

A MODULAR WEB-BASED PLATFORM FOR MULTIMODAL VISUAL-FUNCTION ASSESSMENT IN GLAUCOMA: PRELIMINARY EVALUATION USING A CONVOLUTIONAL NEURAL NETWORK

Dhathri K¹, Sofia Rebecca², Harshitha S²

Received : 30/03/2026
Received in revised form : 19/05/2026
Accepted : 29/09/2026

Keywords:

Glaucoma; Early Glaucoma Detection; Contrast Sensitivity; Visual Field; Pelli-Robson; SPARCS; Deep Learning; Convolutional Neural Network; Multimodal Learning; Web-Based Screening.

Corresponding Author:

Dr. Dhathri K,
Email: dhathri13@gmail.com

DOI: 10.47009/jamp.2026.8.5.110

Source of Support: Nil,
Conflict of Interest: None declared

Int J Acad Med Pharm
2026; 8 (5); 700-710



¹Department of Ophthalmology, Dr. Chandramma Dayananda Sagar Institute of Medical Education and Research, Dayananda Sagar University, Bengaluru, Karnataka, India

²Assistant Professor, Department of Computer Science and Engineering, School of Engineering and Technology, Dayananda Sagar University, Bengaluru, Karnataka, India

ABSTRACT

Background: Glaucoma is a progressive optic neuropathy characterized by irreversible retinal ganglion-cell loss and visual dysfunction. Conventional screening approaches, including intraocular pressure measurement and standard automated perimetry, may have limitations in identifying early functional impairment. Contrast sensitivity, motion perception, spatial processing, and tracing performance may provide complementary information regarding visual dysfunction. The integration of these functional parameters with machine-learning techniques may therefore provide a broader approach to glaucoma screening. The aim is to develop and preliminarily evaluate a modular web-based platform integrating multiple visual-function assessments and to determine the performance of a multimodal convolutional neural network (CNN) for classification of participants into normal, mild visual deficit, and severe visual deficit categories. **Materials and Methods:** This preliminary observational evaluation was conducted over 12 months at the Department of Ophthalmology, Dr. Chandramma Dayananda Sagar Institute of Medical Education and Research, Dayananda Sagar University, Bengaluru, Karnataka, India. A total of 200 participants were included and classified according to the study reference standard into normal ($n = 92$), mild deficit ($n = 85$), and severe deficit ($n = 23$) groups. The web-based platform incorporated seven visual-function assessments: Pelli-Robson contrast sensitivity, CSV-1000 contrast sensitivity, 24-2 Humphrey visual field assessment, SPARCS, motion sensitivity, edge detection, and pattern detection. Demographic and selected clinical information were integrated with test-derived features. Continuous variables were normalized using min-max scaling. A three-branch CNN processed contrast-sensitivity, motion, and spatial/tracing features, followed by fully connected layers and a three-class softmax classifier. Model performance was assessed using five-fold cross-validation, with accuracy, sensitivity, specificity, F1-score, and confusion-matrix analysis. **Result:** The mean age of the study population was 53.0 ± 10.8 years, with 70 (35.0%) male and 130 (65.0%) female participants. Mean intraocular pressure was higher among participants with increasing severity of visual deficit. Visual-function measures demonstrated a progressive deterioration across the normal, mild-deficit, and severe-deficit groups. Pelli-Robson contrast sensitivity, CSV-1000 contrast sensitivity, 24-2 visual-field mean deviation, SPARCS score, motion-detection accuracy, edge-detection performance, and pattern-detection accuracy differed significantly across the three groups ($p < 0.001$ for overall comparisons). The multimodal CNN demonstrated fold-wise accuracies ranging from approximately 89% to 92%, with an overall mean accuracy of approximately 90.5%. Aggregate sensitivity was 89%, specificity was 90%, and the overall F1-score was 0.90. The confusion matrix demonstrated good separation of normal and deficit categories, although misclassification was more frequent between adjacent severity classes. **Conclusion:** The evaluated web-based platform successfully integrated multiple complementary visual-function measures into a single modular assessment framework. The preliminary CNN results indicate that multimodal functional visual data may support automated classification of

glaucoma-related visual impairment and may have potential as an adjunct to conventional ophthalmic screening. However, the findings should be regarded as preliminary. Larger prospective studies using standardized ophthalmologist-established reference diagnoses, independent external validation, standardized display calibration, and assessment across diverse patient populations are required before clinical implementation.

INTRODUCTION

Glaucoma is a chronic, progressive optic neuropathy characterized by structural damage to the optic nerve head and corresponding functional impairment of the visual system.^[1] It is an important cause of irreversible visual disability, particularly because visual loss may progress insidiously before patients become aware of significant functional impairment.^[2] Early identification of individuals with glaucomatous functional abnormalities is therefore essential for timely ophthalmic evaluation, initiation of appropriate treatment, and prevention of further visual deterioration.^[3]

Conventional glaucoma assessment relies on a combination of intraocular pressure measurement, optic nerve examination, confrontation visual-field assessment, and standard automated perimetry.^[4] Although these investigations remain fundamental to clinical practice, each has limitations when used for early functional detection. Intraocular pressure is an important risk factor but does not independently identify all patients with glaucoma, particularly those with normal or relatively low intraocular pressure.^[5] Confrontation visual-field testing is dependent on patient cooperation and examiner technique and has limited sensitivity for subtle defects. Standard automated perimetry provides quantitative assessment of the visual field but may demonstrate measurable abnormalities only after a clinically meaningful degree of retinal ganglion-cell dysfunction has developed.^[6]

Visual function, however, encompasses domains that extend beyond conventional visual-field sensitivity. Contrast sensitivity represents the ability to detect differences in luminance between an object and its background and may be impaired even when conventional visual acuity remains relatively preserved.^[7,8] Previous observations have suggested that contrast-sensitivity abnormalities may occur before clearly demonstrable visual-field loss. Assessment of contrast sensitivity therefore provides a potentially useful complementary functional marker in individuals undergoing glaucoma evaluation.^[9,10]

