference to [24].

i) Wavelet band-pass filtering: to avoid the influence of noise and highlight the difference of different bandwidths, we used the ''Sym8'' wavelet basis function to decompose and reconstruct the images in this work. We use 'W' to denote the ratio of high and low frequency coefficients.
ii) Isotropic resampling: to make the extracted texture features more abundant, this operation was taken to obtain images with different resolutions. We use 'S' to denote the size of isotropic resampling.
iii) Quantization and gray level transformation: to reduce computational complexity and extract more texture features, this operation was taken to convert images to different gray levels. Two important parameters in this operation, namely, the quantization algorithm and number of gray levels, were denoted by 'Algo' and 'Ng', respectively. With respect to the quantization algorithm, two algorithms were chosen in our work, namely, equal-probability [24] and Lloyd–Max quantization [25], [26].

TABLE 4. Parameters used in gray matrix-based texture features.

Parameters Name	Values
Wavelet band-pass filtering	W = [1/2, 1, 3/2]
Isotropic voxel size	S = {in-pR*, 1, 2, 3}
Quantization algorithm	Algo = {Equal, Lloyd}
Number of gray level	Ng = [8, 16, 32, 64]

*in-pR, Ng denote initial in-plane resolution, and gray levels, respectively.

The same preprocessing procedure with different parameters resulted in different features. All parameters are shown in Table 4. It is remarkable that procedure iii) was used on GLCM, GLRLM, GLSZM, and NGDTM only.

Local binary patterns (LBPs) [27] are a kind of common local texture descriptor; it encodes the pixel points by comparing the pixel values between the central pixel and the surrounding pixels, and concatenates them in a certain order. Finally, the histogram is calculated to obtain the characteristics of the local texture. LBPs have been widely used in pattern recognition, face recognition, medical image classification, and other fields for its excellent description performance and robustness. Following their success, many varieties of LBPs and the features inspired by LBP's coding have been developed. Reference [28] proposed a novel variant of an LBP named Median Robust Extended Local Binary Pattern (MRELBP), which can calculate multiscale LBP descriptors by comparing the median of the image on a novel sampling scheme effectively, so that it can capture more texture information of micro and macro structure to make features more robust. Binary Pattern of Phase Congruency (BPPC) [29] combines phase congruency and LBPs to highlight the edges and lines to get multiscale local features. Yang *et al.* [30] advanced the IWBC, which improved the performance in some uneven areas compared with the weber local descriptor (WLD). Gradient directional pattern (GDP) [31], [32] is a kind of feature descriptor that encodes the texture information of the local area by quantifying the gradient direction angle. The obtained GDP features retain more information than the method based on gray level, because the gradient operator can effectively enhance the edge information of the image. A local frequency descriptor (LFD) [33] is a descriptor that represents local phase and amplitude information; it was first proposed for low-resolution face image recognition, and achieved relatively robust performance. The LFD is also included in the list of texture feature extraction because there are also low-resolution images caused by photographic conditions in medical images. Monastic binary coding (MBC) [34], [35] is a kind of texture feature descriptor based on monogenic signal analysis. The original image can be decomposed into three complementary components: amplitude, direction, and phase through monogenic signal analysis and a Reese transform. Then, a series of MBC features can be obtained by coding the selected neighborhood, where have lower time and space complexity than those based

on the Gabor transform. LTrP [36], [37] encodes the transition of intensity change in different directions over a local area.

The above 13 kinds of radiomics features include the information of image gray, gradient, direction, phase, frequency and so on, which can describe the fundus image more comprehensively.

C. FEATURE SELECTION

Due to the high dimensions of radiomics features and the large amount of redundant information they contain, feature selection operation is needed so as to screen out the best feature combination. The feature subset needs to be concise, representative, and distinguishable. Concise meaning that the number of features we choose could not be too much. Otherwise, the complexity of the model will be increased. Representative meaning that the features we choose must be the main component of all the features and highly relevant to the categories. Distinguishable meaning that the features we choose need to have low correlation between each other to prevent redundant information and overfitting. In conclusion, this paper proposed a two-step feature selection model. Firstly, the basic feature selection algorithm was used to preliminarily screen the features to obtain a preliminary feature subset containing K features. For the obtained preliminary feature subset, forward feature selection was used to further screen the optimal feature combination. In this process, the bootstrap+0.632 method was used to ensure the stability of the model.

1) LASSO REGRESSION FEATURE SELECTION

Since the feature subset in 2.2 contains many kinds of features, their value ranges are also different. Normalization is needed for the feature subset to eliminate the influence of different dimensions. The normalization formula used is as follows:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

where X denotes the features, X_{min} and X_{max} denote the minimum and maximum value of that dimension feature, respectively, and X_{norm} denotes the normalized feature. Lasso regression [38] refers to adding an L_1 regularization term on the basis of the original loss function. Regularization function is the embodiment of the structural risk minimization strategy, which is to select the model with a smaller loss function and model complexity at the same time. L_1 norm can help the model to obtain a sparse solution more easily, so as to filter features.

The basic loss function used in our experiment was the mean square error (MSE), and the L_1 regularization term was added to the MSE to form the loss function J (w) used in our work:

$$J(\mathrm{w}) = \frac{1}{m}\sum_{i}^{m}(y_i - w^T x_i)^2 + \lambda \|w\|_1 \ (\lambda > 0)$$

where w, m, x, and y denote weight, number, value and label of the feature respectively. λ denotes the penalty coefficient, which is set by experiment. Considering the computational complexity and Lasso selection results, we set the size of feature subset to 30 in this work.

2) SUPPORT VECTOR MACHINES

In our work, we used a cost-support vector machine (C-SVM) [39], [40] as the prediction model for DF prediction. C-SVM has two main parameters: penalty factor C and the type of kernel function. In addition, there is another significant parameter γ when the type of kernel function is 'RBF.' C is the penalty coefficient, that is, tolerance to error. The larger the C, the less tolerance of error. The generalization ability may be poor if C is too large or too small. So, it is appropriate to use a grid search method to determine the range of C and γ values. Then, after follow-up experiments to find the best combination of parameters, a 5-fold cross validation was also used to select the parameters in our work. More experimental details will be described in section 3.2.

