



Multimodal and multi-omics-based deep learning model for screening of optic neuropathy

Ye-ting Lin, Qiong Zhou^{*}, Jian Tan, Yulin Tao

Department of Ophthalmology, The First Affiliated Hospital of Nanchang University, China

ARTICLE INFO

Keywords:

Multimodal₁
Multi-omics₂
Deep Learning₃
Diabetic optic neuropathy₄
Glaucomatous optic neuropathy₅,
Optic neuritis₆

ABSTRACT

Purpose: To examine the use of multimodal data and multi-omics strategies for optic nerve disease screening.

Methods: This was a single-center retrospective study. A deep learning model was created from fundus photography and infrared reflectance (IR) images of patients with diabetic optic neuropathy, glaucomatous optic neuropathy, and optic neuritis. Patients who were seen at the Ophthalmology Department of First Affiliated Hospital of Nanchang University in Jiangxi Province from November 2019 to April 2023 were included in this study. The data were analyzed in single and multimodal modes following the traditional omics, Resnet101, and fusion models. The accuracy and area-under-the-curve (AUC) of each model were compared.

Results: A total of 312 images fundus and infrared fundus photographs were collected from 156 patients. When multi-modal data was used, the accuracy of the traditional omics mode, Resnet101, and fusion models with the training set were 0.97, 0.98, and 0.99, respectively. The accuracy of the same models with the test sets were 0.72, 0.87, and 0.88, respectively. We compared single- and multi-mode states by applying the data to the different groups in the learning model. In the traditional omics model, the macro-average AUCs of the features extracted from fundus photography, IR images, and multimodal data were 0.94, 0.90, and 0.96, respectively. When the same data were processed in the Resnet101 model, the scores were 0.97 equally. However, when multimodal data was utilized, the macro-average AUCs in the traditional omics, Resnet101, and fusion models were 0.96, 0.97, and 0.99, respectively.

Conclusion: The deep learning model based on multimodal data and multi-omics strategies can improve the accuracy of screening and diagnosing diabetic optic neuropathy, glaucomatous optic neuropathy, and optic neuritis.

1. Introduction

Glaucoma and diabetic retinopathy are two of the most common causes of vision loss worldwide [1]. Both conditions cause irreversible damage to the optic nerve. Optic nerve damage happens especially early in diabetic retinopathy. As such, it is important to diagnose diabetic retinopathy as soon as possible to help mitigate vision loss. At present, optic neuropathy is primarily diagnosed by experienced ophthalmologists, and the lack of ophthalmologists in underdeveloped regions may significantly limit the number of

^{*} Corresponding author.

E-mail addresses: 710013268@qq.com (Y.-t. Lin), qiongzD06@126.com (Q. Zhou), 921103364@qq.com (J. Tan), yulin_tao2022@163.com (Y. Tao).

<https://doi.org/10.1016/j.heliyon.2023.e22244>

Received 8 June 2023; Received in revised form 6 November 2023; Accepted 7 November 2023

Available online 15 November 2023

2405-8440/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

patients that are screened and diagnosed in those areas. The emergence of artificial intelligence (AI) has allowed physicians to diagnose eye diseases based on images. AI can screen large amounts of data and diagnose diabetic retinopathy [2], glaucoma [3], and age-related macular degeneration [4] with expert-level diagnostic accuracy. However, AI-based screening for optic neuropathy remains rare. Future applications of AI in ophthalmology can even extend to include the analysis of multimodal data to improve predictive accuracy.

This study aims to develop a deep learning (DL)-based algorithm to simultaneously screen glaucomatous optic neuropathy (GON), diabetic optic neuropathy (DON), and optic neuritis (ON) using multimodal data and multi-omics strategies, as well as compare the diagnostic performance of different models.

2. Materials and methods

2.1. Image collection and pre-processing

This retrospective study was approved by the Ethics Review Committee of First Affiliated Hospital of Nanchang University (Ethical number : IIT2023281). All procedures were carried out in accordance with the principles of the Declaration of Helsinki. We retrospectively collected fundus photography and infrared reflectance (IR) images of patients with DON, GON, and ON, who were admitted to the Ophthalmology Department of First Affiliated Hospital of Nanchang University from November 2019 to April 2023. Fundus and infrared fundus photographs were performed simultaneously. The following inclusion and exclusion criteria were also established: 1) Photographs with a clinical diagnosis of DON, GON, or ON were included in this study; 2) Photographs that depicted optic neuropathy caused by other diseases, such as hypertensive retinopathy, were excluded from this study; and 3) Patients with poor image quality were excluded from this study.

Seven patients were excluded according to the criteria, resulting in 156 patients. A total of 312 images were collected from these patients. A total of 312 images (fundus photographs and infrared fundus photographs) were collected from these patients. All images were obtained using the same fundus photograph (KOWA Nonmyd WX; KOWA Company Ltd., Tokyo, Japan) and infrared reflectance (IR) (Heidelberg Spectralis Optical Coherence Tomography; Heidelberg Engineering, Dosenheim, Germany) machines. All images were acquired by the same ophthalmologist. The fundus and IR images were deidentified and exported in JPG format (768 × 868 pixels and 24 bit/pixel and 1876 × 1876 pixels and 24 bit/pixel, respectively). Two ophthalmologists reviewed the images and classified each as DON, GON, or ON. All images were then randomly assigned to the training and test sets following an 8:2 ratio. ITK-SNAP version 3.8.0 software (NYU Tandon School of Engineering, NY, USA) was used to manually mark the region of interest (ROI) in each image.

2.2. Model development and feature extraction

This study developed a DL neural network-based model that can predict and classify optic neuropathy. All data inputs were

