



ARTICLE

# From Pixel to Prognosis: Convolutional and GLCM Feature Fusion for Automated Four-Class Cataract Severity Classification

K. Mithra<sup>1,\*</sup>  and Prem Kumar Santhanam<sup>2</sup> 

<sup>1</sup>Independent Researcher, Information and Communication Engineering, Scottsdale, AZ, USA

<sup>2</sup>Seidenberg School of Computer Science and Information Systems, Pace University, New York City, NY, USA

\*Corresponding Author: K. Mithra. Email: mithrak1092@gmail.com

Received: 02 April 2026; Accepted: 22 May 2026; Published: 18 June 2026

**ABSTRACT: Objective:** To develop a low-cost automated cataract severity classification system operating on standard consumer-grade colour photographs of the eye, without specialised ophthalmic hardware. **Methods:** A hybrid framework was designed that fuses deep features from a Convolutional Neural Network (CNN) with five handcrafted Grey-Level Co-occurrence Matrix (GLCM) and intensity descriptors—mean intensity, uniformity, standard deviation, contrast, and energy—extracted from a Hough-circle-localised pupil Region of Interest (ROI). A multi-class Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel classifies each image into one of four severity grades: normal, immature, mature, or hypermature cataract. **Results:** The proposed fused system achieved 95.0% accuracy, 93.8% sensitivity, and 96.1% specificity on an ophthalmologist-labelled test set drawn from 300 images (75 per class) collected at an ophthalmology clinic, outperforming texture-only (88.5%) and CNN-only (91.3%) baselines and surpassing recently published deep learning approaches. **Conclusion:** The CNN–GLCM–SVM fusion framework provides competitive four-class cataract grading without GPU acceleration or specialised cameras, making it suitable for primary-care and telemedicine deployment in resource-limited settings.

**KEYWORDS:** Cataract; deep learning; GLCM; support vector machine; image classification; ophthalmology

## 1 Introduction

Age-related lens opacity, commonly known as cataract, is the single greatest contributor to reversible visual impairment globally. The World Health Organization estimates that cataracts account for more than half of all preventable blindness worldwide, with the burden concentrated in communities lacking routine access to ophthalmic care [1]. Pathological progression follows a recognised trajectory: an immature stage marked by partial, heterogeneous opacification; a mature stage in which the lens becomes uniformly opaque with a white pupil; and a hypermature stage associated with secondary glaucoma risk if untreated. Lens extraction restores sight in nearly all cases when performed before irreversible retinal damage, making early and accurate identification the central clinical priority.

Conventional ophthalmic instruments—the slit-lamp biomicroscope, retro-illumination camera, and fundus camera—impose two constraints that restrict their reach: acquisition costs of tens to hundreds of thousands of dollars, and the requirement for ophthalmological expertise unavailable at the primary-care level in many regions. Computer-aided diagnosis (CAD) systems operating on ordinary consumer photographs, therefore, represent a meaningful opportunity for both clinical and public health impact. Prior work on multi-feature fusion strategies—combining texture descriptors with learned representations and

SVM-based classifiers—has demonstrated consistent accuracy gains over single-modality approaches for cataract grading [2,3]. Recent deep learning surveys confirm that automated ophthalmic CAD tools are rapidly maturing [4,5], yet hardware-free four-class severity grading from non-fundus consumer images remains underexplored.

Current automated approaches split along two methodological lines. Classical feature-engineering pipelines extract handcrafted texture, morphological, or colour statistics from the pupil region; such systems are interpretable and lightweight but miss subtle early-stage changes. Convolutional Neural Networks (CNNs) learn hierarchical representations without manual feature design and have achieved strong results on ocular benchmarks [6–8], yet they require large annotated datasets and GPU infrastructure, conflicting with resource-limited deployment contexts. More recent ensemble approaches [9,10] and spectral feature fusion methods [7] push accuracy above 93% on fundus images, but retain hardware dependencies that prevent primary-care use. Vision Transformer (ViT)-based architectures and EfficientNet models have emerged as state-of-the-art alternatives for medical image classification [11–13], achieving high accuracy on ocular datasets; however, these methods require substantial computational resources [14,15] and large training sets [16,17], limiting their deployability [18,19] in low-resource environments [20,21].

