Methodological Article

Isfahan Artificial Intelligence Event 2023: Macular Pathology Detection Competition

Abstract

Background: Computer-aided diagnosis (CAD) methods have become of great interest for diagnosing macular diseases over the past few decades. Artificial intelligence (AI)-based CADs offer several benefits, including speed, objectivity, and thoroughness. They are utilized as an assistance system in various ways, such as highlighting relevant disease indicators to doctors, providing diagnosis suggestions, and presenting similar past cases for comparison. **Methods:** Much specifically, retinal AI-CADs have been developed to assist ophthalmologists in analyzing optical coherence tomography (OCT) images and making retinal diagnostics simpler and more accurate than before. Retinal AI-CAD technology could provide a new insight for the health care of humans who do not have access to a specialist doctor. AI-based classification methods are critical tools in developing improved retinal AI-CAD technology. The Isfahan AI-2023 challenge has organized a competition to provide objective formal evaluations of alternative tools in this area. In this study, we describe the challenge and those methods that had the most successful algorithms. **Results:** A dataset of OCT images, acquired from normal subjects, patients with diabetic macular edema, and patients with other macular disorders, was provided in a documented format. The dataset, including the labeled training set and unlabeled test set, was made accessible to the participants. The aim of this challenge was to maximize the performance measures for the test labels. Researchers tested their algorithms and competed for the best classification results. **Conclusions:** The competition is organized to evaluate the current AI-based classification methods in macular pathology detection. We received several submissions to our posted datasets that indicate the growing interest in AI-CAD technology. The results demonstrated that deep learning-based methods can learn essential features of pathologic images, but much care has to be taken in choosing and adapting appropriate models for imbalanced small datasets.

Keywords: *Age-related macular degeneration, choroidal neovascularization, diabetic macular edema, Isfahan artificial intelligence challenge, macular hole, optical coherence tomography*

Submitted: 19-Jul-2024 **Revised:** 11-Aug-2024 **Accepted:** 11-Aug-2024 **Published:** 23-Jan-2025

Farnaz Sedighin[1], Maryam Monemian[1], Zahra Zojaji[2], Ahmadreza Montazerolghaem[2], Mohammad Amin Asadinia[3], Seyed Mojtaba Mirghaderi[4], Seyed Amin Naji Esfahani[5], Mohammad Kazemi[3], Reza Mokhtari[6], Maryam Mohammadi[6], Mohadese Ramezani[6], Mahnoosh Tajmirriahi[1], Hossein Rabbani[1]

[1]Medical Image and Signal Processing Research Center, Department of Bioelectrics and Biomedical Engineering, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Iran, [2]Faculty of Computer Engineering, University of Isfahan, Isfahan, Iran, [3]Department of Electrical Engineering, University of Isfahan, Isfahan, Iran, [4]Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran, [5]Electrical Engineering Faculty, K. N. Toosi University, Tehran, Iran, [6]Department of Mathematical Sciences, Isfahan University of Technology, Isfahan, Iran

Address for correspondence:
Dr. Mahnoosh Tajmirriahi, Medical Image and Signal Processing Research Center, Department of Bioelectrics and Biomedical Engineering, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan - 81746734641, Iran.
E-mail: mata.riahi@yahoo.com

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.jmss_47_24

Quick Response Code:

Introduction

Optical coherence tomography (OCT) imaging technique is utilized to obtain high-resolution cross-sectional tomographic images of the retina. Various ocular and systemic diseases can be detected and monitored through the use of retinal OCT images. Common progressive disorders such as age-related macular degeneration (AMD), diabetic macular edema (DME), choroidal neovascularization (CNV), and macular hole (MH) can be identified and followed up through OCT images.[1-3] The International Agency for the Prevention of Blindness report highlights AMD as the third most common cause of visual impairment and irreversible blindness among individuals aged 50 years and above.[4] Shockingly, 8.1 million individuals worldwide suffer from untreated AMD-related vision loss. DME is a significant contributor to vision impairment in young adults in developed nations, stemming from diabetic retinopathy in individuals with diabetes. It stands as the primary cause of blindness in people under 50 years old. Unfortunately, these conditions are occasionally misidentified or detected belatedly, resulting in lasting and irreversible vision impairment. Consequently, the associated expenses can be substantial. Early detection, screening, and timely intervention can help slow down the progression of these macular