3) BOOTSTRAP+0.632

The partition of the training set and test set also plays a key role in the performance of the model when training the prediction model. The same model may behave differently in the partition of different datasets. In order to eliminate the model performance instability caused by the different partition of datasets, we adopt the imbalance-adjusted bootstrap+0.632 [41] method to achieve this purpose, which is sampling with return. For example, assuming that D is the sample set with d samples, we sample one at a time with return, d times. Let D^* denote the sample set selected, named the internal training set, and D^{γ} denote the sample set that did not appear, named the internal test set. Since the probability of selecting each sample is 1/d and the corresponding probability of not being selected is 1-1/d, the probability of never being sampled ford times is (1-1/d) d. If d is relatively large, the probability is approximately 0.368, that is, 36.8% of the samples are used as the test set, and the proportion of the corresponding training set is 63.2%. The nBoot internal training set and corresponding internal test set can be obtained after repeating the operation above nBoot times, which is set to 200 in our work. The evaluation index of this method will be explained in section 3.1.1 in detail.

4) FORWARD SELECTION

In order to fully consider the diversity of the combination of features, we used the method of forward selection to further screen features based on $AUC_{0.632}$ of SVM. Considering the conciseness of the final model, we plan to control the order of the model within 20, that is to say, we will ultimately select up to 20 features to build the prediction model. The pseudocode of the forward selection is shown in Fig. 5 below.

D. ESTABLISHMENT OF PREDICTION MODEL

Regarding the final prediction model, we once again chose SVM owing to its excellent performance. Finally, the model was tested on the pre-divided test set, and the specific

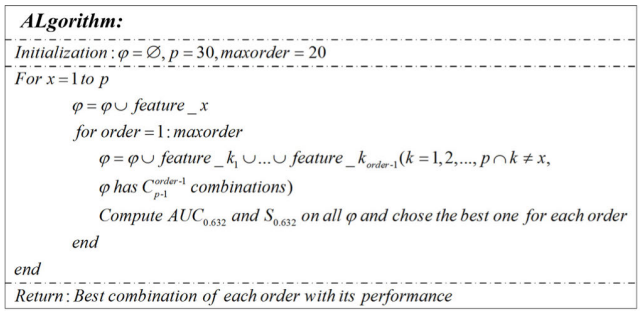

FIGURE 5. The pseudocode of forward selection.

performance indicators of the model will be analyzed and discussed in followed section.

III. EXPERIMENT SETUP AND RESULTS ANALYSES

In this section, we introduce all details of the experiment and analyze the results. We first introduce the experimental steps and tools of implementation. Results of the methods proposed in this paper are indicated in sequence.

A. EXPERIMENTAL STEPS

1) EVALUATION

In order to evaluate our experimental results more objectively, seven evaluation criteria are adopted in this paper: accuracy (acc), sensitivity (sens), sensitivity$_{0.632}$, specificity (spec), specificity$_{0.632}$, area under the curve (AUC), and AUC$_{0.632}$. Among them, the accuracy can represent the overall classification accuracy of the model, and it is also the most commonly used evaluation index. However, when the number distribution of classification objects is unbalanced like the dataset used in our work, it cannot fully represent the advantages and disadvantages of the model. Sensitivity can show the sensitivity of the model to the positive samples. The higher the sensitivity, the lower the missed diagnosis rate. The higher the specificity, the lower the misdiagnosis rate. AUC refers to the area under the ROC (receiver operating characteristic) curve. The horizontal axis of the ROC curve is the false positive rate (FPR) and the vertical axis is the true positive rate (TPR). The specific definitions of each index are shown in the following formulas:

$$\text{spec} = \frac{TN}{TN + FP}, \quad sens = TPR = \frac{TP}{TP + FN}$$
$$FPR = \frac{FP}{TN + FP}, \quad \text{acc} = \frac{TP + TN}{TP + FP + TN + FN}$$

where TN and TP denote true negatives and positives, respectively. Meanwhile, FN and FP denote false negatives and positives, respectively. AUC$_{0.632}$, specificity$_{0.632}$, and sensitivity$_{0.632}$ are the evaluation indices formed by the bootstrap+0.632 self-service sampling method, which are used to evaluate the model performance when training the model. The specific definitions are as follows: $AUC_{0.632}$, as shown at the bottom of the next page,

where AUC(X^{*d}, $X^{\gamma d}$) and S(X^{*d}, $X^{\gamma d}$) denote the AUC and S of the model which was trained on X^{*d} (internal training set) and tested on $X^{\gamma d}$ (internal test set). Accordingly, AUC(X, X) and S(X, X) are similar. In addition, 5-fold cross validation [42], [43] is used to ensure the stability and reliability of the results.

2) IMPLEMENTATION

The proposed algorithms were implemented and verified in MATLAB (2017a). CNNs were constructed and trained in python (3.6) and keras (2.2.4). All the experiments were carried out on a computer with Intel(R) core(TM) i5-6500 CPU @ 3.20GHz, 8 GB memory, and one NVIDIA Titan X GPU. The computer operating system was Windows 7 Professional (x64).

B. RESULTS

1) COMPARISON WITH OTHER METHODS

In order to highlight the advantages of the proposed model, we compare some other algorithms and discuss them in this section. Three algorithms were considered as the basic feature selection algorithm, which are Lasso, ReliefF [44], and Mutinffs [45]. Lasso is affected by the parameter value 'λ' and ReliefF is affected by the parameter value 'k'. In this paper, the feature selection algorithm and its specific parameters were determined by pre-experiment. The specific steps are as follows. Three algorithms were used to select features on the pre-divided training set containing 1424 fundus images. Then, a C-SVM was trained on the selected features and tested on the pre-divided test set containing 611 fundus images. The RBF kernel was used as the kernel function. Since it is the pre-experiment of the selection algorithms, we did not search the grid to find the parameters, but set the default parameters uniformly as $C = 1000$ and $\gamma = -0.2$. Considering the uneven distribution of the two types of data, we added weight information to the selection of features and training of C-SVM. The setting of weight is shown below:

$$\frac{weight_{positive}}{weight_{negative}} = \frac{Number_{positive}}{Number_{negative}}.$$

As shown in Fig. 6, the best parameter 'k' of the ReliefF algorithm is 16. There are two kinds of parameter values of the Lasso algorithm, 'LambdaMinMSE' and 'Lambda1SE.' shown in Fig. 7. In this paper, 'Lambda1SE' was selected as the optimal parameter, so that more feature information can be obtained within the allowable error range. Then parameter 'λ' of Lasso was set to 0.075, and 33 features were selected. In addition, we directly used the Lasso regression model to predict DF, and compared with the model proposed in this paper. The specific results are shown in Table 5.