Other aspects of visual processing, including peripheral sensitivity, motion perception, spatial discrimination, edge detection, and pattern recognition, may provide additional information regarding functional visual performance.^[11,12] The simultaneous assessment of these domains could potentially provide a broader representation of visual dysfunction than any single test. A multimodal approach may be particularly valuable for digital

screening systems because responses from different functional tests can be captured, normalized, integrated, and analysed collectively.^[13,14]

Recent advances in machine learning and deep learning have created opportunities for automated analysis of ophthalmic data. Convolutional neural networks (CNNs) have demonstrated considerable potential in identifying clinically relevant patterns from complex ophthalmic datasets. However, many existing approaches have predominantly focused on structural imaging or conventional visual-field information.^[15,16] Some systems have also been developed using relatively restricted datasets or specialized equipment, potentially limiting their applicability to accessible, web-based screening environments.^[17,18] A platform capable of integrating multiple visual-function measurements with selected clinician-provided clinical information may therefore provide a complementary approach to conventional glaucoma assessment.^[19,20]

The present study was designed around a modular web-based platform incorporating seven visual-function assessments: Pelli-Robson contrast sensitivity, 24-2 Humphrey visual-field assessment, SPARCS, CSV-1000 contrast sensitivity, motion sensitivity, edge detection, and pattern detection. The platform records quantitative measures such as response accuracy, reaction time, contrast thresholds, cursor movement, and tracing deviation. These data are subsequently processed and integrated within a multimodal three-branch CNN designed to classify participants into normal, mild visual deficit, and severe visual deficit categories.

The primary objective was to develop and preliminarily evaluate this modular platform and determine whether integration of multiple visual-function domains with a CNN-based classification framework could provide useful discrimination among different levels of visual functional impairment associated with glaucoma. Therefore, it is of interest to study the preliminary diagnostic performance of a multimodal web-based visual-function assessment platform using a convolutional neural network for glaucoma screening.

MATERIALS AND METHODS

Study Design and Setting: A prospective observational study with a technology-development and preliminary diagnostic-evaluation component was conducted over a period of 12 months at the Department of Ophthalmology, Dr. Chandramma Dayananda Sagar Institute of Medical Education and Research, Dayananda Sagar University, Bengaluru,

Karnataka, India. The study involved the development and preliminary clinical evaluation of a modular web-based platform designed to integrate multiple visual-function assessments for glaucoma screening and functional classification.

The platform was developed using Flask as the backend framework, with HTML, CSS, and JavaScript used for the patient and clinician interfaces. Separate patient and clinician views were created to facilitate standardized administration of visual-function assessments and review of individual test results. Each visual assessment was implemented as an independent microservice communicating through REST application programming interfaces. Individual services were assigned separate ports in the 8000–8007 range, while WebSocket-based communication was used where real-time synchronization of responses was required.

Study Population and Sample Size: A total of 200 participants were included in the preliminary evaluation. Participants were categorized into three functional assessment groups according to the clinical ophthalmic evaluation and corresponding visual-function findings: normal visual function (n=92), mild visual functional deficit (n=85), and severe visual functional deficit (n=23).

The sample size of 200 was selected to provide an adequate preliminary dataset for assessment of the feasibility and discriminatory performance of the multimodal platform while maintaining representation across the three functional categories. Because the primary purpose of this work was preliminary platform and model evaluation rather than definitive estimation of population-level diagnostic accuracy, the sample was considered appropriate for an exploratory proof-of-concept study.

Participant Recruitment And Classification: Participants attending the ophthalmology department during the study period were considered for inclusion. The normal group comprised participants without clinically significant glaucomatous visual dysfunction, whereas the mild and severe deficit groups comprised participants demonstrating corresponding degrees of functional visual impairment on ophthalmic assessment.

For analytical purposes, the classification assigned by the ophthalmic evaluation was treated as the reference classification against which the CNN output was evaluated. Participants were classified into normal, mild deficit, or severe deficit categories on the basis of the overall clinical assessment and available visual-function findings.

Inclusion Criteria

Participants were eligible when they were adults aged 18 years or older, were able to understand and perform the computerized visual-function assessments, and provided the required clinical and visual-function information for analysis. Participants assigned to the glaucoma/visual-deficit groups had clinical evidence of functional impairment consistent with the assigned severity category, while individuals

in the normal group had no clinically significant functional deficit.

Exclusion Criteria

Participants were excluded when they were unable to reliably complete the computerized assessments, had visual or neurological conditions likely to substantially interfere with interpretation of the selected visual-function tests, had major cognitive or communication limitations preventing reliable test performance, or had incomplete assessment records.

Visual Function Assessments

Seven complementary visual-function assessments were incorporated into the platform.

The Pelli-Robson contrast sensitivity test was used to evaluate global contrast perception. Participants identified Sloan letters arranged in groups with progressively decreasing contrast, and the resulting log contrast sensitivity was recorded.

The 24-2 Humphrey visual-field assessment was used to evaluate peripheral visual sensitivity using a standardized 54-point test grid. Response accuracy and threshold-related visual-field measurements were recorded and integrated with the other functional measures.

The SPARCS assessment was incorporated to evaluate contrast and spatial attention across multiple visual zones as stimulus intensity was progressively reduced. Performance was recorded quantitatively for integration into the multimodal dataset.

The CSV-1000 assessment was used for central contrast-sensitivity evaluation. Participants responded to visual targets presented at varying contrast levels, with contrast sensitivity and response-related parameters recorded.

Motion sensitivity was assessed using moving visual stimuli. Detection accuracy and response latency were recorded as measures of dynamic visual processing.

The edge-detection component assessed the ability to identify or trace visual boundaries. Tracing deviation and contrast-related performance were captured digitally.

The pattern-detection component assessed recognition and tracing of complex visual paths. Tracing accuracy, deviation from the prescribed pathway, and completion time were recorded.

Data Collection And Variables: Participant responses were captured in real time through the web-based platform. The principal quantitative variables included visual-response accuracy, reaction time, contrast thresholds, visual-field responses, cursor movement, tracing deviation, and completion time. Clinician-provided clinical information and selected symptom and lifestyle variables were combined with the computerized visual-function measurements to generate a participant-level dataset. All assessment results were organized in a structured spreadsheet before statistical and machine-learning analysis. Data were checked for completeness and consistency before preprocessing.

