

Received April 24, 2020, accepted May 9, 2020, date of publication May 12, 2020, date of current version May 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2994047

# Retinal Image Classification by Self-Supervised Fuzzy Clustering Network

YUEGUO LUO<sup>ID1</sup>, JING PAN<sup>2</sup>, SHAOSHUAI FAN<sup>ID3</sup>, ZEYU DU<sup>ID3</sup>, AND GUANGHUA ZHANG<sup>ID3,4</sup>

<sup>1</sup>College of Big Data and Intelligent Engineering, Yangtze Normal University, Chongqing 408100, China

<sup>2</sup>Department of Physics and Chemistry, Taiyuan University, Taiyuan 030000, China

<sup>3</sup>Shanxi Intelligence Institute of Big Data Technology and Innovation (SIBD), Taiyuan 030000, China

<sup>4</sup>Department of Computer Science and Engineering, Taiyuan University, Taiyuan 030000, China

Corresponding author: Guanghua Zhang (jieming1022@126.com)

This work was supported in part by the Natural Science Foundation of Chongqing, China, under Grant cstc2019jcyj-msxmX0385, in part by the Educational Science Foundation of Chongqing, China, under Grant KJQN201801426, in part by the Research Funds of Shanxi Transformation and Comprehensive Reform Demonstration Zone under Grant 2018KJCX04, and in part by the Key Research and Development Program of Shanxi under Grant 201903D311009.

**ABSTRACT** Diabetic retinal image classification aims to conduct diabetic retinopathy automatically diagnosing, which has achieved considerable improvement by deep learning models. However, these methods all rely on sufficient network training by large scale annotated data, which is very labor-expensive in medical image labeling. Aiming to overcome these drawbacks, this paper focuses on embedding self-supervised framework into unsupervised deep learning architecture. Specifically, we propose a Self-supervised Fuzzy Clustering Network (SFCN) by a feature learning module, reconstruction module, and a fuzzy self-supervision module. The feature learning and reconstruction modules ensure the representative ability of the network, and fuzzy self-supervision module is in charge of further providing the training direction for the whole network. Furthermore, three losses of reconstruction, self-supervision, and fuzzy supervision jointly optimize the SFCN under an unsupervised manner. To evaluate the effectiveness of the proposed method, we implement the network on three widely used retinal image datasets, which results demonstrate the satisfied performance on unsupervised retinal image classification task.

**INDEX TERMS** Retinal image classification, self-supervised, fuzzy clustering, unsupervised learning.

## I. INTRODUCTION

Retinal images classification of diabetic retinopathy is a significant application in diagnosing diabetic retinopathy. Diabetes has a wide and profound impact on a large number of people around the world by the report of the World Health Organization. According to a survey, 1.5 million people died of diabetes in 2012. It is estimated that 347 million people will develop diabetes in 2014 and the proportion of people with diabetes in the world will rise from 2.8% in 2000 to 4.4% by 2030. People with diabetes are usually over 30 years old, and severe diabetes can lead to diabetic retinopathy which has early and controllable retinal abnormalities such as micro-aneurysms and other small lesions caused by thin retinal capillaries. The summary survey from 35 studies found that diabetic retinopathy was 34.6%, vision-threatening diabetic retinopathy accounted for 10.2% and diabetic macular edema accounts for 6.8% in the overall incidence of

The associate editor coordinating the review of this manuscript and approving it for publication was Tallha Akram<sup>ID</sup>.

the disease. The sooner diabetic retinopathy is detected, the less likely there is vision loss. However, the early symptoms of this disease are not obvious, which makes it impossible to effectively treat the late detection. Moreover, the discovery of this disease requires a great deal of experience and sufficient capacity of the doctor, because the diagnosing situation in retinal images is challenging (Figure 1). Therefore, the demand for an effective classification and screening of retinopathy is even more urgent to help the timely detection and effective treatment of the disease.

Deep learning based methods have been widely introduced in the medical image analysis field, and their ability to recognize and distinguish information can well diagnose diseases [1], [14], [28]. When the number of retinal specialists is insufficient for all patients to be examined individually, the increased prevalence of the disease may cause great difficulties in the treatment of retinal diseases. In previous researches, some diabetic retinal diseases, such as 22-27 glaucoma and 27-30 retinopathy of prematurity, helped human experts by effectively processing large amounts of data with



**FIGURE 1.** Several retinal images with different retinopathy grades. From them, it can be seen that the variations of illumination, retinopathy and view points bring great challenges for unsupervised retinal image classification.

the help of deep learning technology [2], [9], [18], [33]. It can be seen that deep learning has great prospects in the medical field, but the premise is that the diagnosing performance of the method based on deep learning is higher than that of conventional ophthalmologists.

In recent years, the utilization of deep learning plays a remarkable role in retinal image classification methods for diabetic retinopathy diagnosis. Alqudah [5] presented an automated convolutional neural network architecture for a multi-class classification system for retinal diseases, and it achieved performable results of 95.3% by a softmax classifier. Shanti and Sabeenian [34] focused on the classification of diabetic retinopathy fundus images to the severity of the disease using convolutional neural network with the application of suitable polling, softmax, and rectified linear activation layers to obtain a high level of accuracy. Bourouis *et al.* [9] proposed a robust hybrid probabilistic learning approach that appropriately combines the advantages both of the generative and discriminative model for challenging problem in retinal image classification. It obtained a better results and showed the flexibility and the merits of deep learning applications in retinal image analysis. Though sufficient training on large amount of annotated data, existing deep learning based retinal image classification methods have achieved satisfactory performance according to the review research [33], where more successful applications can be found.

However, there is a main drawback in existing deep learning retinal image diagnose works. That is, they require large amounts of labeled data to supervise the network learning. Manual labeling of retinopathy information is a complicated task, requiring many experienced experts to manually annotate a large number of retinopathy images. Due to different personal habits, it may cause inconsistencies in image labeling, discomfort and application in scientific research. In order to overcome this limitation, this paper propose a Self-supervised Fuzzy Clustering Network (SFCN)

to conduct retinal image classification without any annotations. In detail, SFCN consists three main components to learn the robust classification model by a self-supervised framework. The first part is the feature extraction module achieved by stacked convolutional layers, aiming at learning representation with the input of retinal images; The second one is reconstruction module for the learned feature to guarantee the representative ability of feature extraction; The last one is the fuzzy self-supervision module to provide the training direction of the whole network. In summary, SFCN employs the fuzzy clustering results to supervise the network and constrain the network to output the probability of each retinal image belonging every clusters.

## II. RELATED WORK

This section reviews the related works of retina image classification, which is partitioned by three aspects, including retina image classification, unsupervised medical image classification, and a brief introduction of self-supervision.

### A. RETINA IMAGE CLASSIFICATION

Diabetic retinopathy is an important diseases of eye complications caused by diabetes. Early detection helps protect vision and patients need to be checked semi-annually to determine eye conditions. We propose an deep learning approach for processing retinal images to help people effectively treat retinal disease in this paper. There are already many medical image analysis algorithms for diagnosing diabetic retinal image lesions, which can identify and distinguish diabetic retinas by examining the occurrence of bleeding, lesions, micro-aneurysms, exudates, and other symptoms in the fundus image.