How to cite this article: Sedighin F, Monemian M, Zojaji Z, Montazerolghaem A, Asadinia MA, Mirghaderi SM, *et al*. Isfahan Artificial Intelligence Event 2023: Macular Pathology Detection Competition. J Med Signals Sens 2025;15:3.

diseases. Therefore, screening for these disorders holds great importance; however, it is frequently overlooked in developing countries due to its high cost. Conversely, the increasing prevalence of such macular pathologies highlights the need for a comprehensive vision screening program for all individuals. Utilizing CADs which were developed with advanced technology could enhance the accessibility of the screening.

In clinical settings, the presence of pathologic morphological and structural changes in the retinal layers, known as biomarkers, can offer crucial insights into the progression of a disease.[5] However, the identification and interpretation of these biomarkers can be time-consuming and challenging due to their abundance, size, shape, and complexity, as well as the vast amount of data present in OCT imaging. Therefore, this greatly depends on the clinicians' expertise. Consequently, the automated classification of OCT images has emerged as a prominent area of interest in recent years to develop automatic screening tools.

New technologies utilizing artificial intelligence (AI) through machine learning (ML) techniques are expected to be advantageous in terms of early detection and suitable treatment, potentially averting the development of severe ocular conditions. In recent years, ML has found extensive application in the analysis of ophthalmic images. The application of ML in macular pathology detection can be categorized into two main areas. The first category involves detecting and grading diseases, aiming to classify the disease identified in an image or a series of images. Other studies go beyond simple detection by employing multiclass classification to assess the severity level of a particular pathology. Image analysis using ML systems requires extracting features that depict the input data in a manner appropriate for classification. Features can be manually crafted, such as by an algorithm developer, extracted using preset filter banks, or even learned during training to execute task-specific processing of the input data. Automatic feature extraction has gained prominence in recent times. The employment of convolutional neural networks (CNNs) for automatic feature extraction has emerged as the favored method in OCT image analysis.[6] However, there are several issues regarding these methods. They usually need large, diverse, and well-balanced training datasets to prevent overfitting and provide the possibility for generalization. In addition, OCT images acquired from imaging devices usually have various qualities and the automatic screening tools must have effective performance independent of capturing device. Therefore, studies are still ongoing to find an efficient ML algorithm for the development of accurate AI-CADs to assist ophthalmologists in analyzing OCT images and making retinal diagnostics simpler and more accurate than before.

In this regard, the macular pathology detection competition was hosted by the Isfahan National Elite Foundation in Isfahan AI (IAI-2023) events and aimed to foster the development of retinal AI-CAD technology through the fair evaluation of various AI-based techniques. In this competition, a dataset of labeled OCT images ("training set") was provided in a well-documented format for the competitors. Researchers managed to use the training data for the development of their methods and then reported their algorithm outputs for evaluation using the test data which were later provided to them. The test data labels were kept unrevealed, guaranteeing an impartial evaluation of performance that was not biased by the selection of methods or parameter choices customized for the data.

Ranking of Competition Results

The outcomes of submitted contests should not be embraced promptly since they might not provide a fully impartial assessment of excellence due to diverse factors. To conduct an impartial assessment, the final phase of the competition took place live and was evaluated based on five criteria of equal importance. These criteria were used to determine the rankings of the three finalist teams.

The five equally weighted criteria were as follows:

1. Innovation of the proposed approach
2. Performance in the first stage competition (based on the best submission and the score)
3. Performance in the last submission (second phase held live at Abbasi Hotel, Isfahan, Iran) on the dataset.
4. Quality and clarity of the final report
5. Quality and clarity of the presentation.