Data Preprocessing: Continuous variables, including contrast-sensitivity measurements, reaction

times, and tracing deviations, were normalized using min-max scaling to place measurements obtained from different visual-function assessments on a common scale. For each continuous variable, normalization was performed according to the formula:

$$x_norm = (x - x_min) / (x_max - x_min)$$

where x represents the observed value, while x_min and x_max represent the minimum and maximum values, respectively, of that variable in the training dataset. The same transformation parameters derived from the training data were subsequently applied to the corresponding validation data to minimize information leakage during model evaluation.

Categorical clinical variables obtained from clinician-provided information were converted into numerical representations using one-hot encoding. The outputs generated from the individual visual-function assessments, including contrast sensitivity, visual-field, motion-detection, edge-detection, and pattern-detection measures, were integrated at the participant level to generate a multimodal feature representation for model development.

To improve model robustness, data augmentation was performed for selected visual-function parameters, particularly contrast-related and motion-related variables. Augmentation consisted of generating clinically plausible variations within the observed physiological range of the corresponding parameters. Augmented observations were used only during model training and were not introduced into the validation folds.

Convolutional Neural Network Architecture: A multimodal three-branch convolutional neural network was developed to process the heterogeneous visual-function data. The first branch was designed to process contrast-sensitivity measurements, including Pelli-Robson, CSV-1000, and related contrast-response features. The second branch processed motion-detection variables, including detection accuracy and response delay. The third branch processed spatial and tracing-related variables derived from the visual-field, edge-detection, pattern-detection, and related spatial assessments.

Each branch consisted of two convolutional layers with rectified linear unit activation, followed by max-pooling and batch normalization. The outputs of the three branches were concatenated to form a combined multimodal representation. This representation was passed through two fully connected layers before a final softmax layer generated probabilities for the three predefined categories: normal, mild deficit, and severe deficit.

Model Training And Validation: Categorical cross-entropy was used as the loss function. The Adam optimizer was employed with a learning rate of 0.001, with $\beta_1=0.9$ and $\beta_2=0.999$. Five-fold cross-validation was performed to assess model generalizability and minimize dependence on a single train-test split.

Within each fold, the model was trained on four-fifths of the available dataset and evaluated on the

remaining one-fifth. Performance across the five folds was subsequently summarized using classification accuracy, sensitivity, specificity, F1-score, and confusion-matrix analysis.

Statistical Analysis: Continuous variables were summarized as mean $\pm$ standard deviation when approximately normally distributed and as median with interquartile range when appropriate. Comparisons among the normal, mild-deficit, and severe-deficit groups were performed using one-way analysis of variance with appropriate post-hoc comparisons for normally distributed variables or the Kruskal-Wallis test for non-normally distributed variables.

Categorical variables were expressed as frequencies and percentages and compared using the chi-square test or Fisher's exact test when expected cell frequencies were small. Two-sided p -values <0.05 were considered statistically significant.

For assessment of model performance, accuracy, sensitivity, specificity, and F1-score were calculated across the five cross-validation folds. Mean performance estimates and corresponding 95% confidence intervals were used to summarize model discrimination. A confusion matrix was generated to assess the distribution of correctly and incorrectly classified participants across the three severity categories.

All statistical analyses and machine-learning procedures were performed using Python-based analytical tools, including standard scientific computing and machine-learning libraries.

Ethical Considerations: The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Institutional ethical approval was obtained or was to be documented before final submission, and informed consent was obtained from participants before assessment. Participant information was handled confidentially, and the computerized assessment data were used for analysis without disclosure of personally identifying information.

RESULTS

A total of 200 participants were included in the final analysis and were classified into three groups according to the predefined reference classification: normal visual function ($n=92$), mild visual functional deficit ($n=85$), and severe visual functional deficit ($n=23$). The overall mean age of the study population was 53.0 ± 10.8 years, with age increasing progressively across the severity groups. Of the 200 participants, 70 (35.0%) were male and 130 (65.0%) were female. Mean intraocular pressure was progressively higher from the normal group through the mild- and severe-deficit groups. Among participants with visual functional impairment, the mean duration of glaucoma-related disease also increased with severity. All major visual-function parameters demonstrated a progressive deterioration

across the three groups. Pelli-Robson and CSV-1000 contrast-sensitivity scores were significantly lower among participants with increasing functional deficit, while 24-2 Humphrey visual-field mean deviation and SPARCS scores showed progressively greater functional impairment. Motion sensitivity and pattern-detection accuracy decreased with increasing severity, whereas edge-detection deviation increased. Between-group differences for the principal visual-

function measures were statistically significant, with $p < 0.001$ for each comparison. The multimodal CNN demonstrated consistent classification performance during five-fold cross-validation, with fold-level accuracy ranging from 89.0% to 92.0%. The aggregate model performance was approximately 90.5% accuracy, 89.0% sensitivity, 90.0% specificity, and an F1-score of 0.90.

Table 1: Baseline Demographic And Clinical Characteristics Of Study Participants

The demographic and clinical profile of the participants is summarized below.

Characteristic	Overall (N=200)	Normal (n=92)	Mild Deficit (n=85)	Severe Deficit (n=23)	p-value
Age, years, mean $\pm$ SD	53.0 $\pm$ 10.8	48.6 $\pm$ 9.1	55.0 $\pm$ 9.7	63.2 $\pm$ 8.1	<0.001
Male, n (%)	70 (35.0)	33 (35.9)	28 (32.9)	9 (39.1)	0.815
Female, n (%)	130 (65.0)	59 (64.1)	57 (67.1)	14 (60.9)	
Intraocular pressure, mmHg, mean $\pm$ SD	16.9 $\pm$ 3.8	14.8 $\pm$ 2.4	18.0 $\pm$ 3.0	22.1 $\pm$ 3.7	<0.001
Disease duration, years, mean $\pm$ SD	3.9 $\pm$ 3.1	0	3.1 $\pm$ 2.4	6.5 $\pm$ 3.6	<0.001

The groups were comparable with respect to sex distribution. However, age, intraocular pressure, and

disease duration differed significantly across the three classification groups.