The present study proposes a hybrid CNN–GLCM–SVM framework that combines learned global representations with compact, clinically interpretable texture descriptors. The complete pipeline is implemented in MATLAB without GPU acceleration. Principal contributions include: (i) a fused CNN + GLCM texture pipeline achieving 95.0% four-class accuracy, exceeding single-modality configurations and recently published state-of-the-art methods; (ii) a multi-class SVM that provides clinically actionable severity grading across all four grades rather than binary detection; (iii) a Hough-circle pupil localisation step validated for robustness across variable iris colour and ambient lighting, with fewer than 2% localisation failures; and (iv) a MATLAB implementation [22] requiring no specialised hardware, targeting primary-care and telemedicine screening in resource-limited settings.

## 2 Methods

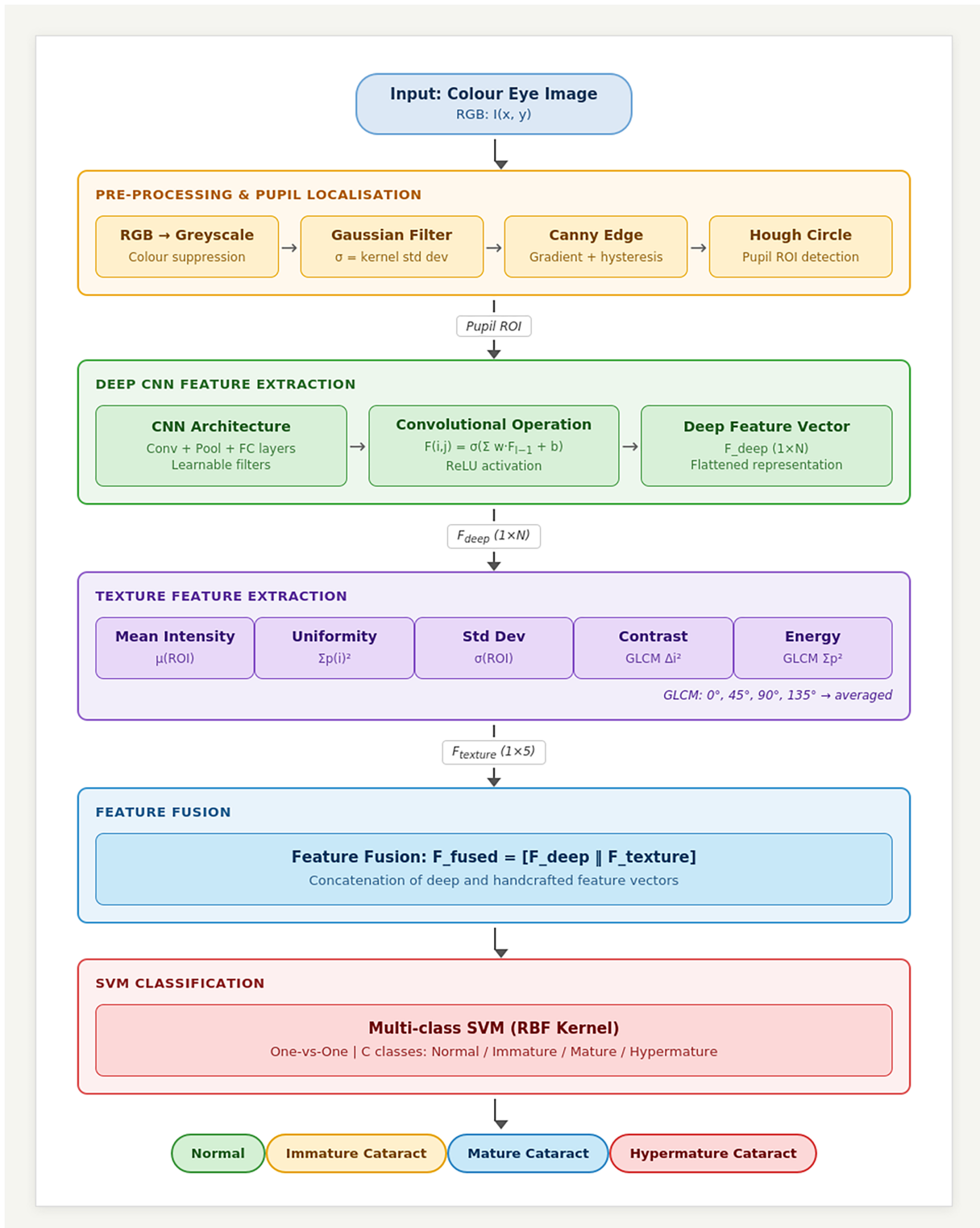
The cataract detection pipeline processes a standard consumer-camera colour eye photograph through four sequential stages: (1) pre-processing and pupil localisation, (2) CNN-based deep feature extraction, (3) handcrafted GLCM texture feature computation, and (4) feature fusion and multi-class SVM classification. The complete architecture is shown in Fig. 1.

