Chinese Journal of Electronics
vol. 33, no. 4, pp. 1023–1033, July 2024
https://doi.org/10.23919/cje.2023.00.081

RESEARCH ARTICLE

FGM-SPCL: Open-Set Recognition Network for Medical Images Based on Fine-Grained Data Mixture and Spatial Position Constraint Loss

Ruru ZHANG[1,2], Haihong E[1,2], Lifei YUAN[3], Yanhui WANG[3], Lifei WANG[3], and Meina SONG[1,2]

1. *School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China*
2. *Engineering Research Center of Information Networks, Ministry of Education, Beijing 100876, China*
3. *Hebei Eye Hospital, Xingtai 054001, China*

Corresponding author: Meina SONG, Email: mnsong@bupt.edu.cn
Manuscript Received March 15, 2023; Accepted August 7, 2023

Abstract — The current intelligent auxiliary diagnosis models all follow the closed-set recognition setting. After the model is deployed online, the input data is often not completely controlled. Diagnosing an untrained disease as a known category would lead to serious medical malpractice. Therefore, realizing the open-set recognition is significant to the safe operation of the intelligent auxiliary diagnosis model. Currently, most open-set recognition models are studied for natural images, and it is very challenging to obtain clear and concise decision boundaries between known and unknown classes when applied to fine-grained medical images. We propose an open-set recognition network for medical images based on fine-grained data mixture and spatial position constraint loss (FGM-SPCL) in this work. Considering the fine graininess of medical images and the diversity of unknown samples, we propose a fine-grained data mixture (FGM) method to simulate unknown data by performing a mixing operation on known data to expand the coverage of unknown data difficulty levels. In order to obtain a concise and clear decision boundary, we propose a spatial position constraint loss (SPCL) to control the position distribution of prototypes and samples in the feature space and maximize the distance between known classes and unknown classes. We validate on a private ophthalmic OCT dataset, and extensive experiments and analyses demonstrate that FGM-SPCL outperforms state-of-the-art models.

Keywords — Few-shot class-incremental learning, Embedding augmentation, Classifier adaptation, Image classification.

Citation — Ruru ZHANG, Haihong E, Lifei YUAN, *et al.*, "FGM-SPCL: Open-Set Recognition Network for Medical Images Based on Fine-Grained Data Mixture and Spatial Position Constraint Loss," *Chinese Journal of Electronics*, vol. 33, no. 4, pp. 1023–1033, 2024. doi: 10.23919/cje.2023.00.081.

I. Introduction

In recent years, deep learning has achieved or exceeded human-level performance in medical image recognition and classification tasks [1]–[3]. Most current intelligent auxiliary diagnosis models are developed for one or a few specific diseases in a closed-set recognition setting, which assumes that all test classes are known or seen during training. However, actual medical scenarios are usually dynamic and open. After the model is deployed and launched, it is inevitable to encounter situations such as disease samples from unknown sources, changes in hardware equipment, and unqualified differentiated samples. Since these types of samples are not in the training set, traditional deep neural networks cannot detect unknown classes while assigning high confidence scores to classify them as one of the known classes, resulting in misdiagnosis or missed diagnosis. To address these issues, researchers have proposed open-set recognition (OSR) to robustly identify unknown diseases in test samples, while correctly classifying known diseases.

Since the unknown class has never been seen in the training process of OSR, improving the distribution differences between unknown and known classes in the fea-

Associate Editor: Prof. Li ZHUO, Beijing University of Technology.

ture space and clarifying the decision boundary between known and unknown classes is a crucial issue. One of the most common research approaches is to learn the boundaries of known class regions and identify the remaining feature space as unknown. Among them, spatial location constraint prototype loss (SLCPL) [4] pointed out that unknown categories usually tend to be distributed in the central area of the feature space, so the study controlled the distribution of known category features at the edge of the feature space to clarify the decision boundary between known and unknown categories. However, focusing on controlling the potential spatial distribution of known categories may not be enough to reduce open space risk solely based on limited known categories during testing. Therefore, researchers have made great efforts to simulate virtual open sets [5]–[7] or virtual open spaces [8]. Among them, using generative adversarial network (GAN) generators [5], [9] to generate virtual data or features is a common method. However, difficulty-aware simulator (DIAS) [6] pointed out that images generated using GAN are not very challenging for classifiers, and introduced the Copycat method to generate samples of different levels of difficulty to simulate open sets. However, since it is impossible to accurately obtain the distribution of unknown classes, simply separating the generated virtual images from known classes does not guarantee the effectiveness for other unknown samples. For this reason, reference [8] combined the above two ideas, proposed adversarial reciprocal point learning (ARPL) to model the potential open space of each known class in the feature space, and proposed an adversarial marginal constraint to reduce the open space risk. Meanwhile, an instantiated adversarial enhancement method is designed to generate diverse virtual open-set samples.

Although researchers have made great contributions to the OSR task, there are still many problems to be explored in medical scenarios. We find that for fine-grained natural images, there are visually distinguishable features between categories (e.g., dog, car, airplane, etc.), making detecting unknown samples relatively easy. While for fine-grained medical images, the unknown samples span a wider area and are more similar to the features of known samples. Therefore, the medical imaging OSR task faces the following problems: 1) The deep learning model is highly dependent on discriminative features to distinguish categories, and the virtual unknown classes generated by the method based on feature simulation cannot cover the area of fine-grained open data. Thus, classifiers learned through them are still vulnerable to unknown classes with similar semantics. 2) Due to the mismatch between the unknown class and the model convolution kernel, the unknown class usually tends to be distributed in the central area of the feature space [4], so most studies have improved the open-set detection performance by making the known class far away from the center of the space. But for fine-grained images, the distribution of unknown classes is relatively wide, and the decision boundary in the feature space is relatively blurred. Therefore, maximizing the decision boundary between known and unknown classes is the key challenge we are currently facing.