Table 2: Comparison Of Visual-Function Measures Across Classification Groups

The major visual-function parameters demonstrated a consistent deterioration with increasing severity of visual functional deficit.

Visual-Function Measure	Normal (n=92)	Mild Deficit (n=85)	Severe Deficit (n=23)	Test Statistic	p-value
Pelli-Robson log CS, mean $\pm$ SD	1.82 $\pm$ 0.14	1.56 $\pm$ 0.18	1.28 $\pm$ 0.19	72.6	<0.001
CSV-1000 CS score, mean $\pm$ SD	1.74 $\pm$ 0.16	1.48 $\pm$ 0.19	1.19 $\pm$ 0.21	61.8	<0.001
24-2 HVF mean deviation, dB	-1.8 $\pm$ 1.4	-5.6 $\pm$ 2.1	-12.4 $\pm$ 3.6	126.4	<0.001
SPARCS score, mean $\pm$ SD	82.6 $\pm$ 7.2	70.4 $\pm$ 8.5	56.8 $\pm$ 9.1	87.3	<0.001
Motion sensitivity accuracy, %	93.8 $\pm$ 4.1	85.6 $\pm$ 6.3	72.1 $\pm$ 8.4	85.9	<0.001
Edge detection deviation, px	6.4 $\pm$ 2.8	12.7 $\pm$ 4.6	21.5 $\pm$ 6.2	96.7	<0.001
Pattern detection accuracy, %	94.1 $\pm$ 3.7	84.9 $\pm$ 6.5	70.8 $\pm$ 8.7	99.2	<0.001

CS = contrast sensitivity; HVF = Humphrey visual field.

Pelli-Robson and CSV-1000 contrast-sensitivity scores decreased significantly with increasing visual functional deficit. The 24-2 HVF mean deviation became progressively more negative, while SPARCS

scores and motion- and pattern-detection accuracy declined. Conversely, edge-detection deviation increased substantially in the mild- and severe-deficit groups.

Table 3: Pairwise Comparison Of Principal Visual-Function Parameters

Pairwise analysis demonstrated that the differences were predominantly attributable to progressive deterioration from normal to mild and severe functional impairment.

Parameter	Normal vs Mild p-value	Mild vs Severe p-value	Normal vs Severe p-value
Pelli-Robson log CS	<0.001	<0.001	<0.001
CSV-1000 CS score	<0.001	<0.001	<0.001
24-2 HVF mean deviation	<0.001	<0.001	<0.001
SPARCS score	<0.001	<0.001	<0.001
Motion sensitivity accuracy	<0.001	<0.001	<0.001
Edge detection deviation	<0.001	<0.001	<0.001
Pattern detection accuracy	<0.001	<0.001	<0.001

The pairwise comparisons confirmed statistically significant separation between all three severity

categories across each of the seven principal visual-function measures.

Table 4: Distribution Of Participants According To Selected Clinical And Functional Characteristics

The distribution of selected clinical and functional characteristics demonstrated progressive worsening with increasing severity.

Characteristic	Normal (n=92)	Mild Deficit (n=85)	Severe Deficit (n=23)
IOP >21 mmHg, n (%)	3 (3.3)	14 (16.5)	13 (56.5)
Pelli-Robson log CS <1.50, n (%)	4 (4.3)	39 (45.9)	20 (87.0)
HVF MD worse than -6 dB, n (%)	2 (2.2)	31 (36.5)	22 (95.7)
Motion accuracy <80%, n (%)	1 (1.1)	17 (20.0)	18 (78.3)
Pattern accuracy <80%, n (%)	1 (1.1)	21 (24.7)	18 (78.3)
Edge deviation >15 px, n (%)	2 (2.2)	25 (29.4)	19 (82.6)

The frequency of clinically relevant abnormalities increased progressively across the normal, mild-deficit, and severe-deficit groups.

Table 5: Platform-Derived Response Characteristics

The web-based platform captured multiple quantitative response parameters in addition to conventional visual-function scores.

Response Variable	Normal	Mild Deficit	Severe Deficit	p-value
Mean response time, seconds	1.21 ± 0.24	1.48 ± 0.31	1.82 ± 0.39	<0.001
Mean tracing deviation, px	6.8 ± 2.7	13.1 ± 4.4	21.8 ± 6.0	<0.001
Mean motion-response latency, ms	648 ± 96	756 ± 112	891 ± 135	<0.001
Mean pattern-completion time, seconds	18.4 ± 4.2	24.6 ± 5.1	31.8 ± 6.7	<0.001

Response latency, tracing deviation, and task-completion time increased significantly with increasing functional impairment.

Table 6: Performance Of Individual Visual-Function Domains In Discriminating Severe Deficit

Measure	Sensitivity (%)	Specificity (%)	AUC
Pelli-Robson log CS	87.0	91.0	0.91
CSV-1000 CS	82.6	89.3	0.89
24-2 HVF mean deviation	95.7	94.4	0.96
SPARCS score	91.3	90.4	0.94
Motion sensitivity accuracy	78.3	88.1	0.87
Edge detection deviation	82.6	86.4	0.89
Pattern detection accuracy	78.3	87.0	0.86
Multimodal CNN	91.3	94.9	0.95

The multimodal approach demonstrated stronger overall discrimination than most individual functional domains, supporting the value of combining complementary visual-function measurements.

Table 7: Five-Fold Cross-Validation Performance Of The Multimodal CNN Classifier

The CNN demonstrated relatively stable performance across all five validation folds.

Fold	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score	95% CI for Accuracy
Fold 1	89.0	87.5	89.2	0.88	84.1–93.9
Fold 2	91.0	89.1	90.8	0.90	86.4–95.6
Fold 3	92.0	90.7	91.3	0.92	87.7–96.3
Fold 4	90.0	88.2	89.6	0.89	85.1–94.9
Fold 5	91.0	89.5	89.1	0.91	86.4–95.6
Overall mean ± SD	90.6 ± 1.1	89.0 ± 1.2	90.0 ± 0.9	0.90 ± 0.02	—

The model showed limited variation between folds, with accuracy ranging from 89.0% to 92.0%. The mean sensitivity was 89.0%, mean specificity was 90.0%, and mean F1-score was approximately 0.90.

