



OPEN

A deep learning model for diagnosis of inherited retinal diseases

Freshteh Jafarbeglou^{1,14}, Hamid Ahmadi^{2,14}, Farnaz Soleimani¹, Ali Karimi³, Narsis Daftarian⁴, Sahba Fekri², Tahmineh Motevasseli², Morteza Naderan², Babak Kamali Doust Azad⁵, Abbas Sheikhtaheri⁶, Farid Khorrami⁷, Hemn Baghban Jaldian⁸, Kia Bayat⁹, Saeideh Shahbazi⁸, Mahdi Yazdanpanah¹⁰, Tahereh Sabbaghi¹¹, Kourosh Sheibani¹², Hadi Ghattan Kashani¹, Masoud Shariat Panahi¹✉ & Hamideh Sabbaghi^{8,13}✉

To evaluate the performance of a multi-input deep learning (DL) model in detecting two common inherited retinal diseases (IRDs), i.e. retinitis pigmentosa (RP) and Stargardt disease (STGD), and differentiating them from healthy eyes. This cross-sectional study includes 391 cases, consisting of 158 subjects with RP, 62 patients with STGD, and 171 healthy individuals. The image dataset is publicly available at <http://en.riovs.sbm.ac.ir/Access-to-Inherited-Retinal-Diseases-Image-Bank>. Separate networks using the same hyperparameters were trained and tested on the dataset. Two single-input MobileNetV2 networks were employed for color fundus photography (CFP) and infrared (IR) images, and a multi-input MobileNetV2 network was applied using both imaging modalities simultaneously. The single-input MobileNetV2 achieved 94.44% diagnostic accuracy using CFP, and 94.44% accuracy employing IR images, respectively. The multi-input MobileNetV2 network outperformed both single-input networks with an accuracy of 96.3%. The impact of single-input and multi-input architectures was further evaluated on state-of-the-art neural network models and machine learning algorithms. The deep learning networks utilized in this study achieved high performance for detection of IRDs. Application of a multi-input network employing both CFP and IR image inputs improves the overall performance of the model and its diagnostic accuracy.

Keywords Artificial intelligence, Machine learning algorithms, Deep learning architectures, Inherited retinal diseases, Retinitis pigmentosa, Stargardt disease

Inherited retinal diseases (IRDs) are rare retinal disorders that result in gradual degeneration of the retinal pigment epithelium (RPE) and photoreceptors^{1,2}. IRD patients commonly complain of gradual loss of vision, nyctalopia or hemeralopia, color vision deficiency and restricted field of vision. Retinitis pigmentosa (RP) is the most common type of diffuse photoreceptor dystrophy, with nearly 1.5 million new cases reported annually^{3,4}. Furthermore, Stargardt disease (STGD) is the most common type of macular dystrophy with an estimated incidence of one per 10,000 individuals^{3,5}.

¹School of Mechanical Engineering, College of Engineering, University of Tehran, Tehran, Iran. ²Ophthalmic Research Center, Research Institute for Ophthalmology and Vision Science, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ³School of Mechanical Engineering, College of Engineering, University of Tehran, Tehran, Iran. ⁴Experimental Medicine, Department of Medicine, The University of British Columbia Faculty of Medicine, Vancouver, British Columbia, Canada. ⁵School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran. ⁶Department of Health Information Management, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran. ⁷Department of Health Information Technology, Hormozgan University of Medical Sciences, Bandar Abbas, Iran. ⁸Departments of Optometry, School of Rehabilitation, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ⁹School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ¹⁰Amirkabir University of Technology, Tehran, Iran. ¹¹Department of Information Technology, University of Tarbiat Modares, Tehran, Iran. ¹²Basir Eye Health Research Center, Iran University of Medical Sciences, Tehran, Iran. ¹³Ophthalmic Epidemiology Research Center, Research Institute for Ophthalmology and Vision Science, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ¹⁴Freshteh Jafarbeglou and Hamid Ahmadi have contributed equally to this work. ✉email: mshariatp@ut.ac.ir; sabbaghi.opt@gmail.com

Modern procedures for the diagnosis of retinal diseases substantially rely on multimodal imaging in addition to retinal examination. Common retinal imaging including color fundus photography (CFP) and infrared (IR) imaging, could improve diagnostic accuracy by underscoring known retinal patterns. Vessel attenuation and waxy pallor of optic disc is considered the most common clinical manifestations in RP patients. In the final steps of disease progression melanin pigment containing bone spicules appear in the mid-peripheral or peripheral retina in CFP⁶. For STGD, pathological changes on CFP include macular atrophy and retinal flecks⁶.

Recently, AI-based algorithms have been widely applied in the automatic detection and grading of various diseases using medical image analysis; in particular, they have been extensively applied for retinal images in ophthalmology^{7–11}. Meanwhile, to compensate for limited number of data, transfer learning techniques, VGG^{12,13}, Inception¹⁴, and ResNet^{15,16} have so far been employed as state-of-the-art deep neural network architectures, especially for CFP and optical coherence tomography (OCT) images. Other lightweight architectures, such as MobileNetV2¹⁷, have also been used. Fundus autofluorescence (FAF)¹⁵, OCT, and ultra-widefield images¹⁸ have been applied to detect various IRDs. In the study by Masumoto et al., high sensitivity was achieved in detecting RP on ultra-widefield pseudo color and FAF images¹⁸. The combination of convolutional neural networks (CNNs), OCT¹⁹, and FAF imaging¹⁵, has achieved high accuracy in detecting STGD. Multi-class and multi-labeled models^{20–22} have been employed to significantly improve the performance of existing methods and cover a broader range of retinal disorders. Although some studies have been predicting or detecting IRDs, none of them has simultaneously studied different types of IRDs using various retinal imaging modalities. Designing an intelligent model could improve the quality and accuracy of the health care system in order to detect patients with any type of IRDs by initial screening at the lowest expense and particularly be employed in telemedicine for remote areas. In this study, we proposed a novel method based on MobileNetV2 using a multi-input model for simultaneous processing of both CFP and IR images.

