

Received 18 October 2022, accepted 31 October 2022, date of publication 9 November 2022, date of current version 16 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3220776

## RESEARCH ARTICLE

# Graph Adversarial Transfer Learning for Diabetic Retinopathy Classification

JINGBO HU<sup>1,2</sup>, HUAN WANG<sup>3</sup>, LE WANG<sup>4</sup>, AND YE LU<sup>3</sup>

<sup>1</sup>School of Electronic and Electrical Engineering, Baoji University of Arts and Sciences, Baoji 721016, China

<sup>2</sup>College of Computer Information and Engineering, Nanchang Institute of Technology, Nanchang 330044, China

<sup>3</sup>School of Computer Science and Technology, Baoji University of Arts and Sciences, Baoji 721016, China

<sup>4</sup>School of Chemistry and Chemical Engineering, Baoji University of Arts and Sciences, Baoji, Shaanxi 721013, China

Corresponding author: Huan Wang (huanw0813@126.com)

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 82104682 and Grant 61971005, in part by the Shaanxi Science and Technology Commission under Grant 2022SF-208, in part by the Innovation Capability Support Program of Shaanxi under Grant 2022TD-63, and in part by the Innovation Research and Development Team of BUAS on Qinling Medicinal Plants and Functional Organic Molecules.

**ABSTRACT** Diabetic retinopathy (DR) is an essential factor that has caused vision loss and even blindness in middle-aged and older adults. A system that can automatically perform DR diagnosis can help ophthalmologists save a lot of tedious work, such as DR grading or lesion detection. At the same time, patients can find their diseases earlier and perform the correct treatment. However, most of the existing methods require many DR annotations to train the model, and the DR data will vary to different degrees due to various shooting tools. The above problems lead to the inefficient use of existing data in the experiment, limiting actual deployment. To alleviate this problem, we propose a novel Graph Adversarial Transfer Learning (GATL) for DR diagnosis in a deep model through transfer learning, including intra-domain alignment and inter-domain alignment. The proposed GATL enjoys several merits. First, our GATL adopts the self-supervised training to save the annotating cost in the target domain thus this domain adaptation method can significantly reduce annotation cost compared to the supervised approaches. Second, we introduce the graph neural network to extract potential features between unknown samples. Third, to enhance the robustness of the model, we use adversarial training to perform both inter-domain and intra-domain alignment to further improve the model's classification accuracy. GATL achieved 94.3%, 97.5%, and 91.1% in accuracy, sensitivity, and specificity in the APTOS dataset and 92.7%, 95.7%, and 89.7% in the EyePACS dataset, respectively. Extensive experimental results on two challenging benchmarks, including APTOS 2019 and EyePACS, demonstrate that the proposed GATL performs favorably against baseline DR classification methods.

**INDEX TERMS** Diabetic retinopathy classification, graph adversarial network, transfer learning, intra-domain, inter-domain.

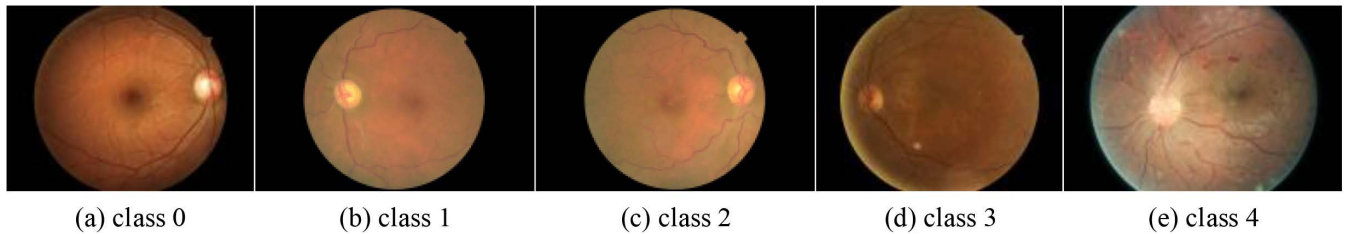
## I. INTRODUCTION

Diabetic retinopathy (DR) is one of the eye diseases with a rapidly increasing incidence in recent years. It is a complication of diabetes, which can cause damage to the blood vessels in the eye and eventually lead to vision loss, even blindness [1], [2]. In modern social life, high blood pressure and high blood sugar have become factors that plague the health of most middle-aged and older people. This further

increases the incidence of DR. Therefore, the early detection of diabetic retinopathy is essential, significantly improving the effectiveness of treatment. However, in clinical testing, the early diagnosis of DR is relatively tricky. Even professional ophthalmologists need to spend a lot of time and energy to compare color images. Therefore, developing a system for automatic detection or classification of DR is very meaningful and necessary.

An automated diagnostic system will review retinal images and perform DR classification based on the learned experience. In some public retinal datasets, DR is usually classified

The associate editor coordinating the review of this manuscript and approving it for publication was Inês Domingues<sup>id</sup>.



**FIGURE 1.** Illustration of different DR grades.

by experts into the following five categories [3]: R0-no DR, R1-mild, R2-moderate, R3-severe, and R4-proliferative. In detail, the type of DR is affected by several factors, such as the number of lesions, area size, and appearance in fundus images. At the same time, the above five categories can also be divided into two different categories according to the standard [3], [4], [5]. For instance, Figure 1 provides an illustration of five DR grades in the Kaggle DR dataset [6]. Notably, the features of DR lesions in fundus images are complex and diverse. Therefore, the diagnostic system needs to extract critical features with discriminative from the sample images to perform DR screening accurately. In recent years, deep learning has achieved excellent results in many fields, and there are also many studies that apply deep learning to DR classification [7], [8], [9]. Dutta et al. [7] introduce an automated knowledge model to identify the critical antecedents of DR, which identify the target class thresholds weighted Fuzzy C-means algorithm. Kassani et al. [8] propose a new feature extraction method using a modified Xception architecture for the diagnosis of DR disease, which is based on deep layer aggregation that combines multilevel features from different convolutional layers of Xception architecture. Pratt et al. [9] develop a network with CNN architecture and data augmentation, which can identify the intricate features involved in the classification task such as micro-aneurysms, exudate, and hemorrhages on the retina and consequently provide a diagnosis automatically and without user input. There are some researches based on pixel-level supervision [10], or patch-level supervision [5], [11], [12] also have been proposed. Since the annotation of fundus images requires manual annotation by experienced domain experts, the advanced methods [13], [14], [15] will be limited in flexibility and scalability in actual deployment.

Although the above studies all demonstrate the effectiveness of their methods on a single dataset, the problem of domain adaptation is ignored in practical application scenarios. For example, in clinical experiments, imaging equipment for acquiring various image data may be provided by multiple manufacturers. These devices have specific differences in software and hardware, resulting in different levels of image quality gaps, and annotating sufficient data in new scenarios requires cost-expensive professional labor. Therefore, when a model trained in a fixed source domain is applied to another unknown domain, its diagnostic accuracy usually

drops significantly. One way to address this problem is to fine-tune the model trained on the source domain with enough labeled samples from the target domain to align the inter-domain differences. However, the annotation data of medical images is minimal, manual annotation requires a lot of time and energy from experts, and the economic cost is high. Therefore, it is necessary to make effective use of label data. As a result, the method of semi-supervised field adaptation is of great significance for DR classification.

Traditional domain adaptation methods usually use Convolutional Neural Networks (CNN) as the backbone to extract the internal features of the samples for classification. This ignores the potential connections between retinal images. In other words, GNN can ignore the invalid features in the CNN-extracted features and achieve the effect of noise reduction. There is already a large body of literature available on this application [16], [17], [18], [19], [20], [21], [22]. Therefore, the Graph Neural Network (GNN) can extract the potential features between images to use more practical information for DR detection. Nevertheless, the problem of model instability is often encountered during the experiment. Through experience, the model under adversarial training is often very robust, and adding adversarial loss is an excellent way to enhance the stability of the model. In addition, if the model is simply constrained by the method of intra-domain alignment, the model trained with a large amount of source domain data may not be suitable for the target domain. Therefore, adding a method of inter-domain alignment can further improve the model effect. Thus, we propose a novel Graph Adversarial Transfer Learning (GATL) method for diabetic retinopathy classification in this paper. GATL not only focuses on the pixel characteristics of the retinal image itself, but it also extracts the potential relationships between samples, making the most of the original data. At the same time, GATL uses the high-confidence prediction value of the target domain as a pseudo-label and uses the confrontation training of two classifiers to promote further the intra-domain alignment of the source domain and the target domain. Finally, the newly designed discriminator will narrow the relationship between domains and align the data distribution of the source domain and target domain.