Based on the above problems, we propose a new framework OSR network for medical images based on fine-grained data mixture and spatial position constraint loss (FGM-SPCL). We design a simple yet effective fine-grained data mixture (FGM) approach to predict unknown class instances. Specifically, we propose to perform a blending operation between known class data to obtain virtual unknown samples. These samples can be close to or far from known class samples, which aligns more with the real-world distribution of unknown classes. At the same time, this method can promote the model to learn more discriminative features, realize the correction of known class distribution boundaries, and learn a more compact and robust feature space. To further increase the distance between known and unknown classes, we propose a spatial position constraint loss (SPCL), which includes a prototype position constraint loss (PPCL) and an asymmetric instance contrastive loss (AICL). First, we use PPCL to constrain the location distribution of prototypes and samples in the feature space, forcing all known classes to be distributed in the peripheral area of the space and virtual unknown classes to be distributed between the center of the feature space and the known class distribution area. In this way, the virtual unknown data can play a role in isolating the real unknown data and the known data. Then, we use AICL to better cluster samples of known classes and push them away from virtual unknown samples to clarify the boundary between the known and unknown classes. In addition, we collaborated with a Grade III Level A ophthalmology hospital to collect an ophthalmic OCT dataset and designed various experiments on this dataset. At the same time, to verify the generality of the model, we conduct experiments on an additional public medical dataset and three natural image datasets. Experimental results show that the performance of FGM-SPCL for fine-grained OSR reaches the state-of-the-art level, reflecting the excellent robustness of the model. Our main contributions are summarized as follows:

1) A simple and effective FGM method is proposed to predict open-set data to obtain virtual unknown class samples covering the range of open areas;

2) An SPCL is proposed to maximize the distance between the known class and the unknown class to obtain a concise and clear decision boundary;

3) Experimental results on two medical datasets and three natural image datasets show that FGM-SPCL achieves state-of-the-art performance, reflecting the robustness of the model to unknown samples.

II. Related Work

Early work on the OSR task mainly used traditional methods such as support vector machines [10], extreme

value theory (EVT), and nearest-to-class mean classifiers [11]–[13]. Then, deep learning algorithms are widely adopted with the development of convolutional neural network (CNN). Regarding the study of OSR in this paper, its related work can be roughly divided into two categories: OSR based on prototype classifiers and OSR based on generative models.

1. OSR based on generative models

Generator methods use GANs, auto-encoders [14], flow-based models [15], and data augmentation-based methods [7] to generate unknown or known samples to help the classifier learn the decision boundary between known and unknown classes. Reference [16] proposed a classification-reconstruction learning for open-set recognition (CROSR), which exploits latent representations for reconstruction and enables unknown detection without compromising the classification accuracy of known classes. Reference [17] proposed the class conditioned auto-encoder (C2AE) algorithm, which uses a class-conditional auto-encoder with a novel training and testing method to model the reconstruction error by EVT to find the threshold for identifying known/unknown class samples. Reference [14] proposed the conditional Gaussian distribution learning (CGDL) algorithm to classify known samples by forcing different latent features to approximate different Gaussian models. Reference [15] proposed the OpenHybrid framework, which implements an embedding space jointly learned by a classifier and a flow-based density estimator. Reference [7] proposed PlaceholdeRs for open-set recognition (PROSER), which prepares for unknown categories by assigning placeholders to data and classifiers. Concretely, the method efficiently generates new classes through manifold mixing and adaptively sets the value of the retained open-set classifier during training. Reference [8] proposed ARPL. Specifically, each interchange point is learned via an open space with the corresponding known category, and the open space is augmented by generated fake samples. To obtain open-set data with different difficulties, Reference [6] proposed a DIAS, which can generate fake samples with varying levels of difficulty to simulate the real world. This method generates medium-difficulty samples via GAN and easy and hard samples via the Copycat form. However, the implementation process of this method is more complicated, and it also increases the complexity of the model.

2. OSR based on prototype classifiers

In order to obtain more discriminative feature representations, some methods exploit prototypes to represent each known class in the embedded feature space, encourage the features of training samples to be close to the corresponding prototypes, and identify open-set images based on the distance to the prototypes. Reference [18] proposed a convolutional prototype network (CPN) and designed a new discriminative loss and generation loss to increase the inter-class distance and reduce the intra-class distance. Reference [19] proposed prototype-based open deep network (P-ODN), which jointly trains prototypes and prototype radii to guide CNN to obtain more discriminative features, and applies a multi-class triplet thresholding method based on the distance metric between features and prototypes to detect unknowns. Reference [20] proposed generalized convolutional prototype learning (GCPL) with a prototype loss as a regularization method to improve the intra-class compactness of feature representations. Prototype-based methods are crucial to the selection of prototype sets, so [21] proposed a new prototype mining and learning (PMAL) framework considering the high quality and diversity of prototype sets. The framework selects prototype sets based on data uncertainty learning and a diversity-based prototype set filtering strategy. In order to further enlarge the decision boundary between known classes and unknown classes, Reference [4] pointed out that unknown classes usually tended to be distributed in the central area of the feature space and proposed spatial location constraint prototype loss to control the distribution of known class prototypes in the feature space edge area. Reference [22] proposes a method of reciprocal point learning (RPL), which constructs a bounded space based on reciprocal points, and at the same time, it can indirectly introduce unknown information into a feature space to learn more compact and discriminative representations. However, the above studies are still unable to achieve a clearer decision boundary for the OSR of fine-grained medical images.

We propose a simple and effective FGM method to generate virtual unknown samples spanning open-set intervals. And through the SPCL, the decision boundary between the known class and the unknown class is maximized on the basis of the FGM method. This significantly improves the robustness of the model.

III. Methods

1. Problem statement

We formalize the OSR problem and emphasize its differences from closed-set recognition. First, the model is trained using $\mathcal{D}_{\text{tr}} = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^{N} \subset \mathcal{X} \times C$, where $\boldsymbol{x}_i$ is an image, $y_i \in \{1, 2, \ldots, C\}$ is its class label, N is the total number of training samples, $\mathcal{X}$ is the input space of the images, and C is the set of "known" classes. In the closed-set assumption, use $\mathcal{D}_{\text{te}} = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^{M} \subset \mathcal{X} \times C$ for testing. Here, $\mathcal{D}_{\text{tr}}$ and $\mathcal{D}_{\text{te}}$ come from the same known dataset $\mathcal{D}_{\text{in}}$, and M is the total number of test samples. Conversely, in the open-set setting, the test images may also come from the unknown dataset $\mathcal{D}_u$, using $\mathcal{D}_{\text{teopen}} = \mathcal{D}_{\text{te}} \cup \mathcal{D}_u = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^{M'} \subset \mathcal{X} \times (C \cup U)$ to test. Here M' is the total number of test samples of known and unknown classes, U is the set of 'unknown' classes, and the known class and the unknown class test sets are represented by $\mathcal{D}_{\text{te}}$ and $\mathcal{D}_u$, respectively.

2. Overview

Prototype-based classifiers can reduce the intra-class distance and increase the inter-class distance, thereby improving the robustness of the classifier, which is proved to be very effective for learning fine-grained features [23]. Therefore, we design our network framework based on the prototype classifier, as shown in Figure 1.