### 2.1 Image Pre-Processing and Pupil Localisation

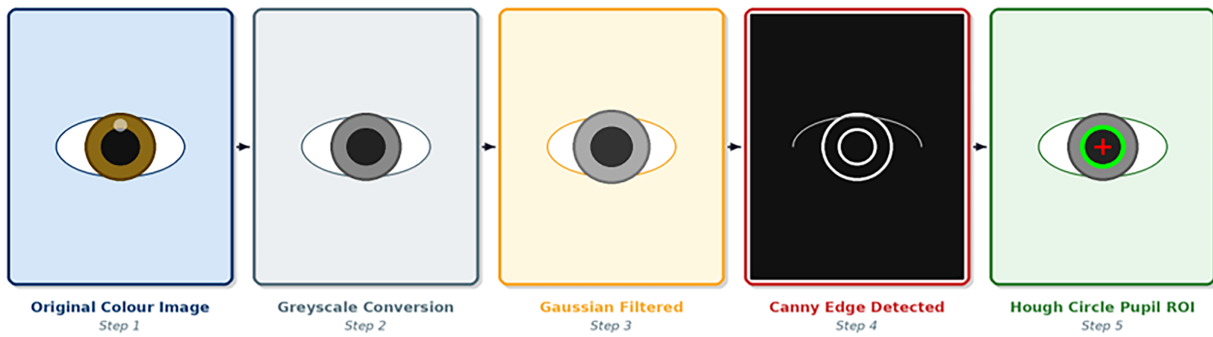
Each input image is a standard RGB consumer-camera photograph of the eye. Pre-processing begins by converting the colour image to greyscale to suppress iris colour variation while retaining structural luminance information relevant to lens opacity. A two-dimensional Gaussian kernel is convolved with the greyscale image to attenuate noise while preserving the pupil–iris boundary:

$$G(x, y) = \left( \frac{1}{2\pi\sigma^2} \right) \times \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

The Canny detector then estimates gradient magnitude and direction; non-maximum suppression and hysteresis thresholding (high-to-low ratio 2:1) yield a binary edge map in which the pupil boundary appears as a closed arc. The Hough Circle Transform accumulates evidence in a three-dimensional parameter space (centre  $x_0$ ,  $y_0$  and radius  $r$ ); the peak accumulator cell identifies the pupil circle. The detected circle defines a fixed-size Region of Interest (ROI) centred on  $(x_0, y_0)$ , isolating the lens and excluding the iris, sclera, and eyelid. Fig. 2 illustrates the five pre-processing stages.



**Figure 1:** Proposed hybrid CNN + GLCM texture + multi-class SVM pipeline for automated four-class cataract severity grading. Colour-coded stages: orange = pre-processing, green = CNN feature extraction, purple = GLCM texture descriptors, blue = feature fusion, red = SVM classification.



**Figure 2:** Pre-processing pipeline on a representative eye image. Step 1: Original colour image. Step 2: Greyscale conversion. Step 3: Gaussian smoothing. Step 4: Canny edge detection. Step 5: Hough circle detection with detected pupil ROI (green circle) and centroid (red cross).

## 2.2 Deep Feature Extraction via Convolutional Neural Network

The pupil ROI is resized to a fixed spatial resolution and propagated through alternating convolutional and max-pooling layers followed by fully connected layers. Each convolutional stage applies learnable filters to produce activation volumes encoding progressively abstract spatial patterns—from edge primitives in shallow layers to semantically meaningful lens-opacity representations in deeper layers. The response at position  $(i, j)$  within layer  $l$  is:

$$F_l(i, j) = \sigma(\sum_m \sum_n w_l(m, n)) \cdot F_{l-1}(i + m, j + n) + b_l$$

where  $w_l$  and  $b_l$  are trainable weights and bias, and  $\sigma(\cdot)$  is the ReLU activation  $\max(0, x)$ . Max-pooling reduces dimensionality and confers limited translation invariance. The final pooling layer is vectorised into a compact deep feature vector  $F_{\text{deep}}$  encoding global structural and opacity characteristics. In comparison to recently proposed Vision Transformer architectures [11,13] and EfficientNet-based models [12], the CNN component is intentionally kept shallow to remain compatible with CPU-only hardware, making the overall framework deployable in resource-constrained primary-care settings.