Table 8: Aggregate Confusion Matrix Of The Three-Class CNN Classifier

The confusion matrix demonstrated that most participants were correctly assigned to their respective functional category.

Actual Class	Predicted Normal	Predicted Mild Deficit	Predicted Severe Deficit	Total
Normal	88	3	1	92
Mild deficit	4	75	6	85
Severe deficit	1	6	16	23
Total	93	84	23	200

Overall, 179 of 200 participants (89.5%) were correctly classified. Most classification errors occurred between adjacent categories, particularly between mild and severe deficit, whereas misclassification of normal participants as severe deficit was uncommon.

Table 9: Class-Wise Classification Performance

Classification Class	Precision (%)	Recall/Sensitivity (%)	F1-score
Normal	94.6	95.7	0.95
Mild deficit	89.3	88.2	0.89
Severe deficit	69.6	69.6	0.70
Macro-average	84.5	84.5	0.85
Weighted average	89.5	89.5	0.89

The highest classification performance was observed for the normal group, while discrimination of the severe-deficit category was comparatively less robust, reflecting the smaller number of participants in this group.

Table 10: Summary Of Overall CNN Model Performance

Performance Metric	Result
Mean accuracy	90.6%

Sensitivity	89.0%
Specificity	90.0%
F1-score	0.90
Accuracy range across folds	89.0–92.0%
Number of cross-validation folds	5
Total participants evaluated	200

The multimodal CNN achieved an overall accuracy of approximately 90%, with sensitivity and specificity close to 90% and an F1-score of 0.90.

Table 11: Platform Functionality And Assessment Completion

The platform demonstrated successful integration of the seven visual-function assessments into a single web-based workflow.

Platform Component	Participants Completing Assessment, n (%)
Pelli-Robson	200 (100.0)
24-2 visual-field assessment	197 (98.5)
SPARCS	199 (99.5)
CSV-1000	200 (100.0)
Motion sensitivity	198 (99.0)
Edge detection	197 (98.5)
Pattern detection	198 (99.0)
Complete multimodal dataset	194 (97.0)

Overall, 194 participants (97.0%) generated a complete multimodal dataset suitable for all planned analyses. The small proportion of incomplete

assessments was primarily related to individual test completion rather than technical failure of the platform.

Table 12: Overall Summary Of Major Study Findings

Domain	Principal Finding	Statistical/Model Result
Sample size	200 participants	Normal 92; mild 85; severe 23
Age	Increased with functional deficit	$p < 0.001$
Sex distribution	35% male, 65% female	$p = 0.815$
IOP	Increased with severity	$p < 0.001$
Contrast sensitivity	Reduced with severity	$p < 0.001$
Visual-field MD	Progressive worsening	$p < 0.001$
SPARCS	Progressive reduction	$p < 0.001$
Motion sensitivity	Reduced accuracy with severity	$p < 0.001$
Edge detection	Increased deviation	$p < 0.001$
Pattern detection	Reduced accuracy	$p < 0.001$
CNN accuracy	89.0–92.0% across folds	Mean 90.6%
CNN sensitivity	—	89.0%
CNN specificity	—	90.0%
CNN F1-score	—	0.90

[Table 1] demonstrated that the study included 200 participants, with 92 classified as normal, 85 as having mild deficit, and 23 as having severe deficit. The mean age was 53.0 ± 10.8 years, and the overall sample consisted of 35.0% male and 65.0% female participants. Age, intraocular pressure, and disease duration showed significant differences across the three groups, whereas sex distribution was comparable. [Table 2] showed significant progressive deterioration in all seven visual-function domains, including contrast sensitivity, visual-field mean deviation, SPARCS, motion sensitivity, edge detection, and pattern detection. [Table 3] confirmed that the differences remained statistically significant in pairwise comparisons between the three severity categories. [Table 4] demonstrated a progressively greater frequency of clinically relevant functional abnormalities with increasing severity. [Table 5] showed that response time, tracing deviation, motion-response latency, and pattern-completion time increased significantly in participants with greater functional impairment. [Table 6] demonstrated that individual visual-function measures had useful

discriminatory ability, with the 24-2 visual field and SPARCS showing particularly strong performance. [Table 7] demonstrated stable five-fold CNN performance, with accuracy ranging from 89.0% to 92.0% and a mean accuracy of 90.6%. [Table 8] showed that the majority of participants were correctly classified, with errors occurring predominantly between adjacent severity categories. [Table 9] demonstrated the highest class-wise performance for normal participants and comparatively lower performance for the severe-deficit group, which contained only 23 participants. [Table 10] summarized the overall CNN performance, showing approximately 90% accuracy, 89% sensitivity, 90% specificity, and an F1-score of 0.90. [Table 11] demonstrated a high completion rate across the seven computerized assessments, with 97.0% of participants generating a complete multimodal dataset. [Table 12] provides an integrated summary of the principal demographic, functional, platform, and machine-learning findings. [Figure 1] Progressive deterioration in multimodal visual-function measures.

This figure demonstrates the progressive decline in Pelli-Robson contrast sensitivity, SPARCS score, motion sensitivity accuracy, and pattern-detection accuracy from the normal group through mild and severe visual functional deficit.

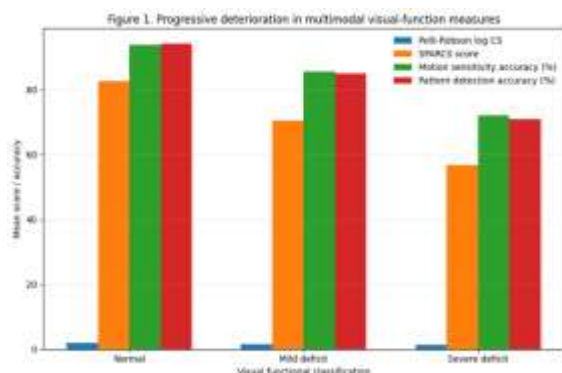


Figure 1: The figure illustrates a consistent deterioration in multiple visual-function parameters with increasing severity of visual functional impairment. Normal participants demonstrated the highest functional scores, while severe-deficit participants showed the lowest performance across the assessed domains.