First, we obtain virtual unknown classes with a broader range of difficulty through FGM and participate in model training together with known classes. We use ResNet-34 [24] and Classifier32 [4], [8] as the backbone network F and map the image to a feature vector $\boldsymbol{F}(\boldsymbol{x})$ (the feature of the last global pooling layer is used as the final feature vector). Among them, Classifier32 is a commonly used backbone network in OSR-related algorithms. It is worth noting that after obtaining the virtual unknown class $\mathcal{D}_{\text{virtual}}$ by $\boldsymbol{x} = \text{FGM}(\boldsymbol{x})$, the total number of categories is expanded from C to N, and we set a prototype for each category, that is, $\boldsymbol{O} = \{\boldsymbol{O}^k, k = 1, 2, \ldots, N\}$. The N prototype points are initialized from a uniform distribution. For any data $\boldsymbol{x}$, label y belongs to k classes, and its prototypical contrastive loss (PCL) can be expressed as

$$\text{PCL}(\boldsymbol{x}) = -\log \frac{\exp\left(-\text{d}\left(\tilde{\boldsymbol{F}}(\boldsymbol{x}), \tilde{\boldsymbol{O}}^k\right)\right)}{\sum_{i=1}^{N} \exp\left(-\text{d}\left(\tilde{\boldsymbol{F}}(\boldsymbol{x}), \tilde{\boldsymbol{O}}^i\right)\right)} \tag{1}$$

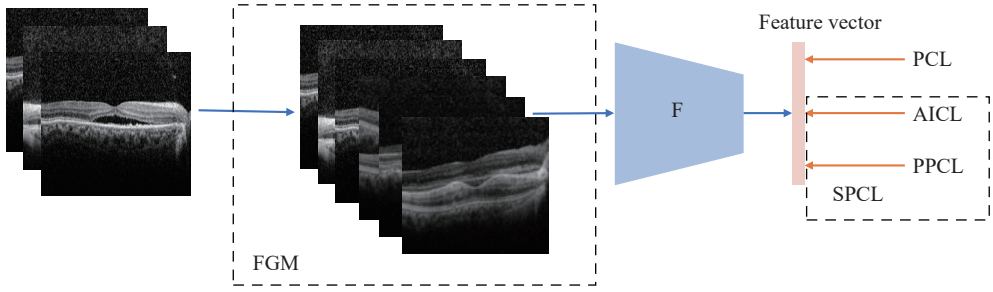


Figure 1 Overall network architecture.

Among them, $\tilde{\boldsymbol{F}}(\boldsymbol{x}) = \boldsymbol{F}(\boldsymbol{x})/\|\boldsymbol{F}(\boldsymbol{x})\|_2$ and $\tilde{\boldsymbol{O}}^k = \boldsymbol{O}^k/\left\|\boldsymbol{O}^k\right\|_2$ means feature vectors and prototypes are normalized. $\text{d}\left(\tilde{\boldsymbol{F}}(\boldsymbol{x}), \tilde{\boldsymbol{O}}^k\right)$ is the Euclidean distance between $\tilde{\boldsymbol{F}}(\boldsymbol{x})$ and $\tilde{\boldsymbol{O}}^k$.

Based on PCL($\boldsymbol{x}$), an SPCL is proposed, including AICL and PPCL. Therefore, the overall loss function of the model is expressed as

$$\mathcal{L}(\boldsymbol{x}) = \text{PCL}(\boldsymbol{x}) + \text{AICL}(\boldsymbol{x}) + \text{PPCL}(\boldsymbol{x}) \tag{2}$$

By minimizing $\mathcal{L}(\boldsymbol{x})$, each sample is clustered around the corresponding prototype, the known classes are distributed in the peripheral area of the space, and the virtual unknown classes are distributed between the center of the feature space and the known class distribution area. It makes the decision boundary between known and unknown classes more explicit. The distance between the sample and the nearest prototype is the criterion. If the distance is below a threshold, the sample is considered to belong to the corresponding category. Otherwise, it is considered unknown.

3. Fine-grained data mixture (FGM)

The goal of FGM is to transform the closed medical image training set into an open training set. This module should have two main features: 1) For fine-grained medical images, produce a virtual unknown dataset that can cover a wider open space; 2) The generation process should be simple and fast without any additional model complexity.

Mixup has become very popular in the field of data augmentation and has been proved to be effective in alleviating adversarial disturbances in neural networks, better-estimating uncertainty, and helping models create better decision boundaries between different categories. Specifically, each new example consists of two randomly sampled examples $(\boldsymbol{x}_i, y_i)$ and $(\boldsymbol{x}_j, y_j)$, weighted by linear interpolation as follows:

$$\begin{aligned} \hat{\boldsymbol{x}} &= \lambda \boldsymbol{x}_i + (1-\lambda)\boldsymbol{x}_j \\ \hat{y} &= \lambda y_i + (1-\lambda) y_j \end{aligned} \tag{3}$$

where $\lambda \in [0, 1]$ is sampled from a Beta distribution, controlling the contribution of each sample to the mixture. This generates new instances across multiple decision boundaries, driving the feature vectors of known classes to be more compact.

Inspired by Mixup, we randomly sample two samples, $\boldsymbol{x}_a$ and $\boldsymbol{x}_b$, from any two different classes, a and b, and generate virtual unknown class samples by weighted linear interpolation: $\boldsymbol{x}_{ab} = \lambda \boldsymbol{x}_a + (1-\lambda)\boldsymbol{x}_b$. The difference from Mixup is that we hope the mixed samples produced by every two categories are an independent virtual class. For example, if there are C categories, $C \times (C-1)/2$ virtual classes will be generated after every two classes are mixed. Therefore, we changed a C-class classification task into $N = C + C \times (C-1)/2$-class classification. To demonstrate the effect of mixing samples, we visualize seen, virtual, and unknown classes in the feature space, as shown in Figure 2. Gray areas

represent real samples of unknown classes, and pink areas represent generated virtual samples, the remaining 8 colored clusters correspond to each known class. It can be found that the virtual samples can cover the area where the real unknown samples are located, and also cover the area where the known samples are located. In this way, not only the possible distribution of unknown classes can be predicted, but also in order to better distinguish known classes from virtual unknown classes, the model will be prompted to learn more discriminative information for each class, thereby correcting the known class decision boundary, and further improve the recognition accuracy of known classes.

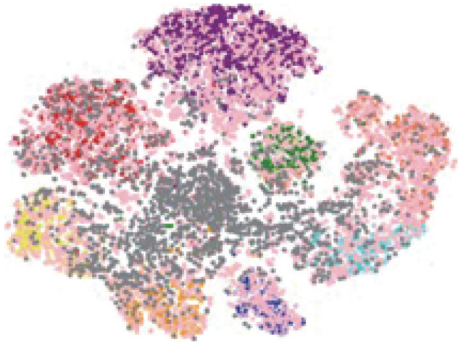


Figure 2 Visualize the feature distribution of real unknown classes and virtual unknown classes.