The main contributions of our method are as follows:

(1) We provide a brand-new graph-based domain-adaptive approach to investigate additional fundus image

characteristics, and we employ transfer learning to successfully make use of additional data for DR classification of unidentified data.

(2) We train the model by designing two classifiers for adversarial purposes to improve the robustness of the model and also to improve the classification performance of the model.

(3) We design a discriminator and obfuscate the discriminator's judgment of the source and target domain data, allowing the model to be further compared between domains on top of the intra-domain comparison.

## II. RELATED WORK

This section first introduces some recent research results on the classification of diabetic retinopathy and then summarizes the successful applications of some domain adaptation methods in medical images.

### A. RETINAL IMAGE CLASSIFICATION

With the continuous development of computer vision, some computer-assisted retinal image research has also achieved great success [23], [24], [25], [26], [27], [28], [29], [30], [31]. For example, Chen et al. [23] propose a general deep learning model for DR classification, which uses a 2-stage training method to solve the overfitting problem. By the way, they also provide a simple method of addressing the imbalance of DR databases. Erciyas et al. [24] develop a deep learning-based method in which diabetic retinopathy lesions are detected automatically and independently of datasets, and the detected lesions are classified. Vives et al. [25] present a bio-inspired approach on synaptic metaplasticity in convolutional neural networks to detect diabetic retinopathy. Jadhav et al. [26] design an optimal feature selection-based diabetic retinopathy detection method which can develop automated DR detection by analyzing the retinal abnormalities like hard exudates, hemorrhages, Microaneurysm, and soft exudates. Canayaz et al. [27] propose an approach based on feature selection with wrapper methods used for fundus images. Abdelmaksoud et al. [28] introduce E-DenseNet, a hybrid deep learning method. Based on transfer learning, they combine the EyeNet and DenseNet models. Zhang et al. [29] design a Source-Free Transfer Learning technique for detecting referable DR that uses unannotated retinal pictures and only uses the source model during the training process. Gangwar et al. [30] tackle the challenge of automated diabetic retinopathy diagnosis and suggest a unique deep learning hybrid solution. Yi et al. [31] suggest the network known as RA-EfficientNet, in which a residual attention (RA) block is added to EfficientNet in order to extract additional features and address the issue of minute changes between lesions.

However, none of the above methods considers domain adaptation issues. The fundus data taken by different equipment are different to a certain extent. Moreover, in practice, annotated data is challenging to obtain due to the high cost,

requiring limited available data to design domain adaptation methods.

### B. DOMAIN ADAPTATION METHODS IN MEDICAL IMAGE CLASSIFICATION

In deep learning research, data labels are often the most expensive and difficult to obtain. This problem is particularly manifested in medical images. Therefore, how to efficiently use existing tags to predict unknown data has become a hot issue. Domain Adaptation [32] is one of the popular research directions. In recent years, many successful domain adaptation methods [33], [34], [35], [36], [37] have been widely used in medical images. Wang et al. [33] propose a method called deep adversarial domain adaptation to improve the performance of breast cancer screening using mammography. They aim to extract the knowledge from a public dataset and transfer the learned knowledge to improve the detection performance on the target dataset. Castellanos et al. [34] design a method that combines neural networks and domain adaptation in order to carry out unsupervised document binarization. Abbet et al. [35] propose a method for colorectal cancer tissue phenotyping to Adapt, which takes advantage of self-supervised learning to perform domain adaptation and remove the necessity of a fully-labeled source dataset. Konyakhin et al. [36] present their solution to the Traffic4Cast 2021 Core Challenge, which employs multiple domain adaptation techniques to fight the domain shift. Hong et al. [37] report an unsupervised domain adaptation framework for cross-modality liver segmentation via joint adversarial learning and self-learning.

Inspired by the above migration methods, to effectively use open source data tags and reduce the loss caused by domain transformation, we propose a GATL network for DR classification.

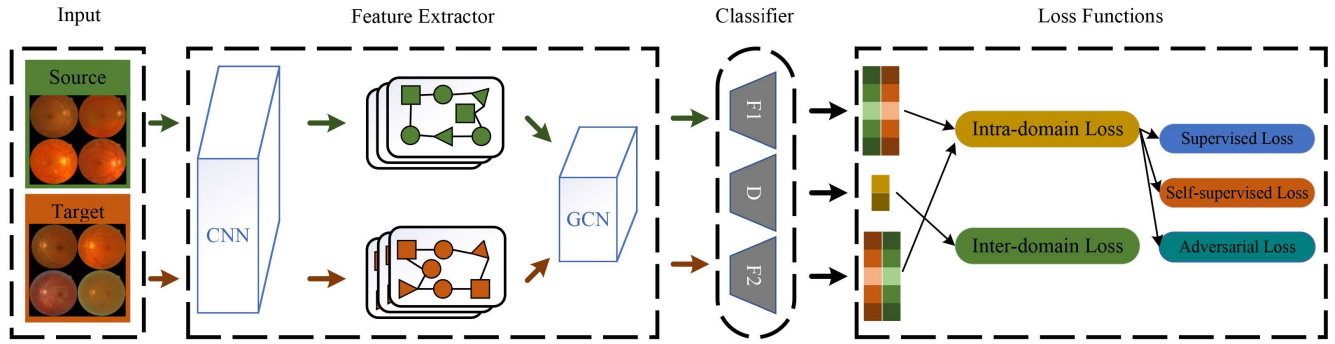
## III. METHOD

This paper introduces a graph adversarial transfer learning method for DR classification. In this section, we divide GATL into two stages. Finally, we report the whole training schema of our GATL.

### A. THE FIRST STAGE OF GATL

We define the input data as  $X = X_S \cup X_T$  with  $N = N_S + N_T$  samples, where  $X_S = \{x_1^s, \dots, x_i^s, \dots, x_{N_S}^s\}$  are source retinal images with the corresponding labels  $Y_S = \{y_1^s, \dots, y_i^s, \dots, y_{N_S}^s\}$ , and  $X_T = \{x_1^t, \dots, x_i^t, \dots, x_{N_T}^t\}$  denotes the unlabeled data from target domain.

The purpose of domain adaptation is to train a classifier  $F$ , which can accurately predict the sample category under the premise of a large amount of source domain data support and can be well applied to target domain data. In our proposed GATL, we use resnet50 as the backbone of the CNN feature extractor  $G_c$  to extract the intrinsic features of the fundus images. At the same time, the GNN feature extractor  $G_g$  will



**FIGURE 2.** GATL framework description: First, fundus images are fed to a feature extractor that combines CNN and GNN architecture. Then, the extracted key features are respectively transmitted to classifiers  $F_1$  and  $F_2$  for adversarial training, and discriminator  $D$  is trained for confusing the source and target domains. Finally, the entire model is trained simultaneously from two perspectives of intra-domain and inter-domain alignment.

explore the potential connections between samples. The bi-classifier  $F_1$  and  $F_2$  will output the prediction vector of each retinal image.

We first feed the source domain data to the CNN feature extractor  $G_c$ , and update the CNN network in the following supervised loss,

$$\min_{\theta_c, \theta_{f1}, \theta_{f2}} \mathcal{L}_{ces}(X_S, Y_S) = \frac{1}{2N_S} \sum_{i=1}^{N_S} \sum_{n=1}^2 \mathcal{L}_{ce}(F_n(G_c(x_i^s)), y_i^s) \quad (1)$$

where  $\mathcal{L}_{ce}$  denotes the standard cross-entropy loss function.  $\theta_c, \theta_{f1}$ , and  $\theta_{f2}$  represent the parameters of  $G_c, F1$  and  $F2$  respectively.

In a short time, we can obtain a feature extractor  $G_c$  with a certain feature representation ability. Then we join the GNN extractor  $G_g$  after  $G_c$ . With the obtained representation  $H_s = \{h_1^s, \dots, h_i^s, \dots, h_{N_S}^s\}$  from  $G_c$ , we use the k-nearest neighbor method to build the graph, the formula is as follows,

$$d_{ij} = \sqrt{\|h_i^s - h_j^s\|^2} \quad (2)$$

where  $d_{ij}$  denotes the similarity between the i-th feature  $h_i^s$  and j-th feature  $h_j^s$  from  $H_s$ . By sorting  $d_{ij}$ , we can obtain the top  $k$  features with high similarity. At the same time, the sort index is  $Idx = \{i_1, \dots, i_m, \dots\}$ , in which  $i_m$  represents the m-th closest neighbor feature to i-th one.

