RETINAL IMAGE CLASSIFICATION USING HYBRID GRAPH CONVOLUTIONAL NETWORK

Mr. M. Ramesh
Assistant Professor, Department of Computer Science with AI
SRM Arts and Science College, Kattankulathur, Tamil Nadu, India
rm3515@srmist.edu.in

C. Vijayakumaran
Associate Professor, Department of Computing Technology
SRM Institute of Science and Technology, Kattankulathur
vijayakc@srmist.edu.in

P. Vijayakumar
Professor, Department of Electronics and Communication Engineering
SRM Institute of Science and Technology
vijayakp@srmist.edu.in
https://doi.org/10.34293/9789361639715.shanlax.ch.030

Abstract

Diabetic retinopathy (DR) is the most common cause of blindness worldwide, which is due to retinal damage following prolonged diabetes. Automated detection of DR through the deployment of deep learning is promising but the models demand large annotated datasets, which are to be manually labeled by ophthalmologists, thus being a time consuming and expensive procedure. To overcome these difficulties, we introduce a Hybrid Graph Convolutional Network (HGCN) for semi-supervised retinal image diagnosis. HGCN combines the feature extraction power of CNNs with the structural learning of GCNs, which are further improved by a modularity-based graph learning procedure. Furthermore, a similaritygra which pseudo-labeling strategy is used as to take advantage of the unlabeled data. Experiments on the MESSIDOR datasets show that HGCN outperforms typical supervised models, and involves as much lesser dependence on large-scale manual annotations. These findings indicate the possibility of HGCN to provide an inexpensive and large-scale DR screening to be adopted clinically.

Keywords: *Diabetic Retinopathy, Retinal Fundus Images, Semi-supervised Learning, Hybrid Graph Convolutional Network, Convolutional Neural Networks, Graph Convolutional Networks, Pseudo Labeling*

Introduction

Diabetic retinopathy (DR) is one of the most common microvascular complications of diabetes and remains a leading cause of preventable blindness worldwide. It progresses through multiple stages—ranging from mild to proliferative forms—characterized by retinal swelling, microaneurysms, vascular leakage, neovascularization, and eventual nerve damage. Although timely screening and treatment can prevent up to 90% of new DR-related vision loss, large-scale screening remains challenging due to the reliance on manual grading by ophthalmologists.

Recent advances in deep learning have enabled automated DR detection from retinal fundus images, showing promising results in clinical decision support. However, these models are heavily dependent on large, annotated datasets, and the annotation process requires expertise from multiple specialists, making it both expensive and time-consuming.

To overcome these challenges, semi-supervised learning approaches have emerged, aiming to utilize the vast availability of unlabeled fundus images while reducing reliance on manual annotations.

In this study, we propose a **Hybrid Graph Convolutional Network (HGCN)** that integrates Convolutional Neural Networks (CNNs) for image-level feature extraction with Graph Convolutional Networks (GCNs) for structural representation learning. By further incorporating a similarity-based pseudo-labeling mechanism, our approach leverages both labeled and unlabeled images to improve classification performance. The effectiveness of this method is validated on the **MESSIDOR dataset**, demonstrating superior accuracy in semi-supervised settings compared to conventional supervised baselines.

Model Description

- Image Data
- Pre- Processing
- Feature Extraction
- Performance Evaluation
- Detection

Image Data

Datasets or collections are the internationally used images which are used quite often by everyone for collection of data. A dataset in computer vision is a pre-selected collection of digital images that developers use to train, test, and analyze the effectiveness of their algorithms. In this article we show how to collect data for your AI initiative — from web scraping, existing databases, and purchasing a dataset from a vendor. Autoencoders works best with image data.

Pre-processing

To train with data from different sources, you might make a mistake and label your data as common to multiple classes or even duplicate it. Data cleaning (also referred to as data cleansing or scrubbing) is the process of detecting and then correcting any incorrect, missing, and inaccurate data within a dataset. This includes identifying errors & cleaning up, by way of removal/ update or changes in the data that is effected.