Methods

Dataset

In this study, a total of 220 patients with two types of IRDs including RP ($n=158$) and STGD ($n=62$) registered at the Iranian National Registry for Inherited Retinal Diseases (IRDReg[®]), were investigated. This combination of data was designed to replicate the real-world prevalence of these diseases ($\frac{Stargardt}{RP} \cong \frac{1}{2.6}$)²³. The gender and age distribution of patients in each of these categories are presented in Table 1. In addition, 171 healthy individuals were included as controls. There was no statistically significant difference among the study subjects.

The study protocol was approved by the Ethics Committee of the Ophthalmic Research Center affiliated to Shahid Beheshti University of Medical Sciences, Tehran, Iran (IR.SBMU.ORB.REC.1396.15) and was in accordance with the Declaration of Helsinki. The study participants signed the consent form and were assured that the gathered data would be anonymous. Dataset of the images is publicly available at <http://en.riovs.sbmua.c.ir/Access-to-Inherited-Retinal-Diseases-Image-Bank>.

Initially, patients with suspected or definitive IRD diagnosis were interviewed to record their visual symptoms. Then, they underwent meticulous retinal examination by a retina specialist using a 90D lens and indirect ophthalmoscopy. In the next step, two imaging techniques including CFP (Visucam Pro NM; Carl Zeiss Meditec AG, Germany) and IR (Heidelberg Engineering, Germany), were applied for all participants.

Data preprocessing techniques, including Crop, Center, Flip (if left-eye image), and Contrast Limited Adaptive Enhancement, were applied to reduce the variance of the images and produce a more uniform image dataset. Figure 1 shows samples of the raw CFP and IR images.

Proposed method

In the present study, retinal image processing and classification were conducted using MobileNetV2. During the training process, the network learns to extract the basic features of the images and thereby, determine the type of disease according to these features. In this type of network, first, the features are extracted using convolutional layers. Then, the extracted high-level features are fed into the dense layers, the output of which is the predicted label of the input retinal images. MobileNetV2 is based on an inverted residual structure where the residual connections are between the bottleneck layers. The architecture of MobileNetV2 contains an initial fully convolutional layer with 32 filters, followed by 19 residual bottleneck layers [19].

The proposed method is a novel approach that simultaneously receives CFP and IR images. As illustrated in Fig. 2, the network's training begins with cropping the CFP and IR images, which are merged into a single image before being fed into the MobileNetV2 network. This innovative method allows the network to leverage the complementary information obtained from both types of images, enhancing its ability to predict the presence of diseases such as STDG and RP or to determine the normalcy of the individual. The network is designed to output the necessary predictions based on the integrated features derived from the combined input images.

Factors	Subject	RP	STGD	Healthy eyes	P-value
Age (yrs)	Mean \pm std	40 \pm 13	30 \pm 12	38 \pm 12	0.41
	Median (min, max)	40 (7, 74)	31 (6, 63)	34 (6, 72)	
Gender (%)	Male	74 (47%)	27 (44%)	73 (43%)	0.14
	Female	84 (53%)	35 (56%)	98 (57%)	

Table 1. Age and gender distribution of the patients who were recruited in this study.

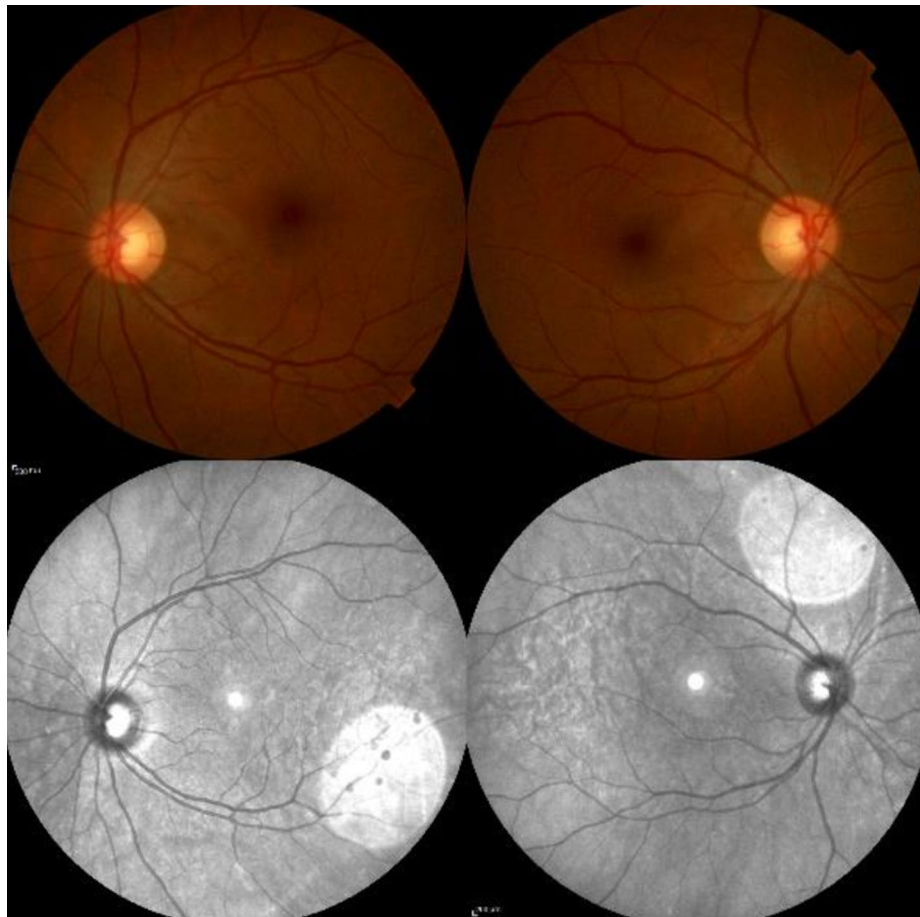


Fig. 1. Two sample raw images of color fundus photograph (CFP) and infrared (IR).