In previous studies, diabetic retinopathy was automatically detected by various models, which are principally separated into two categories (conventional machine learning approaches [20], [23], and deep learning methodologies [11], [40]). For the first category, several traditional models are employed into retinal image analysis. For instances, Jelinek *et al.* [23] deployed points-of-interest and visual dictionary that contains important features required to identify retinal pathology; Gupta and Karandikar [20] conducted morphological operations on retinal images to identify exudates and micro-aneurysms features, and then used multi-class support vector machine and k-nearest neighbor classifier for giving severity or grade of abnormality. Recent researches about retina image analysis are almost employing Convolutional Neural Networks (CNN) to automatically learn CNN features from original images and conducting classification on them. Krause *et al.* [25] proposed a domain enriched deep network consisting a representation network that learns geometric features specific to retinal images, and a custom designed computationally efficient residual task network that utilize the features obtained from the representation layer to perform pixel-level segmentation. For retina image classification, Wu *et al.* [40] presented an attention-based deep learning method for the retinal disease classification in

optical coherence tomography images, especially designed an attention model which pays attention to critical regions containing pathological anomalies. These methods are all under supervised framework with demanding a large amount of annotated data, which limits their flexibilities in realistic applications.

### B. UNSUPERVISED MEDICAL IMAGE CLASSIFICATION

For deep learning models, annotating a large scale medical images meets greater challenges than conventional vision task (e.g face recognition) because it requires the expensive labor from well-trained human experts. Thus, unsupervised medical image classification has more extensive prospect in real scenario.

There exists several proposed deep learning based unsupervised medical image classification models by transfer learning or clustering based methodologies [3], [4], [31], [36]. From the view of transfer learning methods, Ahn *et al.* [3] utilized transferable knowledge across different domains aiming at reducing the reliance on annotated training data by using a new hierarchical unsupervised feature extractor with a convolutional auto-encoder placed a top of a pre-trained convolutional neural network; Tang *et al.* [36] combined active learning and transfer learning for medical data classification, which is iteratively querying a small number of informative unlabeled target samples, and removing the source samples conflicted with distribution of target data. As for clustering based methods, Ahn *et al.* [4] proposed an unsupervised feature learning method to tackle the issue of the large volume of unlabeled medical data that learns feature representations to and then differentiate dissimilar medical images using an ensemble of different convolutional neural networks and  $K$ -means clustering; Perkonigg *et al.* [31] presented a method to identify predictive texture patterns in medical images under unsupervised framework, simultaneously encode and cluster medical image patches in a low-dimensional latent space.

### C. SELF-SUPERVISION

Self-supervision is proposed to leverage the huge amounts of unlabeled data to learn useful representation for various problems, for example, image classification, object detection, video recognition, etc. It has been proved that lots of deep learning methods can benefit from pre-trained models on large labeled datasets. The basic motivation behind self-supervision methods is to replace the expensive labeled data by ‘free’ unlabeled data. Being a generic framework, self-supervision enjoys a wide number of applications, ranging from robotics to image understanding.

A common way to achieve self-supervised learning is to derive easy-to-obtain supervision signals without human annotations, to encourage the learning of useful features for regular tasks. Zhang *et al.* [42] designed self-supervised convolutional subspace clustering network, combining a feature extraction module, subspace clustering and a spectral clustering module into a joint optimization framework. Wang *et al.* [38] developed a HSI feature learning

network to learn consistent features by self-supervision for HIS classification and jointly employed the conditional random field framework to boost the performance of self-supervised feature learning. Chen *et al.* [10] trained a generalized self-supervised embedding network to provide slow and robust representation for downstream tasks by learning from the data itself. Moreover, several self-supervised learning methods employ image rotation. For example, Gidaris *et al.* [19] proposed to learn image features by training ConvNets to recognize the 2d rotation that is applied to the image that it gets as input (RotNet); Feng *et al.* [15] introduced a self-supervised learning method that incorporates rotation invariance into the feature learning framework, one of many good and well-studied properties of visual representation, which is rarely appreciated or exploited by previous deep convolutional neural network based self-supervised representation learning methods.

Inspired by these successful application in traditional computer vision tasks, we embed fuzzy clustering algorithm as the training direction into a self-supervised framework to achieve the unsupervised retinal image classification task. Detail approach will be demonstrated in Section III.

## III. SELF-SUPERVISED FUZZY CLUSTERING NETWORK

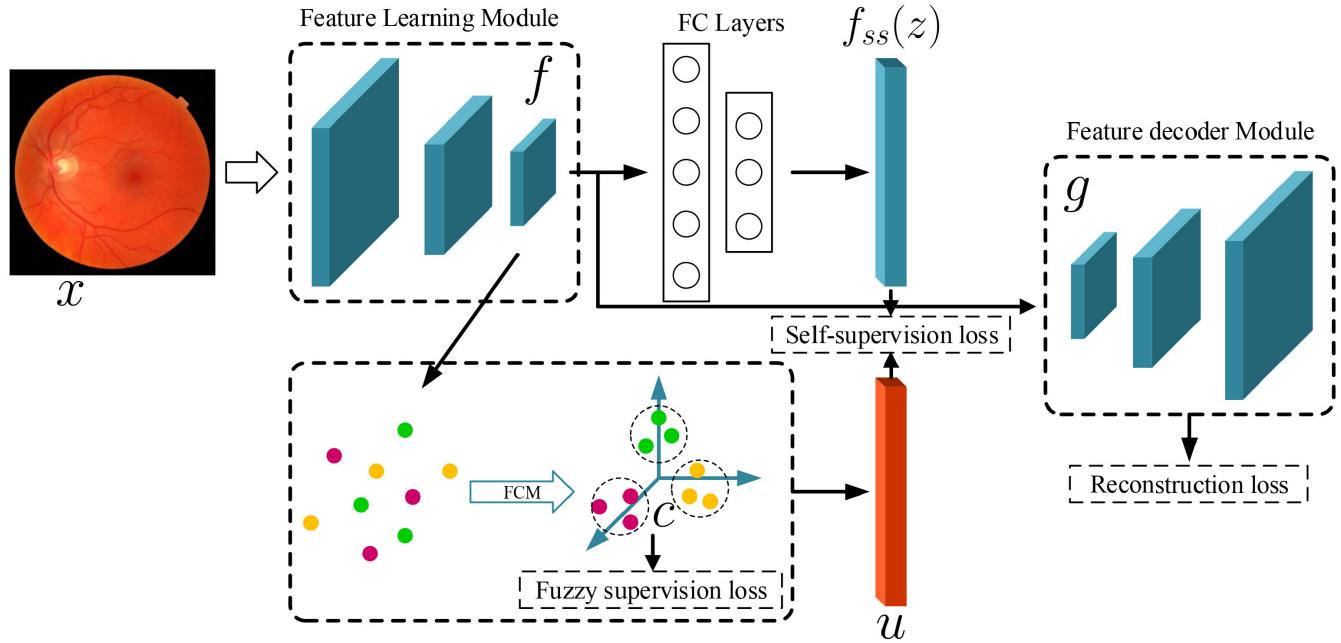
We introduce the proposed Self-supervised Fuzzy Clustering Network (SFCN) in this section, and demonstrate the joint optimization of the self-supervised framework.