Thus, the adjacency matrix  $A$  can be obtained according to the following rules.

$$A_{ij}^k = \begin{cases} 1, & i_m \leq k \\ 0, & i_m > k \end{cases} \quad (3)$$

In this way, we can use  $G_g$  to extract special features in the graph structure. Assist by the labels, we update backbone and replace Eq(1) in the following way,

$$\min_{\theta_c, \theta_g, \theta_{f1}, \theta_{f2}} \mathcal{L}_{ces} = \frac{1}{2N_S} \sum_{i=1}^{N_S} \sum_{n=1}^2 \mathcal{L}_{ce}(F_n(G_g(G_c(x_i^s))), y_i^s) \quad (4)$$

where  $\theta_g$  denotes the parameters of  $G_g$ . The features extracted through the cooperation of  $G_c$  and  $G_g$  will be sent to  $F1$  and  $F2$ . We designed a self-supervised loss  $\mathcal{L}_{ss}$  to enhance the robustness of the model. The formula is as follows,

$$\mathcal{L}_{ss} = |P_1 - P_2|_1 \quad (5)$$

where  $P_1$  and  $P_2$  represent the output predict vectors from  $F_1$  and  $F_2$  for the same source input. In addition,  $|\cdot|_1$  represents the  $l_1$  distance.

### B. THE SECOND STAGE OF GATL

After training the model with a large amount of labeled source domain data, we feed the target domain data to the network and perform high-confidence label filtering on the softmax output of the classifier. The filtered annotations are used as pseudo labels  $Y_T = \{y_1^t, \dots, y_i^t, \dots\}$  for the fundus image of the target domain.

In this way, we can use the target domain data in a self-supervised manner. At the same time, self-supervised training can reduce the fuzzy samples of the decision boundary. The loss function is as follows,

$$\min_{\theta_c, \theta_g, \theta_{f1}, \theta_{f2}} \mathcal{L}_{ces} = \frac{1}{2n_t} \sum_{i=1}^{n_t} \sum_{n=1}^2 \mathcal{L}_{ce}(F_n(G_g(G_c(x_i^t))), y_i^t) \quad (6)$$

where  $n_t$  is the number of selected high-confidence pseudo-label.

The difference from the first stage is that in order to perform inter-domain alignment, we design a discriminator  $D$  to identify whether the image belongs to the source domain or the target domain. The purpose is to confuse the discriminator's recognition of inter-domain samples so that the feature extractor can extract cross-domain features. The inter-domain loss is as follows,

$$\min_{\theta_d, \theta_g, \theta_c} \mathcal{L}_{dis} = |D(G_g(G_c(x_i^s))) - D(G_g(G_c(x_i^t)))|_1 \quad (7)$$

where  $\theta_d$  denotes the parameter of  $D$ . By reducing the sample difference between domains GATL will narrow the distance

between the source domain and the target domain, and further improve the accuracy of the model applied to the target domain.

### C. TRAINING STRATEGY

GATL aims to perform DR classification by aligning the category distribution within and between domains in this work. We divide the GATL training process into the following three steps:

#### Step 1. Training the model by the labeled source data

Use the large amount of source domain data to learn  $G_c$ ,  $G_g$ ,  $F_1$ , and  $F_2$  together to reduce the empirical risk of the source distribution. The model runs in the following way,

$$\min \mathcal{L} = \lambda \mathcal{L}_{ces} - (1 - \lambda) \mathcal{L}_{ss} \quad (8)$$

where  $\lambda$  denotes the balance factor.

#### Step 2. Self-supervised learning for the unlabeled target domain

Join the self-supervised training of the target domain and inter-domain alignment. Freeze the parameters of feature extractors  $G_c$  and  $G_g$ , and then update classifiers  $F_1$  and  $F_2$  to maximize the difference in probability output between them while maintaining classification accuracy. The model runs under the adversarial loss,

$$\min \mathcal{L} = \lambda_1 \mathcal{L}_{ces} - \lambda_2 \mathcal{L}_{ss} - \lambda_3 \mathcal{L}_{dis} \quad (9)$$

where  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  denote the balance factors.

#### Step 3. Further optimization on feature extractors

Freeze the parameters of the classifiers  $F_1$  and  $F_2$ , and update the feature extractors  $G_c$  and  $G_g$  to minimize the difference between the classifiers while minimizing the difference in sample characteristics between domains. The model runs in the following way,

$$\min \mathcal{L} = \lambda \mathcal{L}_{ss} + (1 - \lambda) \mathcal{L}_{dis} \quad (10)$$

After repeating the above steps several times, GATL can effectively classify DR and perform inter-domain alignment based on intra-domain alignment. The model optimization is summarized in Algorithm 1.

## IV. EXPERIMENTS

### A. IMPLEMENTATION DETAILS

The entire experimental process is completed using GeForce 2080ti GPU under the PyTorch framework. According to statistics, the model converges around 40 epochs, at which point the curve of the loss function also reaches a smooth state. It took about 6 hours from start to model convergence. Each retinal image is adjusted to 512\*512 pixels and augmented by random rotation, cropping, and contrast-color augmentation before input to the network. In the GATL training process, we choose Adam as the optimizer, the batch size is 32, and the epoch is set to 50. For the learning rate setting, we chose the strategy of dynamically adjusting the learning rate, which gradually decreases according to the epoch change. The initial learning rate is set to 0.1. In detail, we use

### Algorithm 1 Graph Adversarial Transfer Learning (GATL)

**Input data:** Source data  $X_S$  with corresponding labels  $Y_S$ , target data  $X_T$ , every images are reshaped into  $512 \times 512$  and augmented, batch size =32, learning rate=0.0001, number of epochs  $T$ , the parameters  $\lambda$  and  $k$ .

Initialize network parameters for  $G_c$ ,  $G_g$ ,  $F_1$ ,  $F_2$  and  $D$ ;

**for**  $t(\text{epoch}) = 0$  to  $T_1$  **do** (step 1)

    Extract the feature vectors from  $G_c$  and  $G_g$  for source data  
    Obtain the predicted probabilities from  $F_1$  and  $F_2$  for source features

    Optimize the parameters in  $G_c$ ,  $G_g$ ,  $F_1$  and  $F_2$  by minimizing Eq.8

**end for**

**for**  $t(\text{epoch}) = T_1$  to  $T_2$  **do** (step 2)

    Extract features from  $G_c$  and  $G_g$  for target data

    Obtain the predicted probabilities from  $F_1$  and  $F_2$  for target features

    Freeze  $G_c$  and  $G_g$

    Optimize the parameters in  $F_1$  and  $F_2$  by minimizing Eq.9

**end for**

**for**  $t(\text{epoch}) = T_2$  to  $T_3$  **do** (step 3)

    Freeze  $F_1$  and  $F_2$

    Optimize the parameters in  $G_c$  and  $G_g$  by minimizing Eq. 10

**end for**

**Return:** The trained network parameters

resnet-50 as the backbone, and the classifier is composed of linear layers. For parameter settings, the balance factor  $\lambda$  and KNN parameter  $k$  in graph building are 0.4 and 4. As for the data division, the source data are all fed into the network training, and the target domain is split into training and testing sets with an 80:20 ratio. The available codes and trained models will be available at <https://github.com/huanw0813/GATL> once the paper is published.

### B. DATABASE DESCRIPTION

The proposed model is evaluated on publicly available EyePACS [6] and APTOS 2019 [38] datasets.

**EyePACS** compiled 88,702 color fundus images from the patients' left and right eyes. Multiple devices took these images under different imaging conditions. To uniformly process these fundus images with specific differences, we resize them to 512\*512 pixels in the preprocessing stage. After being annotated by experts, these images were classified into five categories: 0 - No DR, 1 - Mild, 2 - Moderate, 3 - Severe, and 4 - Proliferative DR. In addition, the proportion of various types of images in this dataset is unbalanced, and the specific distribution is shown in Table 1.

**APTOS 2019** is an open-source retinal dataset collated by the Asia Pacific Tele-Ophthalmology Society, in which fundus images were collected under different imaging conditions. There are 3662 images in this dataset, all uniformly pre-processed and rescaled to 512\*512 pixels. Clinicians labeled

**TABLE 1.** The class distribution of datasets.

Label	EyePACS	APTOS 2019
DR 0	65343	1,805
DR 1	6205	370
DR 2	13153	999
DR 3	2087	193
DR 4	1914	295

these images into the following five categories: 0 - No DR, 1 - Mild, 2 - Moderate, 3 - Severe, and 4 - Proliferative DR. The category distribution of the dataset is organized into Table 1.

We divide the samples in these two data sets into binary categories according to the normal/abnormal criteria.