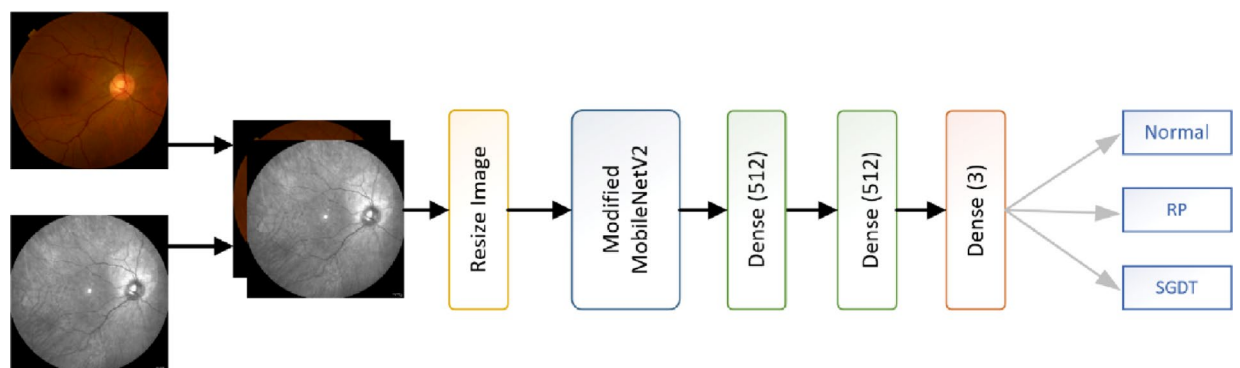


Fig. 2. The schematic overview of the image processing in the present study.

Implementation and training

The current study utilized the TensorFlowTM²⁴ framework as a deep learning (DL) platform. After preprocessing, images were randomly split into a training set and a test set. A CNN architecture based on the MobileNetV2 model was trained to classify the images into normal, RP, and STGD categories. Transfer learning was applied using the initial weights provided by ImageNet²⁵. Our multi-input model was simultaneously implemented and trained on CFP and IR images. MobileNetV2 was used as a feature extractor for all of the input paths, and the same classifier layers and hyperparameters were applied. A total of 250 (79%) RP, 86 (69%) STGD, and 266 (78%) healthy eyes were utilized to train the model. Also, the remaining images were used for the test set. All models were trained for 30 epochs on the same training set and were tested on the same test set. The Adam optimizer was employed with a learning rate of 10^{-5} along with a learning rate scheduler that featured a patience of 1 and a decay ratio of 0.9. The final layer utilized a SoftMax activation function, and weighted cross-entropy loss was

applied during training to address the imbalance in the dataset. All models had an input shape of 224 * 224 * 3, indicating that both image modalities were loaded and processed in RGB format.

In this study, a diverse set of machine learning (ML) algorithms and DL models were employed to evaluate the performance of the proposed approach. Among the ML algorithms, the Extremely Randomized Tree²⁶, Support Vector Machine (SVM)²⁷, XGBoost²⁸, Random Forest²⁹, and LightGBM³⁰ were utilized due to their proven effectiveness in handling structured data and their ability to capture complex patterns. Also several state-of-the-art DL models were incorporated to assess the performance on unstructured data, particularly in image-related tasks. These models include MnasNet-A1³¹, AlexNet³², VGG11³³, ShuffleNetV21 × ³⁴, VGG13³³, InceptionV3³⁵, ResNet50³⁶, VGG16³³, DenseNet121³⁷, and MobileNetV2³⁸. Each of these architectures was chosen for its unique design and capabilities, ranging from lightweight models optimized for mobile devices to deeper networks designed for high accuracy in complex tasks. By leveraging this comprehensive set of algorithms and models, the evaluation process ensures a thorough analysis of the proposed method's performance across various scenarios and data types.

Evaluation metrics

The performance of the models was evaluated based on various metrics, including accuracy, balanced accuracy, precision, F1 score, sensitivity, specificity, and negative predictive value (NPV). The Receiver Operating Characteristics (ROC) and Precision-Recall (PR) curves of all classification models were reported for each approach. Furthermore, Grad-CAM visualization of the class activation functions was implemented to enhance the interpretability of the final multi-input model. This representation shows the regions and areas of the image that the model is concentrating on which helps identify the feature the model considers relevant to its prediction. This step can ensure that the model detects meaningful elements in an image and makes predictions based on meaningful patterns.

Results

The classification performance of various ML algorithms and DL networks is summarized in Table 2, highlighting the superiority of DL models over traditional ML approaches. Among the ML algorithms, LightGBM achieved the highest accuracy (77.16%) and balanced accuracy (67.92%), outperforming other ML methods such as Random Forest and XGBoost. However, the DL models demonstrated significantly better performance, with MobileNetV2 achieving the highest accuracy (96.3%) and balanced accuracy (95.48%), along with an AUC of 99.31%, surpassing all other models. Inception V3 and VGG16 also performed exceptionally well, with equal accuracy accuracies of 95.68% and AUC values of 99.37% and 98.52%, respectively. The results indicate that DL models, particularly MobileNetV2, provide more robust and reliable classification performance compared to traditional ML methods, making them better suited for the automated detection of inherited retinal diseases.

As shown in Table 3, the Multi-input models, which combine CFP and IR images, outperform the Single-input models (using either CFP or IR alone). This superiority stems from the complementary information provided by the two modalities: CFP captures detailed color-based retinal features, while IR highlights deeper structures and vascular patterns. For instance, in the MobileNetV2 model, the Multi-input configuration achieved an accuracy of 96.3% and an AUC of 99.31%, whereas the Single-input models (CFP and IR) both achieved an accuracy of 94.44%. Similarly, in the Inception V3 model, the Multi-input setup achieved an accuracy of 95.68% and an AUC of 99.37%, outperforming CFP (92.59%) and IR (93.21%). These results demonstrate that combining CFP and IR in Multi-input models enhances sensitivity, accuracy, and the ability to detect more complex retinal disease features, offering a significant advantage over Single-input approaches.