### A. APPROACH OVERVIEW

The proposed SFCN is achieved by three main stages, including feature learning directly from unlabeled medical images, self-supervision by fuzzy clustering module, and the application on retinal image classification task. Firstly, the feature learning module is achieved by several stacked convolutional layers to extract CNN representation given a retinal image, and multiple deconvolutional layers are utilized to reconstruct the given retinal image. This module ensures the learned feature representation contains sufficient information to construct itself. Secondly, the fuzzy self-supervision module provides a training supervision for the feature learning module, by the predicted result of fuzzy clustering algorithm. This self-supervision strategy exploits the inherent correlation in unlabeled medical images and makes the output of the network become fuzzy probabilities belonging to each cluster. Finally, we illustrate how to apply the well-trained self-supervised network on retinal image classification task.

### B. FEATURE LEARNING MODULE

The feature learning module is the basic component of our proposed SFCN framework, which is used to extract discriminative features from retinal images. Given a retinal image  $x$ , we can obtain its output feature vector  $z = f(x)$  through an encoder  $f$  with  $L$  stacked convolutional layers with their parameters  $\{w^l, b^l\}_{l=1}^L$  because the convolutional neural network has proved its feature extracting ability in many computer vision tasks.  $w^l$  represents the weight parameters



**FIGURE 2.** Framework of the proposed self-supervised fuzzy clustering network. Our SFCN method contains feature learning module along with fully connected layers, fuzzy self-supervision module, and feature decoder module to achieve the self-supervision process. Through the fuzzy c-means clustering, the last FC layer could output the probability for each retinal image belonging to every clusters, by the constraint of self-supervision loss. To further improve the representation ability for the feature learning module, a reconstruction loss is fixed on the feature decoder module. By the self-supervised fuzzy clustering network, the retinal image classification task can be solved without any annotations.

of  $l$ -th layer and  $b^l$  is the bias parameters of  $l$ -th layer. There are two stream tasks for feature vector  $z$ , which is fed into a decoder to reconstruct original retinal image  $x$ , and provide information for the fuzzy self-supervision module, simultaneously. Explicitly, the feature vector  $z$  is constituted by  $m^L$  feature maps calculated by convolutional layers,

$$z = [h_1^{(L)}, \dots, h_{m^L}^{(L)}] \quad (1)$$

where  $h_1^{(L)}, \dots, h_{m^L}^{(L)}$  represents the raw output of the last layer of  $f$ . This feature vector is learned from unlabeled retinal images, and should exploit the informative data of  $x$  in SFCN framework.

Therefore, we employ a decoder  $g$  of several deconvolutional layers to construct an image  $\hat{x}$ , which should be restore the detail information from the feature vector  $z$ . The reconstruction loss  $L_{re}$  is defined as,

$$L_{re} = \frac{1}{N} \sum_{\mathbb{E}_{x \in X}} \|x - \hat{x}\|^2 \quad (2)$$

where  $N$  denotes the number of training retinal images, and  $X$  is the whole training datasets. The reconstruction loss employs  $l2$  loss to guarantee the stable feature representations for retinal images after the feature learning, which is more sensitive and robust than  $l1$  loss. Through this encoder-decoder architecture,  $z = [h_1^{(L)}, \dots, h_{m^L}^{(L)}]$  maintains the most discriminative feature representation of the image  $x$ . This procedure achieves the robust feature learning in unlabeled medical images.

### C. FUZZY SELF-SUPERVISION MODULE

As an effective generic unsupervised framework, self-supervision can provide a training direction on the unlabeled data. We choose the fuzzy clustering results as the cues of self-supervision, because fuzzy clustering algorithm [8] can learn the probabilities belonging to each cluster center. The estimated probabilities are more realistic than traditional clustering methods, e.g.  $k$ -means, which reason is discussed below. The traditional clustering analysis belongs to “Crisp Partition”, such as  $k$ -means, hierarchical clustering algorithms. They strictly classify each object to be identified into a certain category, with unique decision of the category which the object belongs to. Therefore, this partition of the categories is distinct. However, most objects do not have strict category properties, and they are intermediate condition or category attributes, which is more reasonable to conduct “soft partition”. As for soft partition, fuzzy set theory provides a powerful analytical tool for it, which inspires researchers to adopt fuzzy-based methods to handle clustering task. Fuzzy C-means algorithm is the representative and scalable soft partition method to obtain the uncertainty degree belonging to each category, which can build the fuzzy representation of categories to reflect the nature of the real world. Therefore, integrating fuzzy C-means algorithm into our self-supervised framework is better than conventional crisp partition methods, e.g. K-means, and hierarchical clustering methods.

In detail, fuzzy clustering algorithm treats each sample belonging to every cluster points in a fuzzy conception, which always occurs in realistic application. It partitions each

sample into multiple classes because there often exists noisy, unclear boundaries in the feature space. In fuzzy clustering algorithms, they introduce the membership for each sample belonging to each class by the fuzzy partition conception [8].

Inspired by the successful applications of Fuzzy C-means algorithm [7], [26], [43], we build the self-supervision module by the probabilities of membership in Fuzzy C-Means (FCM) algorithm. Specifically, we denote  $u_i$  as the probability of  $x$  belonging to  $i$ -th cluster, which is iteratively learned by FCM. In the self-supervised framework, we expect our network can directly output this probability following the FCM algorithm without any iterative training, which leads an efficient application in realistic retinal image diagnosis. This expectation can be achieved by a proposed self-supervision loss  $L_{ss}$ ,

$$L_{ss} = \frac{1}{N} \sum_{\mathbb{E}_{x \in X}} \|f_{ss}(z) - u\|^2 \quad (3)$$

where  $f_{ss}$  consists two fully connected layers for reducing the dimension of  $z$  into the number of clusters  $N_c$ , and  $u = [u_1, \dots, u_i, \dots, u_{N_c}]$  is the concatation of every probabilities estimated by FCM algorithm. Though this loss function, we can observe that the predicted performance  $f_{ss}$  is affected by the membership  $u$ . We will discuss how to obtain an optimal probabilities of membership  $u$ .

In fuzzy C-means algorithms, there are two major components for each feature vector of retinal images, including  $N_c$  clusters  $c = [c_1, \dots, c_i, \dots, c_{N_c}]$ , and the probabilities belonging to each cluster  $u = [u_1, \dots, u_i, \dots, u_{N_c}]$ . Note that,

$$\sum_{i=1}^{N_c} u_i = 1, \quad \forall i = 1, \dots, N_c \quad (4)$$

Based on these two components, we design a fuzzy supervision loss  $L_{fs}$  on the feature vector  $z$  to boost the clustering effectiveness by optimizing the network parameters  $\{w^l, b^l\}_{l=1}^L$ . The fuzzy supervision loss  $L_{fs}$  is defined by,

$$\mathcal{L}_{fs} = \frac{1}{N} \sum_{\mathbb{E}_{x \in X}} \sum_{i=1}^{N_c} u_i^m \|z - c_i\|^2, \quad 1 \leq m < \infty \quad (5)$$

where  $m$  is a weighting exponent parameter.