### C. EVALUATION METRICS

To evaluate the performance of the proposed method, we employ accuracy (ACC), precision (PRE), sensitivity (SEN), and specificity (SPE) for binary normal/abnormal DR grading tasks. In this section, we also use the ROC curve and confusion matrix to evaluate the performance of GATL. The evaluation formulas are as follows,

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (12)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (13)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (14)$$

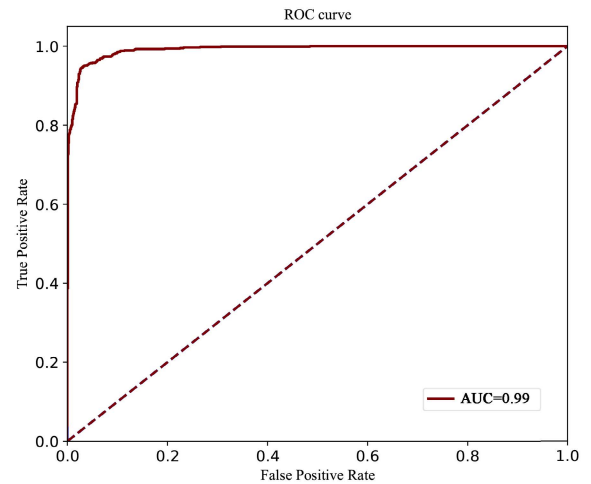
where  $TP$  means that the predicted value and the true value are sick;  $FP$  means that the predicted value is sick and the true value is healthy;  $TN$  indicates that both the predicted value and the true value are healthy;  $FN$  means that the predicted value is healthy, and the true value is sick.

We also included AUC in the evaluation. AUC is a performance metric to measure the quality of a model, and it is defined as the area under the ROC curve. Usually, researchers can utilize the AUC area to measure the effectiveness of the binary classification model, representing the probability that the predicted positive example is ranked ahead of the negative example. The more significant the AUC value, the better the classification effect of the classifier.

### D. EVALUATION ON DR CLASSIFICATION

#### 1) COMPARE WITH BASELINE METHODS ON APTOS DATASETS

In order to further analyze the performance of GATL on DR classification, we have selected some baseline methods for comparison. The comparison results are shown in Tables 2. From Table 2, it can be observed that the proposed GATL achieves the highest accuracy of 94.3% among all the methods, which is 1.9% higher than the supervised methods SE-ResNeXt50 and Vgg-16, 3.1% higher than SFTL,

**FIGURE 3.** ROC curves of the proposed GATL for DR classification on APTOS 2019 dataset, where FPR is False Positive Rate.

2.8% higher than AmResNet50, 3.25% higher than PB-CNN, 2.3% higher than ShallowNet+PI, and 3.6% higher than EfficientNet. This shows that GATL has a great advantage in DR classification accuracy. GATL also achieves the high sensitivity of 97.5% among all the baseline methods, which is 10.4% higher than SE-ResNeXt50, 16.8% higher than EfficientNet, 2.4% higher than SFTL, 8.5% higher than AmResNet50, and 4.9% higher than Vgg-16. This indicates that the misdiagnosis rate of GATL is low, and it can play an excellent effect in clinical applications.

#### 2) COMPARE WITH BASELINE METHODS ON EyePACS DATASETS

We exchanged the source and target domains of the proposed GATL, and chose the corresponding baseline method for comparison. The comparison results are shown in Tables 3. From Table 2, it can be observed that GATL achieves the highest accuracy of 92.7% among all the methods on EyePACS datasets, which is 1.8% higher than Inception V3, 5.36% higher than Bi-channel CNN, 7.2% higher than ViT, and 9% higher than Vgg-16. Also, in terms of sensitivity, GATL achieves the highest value of 95.7%, which is 41.2% higher than VGG16, 1.2% higher than Custom CNN, 18.77% higher than Bi-channel CNN, 2% higher than ViT, and 15.7% higher than Modified VGGNet.

Summarizing the results of the above two tables, we can summarize the following conclusions, (1) GATL still has a high accuracy rate after swapping the source and target domains, which means the proposed model has high robustness due to the adversarial loss. (2) Our domain adaptation method performs better than some supervised methods. This shows that after intra-domain and inter-domain alignment, the classification effect of GATL can be compared with some supervision methods.

#### 3) PERFORMANCE OF ROC VISUALIZATION

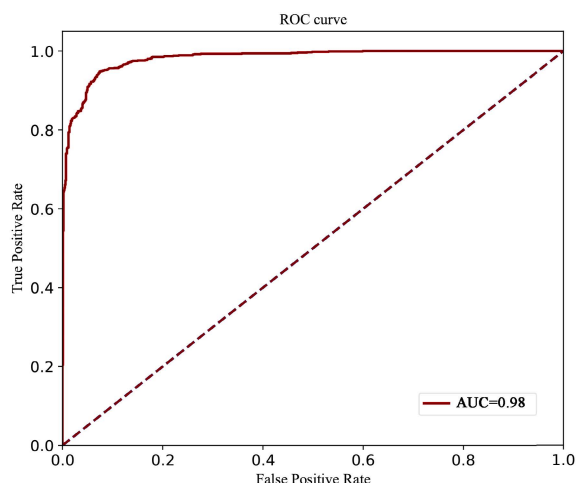
In order to evaluate the performance of the proposed model, we draw ROC curves for two public data sets as shown in

**TABLE 2.** Compared with baseline methods for normal/abnormal DR classification on APTOS 2019 datasets, where FPR is False Positive Rate.

Methods	APTOS 2019				
	Accuracy	Sensitivity	Specificity	Precision	FPR
SE-ResNeXt50 [39]	92.4	87.1	98.2	-	-
EfficientNet [40]	90.7	80.7	97.7	-	-
Vgg-16 [41]	92.4	92.6	92.0	-	-
SFTL [29]	91.2	95.1	85.8	-	-
EfficientNet-B3 [42]	-	98	98	-	-
AmResNet50 [25]	91.5	89	-	-	-
PB-CNN [43]	91.05	-	-	-	-
ShallowNet+PI [44]	92	-	-	-	-
Removing $\mathcal{L}_{dis}$	91.8	95.3	88.6	92.4	4.2
GATL	94.3	97.5	91.1	94.5	2.7

**TABLE 3.** Compared with baseline methods for normal/abnormal DR classification on EyePACS dataset, where FPR is False Positive Rate.

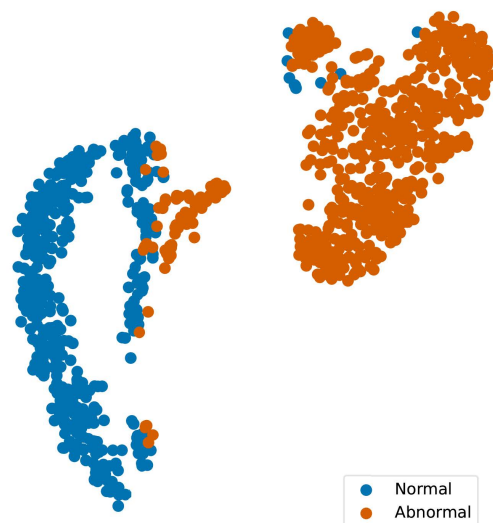
Methods	EyePACS				
	Accuracy	Sensitivity	Specificity	Precision	FPR
Inception V3 [45]	90.9	-	-	-	-
VGG16 [46]	83.7	54.5	93.7	-	-
Custom CNN [47]	-	94.5	90.2	-	-
Modified VGGNet [48]	-	80	92	-	-
Bi-channel CNN [49]	87.37	76.93	93.57	-	-
ViT [50]	85.5	93.7	-	-	-
Removing $\mathcal{L}_{dis}$	89.8	92.5	88.2	90.2	7.5
GATL	92.7	95.7	89.7	92.9	4.6



**FIGURE 4.** ROC curves of the proposed GATL for DR classification on EyePACS dataset.

Figure 3 and 4. ROC curve is an analysis tool of the coordinate scheme in signal detection theory. ROC analysis can not be affected by cost/benefit when making decisions and giving objective and neutral suggestions.

The areas above and below the diagonal in the ROC plot are opposing areas. The diagonal lines represent the results of random classification. The closer the curve is to the upper left corner in the comparative experiment, the better the classifier’s performance. It can be seen from Figure 3 that GATL is very effective in the task of DR classification. The value of AUC reaches 0.99 for the DR binary classification task. Also, in Figure 4, the value of AUC reached 0.98 in the



**FIGURE 5.** TSNE of the proposed GATL for DR classification on APTOS 2019 dataset.

DR classification. This shows that GATL has improved with training in the migration of the two data sets, which proves the effectiveness of GATL.