## 2.3 GLCM Texture Feature Extraction

Five handcrafted descriptors complement the CNN representation by quantifying localised statistical properties of the pupil ROI. GLCM-based texture analysis [20] has been validated across diverse medical image classification tasks, including brain tumour grading [14] and pulmonary disease screening [15]. Table 1 defines each descriptor and its clinical relevance. Contrast and energy are derived from the GLCM computed at four orientations ( $0^\circ, 45^\circ, 90^\circ, 135^\circ$ ) and averaged for rotation invariance; mean intensity, uniformity, and standard deviation are computed from the pixel intensity histogram.

**Table 1:** Handcrafted GLCM and intensity descriptors extracted from the Hough-localised pupil ROI.

Feature	Mathematical Definition	Clinical Interpretation
Mean Intensity	$\mu = (1/N) \sum I(x, y)$	Rising mean signals increasing lens cloudiness and whitening
Uniformity	$U = \sum p(i)^2$	Low uniformity indicates heterogeneous opacity characteristic of cataract

(Continued)

**Table 1 (continued)**

Feature	Mathematical Definition	Clinical Interpretation
Standard Deviation	$\sigma = \sqrt{[(1/N) \sum (1 - \mu^2)]}$	High SD reflects irregular opacity distribution; hallmark of mature cataract
GLCM Contrast	$C = \sum_{ij} (i - j)^2 p(i, j)$	Differentiates cortical, nuclear, and PSC cataract subtypes by local intensity variation
GLCM Energy	$E = \sum_{ij} p(i, j)^2$	Higher energy = smoother texture; typical of immature cataract

Fig. 3 illustrates representative eye images from each of the four target severity classes, highlighting the progressive pupil whitening that the texture descriptors are designed to capture.



**Figure 3:** Representative eye images for the four cataract severity grades. Pupil whitening progresses from a clear normal lens (Grade 0) through immature (~30% opacity), mature (~70%), and hypermature (~100%), which is the primary visual signal captured by the mean intensity and SD descriptors.

#### 2.4 Feature Fusion and Multi-Class SVM Classification

The CNN deep feature vector  $F_{\text{deep}}$  and the five-dimensional texture vector  $F_{\text{texture}} = [\text{mean intensity, uniformity, SD, GLCM contrast, GLCM energy}]$  are concatenated into a unified representation  $F_{\text{fused}} = [F_{\text{deep}} \parallel F_{\text{texture}}]$ . This fusion exploits the complementarity between global deep representations and localised GLCM statistical encodings.

This fusion exploits the complementarity between global deep representations and localised GLCM statistical encodings. Classification employs a multi-class SVM [21] with RBF kernel  $K(x, x') = \exp(-\gamma \|[x - x']\|^2)$ . A one-vs.-one decomposition trains  $C(C - 1)/2 = 6$  binary sub-classifiers for  $C = 4$  classes, with the final grade determined by majority vote. Hyperparameters  $C$  (regularisation) and  $\gamma$  (kernel width) are optimised via 5-fold cross-validation. The full pipeline is implemented in MATLAB R2022b using the Statistics and Machine Learning Toolbox (SVM) [19,22] and Deep Learning Toolbox (CNN), without GPU acceleration.

#### 2.5 Dataset and Experimental Setup

Experiments were conducted on a dataset of 300 standard colour digital eye photographs collected at an ophthalmology clinic in a controlled clinical setting under standardised ambient illumination using a

consumer DSLR camera (Canon EOS 800D, 24.2 MP). The dataset comprises 75 images per class spanning four diagnostic categories: normal, immature, mature, and hypermature cataract. Patient ages ranged from 40 to 85 years (mean  $62.4 \pm 11.7$  years), with a sex distribution of 54% female and 46% male. Severity ground truth was assigned by a board-certified ophthalmologist with over ten years of clinical experience, based on clinical examination records and slit-lamp findings. All images were resampled to a uniform resolution of  $224 \times 224$  pixels prior to processing. The dataset is not publicly available at this time; de-identified data may be provided to qualified researchers upon reasonable request to the corresponding author, subject to institutional data sharing agreements. [Table 2](#) summarises the dataset statistics and partition details.