This objective function is not only the network loss, but also for the FCM algorithm. Given the fixed  $u$  and  $c$ , the network parameters  $\{w^l, b^l\}_{l=1}^L$  can be optimized by back propagation algorithm. In contrast, the training of  $u$  and  $c$  can be conducted while given the learned feature vectors by the fixed network parameters  $\{w^l, b^l\}_{l=1}^L$ . The detail optimizations of  $u$  and  $c$  are followed by,

$$u_i = \frac{1}{\sum_{k=1}^{N_c} \left( \frac{\|z - c_k\|}{\|z - c_i\|} \right)^{\frac{2}{m-1}}} \quad (6)$$

$$c_i = \frac{\sum_{\mathbb{E}_{x \in X}} u_i^m \cdot z}{\sum_{\mathbb{E}_{x \in X}} u_i^m} \quad (7)$$

---

**Algorithm 1** Self-Supervised Fuzzy Clustering Network

---

**Initialization:** The parameters of feature extractor  $f, f_{ss}$  and the image decoder  $g$ , the parameters  $\alpha = 0.8, \beta = 0.7$  in encoder-decoder network; The cluster points  $c$ , membership  $u$ , and parameter  $m = 2$  in FCM algorithm. For experimental settings, batch size = 16, learning rate = 0.001, maximum epochs  $N = 300$ , and FCM updating interval  $I_{fcm} = 10$ .

**Pre-train the Encoder-Decoder with maximum Epochs  $N_p$ :**

```
for  $t \in 1, \dots, N_p$  do
    Train  $f$  and  $g$  by  $L_{re}$ (Eq.2).
    If  $\|L_{re}(t) - L_{re}(t-1)\| \leq \varepsilon_p$ :
        Return the pre-trained  $f$  and  $g$ .
end for
```

**Train the self-supervised framework by the pre-trained model:**

```
for  $t \in 1, \dots, N$  do
    Train  $f$  and  $g$  by  $L_{ss}$ (Eq.3) and  $L_{fs}$ (Eq.4).
    If  $t \% I_{fcm} == 0$ :
        Updating  $u$  and  $c$  by Eq.6 and 7 iteratively,
        until  $\|L_{fs}(t) - L_{fs}(t-1)\| \leq \varepsilon_{fcm}$ .
    If  $\|L(t) - L(t-1)\| \leq \varepsilon$ :
        Return the parameters of SFCN framework.
end for
```

---

Finally, we calculate the objective function Eq.3 when  $u$  and  $v$  are in optimal values by iteratively updating them by Eq.6 and 7. Detailed optimization will be described below.

**D. JOINT OPTIMIZATION OF SFCN**

To train the proposed self-supervised fuzzy clustering network in a unified framework, the final loss function  $L$  of SFCN is fused by the balanced losses of Eq.2, 3, and 5,

$$L = L_{ss} + \alpha L_{re} + \beta L_{fs} \quad (8)$$

where  $\alpha$  and  $\beta$  are balance parameters for different items in the total loss. This loss function ensures that we can obtain optimal network parameters given the fixed clusters  $c$  and probabilities  $u$  learned by FCM algorithm. As for the updating of the optimal  $u$  and  $c$ , we should minimize the fuzzy supervision loss  $L_{fs}$  when we obtain a satisfied network output of feature vector  $z$ .

To train SFCN framework, we design the two-stage optimization algorithm to conduct: (1) train the encoder-decoder network to prepare the initial parameters in SFCN; (2) Update the unified network by the fuzzy self-supervised module in back propagation manner. The optimization are summarized in Algorithm 1.

**E. ALGORITHM TRAINING**

Our self-supervised fuzzy clustering network consists of feature learning module, reconstruction module and fuzzy self-supervision module. The feature learning module is

built by ResNet [21], with one convolution layer, four residual blocks and one fully connected layer. Moreover, the reconstruction module is achieved by the image decoder in Cycle GAN [44], which employs the architecture from Johnson and Fei-Fei [22] containing several residual convolution blocks, and two stride-1/2 deconvolutions to conduct upsample, attached instance normalization. As for fuzzy self-supervision module, we attach two fully connected layers after the feature learning module to output the prediction of  $u$ , which is constrained by the results of Fuzzy C-means Clustering on the extracted features.

In detail, we employ Pytorch framework to implement the network on two NVIDIA GTX 2080Ti GPUs under Ubuntu 16.04 operating system. Each image is resized as  $224 \times 224$ , and batch size is set as 32. In addition, we utilize SGB optimizer to update the network parameters, which learning rates is set as 0.001 with decayed to 0 in 300 epochs. The balance parameters in Eq. 8 is following  $\alpha = 0.8$ ,  $\beta = 0.7$ , the parameter  $m = 2$  in Eq.5 when we obtain the optimal performance. The maximum training epochs are set as 300. The training strategy in datasets of different scales is introduced in Section IV.

#### F. APPLICATION ON RETINAL IMAGES

Through the proposed optimization algorithm for SFCN, we can obtain the optimal network parameters of  $f$  and the following fully connected layer  $f_{ss}$ . Given a retinal image dataset  $X = \{x_1, \dots, x_i, \dots, x_N\}$ ,  $f_{ss}$  outputs their probabilities  $u$  belonging to each cluster. We can obtain each predicted cluster label  $\text{Label}_c$  for  $x_i$  by,

$$\text{Label}_c = \arg \max([u_1, \dots, u_{N_c}]) \quad (9)$$

#### IV. EXPERIMENTS

In this section, we conduct extensive evaluating experiments on three available retinal image datasets, including MESSIDOR [13], DRIVE [35] and DIARETDB1 [24], according to the research [6]. Then, we evaluate the performance of our proposed self-supervised fuzzy clustering network in several aspects (e.g. Accuracy, ROC curve, and AUC). Finally, we discuss the effectiveness of the fuzzy self-supervised module in SFCN methods to further validate the novelties of this paper.

#### A. DATASETS

In experiments, we employ MESSIDOR, DIARETDB1, and DRIVE datasets to conduct SFCN, and we bring a brief introduction of them in this subsection.

**MESSIDOR** [13] is a widely used retinal image dataset in automatically diagnosing of retinal images, which has employed by existing approaches. This dataset is funded by the French Ministry of Research and Defense within a 2004 TECHNO-VISION program. In detail, it contains 1200 eye fundus color digital images acquired 3 ophthalmologic departments using a color video 3CCD camera mounted on a Topcon TRC NW6 non-mydriatic retinograph

with a 45 degree field of view. These images were captured using 8 bits per color plane at  $1440 \times 960$ ,  $2240 \times 1488$  or  $2304 \times 1536$  pixels, and annotated by the medical experts into retinopathy grade and risk of macular edema. In this paper, we evaluate this dataset on the normal/abnormal retinal images classification task. Moreover, training the architecture is implemented with 900 fundus images for this dataset, and the left 300 images are employed as testing data, which are equally sampled from normal and abnormal categories.