4. Spatial position constraint loss (SPCL)

SPCL aims to learn a constrained feature space that reduces the overlap of known and unknown class features. SPCL mainly includes PPCL and AICL.

Prototype position constraint loss (PPCL) Reference [4] pointed out that unknown classes usually tend to be distributed in the central region of the feature space. Therefore, we constrain all known class prototypes to the edge region of the feature space and make the prototypes as far as possible from the feature center. The implementation method is as follows:

$$
\begin{aligned}
&\boldsymbol{O}_{\text{center}} = \frac{1}{N}\sum_{k=1}^{N}\tilde{\boldsymbol{O}}^k \\
&r_k = \mathrm{d}\left(\tilde{\boldsymbol{O}}_I^k, \boldsymbol{O}_{\text{center}}\right) \\
&R_I = \frac{1}{C}\sum_{k=1}^{C} r_k \\
&\mathrm{PPCL}\left(R_I\right) = -\log R_I \\
&\mathrm{PPCL}\left(\boldsymbol{O}_I\right) = \frac{1}{C-1}\sum_{k=1}^{C}\left(r_k - R_I\right)^2
\end{aligned} \tag{4}
$$

Here, $\tilde{\boldsymbol{O}}^k = \frac{\boldsymbol{O}^k}{\|\boldsymbol{O}^k\|_2}$ means to normalize the prototype of the k-th class, and $\boldsymbol{O}_{\text{center}}$ represents the center of all prototypes, which is also the center of feature space. $\boldsymbol{O}_I = \boldsymbol{O}[:C,:]$ represents the prototypes for the known classes. r_k represents the distance of the k-th prototype from the center $\boldsymbol{O}_{\text{center}}$. R_I represents the average Euclidean distance between $\boldsymbol{O}_I$ and $\boldsymbol{O}_{\text{center}}$. We want $\boldsymbol{O}_I$ to be as far away from $\boldsymbol{O}_{\text{center}}$ as possible, and achieve this by $\mathrm{PPCL}\left(R_I\right)$, where a smaller value of $\mathrm{PPCL}\left(R_I\right)$ indicates a larger R_I, meaning that $\boldsymbol{O}_I$ is farther away from $\boldsymbol{O}_{\text{center}}$. To ensure that all prototypes for the known classes are far away from $\boldsymbol{O}_{\text{center}}$, we calculate the variance of the distance between $\boldsymbol{O}_I$ and $\boldsymbol{O}_{\text{center}}$ using $\mathrm{PPCL}\left(\boldsymbol{O}_I\right)$. A smaller variance indicates that the size of r_k, $k \in \{1, 2, \ldots, C\}$ is more similar, i.e., the circular distribution of known class prototypes in the peripheral region of the feature space is more regular, ensuring that all known class prototypes are far away from the center of the feature space.

For fine-grained images, there are more similar semantic features between unknown classes and known classes. Therefore, we hope that the decision boundary between known and unknown classes can be further enlarged. We distribute the virtual unknown classes in the region between the center of the feature space and the distribution of known classes. In this way, the real unknown class distribution space is compressed and isolated from the known class distribution space. The implementation method is as follows:

$$
\begin{aligned}
&r_{k'} = \mathrm{d}\left(\tilde{\boldsymbol{O}}_O^{k'}, \boldsymbol{O}_{\text{center}}\right) \\
&R_O = \frac{1}{N-C}\sum_{k'=1}^{N-C} r_{k'} \\
&\mathrm{PPCL}\left(R_O\right) = -\log R_O \\
&\mathrm{PPCL}\left(\boldsymbol{O}_O\right) = \frac{1}{N-C-1}\sum_{k'=1}^{N-C}\left(r_{k'} - R_O\right)^2 \\
&\mathrm{PPCL}(R) = \max\left(R_O - R_I + m\sqrt{D}, 0\right)
\end{aligned} \tag{5}
$$

Among them, $\boldsymbol{O}_O = \boldsymbol{O}[C:,:]$ represents virtual class prototypes. $r_{k'}, R_O, \mathrm{PPCL}\left(R_O\right)$, and $\mathrm{PPCL}\left(\boldsymbol{O}_O\right)$ have similar meanings to $r_k, R_I, \mathrm{PPCL}\left(R_I\right)$, and $\mathrm{PPCL}\left(\boldsymbol{O}_I\right)$, respectively. $\mathrm{PPCL}(R)$ means to limit R_I to be a little larger than R_O so that in the feature space, the virtual data can isolate the distribution area of known class data and unknown class data to a certain extent. In order to easily set the hyper-parameters, the margin is designed as $m\sqrt{D}$. D represents the feature dimension, and m is a hyper-parameter used to adjust the distribution distance between known classes and virtual unknown classes in the feature space. PPCL can be expressed as

$$
\begin{aligned}
\mathrm{PPCL}(\boldsymbol{x}) =& \alpha\mathrm{PPCL}\left(\boldsymbol{O}_I\right) + \beta\mathrm{PPCL}\left(\boldsymbol{O}_O\right) + \gamma\mathrm{PPCL}(R) \\
&+ \mu\left(\mathrm{PPCL}\left(R_I\right) + \mathrm{PPCL}\left(R_O\right)\right)
\end{aligned} \tag{6}
$$

where, $\alpha, \beta, \gamma, \mu$ are hyper-parameters.

Asymmetric instance contrast loss (AICL) The traditional instance contrast loss strengthens the intra-class compactness for each category in the feature space and pushes out other categories. In contrast, we hope that

only the intra-class compactness of known classes can be strengthened and push away the virtual unknown classes as a whole. In this way, we further widen the distance from unknown classes while calibrating the decision boundary of known classes. The implementation process of AICL is as follows:

$$\begin{aligned}&\mathrm{AICL}(\boldsymbol{x})\\&=-\frac{1}{|\mathcal{P}(\boldsymbol{x})|}\sum_{\boldsymbol{p}\in\mathcal{P}(\boldsymbol{x})}\log\frac{\exp\left(\boldsymbol{F}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{F}(\boldsymbol{p})/\tau\right)}{\sum_{\boldsymbol{a}\in\mathcal{A}(\boldsymbol{x})}\exp\left(\boldsymbol{F}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{F}(\boldsymbol{a})/\tau\right)}\end{aligned}\tag{7}$$