Algorithm		Accuracy	Balanced accuracy	Precision	F1	Specificity	NPV	AUC
ML algorithms	Extremely randomized tree ²⁶	72.22 (71.6–72.84)	63.88	73.46	65.07	84.02	90.91	88.22
	Support vector machine ²⁷	72.84 (71.6–73.46)	63.29	57.59	63.69	81.89	92.06	79.38
	XGBoost ²⁸	75.93 (74.69–76.55)	66.16	59.97	66.51	83.77	93.84	87.15
	Random forest ²⁹	75.93 (74.07–77.17)	66.16	58.99	66.14	85.69	94.02	80.81
	LightGBM ³⁰	77.16 (75.3–78.14)	67.92	77.66	69.81	84.53	94.02	81.17
DL networks	MnasNet-A1 ³¹	83.33 (80.85–86.42)	78.85	82.46	81.96	90.67	93.51	94.93
	AlexNet ³²	87.65 (87.65–90.12)	86.55	87.83	87.52	94.19	93.95	97.58
	VGG11 ³³	92.59 (90.12–93.83)	91.33	92.5	92.47	96.17	96.61	97.26
	ShuffleNetV21 × ³⁴	93.83 (91.36–94.44)	93.09	93.76	93.78	96.86	96.96	98.39
	VGG13 ³³	94.44 (92.59–95.06)	93.22	94.46	94.35	96.97	97.63	98.02
	Inception V3 ³⁵	95.68 (95.06–96.3)	94.6	95.71	95.64	97.5	98.2	99.37
	ResNet50 ³⁶	95.06 (94.44–95.06)	93.73	95.13	95	97.08	98.01	98.25
	VGG16 ³³	95.68 (95.06–96.3)	94.98	95.64	95.65	97.82	97.98	98.52
	DenseNet121 ³⁷	95.06 (94.44–95.68)	94.47	95.04	95.05	97.63	97.56	98.83
	MobileNetV2 ³⁸	96.3 (95.06–96.3)	95.48	96.3	96.27	97.92	98.39	99.31

Table 2. Classification performance of deep learning models and machine learning algorithm for our method (Accuracy, balanced accuracy, precision,, F1 score, specificity, NPV, AUC calculated using a weighted approach).

Algorithm		Input	Accuracy	Balanced accuracy	Precision	F1	Specificity	NPV	AUC
ML algorithms	XGBoost ²⁸	CFP	73.46	63.86	59.9	64.83	81.83	92.42	85.13
		IR	73.46	63.86	59.9	64.83	81.83	92.42	85.13
		Multi-Input	75.93 (74.69–76.55)	66.16	59.97	66.51	83.77	93.84	87.15
	LightGBM ³⁰	CFP	73.46	63.86	58.8	64.49	82.07	92.42	87.69
		IR	75.31	71.26	74.69	74.01	85.63	88.42	91.94
		Multi-Input	77.16 (75.3–78.14)	67.92	77.66	69.81	84.53	94.02	91.17
DL networks	AlexNet ³²	CFP	87.04	83.37	87.13	86.18	92.35	95.06	94.06
		IR	87.04	83.37	87.13	87.04	86.18	92.35	95.06
		Multi-Input	87.65 (87.65–90.12)	86.55	87.83	87.52	94.19	93.95	97.58
	ShuffleNetV2 1 × ³⁴	CFP	90.12	88.97	90.12	90.12	95.26	95.26	98.22
		IR	90.74	88.63	90.78	90.5	94.58	96.07	95.41
		Multi-Input	93.83 (91.36–94.44)	93.09	93.76	93.78	96.86	96.96	98.39
	Inception V3 ³⁵	CFP	92.59	90.22	92.72	92.3	95.54	97.32	98.39
		IR	93.21	92.51	93.27	93.23	96.51	96.57	97.59
		Multi-Input	95.68 (95.06–96.3)	94.6	95.71	95.64	97.5	98.2	99.37
	ResNet50 ³⁶	CFP	93.21	93.33	93.75	93.27	97.37	96.24	98.39
		IR	93.83	91.97	93.91	93.67	96.31	97.66	97.63
		Multi-Input	95.06 (94.44–95.06)	93.73	95.13	95	97.08	98.01	98.25
	VGG16 ³³	CFP	94.44	92.85	94.56	94.35	96.65	97.83	98.83
		IR	92.59	90.59	92.5	92.35	95.93	97.08	97.8
		Multi-Input	95.68 (95.06–96.3)	94.98	95.64	95.65	97.82	97.98	98.52
	DenseNet121 ³⁷	CFP	93.83	92.72	93.76	93.74	96.7	97.19	97.19
		IR	94.44	93.15	94.38	94.37	97.05	97.68	98.83
		Multi-Input	95.06 (94.44–95.68)	94.47	95.04	95.05	97.63	97.56	98.83
	MobileNetV2 ³⁸	CFP	94.44	93.15	94.38	94.37	97.05	97.68	98.83
		IR	94.44	93.59	94.43	94.4	97.36	97.4	99.15
		Multi-Input	96.3 (95.06–96.3)	95.48	96.3	96.27	97.92	98.39	99.31

Table 3. Comparison of classification accuracy (%) across different models and training methods (Single input vs multi input) for retinal disease image classification. Significance value bold.

The Precision-Recall (PR) curves and Receiver Operating Characteristic (ROC) curves for all three models are presented in Fig. 3. These curves illustrate the superior performance of the Multi-input model, which consistently achieved the highest precision and recall across all disease categories. The ROC curves further confirm the Multi-input model’s ability to effectively balance sensitivity and specificity, achieving near-perfect AUC values for all three classes (RP, STGD, and healthy eyes).

To enhance the interpretability of the proposed model, Grad-CAM visualizations were generated to identify the regions of the retinal images that contributed most to the model’s predictions. As shown in Fig. 4, the Multi-input model accurately focused on disease-relevant features, such as bone spicules and vascular attenuation in RP and macular atrophy and retinal flecks in STGD. These visualizations validate that the model makes predictions based on meaningful retinal features, increasing its reliability for clinical applications.

The performance details of the proposed method are summarized in Table 4, which presents the metrics for detecting RP, STGD, and normal eyes. The Multi-input model, which combines both CFP and IR images, significantly outperformed the Single-input models in overall performance. Notably, the precision for detecting normal eyes improved to 100%. Furthermore, the specificity values for RP and STGD detection increased to 95.83% and 98.39%, respectively, as indicated in Table 4. These improvements reflect the benefits of integrating complementary information from CFP and IR imaging modalities.