We also used transfer learning (learning from Imagenet) to train three classic CNN models (Alex-Net [46], VGG-Net16 [47], and Res-Net18 [48]) and compared them with the proposed method. The weights of the layers before the fully collected layer were frozen during fine tuning. When training the network, the pre-divided training set was

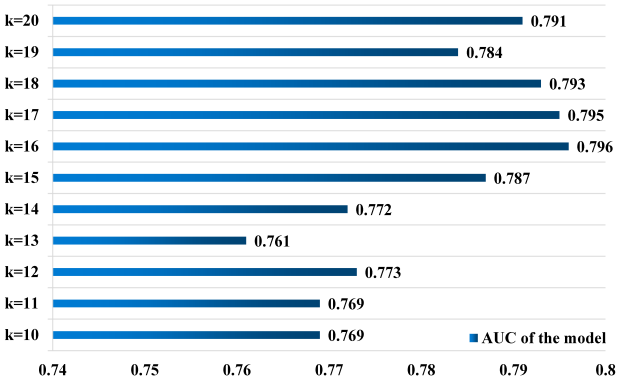

FIGURE 6. AUC of prediction model with different values of parameter 'k' in ReliefF algorithm.

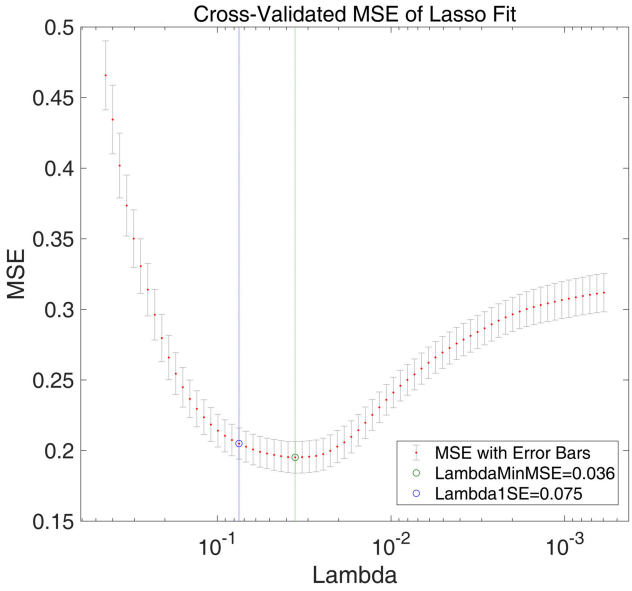

FIGURE 7. The parameter 'λ' of Lasso algorithm was selected through 10-fold cross validation and MSE evaluation.

divided into a training set (1137 images) and verification set (287 images) according to the ratio of 8:2, and the pre-divided test set (611 images) was still used to evaluate the network performance. Since training CNN required a large amount of data, we obtained small patches by randomly cutting the ROI of the original images (1400 × 1400 pixels) for 10 times. Then, we used affine transformation, such as flip and rotation, to further enlarge the data by 5 times, and finally obtained 62535 training data and 15785 validation data. CNN training parameters were set as follows: learning rate = 10^{-5}, loss function = 'focal loss', optimizer = 'SGD', batch = 16, number of iterations = 200.

As shown in Table 5, the performance of Lasso and ReliefF was significantly higher than that of Mutinffs, and Lasso's was slightly better than ReliefF's. In light of the results, Mutinffs does not work well on unbalanced data, that is to say, it prefers to forecast the sample to the majority class. Lasso and ReliefF were used as the basic algorithm of the two-step feature selection model in the following experiments. The results showed that compared with the model constructed by the single feature selection algorithm, the two-step feature selection model proposed in this paper can effectively find the optimal feature combination, and the Lasso-SVM model has the best performance. Although we used focal loss to increase the cost of misclassification of DF, the prediction model based on CNNs still generally performed poorly in the

$$AUC_{0.632} = \frac{1}{D}\sum_{d=1}^{D} \times[(1-\alpha(d)) \cdot AUC(X,X) + \alpha(d) \cdot AUC'(X^{*d}, X^{\gamma \mathrm{d}})]$$

$$where\ AUC'(X^{*d}, X^{\gamma \mathrm{d}}) = \max\{0.5, AUC(X^{*d}, X^{\gamma \mathrm{d}})\},$$

$$\alpha(d) = \frac{0.632}{1 - 0.368 \cdot R(d)},$$

$$and\ R(d) = \begin{cases} 1 \mathrm{if} AUC(X^{*d}, X^{\gamma \mathrm{d}}) \leq 0.5, \\ \dfrac{AUC(X,X) - AUC(X^{*d}, X^{\gamma \mathrm{d}})}{AUC(X,X) - 0.5} & \text{if } 2 > \dfrac{AUC(X,X)}{AUC(X^{*d}, X^{\gamma \mathrm{d}})} > 1, \\ 0 & \text{otherwise.} \end{cases}$$

$$S_{0.632} = \frac{1}{D}\sum_{d=1}^{D} [(1-\alpha(d)) \cdot S(X,X) + \alpha(d) \cdot S(X^{*d}, X^{\gamma d})]$$

$$where\ \alpha(d) = \frac{0.632}{1 - 0.368 \cdot R(d)},$$

$$and\ R(d) = \begin{cases} \dfrac{S(X,X) - S(X^{*d}, X^{\gamma d})}{S(X,X)} & \text{if } \dfrac{S(X,X)}{S(X^{*d}, X^{\gamma d})} > 1, \\ 0 & \text{otherwise.} \end{cases}$$

for S : Sensitivity, Specificity

TABLE 5. Performance of three feature selection methods.

Algorithm	AUC	Sen	Spec	Acc	Order
Mutinffs+SVM	0.7495	0.7111	0.8196	0.8036	30
ReliefF+SVM	0.7960	0.8000	0.8388	0.8331	30
Lasso+SVM	0.8412	0.8889	0.8445	0.8511	30
Lasso Regression	0.8737	0.8111	**0.8848**	0.8747	33
Alex-Net	0.8031	0.7222	0.7486	0.7021	\
Vgg-Net16	0.8124	0.6889	0.8273	0.7758	\
Res-Net18	0.8673	0.7778	0.8733	0.8592	\
ReliefF-SVM(proposed)	0.9136	0.9222	0.8522	0.8625	15
Lasso-SVM(proposed)	**0.9471**	**0.9889**	0.8829	**0.8985**	19

TABLE 6. Comparation between high and low resolution images.