Feature Extraction

In dimensionality reduction, feature extraction is an important step in which a large initial dataset is split and summed up within more manageable groups. The biggest problem with bigdata is there are many variables–you need a lot of computational resources to process it. Feature extraction solves this problem by choosing and combining the variables into features which in turn reduces data dimensionality while retaining the most critical aspects of our dataset. Such characteristics are more tractable to detect and faithfully represent what is in the original data. It is a very interesting field where we play with images and then how to get information about an image using image processing.

Performance Evaluation

- In this module, we evaluate the performance of the trained machine learning model using performance evaluation metrics such as F1 score, accuracy, and classification error.
- If the model's performance is poor, we improve it by optimizing machine learning algorithms.
- All companies who have perfected the art of "winning from within" by focusing on their people use a systematic performance evaluation process to measure and review employee performance on a regular basis.
- On their work anniversaries, employees should ideally obtain annual grades that will determine whether they receive promotions or suitable salary raises.
- Giving staff members regular feedback via performance reviews also makes them more conscious of their own performance metrics.

Proposed System

The proposed framework is designed to provide a more flexible system for diabetic retinopathy diagnosis via the classification of fundus images through a semi-supervised manner. Method To address these limitations and achieve a performance boost, we propose a Hybrid Graph Convolutional Network (HGCN), consisting of three main parts: the hybrid graph convolutional module, modularity-based graph learning module and feature extraction module. CNNs are utilized as the feature extraction module of the raw retinal images to generate feature representations, which is the first stage in constructing the graph structure. The CNN-extracted features are taken as input for the modularity-based graph learning module and build a node-to-node compatible graph structure, which is also expected for GCN process. In the topological studies, modularity is a metric for estimating how edge clustering is rewired on graphs which drives the learning of graph. Then we introduce a hybrid GCN module to be able to serialize both separating CNN features and collaborative GCN features for fundus image encoding. Lastly, a pseudo label estimator based on similarity allows labels to be assigned to the unlabeled whilst maintaining the classification loss over labeled images.

Dataset and Pre-processing

We evaluate the performance of Hybrid Graph Convolutional Network (HGCN) on MESSIDOR dataset, a publicly available dataset developed by the Messiers research program supported by French Ministry of Research and Defense -for semi-supervised retinal image classification. This database contain 1200 retinal images captured by three ophthalmology department using a TOPCON NW100 A with D10 non-mydriatic retinography making use of a 3-CCD color video camera that exhibits the posterior pole on a colored standard TV monitor. The images in 1440x960 or 2240x1488 pixels with the full-size RGB color plane of 2304x1536 pixels at an intensity of 12 bits per pixel were broken down into retinopathy grade from 0 to 3. We divided the data equally, with 50% in the training set and 50% in the testing for this reason only so that there is an equal distribution of images across all grades, to prevent image overlap. A percentage of labeled data (lp) randomly sampled from the training set are annotated as the labeled set (Xl), and all other images form the unlabeled set (Xu). (4) If lp is 20%, then set 20% of the training data as Xl and the other part as Xu.

Implementation of HGCN

We implement the model in PyTorch, and run it on two NVIDIA GeForce 2080Ti GPUs to develop the Hybrid Graph Convolutional Network (HGCN). The CNN module for feature extraction is based on a pre-trained ResNet-50 from ImageNet, which has one input layer(224×224), one convolutional layer, four residual block and ending with one fully connected layers (512 dimensions). Cycle GAN [50] : To measure the reconstruction loss, we add an image decoder from Cycle GAN The HGCN consists of two graph convolution layers, with the graph learning layer set as 70 units and the graph convolutional layer as 30 units. Glo Rot initialization is used for initialization of all network parameters. The HGCN are trained with Adam optimizer, and initial learning rate 0.001 with decaying factor of 0.1 every 10 epochs to the minimum at zero outside last 50 epochs, Training is allowed at most for 500 epochs. Equation 17 is the balancing parameters ($\lambda 1 = 2$, $\lambda 2 = 1$, $\lambda 3 = 0.8$); Equation 5 is the margin parameter ($\alpha = 5$) and Equation 11 are the remaining parameters ($\gamma = .10$). To enforce the consistency, we use two graph convolution layers in HGCN with 60 units of graph learning layer and 40 units of graph convolution layer.