Each criterion was scored on a scale of 1, 2, or 3. The team that performed the best in each criterion was received 3 points and the second-best team received 2 points. The final rankings were determined by the highest total points awarded across all five criteria during the judge deliberations at the end of the competition.

Performance Metrics

The utilized performance metrics include accuracy, precision, recall, F-measure, and macro-f. Accuracy measurement evaluates the accuracy of predictions generated by the model throughout the test dataset.[7] It is determined by dividing the sum of true positives (TP) and true negatives by the total sample size. Precision is determined by the number of correct positive predictions compared to all positive predictions made by the model. This is computed by dividing the number of TPs by the sum of TPs and false positives. Recall, also referred to as sensitivity or TP rate, quantifies the percentage of accurate positive predictions out of all the actual positive instances. It is computed by dividing the number of TPs by the sum of TP and false negatives. The F-measure also called the F1 score is a measurement index that balances precision and recall. It is computed as the harmonic mean of precision and recall. The F1 score proves to be valuable when aiming for a trade-off between high precision and high recall, as it penalizes extreme negative values of either component. The model's predictions are evaluated based on accuracy, which measures the overall correctness. On the other hand, precision and recall assess the quality of positive and negative predictions, respectively. To provide a more comprehensive evaluation of classification models, the F1 score strikes a balance between precision and recall.

The macroaveraged F1 score, also known as the macro-f score, is calculated by finding the arithmetic mean of all the F1 scores for each class. This approach treats all classes equally, regardless of their support values. In an imbalanced dataset where all classes are considered equally important, choosing the macroaverage is a wise decision because it gives equal weight to each class.

In the following, we will detail the dataset of the competition in Section II. The winning teams and their competing methods are also explained in this section, briefly. In Section III, we will conclude the outcomes of the companioning and remark future work.

Description of the Dataset

The dataset provided for this competition was obtained from a costume-made swept-source OCT (SS-OCT) imaging system, which was designed and developed at the Department of Biomedical Engineering, University of Basel.[8,9] The capturing device used in this system has a central wavelength of 1064 nm, a spectral bandwidth of 100 nm, and an A-scan rate of 100 kHz. The data acquisition took place at the Didavaran Eye Clinic in Isfahan, Iran.

The dataset includes OCT images of a total of 191 participants, categorized into three groups: 73 individuals with DME or class number "1," 54 healthy individuals or class number "0," and 64 non-DME patients or class number "2" with AMD, CNV, and MH disorders. Each participant contributed 300 B-scans, with a resolution ranging from 300–1200 × 300 pixels.

Due to the low quality of raw data,[10] the pre-processing including quality assessment, contrast enhancement, noise removal, normalization, and alignment was performed on the images. Finally, the number of B-scans in the dataset varied between 30 to 300 preprocessed images for each subject.[11] We provided sample B-scans of this dataset in Figure 1 for visual inspection.

The overall data were divided into two categories, training and testing, with a ratio about 70/30, and the training data were provided to the teams with corresponding labels. The number of training and test datasets is reported in Table 1.

In the final phase of the competition, test data belonged to 10 individuals whose data had different preprocessing steps [Table 2]. This helped us evaluate the robustness of the participant algorithms against noise, low contrast, and misaligned datasets. Each individual's dataset comprised up to 300 B-scans with a resolution of 300–1200 × 300 pixels

Results and Discussion

In the first stage of the competition, we received five submissions for the test labels. Details of the average values of the evaluation metrics in the first stage evaluation are reported in Figure 2. It can be seen that the submissions managed to achieve more than 65% classification accuracy on the test set. Although excluding an outlier of only 31% accuracy (probably submitted with accidentally confused class labels), the average accuracy of all submissions was 67%.