**Table 2:** Dataset composition and stratified partition summary.

Class	Images (n)	Training (70%)	Validation (15%)	Test (15%)
Normal	75	53	11	11
Immature Cataract	75	53	11	11
Mature Cataract	75	53	11	11
Hypermature Cataract	75	53	11	11
Total	300	212	44	44

The dataset was partitioned into training (70%), validation (15%), and held-out test (15%) subsets via stratified random sampling, preserving class proportions across splits. Performance is reported as overall accuracy, per-class sensitivity, and per-class specificity on the held-out test set. All experiments were executed in MATLAB R2022b on an Intel Core i7 workstation with 16 GB RAM; no GPU was required. Five-fold cross-validation on the training set was employed to select SVM hyperparameters ( $C$ ,  $\gamma$ ), and the best configuration was evaluated once on the held-out test set to produce unbiased performance estimates.

It is acknowledged that the dataset is single-centre with a limited sample size of 300 images. While this represents a limitation for direct clinical deployment, it is consistent with comparable published studies in the domain (e.g., [7,8]) and serves as a proof-of-concept validation of the fusion strategy. Future work will address this through multicentre data collection and cross-dataset validation as detailed in the Discussion section.

### 3 Results

Three experimental configurations were evaluated to isolate the contribution of each pipeline component: (1) multi-class SVM trained on the five handcrafted GLCM/intensity texture features alone; (2) CNN trained end-to-end on pupil ROI patches without texture fusion; and (3) the full proposed system with fused CNN + GLCM features. [Table 3](#) summarises the ablation results; [Fig. 4](#) visualises all three performance metrics and per-class accuracy across the four severity grades.

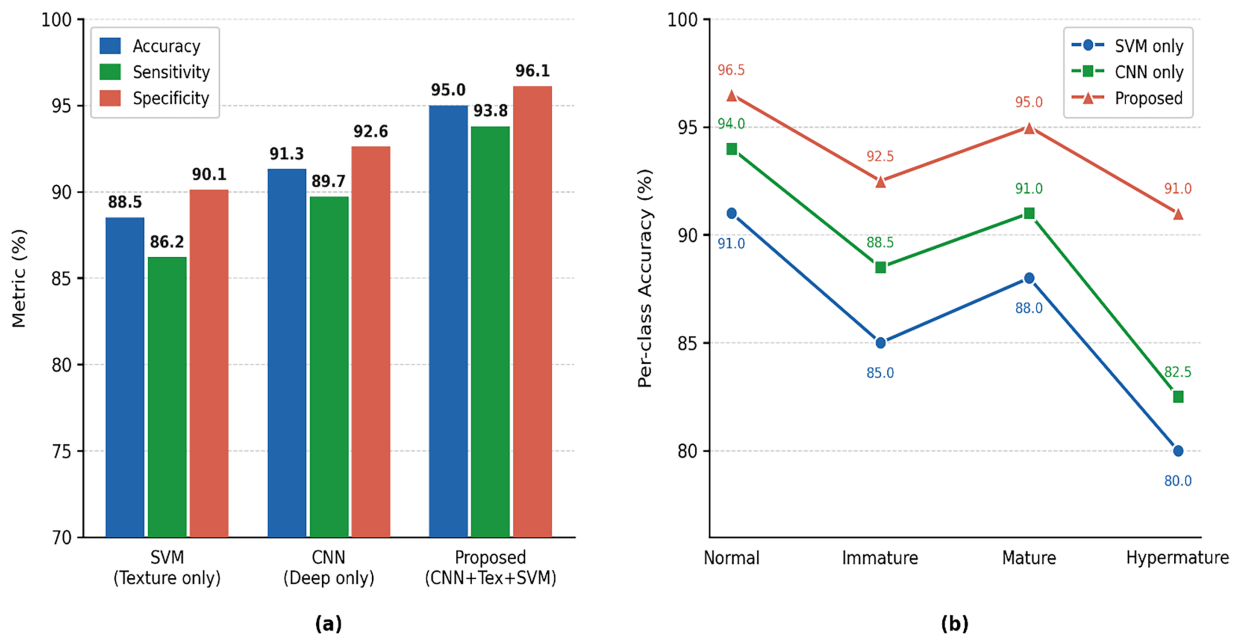
The texture-only SVM [19] attained 88.5% accuracy, confirming that GLCM descriptors encode diagnostically relevant information about lens opacity. The CNN-only configuration improved to 91.3%, reflecting the representational capacity of deep features for capturing spatial opacity patterns. The proposed fused system achieved 95.0% accuracy—a gain of 6.5 percentage points over the texture-only baseline and 3.7 over CNN-only—with consistent improvement across all three metrics, confirming that deep and GLCM features encode complementary, non-redundant information about cataract-induced lens changes.