**DRIVE** [35] is a numerical retinal images used in retinal vessel detection in fundus image, which is employed to conduct the normal/abnormal retinal image classification task in this paper. This dataset is collected from a diabetic retinopathy screening program in the Netherlands. The screening population consists of 400 diabetic subjects between 25-90 years of age. Forty photographs have been randomly selected, 33 do not show any sign of diabetic retinopathy and 7 show signs of mild early diabetic retinopathy. Each image was acquired using a Canon CR5 non-mydriatic 3CCD camera with a 45 degree field of view, and captured using 8 bits per color plane at  $768 \times 584$  pixels. For the division of training and testing data, we select 250 images equally from normal and abnormal classes and the left 150 images are utilized for testing. Note that, because deep neural networks require large amount data to train, we utilize the SFCN model optimized by MESSIDOR to initialize the parameters of SFCN when conduct experiments on DRIVE dataset.

**DIARETDB1** [24] dataset is a standard diabetic retinopathy database to benchmark diabetic retinopathy detection for digital images. In this paper, we employ it for the classification task of diabetic retinopathy. This database consists of 89 color fundus images of which 84 contain at least mild non-proliferative signs of the diabetic retinopathy, and 5 are considered as normal which do not contain any signs of the diabetic retinopathy according to all experiments who participate in the evaluation. These images were captured using the same 50 degree field-of-view digital fundus camera with varying imaging settings, which were independent marked by 4 medical experts. Because this dataset is in a limited scale with 89 images, we also train the SFCN model by initializing the network parameters by the optimized model on MESSIDOR. For this dataset, we use 39 images to conduct testing and 50 images are utilized in training stage.

#### B. EVALUATION CRITERIA

In this subsection, we introduce several evaluation criteria to measure the performance of the proposed self-supervised fuzzy clustering networks.

(1) **Accuracy** indicates the overall classification performance of the model, and it is measured by the ratio of exactly truly predicted sample number in the whole samples, which is calculated by  $\frac{TP+TN}{TP+FP+TN+FN}$ .

(2) **ROC** curve is the Receiver Operating Characteristic curve to measure the representative ability of a classifier, which is widely used in retinal image classification.

**TABLE 1.** Comparison results of accuracy and AUC for normal/abnormal retinal image classification on MESSIDOR dataset.

Models	Training	Accuracy (%)	AUC
Nandy et al. [29](2016)	Supervised	-	0.921
Vo and Verma [32](2016)	Supervised	-	0.862
Quellec et al. [32] (2016)	Supervised	-	0.89
Pedro Costa et al. [12] (2017)	Supervised	-	0.90
Zoom-in-Net [39](2017)	Supervised	90.5	0.921
HCF/CNN [30] (2018)	Supervised	-	0.893
Bourouis et al. [9](2018)	Supervised	100	1.00
DEC [41](2016)	Unsupervised	73.2	0.72
RotNet [19](2018)	Unsupervised	82.3	0.837
SKCN	Modified	82.8	0.815
SHCN	Modified	81.7	0.820
Directly FCM on features	Modified	81.5	0.81
<b>Our proposed SFCN</b>	<b>Modified</b>	<b>87.6</b>	<b>0.85</b>

**TABLE 2.** Comparison results of accuracy for normal/abnormal retinal image classification on DRIVE dataset.

Models	Training	Accuracy (%)
Wang et al. [37] (2002)	Supervised	85.0
Li and Chutatape et al. [27] (2003)	Supervised	85.5
Fleming et al. [16](2007)	Supervised	89.8
Garcia et al. [17](2007)	Supervised	73.6
Bourouis et al. [9](2018)	Supervised	91.3
DEC [41](2016)	Unsupervised	66.4
RotNet [19](2018)	Unsupervised	76.6
SKCN	Modified	75.2
SHCN	Modified	74.8
Directly FCM on features	Modified	77.6
<b>Our proposed SFCN</b>	<b>Modified</b>	<b>81.7</b>

(3) **AUC** is the area under the ROC curve, which is often used to measure the classification ability in the recognition task under the ROC curve. AUC is a widely used evaluation metric for examining performance of the classification model.

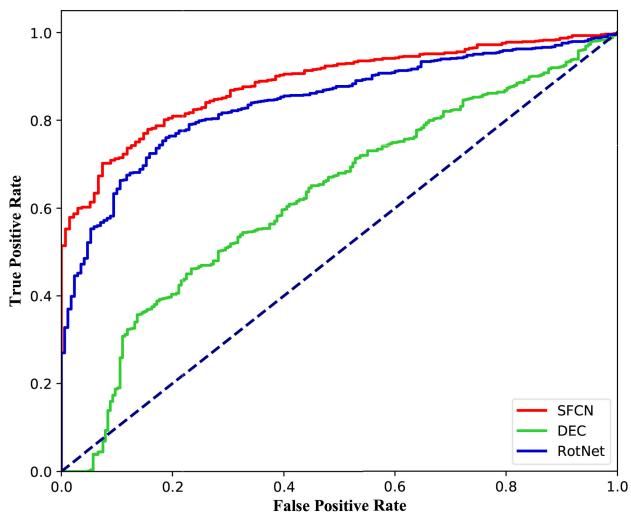
To further demonstrate the effectiveness of SFCN on unsupervised retinal image classification, we also demonstrate the network performance by **confusion matrix, loss and accuracy curves, feature map visualization** and the clustering visualization by t-SNE, which will be described in following paragraphs.

### C. PERFORMANCE EVALUATION

Following the training detail described in Section III, we supplement extensive experiments on the three available retinal image databases, and achieve comparative results on them. We conclude the results in Table 1, 2, and 3, along with

**TABLE 3.** Comparison results of accuracy for normal/abnormal retinal image classification on DIARETDB1 dataset.

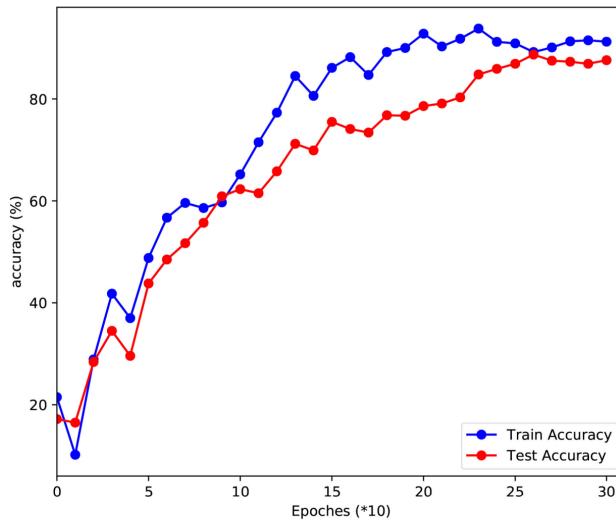
Models	Training	Accuracy (%)
Bourouis et al. [9](2018)	Supervised	98.7
DEC [41](2016)	Unsupervised	70.7
RotNet [19](2018)	Unsupervised	80.1
SKCN	Modified	78.8
SHCN	Modified	78.5
Directly FCM on features	Modified	79.3
<b>Our proposed SFCN</b>	<b>Modified</b>	<b>84.7</b>

**FIGURE 3.** ROC curves of unsupervised retinal image classification methods.

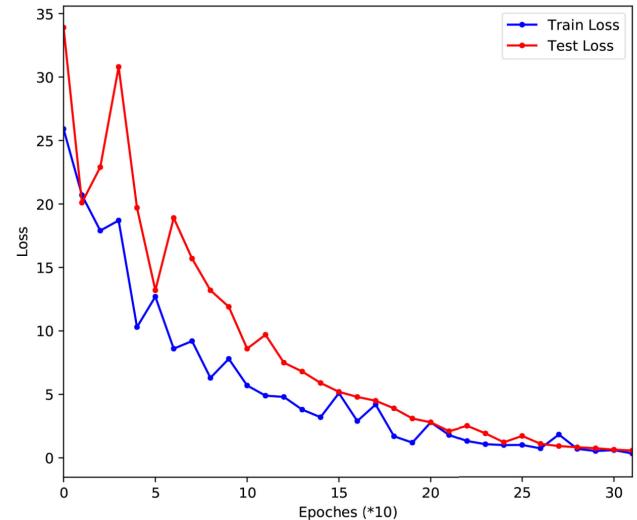
several related compared methods (including supervised and unsupervised frameworks).