The final phase of the competition took place live, and the final test data, as described in Table 2, was given to the participants. The teams who achieved ranks one to two in the final assessment and their proposed methods at a glance are as follows.

1) The first ranked team could achieve an accuracy of 80%. The team members were contributed from Khaje. N. Toosi University of Technology, Isfahan University of technology, and Isfahan University, Isfahan, Iran under the supervision of A. Najii Isfahani. They utilized OpticNet-71[12] with a modified convolutional structure. Original OpticNet-71 introduced an innovative CNN that excels in accurately detecting retinal diseases. The architecture incorporates several advancements, including a new residual unit that encompasses Atrous Separable Convolution, a unique building block, and a mechanism to prevent gradient degradation. Using a modification of OpticNet-71, modified residual units and recursive structure were employed in this first ranked method to improve the performance in both accuracy and runtime. In addition, this modification of OpticNet-71 is resistant to gradient fading and over fitting due to the use of a branch structure as well. The block diagram of the proposed method of the first ranked team is depicted in Figure 3.
2) The second ranked team could achieve an accuracy of 60%. The team members were contributed from from Isfahan University of Technology, Isfahan, Iran under the supervision of R. Mokhtari. They designed a deep learning-driven classification framework that harnesses the powerful capabilities of a transfer learning approach. The method was built upon the ResNet architecture.[13] A residual neural network, also known as ResNet, is a deep learning model that utilizes weight layers to learn residual functions based on the input of each layer. It functions similarly to a highway network, where gates are opened by strongly positive bias weights. This characteristic allows for the training of deep learning models with numerous layers, resulting in improved accuracy as the depth increases. The second-ranked team utilized Res-Net50 and changed the original

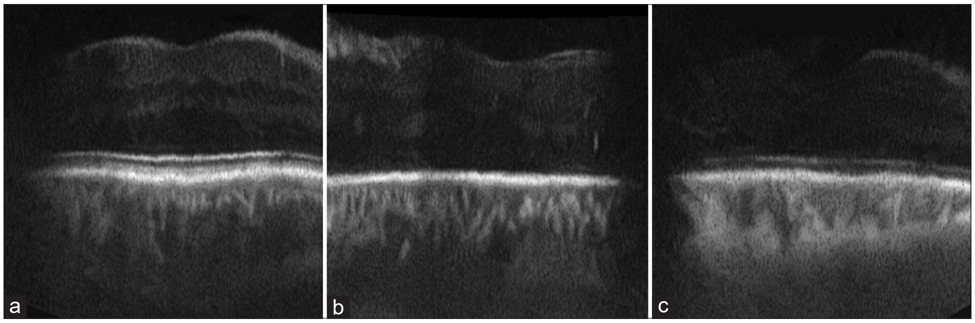

Figure 1: Sample B‑scans of the datasets of this competition. (a) Healthy "0", (b) diabetic macular edema (DME) "1", (c) non‑DME "2"

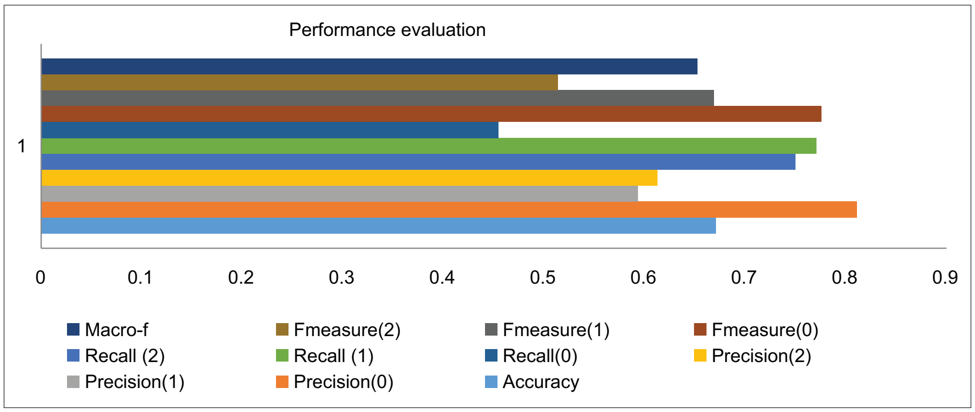

Figure 2: The average values of the measured evaluation metrics in the first stage evaluation. The numbers in the parentheses indicate the corresponding class number