Resolution (pixels)	Model order	Time cost	AUC	Sen	Spec	Acc
1400×1400	17	1215s	0.9486	1	0.8887	0.9051
350×350	19	74s	0.9471	0.9889	0.8829	0.8985

sensitivity index, which indicates that it is difficult for CNNs to capture the features of DF in the prediction.

In addition, we analyzed the effect of different resolution images from two aspects: the efficiency of feature extraction and the performance of the prediction model, using the Lasso-SVM model. It can be seen from Table 6 that the prediction performance of the model constructed from the features extracted from high-resolution and low-resolution images was almost the same, but the time that feature extraction took was nearly 16 times longer for the high-resolution images. Considering both of them, it is advisable to reduce the image resolution, extract features, and build models.

2) FEATURE ANALYSIS AND MODEL VALIDATION

Known by the results of section 3.2.1, it can be seen that the Lasso-SVM model proposed in this paper had the best prediction performance. Further analysis and validation are discussed in this section.

After Lasso filtering, many features were set to 0, and the top 30 features were selected by weight order. The details of the 30 features of the initial feature subset are shown in Table 7. As shown in the table, there are nine types of features selected, of which MBC is the largest proportion.

Fig. 8 shows the model performance curves of different orders obtained by using the proposed method. It can be seen that each index fluctuates greatly at the initial two orders, and then presents according to their respective trends. The specificity$_{0.632}$ of the model keeps increasing and is stable at about 0.93. Meanwhile, the AUC$_{0.632}$ is basically stable at about 0.95. In addition, the sensitivity$_{0.632}$ gradually decreases and remains stable at about 0.96. Considering the stability and performance of the model, we finally determine the order of the model as 19. The features selected in the final model are listed in Table 8.

For the selected features, we used the grid search method to determine the approximate range of model parameters, as shown in Fig. 9, and then the values of SVM parameters were obtained by the experiments as follows: $C = 2^{10}$, $\gamma = 2^{-5}$. To observe the prediction performance of the model more clearly, we drew the SVM decision value sigmoid curves on the training and the test sets (Fig. 10), and listed the corresponding confusion matrix (Fig. 11). The corresponding evaluation indices of the training set and test set are shown in Table 9. At the same time, the ROC curves of the model on the training and test sets are shown in Fig. 12.

The pre-divided training set did not participate in the training process of feature selection, which reflected the generalization ability of the model to a certain extent. However, there

TABLE 7. Features in initial feature subset.

Feature_x	Feature name	Feature_x	Feature name
feature_1	NGTDM	feature_16	LFD
feature_2	GLSZM	feature_17	IWBC
feature_3	NGTDM	feature_18	MBC
feature_4	GLSZM	feature_19	MBC
feature_5	MBC	feature_20	MBC
feature_6	MBC	feature_21	MBC
feature_7	LTrP	feature_22	BPPC
feature_8	MBC	feature_23	IWBC
feature_9	GDP	feature_24	MBC
feature_10	MBC	feature_25	MBC
feature_11	MBC	feature_26	MRELBP
feature_12	MBC	feature_27	MBC
feature_13	LFD	feature_28	MBC
feature_14	MBC	feature_29	MBC
feature_15	MBC	feature_30	MBC

TABLE 8. Features in Lasso-SVM model.

Feature_x	Feature name	Feature_x	Feature name
feature_22	BPPC	feature_24	MBC
feature_2	GLSZM	feature_15	MBC
feature_3	NGTDM	feature_16	LFD
feature_21	MBC	feature_25	MBC
feature_30	MBC	feature_4	GLSZM
feature_1	NGTDM	feature_14	MBC
feature_8	MBC	feature_9	GDP
feature_29	MBC	feature_17	IWBC
feature_20	MBC	feature_11	MBC
feature_13	LFD		

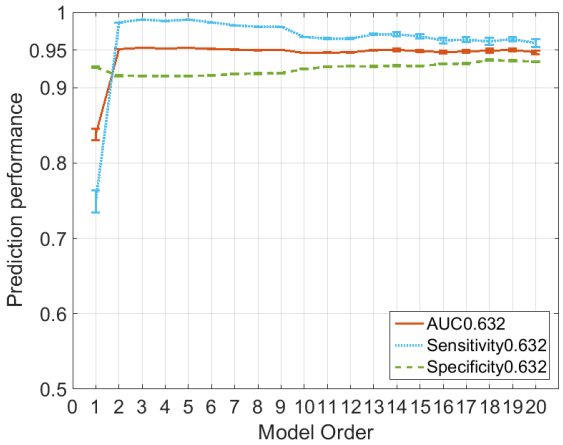

FIGURE 8. Model performance of different order.

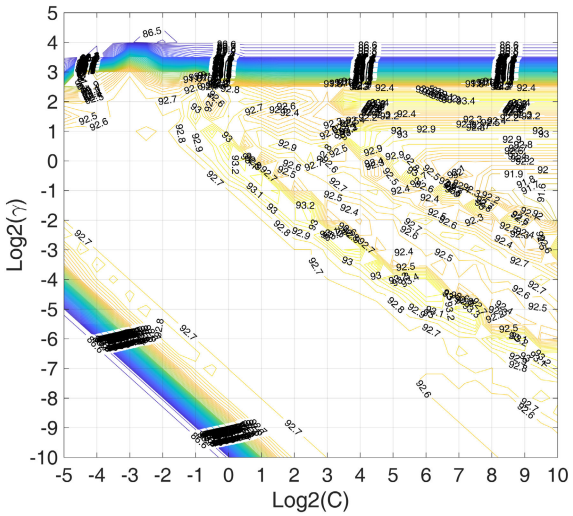

FIGURE 9. Trace of parameter selection (C and γ).