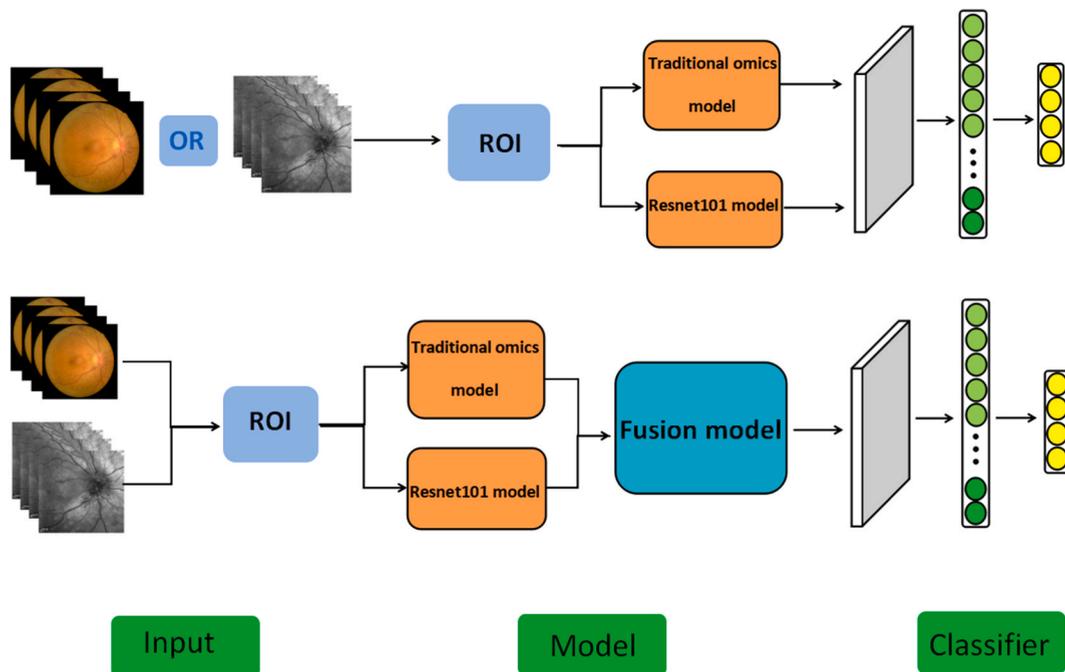


Fig. 1. The flowchart of this study.

anonymized. The data were divided into training and test sets, and the classification of optic neuropathy was assigned as a multi-classification task in the test set. The images of the two modalities were featured separately and then fused and inputted into the neural network for classification. The results of the fusion model were then compared with the results of the traditional omics and Resnet 101 models. The flowchart for this process is shown in Fig. 1.

2.3. Feature extraction of the traditional omics model

The Python package “PyRadiomics 3.0.1” (Python Software Foundation; Wilmington, DE) was used to extract the target features, RF and rad-feature. The features extracted by traditional omics were numerous and quantitative, which may explain some of the resulting heterogeneity [5].Seven radiomic features were extracted: first-order, two-dimensional shape, gray co-incidence matrix, gray running length matrix, gray size band matrix, adjacent gray tone difference matrix, and gray dependence matri [6].Principle component analysis (PCA), least absolute shrinkage and selection operator(LASSO), and Pearson correlation coefficients were used for feature selection and data dimensionality reduction [7].For additional details on the definitions of each of the features, please refer to the PyRadiomics documentary online (<https://pyradiomics.readthedocs.io/en/latest/features.html>). The distribution and gravity of the traditional omics features are shown in Fig. 2.

2.4. Feature extraction of the Resnet101 model

Convolutional neural networks (CNNs), which are composed of input, hidden, and output layers, are currently the most well-known type of deep learning architecture [8]. Resnet [9] is a milestone development among CNNs, the widely used Resnet101 model [10],[11]. was selected for this study. ResNet reduced the error rate of image classification recognition to 3.6 %. Its overall architecture can be divided into three parts: 1) an input stem that uses general convolution with a large stride to reduce image resolution, 2) a stage block with four stages, with each stage composed of multiple building blocks that can use stride or pooling. Each stage typically reduces image resolution, expands width (channel), and carries out a series of residual learning activities; and 3) an output stem that is designed for different tasks. Resnet101 is also known as a DAG network, which is a kind of neural network for DL. Its layers are arranged as directed acyclic graphs, creating a more complex structure that has multiple layers of input and output. ResNet101 also has 101 sub-networks, and each layer responds to or activates a unique input image.

The model network weights were pre-trained and initialized on ImageNet (<http://www.image-net.org/>).To improve performance, the parameters were fine-tuned with the feedback of cross entropy loss function. The average probability of all images was used to produce the DL features, and the outputs of the penultimate FC layer of the convolutional neural network (CNN) were used as DL features. The structure of the CNN was implemented in Python 3.0(Python Software Foundation).

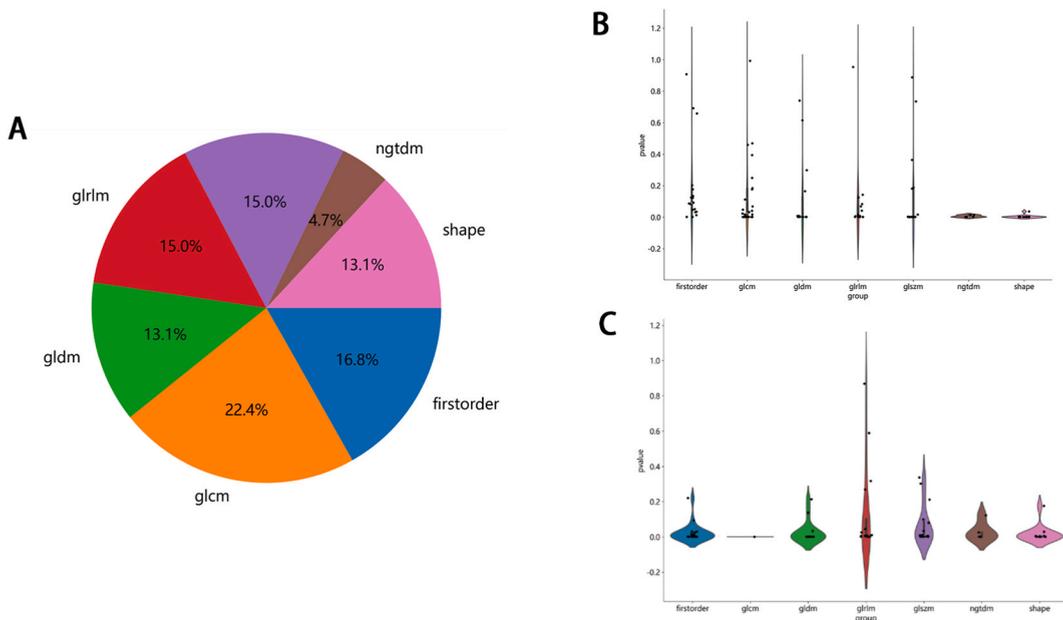


Fig. 2. (A) Pie chart of feature distribution extracted from traditional omics model; (B) A violin diagram of the output distribution of traditional omics features(features extracted from fundus photography);(C) A violin diagram of the output distribution of traditional omics features(features extracted from IR images).