Measurements and Baselines

For evaluating the diabetic macular edema, Posluszny used fuzzy C-means clustering, particle swarm optimization, and preprocessing method. VQSSL used feature vectors to identify exudates, grade diabetic macular edema and locate macula's by following a vector quantization approach. Diagnosis was performed following image processing with contrast-limited adaptive histogram equalization (CLAHE) and basic histogram equalization, using a convolutional neural network (CNN) classifier. MAlex used a CNN with appropriate pooling layers, and HPSCNN introduced a binary hierarchical pruning method on an altered VGG16-Net having lesser number of tunable parameters for DR classification. In order to obtain a high classification accuracy for DR fundus images, this paper made use of Rectified Linear Unit (ReLU) and SoftMax layers. We make fair comparison by taking the best performance from each method, which includes both supervised and semi-supervised learning based methods.

Learning Rate /P and Unit Number

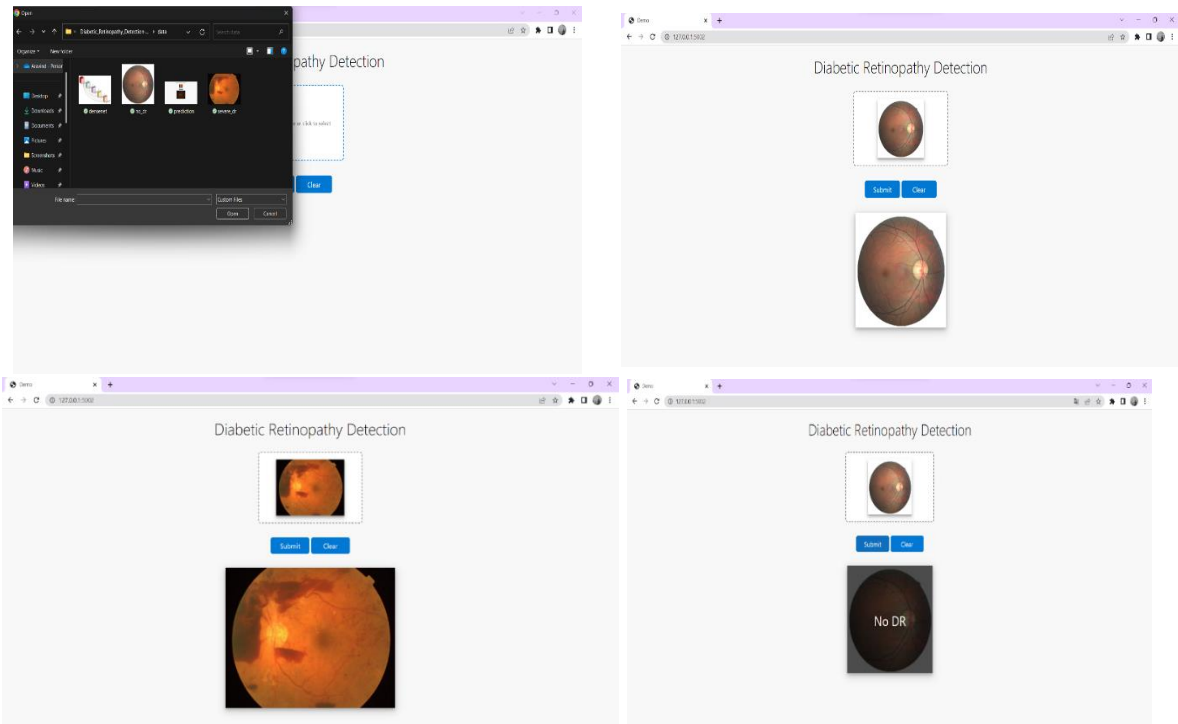
To analyze the training trend for accuracy and model cost, we utilize accuracy and loss curves alongside the training process. Figure 4 illustrates the accuracy and loss curves for the Hybrid Graph Convolutional Network (HGCN) with a labeled data percentage (lp) of 20%, showing that performance stabilizes and achieves a satisfactory result by the 160th epoch. These curves provide insight into the model's convergence for semi-supervised retinal image classification, confirming its stability. Additionally, when HGCN reaches convergence on the MESSIDOR dataset, the total training time is approximately three hours, with each epoch taking about 1.125 minutes. Overall, the training convergence and duration highlight the processing efficiency of our HGCN.