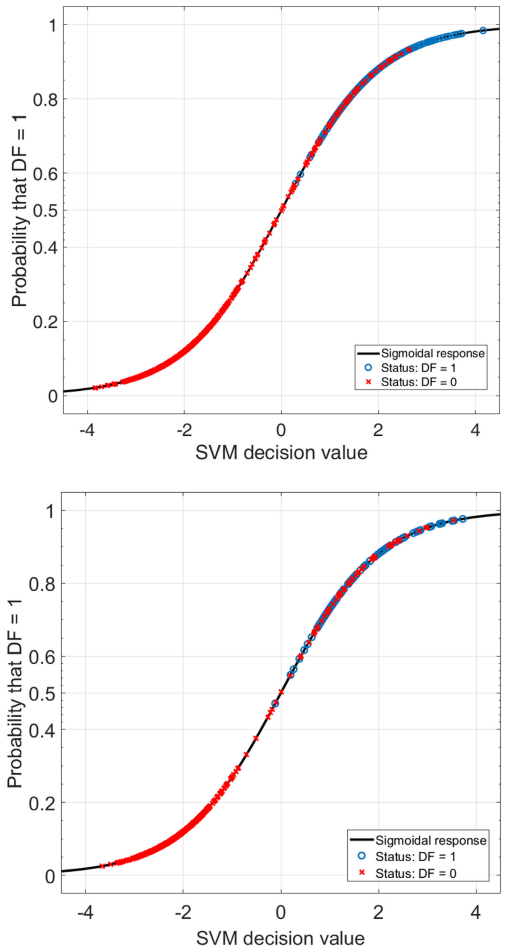

FIGURE 10. SVM decision value sigmoid curve on training (above) and test (bottom) set.

is a risk that the model may be unstable due to the single data partition. Therefore, we combined the two data sets and conducted a 5-fold cross validation to further verify the

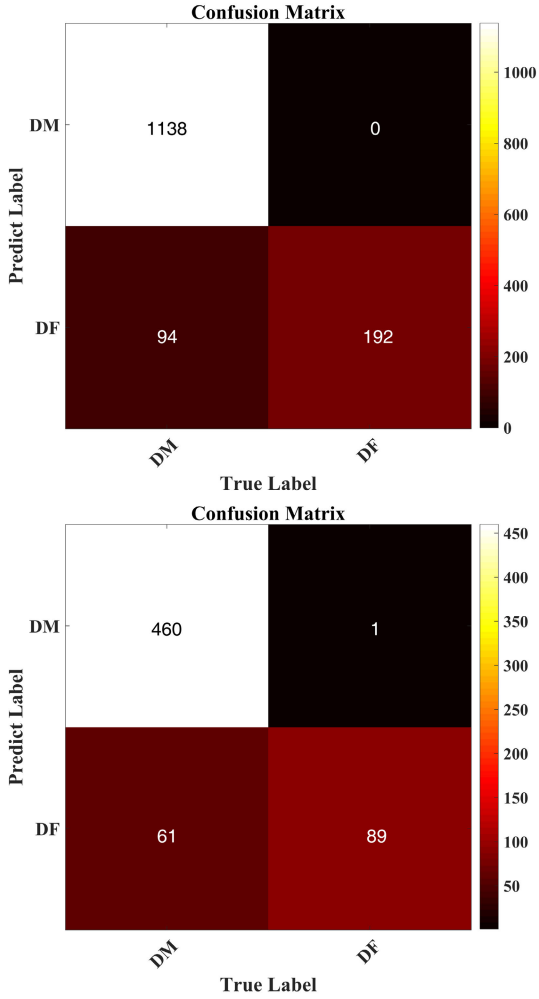

FIGURE 11. Confusion matrix of training (above) and test (bottom) set.

TABLE 9. Index of the proposed model on training and test set.

Index / Data	AUC	Sen	Spec	Acc
Training set	0.9793	1	0.9237	0.9340
Test set	0.9471	0.9889	0.8829	0.8985

TABLE 10. The results of 5-fold cross validation.

(a). Training set

Train	CV 1*	CV 2	CV 3	CV 4	CV 5	Mean
AUC	0.9813	0.9809	0.9815	0.9816	0.9807	**0.9812**
Sen	1	1	1	1	1	**1**
Spec	0.9233	0.9229	0.9223	0.9230	0.9221	**0.9227**
Acc	0.9339	0.9335	0.9330	0.9337	0.9329	**0.9334**

(b). Test set

Test	CV 1	CV 2	CV 3	CV 4	CV 5	Mean
AUC	0.9677	0.9679	0.9672	0.9662	0.9702	**0.9678**
Sen	0.9747	0.9817	0.9788	0.9725	0.9854	**0.9786**
Spec	0.9132	0.9198	0.9172	0.9163	0.9138	**0.9161**
Acc	0.9219	0.9283	0.9258	0.9238	0.9238	**0.9247**

* "CV x" means the xth 5-fold cross validation.

effectiveness of our experiments. The results in Table 10 show that the performance of the model could withstand testing.

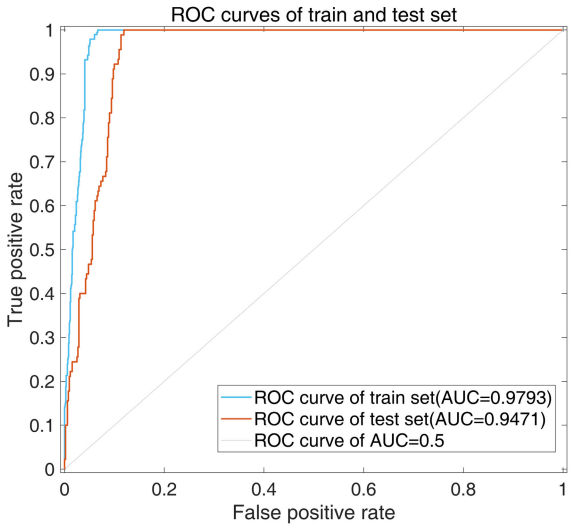

FIGURE 12. ROC curve of training and test set.

TABLE 11. Correlation between features of different index and label.