Among them, $\boldsymbol{F}(\boldsymbol{x})$, $\boldsymbol{F}(\boldsymbol{p})$, and $\boldsymbol{F}(\boldsymbol{a})$ represent the feature vectors of different types of samples. $\boldsymbol{F}(\boldsymbol{x})^{\mathrm{T}}$ represents the transpose operation on $\boldsymbol{F}(\boldsymbol{x})$, $\mathcal{P}(\boldsymbol{x})=\{\boldsymbol{p}\mid y(\boldsymbol{p})=y(\boldsymbol{x}),\boldsymbol{p}\in\mathcal{D}_{\mathrm{in}}\backslash\{\boldsymbol{x}\}\}$ is the positive sample set of $\boldsymbol{x}$, that is, samples with the same ground-truth labels as $\boldsymbol{x}$, $\mathcal{A}(\boldsymbol{x})=\mathcal{D}_{\mathrm{in}}\cup\mathcal{D}_{\mathrm{virtual}}\backslash\{\boldsymbol{x}\}$ is the set of all negative samples about $\boldsymbol{x}$, that is, samples with different ground truth labels from $\boldsymbol{x}$ and all virtual unknown samples. Finally, the contrastive learning loss of all samples is summed and normalized by $\frac{1}{|\mathcal{P}(\boldsymbol{x})|}$, and the scaling temperature τ is added as in contrastive learning.

IV. Experiments

1. Dataset

Ophthalmic OCT dataset We collected electronic medical records (EMRs) of 11829 patients from the Grade III Level A ophthalmology hospital and extracted 7917 ophthalmic OCT images. We built an ophthalmic image labeling system, and 11 ophthalmologists participated in the labeling work. The dataset contains 13 ophthalmic diseases, including epiretinal membranes (ERMs, 2083), central serous chorioretinopathy (CSCR, 1436), age-related macular degeneration (AMD, 876), macular hole (MH, 763), branch retinal vein occlusion (BRVO, 503), macular splitting (MS, 480), retinal detachment (RD, 436), vitreous macular traction syndrome (VMT, 392), retinal artery occlusion (RAO, 304), polypoidal choroidal vasculopathy (PCV, 258), central retinal vein occlusion (CRVO, 162), papilledema (PE, 139), and Harada disease (HD, 85). We take 8 of these categories as known datasets and divide the dataset into 80%–20% training-testing parts. The rest of the classes are treated as unknown class datasets.

HyperKvasir dataset It is one of the largest publicly available gastrointestinal endoscopy datasets under CC BY 4.0 (Creative Commons Attribution 4.0 International) [25]. The dataset includes labeled images, segmented images, unlabeled images, and labeled videos, and we choose labeled images among them for experiments. The dataset contains a total of 10662 labeled images, 23 categories, including BBPS 2–3 (1148), Polyps (1028), Cecum (1009), Dyed lifted polyps (1002), Pylorus (999), Dyed resection margins (989), Z-line (932), Retroflex stomach (764), BBPS 0–1 (646), Ulcerative colitis grade-2 (443), Esophagitis grade A (403), Retroflex rectum (391), Esophagitis grade B–D (260), Ulcerative colitis grade-1 (201), Ulcerative colitis grade-3 (133), Impacted stool (131), Barrett's short segments (53), Barretts (41), Ulcerative colitis grade-0–1 (35), Ulcerative colitis grade-2–3 (28), Ulcerative colitis grade-1–2 (11), Ileum (9), and Hemorrhoids (6). We take 8 of these categories as known datasets and divide the dataset into 80%–20% training-testing parts. The rest of the classes are treated as unknown class datasets.

CIFAR10, CIFAR+10, and CIFAR+50 datasets CIFAR10 [26] contains 60000 images divided into 10 classes. Each class has 6000 images, with 5000 images used for training, and 1000 for testing. CIFAR100 [26] contains 60000 images divided into 100 classes, with 600 images per class. Of these, 500 images are used for training and 100 for testing. In this experiment, CIFAR10 represents a dataset in which 6 classes are randomly selected as known classes and the remaining 4 classes are unknown. CIFAR+10 and CIFAR+50 represent datasets in which 4 classes are randomly selected from CIFAR10 as known classes, and 10 and 50 classes, respectively, are randomly selected from CIFAR100 as unknown classes.

2. Implementation details and evaluation metrics

Ophthalmic OCT dataset experiment setup We use ResNet-34 and Classifier32 as backbone architectures and train the models using the Adam optimizer with an initial learning rate of 0.01. We trained all the models for 200 epochs with a batch size of 128. We resize the input image to $224\times224\times3$ and apply standard data augmentation with random cropping and horizontal cropping. Models are trained only on known categories. In whole experiment, $\alpha=\beta=0.5,\gamma=\mu=0.1$, and $m=0.003$.

HyperKvasir dataset experiment setup We use ResNet-34 as the backbone architecture. In whole experiment, $\alpha=\beta=0.1$, $\gamma=\mu=0.1$, and $m=0.003$. Other settings are the same as the OCT dataset.

CIFAR10, CIFAR+10, and CIFAR+50 datasets experiment setup We use Classifier32 as the backbone architecture. All classifiers are trained with the momentum stochastic gradient descent (SGD-M) optimizer. The initial learning rate of the network is set to 0.1, dropping to one-tenth of the original rate every 30 epochs, and we train the network for 100 epochs with a batch size of 128. We resize the input image to 32×32 and apply standard data augmentation with random cropping and horizontal cropping. In whole experiment, $\alpha=\beta=0.1$, $\gamma=\mu=0.1$, and $m=0.003$.

We use accuracy (Acc) as the performance evaluation metric of the closed set, that is, the known class, and use area under the receiver operating characteristic (AUROC) curve and open-set classification rate (OSCR) as the evaluation metrics of the open set. Among them, AUROC is the most commonly used standard measure to measure the performance of OSR task models. OSCR

weighs correct classification rate (CCR) and false positive rate (FPR). Specifically, let δ be a score threshold. The CCR is the sample fraction that correctly classifies the image as c and has the maximum probability, and the probability is greater than δ:

$$\begin{aligned}&\mathrm{CCR}(\delta)\\&=\frac{|\{\boldsymbol{x}\mid \boldsymbol{x}\in\mathcal{D}_{\mathrm{te}}^{c}\wedge \operatorname{argmax}_{c} P(c\mid \boldsymbol{x})=\hat{c}\wedge P(\hat{c}\mid \boldsymbol{x})\geq\delta\}|}{|\mathcal{D}_{\mathrm{te}}^{c}|}\end{aligned}\tag{8}$$

FPR is the fraction of samples from unknown data $\mathcal{D}_u$ that are classified into any known class c with probability greater than δ:

$$\mathrm{FPR}(\delta)=\frac{|\{\boldsymbol{x}\mid \boldsymbol{x}\in\mathcal{D}_{u}\wedge \max_{c} P(c\mid \boldsymbol{x})\geq\delta\}|}{|\mathcal{D}_{u}|}\tag{9}$$

Therefore, OSCR based on CCR and FPR is an indicator similar to AUROC. The larger the OSCR value of the model, the stronger the recognition ability of the model.