4) PERFORMANCE OF TSNE VISUALIZATION

We also made TSNE diagrams for two datasets simultaneously to show the effect of GATL more intuitively.

TSNE is a data visualization tool for data dimensionality reduction. It can reduce and visualize high-dimensional data

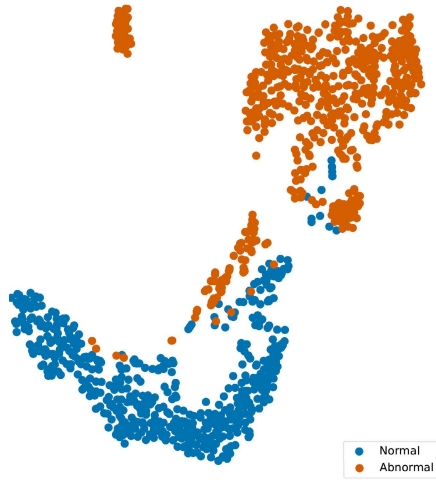


FIGURE 6. TSNE of the proposed GATL for DR classification on EyePACS dataset.

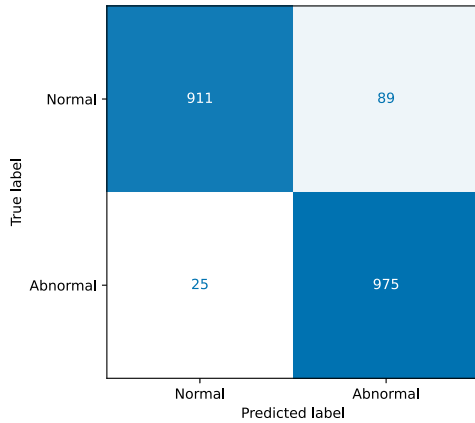


FIGURE 7. Confusion matrix of the proposed GATL for DR classification on APTOS dataset.

and intuitively display the distribution of data samples. As shown in Figure 5, the fundus image features after passing through the GATL network are clearly divided into two categories. The same effect was also shown on another dataset as shown in Figure 6, which shows that GATL uses GCN to extract potential features between samples, which improves the classification accuracy.

5) PERFORMANCE OF CONFUSION MATRIX

In order to analyze the classification performance of GATL and adjust the parameters, we draw confusion matrices as shown in Figures 7 and 8 based on two public datasets. The confusion matrix can quickly visualize the proportion of various misclassified categories into other categories, which can help researchers adjust subsequent models, such as setting weight attenuation for some categories. From Figure 7 and 8 we can observe that the misdiagnosis rate of GATL is very low, which is of great significance in practical applications.

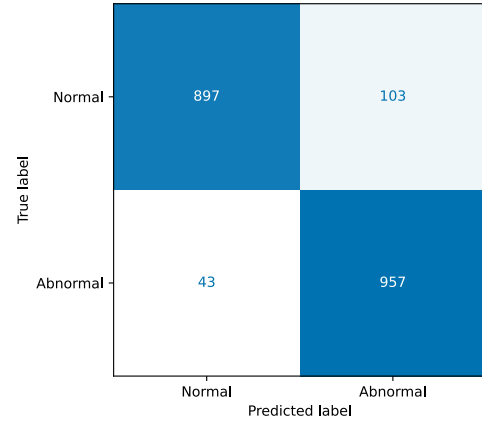


FIGURE 8. Confusion matrix of the proposed GATL for DR classification on EyePACS dataset.

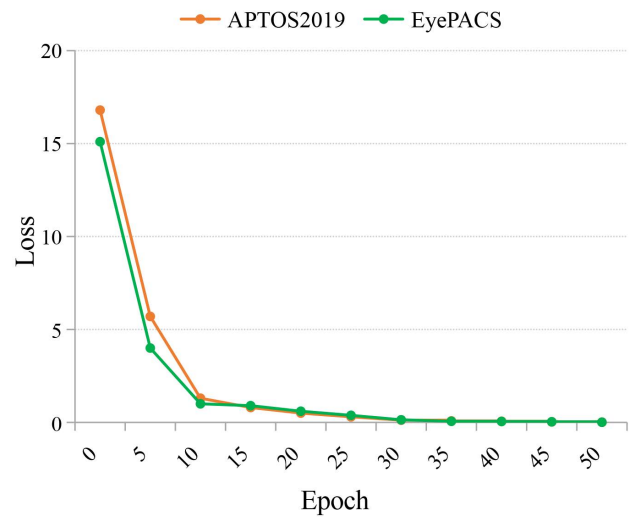


FIGURE 9. The graph of the loss function from start to the model convergence.

6) PERFORMANCE OF LOSS GRAPH

To imply the model training performance, we visualize the graph of loss function from start to model convergence in the training process. Figure 9 summarizes the training loss curves for APTOS2019 and EyePACS datasets, showing the model begins convergence at 10 epochs and remains stable in the following training steps. This visualization further proves the model robustness of our GATL approach, bringing it into correspondence with above-mentioned visualizations.

7) FEATURE MAP VISUALIZATION FOR INTERMEDIATE STEPS

To visualize the sample images in the intermediate steps, we generate feature maps from the first convolutional layer for an abnormal retinal image. As shown in Figure 11, it is key evidence to visualize the shape and texture information seen by the convolution layers. That demonstrates the proposed GATL can explicitly identify and recognize patterns inside the network models.



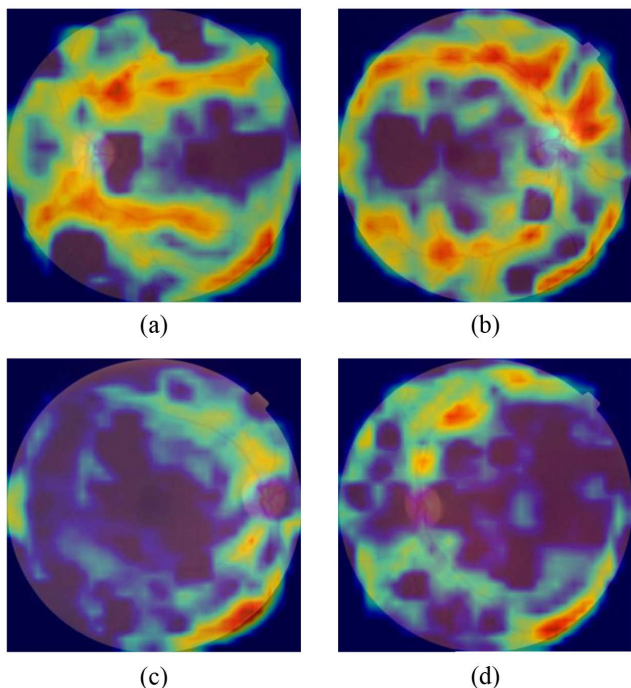


FIGURE 10. CAM images of APTOS 2019, where (a), (b) are abnormal samples and (c), (d) are normal samples.

TABLE 4. Compared with baseline methods for multiple DR classification on APTOS 2019 datasets.

Methods	APTOS 2019		
	Accuracy	Sensitivity	Specificity
ResNet50 [52]	74.6	56.5	85.7
MobileNet [53]	79	76.4	84.6
Vgg [41]	80	85.3	86.6
Bodapati et al. [54]	80.9	-	-
Bodapati et al. [55]	82.5	-	-
GATL	83.5	69.2	97.8

### 8) CLASS ACTIVATION MAP VISUALISATION

To further evaluate the performance of GATL, we visualized the Class Activation Map (CAM) of the features extracted by GATL, as shown in Figure 10. CAM is a tool that helps researchers [51] visualize CNNs. It can clearly show the image regions that the network is focusing on. DR can present with microaneurysms, hemorrhages, hard and soft exudates, and microvascular abnormalities within the retina, which are usually in the vicinity of blood vessels. The CAM diagram shows that the GATL network’s attention is concentrated on the lesion region surrounding the vessel in the DR image, demonstrating that GATL is capable of accurately localizing the vessel location for lesion identification.

### E. FURTHER ANALYSIS

In this section we analyze some factors that affect GATL performance.