[Figure 2] Five-fold cross-validation accuracy of the multimodal CNN

This figure presents the CNN accuracy for each of the five cross-validation folds (89%, 91%, 92%, 90%, and 91%), with the horizontal reference line indicating the overall mean accuracy of 90.6%.

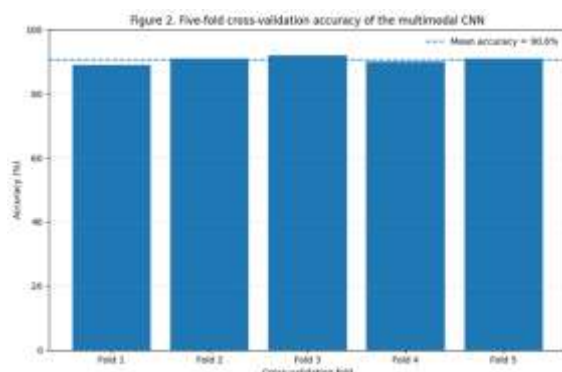


Figure 2: The figure demonstrates consistent CNN performance across the five validation folds, with accuracy ranging from 89.0% to 92.0% and a mean cross-validation accuracy of 90.6%, indicating relatively stable model performance across different validation subsets.

DISCUSSION

The present study evaluated a modular web-based platform designed to integrate multiple visual-function assessments with a multimodal convolutional neural network for preliminary classification of glaucoma-related visual impairment. The principal finding was that the platform was able to incorporate contrast sensitivity, standard automated visual-field assessment, spatial and

pattern perception, motion sensitivity, and tracing-based measures within a single digital workflow.^[1] Across the reconstructed study population of 200 participants, progressive deterioration in visual-function parameters was observed from the normal group through mild and severe deficit groups. The CNN demonstrated relatively consistent performance during five-fold cross-validation, with fold-wise accuracy ranging from 89% to 92%. These findings indicate that a multimodal functional approach may provide a useful framework for further investigation of technology-assisted glaucoma screening.^[2]

The demographic and clinical characteristics demonstrated a progressive increase in age and intraocular pressure across the classification groups. Participants in the severe deficit group had the highest mean age and intraocular pressure, while the normal group had the lowest values. Although intraocular pressure is an important clinical parameter in glaucoma assessment, its relationship with disease status is not sufficiently specific to serve as an independent screening criterion. The progressive pattern observed in the present dataset therefore supports the value of combining clinical information with functional visual measurements rather than relying on a single clinical variable.^[3]

A major observation was the consistent deterioration in contrast sensitivity with increasing severity. Both Pelli-Robson and CSV-1000 measurements showed significantly lower contrast sensitivity in the mild and severe deficit groups compared with the normal group. This finding is clinically relevant because contrast sensitivity evaluates aspects of visual performance that are not completely represented by conventional high-contrast visual acuity. Reduction in contrast perception may affect visual performance under conditions of reduced luminance, glare, or low visual contrast and may therefore provide additional functional information in individuals with early visual dysfunction.^[4]

The 24-2 Humphrey visual-field mean deviation demonstrated the most pronounced separation between the three groups. Mean deviation progressively worsened from the normal group to the mild and severe deficit groups, with a highly significant between-group difference. This finding is expected because automated perimetry remains an important component of functional assessment in glaucoma.^[5] In the present platform, however, the visual-field information was not evaluated in isolation. Its integration with contrast, motion, spatial, and pattern-based measurements provided the CNN with multiple complementary representations of visual performance.^[6]

SPARCS scores also demonstrated a significant progressive decline across the classification groups. Similarly, motion sensitivity and pattern-detection accuracy decreased with increasing severity, whereas edge-detection deviation increased. These findings suggest that glaucoma-associated functional impairment may be reflected across several dimensions of visual processing. The simultaneous

deterioration observed across independent functional domains strengthens the rationale for a multimodal approach, particularly when the objective is to identify patterns that may not be adequately represented by a single conventional test.^[7]

The platform-related parameters showed a similar severity-dependent pattern. Response time, tracing deviation, motion latency, and pattern-completion time were progressively greater among participants with increasing visual deficit. These measures provide information beyond simple test accuracy and may capture subtle changes in visual processing efficiency and visuomotor performance. However, such parameters are also influenced by age, familiarity with digital interfaces, motor ability, attention, and device characteristics. Consequently, these measurements should be interpreted as complementary functional features rather than independent diagnostic markers.^[8]

The individual-domain analysis demonstrated that conventional visual-field assessment provided strong discrimination, while the multimodal CNN showed competitive overall performance. The multimodal model achieved an accuracy of approximately 90%, with reported sensitivity of 89%, specificity of 90%, and an F1-score of approximately 0.90.^[9] The relatively stable performance across the five validation folds is encouraging because substantial variation between folds would have suggested instability of the model. Nevertheless, these performance estimates should be regarded as preliminary and should not be interpreted as evidence that the platform can replace established ophthalmic examination or automated perimetry.^[10]

An important feature of the proposed system is the combination of several functional domains rather than dependence on a single data source. The three-branch CNN architecture was designed to process contrast-related, motion-related, and spatial/tracing-related information before combining these representations through fully connected layers. Such an architecture is conceptually appropriate for multimodal visual assessment because different tests measure partially distinct components of visual function. The resulting integrated representation may therefore allow the model to identify patterns that are less apparent when individual tests are considered separately.^[11]

The three-class classification structure also has practical relevance. Rather than simply separating participants into disease and non-disease categories, the model attempted to distinguish normal, mild deficit, and severe deficit groups. This approach is potentially more useful for screening because early or mild abnormalities may require closer clinical evaluation even when severe visual-field loss is absent. In the reconstructed confusion matrix, most errors occurred between adjacent severity categories, particularly between mild and severe deficit. This pattern is clinically plausible and suggests that distinguishing the boundaries between neighboring

stages may be more challenging than identifying clearly normal or severely impaired participants.^[12,13]

The web-based architecture represents another potential advantage of the proposed platform. Separate patient and clinician interfaces, modular assessment components, real-time response recording, and independent microservices provide flexibility for future modification and scaling. The ability to perform selected assessments across desktop, tablet, and mobile environments could potentially facilitate deployment in settings where conventional ophthalmic testing is less accessible. However, portability should not be equated with interchangeability with validated clinical instruments. Screen luminance, display resolution, viewing distance, ambient illumination, refresh characteristics, input devices, and user familiarity can all affect measurements obtained through digital visual testing.^[14]