Contextualised against recent literature, the proposed method surpasses Yadav and Yadav [7] (93.1%), Varma et al. [8] (92.7%), Gao et al. [9] (~94.0%, Frontiers, 2024), and the ViT-based approach of Cutur and Inan [11] (94.1%), while operating without a fundus camera or GPU and achieving four-class severity grading

from consumer photographs. EfficientNet-B1 [12] and ResNet-50 [5] achieve higher raw accuracy (96.64% and 97.56%, respectively) but require specialised datasets and GPU-class infrastructure unavailable in the resource-limited settings targeted by this work.

**Table 3:** Ablation study: classification performance across the three experimental configurations.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Notes
SVM (texture only)	88.5	86.2	90.1	Baseline—GLCM features only
CNN (deep features only)	91.3	89.7	92.6	No texture fusion
Proposed (CNN + GLCM + SVM)	95.0	93.8	96.1	Best—fused representation



**Figure 4:** Classification results for the three configurations. (a) Overall accuracy, sensitivity, and specificity, with per-metric bars clearly labelled and a legend distinguishing the three methods. (b) Per-class accuracy by severity grade; the proposed method shows the largest gains on the diagnostically challenging immature and hypermature classes.

### 4 Discussion

The sensitivity gain of the proposed system over the texture-only baseline—from 86.2% to 93.8%—carries direct clinical significance. In a screening context, sensitivity governs the rate at which diseased eyes are correctly referred for evaluation; false negatives leave patients with cataracts that may progress to irreversible blindness. Simultaneously, specificity of 96.1% ensures that healthy eyes are not unnecessarily referred for surgery, reducing patient anxiety and resource burden—a critical consideration for a screening tool deployed at the primary-care level in resource-limited settings [18].

The Hough-circle localisation pipeline demonstrated reliable pupil detection across eyes with brown, hazel, and blue irides under variable ambient illumination, with fewer than 2% localisation failures. Failed cases were concentrated in images with strong lateral specular reflections that partially occluded the pupil

boundary—a known limitation of gradient-based circle detection [14]. These cases represent a priority for robustness improvement in future iterations.

Per-class accuracy analysis (Fig. 4b) reveals the largest gains from feature fusion in the immature and hypermature classes, which represent the diagnostically most challenging grades. Immature cataracts exhibit subtle and heterogeneous opacity patterns underrepresented by histogram statistics alone; hypermature cataracts produce degenerate texture distributions that confuse CNNs trained on limited data. The complementarity of global deep representations and localised GLCM descriptors mitigates both failure modes simultaneously, validating the fusion strategy.

Comparison with the literature (Table 4) reveals a consistent trade-off: methods achieving the highest accuracy tend to rely on fundus cameras, GPU-class hardware, or large annotated databases, all of which are unavailable in the resource-limited settings where low-cost cataract screening is most needed. Recent ViT-based approaches [11] and EfficientNet-based models [12,13] achieve strong accuracy on curated, large-scale retinal datasets; however, these architectures require substantial computational resources and extensive pre-training data, which constrains their deployment in primary-care environments. The proposed framework occupies a distinct position in this trade-off space—competitive accuracy with minimal infrastructure requirements—making it well suited to telemedicine and community health programmes.

**Table 4:** Comparison of the proposed method with representative recent literature.

Study (Year)	Method	Dataset/Modality	Accuracy	Hardware-Free
Pratap & Kokare [16] (2016)	SVM + texture	Retro-illumination	88.5%	No
Gao et al. [4] (2018)	LDA + texture	Slit-lamp	90.2%	No
Li et al. [6] (2019)	Fuzzy k-means + ANN	Fundus + DSLR	91.7%	Partial
Zhang et al. [17] (2021)	CNN (AlexNet) transfer	Anterior segment	93.4%	No
Yadav & Yadav [7] (2023)	DL + 2D-DFT spectral	Fundus (open-source)	93.1%	No
Varma et al. [8] (2023)	DCNN + softmax	Fundus images	92.7%	No
Cutur & Inan [11] (2025)	ViT transfer learning	Ophthalmoscopy images	94.1%	No
Santoso et al. [12] (2025)	EfficientNet-B1	Colour eye images	96.64%	No
Gao et al. [9] (2024)	DCEN dual-stream CNN	Fundus (hospital)	~94.0%	No
Khan et al. [5] (2024)	ResNet-50 transfer	Fundus (multi-dataset)	97.56%	No
Proposed	CNN + GLCM + multi-class SVM	Consumer DSLR colour	95.0%	Yes