#### 1) THE IMPACT OF PRE-TRAINING

In order to explore the effect of pre-training on the model effect, we compared the effect of pre-training and

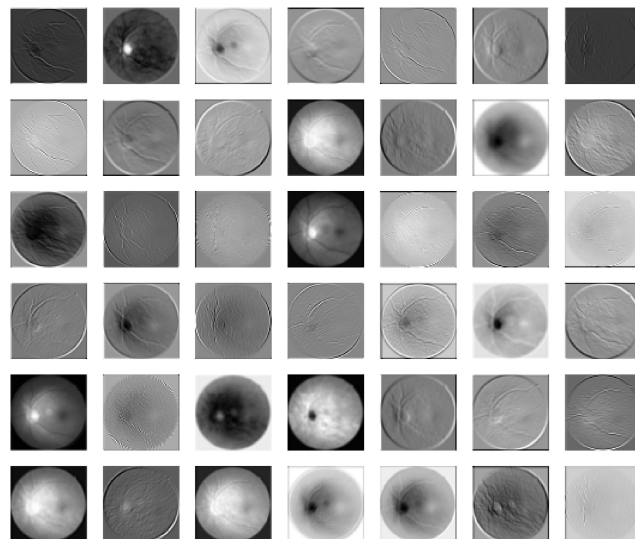


FIGURE 11. The feature map visualization for an abnormal sample.

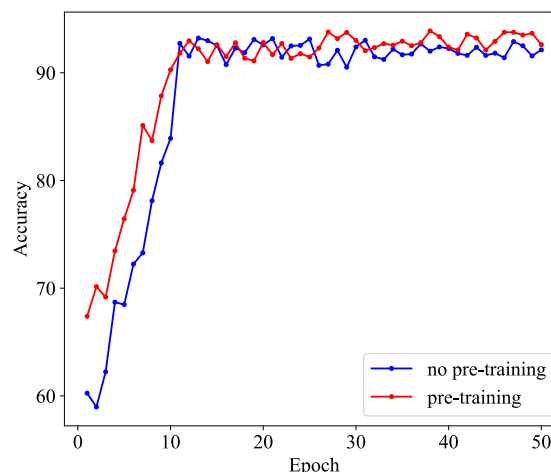


FIGURE 12. Pre-training influence diagram.

non-pre-training models. As shown in Figure 12, the red line represents the change process of the classification accuracy with the increase of epoch after adding ImageNet pre-training. The blue line represents the result of not adding pre-training. From 12, it can be observed that the accuracy of the pre-trained GATL is higher than that of the non-pre-trained model at the beginning of training, which is because the pre-trained network model already has the potential to extract simple features. After a period of training, the accuracy of the pre-training model has gradually grown from 68% to 90%. At the same time, the accuracy of the un-pre-trained model after training increased from 60% to 83%. This means that pre-trained models can converge faster and achieve good results. In subsequent training, the accuracy of both the pre-trained GATL and the non-pre-trained network gradually leveled off, but the accuracy of the pre-trained GATL was higher than that of the non-pre-trained model. This demonstrates that the pre-trained network can extract essential characteristics

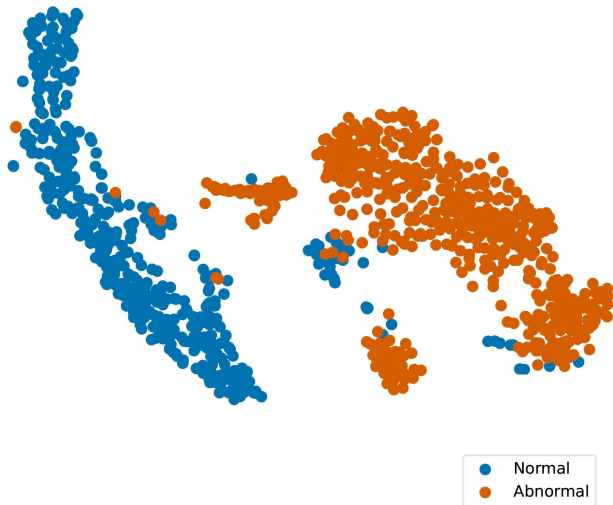


FIGURE 13. The TSNE diagram after removing the graph neural network.

and, to some part, demonstrates the superiority of transfer learning itself, implying that using transfer learning to DR classification is efficient. GATL can effectively use the source domain data, transfer the knowledge of the known domain to the new domain, and ensure that the good effect of the model can be further improved. In the diagnosis of some clinical diseases, experts need to spend a lot of energy to screen the medical results of patients, which causes labeled data to be precious and challenging to obtain. Our GATL reduces the model’s dependence on labeled data to some extent.

2) IMPACT OF THE GRAPH NEURAL NETWORK

We then analyze the influence of the GNN module. We removed the GNN module when performing experiments on the EyPACS data set and finally reduced the dimensions of the extracted features and made TSNE maps, as shown in Figure 13. It can be observed that after removing the GNN module, the number of fuzzy samples has increased a lot. This shows that the graph neural network is of great significance for the extraction of latent features. The GNN module plays an essential role in GATL.

3) THE EFFECTIVENESS OF DISCRIMINATOR  $D$

The discriminator  $D$  provides the domain-differentiation ability, which is optimized by loss  $\mathcal{L}_{dis}$ . To evaluate the effectiveness of discriminator  $D$ , this paper removes  $\mathcal{L}_{dis}$  in the final objective function, and it obtains accuracies of 91.8% and 92.7% on APTOS2019 and EyePACS datasets (Table 2 and Table 3), respectively. That states the discriminator improves the accuracy performance of 2.5% and 2.9% on both datasets, and verifies the necessary of the discriminator  $D$ .

4) THE IMPACT OF MULTIPLE CLASSIFICATIONS

To explore whether the proposed GATL is generalizable in a multiclassification task, we conducted experiments on DR

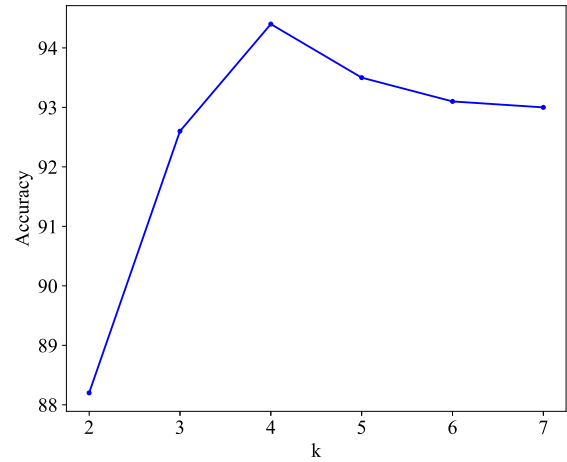


FIGURE 14. Variation in GATL accuracy at different k values.

TABLE 5. Accuracy of GATL for different values of the balance coefficient  $\lambda$  in step 1.

step1	APTOS 2019	
$\lambda$	Accuracy	Sensitivity
0.1	86.6	83
0.2	89	87.2
0.3	89.8	87.2
0.4	91.2	90.8
0.5	89.3	88.6

TABLE 6. Accuracy of GATL for different values of the balance coefficient  $\lambda_1, \lambda_2,$  and  $\lambda_3$  in step 2.

step2			APTOS 2019	
$\lambda_1$	$\lambda_2$	$\lambda_3$	Accuracy	Sensitivity
0.3	0.3	0.4	87.2	88.2
0.35	0.35	0.3	89.6	90.4
0.4	0.4	0.2	90.2	90.8
0.45	0.45	0.1	89.4	89.6

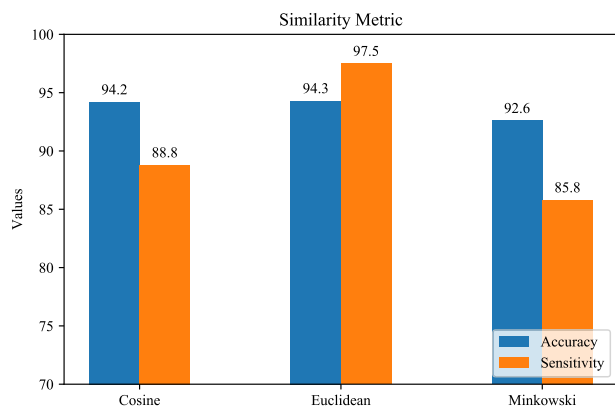
with five classifications and collated the results into Table 4 for comparison with other methods. It can be observed that GATL also achieved good results on the multiclassification task. 83.5%, 69.2% and 97.8% were achieved in terms of accuracy, specificity and sensitivity, respectively. Of these, GATL achieved the best results in terms of accuracy and specificity, suggesting that GATL has general ability and a low rate of misdiagnosis.

5) THE IMPACT OF K

In our model,  $k$  is a significant adjustable parameter of the relationships initializing module, which has a huge impact on accuracy. In order to explore the effect of changing  $k$  values on the model, we took different  $k$  values for several experiments when performing binary classification under the APTOS dataset. Limited by the memory of the graphics card, we only set  $k$  empirically between [2], [7], and the results are shown in Figure 14. It can be observed that the accuracy of GATL reaches its highest value when  $k$  is 4, so for the experiments, we set  $k = 4$ .