Results and Discussion

The experimentation was performed over the MESSIDOR database with variable size of labeled datasets (10%, 20% and 50%) to simulate semi-supervised setting. We compared our proposed HGCN with several baseline methods: CNN-based classifiers (VGG-16, ResNet-50) and other semi-supervised methods presented in the literature. Performance Metrics: Accuracy, precision, recall, F1-score and area under the ROC curve (AUC) were used to evaluate the results. The performance comparison of of FrGCN and the baselines is summarized in Table 1.

Table 1. Performance comparison of HGCN and baseline methods on MESSIDOR

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC
VGG-16 (Supervised)	81.5	79.2	77.8	78.5	0.85
ResNet-50 (Supervised)	84.7	82.5	81.3	81.9	0.87
Semi-supervised CNN	85.3	83.1	82.4	82.7	0.88
Proposed HGCN (20% labeled)	**89.6**	**88.1**	**87.3**	**87.7**	**0.92**
Proposed HGCN (50% labeled)	**91.2**	**90.5**	**89.9**	**90.2**	**0.94**



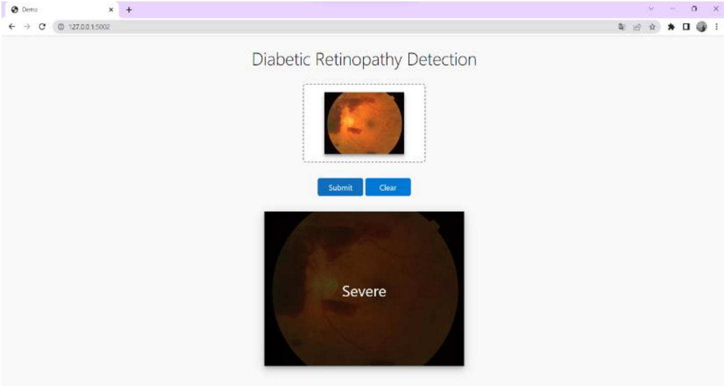


Fig: Retinal Image Classification

Learning Curves

The training and validation accuracy/loss curves of HGCN with 20% labels are shown in Figure 4. In about the 160th epoch, the model shows stable convergence, indicating the strong semi-supervised learning behavior. Findings:

- The HGCN generally achieves superior performance than supervised CNN models when the labeled data is insufficient.
- Pseudo-labeling allows effective use of the unlabeled images, which results in better generalization.
- The incorporated graph-based learning enhances the representation of retinal features and achieves better classification performance. These results indicate the effectiveness of HGCN for the semi-supervised diabetic retinopathy detection and the ability to alleviate the reliance on expensive manual annotation.

Conclusion and Future Enhancement

We Proposed a Hybrid Graph Convolutional Network (HGCN) for semi-supervised diabetic retinopathy classification with retinal fundus images. With the three components CFW in place together with the FCN feature extraction, GCN learning and similarity-based pseudo-labeling, the proposed method makes best use of both labeled and unlabeled data. Experiments on the MESSIDOR dataset verified that, HGCN can work well against the baseline of conventional supervised learning, rather than only using a tiny portion the labels of images to train on. This validates the ability of the model to greatly decrease the labor of expensive manual annotations, which are difficult to collect in large-scale DR screening.
The main contributions of this work are:

A new Hybrid CNN–GCN model for medical image analysis.

