Hybrid Handcrafted Feature Extraction Combined with an Assisted Region of Interest Protocol for Multi-Class Age-Related Macular Degeneration Classification

Leonardo Petra Refialy
Department of Computer Sciences and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia | Department of Informatics, Universitas Kristen Indonesia Maluku, Maluku, Indonesia
leonardopetrarefialy@mail.ugm.ac.id

Sri Hartati
Department of Computer Sciences and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia
shartati@ugm.ac.id (corresponding author)

Sigit Priyanta
Department of Computer Sciences and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia
seagatejogja@ugm.ac.id

Supanji Supanji
Department of Ophthalmology, Universitas Gadjah Mada, Yogyakarta, Indonesia
supanji@ugm.ac.id

Mohammad Eko Prayogo
Department of Ophthalmology, Universitas Gadjah Mada, Yogyakarta, Indonesia
mohammad.eko.prayogo@ugm.ac.id

Firman Setya Wardhana
Department of Ophthalmology, Universitas Gadjah Mada, Yogyakarta, Indonesia
firman.sw@ugm.ac.id

Received: 1 January 2026 | Revised: 6 February 2026 | Accepted: 13 February 2026

 DOI: https://doi.org/10.48084/etasr.17293

ABSTRACT

Age-related Macular Degeneration (AMD) is a major cause of vision impairment and may progress to blindness; thus, accurate differentiation between Dry and Wet AMD is essential for timely clinical decision-making. This study proposes a lightweight multi-class AMD classification framework from fundus images that combines an anatomically consistent Assisted Region of Interest (Assisted-ROI) protocol with a hybrid handcrafted representation. Four public datasets (ODIR, ADAM, FIVES, and RFMiD) were merged to construct a curated dataset of 597 images across three classes (Normal, Dry AMD, and Wet AMD). AMD cases were re-annotated into Dry and Wet subtypes by three ophthalmologists, while Normal labels were retained. The experiments used 5-fold cross-validation under three input scenarios: full-fundus (no ROI), Optic Disc (OD)-guided automatic ROI, and the proposed Assisted-ROI. The no-ROI baseline achieved 80.91% accuracy, the OD-guided ROI achieved 85.27%, and Assisted-ROI delivered the best performance with 90.63% accuracy. Feature ablation under identical controlled settings showed that early fusion of LBP-RIU, Haralick/GLCM texture features, and RGB color moments yielded the most discriminative representation (90.63% accuracy; 89.06% macro-F1), outperforming the best single descriptor (LBP-RIU: 83.09%). Overall, anatomically consistent ROI standardization and complementary

handcrafted feature fusion improve robust multi-class AMD recognition on heterogeneous multi-source fundus data under reduced computational cost.

Keywords-AMD classification; assisted ROI; handcrafted features fusion; LBP_RIU; GLCM; color moments; SVM