2.5. Feature extraction of the fusion model

The features extracted from the Resnet101 model were fused with those extracted from the traditional omics model. Feature fusion [12] is the process of fusing two feature vectors into one feature vector. Feature fusion improves image information, which makes the fused feature vector more discriminative. Two commonly used feature level fusion methods include concat and add. The concat method increases the dimensions that describe image features but does not change the amount of information for each feature (Fig. 3). Comparatively, the add method increases the information for each image feature information but does not change the dimensions of each feature. Our study used the concat method to average the output probabilities of the traditional omics and DL features.

2.6. LASSO model construction

Computer-generated randomized sampling assigned 80 % and 20 % of the images to the training and test sets, respectively. A variable selection method proposed by statistician Robert Tibshirani in 1996 -LASSO [13]. LASSO regression removes any unimportant features by penalizing the regression coefficients of relevant parameter sizes, which also functions as an additional feature screening method. LASSO works by constructing a first-order penalty function and determining the coefficient of the selected variables for feature screening, which creates a refined model. Ultimately, LASSO regression shrinks the coefficient estimate to zero, depending on the extent of the additional parameter, λ . To determine the best value for λ , a five-fold cross validation was performed, and λ was chosen based on the optimal criterion [14].

2.7. Prediction model building

Each feature group was individually normalized with a z-score to combine the features of different magnitudes into one value. Non-zero coefficients were used as useful predictors in each feature group, and LASSO regression was used for feature selection in the training cohort. The machine learning classifiers, GradientBoosting and LightGBM, were used for predictive classification.

2.8. Visualization

We generated heat maps using the grad-CAM technique to visualize neural network outputs and identify which image region was likely involved in the classification of optic neuropathy [15]. Involved regions were highlighted in the heat maps to indicate the regions that contributed more significantly to the predictive outputs.

2.9. Statistical analysis

The chi-square test was used for qualitative data. A two-sided P-value <0.05 was considered statistically significant. All extracted features were dimensionally reduced by LASSO regression to improve the accuracy and fit of the modeling. The extracted features were then used to construct the prediction model with the machine learning methods, GradientBoosting and LightGBM. The performance of the study model was evaluated through AUC, accuracy, precision, sensitivity, and specificity measurements. Selected data were also presented through confusion matrices and heat maps.

3. Results

3.1. Characteristics of participants and images

A total of 156 patients (312 images) were included in the study. The numbers of patients with DON, GON, and ON were 55, 64, and 37, respectively. Each patient had a fundus photograph and an IR image. As shown in Table 1, there were 110, 128, and 74 images of DON, GON, and ON, respectively. The training and test sets had 248 (80 %) and 64 (20 %) images, respectively.

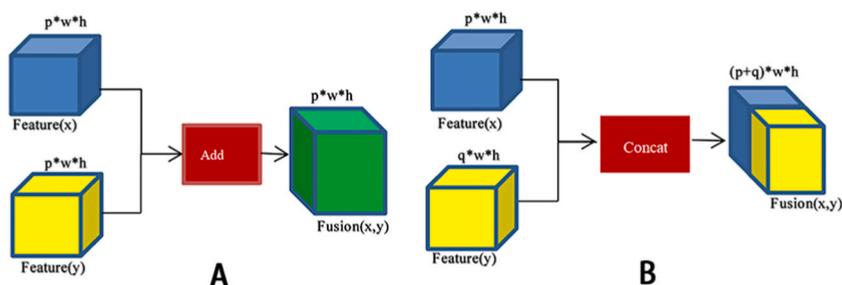


Fig. 3. Two feature fusion methods; (A) Add. (B) Concat.

Table 1
Baseline characteristics of participants and images in dataset.

Clinical characteristics	Participants, No.(%)		P value
Group	Train (n = 124)	Test(n = 32)	–
Age, median±SD(range), years	52.42 ± 16.69	50.91 ± 16.16	0.87
Male	63(50.80 %)	16(50 %)	0.94
Female	61(49.20 %)	16(50 %)	
Eye(Right)	66(53.23 %)	16(50 %)	0.75
Eye(Left)	58(46.77 %)	16(50 %)	
Images, No.	Fundus photography	IR images	–
Glaucoma optic neuropathy	64(21 %)	64(21 %)	–
Diabetes optic neuropathy	55(18 %)	55(18 %)	–
Optic neuritis	37(19 %)	37(19 %)	–

3.2. LASSO model and feature screening

LASSO regression was used to screen the features of the fusion model, and regularization was introduced to screen the weight of each feature. Each feature was assigned a coefficient that represented its weight size. The larger the penalty coefficient, λ , the smaller the weight. The optimal λ value identified for the fusion model was 0.0592 (Fig. 4).

3.3. Comparison of model predictive performance

3.3.1. Accuracy, precision, sensitivity, and specificity

We utilized the GradientBoosting classification model to measure accuracy. When multimodal data comprised the training sets in the traditional omics, Resnet101, and fusion models, the accuracy scores were 0.97, 0.98, and 0.99, respectively. When multimodal modal data comprised the test sets in the same models, the scores were 0.72, 0.87, and 0.88, respectively. We also utilized the LightGBM classification model to test accuracy. In the training set, the accuracy of the three models were 0.87, 0.87, and 0.89, respectively. In the test set, the accuracy of the three models were 0.81, 0.84, and 0.88, respectively. These data are shown in Table 2.

In the Fusion model, we calculated the precision, sensitivity and specificity of each class using the confusion matrix shown in Table 3.

The calculation formula was as follows [16].

A true positive (TP) result indicates that a patient with disease A was correctly detected as disease A. A true negative (TN) result indicates that a person with disease A was detected as non-disease A. A false positive (FP) result represents an incorrect data point, wherein a person with non-disease A was detected as positive for disease A. A false negative (FN) result represents an incorrect data point, wherein a person with disease A is detected as a person without disease A.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{Specificity} = \text{TN} / (\text{FP} + \text{TN})$$