Discussion

The current study demonstrated the excellent performance of a DL model in detecting RP and STGD based on CFP and IR images. These two entities were selected due to being the most prevalent types of IRDs in different populations and the fact that the clinical diagnosis of RP may not be easily possible very early or advanced stages of the disease due to similarities to other retinal pathologies. Furthermore, this automated model might improve reaching a clinical diagnosis given the widespectrum of causative genes reported in RP.

Based on recommendations of the American Academy of Ophthalmology (AAO)³⁹, non-invasive retinal imaging is promising in the diagnosis of IRDs. Herein, we investigated the utility of IR images alongside CFP for detection of STGD and RP; this approach has not been previously reported. As shown in Table 5, although numerous studies have investigated automated detection of different types of IRDs using different retinal images^{17–19,40,41}, none of them has investigated the use of IR images.

In the current study, the MobileNetV2 architecture was trained simultaneously using CFP and IR images to diagnose RP and STGD. The single-input analysis using CFP had an acceptable accuracy of 94.44%, AUC of

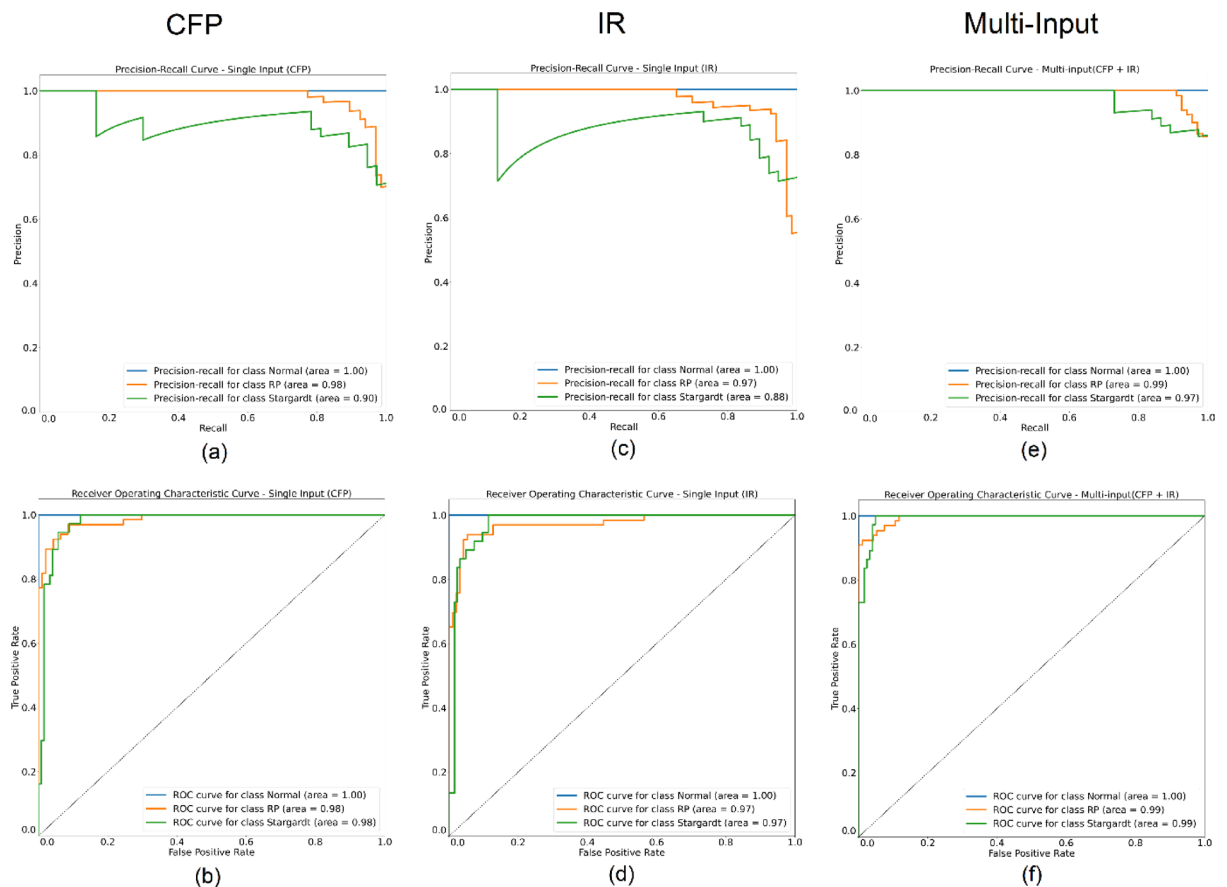


Fig. 3. Precision-Recall Curves and ROC Curves for three trained models on color fundus photographs, CFP (a & b), infrared, IR (c & d), and multi-input (e & f) images.