I. INTRODUCTION

Age-related Macular Degeneration (AMD) is a degenerative retinal disease and a leading cause of irreversible vision impairment that may progress to blindness in older adults. The global prevalence is projected to increase from 196 million in 2020 to 288 million by 2040 [1]. Clinically, AMD is commonly categorized into non-exudative (Dry) AMD, characterized by drusen accumulation in the macular region, and exudative (Wet) AMD, which involves choroidal neovascularization and retinal hemorrhage [2]. Early detection and reliable differentiation between these subtypes are essential because Wet AMD often requires urgent intervention (e.g., anti-VEGF therapy) to reduce the risk of permanent central vision loss [3]. Despite this urgency, routine screening and subtype assessment from color fundus images still rely heavily on expert visual grading, which is time-consuming and may introduce inter-observer variability and subjectivity [4]. In large-scale screening programs, manual workload can further lead to grader fatigue, while service demand continues to exceed the availability of retinal specialists [5]. To mitigate these constraints, Computer-Aided Diagnosis (CAD) has been extensively explored for retinal screening. Recent surveys indicate that deep learning has become the dominant paradigm in ophthalmic image classification, among others, due to its strong diagnostic performance [6], yet its deployment is still challenged by limited labeled data and substantial variability in real-world imaging conditions. In fundus imaging, deep networks have demonstrated promising results for automated retinal disease detection [7], and several studies have reported AMD grading and detection using convolutional architectures and transfer learning schemes [8, 9]. Beyond CNN-only models, hybrid architectures that combine CNN feature extraction with recurrent modules (e.g., LSTM) can outperform classical pipelines and CNN baselines in comparative evaluations [10]. However, CNN-LSTM frameworks are often computationally demanding and typically require large, well-curated labeled datasets, which may limit practicality when training data are scarce or heterogeneous. Feature-engineered machine-learning pipelines remain relevant as lightweight and more transparent alternatives. Nevertheless, handcrafted approaches often struggle with robustness and generalization on heterogeneous multi-source fundus images, where variations in optics, resolution, illumination, and disease presentation can alter macula-centered appearance and reduce cross-dataset consistency. Regarding descriptor suitability, in [11], shape-oriented representations (Hu moments) were employed, which may be suboptimal for AMD lesions that are largely amorphous and texture-dominant (drusen or exudates). Similarly, in [12], PHOG with entropy/nonlinear descriptors was adopted for AMD versus healthy screening, but such features may provide limited separability when the task shifts to fine-grained Dry-Wet subtype discrimination due to overlapping visual cues. In [13], multiple handcrafted descriptors (gradient/texture and histogram-based features) were fused with an SVM to enhance discrimination, estimating the macular ROI via Optic Disc (OD) localization using a fixed OD-macula offset with intermediate preprocessing. However, this multi-stage pipeline was sensitive to heterogeneous acquisition, as OD localization or segmentation errors may propagate, leading to inconsistent macular sampling and unstable feature representations, and it is primarily tailored to binary AMD-healthy classification rather than explicit subtype separation. These issues highlight that robustness depends not only on the choice of descriptors but also on how reliably the macular region and lesion cues are standardized across sources. Similarly, in [14], OD-guided ROI cropping/normalization with compact descriptors (Mean/Range, GLCM, CHKM) was utilized, but still relied on drusen segmentation before feature extraction. Although efficient, the segmentation stage was sensitive to cross-source variability (illumination, sensor/FOV, severity), so thresholding/morphology can miss subtle drusen or introduce false positives that propagate into unstable mask-based features and discard fine macular cues through binarization. Moreover, Mean/Range and CHKM mainly capture global intensity-color trends, and fixed-setting GLCM yields region-aggregated texture, which may under-represent localized AMD micro-patterns needed for Dry-Wet subtype discrimination.

Motivated by error propagation from landmark/segmentation steps and the limited sensitivity of compact descriptors to subtle macular cues, this study proposes an anatomically consistent macular standardization and a lightweight yet discriminative representation for multi-class AMD classification. The proposed Assisted-ROI protocol for reproducible macular localization enables controlled comparison with full-fundus and OD-guided ROI baselines, and employs a hybrid handcrafted fusion of LBP-RIU, Haralick/GLCM, and RGB color moments to capture macula-centered micro-texture and color statistics, thereby improving robustness and class separability across heterogeneous sources.

II. MATERIALS AND METHODΣ

A. Dataset

This study compiled a multi-source fundus dataset from four public datasets: ODIR [15], ADAM [16], FIVES [17], and RFMiD [18]. The merged dataset contains 597 images, comprising 287 Normal cases and 310 AMD cases. To harmonize ground truth across sources, all AMD images were re-annotated by three ophthalmologists and categorized into Dry AMD (139 images) and Wet AMD (171 images), where the final label was determined using majority voting.

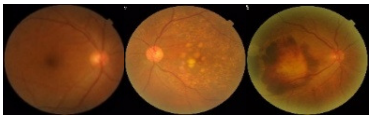

Fig. 1. Sample fundus images from the dataset representing (from left to right) Normal, Dry AMD, Wet AMD.

B. Proposed Method

To ensure consistent spatial resolution and illumination normalization during preprocessing, each fundus image was resized to 512×512 and contrast-enhanced using CLAHE [19]. The proposed pipeline applies an Assisted-ROI step after preprocessing to standardize macular sampling under heterogeneous acquisition conditions. For each preprocessed fundus image $I \in \mathbb{R}^{H\times W\times 3}$, a fixed-size ROI of 224×224 pixels is positioned over the macular region by interactively adjusting the top-left coordinate (x_m, y_m), where x_m denotes the horizontal (left-to-right) and y_m the vertical (top-to-bottom) position. The adjustment is constrained to $x_m \in [0,\ W-224]$ and $y_m \in [0,\ H-224]$ to keep the ROI fully inside the image. The resulting patch $I_{\text{ROI}} = I[y_m{:}\,y_m+224, x_m{:}\,x_m+224]$ is then used for downstream feature extraction, and the ROI coordinates (x_m, y_m, 224, 224) are recorded for traceability and reproducibility. The Assisted-ROI protocol is designed to provide anatomically consistent macular sampling for controlled evaluation rather than a fully automated screening solution.