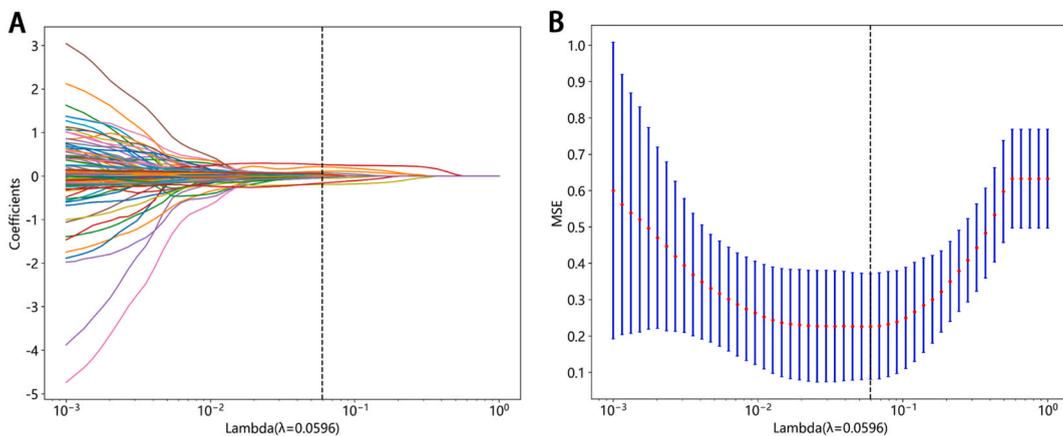


Fig. 4. Feature selection by LASSO model. (A) LASSO coefficient profiles of the 104 variables(Fusion model); (B) The curve of the binomial deviation with respect to the parameter λ , where the vertical dotted line at the right shows the optimal value of λ ($\lambda = 0.0596$)(Fusion model).

Table 2
The accuracy of the three models for the train and test set.

Model	Type of data	Task	Accuracy	
			GradientBoosting	LightGBM
Traditional omics model	Multimodal	label-train	0.97	0.87
		label-test	0.72	0.81
Resnet101 model	Multimodal	label-train	0.98	0.87
		label-test	0.87	0.84
Fusion model	Multimodal	label-train	0.99	0.89
		label-test	0.88	0.88

Table 3
The Precision, Sensitivity, and Specificity of each class in the Fusion model.

Class	Precision	Sensitivity	Specificity
0(GON)	0.92	0.92	0.95
1(DR)	0.85	1.00	1.00
2(ON)	0.83	0.63	0.89

3.3.2. Receiver operating characteristic (ROC) curves and AUC values

3.3.2.1. Multimodal data vs. single-mode data. We utilized the GradientBoosting classification method to classify our data. In the traditional omics model, the macro-average AUCs of the fundus photographs, IR images, and multimodal features of our data were

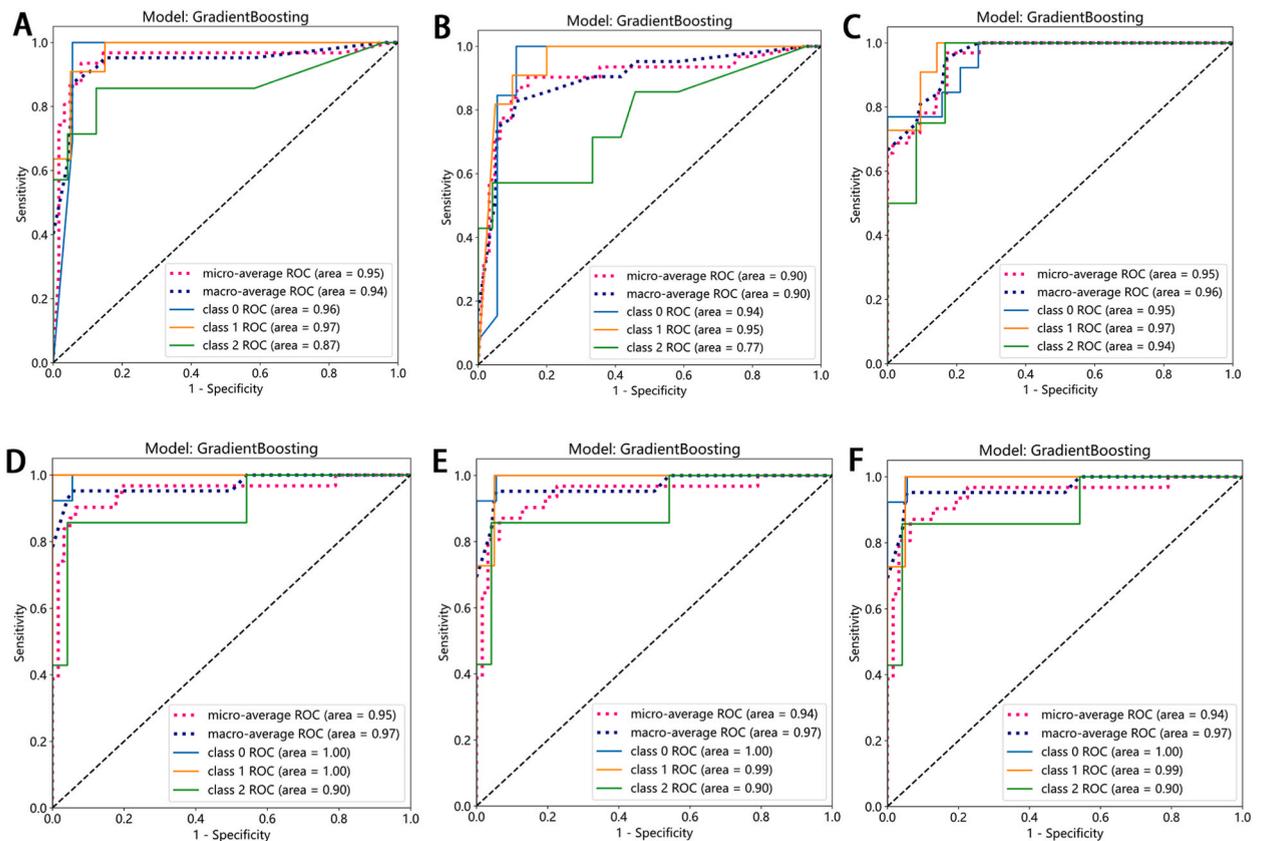


Fig. 5. Receiver operating characteristic curves (ROC) and area under curve (AUC). Performance comparison of single mode and multimodal data in the same omics model. (A)(B) (C) Traditional omics model; (D)(E) (F) Resnet101 model; (A)(D) Fundus photography was used as a single mode data input; (B)(E) IR images were used as a single mode data input; (C)(F) Fundus photography and IR images were used as multimodal data input. The ROC curves in each figure represent the glaucomatous optic neuropathy (blue), diabetic optic neuropathy (orange), optic neuritis (green). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

0.94, 0.90, and 0.96, respectively. Comparatively, the AUCs of the same variables in the Resnet101 model were all 0.97. These data are shown in Fig. 5.