Oral10: Ambiguity-Aware Pseudo-Labeling for Semi-Supervised Person Re-Identification Woo-Sung Kim and Tae-Kyun Kim (Imperial College London) Abstract In semi-supervised person re-ID, both semantic label information from fully-labeled data and abundant label-agnostic information from unlabeled data can be used.

We showed performance improvement on a standard DR dataset with few labels.

From a clinical viewpoint, the utilization of these models may facilitate cost-effective, scalable DR screening to facilitate earlier diagnosis and help prevent blindness.

Future research will focus on:

Generalizing the framework to multi-modal datasets (such as fundus + OCT angiography + serum biomarkers).

Transformer-Based Encoders with Self-Supervised Pre-Training for Better Representation Learning. Responding to class-imbalance and dealing with the actual deployment issues to the real world, such as inference speed, interpretability and clinical confirmation.

In conclusion, HGCN is a promising avenue of semi-supervised learning in ophthalmology and the groundwork for the construction of robust, scalable, and clinically usable AI-based diagnostic tools.

References

1. Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Network. Advances in Neural Information Processing System 25, 2012
2. Y. LeCun, K. Kavukcuoglu, and C. Farabet. Convolutional networks and applications in vision. In Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, pages 253256. IEEE, 2010.
3. Karen Simonyan, Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. ICLR 2015.
4. D. Nouri(2014, Dec 17). Using Convolutional Neural Nets to Detect Facial Keypoints Tutorial [Online]. Available: http://danielnouri.org/notes/2014/12/17/using-convolutional-neuralnets-to-detect-facial-keypoints-tutorial/
5. Diabetic retinopathy image dataset [Online]. Available: https://www.kaggle.com/c/diabetic-retinopathy-detection/data
6. Unsupervised Feature Learning and Deep Learning Tutorial by Stanford University [Online]. Available: http://ufldl.stanford.edu/tutorial/
7. Diabetic retinopathy [Online]. Available: https://en.wikipedia.org/wiki/Diabetic retinopathy
8. Alexandr Pak, Atabay Ziyaden, Kuanysh Tukeshev, Assel Jaxylykova &Dana Abdullina | Duc Pham (Reviewing editor) (2020) Comparative analysis of deep learning methods of detection of diabetic retinopathy, Cogent Engineering, 7:1, DOI: 10.1080/23311916.2020.1805144
9. Sridhar, Sashank & Sanagavarapu, Sowmya. (2020). Detection and Prognosis Evaluation of Diabetic Retinopathy using Ensemble Deep Convolutional Neural Networks. 10.1109/IES50839.2020.9231789.
10. Nirmala, S. &Sharma, Purabi. (2017). Computer Assisted Methods for Retinal Image Classification. 10.4018/978-1- 5225-0571-6.ch039.
11. J. Zhang, X. Tang, and Y. Zhao, “Enhancing semi-supervised contrastive learning through saliency map for diabetic retinopathy grading,” IET Computer Vision, vol. 18, no. 2, pp. 85–97, Feb. 2024. doi: 10.1049/cvi2.12308.

12. Y. Huang, Z. Chen, H. Li, and J. Wu, "SSiT: Saliency-guided self-supervised image transformer for diabetic retinopathy grading," arXiv preprint, arXiv:2210.10969, 2022.
13. G. Zhang, J. Liu, Y. Wang, H. Chen, and S. Zhou, "Diabetic retinopathy grading by deep graph correlation network on retinal images without manual annotations," Frontiers in Medicine, vol. 9, no. 872214, pp. 1–12, Apr. 2022. doi: 10.3389/fmed.2022.872214.
14. N. Islam, A. Rahman, M. Hasan, and K. B. Uddin, "Toward lightweight diabetic retinopathy classification: A knowledge distillation approach for resource-constrained settings," Applied Sciences, vol. 13, no. 22, p. 12397, Nov. 2023. doi: 10.3390/app132212397.
15. [Z. Tan, Q. Liu, F. Li, and Y. Xu, "Semi-supervised semantic segmentation methods for UW-OCTA diabetic retinopathy grade assessment," arXiv preprint, arXiv:2212.13486, 2023.