Feature_x	Spearman coefficient		Cross entropy	MIC
	r value	*p* value		
feature_1	0.4575	7.7043e-106	7.0400	0.3525
feature_2	0.4628	1.5672e-108	13.1020	0.3056
feature_3	0.4875	5.7851e-122	6.4296	0.3586
feature_4	0.2869	7.4200e-40	4.7479	0.2173
feature_8	0.0135	0.5421	14.3499	0.0120
feature_9	0.0272	0.2195	12.4798	0.0220
feature_11	0.0600	0.0068	16.9948	0.0174
feature_13	0.0433	0.0510	12.3310	0.0193
feature_14	0.0207	0.3500	14.1200	0.0171
feature_15	0.0065	0.7680	13.8064	0.0066
feature_16	0.0159	0.4729	13.2334	0.0204
feature_17	0.0243	0.2740	13.0886	0.0165
feature_20	0.0390	0.0790	13.3677	0.0226
feature_21	0.0515	0.0201	12.4459	0.0534
feature_22	0.0893	0.0001	12.3114	0.0579
feature_24	0.0265	0.2327	15.6442	0.0133
feature_25	0.0197	0.3736	15.4821	0.0037
feature_29	0.0027	0.9042	12.5419	0.0139
feature_30	0.0325	0.1433	10.0550	0.0405
G(x)	**0.5623**	**5.0389e-170**	**2.3202**	**0.4742**

In order to analyze the role of each feature in the logistic regression model and prove the superiority of the model, Spearman coefficient, cross entropy [49], and maximum information coefficient (MIC) [50] were calculated on the features. It is clear that the response of the SVM model (G(x)) has better performance than any other single feature shown in Table 11. Moreover, the traditional texture features based on a gray matrix still remain as good indicators, while some new texture features do not perform well in each evaluation index, but the indices of the model that trains them together are better than all the individual features. This shows that the new features play an active role in the construction of a high-dimensional feature classifier.

IV. DISCUSSION AND CONCLUSION

Diabetic foot, as a serious complication of type II diabetes, has the characteristics of being difficult to check, expensive treatment, and serious consequences, leading to lifelong struggles for patients and their families. The traditional method of DF examination is tedious and not timely. In order to explore a more convenient and efficient diagnostic method, we focused on the fundus image of patients, and proposed an SVM classification model based on radiomics features. Fundus imaging is a routine method of ophthalmic examination, which has the characteristics of low cost and simple operation, that is, it can be obtained immediately without invasion. In the feature extraction stage, we extract 11882 radiomics features, including many texture features that are used for the first time in the medical image field. After verification, some of them were proved to improve the classification performance. A two-step feature selection model was adopted to reduce the model complexity and prevent over fitting. Regarding the first feature selection step, we compared three feature selection algorithms through pre-experiment, and finally decided to use Lasso and ReliefF for basic feature selection method to get a feature subset containing 30 radiomics features. On the basis of these 30 radiomics features, a forward selection method based on SVM was designed to further screen the most concise and optimal radiomics featurescombination. In this process, the bootstrap+0.632 method was used to generate 200 sets of internal training and test sets to ensure the stability of the model. Finally, considering the performance and complexity of the model, we chose 19 features to train the SVM model, and obtained excellent prediction results in the pre-divided and 5-fold cross-validated test set. In addition, we compared the DF prediction model based on the CNN and single feature selection algorithm. The model proposed in this paper had obvious advantages in each index, and further illustrates the importance of the optimal feature combination.

On the basis of commonly used texture features based on the gray matrix, we tried using some new features innovatively. Although the performance of these features was not satisfactory in some correlation evaluation indices, our experiments proved that they can achieve good results by combining them reasonably and training the classifier. The addition of new features provides some new information for the establishment of the model, such as phase and gradient information. It has been proved that this information has high value for DF prediction based on fundus images. The value of these features for other medical problems remains to be verified, but our work has presented new ideas for future research.

However, our model still has some limitations. For example, our data are all from the same hospital. Using multicenter data to verify the model is the key to its clinical application. In addition, this paper mainly focused on the texture features of fundus images, and was not combined with clinical features. How to combine clinical features with radiomics featuresis the next direction we need to explore.

Our next work will focus on the integration of deep learning technology and radiomics. As a new technology, deep learning has made good achievements in various fields

with its powerful performance. In the future, we will apply ourselves to combine the deep learning technology with radiomics features to improve the prediction performance. Meanwhile, we plan to integrate this method into clinical examination, and take the result of the model proposed in this paper as an examination index to assist clinicians in comprehensive diagnosis and analysis of patients.

CONFLICTS OF INTEREST

None declared.