3. State-of-the-art methods

In order to better evaluate the overall performance of our model, in addition to the introduced baselines, we also compare FGM-SPCL with state-of-the-art methods in the OSR domain, including Softmax, G-OpenMax [9], open-set recognition with counterfactual images (OSRCI) [27], CROSR [16], C2AE [17], DIAS [6], RPL [22], GCPL [20], SLCPL [4], ARPL+confusing samples (CS) [8], and maximum logit score (MLS) [28].

The Softmax model is trained with the cross-entropy loss on the known classes, and during testing, the maximum value of the Softmax probability vector is used to determine whether the input belongs to a known class. MLS uses the maximum logit score as an open-set score rule. GCPL learns generalized convolutional prototypes that effectively improve intra-class feature compactness, bringing the same samples closer around corresponding prototypes and thus calibrating known class decision boundaries. The PCL baseline model, compared to the calculation process of the prototype contrastive loss in the GCPL model, the prototype in the PCL uses uniform distribution initialization and normalization operations. RPL introduces the concept of reciprocal points to model the potential open spaces of each known class in the feature space. SLCPL controls the distribution of known class prototypes in the marginal regions of the feature space.

In generative models, OSRCI generates images similar to those in the training set but not belonging to any known class and uses them to train open-set classifiers. CROSR combines supervised learning prediction and unsupervised reconstruction of latent representations to reassign probability distribution. C2AE uses class-conditional auto-encoders to obtain decision boundaries from reconstruction errors through EVT. DIAS generates virtual samples of varying difficulty using a GAN model and Copycat. ARPL+CS further introduces adversarial margin constraints based on the RPL and adds an instantiated adversarial enhancement method to generate diverse and confusing training samples based on the adversarial mechanism between reciprocal points and known classes.

4. Comparison with state-of-the-art methods

The experimental results on the ophthalmic OCT dataset are shown in Table 1. Due to the particularity of the generative models ARPL+CS and DIAS, we used the original backbone network Classifier32ABN [8] in these two models for experiments, and the FGM-SPCL network also used Classifier32 as the backbone network for experiments. Compared with DIAS, our model (PCL+FGM) improves Acc by 3.1% and AUROC by 5.68% with only FGM added. It is proved that FGM's method of generating virtual unknown classes is simpler and more effective. Compared with ARPL+CS, FGM-SPCL also combines virtual class generation and clear decision boundary, its Acc is increased by 5.01%, and AUROC is increased by 11.93%, respectively, which shows the superiority of the FGM-SPCL method. For the prototype classifier models PCL, GCPL, RPL, and SLCPL, and the linear classifier-based MLS, we used ResNet-34 as the backbone architecture and pre-trained on the ImageNet dataset. Compared with the baseline model PCL, the effect of the FGM-SPCL model has been significantly improved after the introduction of FGM and SPCL, and two key indicators: AUROC increased by 5.68% and OSCR increased by 7.54%. Compared with MLS and GCPL, models RPL, SLCPL, and FGM-SPCL adopt certain strategies to further improve the semantic difference between known and unknown classes, and the performance of the models has been improved. It is proved that a clear decision boundary can further enhance the ability of OSR ability. Compared with RPL and SLCPL, FGM-SPCL generates virtual unknown samples to participate in model training through FGM. This method

Table 1 Experimental results on the ophthalmic OCT

Method	Acc	AUROC	OSCR
F: Classifier32			
ARPL+CS [8]	88.54	74.94	–
DIAS [6]	88.75	78.70	–
PCL+FGM	91.85	84.38	78.72
FGM-SPCL	**93.55**	**86.87**	**82.11**
F: ResNet34			
PCL	92.70	82.07	75.90
MLS [28]	92.29	81.59	75.37
GCPL [20]	93.13	78.94	72.97
SLCPL [4]	92.71	82.67	77.71
RPL [22]	93.54	83.43	76.25
FGM-SPCL	**93.75**	**87.75**	**83.44**

not only improves the intra-class compactness but also calibrates the decision boundary of known classes. At the same time, SPCL is used to constrain the distribution range of known classes and virtual unknown classes, which helps to maximize the decision boundary between known classes and unknown classes.

The experimental results on the HyperKvasir dataset are shown in Table 2. Experimental results show that the FGM-SPCL model achieves state-of-the-art or comparable results on the HyperKvasir dataset, further showing that FGM-SPCL achieves better and more stable results when dealing with real-world medical data problems.

Table 2 Experimental results on the HyperKvasir

Method	Acc	AUROC	OSCR
PCL	98.73	81.12	79.83
MLS [28]	98.92	82.46	81.10
GCPL [20]	98.66	82.91	78.72
SLCPL [4]	**98.98**	83.53	82.12
RPL [22]	98.66	76.54	75.22
FGM-SPCL	98.92	**88.09**	**87.44**

The experimental results on the CIFAR10, CIFAR+10, and CIFAR+50 are shown in Table 3. The experiments were conducted on the backbone network Classifier32, and the performance of all baseline models was evaluated using the AUROC metric. During the experiment, it was found that PCL($\boldsymbol{x}$) and PPCL($\boldsymbol{x}$) had better performance when the feature vectors and prototype vectors were not normalized. Therefore, we only normalized the feature and prototype vectors when computing the AICL loss. The experimental results on the three datasets demonstrate that FGM-SPCL achieved the best performance compared to state-of-the-art discriminative models and generative models.