It can be seen that the proposed self-supervised fuzzy clustering network achieves 87.6%, 81.7%, and 84.7% accuracies on MESSIDOR, DRIVE, and DIARETDB1 datasets, respectively. For MESSIDOR datasets, we compare our SFCN approach with seven supervised methods and one unsupervised method, and the best supervised method [9] obtains 100% accuracy and 1.00 AUC, which has a limited superiority to our SFCN. In contrast, our SFCN achieves the 87.6% accuracy and 0.85 AUC, while the best result (RotNet) of compared unsupervised methods obtains 82.3% accuracy and 0.837 AUC. Moreover, the ROC curves of unsupervised methods are illustrated in Figure 3, which can be seen that our SFCN has the best performance among them.

Though our method seems worst than the best supervised method, it is superior to the DEC and RotNet model, and leaves an acceptable distance between supervised methods. Note that, DEC method is a classical clustering method based on deep learning, and RotNet is a rotation based self-supervised learning approach. The comparison on MESSIDOR illustrates that our SFCN is not only outperform the clustering method (DEC),



(a) Accuracy on training and testing data



(b) Loss on training and testing data

**FIGURE 4.** Loss and Accuracy curves on MESSIDOR dataset. We choose the loss and accuracy value in each 10 epochs to clearly shows the tendency in network training.

but also for rotation-based self-supervised learning methods (RotNet). From Table 2 and 3, the same compared results can be concluded by their accuracies on DRIVE and DIARETDB1 datasets. That demonstrates our proposed SFCN model can reach a satisfied performance compared with supervised and unsupervised methods, while our approach does not require any labeled data to train such an efficient network. In comparison, it can be seen that our SFCN achieves better performance than a supervised method [17], which is a hand-crafted feature based method. That benefits from the robust feature learning ability of convolutional neural network. Moreover, the main reason of the performable results of SFCN is that our method combines the advantages of self-supervised learning based on deep convolutional neural network and the fuzzy C-means clustering with soft partition strategy, which is more realistic than the crisp division methods.

To further show the training efficiency of our network, we further introduce four criteria of the network, including confusion matrix, loss and accuracy curves, feature map visualization and t-SNE clustering plot, taking the performance on MESSIDOR dataset as a representation.

**Confusion Matrix** is an error matrix to judge the advantage of the classifier, indicating the predicted results whether they belongs to the real class. Figure 5 is the confusion matrix of MESSIDOR dataset, which we randomly choose 100 image per class. From the confusion matrix, the vertical coordinates are true annotations, and the horizontal coordinates denote the predicted labels of the SFCN. Our SFCN obtains 90 true predicted abnormal samples, and achieves 86 true predicted normal retinal samples, which carries out 88% accuracy on the chosen data. As for abnormal images,

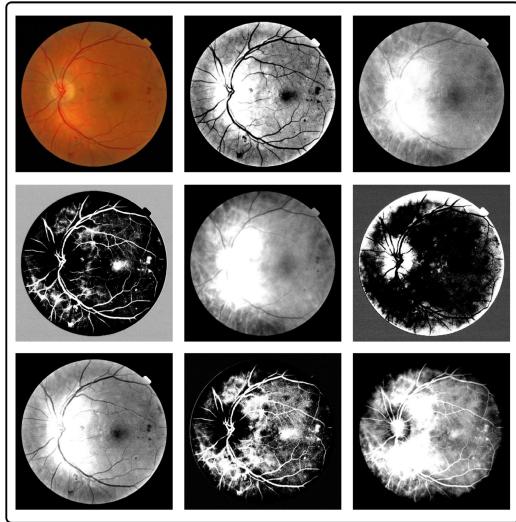
		Predicted Annotations	
		Abnormal	Normal
True Annotations	Abnormal	90	10
	Normal	14	86

**FIGURE 5.** Confusion matrix on MESSIDOR dataset. To better visualization, we randomly choose 100 images from abnormal and normal class respectively to conduct confusion matrix generation.

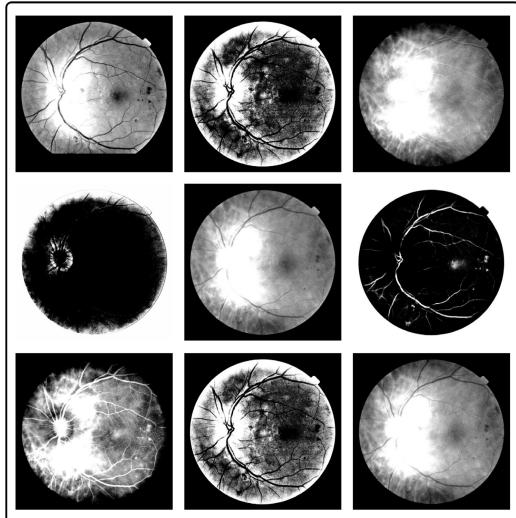
our model can diagnose 90% of them, which is a preferable performance in clinical applications.

**Loss and Accuracy curves** are the reflections of the network training efficiency. They can illustrate the variations of loss and accuracy along with training epochs. We draw the curves both for training and testing data in Figure 4. From the curves, the loss is reducing obviously along with training until it reaches 250 epochs, and the accuracy also achieves the optimal result in 250 epochs. The training trend of the model demonstrates our SFCN network can be easily convergence and has a desirable training efficiency.

**Feature Map** is a measure to evaluate the convolutional layers, which is a significant tool to reflect the learning ability inside CNN. We choose several feature maps produced

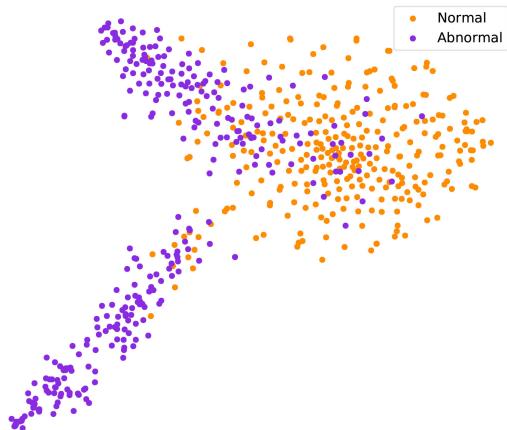


**FIGURE 6.** Feature map of the first convolutional layer on MESSIDOR dataset. We choose 8 representative feature maps from all generated feature maps for a retinal image.