REFERENCES

[1] G. Roglic, "WHO global report on diabetes: A summary," *Int. J. Noncommun. Dis.*, vol. 1, no. 1, pp. 3–8, Jan. 2016.
[2] P. Zhang, J. Lu, Y. Jing, S. Tang, D. Zhu, and Y. Bi, "Global epidemiology of diabetic foot ulceration: A systematic review and meta-analysis," *Ann. Med.*, vol. 49, no. 2, pp. 106–116, Feb. 2017, doi: 10.1080/07853890.2016.1231932.
[3] P. Sun, D. Wang, V. C. Mok, and L. Shi, "Comparison of feature selection methods and machine learning classifiers for radiomics analysis in glioma grading," *IEEE Access*, vol. 7, pp. 102010–102020, 2019, doi: 10.1109/ACCESS.2019.2928975.
[4] Y. Guo, Y. Hu, M. Qiao, Y. Wang, J. Yu, J. Li, and C. Chang, "Radiomics analysis on ultrasound for prediction of biologic behavior in breast invasive ductal carcinoma," *Clin. Breast Cancer*, vol. 18, no. 3, pp. e335–e344, Jun. 2018, doi: 10.1016/j.clbc.2017.08.002.
[5] A. Rahmim, K. P. Bak-Fredslund, S. Ashrafinia, L. Lu, C. R. Schmidtlein, R. M. Subramaniam, A. Morsing, S. Keiding, J. Horsager, and O. L. Munk, "Prognostic modeling for patients with colorectal liver metastases incorporating FDG PET radiomic features," *Eur. J. Radiol.*, vol. 113, pp. 101–109, Apr. 2019, doi: 10.1016/j.ejrad.2019.02.006.
[6] J. Antunes, S. Viswanath, M. Rusu, L. Valls, C. Hoimes, N. Avril, and A. Madabhushi, "Radiomics analysis on FLT-PET/MRI for characterization of early treatment response in renal cell carcinoma: A proof-of-concept study," *Transl. Oncol.*, vol. 9, no. 2, pp. 155–162, Apr. 2016, doi: 10.1016/j.tranon.2016.01.008.
[7] X. Li, Y. Lu, J. Xiong, D. Wang, D. She, X. Kuai, D. Geng, and B. Yin, "Presurgical differentiation between malignant haemangiopericytoma and angiomatous meningioma by a radiomics approach based on texture analysis," *J. Neuroradiol.*, vol. 46, no. 5, pp. 281–287, Sep. 2019, doi: 10.1016/j.neurad.2019.05.013.
[8] R. Biswas, U. Jia, and M. J. Hasan, "A new approach of iris detection and recognition," *Int. J. Electr. Comput. Eng.*, vol. 7, no. 5, pp. 2530–2536, 2017, doi: 10.11591/ijece.v7i5.pp2530-2536.
[9] A. Onan and S. Korukoğlu, "A feature selection model based on genetic rank aggregation for text sentiment classification," *J. Inf. Sci.*, vol. 43, no. 1, pp. 25–38, Feb. 2017, doi: 10.1177/0165551515613226.
[10] A. Onan, S. Korukoğlu, and H. Bulut, "A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification," *Expert Syst. Appl.*, vol. 62, pp. 1–16, Nov. 2016, doi: 10.1016/j.eswa.2016.06.005.
[11] A. Onan, S. Korukoğlu, and H. Bulut, "Ensemble of keyword extraction methods and classifiers in text classification," *Expert Syst. Appl.*, vol. 57, pp. 232–247, Sep. 2016, doi: 10.1016/j.eswa.2016.03.045.
[12] A. Onan, "An ensemble scheme based on language function analysis and feature engineering for text genre classification," *J. Inf. Sci.*, vol. 44, no. 1, pp. 28–47, Feb. 2018, doi: 10.1177/0165551516677911.
[13] A. Onan, S. Korukoğlu, and H. Bulut, "A hybrid ensemble pruning approach based on consensus clustering and multi-objective evolutionary algorithm for sentiment classification," *Inf. Process. Manage.*, vol. 53, no. 4, pp. 814–833, Jul. 2017, doi: 10.1016/j.ipm.2017.02.008.
[14] A. Onan, "Classifier and feature set ensembles for Web page classification," *J. Inf. Sci.*, vol. 42, no. 2, pp. 150–165, Apr. 2016, doi: 10.1177/0165551515591724.
[15] A. Onan, "A fuzzy-rough nearest neighbor classifier combined with consistency-based subset evaluation and instance selection for automated diagnosis of breast cancer," *Expert Syst. Appl.*, vol. 42, no. 20, pp. 6844–6852, Nov. 2015, doi: 10.1016/j.eswa.2015.05.006.
[16] H. Wang, W. Hsu, K. G. Goh, and M. L. Lee, "An effective approach to detect lesions in color retinal images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Hilton Head, SC, USA, vol. 2, Jun. 2000, pp. 181–186, doi: 10.1109/CVPR.2000.854775.
[17] P. Bharali, J. P. Medhi, and S. R. Nirmala, "Detection of hemorrhages in diabetic retinopathy analysis using color Fundus images," in *Proc. IEEE 2nd Int. Conf. Recent Trends Inf. Syst. (ReTIS)*, Kolkata, India, Jul. 2015, pp. 237–242, doi: 10.1109/ReTIS.2015.7232884.
[18] J. B. Zimmerman, S. B. Cousins, K. M. Hartzell, M. E. Frisse, and M. G. Kahn, "A psychophysical comparison of two methods for adaptive histogram equalization," *J. Digit. Imag.*, vol. 2, no. 2, pp. 82–91, May 1989, doi: 10.1007/BF03168024.
[19] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973, doi: 10.1109/TSMC.1973.4309314.
[20] B. W. Park, J. K. Kim, C. Heo, and K. J. Park, "Reliability of CT radiomic features reflecting tumour heterogeneity according to image quality and image processing parameters," *Sci. Rep.*, vol. 10, no. 1, Mar. 2020, Art. no. 3852, doi: 10.1038/s41598-020-60868-9.
[21] V. B. Dasarathy and E. B. Holder, "Image characterizations based on joint gray level—Run length distributions," *Pattern Recognit. Lett.*, vol. 12, no. 8, pp. 497–502, Aug. 1991, doi: 10.1016/0167-8655(91)80014-2.
[22] G. Thibault, J. Angulo, and F. Meyer, "Advanced statistical matrices for texture characterization: Application to cell classification," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 3, pp. 630–637, Mar. 2014, doi: 10.1109/TBME.2013.2284600.
[23] M. Amadasun and R. King, "Textural features corresponding to textural properties," *IEEE Trans. Syst., Man, Cybern.*, vol. 19, no. 5, pp. 1264–1274, Sep. 1989, doi: 10.1109/21.44046.
[24] M. Vallieres, C. R. Freeman, S. R. Skamene, and I. E. Naqa, "A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities," *Phys. Med. Biol.*, vol. 60, no. 14, pp. 5471–5496, Jul. 2015, doi: 0.1088/0031-9155/60/14/5471.
[25] J. Max, "Quantizing for minimum distortion," *IEEE Trans. Inf. Theory*, vol. IT-6, no. 1, pp. 7–12, Mar. 1960, doi: 10.1109/TIT.1960.1057548.
[26] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 129–137, Mar. 1982, doi: 10.1109/TIT.1982.1056489.
[27] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002, doi: 10.1109/TPAMI.2002.1017623.
[28] L. Liu, S. Lao, P. W. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Median robust extended local binary pattern for texture classification," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1368–1381, Mar. 2016, doi: 10.1109/TIP.2016.2522378.
[29] S. Shojaeilangari, W.-Y. Yau, J. Li, and E.-K. Teoh, "Feature extraction through binary pattern of phase congruency for facial expression recognition," in *Proc. 12th Int. Conf. Control Automat. Robot. Vis. (ICARCV)*, Guangzhou, China, Dec. 2012, pp. 166–170, doi: 10.1109/ICARCV.2012.6485152.
[30] B.-Q. Yang, T. Zhang, C.-C. Gu, K.-J. Wu, and X.-P. Guan, "A novel face recognition method based on IWLD and IWBC," *Multimedia Tools Appl.*, vol. 75, no. 12, pp. 6979–7002, Jun. 2016, doi: 10.1007/s11042-015-2623-4.
[31] F. Ahmed, "Gradient directional pattern: A robust feature descriptor for facial expression recognition," *Electron. Lett.*, vol. 48, no. 19, pp. 1203–1204, Sep. 2012, doi: 10.1049/el.2012.1841.
[32] W. J. Chu, "Facial expression recognition based on local binary pattern and gradient directional pattern," in *Proc. IEEE GreenCom iThings CPSCom*, Beijing, China, Aug. 2013, pp. 1458–1462, doi: 10.1109/GreenCom-iThings-CPSCom.2013.257.
[33] Z. Lei, T. Ahonen, M. Pietikäinen, and S. Z. Li, "Local frequency descriptor for low-resolution face recognition," in *Proc. IEEE Int. Conf. Automat. Face Gesture Recognit. (FG)*, Santa Barbara, CA, USA, Mar. 2011, pp. 161–166, doi: 10.1109/FG.2011.5771391.
[34] M. Yang, L. Zhang, S. C.-K. Shiu, and D. Zhang, "Monogenic binary coding: An efficient local feature extraction approach to face recognition," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 6, pp. 1738–1751, Dec. 2012, doi: 10.1109/TIFS.2012.2217332.
[35] X. X. Xia, Z. L. Ying, and W. J. Chu, "Facial expression recognition based on monogenic binary coding," *Appl. Mech. Mater.*, vols. 511–512, pp. 437–440, Feb. 2014.