3.3.2.2. Multiple omics model the traditional omics vs. Resnet101 vs. fusion model. We also utilized the GradientBoosting classification method in the fusion model. The AUCs for GON, DON, and ON in the fusion model were 0.99, 1.00, and 0.96 respectively. Comparatively, the scores for each of these diagnoses in the traditional omics model were 0.95, 0.97, and 0.94, respectively. The corresponding scores in the Resnet101 model were 1.00, 0.99, and 0.90, respectively.

When the LightGBM classification was used as a classifier, the AUCs for GON, DON, and ON in the fusion model were 0.98, 0.99, and 0.94, respectively. In comparison, the AUCs for the same conditions in the traditional omics model were 0.96, 0.95, and 0.94, respectively. The AUCs for the same conditions in the Resnet101 model were 1.00, 0.95, and 0.92, respectively. These data are depicted in Fig. 6.

3.3.3. Confusion matrix

We distributed the classification results of the test set into a 6 × 6 confusion matrix (Fig. 7). When GradientBoosting was used for predictive classification, the fusion model demonstrated better recognition performance than the traditional omics model. Specifically, the fusion model identified 28 correct images, whereas the traditional omics and Resnet101 models only identified 23 and 27 correct images, respectively. The total number of images in the test set was 32. Comparatively, when LightGBM was used as the classification method, the traditional omics, Resnet101, and fusion models identified 26, 26, and 28 correct images, respectively.

3.4. Heat map

As seen in the heat maps, the central region of each image was the primary ROI (Fig. 8).

4. Discussion

Our study demonstrated that multimodal data and a multi-omics DL system based on fundus photography and IR images could accurately diagnose DON, GON, and ON. These performed better than algorithms with a single-omics approach.

AI algorithms for glaucomatous optic neuropathy have been studied extensively. The AUC of DL algorithms using color fundus

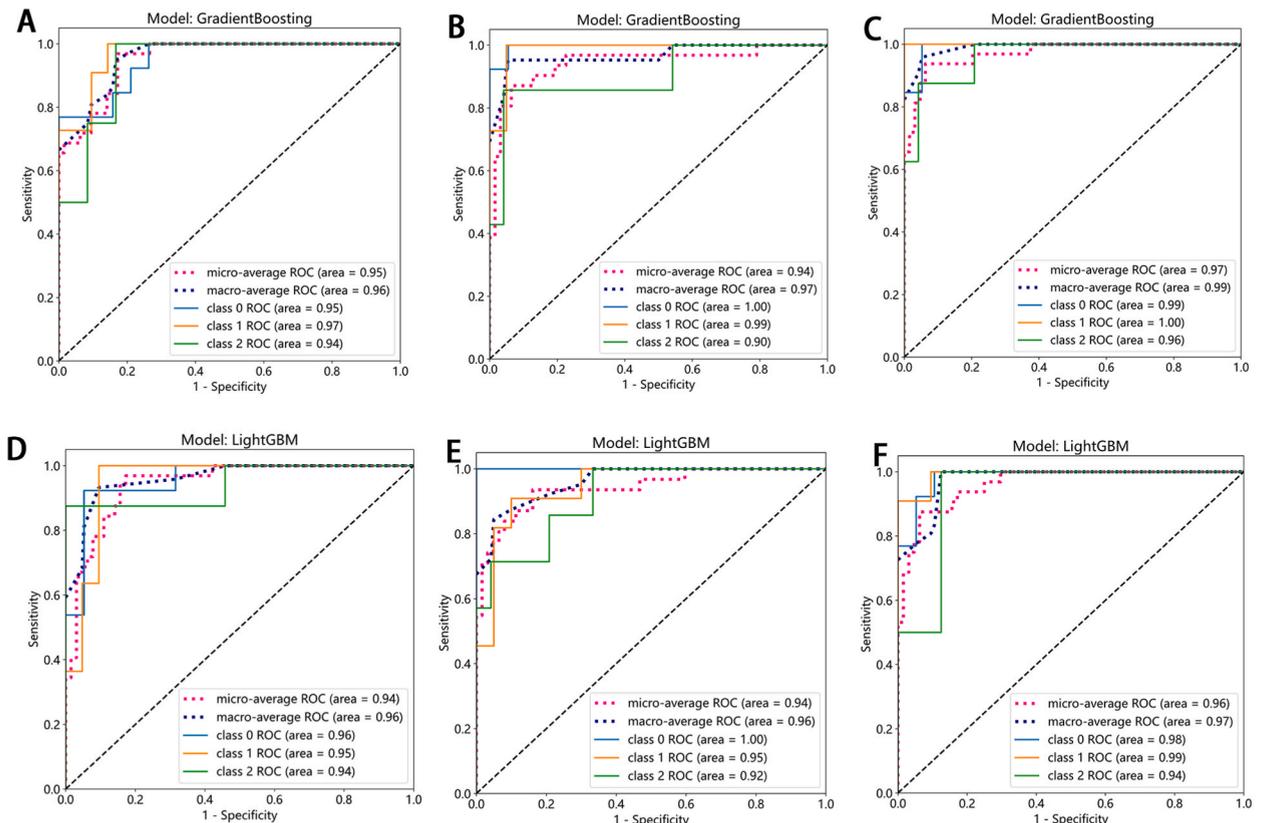


Fig. 6. The AUCs of multimodal data in different omics models. (A)(D) Traditional omics model; (B)(E) Resnet101 model; (C)(F) Fusion model.