**FIGURE 7.** Feature map of the second convolutional layer on MESSIDOR dataset. We choose 9 representative feature maps from all generated feature maps for the same retinal image in Figure 6.

by the first convolutional layer in the feature learning module to do visualization, as shown in Figure 6. The first picture in Figure 6 is the original retinal image, and others are its feature maps obtained by the first convolutional layer. To better understanding the feature maps for the convolutional layers, we also visualize the feature maps of the second convolutional layer, which is for the same retinal image (as shown in Figure 7). It can be seen the feature maps present the clear components about the diagnosis, such as the vessel trail, retinal edge, and the lesion area both of the first and second convolutional layers. That further proves the feature extracting ability of the proposed self-supervised fuzzy clustering network. Moreover, the feature visualization of the last layer of the feature learning module are analyzed by t-SNE, which performance can be observed in next paragraph.



**FIGURE 8.** t-SNE plot on MESSIDOR dataset. In detail, we randomly employ 300 samples from normal and abnormal retinal images individually, to learn their features, then utilize t-SNE on them to visualize in a 2-D dimensional space.

t-SNE is a dimension-reducing measure combining t-distribution and stochastic neighbor embedding to transform the highly dimensional feature into 2-D feature space. From t-SNE plot, it can present the feature representation ability of the last layer in feature learning module. We draw the t-SNE plot in Figure 8, and the result shows our SFCN model can clearly partition the abnormal and normal samples in the 2-D space. This is another crucial evidence to prove the robustness of the proposed classifier.

#### D. FURTHER ANALYSIS

##### 1) EVALUATION OF THE FUZZY C-MEANS ALGORITHM

To evaluate the effectiveness of using fuzzy C-means algorithm as the self-supervision guiding, we replace the fuzzy C-means algorithm by two crisp partition strategies, including K-means, and hierarchical clustering. The crisp clustering results of K-means and hierarchical clustering on the learned features are introduced into Eq.3, which are named by Self-supervised K-means Clustering Network (SKCN), and Self-supervised Hierarchical Clustering Network (SHCN), respectively. Their results are reported in Tables 1-3. For MESSIDOR dataset, SKCN and SHCN achieves accuracies of 82.8% and 81.7% respectively, with a distance of SFCN at least 4%, and the AUCs of them are also less than our proposed SFCN. Moreover, the accuracies on DRIVE and DIARETDB1 datasets also present that our SFCN is superior to the modified methods SKCN and SHCN. This comparison demonstrates the fuzzy C-means is more helpful than traditional script partition strategies, which is more realistic for the object category-attributes.

##### 2) EVALUATION OF SELF-SUPERVISION LOSS

To validate the effectiveness of the proposed fuzzy self-supervision module, we modify the SFCN by removing the self-supervision loss, and directly attach the FCM on the features obtained by feature learning module, which results

are shown in Table 1, 2 and 3. The complete SFCN achieves an improvement of 6.1%, 4.1%, and 5.4% than directly using FCM, on MESSIDOR, DRIVE and DIARTDB1 datasets, separately. In addition, this also obtains a promotion of 0.04 AUC on MESSIDOR dataset. These evaluated results are further proved that the proposed self-supervision mechanism can effectively improve the performance of retinal image classification in automatically abnormal diagnosing problem.

## V. CONCLUSION

In this paper, we propose a Self-supervised Fuzzy Clustering Network (SFCN) to solve the diabetic retinopathy classification task, which has attracted much research interest by supervised models. However, these existing methods require tremendous professional labor to annotate large amount data, which is very expensive in medical image analysis. Therefore, our proposed SFCN model focus on unsupervised retinal image classification by a feature learning module, reconstruction model, and a fuzzy self-supervision module to learn the unsupervised feature space. Extensive validating experiments are implemented on MESSIDOR, DRIVE, and DIARTDB1 datasets, which the analysis of results demonstrate the effectiveness of our SFCN approach.