[36] T. Jabid and O. S. Chae, "Local transitional pattern: A robust facial image descriptor for automatic facial expression recognition," in *Proc. Int. Conf. Comput. Converg. Technol.*, Seoul, South Korea, 2011, Art. no. 333344.
[37] T. Jabid and O. Chae, "Facial expression recognition based on local transitional pattern," *Information*, vol. 15, no. 5, p. 2007, 2012.
[38] H. Zou, "The adaptive lasso and its oracle properties," *J. Amer. Stat. Assoc.*, vol. 101, no. 476, pp. 1418–1429, Jan. 2012, doi: 10.1198/016214506000000735.
[39] B. Sahiner, H. P. Chan, and L. Hadjiiski, "Classifier performance prediction for computer-aided diagnosis using a limited dataset," *Med. Phys.*, vol. 35, no. 4, pp. 1559–1570, Apr. 2008, doi: 10.1118/1.2868757.
[40] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, p. 27, Apr. 2011, doi: 10.1145/1961189.1961199.
[41] B. Efron and R. Tibshirani, "Improvements on cross-validation: The 632+ bootstrap method," *J. Amer. Stat. Assoc.*, vol. 92, no. 438, pp. 548–560, Jun. 1997, doi: 10.1080/01621459.1997.10474007.
[42] M. J. Hasan and J. M. Kim, "A hybrid feature pool-based emotional stress state detection algorithm using EEG signals," *Brain Sci.*, vol. 9, no. 12, p. 376, Dec. 2019, doi: 10.3390/brainsci9120376.
[43] M. J. Hasan, D. Shon, K. Im, H.-K. Choi, D.-S. Yoo, and J.-M. Kim, "Sleep state classification using power spectral density and residual neural network with multichannel EEG signals," *Appl. Sci.*, vol. 10, no. 21, p. 7639, Oct. 2020, doi: 10.3390/app10217639.
[44] M. Robnik-Šikonja and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," *Mach. Learn.*, vol. 53, nos. 1–2, pp. 23–69, Oct. 2003, doi: 10.1023/A:1025667309714.
[45] Z. Marco and H. Marcus, "Robust feature selection by mutual information distributions," in *Proc. 18th Conf. Uncertainty Artif. Intell.* San Francisco, CA, USA: Morgan Kaufmann, 2002, pp. 577–584.
[46] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.
[47] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, 2015, pp. 1–14. [Online]. Available: http://arxiv.org/abs/1409.1556
[48] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
[49] P.-T. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," *Ann. Oper. Res.*, vol. 134, no. 1, pp. 19–67, Feb. 2005, doi: 10.1007/s10479-005-5724-z.
[50] D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti, "Detecting novel associations in large data sets," *Science*, vol. 334, no. 6062, pp. 1518–1524, Dec. 2011, doi: 10.1126/science.1205438.

XIAOKANG LIANG received the B.Eng. degree from Southern Medical University, in 2018, where he is currently pursuing the master's degree in engineering with the School of Biomedical Engineering. His research interests include radiomics and medical image processing.

EBTESAM N. ALSHEMMARY received the M.Sc. and Ph.D. degrees from the University of Technology, Baghdad, in 1999 and 2007, respectively. She was a Rapporteur with the Computer Department, Faculty of Education for Girls, University of Kufa, Iraq, where she is currently the IT-RDC Manager. Her specializations are computer and control engineering and mechatronics engineering (robotics and image processing). She is especially well-known for profound contributions to different fields of computer programs, biometrics, and automatic identification systems. She has published research in international and local journal in computer science. Her research interests include statistical analyses of health, medical studies, and information technology. She was a member of many committees with the University of Kufa, such as scientific and examination committees. She is a member of the Editorial and the Reviewer Board of the *International Journal of Electronics and Computer Science Engineering* (IJECSE). She is also a member of the Reviewer Board of *European Journal of Science and Engineering* (EJSE), a Reviewer of *IET Image Processing* journal, and the Editorial Board of the *International Journal of Advanced Computer Science and Information Technology* (IJACSIT).

MING MA received the M.S. degree in ophthalmology from Southern Medical University, Guangzhou, China, in 2014. He is currently an Attending Doctor with the Department of Ophthalmology, Nanfang Hospital, Southern Medical University, Guangzhou, China. His research focuses on corneal transplantation.

SHENGWU LIAO received the master's degree in hospital management from Nankai University, in 2018. He is currently working with the Southern Hospital of Southern Medical University, as an Associate Researcher. His research interests include mutual network medical, intelligent medical treatment, and chronic disease health management.

WEI ZHOU received the master's degree in oncology from Tongji University, Shanghai, China, in 2011. She is currently an Associate Chief Physician with Nanfang Hospital, working on medical quality management and improvement.

ZHENTAI LU received the M.S. degree in mathematics from Sun Yat-sen University, Guangzhou, China, and the Ph.D. degree from the School of Biomedical Engineering, Southern Medical University, Guangzhou, in 2008. He is currently a Professor with the School of Biomedical Engineering, Southern Medical University. His research interests include pattern recognition, machine learning, and image processing.