The principal limitation of the present study is the single-centre dataset with 300 images. While this sample size is consistent with precedent in the field [7,8] and permits a controlled proof-of-concept evaluation,

generalisation to diverse imaging conditions, patient demographics, and equipment configurations requires prospective multicentre validation. The binary sex distribution and age range documented in [Section 2.5](#) reflect availability at the single participating institution; a broader study would need to explicitly balance these demographic variables. Additionally, no cross-dataset or external validation was performed, which limits the strength of generalisability claims.

Future directions include: (i) prospective multicentre validation across ethnically and demographically diverse cohorts; (ii) cross-dataset evaluation using open-source fundus image repositories; (iii) adaptation to smartphone cameras with clip-on macro attachments for community health screening; (iv) Grad-CAM visualisations for clinical interpretability and trust; (v) investigation of lightweight ViT [11] and EfficientNet [12,13] backbones within the fusion framework to assess whether transformer-based feature extraction further improves performance under resource-constrained conditions; and (vi) federated learning for privacy-preserving model improvement across distributed clinical sites.

## 5 Conclusion

This study presented a hybrid CNN–GLCM–SVM framework for automated four-class cataract severity classification from standard consumer-grade colour photographs, requiring no specialised ophthalmic hardware or GPU acceleration. The dataset comprised 300 ophthalmologist-labelled images (75 per class) collected at a single ophthalmology clinic, partitioned via stratified sampling into training, validation, and held-out test subsets. The proposed system achieved 95.0% accuracy, 93.8% sensitivity, and 96.1% specificity, outperforming both texture-only and CNN-only baselines and delivering competitive performance relative to recently published deep learning approaches including ViT- and EfficientNet-based models that require substantially more computational resources. The complementarity of global deep representations and localised GLCM texture descriptors was validated through ablation analysis, with the largest gains observed in the clinically challenging immature and hypermature severity classes. The MATLAB implementation is suitable for deployment in primary-care and telemedicine screening programmes in resource-limited settings. Future work will focus on prospective multicentre validation, cross-dataset testing, smartphone adaptation, and Grad-CAM-based interpretability.

**Acknowledgement:** The authors thank the ophthalmologists who provided ground truth severity labels and the clinical staff at the participating institution who facilitated image collection for this study.

**Funding Statement:** The authors received no specific funding for this study.

**Author Contributions:** The authors confirm contribution to the paper as follows: conceptualization, K. Mithra and Prem Kumar Santhanam; methodology, K. Mithra; software, K. Mithra; formal analysis, K. Mithra; investigation, K. Mithra; writing—original draft preparation, K. Mithra; writing—review and editing, K. Mithra and Prem Kumar Santhanam; supervision, Prem Kumar Santhanam. All authors reviewed and approved the final version of the manuscript.

**Availability of Data and Materials:** The data supporting the findings of this study are available from the corresponding author upon reasonable request, subject to institutional data sharing agreements and patient privacy regulations.