**TABLE 7. Accuracy of GATL for different values of the balance coefficient  $\lambda$  in step 3.**

step3	APTOS 2019	
$\lambda$	Accuracy	Sensitivity
0.1	91.2	93.4
0.2	91.6	92.8
0.3	92.6	94.6
0.4	94.3	97.5
0.5	93.7	96.8



**FIGURE 15. The performance of GATL when using different similarity metrics.**

6) THE IMPACT OF  $\lambda$

To analyze the effect of the balance coefficients on the experiment, we adjusted and recorded the accuracy of GATL with different parameters during different training sessions as shown in Tables 5, 6, and 7. As can be observed from Table 5, the model achieves its best results in Step 1 when  $\lambda = 0.4$ . From Table 6, it can be observed that in Step 2, to maximize the differentiation of classifiers  $F1$  and  $F2$  while ensuring accuracy, the model achieves the highest accuracy and sensitivity when  $\lambda_1 = 0.4$ ,  $\lambda_2 = 0.4$ , and  $\lambda_3 = 0.2$ . In Step 3, as seen in Table 7, GATL achieves the best results when  $\lambda = 0.4$ .

7) THE IMPACT OF SIMILARITY METRICS

Many studies have found that in some data the Euclidean distance does not reflect the true distance between features very well [56], [57], [58]. In order to analyze which similarity measure is more suitable for GATL in DR data, we experimented with three similarity metrics on APTOS dataset and recorded the variation in model performance in Figure 15. It can be observed that the use of cosine and Euclidean distances has little effect on model accuracy, but the sensitivity is relatively low when GATL uses cosine distances to calculate feature similarity. This implies that the numerical differences are more prominent for lesion characteristics in DR images than the dimensional differences. In addition to this, we experimented with the Minkowski distance, but the adjustment of the parameter  $p$  in the Minkowski distance was difficult. In the end, we empirically set  $p$  to 1.5, and the results are shown in Figure 15. The experimental findings

demonstrate that the Euclidean distance is more suited for determining how related DR image features are.

**V. CONCLUSION AND DISCUSSION**

This study proposes a novel graph adversarial transfer learning method to achieve DR classification via the domain adaptation structure. In order to successfully employ the extra information for DR classification of unidentified data, we first provide a unique graph-based domain adaption strategy for investigating additional fundus image characteristics. Second, we create two adversarial classifiers for use in training the model. This enhances the model’s classification performance in addition to strengthening its robustness. Finally, we design a discriminator and obfuscate the discriminator’s judgment of the source and target domain data such that the model may be aligned between domains on top of the intra-domain alignment. Specifically, GATL combines the characteristics of convolutional neural networks and graph neural networks to extract the pixel features and potential features of fundus images. Under the guidance of adversarial training, the goal is to align within and between domains. Extensive results on two public datasets show that our GATL method is better than some baseline DR classification methods.

GATL is well compatible with a wide range of data, making more efficient use of existing data while still maintaining a high accuracy rate for classifying DR. However, the only regret about GATL is that it cannot take into account the fine classification of lesions. Because of the small size and similarity of DR lesions, GATL can only determine with high accuracy whether a lesion is present or not, but further analysis is required to determine the degree of illness. We hope to overcome this problem in our next research work and contribute to the fine classification of DR.

**REFERENCES**

- [1] K. Little, M. Llorián-Salvador, S. Scullion, C. Hernández, O. Simó-Servat, A. Del Marco, E. Bosma, M. Vargas-Soria, M. J. Carranza-Naval, and T. Van Bergen, “Common pathways in dementia and diabetic retinopathy: Understanding the mechanisms of diabetes-related cognitive decline,” *Trends Endocrinol. Metabolism*, vol. 33, no. 1, pp. 50–71, 2022.
- [2] E. Kooroor, S. K. Chauhan, and A. Hajrasouliha, “Role of inflammatory cells in pathophysiology and management of diabetic retinopathy,” *Surv. Ophthalmol.*, vol. 67, pp. 1563–1573, Jul. 2022.
- [3] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, and J. Cuadros, “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs,” *J. Amer. Med. Assoc.*, vol. 316, no. 22, pp. 2402–2410, 2016.
- [4] 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.-Assist. Intervent.* Quebec City, QC, Canada: Springer, 2017, pp. 267–275.
- [5] Y. Yang, F. Shang, B. Wu, D. Yang, L. Wang, Y. Xu, W. Zhang, and T. Zhang, “Robust collaborative learning of patch-level and image-level annotations for diabetic retinopathy grading from fundus image,” *IEEE Trans. Cybern.*, vol. 52, no. 11, pp. 11407–11417, Nov. 2021.
- [6] Eyepacs Dataset. Accessed: Mar. 12, 2022. [Online]. Available: <https://www.kaggle.com/c/diabetic-retinopathy-detection/data>
- [7] S. Dutta, B. C. Manideep, S. M. Basha, R. D. Caytiles, and N. C. S. N. Iyengar, “Classification of diabetic retinopathy images by using deep learning models,” *Int. J. Grid Distrib. Comput.*, vol. 11, no. 1, pp. 89–106, Jan. 2018.