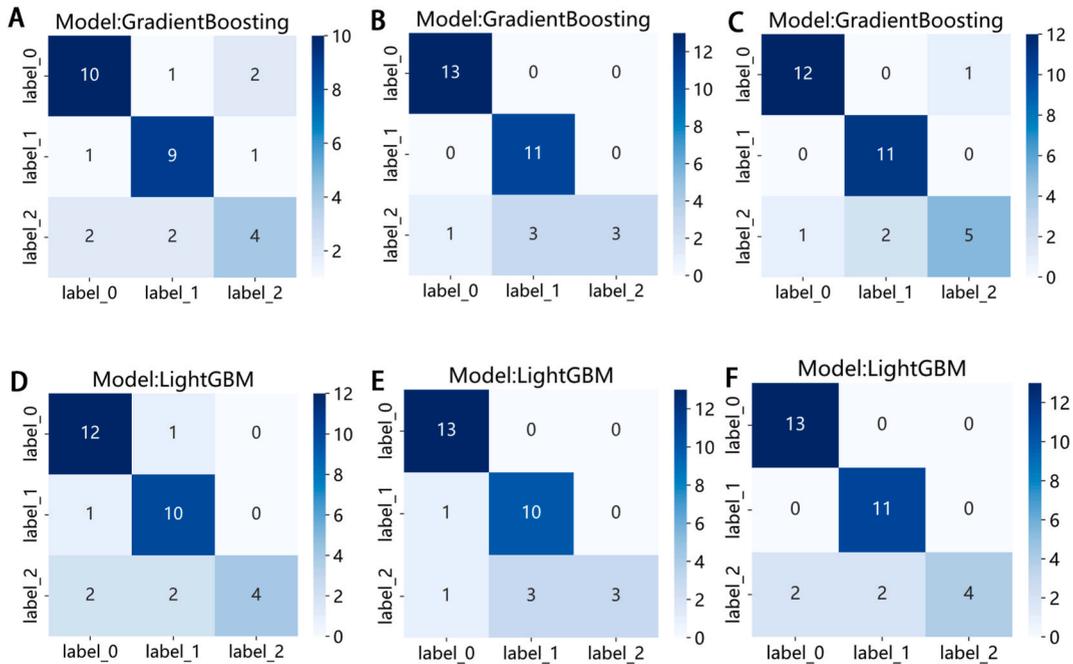


Fig. 7. The Confusion matrix of multimodal data in different omics models. “label = 0 ” means “glaucomatous optic neuropathy”; “label = 1 ” means “diabetic optic neuropathy”; “label = 2 ” means “optic neuritis”; (A)(B) (C) When Gradient Boosting was used as the prediction classification model in the test set, the number of correct predictions by traditional omics model, Resnet101 model and Fusion model were 23、 27 and 28 respectively; (D)(E) (F) When LightGBM was used as the prediction classification model in the test set, the number of correct predictions by the three model were 26、 26 and 28 respectively.

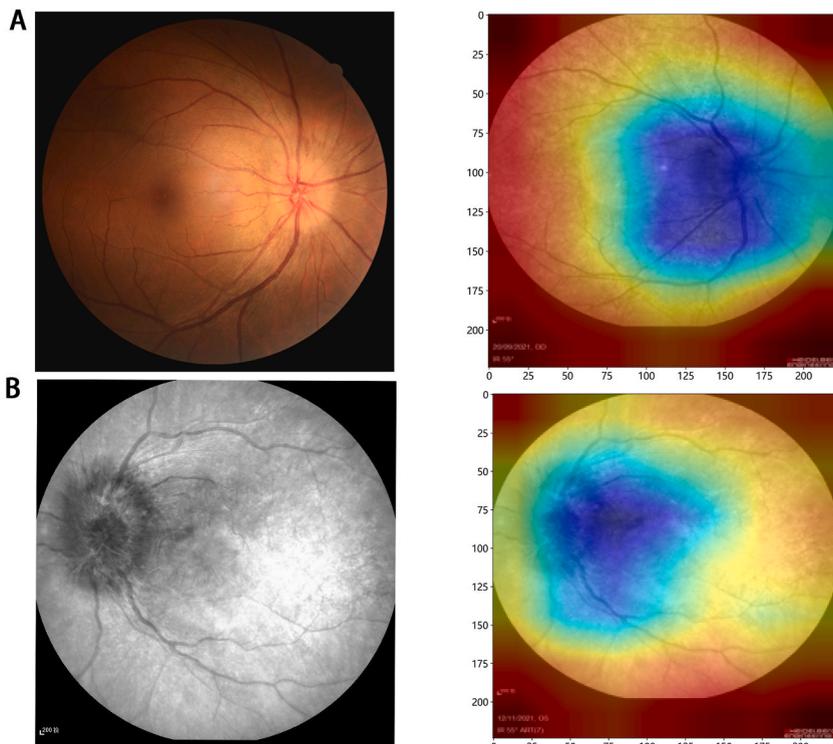


Fig. 8. (A)Fundus photography and its heat map;(B) IR image and its heat map; The blue area is the synaptic focus of ROI. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

photography reached 0.98 in a previous study [17], and the AUC of a DL algorithm based on spectral-domain optical coherence tomography (OCT) images of the optic disc reached 0.97 [18]. OCT-angiography also adds additional depth to the understanding of optic neuropathy by providing *in vivo* data on the peripheral capillaries and macular retinal plexus of the eye, the AUC for diagnosing glaucomatous optic neuropathy was 0.94 [19]. Kihara et al. developed a multi-modal diagnostic system based on the fusion of IR optic disc images and visual field data [20]. Their study demonstrated that a multi-modal system was significantly superior to single-mode systems. Kihara's DL system, which integrated OCT and visual field images, also showed excellent diagnostic performance, with an AUC of 0.90–0.95 [21]; [22]. The development of multimodal algorithms is a developing trend.

AI systems have also demonstrated excellent performance in distinguishing between retinal diseases. A DL system based on colored fundus photographs can distinguish more than 10 types of retinal diseases, including diabetic retinopathy, macular hole, and hypertensive retinopathy [23]. Another DL system that is also based on colored fundus images can distinguish between diabetic retinopathy and GON [24]. Other DL systems can also be structured to differentiate between other types of optic neuropathy, such as optic neuritis, DON, and ischemic optic neuropathy [25]. However, it is important to note that these existing studies utilize single-mode data. The use of multimodal data in DL systems still requires future examination.

Systems that use multimodal data seem superior because they can utilize early, intermediate, and late fusion to process multiple data features in different modes. This provides multi-dimensional perspectives for feature extraction, which can improve model performance [26].

More specifically, early fusion connects the data features in the different modes as input features directly. These are then entered into the first layer of the DL architecture for learning analysis [27]. The major limitation of early fusion is that the resulting neural networks cannot distinguish the features from different modes. Comparatively, intermediate fusion inputs the data of each mode into the neural network for learning, learns the correlation within the modes, and then fuses the features of the different modes in the middle layer of the neural network for further learning and to make an output prediction [28]. Late fusion trains and averages the data in each modality separately [29].

Kihara et al. and Huang et al. [29] utilized the late fusion method to analyze IR images of patients with glaucoma. Both authors fused two separate optic nerve OCT images into one output. The resulting neural network demonstrated good predictive performance. In a related study, Huang et al. utilized early fusion to analyze multiple magnetic resonance imaging sequences in an attempt to classify demyelinating diseases [30]. The same author also utilized early fusion to analyze fundus photos and visual field data [31]. The AUC for detecting GON in the latter study was 0.94, which was higher than the AUC for single-mode data-based models (fundus photos AUC, 0.90; visual field AUC, 0.89). Our study also utilized early fusion, which can simultaneously learn cross-modal and intra-modal correlation at low levels of abstraction. While the Resnet101 model performed better than the single-mode traditional omics model, the difference was not significant. This may be related to the way data features were extracted by the Resnet101 model. Early fusion-based models are also incapable of learning the marginal features of different data [32]. Future research can be conducted to compare the performance of different fusion methods.