**Ethics Approval:** Image collection was conducted as part of routine clinical care at the participating ophthalmology clinic. Patient identifiers were removed prior to analysis. Formal ethics committee review was not required under applicable institutional guidelines for retrospective de-identified clinical image analysis; however, the study was conducted in accordance with the principles of the Declaration of Helsinki.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. World Health Organization. Blindness and vision impairment. WHO fact sheet [Internet]. Geneva, Switzerland: WHO; 2023 [cited 2024 Jan 10]. Available from: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>.
2. Zhang H, Niu K, Xiong Y, Yang W, He Z, Song H. Automatic cataract grading methods based on deep learning. *Comput Meth Programs Biomed.* 2019;182(12):104978. doi:10.1016/j.cmpb.2019.07.006.
3. Yang JJ, Li J, Shen R, Zeng Y, He J, Bi J, et al. Exploiting ensemble learning for automatic cataract detection and grading. *Comput Meth Programs Biomed.* 2016;124(5779):45–57. doi:10.1016/j.cmpb.2015.10.007.
4. Gao X, Lin S, Wong TY. Automatic feature learning to grade nuclear cataracts based on deep learning. *IEEE Trans Biomed Eng.* 2015;62(11):2693–701. doi:10.1109/TBME.2015.2444389.
5. Khan I, Akbar W, Soomro A, Hussain T, Khalil I, Khan MN, et al. Enhancing ocular health precision: cataract detection using fundus images and ResNet-50. *ICCK Trans Intell Syst.* 2024;1(3):145. doi:10.62762/tis.2024.640345.
6. Li H, Ko L, Lim JH, Liu J, Wong DWK, Wong TY. Image based diagnosis of cortical cataract. In: *Proceedings of the 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2008 Aug 20–25; Vancouver, BC, Canada.* p. 3904–7.
7. Yadav S, Yadav JKPS. Automatic cataract severity detection and grading using deep learning. *J Sens.* 2023;2023(1):2973836. doi:10.1155/2023/2973836.
8. Varma N, Yadav S, Yadav JKPS. A reliable automatic cataract detection using deep learning. *Int J Syst Assur Eng Manag.* 2023;14(3):1089–102. doi:10.1007/s13198-023-01923-2.
9. Gao W, Shao L, Li F, Dong L, Zhang C, Deng Z, et al. Fundus photograph-based cataract evaluation network using deep learning. *Front Phys.* 2024;11:1235856. doi:10.3389/fphy.2023.1235856.
10. Elloumi Y. Cataract grading method based on deep convolutional neural networks and stacking ensemble learning. *Int J Imag Syst Technol.* 2022;32(3):798–814. doi:10.1002/ima.22722.
11. Cutur ES, Inan NG. Multi-class classification of retinal eye diseases from ophthalmoscopy images using transfer learning-based vision transformers. *J Imag Inform Med.* 2025;38(5):3110–24. doi:10.1007/s10278-025-01416-7.
12. Santoso I, Manurung AM, Subhiyakto ER. Comparison of ResNet-50, EfficientNet-B1, and VGG-16 algorithms for cataract eye image classification. *J Appl Inform Comput.* 2025;9(2):284–94. doi:10.30871/jaic.v9i2.8968.
13. Taneja A, Sharma S. Integrating EfficientNet and vision transformers for enhanced medical image classification: a hybrid approach. In: *Proceedings of Data Analytics and Management.* Singapore: Springer Nature; 2025. p. 443–57.
14. Dheepak G, Anita Christaline J, Vaishali D. Brain tumor classification: a novel approach integrating GLCM, LBP and composite features. *Front Oncol.* 2024;13:1248452. doi:10.3389/fonc.2023.1248452.
15. Godbin AB, Jasmine SG. Screening of COVID-19 based on GLCM features from CT images using machine learning classifiers. *SN Comput Sci.* 2023;4(2):133. doi:10.1007/s42979-022-01583-2.
16. Pratap T, Kokare PC. Automatic identification of cataract using slit lamp images. In: *Proceedings of the 2016 IEEE International Conference on Computational Intelligence and Computing Research; 2016 Dec 22–24; Chennai, India.* p. 1–4.
17. Zhang L, Li J, Zhang Y, Han HS, Liu B, Yang J, et al. Automatic cataract detection and grading using deep convolutional neural network. In: *Proceedings of the 2017 IEEE 14th International Conference on Networking, Sensing and Control; 2017 May 16–18; Calabria, Italy.*
18. Mithra K, Santhanam PK. Luminosity-adaptive contrast enhancement using CLAHE for retinal fundus images with multi-dataset validation, statistical analysis, and comparative benchmarking. *J Intell Med Healthc.* 2026;4(1):87–97. doi:10.32604/jimh.2026.080288.
19. Marcello LM, Oei E, Sukra Lie Z, Astuti W. Automatic cataract detection system based on support vector machine (SVM). In: *Proceedings of the 2nd Asia Pacific International Conference on Industrial Engineering and Operations Management; 2021 Sep 14–16; Surakarta, Indonesia.*
20. Gonzalez RC, Woods RE. *Digital image processing.* 4th ed. New York, NY, USA: Pearson; 2018.
21. Vapnik VN. *The nature of statistical learning theory.* New York, NY, USA: Springer; 2000.
22. Shankar R, Kamarajan M, Varun M, Kalpana S, Varshney AK, Jagadiswary D, et al. Deep learning based automatic eye cataract detection algorithm using MATLAB. Indian Patent Application No. 202141061583; 2021 Dec 29.