- [8] S. H. Kassani, P. H. Kassani, R. Khazaeinezhad, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Diabetic retinopathy classification using a modified xception architecture," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2019, pp. 1–6.
- [9] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," *Proc. Comput. Sci.*, vol. 90, pp. 200–205, Jan. 2016.
- [10] Y. Zhou, X. He, L. Huang, L. Liu, F. Zhu, S. Cui, and L. Shao, "Collaborative learning of semi-supervised segmentation and classification for medical images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 2079–2088.
- [11] Z. Lin, R. Guo, Y. Wang, B. Wu, T. Chen, W. Wang, D. Z. Chen, and J. Wu, "A framework for identifying diabetic retinopathy based on anti-noise detection and attention-based fusion," in *Proc. Int. Conf. Med. Image Comput.-Assist. Intervent. Granada, Spain: Springer*, 2018, pp. 74–82.
- [12] Y. Yang, T. Li, W. Li, H. Wu, W. Fan, and W. Zhang, "Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks," in *Proc. Int. Conf. Med. Image Comput.-Assist. Intervent. Quebec City, QC, Canada: Springer*, 2017, pp. 533–540.
- [13] F. Li, Y. Wang, T. Xu, L. Dong, L. Yan, M. Jiang, X. Zhang, H. Jiang, Z. Wu, and H. Zou, "Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs," *Eye*, vol. 36, no. 7, pp. 1433–1441, Jul. 2022.
- [14] M. Saini and S. Susan, "Diabetic retinopathy screening using deep learning for multi-class imbalanced datasets," *Comput. Biol. Med.*, vol. 149, Oct. 2022, Art. no. 105989.
- [15] L. Dai, L. Wu, H. Li, C. Cai, Q. Wu, H. Kong, R. Liu, X. Wang, X. Hou, Y. Liu, X. Long, Y. Wen, L. Lu, Y. Shen, Y. Chen, D. Shen, X. Yang, H. Zou, B. Sheng, and W. Jia, "A deep learning system for detecting diabetic retinopathy across the disease spectrum," *Nature Commun.*, vol. 12, no. 1, pp. 1–11, Dec. 2021.
- [16] Y. Zhang, W. Deng, M. Wang, J. Hu, X. Li, D. Zhao, and D. Wen, "Global-local GCN: Large-scale label noise cleansing for face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 7731–7740.
- [17] C. Qi, J. Zhang, H. Jia, Q. Mao, L. Wang, and H. Song, "Deep face clustering using residual graph convolutional network," *Knowl.-Based Syst.*, vol. 211, Jan. 2021, Art. no. 106561.
- [18] D. Mahapatra, "Unsupervised domain adaptation using feature disentanglement and GCNs for medical image classification," 2022, *arXiv:2206.13123*.
- [19] S. Duan, P. Huang, M. Chen, T. Wang, X. Sun, M. Chen, X. Dong, Z. Jiang, and D. Li, "Semi-supervised classification of fundus images combined with CNN and GCN," *J. Appl. Clin. Med. Phys.*, pp. 1–13, Aug. 2022.
- [20] B. Yang, H. Pan, J. Yu, K. Han, and Y. Wang, "Classification of medical images with synergic graph convolutional networks," in *Proc. IEEE 35th Int. Conf. Data Eng. Workshops (ICDEW)*, Apr. 2019, pp. 253–258.
- [21] Q. Liu, L. Xiao, J. Yang, and Z. Wei, "CNN-enhanced graph convolutional network with pixel-and superpixel-level feature fusion for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 10, pp. 8657–8671, Oct. 2020.
- [22] Z.-M. Chen, X.-S. Wei, P. Wang, and Y. Guo, "Multi-label image recognition with graph convolutional networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 5177–5186.
- [23] P.-N. Chen, C.-C. Lee, C.-M. Liang, S.-I. Pao, K.-H. Huang, and K.-F. Lin, "General deep learning model for detecting diabetic retinopathy," *BMC Bioinf.*, vol. 22, no. S5, pp. 1–15, Nov. 2021.
- [24] A. Erciyas and N. Barşı, "An effective method for detecting and classifying diabetic retinopathy lesions based on deep learning," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–13, May 2021.
- [25] V. Vives-Boix and D. Ruiz-Fernández, "Diabetic retinopathy detection through convolutional neural networks with synaptic metaplasticity," *Comput. Methods Programs Biomed.*, vol. 206, Jul. 2021, Art. no. 106094.
- [26] A. S. Jadhav, P. B. Patil, and S. Biradar, "Optimal feature selection-based diabetic retinopathy detection using improved rider optimization algorithm enabled with deep learning," *Evol. Intell.*, vol. 14, pp. 1–18, Dec. 2020.
- [27] M. Canayaz, "Classification of diabetic retinopathy with feature selection over deep features using nature-inspired wrapper methods," *Appl. Soft Comput.*, vol. 128, Oct. 2022, Art. no. 109462.
- [28] E. Abdelmaksoud, S. Barakat, and M. Elmogy, "A computer-aided diagnosis system for detecting various diabetic retinopathy grades based on a hybrid deep learning technique," *Med. Biol. Eng. Comput.*, vol. 60, pp. 1–24, May 2022.
- [29] C. Zhang, T. Lei, and P. Chen, "Diabetic retinopathy grading by a source-free transfer learning approach," *Biomed. Signal Process. Control*, vol. 73, Mar. 2022, Art. no. 103423.
- [30] A. K. Gangwar and V. Ravi, "Diabetic retinopathy detection using transfer learning and deep learning," in *Evolution in Computational Intelligence*. Singapore: Springer, 2021, pp. 679–689.
- [31] S.-L. Yi, X.-L. Yang, T.-W. Wang, F.-R. She, X. Xiong, and J.-F. He, "Diabetic retinopathy diagnosis based on RA-EfficientNet," *Appl. Sci.*, vol. 11, no. 22, p. 11035, Nov. 2021.
- [32] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Trans. Neural Netw.*, vol. 22, no. 2, pp. 199–210, Feb. 2010.
- [33] Y. Wang, Y. Feng, L. Zhang, Z. Wang, Q. Lv, and Z. Yi, "Deep adversarial domain adaptation for breast cancer screening from mammograms," *Med. Image Anal.*, vol. 73, Oct. 2021, Art. no. 102147.
- [34] F. J. Castellanos, A.-J. Gallego, and J. Calvo-Zaragoza, "Unsupervised neural domain adaptation for document image binarization," *Pattern Recognit.*, vol. 119, Nov. 2021, Art. no. 108099.
- [35] C. Abbet, L. Studer, A. Fischer, H. Dawson, I. Zlobec, B. Bozorgtabar, and J.-P. Thiran, "Self-rule to multi-adapt: Generalized multi-source feature learning using unsupervised domain adaptation for colorectal cancer tissue detection," *Med. Image Anal.*, vol. 79, Jul. 2022, Art. no. 102473.
- [36] V. Konyakhin, N. Lukashina, and A. Shpilman, "Solving Traffic4Cast competition with U-Net and temporal domain adaptation," 2021, *arXiv:2111.03421*.
- [37] J. Hong, S. C.-H. Yu, and W. Chen, "Unsupervised domain adaptation for cross-modality liver segmentation via joint adversarial learning and self-learning," 2021, *arXiv:2109.05664*.
- [38] *Aptos 2019 Blindness Detection*. Accessed: Mar. 16, 2022. [Online]. Available: <https://www.kaggle.com/c/aptos2019-blindness-detection/>
- [39] J. Hu, L. Shen, and G. Sun, "Squeeze- and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.
- [40] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 6105–6114.
- [41] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
- [42] A. Sugeno, Y. Ishikawa, T. Ohshima, and R. Muramatsu, "Simple methods for the lesion detection and severity grading of diabetic retinopathy by image processing and transfer learning," *Comput. Biol. Med.*, vol. 137, Oct. 2021, Art. no. 104795.
- [43] G. T. Zago, R. V. Andreão, B. Dorizzi, and E. O. Teatini Salles, "Diabetic retinopathy detection using red lesion localization and convolutional neural networks," *Comput. Biol. Med.*, vol. 116, Jan. 2020, Art. no. 103537.
- [44] W. Chen, B. Yang, J. Li, and J. Wang, "An approach to detecting diabetic retinopathy based on integrated shallow convolutional neural networks," *IEEE Access*, vol. 8, pp. 178552–178562, 2020.
- [45] M. T. Hagos and S. Kant, "Transfer learning based detection of diabetic retinopathy from small dataset," 2019, *arXiv:1905.07203*.
- [46] G. García, J. Gallardo, A. Mauricio, J. López, and A. D. Carpio, "Detection of diabetic retinopathy based on a convolutional neural network using retinal fundus images," in *Proc. Int. Conf. Artif. Neural Netw. Alghero, Italy: Springer*, 2017, pp. 635–642.
- [47] S. M. S. Islam, M. M. Hasan, and S. Abdullah, "Deep learning based early detection and grading of diabetic retinopathy using retinal fundus images," 2018, *arXiv:1812.10595*.
- [48] A. Rakhlin, "Diabetic retinopathy detection through integration of deep learning classification framework," *BioRxiv*, Jan. 2018, Art. no. 225508.
- [49] R. R. Kumar, R. Pandian, T. P. Jacob, A. Pravin, and P. Indumathi, "Detection of diabetic retinopathy using deep convolutional neural networks," in *Computational Vision and Bio-Inspired Computing*. Singapore: Springer, 2021, pp. 415–430.
- [50] S. Yu, K. Ma, Q. Bi, C. Bian, M. Ning, N. He, Y. Li, H. Liu, and Y. Zheng, "MIL-VT: Multiple instance learning enhanced vision transformer for fundus image classification," in *Proc. Int. Conf. Med. Image Comput.-Assist. Intervent. Istanbul, Turkey: Springer*, 2021, pp. 45–54.
- [51] M. Lucas, M. Lerma, J. Furst, and D. Raicu, "Visual explanations from deep networks via Riemann–Stieltjes integrated gradient-based localization," 2022, *arXiv:2205.10900*.
- [52] 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.

[53] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[54] V. Naralasetti, S. N. Shareef, and S. Hakak, "Blended multi-modal deep convnet features for diabetic retinopathy severity prediction," *Electronics*, vol. 9, no. 6, p. 914, 2020.

[55] J. D. Bodapati, N. S. Shaik, and V. Naralasetti, "Composite deep neural network with gated-attention mechanism for diabetic retinopathy severity classification," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 10, pp. 9825–9839, Oct. 2021.

[56] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, "SphereFace: Deep hypersphere embedding for face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 212–220.

[57] J. Wu, Y. Zhang, R. Fu, Y. Liu, and J. Gao, "An efficient person clustering algorithm for open checkout-free groceries," 2022, *arXiv:2208.02973*.

[58] P. Fuangkhan, "Effect of the distance functions on the distance-based instance selection for the feed-forward neural network," *Evol. Intell.*, vol. 15, no. 3, pp. 1991–2015, Sep. 2022.



**HUAN WANG** received the master's and Ph.D. degrees in system analysis and integration from Yunnan University, in 2008 and 2013, respectively. She is currently an Associate Professor with the Baoji University of Arts and Sciences. Her research interests include complex networks and complex system research.



**LE WANG** received the Ph.D. degree in biomedical engineering from Xi'an Jiaotong University, in 2015. He is currently a Researcher with the Baoji University of Arts and Sciences. His research interests include mass spectrometry multiomics analysis and drug bioinformatics.



**JINGBO HU** received the master's degree in computer architecture from Yunnan University, in 2010. He is currently an Associate Professor with the Baoji University of Arts and Sciences. His research interests include target detection and complex system research.



**YE LU** received the master's and doctor's degrees from the Lanzhou University of Technology, in 2011 and 2014, respectively. He is currently a Lecturer with the Baoji College of Arts and Sciences. His current research interests include industrial network security and the application of block chain in access control.

...