Graphic feature extraction was the most important difference between the different omics models. Extraction of first-order, second-order and higher-order features in the traditional omics model relied on ROI segmentation. First-order features [7] mainly describe overall intensity and change, whereas second-order features [5] demonstrated the relationship between voxels. A CNN relies on neural networks to extract image information. As such, it can extract different features from a single image. This eliminates the need for image segmentation, which typically reduces feature accuracy [33]. At the same time, a local filter can slide over the input space in the CNN to preserve local image correlation of the image, as well as share the filter. This reduces the number of weights. Finally, the design of the neural network can be modified as needed, which provides flexibility. Our study also confirmed that deep learning models can diagnose different diseases with higher accuracy than a traditional omics model.

Since each omics platform has specific limitations, an integrated approach based on multiple omics may yield more coherent features [34]. Our study compares 1) the predictive performance of the traditional omics and DL models in extracting multi-modal data features and 2) the predictive performance of the fusion model with two extraction features. The results of our study showed that the predictive performance of the fusion model is better than the two previous models. This is consistent with existing literature, which states that integrated models achieve better model performance than single-data models [35,36].

Our study has some limitations. Our study population was small and recruited from the same hospital. As such, our sample may not be representative of the general population. Future research should consider joint multi-center studies. More multi-modal data, such as optic nerve fiber layer thickness, and more data sources, such as OCT or visual field perimeters, may also increase the accuracy of future multi-modal fusion systems, improve model performance, and provide simpler and more effective screening tools for clinics.

5. Conclusion

The use of multimodal and multi-omics-based strategies can improve the accuracy of screening for DON, GON, and ON.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Review Committee of First Affiliated Hospital of Nanchang University (Ethical number : IIT2023281) .

Funding

This work was supported by National Natural Science Foundation of China [grant numbers 82260211 , 81460092]; The Central Government Guides Local Science and Technology Development Foundation [grant number 20211ZDG02003]; The Key research and development project in Jiangxi Province [grant number 20203BBG73058 and 20192BBGL70033]; The Chinese medicine science and technology project in Jiangxi province [grant number 2020A0166]; The funding sources: National Natural Science Foundation of China; Ministry of Science and Technology of the People's Republic of China; Science and Technology Department of Jiangxi Province, China.

Consent for publication

We statement that the details of any images, videos, recordings, etc can be published, and that the persons providing consent have been shown the article contents to be published.

Data availability statement

Data will be made available on request.

CRediT authorship contribution statement

Ye-ting Lin: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Qiong Zhou:** Conceptualization, Funding acquisition, Methodology, Project administration, Validation. **Jian Tan:** Data curation. **Yulin Tao:** Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Study GbaVICVLEGotGBOD, Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study, *Lancet Global health* 9 (2) (2021) e144–e160, [https://doi.org/10.1016/s2214-109x\(20\)30489-7](https://doi.org/10.1016/s2214-109x(20)30489-7). Feb.
- [2] L. Dai, L. Wu, H. Li, et al., A deep learning system for detecting diabetic retinopathy across the disease spectrum, *Nat. Commun.* 12 (1) (2021) 3242, <https://doi.org/10.1038/s41467-021-23458-5>. May 28.
- [3] K.E. Kim, J.M. Kim, J.E. Song, C. Kee, J.C. Han, S.H. Hyun, Development and validation of a deep learning system for diagnosing glaucoma using optical coherence tomography, *J. Clin. Med.* 9 (7) (2020), <https://doi.org/10.3390/jcm9072167>. Jul 9.
- [4] K.A. Thakoor, J. Yao, D. Bordbar, et al., A multimodal deep learning system to distinguish late stages of AMD and to compare expert vs. AI ocular biomarkers, *Sci. Rep.* 12 (1) (2022) 2585, <https://doi.org/10.1038/s41598-022-06273-w>. Feb 16.
- [5] P. Lambin, E. Ríos-Velázquez, R. Leijenaar, et al., Radiomics: extracting more information from medical images using advanced feature analysis, *Eur. J. Cancer* 48 (4) (2012) 441–446, <https://doi.org/10.1016/j.ejca.2011.11.036>. Mar.
- [6] V. Parekh, M.A. Jacobs, Radiomics: a new application from established techniques, *Expert review of precision medicine and drug development* 1 (2) (2016) 207–226, <https://doi.org/10.1080/23808993.2016.1164013>.
- [7] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures, they are data, *Radiology*. Feb 278 (2) (2016) 563–577, <https://doi.org/10.1148/radiol.2015151169>.
- [8] S.M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, M.K. Khan, Medical image analysis using convolutional neural networks: a review, *J. Med. Syst.* 42 (11) (2018) 226, <https://doi.org/10.1007/s10916-018-1088-1>. Oct 8.
- [9] K. He Xz, S. Ren, J Sun Deep Residual Learning for Image Recognition. *Computer Vision And Pattern Recognition (CVPR)*, 2016, pp. 770–778.
- [10] K.M. He, X.Y. Zhang, S.Q. Ren, J. Sun, Ieee. Deep Residual Learning for Image Recognition, 2016, pp. 770–778, <https://doi.org/10.1109/cvpr.2016.90>. Jun 27–30.
- [11] L.Q. Zhou, X.L. Wu, S.Y. Huang, et al., Lymph node metastasis prediction from primary breast cancer US images using deep learning, *Radiology* 294 (1) (2020) 19–28, <https://doi.org/10.1148/radiol.2019190372>. Jan.
- [12] J. Liu, W. Sun, X. Zhao, J. Zhao, Z. Jiang, Deep feature fusion classification network (DFFCNet): towards accurate diagnosis of COVID-19 using chest X-rays images, *Jul, Biomed. Signal Process Control* 76 (2022), 103677, <https://doi.org/10.1016/j.bspc.2022.103677>.
- [13] R. Tibshirani, Regression shrinkage and selection via the lasso, *J. Roy. Stat. Soc. B* 58 (1) (1996) 267–288, <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>.
- [14] J. Kang, Y.J. Choi, I.K. Kim, et al., LASSO-based machine learning algorithm for prediction of lymph node metastasis in T1 colorectal cancer, *Cancer Res. Treatment* 53 (3) (2021) 773–783, <https://doi.org/10.4143/crt.2020.974>. Jul.
- [15] H. Jiang, J. Xu, R. Shi, et al., A multi-label deep learning model with interpretable grad-CAM for diabetic retinopathy classification, in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference 2020, 2020, pp. 1560–1563, <https://doi.org/10.1109/embc44109.2020.9175884>.
- [16] H. Panwar, P.K. Gupta, M.K. Siddiqui, R. Morales-Menendez, P. Bhardwaj, V. Singh, A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images, *Chaos, solitons, and fractals* 140 (2020), 110190, <https://doi.org/10.1016/j.chaos.2020.110190>. Nov.
- [17] Z. Li, C. Guo, D. Lin, et al., Deep learning for automated glaucomatous optic neuropathy detection from ultra-widefield fundus images, *The British journal of ophthalmology* 105 (11) (2021) 1548–1554, <https://doi.org/10.1136/bjophthalmol-2020-317327>. Nov.
- [18] A.R. Ran, C.Y. Cheung, X. Wang, et al., Detection of glaucomatous optic neuropathy with spectral-domain optical coherence tomography: a retrospective training and validation deep-learning analysis, *The Lancet Digital health* 1 (4) (2019) e172–e182, [https://doi.org/10.1016/s2589-7500\(19\)30085-8](https://doi.org/10.1016/s2589-7500(19)30085-8). Aug.