Table 3 Experimental results on the CIFAR10, CIFAR+10, and CIFAR+50

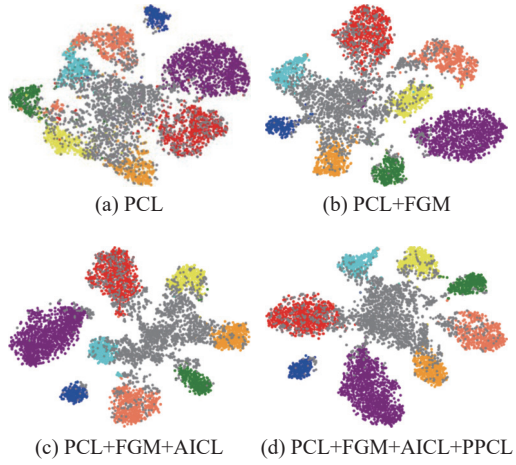
Method	CIFAR10	CIFAR+10	CIFAR+50
PCL	90.83	97.08	95.00
Softmax	67.7	81.6	80.5
G-OpenMax [9]	67.5	82.7	81.9
OSRCI [27]	69.9	83.8	82.7
CROSR [16]	88.3	91.2	90.5
C2AE [17]	89.5	95.5	93.7
RPL [22]	86.1	85.6	85.0
GCPL [20]	82.8	88.1	87.9
SLCPL [4]	86.1	91.6	88.8
ARPL+CS [8]	91.0	97.1	95.1
FGM-SPCL	**92.73**	**97.57**	**96.01**

5. Ablation experiments

The proposed method comprises three elements: FGM, PPCL, and AICL. Here, we perform ablation experiments and further discuss the impact of each element. Table 4 shows the experimental results for the four variants, Figure 3 shows the visualization results.

Table 4 Results of ablation experiments on ophthalmic OCT dataset

Method	Acc	AUROC	OSCR
PCL	92.70	82.07	75.90
PCL+FGM	93.33	85.19	79.96
PCL+FGM+AICL	**94.17**	86.20	82.14
PCL+FGM+PPCL	93.75	86.52	82.11
FGM-SPCL	93.75	**87.75**	**83.44**



(a) PCL (b) PCL+FGM (c) PCL+FGM+AICL (d) PCL+FGM+AICL+PPCL

Figure 3 FGM, PPCL, and AICL effect visualization.

The experimental results show that with the sequential addition of the three elements, the comprehensive performance of the model on the three indicators has been continuously improved, which proves the effectiveness of the three elements. At the same time, combined with the visualization results, it can be found that:

1) From the visualization results shown in Figures 3 (a) and (b), it can be seen that the FGM module helps the model learn a more compact feature representation space, which is more conducive to the improvement of the decision-making performance of closed-set classifiers. The number of samples in the open set distributed in the region of known class features is reduced and is further compressed to a position that tends to the center of the feature space. At the same time, the decision boundaries between known classes and known classes and between known classes and unknown classes are calibrated. Compared with PCL, the recognition accuracy (Acc) of known classes is improved by 0.63%, the OSR performance AUROC is improved by 3.12%, and OSCR is improved by 4.06%.

2) As can be seen from the visualization results shown in Figures 3(b) and (c), after adding the AICL function, the known class features are more compact, and the unknown classes are further pushed away, which enhances the OSR ability. Compared with PCL+FGM, the

Acc of known classes is further improved by 0.84%, the OSR performance AUROC is improved by 1.01%, and the OSCR is improved by 2.18%.

3) As can be seen from the visualization results shown in Figures 3(c) and (d), after adding the PPCL loss function, all known classes are distributed in the peripheral area of the embedding space. The virtual unknown data has a certain isolation effect between the real unknown data and the known data and widens the distance between the known and unknown data. It makes the decision boundary between the known class and the unknown class more concise and clearer. Compared with PCL+FGM+AICL, the OSR performance of AUROC is improved by 1.55%, and the OSCR is improved by 1.30%.

Tables 5 and 6 present the ablation experiment results on the HyperKvasir dataset and CIFAR10 dataset, achieving consistent performance with the OCT dataset. The results indicate that with the sequential introduction of the three components, the model's comprehensive performance on the three indicators continuously improves, further demonstrating the effectiveness and universality of the three components.

Table 5 Results of ablation experiments on HyperKvasir dataset

Method	Acc	AUROC	OSCR
PCL	98.73	81.12	79.83
PCL+FGM	98.86	83.17	82.16
PCL+FGM+AICL	98.92	85.84	85.21
PCL+FGM+PPCL	**98.98**	85.72	84.14
FGM-SPCL	98.92	**88.09**	**87.44**

Table 6 Results of ablation experiments on CIFAR10 dataset

Method	Acc	AUROC	OSCR
PCL	95.39	90.83	87.99
PCL+FGM	95.73	91.73	88.94
PCL+FGM+PPCL	96.32	92.23	89.71
FGM-SPCL	**96.74**	**92.73**	**90.43**

6. Effect of hyperparameters

In order to make a trade-off between OSR performance and closed-set accuracy, we tested the effect of different α and β on the experimental results. α and β, respectively, control the regularity of the location distribution of known classes and virtual classes in space. On the one hand, when α and β are too small, the model cannot gather all category prototypes to the peripheral area of the feature space well. On the other hand, given larger values of α and β, the classification accuracy will be affected due to excessively restricting the spatial location. The results are shown in Table 7.

With the increase of α and β, the performance of AUROC and OSCR is getting better and better. The results show that by moderately increasing the constraint on the prototype position, the semantic difference between the known and unknown classes can be improved, the decision boundary can be widened, and the OSR performance can be improved. But at the same time, the Acc decreased slightly, probably because the control of the position of the prototype points will limit the learning of the semantic features of the known images, resulting in a downward trend in the accuracy rate.

Table 7 Effect of different hyperparameters

α	β	Acc	AUROC	OSCR
0.5	0.5	93.75	**87.75**	**83.44**
0.5	0.2	93.96	87.29	82.60
0.2	0.2	**94.17**	86.76	82.53

V. Conclusions

In this paper, we propose an open-set recognition network for medical images based on fine-grained data mixture and spatial position constraint loss, which improves the robustness of the intelligent auxiliary diagnosis system for fine-grained image recognition tasks. First, the proposed fine-grained data mixture module expands the coverage of virtual unknown data identification difficulty levels and calibrates the decision boundary. Secondly, the proposed spatial position constraint loss maximizes the distance between the known and unknown classes. A large number of experiments and analyses have proved the effectiveness of this method.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 62176026), the Beijing Natural Science Foundation (Grant No. M22009), the Key Research and Development Program of Hebei (Grant No. 22377775D), and the Engineering Research Center of Information Networks, Ministry of Education.

References

[1] L. Schiff, B. Migliori, Y. Chen, *et al.*, "Integrating deep learning and unbiased automated high-content screening to identify complex disease signatures in human fibroblasts," *Nature Communications*, vol. 13, no. 1, article no. 1590, 2022.

[2] G. Y. Wang, X. H. Liu, J. Shen, *et al.*, "A deep-learning pipeline for the diagnosis and discrimination of viral, non-viral and COVID-19 pneumonia from chest X-ray images," *Nature Biomedical Engineering*, vol. 5, no. 6, pp. 509–521, 2021.

[3] K. Zhang, X. H. Liu, J. Shen, *et al.*, "Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography," *Cell*, vol. 181, no. 6, pp. 1423–1433.e11, 2020.