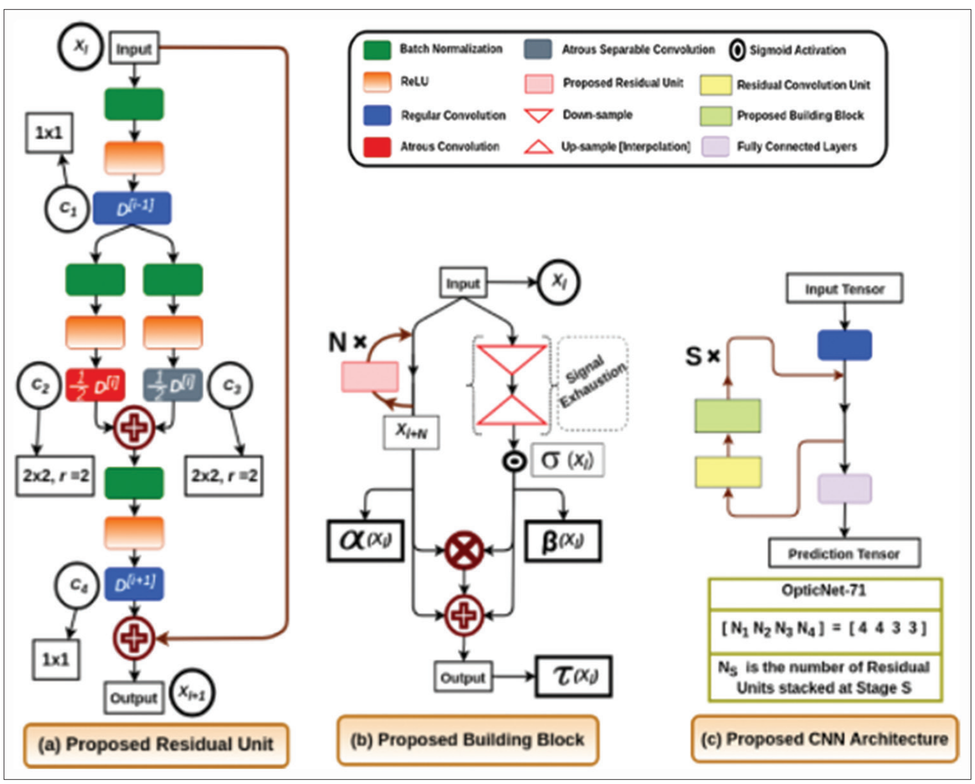

Figure 3: Block diagram of the model that acquired the first ranked team. (a) Proposed residual unit, (b) proposed building block, (c) proposed convolutional neural network architecture