- [19] R. Bunod, M. Lubrano, A. Pirovano, et al., A deep learning system using optical coherence tomography angiography to detect glaucoma and anterior ischemic optic neuropathy, *J. Clin. Med.* (2) (2023) 12, <https://doi.org/10.3390/jcm12020507>. Jan 7.
- [20] Y. Kihara, G. Montesano, A. Chen, et al., Policy-driven, multimodal deep learning for predicting visual fields from the optic disc and OCT imaging, *Ophthalmology* 129 (7) (2022) 781–791, <https://doi.org/10.1016/j.ophtha.2022.02.017>. Jul.
- [21] J. Xiong, F. Li, D. Song, et al., Multimodal machine learning using visual fields and peripapillary circular OCT scans in detection of glaucomatous optic neuropathy, *Ophthalmology* 129 (2) (2022) 171–180, <https://doi.org/10.1016/j.ophtha.2021.07.032>. Feb.
- [22] S. Yi, G. Zhang, C. Qian, Y. Lu, H. Zhong, J. He, A multimodal classification architecture for the severity diagnosis of glaucoma based on deep learning, *Front. Neurosci.* 16 (2022), 939472, <https://doi.org/10.3389/fnins.2022.939472>.
- [23] L. Dong, W. He, R. Zhang, et al., Artificial intelligence for screening of multiple retinal and optic nerve diseases, *JAMA Netw. Open* 5 (5) (May 2022), e229960, <https://doi.org/10.1001/jamanetworkopen.2022.9960>.
- [24] S. Keel, J. Wu, P.Y. Lee, J. Scheetz, M. He, Visualizing deep learning models for the detection of referable diabetic retinopathy and glaucoma, *JAMA ophthalmology* 137 (3) (2019) 288–292, <https://doi.org/10.1001/jamaophthalmol.2018.6035>. Mar 1.
- [25] J.M. Ahn, S. Kim, K.S. Ahn, S.H. Cho, U.S. Kim, Accuracy of machine learning for differentiation between optic neuropathies and pseudopapilledema, *BMC Ophthalmol.* 19 (1) (2019) 178, <https://doi.org/10.1186/s12886-019-1184-0>. Aug 9.
- [26] A. Ma'ayan, Complex systems biology, *J. R. Soc., Interface* 14 (134) (2017), <https://doi.org/10.1098/rsif.2017.0391>. Sep.
- [27] T. Baltrusaitis, C. Ahuja, L.P. Morency, Multimodal Machine Learning: a Survey and Taxonomy, *IEEE transactions on pattern analysis and machine intelligence* 41 (2) (2019) 423–443, <https://doi.org/10.1109/tpami.2018.2798607>. Feb.
- [28] Z. Huang, X. Zhan, S. Xiang, et al., SALMON: survival analysis learning with multi-omics neural networks on breast cancer, *Front. Genet.* 10 (2019) 166, <https://doi.org/10.3389/fgene.2019.00166>.
- [29] S.C. Huang, A. Pareek, R. Zamanian, I. Banerjee, M.P. Lungren, Multimodal fusion with deep neural networks for leveraging CT imaging and electronic health record: a case-study in pulmonary embolism detection, *Sci. Rep.* 10 (1) (2020), 22147, <https://doi.org/10.1038/s41598-020-78888-w>. Dec 17.
- [30] C. Huang, W. Chen, B. Liu, et al., Transformer-based deep-learning algorithm for discriminating demyelinating diseases of the central nervous system with neuroimaging, *Front. Immunol.* 13 (2022), 897959, <https://doi.org/10.3389/fimmu.2022.897959>.
- [31] X. Huang, J. Sun, K. Gupta, et al., Detecting glaucoma from multi-modal data using probabilistic deep learning, *Front. Med.* 9 (2022), 923096, <https://doi.org/10.3389/fmed.2022.923096>.
- [32] S.R. Stahlschmidt, B. Ulfenborg, J. Synnergren, Multimodal deep learning for biomedical data fusion: a review, *Briefings in bioinformatics* (2) (2022) 23, <https://doi.org/10.1093/bib/bbab569>. Mar 10.
- [33] Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature.* 521 (7553) (2015) 436–444, <https://doi.org/10.1038/nature14539>. May 28.
- [34] M.D. Ritchie, E.R. Holzinger, R. Li, S.A. Pendergrass, D. Kim, Methods of integrating data to uncover genotype-phenotype interactions, *Nature reviews Genetics.* Feb 16 (2) (2015) 85–97, <https://doi.org/10.1038/nrg3868>.
- [35] O.B. Poirion, Z. Jing, K. Chaudhary, S. Huang, L.X. Garmire, DeepProg: an ensemble of deep-learning and machine-learning models for prognosis prediction using multi-omics data, *Genome Med.* 13 (1) (2021) 112, <https://doi.org/10.1186/s13073-021-00930-x>. Jul 14.
- [36] H. Chai, X. Zhou, Z. Zhang, J. Rao, H. Zhao, Y. Yang, Integrating multi-omics data through deep learning for accurate cancer prognosis prediction, *Comput. Biol. Med.* 134 (2021), 104481, <https://doi.org/10.1016/j.combiomed.2021.104481>. Jul.