[4] Z. H. Xia, P. H. Wang, G. G. Dong, *et al.*, "Spatial location constraint prototype loss for open set recognition," *Computer Vision and Image Understanding*, vol. 229, article no. 103651, 2023.

[5] S. Kong and D. Ramanan, "OpenGAN: Open-set recognition via open data generation," in *Proceedings of 2021 IEEE/CVF International Conference on Computer Vision*, Montreal, QC, Canada, pp. 793–802, 2021.

[6] W. Moon, J. Park, H. S. Seong, *et al.*, "Difficulty-aware simulator for open set recognition," in *Proceedings of the 17th*

European Conference on Computer Vision, Tel Aviv, Israel, pp. 365–381, 2022.

[7] D. W. Zhou, H. J. Ye, and D. C. Zhan, "Learning placeholders for open-set recognition," in *Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, TN, USA, pp. 4399–4408, 2021.

[8] G. Y. Chen, P. X. Peng, X. Q. Wang, *et al.*, "Adversarial reciprocal points learning for open set recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 8065–8081, 2021.

[9] Z. Y. Ge, S. Demyanov, Z. Chen, *et al.*, "Generative OpenMax for multi-class open set classification," in *Proceedings of British Machine Vision Conference 2017*, London, UK, 2017.

[10] W. J. Scheirer, A. de Rezende Rocha, A. Sapkota, *et al.*, "Toward open set recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1757–1772, 2013.

[11] A. Bendale and T. Boult, "Towards open world recognition," in *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, pp. 1893–1902, 2015.

[12] P. R. M. Júnior, R. M. De Souza, R. de O. Werneck, *et al.*, "Nearest neighbors distance ratio open-set classifier," *Machine Learning*, vol. 106, no. 3, pp. 359–386, 2017.

[13] H. Zhang and V. M. Patel, "Sparse representation-based open set recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 8, pp. 1690–1696, 2017.

[14] X. Sun, Z. N. Yang, C. Zhang, *et al.*, "Conditional Gaussian distribution learning for open set recognition," in *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, pp. 13477–13486, 2020.

[15] H. J. Zhang, A. Li, J. Guo, *et al.*, "Hybrid models for open set recognition," in *Proceedings of the 16th European Conference on Computer Vision*, Glasgow, UK, pp. 102–117, 2020.

[16] R. Yoshihashi, W. Shao, R. Kawakami, *et al.*, "Classification-reconstruction learning for open-set recognition," in *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, pp. 4011–4020, 2019.

[17] P. Oza and V. M. Patel, "C2AE: Class conditioned auto-encoder for open-set recognition," in *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, pp. 2302–2311, 2019.

[18] H. M. Yang, X. Y. Zhang, F. Yin, *et al.*, "Convolutional prototype network for open set recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 5, pp. 2358–2370, 2022.

[19] Y. Shu, Y. M. Shi, Y. W. Wang, *et al.*, "P-ODN: Prototype-based open deep network for open set recognition," *Scientific Reports*, vol. 10, no. 1, article no. 7146, 2020.

[20] C. G. Lyu, Y. X. Chen, Z. J. Chen, *et al.*, "Automatic epilepsy detection based on generalized convolutional prototype learning," *Measurement*, vol. 184, article no. 109954, 2021.

[21] J. Lu, Y. L. Xu, H. Li, *et al.*, "PMAL: Open set recognition via robust prototype mining," in *Proceedings of the 36th AAAI Conference on Artificial Intelligence*, Virtual Event, pp. 1872–1880, 2022.

[22] G. Y. Chen, L. M. Qiao, Y. M. Shi, *et al.*, "Learning open set network with discriminative reciprocal points," in *Proceedings of the 16th European Conference on Computer Vision*, Glasgow, UK, pp. 507–522, 2020.

[23] D. Mehta, Y. Gal, A. Bowling, *et al.*, "Out-of-distribution detection for long-tailed and fine-grained skin lesion images," in *Proceedings of the 25th International Conference on Medical Image Computing and Computer Assisted Intervention*, Singapore, Singapore, pp. 732–742, 2022.

[24] K. M. He, X. Y. Zhang, S. Q. Ren, *et al.*, "Deep residual learning for image recognition," in *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, pp. 770–778, 2016.

[25] H. Borgli, V. Thambawita, P. H. Smedsrud, *et al.*, "HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy," *Scientific Data*, vol. 7, no. 1, article no. 283, 2020.

[26] A. Krizhevsky, "Learning multiple layers of features from tiny images," *Technical report*, University of Toronto, Available at: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf, 2009.

[27] L. Neal, M. Olson, X. Fern, *et al.*, "Open set learning with counterfactual images," in *Proceedings of the 15th European Conference on Computer Vision*, Munich, Germany, pp. 620–635, 2018.

[28] S. Vaze, K. Han, A. Vedaldi, *et al.*, "Open-set recognition: A good closed-set classifier is all you need," in *Proceedings of the 10th International Conference on Learning Representations*, Virtual Event, 2022.

Ruru ZHANG was born in 1992. She is currently a Ph.D. candidate at Beijing University of Posts and Telecommunications (BUPT), Beijing, China. Her research interests include artificial intelligence and medical image processing.
(Email: zrr@bupt.edu.cn)

Haihong E was born in 1982. She received the Ph.D. degree in computer science and technology from Beijing University of Posts and Telecommunications, Beijing, China. She is currently a Professor at Beijing University of Posts and Telecommunications. She has been engaged in research and teaching in the fields of big data, artificial intelligence, and cloud native services for a long time.
(Email: ehaihong@bupt.edu.cn)

Lifei YUAN was born in 1986. He received the M.S. degree in ophthalmology from Tianjin Medical University, Tianjin, China. His main research interest is fundus disease.
(Email: yuanlifei2341@163.com)

Yanhui WANG was born in 1981. She received the M.S. degree in ophthalmology from Tianjin Medical University, Tianjin, China. Her main research interest is fundus disease.
(Email: yanhuiwang1981@163.com)

Lifei WANG was born in 1976. She received the Ph.D. degree in ophthalmology from Sun Yat-Sen University, Guangzhou, China. She is currently Secretary of the Party Committee of Hebei Provincial Eye Hospital. Her main research interest is fundus disease.
(Email: wlfhb@126.com)

Meina SONG was born in 1974. She received the Ph.D. degree from BUPT, Beijing, China. She is currently a Professor at BUPT, and Director of the Engineering Research Center of Information Networks, Ministry of Education, Beijing, China. Her main research interests include artificial intelligence and its application in the fields of finance and medicine.
(Email: mnsong@bupt.edu.cn)