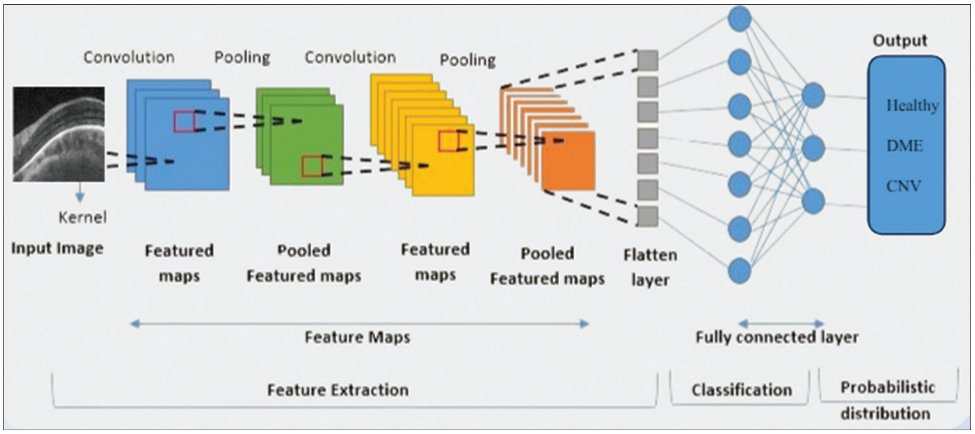

Figure 4: Block diagram of the model that acquired the second ranked team

Table 1: Dataset composition in each stage of competition

Dataset	Number of participants		
	Healthy - 0	DME - 1	Non-DME - 2
Training	35	28	23
Test 1	15	12	11
Final test	3	3	4

DME - Diabetic macular edema

Table 2: Number of final test data divided by different preprocessing steps

	Raw data	Denoised data	Enhanced data	Aligned data
Healthy - 0	1	1	1	0
DME - 1	0	1	1	1
Non-DME - 2	1	2	0	1

DME - Diabetic macular edema

Relu activation functions to Leaky-Relu to gain better performance than the original network. In addition, to address the inherent bias present in the dataset, implementing an oversampling methodology that leverages the principles of interpolation is suggested. They used this oversampling method to increase the number of data and make the datasets of the classes be more balanced. The block diagram of their proposed method is depicted in Figure 4.

Conclusion and Outlook

Looking at all the winning algorithms of IAI-2023 reveals several very interesting aspects. (1) All classification methods are based on deep learning-based algorithms. The most popular methods are CNNs[14] and Resnet blocks.[13] (2) In all methods, the recall of Class 2 has the lowest value which means that many of the nondiabetic patients are missed during classification. This may be caused due to imbalanced datasets and indicates that the methods could not overcome this challenge. (3) In all methods, the classification of Class 0 (healthy individuals) has achieved the highest performance, which may caused by the larger number of data in this class data than the others, creating a skewed distribution. Indeed, for each DME or Non-DME patient, only a small portion of B-scans are pathologic which leads to an intra-subject data imbalance as well. B-scan-based methods suffer from this issue and need to pay more attention to labeling.

Overall, the challenges of the current small and imbalanced training dataset were not properly addressed by the competitors. Additionally, key topics such as a thorough validation of the approaches concerning computational complexity and the feasibility of real-time, portable implementation were not addressed in this competition. This could be a new and ambitious objective of a future IAI competition.

The datasets and their descriptions are available on the Medical Image and Signal Processing webpage https://misp.mui.ac.ir/en/oct-basel-data-0. Other researchers interested in OCT analysis are welcome to test their algorithms on these datasets and to report their results. However, due to the current availability of the labels of the test data, future classification results of the competition data cannot fairly be compared to the original submissions.

Financial support and sponsorship

This work was supported by the Isfahan Elite Foundation (IEF), which sponsored the Isfahan Artificial Intelligence Event 2023 (IAI2023). The IEF organized the event and provided financial support for the 10 challenges, including Challenge X: Macular Pathology Detection. Several winners received prizes from the IEF.

Conflicts of interest

The authors declare the following potential conflicts of interest:

- ZZ and AM were the organizers of the Isfahan AI (IAI) 2023 competitions on behalf of Isfahan Elite Foundation (IEF), which included 10 challenges.

- FS, MM, MT, and HR served as scientific committee members for Challenge X: Macular Pathology Detection. They were responsible for evaluating the methodologies and results of all participant teams.
- MAA, SMM, SANE, MK, RM, MM and MR are members of the winning teams in this challenge.

None of the organizers and scientific committee members (FS, MM, MT, HR, ZZ and AM) contributed to the development of the methods used by the participating teams. The final decision regarding the winners was made by the policy council members based on the following criteria:

- Technical contribution in developed models by teams,
- The results on initial and final test data of each team,
- The submitted reports and teams' presentations.