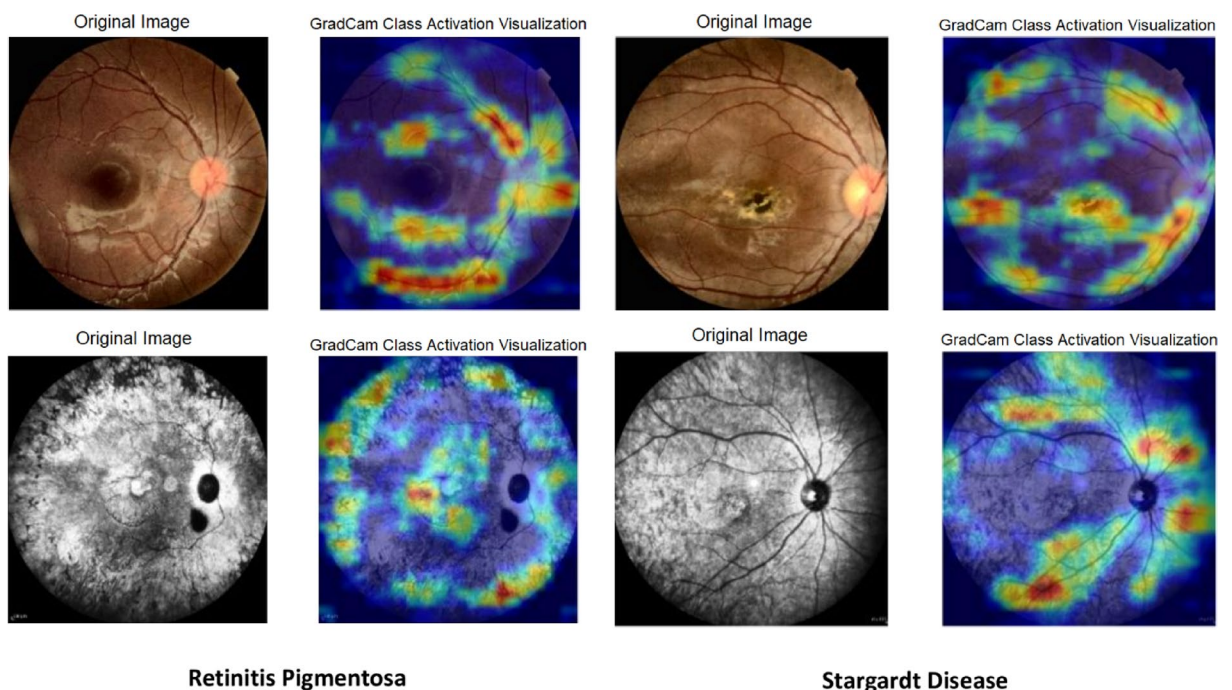


Fig. 4. Sample visualizations of the Grad-Cam class activations for both RP and Stargardt disease derived from the multi-input model.

		RP	STGD	Healthy eyes
Training mode	Support	66	38	58
	Train images	250	86	284
	Test images	66	38	58
	Total images	316	124	342
	Total patient	158	62	171
	Prevalence	0.4074	0.2345	0.3580
Single-input (CFP)	Precision	0.9412	0.9143	0.9661
	Sensitivity	0.9697	0.8421	0.9828
	Specificity	0.9583	0.9758	0.9808
	NPV	0.9787	0.9528	0.9903
	F1	0.9552	0.8767	0.9744
	AUC	0.9813	0.9644	0.9973
Single-input (IR)	Precision	0.8919	0.9688	0.9508
	Sensitivity	0.8684	0.9394	1
	Specificity	0.9677	0.9792	0.9712
	NPV	0.96	0.9592	1
	F1	0.88	0.9538	0.9748
	AUC	0.9824	0.9902	0.9988
Multi-input (CFP + IR)	Precision	0.9412	0.9444	1
	Sensitivity	0.9697	0.8947	1
	Specificity	0.9583	0.9839	1
	NPV	0.9787	0.9683	1
	F1	0.9552	0.9189	1
	AUC	0.9905	0.9868	1

Table 4. Class-wise performance metrics for RP, STGD, and healthy eyes using different inputs. Significance value bold.

References	Image type	Disease	Model	Results
Ta-Ching Chen et al. ⁴⁰	Color Fundus (1670)	RP	Xception	AUC-ROC: 96.74%/ Sensitivity: 95.71% Accuracy: 96.00
Guo et al. ¹⁷	Color Fundus (250)	Glaucoma, maculopathy, pathological myopia, RP	MobileNetV2	Accuracy: 96.2% Sensitivity: 90.4% Specificity: 97.6%
Fujinami -Yokokawa et al. ⁴¹	Color Fundus and FAF (417)	Stargardt, RP, occult macular dystrophy	InceptionV3	Fundus: Accuracy: 88.2% Sensitivity/Specificity: 88.3%/97.4% FAF: Accuracy: 81.3% Sensitivity/Specificity: 81.8%/95.5%
Shah et al. ¹⁹	OCT (102)	STGD	CNN	Accuracy: 99.6% Sensitivity: 99.8% Specificity: 98.0%
Masumoto et al. ¹⁸	ultrawide-field pseudocolor and ultrawide-field autofluorescence (373)	RP	CNN	Pseudocolor : Sensitivity: 99.3% Specificity: 99.1% Autofluorescence: Sensitivity: 100% Specificity: 99.5%
This study	Color Fundus and IR (233)	RP, Stargardt	Multi-Input MobileNetV2	Accuracy: 96.3% Sensitivity: 96.3% Specificity: 97.92% AUC-ROC: 99.31%

Table 5. The classification results of the similar studies reported in the literature and the model proposed in this study.

98.83%, and average precision of 94.38%. Corresponding values utilizing a single input model with IR images achieved an overall accuracy of 94.44%, AUC of 99.15%, and average precision of 94.43%. The single-input IR model had higher sensitivity for STGD and higher specificity for RP detection. Numerous experiments showed that the proposed multi-input approach resulted in higher diagnosis accuracy than classical models that use either CFP or IR retinal images as their single input.

In the current study, RP detection using a single-input CFP model showed higher sensitivity compared to the model used by Chen et al.⁴⁰ conducted on CFP images from a large sample of 1153 RP and 517 healthy eyes. These authors utilized a fine-tuned Xception architecture model to distinguish RP patients from normal eyes with the sensitivity of 95.71%, AUC-ROC of 96.74% and accuracy of 96%. Chen et al. [37] trained and validated their model based on registered images in the Taiwan inherited retinal degeneration project employing various instruments and technicians. Our data were collected from the IRDReg[®] program in which all patients were examined by one retina specialist (N.D) and one optometrist (H.S), and a single experienced technician conducted all CFP and IR imaging.

Another study on RP patients conducted by Masumoto et al.¹⁸ investigated the processing of ultrawide- field pseudo color and ultra-wide field autofluorescence images using a VGG16 model for RP detection. They found that both imaging modalities achieved high (>99%) values for AUC-ROC, sensitivity, and specificity metrics. This slight difference with our figures can be attributed to the extensive field of data capture by ultrawide- field pseudo color and autofluorescence imaging.

Shah et al.¹⁹ specifically performed a study on patients with STGD regarding disease detection and severity utilizing two methods: VGG19 architecture and a customized LeNet model⁴². Their model was capable of distinguishing STGD patients from healthy controls using optical coherence tomography (OCT) images and yielding accuracy of 99%, sensitivity of 99.8%, and specificity of 98% for the first model. This level of performance may be attributed to this fact that OCT is focused on retinal structures in the macula, the area most severely involved in STGD.