## REFERENCES

- [1] A. M. Abdel-Zaher and A. M. Eldeib, "Breast cancer classification using deep belief networks," *Expert Syst. Appl.*, vol. 46, pp. 139–144, Mar. 2016.
- [2] M. D. Abràmoff, Y. Lou, A. Erginay, W. Clarida, R. Amelon, J. C. Folk, and M. Niemeijer, "Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning," *Investigative Ophthalmol. Vis. Sci.*, vol. 57, no. 13, pp. 5200–5206, Oct. 2016.
- [3] E. Ahn, A. Kumar, D. Feng, M. Fulham, and J. Kim, "Unsupervised deep transfer feature learning for medical image classification," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1915–1918.
- [4] E. Ahn, A. Kumar, D. Feng, M. Fulham, and J. Kim, "Unsupervised feature learning with K-means and an ensemble of deep convolutional neural networks for medical image classification," 2019, *arXiv:1906.03359*. [Online]. Available: <http://arxiv.org/abs/1906.03359>
- [5] A. M. Alqudah, "AOCT-NET: A convolutional network automated classification of multiclass retinal diseases using spectral-domain optical coherence tomography images," *Med. Biol. Eng. Comput.*, vol. 58, no. 1, pp. 41–53, Jan. 2020.
- [6] J. Amin, M. Sharif, M. Yasmin, H. Ali, and S. L. Fernandes, "A method for the detection and classification of diabetic retinopathy using structural predictors of bright lesions," *J. Comput. Sci.*, vol. 19, pp. 153–164, Mar. 2017.
- [7] A. M. Anter and M. Ali, "Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems," *Soft Comput.*, vol. 24, no. 3, pp. 1565–1584, Feb. 2020.
- [8] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Comput. Geosci.*, vol. 10, nos. 2–3, pp. 191–203, Jan. 1984.
- [9] S. Bourouis, A. Zagaria, N. Bouguila, and R. Alroobaea, "Deriving probabilistic SVM kernels from flexible statistical mixture models and its application to retinal images classification," *IEEE Access*, vol. 7, pp. 1107–1117, 2019.
- [10] D. Chen, Y. Chen, Y. Li, F. Mao, Y. He, and H. Xue, "Self-supervised learning for few-shot image classification," 2019, *arXiv:1911.06045*. [Online]. Available: <http://arxiv.org/abs/1911.06045>
- [11] V. Cherukuri, V. Kumar B. G., R. Bala, and V. Monga, "Deep retinal image segmentation with regularization under geometric priors," *IEEE Trans. Image Process.*, vol. 29, pp. 2552–2567, 2020.
- [12] P. Costa and A. Campilho, "Convolutional bag of words for diabetic retinopathy detection from eye fundus images," *IPSJ Trans. Comput. Vis. Appl.*, vol. 9, no. 1, pp. 1–6, 2017.
- [13] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordóñez, P. Massin, A. Erginay, B. Charton, and J.-C. Klein, "Feedback on a publicly distributed image database: The Messidor database," *Image Anal. Stereology*, vol. 33, no. 3, pp. 231–234, Aug. 2014.
- [14] A. Esteve, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Feb. 2017.
- [15] Z. Feng, C. Xu, and D. Tao, "Self-supervised representation learning by rotation feature decoupling," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 10364–10374.
- [16] A. D. Fleming, S. Philip, K. A. Goatman, G. J. Williams, J. A. Olson, and P. F. Sharp, "Automated detection of exudates for diabetic retinopathy screening," *Phys. Med. Biol.*, vol. 52, no. 24, p. 7385, 2007.
- [17] M. Garcia, R. Hornero, C. I. Sanchez, M. I. Lopez, and A. Diez, "Feature extraction and selection for the automatic detection of hard exudates in retinal images," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 4969–4972.
- [18] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, Jul. 2017.
- [19] S. Gidaris, P. Singh, and N. Komodakis, "Unsupervised representation learning by predicting image rotations," in *Proc. 6th Int. Conf. Learn. Represent. (ICLR)*, Vancouver, BC, Canada, Apr./May 2018, pp. 1–16.
- [20] S. Gupta and A. Karandikar, "Diagnosis of diabetic retinopathy using machine learning," *J. Res. Develop.*, vol. 3, no. 2, pp. 1–6, 2015.
- [21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [22] A. A. J. Johnson and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 694–711.
- [23] H. F. Jelinek, A. Rocha, T. Carvalho, S. Goldenstein, and J. Wainer, "Machine learning and pattern classification in identification of indigenous retinal pathology," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2011, pp. 5951–5954.
- [24] R. Kälviäinen and H. Uusitalo, "Diaretbd1 diabetic retinopathy database and evaluation protocol," in *Medical Image Understanding and Analysis*, vol. 2007. Edinburgh, U.K.: Citeseer, 2007, p. 61.
- [25] J. Krause, V. Gulshan, E. Rahimy, P. Karth, K. Widner, G. S. Corrado, L. Peng, and D. R. Webster, "Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy," *Ophthalmology*, vol. 125, no. 8, pp. 1264–1272, Aug. 2018.
- [26] S. V. Aruna Kumar, B. S. Harish, B. S. Mahanand, and N. Sundararajan, "An efficient meta-cognitive fuzzy C-Means clustering approach," *Appl. Soft Comput.*, vol. 85, Dec. 2019, Art. no. 105838.
- [27] H. Li and O. Chutatape, "A model-based approach for automated feature extraction in fundus images," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, Oct. 2003, p. 394.
- [28] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.
- [29] J. Nandy, W. Hsu, and M. Li Lee, "An incremental feature extraction framework for referable diabetic retinopathy detection," in *Proc. IEEE 28th Int. Conf. Tools with Artif. Intell. (ICTAI)*, Nov. 2016, pp. 908–912.
- [30] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "An ensemble deep learning based approach for red lesion detection in fundus images," *Comput. Methods Programs Biomed.*, vol. 153, pp. 115–127, Jan. 2018.
- [31] M. Perkonigg, D. Sobotka, A. Ba-Salamah, and G. Langs, "Unsupervised deep clustering for predictive texture pattern discovery in medical images," 2020, *arXiv:2002.03721*. [Online]. Available: <http://arxiv.org/abs/2002.03721>
- [32] G. Quellec, M. Lamard, A. Erginay, A. Chabouis, P. Massin, B. Cochener, and G. Cazuguel, "Automatic detection of referral patients due to retinal pathologies through data mining," *Med. Image Anal.*, vol. 29, pp. 47–64, Apr. 2016.
- [33] R. Raman, S. Srinivasan, S. Virmani, S. Sivaprasad, C. Rao, and R. Rajalakshmi, "Fundus photograph-based deep learning algorithms in detecting diabetic retinopathy," *Eye*, vol. 33, no. 1, pp. 97–109, Jan. 2019.
- [34] T. Shanthi and R. S. Sabeanian, "Modified alexnet architecture for classification of diabetic retinopathy images," *Comput. Electr. Eng.*, vol. 76, pp. 56–64, Jun. 2019.

- [35] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.
- [36] X. Tang, B. Du, J. Huang, Z. Wang, and L. Zhang, "On combining active and transfer learning for medical data classification," *IET Comput. Vis.*, vol. 13, no. 2, pp. 194–205, Mar. 2019.
- [37] H. Wang, W. Hsu, K. Guan Goh, and M. Li Lee, "An effective approach to detect lesions in color retinal images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2, Jun. 2000, pp. 181–186.
- [38] Y. Wang, J. Mei, L. Zhang, B. Zhang, P. Zhu, Y. Li, and X. Li, "Self-supervised feature learning with CRF embedding for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 2628–2642, May 2019.
- [39] Z. Wang, Y. Yin, J. Shi, W. Fang, H. Li, and X. Wang, "Zoom-in-net: Deep mining lesions for diabetic retinopathy detection," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Quebec City, QC, Canada: Springer, 2017, pp. 267–275.
- [40] J. Wu, Y. Zhang, J. Wang, J. Zhao, D. Ding, N. Chen, L. Wang, X. Chen, C. Jiang, X. Zou, X. Liu, H. Xiao, Y. Tian, Z. Shang, K. Wang, X. Li, G. Yang, and J. Fan, "AttenNet: Deep attention based retinal disease classification in OCT images," in *Proc. Int. Conf. Multimedia Modeling*. Daejeon, South Korea: Springer, 2020, pp. 565–576.
- [41] J. Xie, R. Girshick, and A. Farhadi, "Unsupervised deep embedding for clustering analysis," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 478–487.
- [42] J. Zhang, C.-G. Li, C. You, X. Qi, H. Zhang, J. Guo, and Z. Lin, "Self-supervised convolutional subspace clustering network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 5473–5482.
- [43] K. Zhou and S. Yang, "Effect of cluster size distribution on clustering: A comparative study of k-means and fuzzy c-means clustering," *Pattern Anal. Appl.*, vol. 23, no. 1, pp. 455–466, Mar. 2019.
- [44] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2223–2232.



**JING PAN** was born in 1986. She received the Ph.D. degree from Southwest University, in 2015. She is currently a Professor with the Department of Physics and Chemistry, Taiyuan University. She has a wealth of research experience in the synthesis of nanomaterials, battery electrode modified, math, optimization, computer vision, and coding.



**SHAO SHUAI FAN** is currently pursuing the bachelor's degree with the School of Cyber Security and Computer, Hebei University, Baoding, China. He is also an Intern with SIBD. His research interests include deep learning and medical image classification.



**ZEYU DU** is currently pursuing the bachelor's degree with the School of Cyber Security and Computer, Hebei University, Baoding, China. He is also an Intern with SIBD. His research interests include medical image analysis and pattern recognition.



**GUANGHUA ZHANG** was born in 1986. He received the Ph.D. degree in computer science and technology from Chongqing University, in 2017. He is currently the Director of the Shanxi Intelligent Big Data Industry Technology Innovation Research Institute. He is also a Professor with the Department of Computer Science and Engineering, Taiyuan University. His research interests include machine learning, multispectral image processing, computer vision, and medical images analysis.



**YUEGUO LUO** received the Ph.D. degree from Chongqing University, China, in 2017. Since 2004, he has been working as a Teacher with Yangtze Normal University, Fuling, Chongqing. He is currently an Associate Professor. His research interests include computational intelligence and biology computing.