The authors have disclosed these relationships to ensure transparency and maintain the integrity of the research.

Acknowledgements

We extend our appreciation to the policy council members, Prof. Behrouz Minaei Bidgoli from Iran University of Science and Technology, Prof. Mohammad Hassan Moradi from Amirkabir University of Technology, Prof. Heshaam Faili from University of Tehran, Prof. Emad Fatemizadeh from Sharif University of Technology, and Prof. Hossein Rabbani from Isfahan University of Medical Sciences, for making the final decisions regarding the winners of all 10 challenges.

Finally, we thank the anonymous reviewers for their valuable feedback and contributions to the improvement of this paper.

References

1. Hajizadeh F, Kafieh R, Tajmirriahi M. Introduction to Optical Coherence Tomography. In: Hajizadeh, F. (eds) Atlas of Ocular Optical Coherence Tomography. Springer, Cham; 2022. doi: 10.1007/978-3-031-07410-3_1.
2. Darooei R, Nazari M, Kafieh R, Rabbani H. Loss-modified transformer-based u-net for accurate segmentation of fluids in optical coherence tomography images of retinal diseases. J Med Signals Sens 2023;13:253-60.
3. Esmaeili M, Dehnavi AM, Rabbani H, Hajizadeh F. Three-dimensional segmentation of retinal cysts from spectral-domain optical coherence tomography images by the use of three-dimensional curvelet based K-SVD. J Med Signals Sens 2016;6:166-71.
4. Steinmetz JD, Bourne RR, Briant PS, Flaxman SR, Taylor HR, Jonas JB, *et al*. 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 Glob Health 2021;9:e144-60.
5. Hanson RL, Airody A, Sivaprasad S, Gale RP. Optical coherence tomography imaging biomarkers associated with neovascular age-related macular degeneration: A systematic review. Eye (Lond) 2023;37:2438-53.
6. Sarhan MH, Nasseri MA, Zapp D, Maier M, Lohmann CP, Navab N, *et al*. Machine learning techniques for ophthalmic data processing: A review. IEEE J Biomed Health Inform 2020;24:3338-50.
7. Hicks SA, Strümke I, Thambawita V, Hammou M, Riegler MA, Halvorsen P, *et al*. On evaluation metrics for medical applications of artificial intelligence. Sci Rep 2022;12:5979.
8. Tajmirriahi M, Amini Z, Hamidi A, Zam A, Rabbani H. Modeling of retinal optical coherence tomography based on stochastic differential equations: Application to denoising. IEEE Trans Med Imaging 2021;40:2129-41.
9. Tajmirriahi M, Amini Z, Rabbani H. Local Self-Similar Solution of ADMM for Denoising of Retinal OCT Images. IEEE Transactions on Instrumentation and Measurement; 2023.
10. Tajmirriahi M, Rostamian R, Amini Z, Hamidi A, Zam A, Rabbani H. Stochastic differential equations for automatic quality control of retinal optical coherence tomography images. Annu Int Conf IEEE Eng Med Biol Soc 2022;2022:3870-3.
11. Amini Z, Rabbani H. Optical coherence tomography image denoising using Gaussianization transform. Journal of Biomedical Optics 2017;22:086011-.
12. Kamran SA, Saha S, Sabbir AS, Tavakkoli A. Optic-net: A novel convolutional neural network for diagnosis of retinal diseases from optical tomography images. In: 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA). IEEE; 2019. p. 964-71.
13. Targ S, Almeida D, Lyman K. Resnet in resnet: Generalizing residual architectures. arXiv preprint arXiv:1603.08029. 2016.
14. Zang B, Ding L, Feng Z, Zhu M, Lei T, Xing M, *et al*. CNN-LRP: Understanding convolutional neural networks performance for target recognition in SAR images. Sensors (Basel) 2021;21:4536.