An automated model was developed by Fujinami-Yokokawa et al.⁴¹ for prediction of the most common causative genes in RP, STGD and Occult macular dystrophy by analyzing CFP and FAF images. The overall accuracy, sensitivity, and specificity of CFP was 88.2%, 88.3%, and 97.4%, respectively with Inception V3 architecture, trained and tested, with 4-fold cross validation method, while our model could automatically detect the RP and STGD more accurately. It could be concluded that the sensitivity of >90% was achieved by analyzing of the CFP images in RP patients with mutation in *RP11* gene and high specificity (>90%) was identified in detection of RP and STGD patients having mutation in *ABCA4*, *EYS* and *RP11*. In a different study published by the same author analyzing the OCT images on a small population, the mean accuracy and sensitivity of 100% for *ABCA4*, >78% for *RP11*, and >89% for *EYS* were identified. The analysis of the causative genes by automated models is going to be investigated in the future studies of IRDReg[®] program.

In a study conducted by Miere et al.¹⁵ on patients with RP, STGD and Best disease by automated analyzing of FAF images, the overall accuracy was 95% for all studied IRDs. The sensitivity, specificity, and AUC-ROC were obtained greater than 96% for detection of RP and STGD. In our study, the sensitivity and specificity of IR and CFP images for detection of RP and STGD were more than 90%. This difference in results might be attributed to the different imaging techniques, variation in parameters and architectures used for classification.

In another study conducted by Miere et al.⁴³, the single input ResNet101 model findings of FAF images with highest accuracy of 92% and AUC-ROC of 98.1% were reported for differential diagnosis of advanced age-related macular degeneration (AMD) with geographic atrophy from late onset STGD and pseudo- Stargardt pattern dystrophy. This study mainly focused on the differential diagnosis of AMD comparison with the above mentioned IRD types and no results were reported in detection of STGD disease for further comparison.

Deep neural networks have been extensively used for medical image analysis and diagnosis. The complex architecture of such networks allows them to achieve a level of abstraction where high-level information is extracted from raw images. This high-level information is then used for classification of the input image. MobileNetV2 was chosen over other network architectures in this study due to the highest accuracy and precision compared with the other popular architectures, such as DenseNet121, ResNet50, InceptionV3, and VGG16.

The increasing amount of data for diagnosis can be time-consuming and overwhelming for medical professionals. This issue is under extensive investigation by ongoing studies and AI applications can be particularly be beneficial in medicine, particularly in the field of ophthalmology. AI-based automated models enable more accurate diagnosis for referring IRD patients by analyzing their condition. This could also be influential in the early diagnosis of other family members suspicious of IRD diseases with less dependency on a retina specialist^{40,41}.

Automated AI models can be effective in decision support systems to telescreen and detect patients with different IRDs in remote locations with no access to retina specialists. These AI modalities may also lead to more accurate IRD diagnosis, foster personalized medicine, assist possible gene therapy, and eventually help provide better therapeutic outcomes¹⁷.

The current study demonstrated that the suggested multi-input MobileNetV2 outperformed its single-input alternatives where only IR or CFP images were used separately. Its efficiency was comparable to other state-of-the-art neural networks and machine learning algorithms. One of the strengths of the current study was that all images were collected from the patients in the Iranian National Registry for Inherited Retinal Diseases (IRDReg[®]) in which the retinal imaging was conducted by a single experienced technician using a same imaging device. The proposed multi-input approach seems to enjoy the advantage of high diagnostic accuracy and acceptably low amount of data required to train the network. By merging features of IR and CFP images in the model, RP and SGDT were diagnosed more accurately than using single input IR and CFP image models.

One of the limitations of the present study was that the DL model could not grade disease severity in RP and STGD. This model may have the potential to get improved and diagnose of other types of IRDs, detect disease

progression over time and also classify disease severity. In the future studies, other imaging modalities can be investigated alongside CFP and IR images.

Conclusion

We proposed a multi-input deep learning neural network model based on MobileNetV2 to classify IRDs. In this approach, each individual's CFP and IR images were fed into the model to get a prediction. Our results indicate that by accumulating visual features extracted from both of these imaging modalities, the multi-input model outperforms the single-input models separately trained on each set of CFP or IR images. This configuration makes use of the complimentary data of the two imaging modalities to help diagnose the condition.

Data availability

Data could be available by sending a request to the corresponding author.