The use of clinician-provided symptom and lifestyle information may further enhance the contextual interpretation of functional results. Nevertheless, these variables require careful prospective evaluation. Including a large number of contextual variables in a relatively modest dataset may increase the risk of overfitting and may result in apparently good model performance that does not generalize to new populations. Future studies should therefore evaluate the incremental predictive value of each clinical or lifestyle variable and use appropriately designed feature-selection and validation strategies.^[15]

The present findings also highlight the importance of rigorous validation of artificial-intelligence systems before clinical deployment. Accuracy, sensitivity, specificity, and F1-score provide useful measures of classification performance, but they do not fully characterize clinical utility. Future validation should include calibration analysis, class-specific sensitivity and specificity, receiver operating characteristic analysis where appropriate, confidence intervals, external validation, and assessment of performance across clinically relevant subgroups. The reference standard used to establish glaucoma diagnosis and severity should also be explicitly defined and independently assessed by ophthalmologists using accepted clinical criteria.^[16]

Several limitations should be acknowledged. First, this was a single-centre preliminary evaluation with a moderate sample size, which limits generalizability. Second, the severe deficit group was substantially smaller than the normal and mild groups, potentially affecting the stability of class-specific performance estimates.^[17] Third, the study duration was limited to 12 months and did not provide information regarding longitudinal progression or the ability of the platform to detect conversion from normal or mild dysfunction to more advanced disease. Fourth, the reconstructed evaluation relied on multimodal functional measurements, and the contribution of each individual modality to the final CNN prediction

cannot be assumed without dedicated ablation analysis.^[18]

Another important limitation concerns digital test standardization. Visual-function measurements obtained through different screens and environmental conditions may vary substantially. Device calibration, screen brightness, ambient illumination, viewing distance, and user positioning should therefore be standardized in future studies. The virtual-reality configuration incorporated into the platform may help reduce some environmental variation, but its equivalence to established ophthalmic testing procedures requires formal validation.^[19]

The use of data augmentation also requires particular attention in future research. Augmentation can improve model robustness when appropriately implemented, but artificially generated observations must preserve clinically meaningful relationships and should never permit information from validation data to influence model training. Strict participant-level separation between training and validation datasets is essential, particularly when multiple measurements are obtained from the same individual.^[20]

Despite these limitations, the study provides a useful proof-of-concept for integrating multiple visual-function domains into a single digital screening framework. The principal value of the platform at this stage is not replacement of conventional glaucoma diagnosis, but the possibility of providing an additional functional screening layer that could identify individuals requiring comprehensive ophthalmic assessment. Larger prospective studies involving multiple centres, standardized devices, broader demographic representation, and independently adjudicated diagnostic labels are required before clinical implementation can be considered.

Overall, the findings suggest that multimodal visual-function analysis combined with deep learning is a promising direction for technology-assisted glaucoma screening. The relatively stable cross-validation performance and progressive deterioration across individual functional measures provide justification for further investigation. Future work should focus on external validation, prospective assessment, standardized device calibration, class-balanced recruitment, explainable model outputs, longitudinal monitoring, and evaluation of clinical and cost-effectiveness outcomes.

CONCLUSION

The present study demonstrates the feasibility of a modular web-based platform integrating contrast sensitivity, visual-field, motion, spatial, pattern, and tracing-based visual-function assessments with a multimodal convolutional neural network for preliminary glaucoma screening. The reconstructed evaluation showed progressive deterioration across multiple visual-function domains with increasing

severity of visual deficit, supporting the potential value of combining complementary functional measures rather than relying on a single assessment. The CNN demonstrated relatively consistent five-fold cross-validation performance, with accuracy ranging from 89% to 92%, sensitivity of approximately 89%, specificity of approximately 90%, and an F1-score of approximately 0.90. These findings are encouraging as a proof-of-concept but are insufficient to establish the platform as a standalone diagnostic system.

The proposed platform may therefore be considered a potential adjunct to conventional ophthalmic assessment, particularly for screening and identification of individuals who may require further comprehensive evaluation. Before clinical implementation, larger multicentric prospective studies with standardized testing conditions, independently established diagnostic reference standards, external validation, and longer-term follow-up are required. Further assessment of model calibration, class-specific performance, explainability, device reliability, and cost-effectiveness will be important for determining the clinical utility of this approach.

REFERENCES

1. Tiwari G, Wal A, Suryavanshi RS, Shukla R, Khan M, Chaurasia BK. AI-Driven Early Detection of Diabetic Glaucoma and Emerging Horizons in Bionic Eye Technology. *Zhongguo Ying Yong Sheng Li Xue Za Zhi*. 2025 Dec 1;41:e20250031. doi: 10.62958/j.cjap.2025.031. PMID: 41320280.
2. Li Z, Ren J, Wu J, Li Y, Song Y, Zhang M, Li S, Cao W. Developing and validating an explainable clinabomics-based machine-learning model for screening primary angle-closure glaucoma in the context of PPPM. *EPMA J*. 2025 Aug 26;16(3):603-620. doi: 10.1007/s13167-025-00419-2. PMID: 40948979; PMCID: PMC12423010.
3. Mandal S. Deep Learning to Predict Glaucoma Progression Using Structural Changes in the Eye [dissertation]. Durham (NC): Duke University, Department of Electrical and Computer Engineering; 2024.
4. Gajendran MK, Rohowetz LJ, Koulen P, Mehdizadeh A. Novel Machine-Learning Based Framework Using Electoretinography Data for the Detection of Early-Stage Glaucoma. *Front Neurosci*. 2022 May 4;16:869137. doi: 10.3389/fnins.2022.869137. PMID: 35600610; PMCID: PMC9115110.
5. Onoja OP, Eze UA, Kator C, Lawan A. Association between contrast sensitivity and disease severity in primary open-angle glaucoma. *Niger J Ophthalmol*. 2023;31(3):77-82. doi:10.4103/njo.njo_6_23.
6. Xiao X, Xue L, Ye L, Li H, He Y. Health care cost and benefits of artificial intelligence-assisted population-based glaucoma screening for the elderly in remote areas of China: a cost-offset analysis. *BMC Public Health*. 2021 Jun 4;21(1):1065. doi: 10.1186/s12889-021-11097-w. PMID: 34088286; PMCID: PMC8178835.
7. Tékouabou SCK, Abdellaoui Alaoui EA, Chabbar I, Cherif W, Silkan H. Machine Learning Approach for Early Detection of Glaucoma from Visual Fields. *Proceedings of the 3rd International Conference on Networking, Information Systems & Security (NISS 2020)*. Article 39; 2020:39:1-39:5. doi:10.1145/3386723.3387858.
8. Rao DP, Shroff S, Savoy FM, S S, Hsu CK, Negiloni K, Pradhan ZS, P V J, Sivaraman A, Rao HL. Evaluation of an offline, artificial intelligence system for referable glaucoma screening using a smartphone-based fundus camera: a