Received: 7 February 2024; Accepted: 28 May 2025

Published online: 02 July 2025

References

1. Sabbaghi, H. et al., *The first inherited retinal disease registry in Iran: research protocol and results of a pilot study*. (2020).
2. Ben-Yosef, T. Inherited retinal diseases. *MDPI* **23**, 13467 (2022).
3. Sabbaghi, H. et al. Choroidal thickness in different types of inherited retinal dystrophies. *J. Ophthalmic Vis. Res.* **15**, 351 (2020).
4. Ayton, L. N., Guymer, R. H. & Luu, C. D. Choroidal thickness profiles in retinitis pigmentosa. *Clin. Exp. Ophthalmol.* **41**, 396–403 (2013).
5. Sodi, A. et al. EDI OCT evaluation of choroidal thickness in Stargardt disease. *Plos One* **13**, e0190780 (2018).
6. Pichi, F., Abboud, E. B., Ghazi, N. G. & Khan, A. O. Fundus autofluorescence imaging in hereditary retinal diseases. *Acta Ophthalmol.* **96**, e549–e561 (2018).
7. Jagan Mohan, N., Murugan, R., Goel, T., Mirjalili, S. & Roy, P. A novel four-step feature selection technique for diabetic retinopathy grading. *Phys. Eng. Sci. Med.* **44**, 1351–1366 (2021).
8. Mohan, N. J., Murugan, R., Goel, T. & Roy, P. Fast and robust exudate detection in retinal fundus images using extreme learning machine autoencoders and modified Kaze features. *J. Digit. Imaging.* **35**, 496–513 (2022).
9. Mohan, N. J., Murugan, R., Goel, T. & Roy, P. DRFL: Federated learning in diabetic retinopathy grading using fundus images. *IEEE Trans. Parallel Distrib. Syst.* **34**, 1789–1801 (2023).
10. Mohan, N. J. et al. Optimal hybrid feature selection technique for diabetic retinopathy grading using fundus images. *Sādhanā* **48**, 102 (2023).
11. Mohan, N. J., Murugan, R., Goel, T., Tanveer, M. & Roy, P. An efficient microaneurysms detection approach in retinal fundus images. *Int. J. Mach. Learn. Cybernet.* **14**, 1235–1252 (2023).
12. Gómez-Valverde, J. J. et al. Automatic glaucoma classification using color fundus images based on convolutional neural networks and transfer learning. *Biomedical Opt. Express.* **10**, 892–913 (2019).
13. Nagasato, D. et al. Deep-learning classifier with ultrawide-field fundus ophthalmoscopy for detecting branch retinal vein occlusion. *Int. J. Ophthalmol.* **12**, 94 (2019).
14. Li, F. et al. Automatic detection of diabetic retinopathy in retinal fundus photographs based on deep learning algorithm. *Transl. Vis. Sci. Technol.* **8**, 4–4 (2019).
15. Miere, A. et al. Deep learning-based classification of inherited retinal diseases using fundus autofluorescence. *J. Clin. Med.* **9**, 3303 (2020).
16. Hemelings, R. et al. Accurate prediction of glaucoma from colour fundus images with a convolutional neural network that relies on active and transfer learning. *Acta Ophthalmol.* **98**, e94–e100 (2020).
17. Guo, C., Yu, M. & Li, J. Prediction of different eye diseases based on fundus photography via deep transfer learning. *J. Clin. Med.* **10**, 5481 (2021).
18. Masumoto, H. et al. Accuracy of a deep convolutional neural network in detection of retinitis pigmentosa on ultrawide-field images. *PeerJ* **7**, e6900 (2019).
19. Shah, M., Roomans Ledo, A. & Rittscher, J. Automated classification of normal and Stargardt disease optical coherence tomography images using deep learning. *Acta Ophthalmol.* **98**, e715–e721 (2020).
20. Nagasawa, T. et al. Accuracy of ultrawide-field fundus ophthalmoscopy-assisted deep learning for detecting treatment-naïve proliferative diabetic retinopathy. *Int. Ophthalmol.* **39**, 2153–2159 (2019).
21. Sahlsten, J. et al. Deep learning fundus image analysis for diabetic retinopathy and macular edema grading. *Sci. Rep.* **9**, 10750 (2019).
22. Redd, T. K. et al. Evaluation of a deep learning image assessment system for detecting severe retinopathy of prematurity. *Br. J. Ophthalmol.* **103**, 580–584 (2019).
23. Arai, Y. et al., Retinitis pigmentosa with EYS mutations is the most prevalent inherited retinal dystrophy in Japanese populations. *J. Ophthalmol.* **2015** (2015).
24. Abadi, M. et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467* (2016).
25. Deng, J. et al. Imagenet: A large-scale hierarchical image database. in *IEEE conference on computer vision and pattern recognition* 248–255 (IEEE, 2009).
26. Geurts, P., Ernst, D. & Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **63**, 3–42 (2006).
27. Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J. & Scholkopf, B. Support vector machines. *IEEE Intell. Syst. Their Appl.* **13**, 18–28 (1998).
28. Chen, T., Guestrin, C. & Xgboost A scalable tree boosting system. in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* 785–794 (2016).
29. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
30. Ke, G. et al. Lightgbm: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **30** (2017).
31. Tan, M. et al. Mnasnet: Platform-aware neural architecture search for mobile. in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* 2820–2828 (2019).
32. Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **25** (2012).
33. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).
34. Ma, N., Zhang, X., Zheng, H. T. & Sun, J. Shufflenet v2: Practical guidelines for efficient cnn architecture design. in *Proceedings of the European conference on computer vision (ECCV)* 116–131 (2018).

35. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. Rethinking the inception architecture for computer vision. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 2818–2826 (2016).
36. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 770–778 (2016).
37. Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 4700–4708 (2017).
38. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. & Chen, L. C. Mobilenetv2: Inverted residuals and linear bottlenecks. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 4510–4520 (2018).
39. Stuart, A. Inherited Retinal Diseases. American Academy of Ophthalmology. <https://www.aao.org/eyenet/article/inherited-retinal-diseases> (2018).
40. Chen, T. C. et al. Artificial intelligence–assisted early detection of retinitis pigmentosa—the most common inherited retinal degeneration. *J. Digit. Imaging*. **34**, 948–958 (2021).
41. Fujinami-Yokokawa, Y. et al. Prediction of causative genes in inherited retinal disorder from fundus photography and autofluorescence imaging using deep learning techniques. *Br. J. Ophthalmol.* **105**, 1272–1279 (2021).
42. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324 (1998).
43. Miere, A. et al. Deep learning-based classification of retinal atrophy using fundus autofluorescence imaging. *Comput. Biol. Med.* **130**, 104198 (2021).

Acknowledgements

This article has been taken from the disease registry, titled “The Iranian National Registry for Inherited Retinal Dystrophy (IRDReg[®])” and the code number of IR.SBMU.ORB.REC.1396.15, supported by the Deputy of Research and Technology at Shahid Beheshti University of Medical Sciences (<http://dregistry.sbm.ac.ir>).

Author contributions

E.J.: Data analysis; Software development; Manuscript writing. H.A.: Patient examination; Data interpretation; Manuscript writing. F.S.: Data analysis; Software development; Manuscript writing. A.K.: Data analysis; Manuscript writing. N.D.: Patient examination; Data interpretation; Manuscript writing. S.F.: Patient examination. T.M.: Patient examination. M.N.: Patient examination. B.K.D.A.: Software development. A.S.T.: Software development. F.K.: Software development. H.B.J.: Manuscript writing. K.B.: Manuscript writing. S.S.: Patient examination. M.Y.: Software development. T.S.: Software development. K.S.: English editing. H.G.K.: Data analysis; Software development; Manuscript writing. M.S.P.: Data analysis; Software development; Manuscript writing. H.S.: Data analysis; Software development; Manuscript writing.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.S.P. or H.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025