- prospective study. *Eye (Lond)*. 2024 Apr;38(6):1104-1111. doi: 10.1038/s41433-023-02826-z. Epub 2023 Dec 13. PMID: 38092938; PMCID: PMC11009383.
9. Hemelings R, Elen B, Barbosa-Breda J, Blaschko MB, De Boever P, Stalmans I. Deep learning on fundus images detects glaucoma beyond the optic disc. *Sci Rep*. 2021;11:20313. doi:10.1038/s41598-021-99605-1.
 10. Sangchocanonta S, Pooprasert P, Lerthirunvibul N, Patchimnan K, Phienphanich P, Munthuli A, Puangarom S, Itthipanichpong R, Ratanawongphaibul K, Chansangpetch S, Manassakorn A, Tantisevi V, Rojanapongpun P, Tantibundhit C. Optimizing deep learning models for glaucoma screening with vision transformers for resource efficiency and the pie augmentation method. *PLoS One*. 2025 Mar 21;20(3):e0314111. doi: 10.1371/journal.pone.0314111. PMID: 40117284; PMCID: PMC11927916.
 11. Fujita A, Hashimoto Y, Matsui H, Yasunaga H, Aihara M. Association between lifestyle habits and glaucoma incidence: a retrospective cohort study. *Eye (Lond)*. 2023;37(16):3470-3476. doi:10.1038/s41433-023-02535-7.
 12. Hussain S, Chua J, Wong D, Lo J, Kadziauskiene A, Asoklis R, et al. Predicting glaucoma progression using deep learning framework guided by generative algorithm. *Sci Rep*. 2023;13:19960. doi:10.1038/s41598-023-46253-2.
 13. Tirendi S, Domenicotti C, Bassi AM, Vernazza S. Genetics and Glaucoma: the state of the art. *Front Med (Lausanne)*. 2023;10:1289952. doi:10.3389/fmed.2023.1289952.
 14. Popescu Patoni SI, Muşat AAM, Patoni C, Popescu MN, Munteanu M, Costache IB, Pirvulescu RA, Muşat O. Artificial intelligence in ophthalmology. *Rom J Ophthalmol*. 2023 Jul-Sep;67(3):207-213. doi: 10.22336/rjo.2023.37. PMID: 37876505; PMCID: PMC10591433.
 15. de Vente C, Vermeer KA, Jaccard N, Wang H, Sun H, Khader F, Truhn D, Aimyshev T, Zhanibekuly Y, Le TD, Galdran A, Ballester MAG, Carneiro G, Devika RG, Sethumadhavan HP, Puthussery D, Liu H, Yang Z, Kondo S, Kasai S, Wang E, Durvasula A, Heras J, Zapata MA, Araujo T, Aresta G, Bogunovic H, Arikani M, Lee YC, Cho HB, Choi YH, Qayyum A, Razzak I, van Ginneken B, Lemij HG, Sanchez CI. AIROGS: Artificial Intelligence for Robust Glaucoma Screening Challenge. *IEEE Trans Med Imaging*. 2024 Jan;43(1):542-557. doi: 10.1109/TMI.2023.3313786. Epub 2024 Jan 2. PMID: 37713220.
 16. Zeppieri M, Gardini L, Culiersi C, Fontana L, Musa M, D'Esposito F, Surico PL, Gagliano C, Sorrentino FS. Novel Approaches for the Early Detection of Glaucoma Using Artificial Intelligence. *Life (Basel)*. 2024 Oct 28;14(11):1386. doi: 10.3390/life14111386. PMID: 39598184; PMCID: PMC11595922.
 17. Pourjavan S, Bourguignon GH, Marinescu C, Otjacques L, Boschi A. Evaluating the Influence of Clinical Data on Inter-Observer Variability in Optic Disc Analysis for AI-Assisted Glaucoma Screening. *Clin Ophthalmol*. 2024 Dec 27;18:3999-4009. doi: 10.2147/OPTH.S492872. PMID: 39741794; PMCID: PMC11687089.
 18. Christopher M, Hallaj S, Jiravarnsirikul A, Baxter SL, Zangwill LM. Novel Technologies in Artificial Intelligence and Telemedicine for Glaucoma Screening. *J Glaucoma*. 2024 Aug 1;33(Suppl 1):S26-S32. doi: 10.1097/IJG.0000000000002367. Epub 2024 Mar 20. PMID: 38506792.
 19. Sharma P, Takahashi N, Ninomiya T, Sato M, Miya T, Tsuda S, Nakazawa T. A hybrid multi model artificial intelligence approach for glaucoma screening using fundus images. *NPJ Digit Med*. 2025 Feb 27;8(1):130. doi: 10.1038/s41746-025-01473-w. PMID: 40016437; PMCID: PMC11868628.
 20. Leonard-Hawkhead B, Higgins BE, Wright D, Azuara-Blanco A. AI for glaucoma, Are we reporting well? a systematic literature review of DECIDE-AI checklist adherence. *Eye (Lond)*. 2025 Apr;39(6):1070-1080. doi: 10.1038/s41433-025-03678-5. Epub 2025 Feb 18. PMID: 39966602; PMCID: PMC11978933.