To assess the impact of macular localization on handcrafted feature learning, three ROI input scenarios were evaluated after preprocessing: (i) Full fundus (no ROI), (ii) Assisted-ROI (proposed), and (iii) The automatic ROI that localizes the macula using an OD-guided mapping. First, a retinal mask is generated from the green channel using thresholding ($T = 20$) followed by erosion and dilation; the green channel is used because it typically provides higher contrast for retinal structures and vessels, facilitating mask formation and OD localization. Next, the OD center c_{OD} is estimated as the maximum-intensity location in a Gaussian-blurred (15×15) green image constrained by the mask. An image-adaptive OD scale proxy is defined as $r_{OD} = 0.08\, w_c$, where w_c denotes the width of the main retinal contour's bounding box, to normalize the search with respect to varying retinal sizes and fields-of-view. A macula search region is then placed on the temporal side relative to the detected OD by shifting the search center along the x-axis by $\pm 5 r_{OD}$, which heuristically approximates the typical OD-fovea separation in disc-diameter units while remaining robust to scale variation. The macula center is selected as the minimum-intensity point within the masked search region (with a fallback when the local mask is empty), since the fovea/macula region often appears darker than surrounding tissue in the green channel under standard imaging. Finally, a fixed 224×224 patch centered at the estimated macula location is cropped with boundary checks to ensure it remains fully inside the image. For all scenarios, the resulting input (full image or ROI patch) was processed using the same downstream feature extraction and classifier settings to ensure a fair comparison.

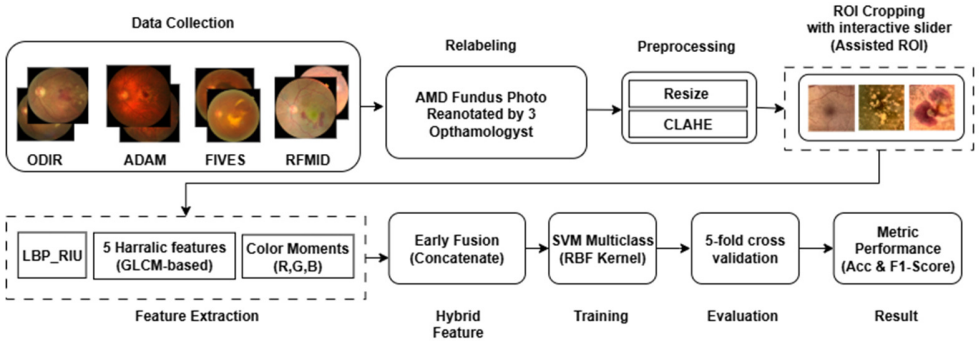

Fig. 2. The proposed model for hybrid handcrafted feature extraction for AMD classification.

Handcrafted features were extracted to integrate texture and color features relevant to AMD. Local texture was represented by uniform LBP-RIU (P = 16, R = 2), summarized into an 18-bin normalized histogram [20]. Second-order texture was described using five Haralick descriptors: angular second moment, contrast, correlation, variance, and inverse difference moment computed from GLCM at distance 1 and averaged across orientations [21]. Color information was captured by RGB color moments (mean, standard deviation, and skewness per channel), yielding nine features [22].

The final representation was formed by early fusion (vector concatenation) of $f_{LBP} \in \mathbb{R}^{18}$, $f_{Har} \in \mathbb{R}^{5}$, and $f_{Col} \in \mathbb{R}^{9}$. The final feature representation was obtained through the vector concatenation process using the formula below:

$$f = [f_{\text{LBP}} \| f_{\text{Har}} \| f_{\text{Col}}] \in \mathbb{R}^{32} \quad (1)$$

Before classification, features were standardized using StandardScaler; in each fold, scaling parameters were fitted on the training features and then applied to both training and validation features. Multiclass classification was performed using an SVM with an RBF kernel (C = 10, γ = 0.01) [23]. To address class imbalance, SMOTE was applied only to the standardized training partition within each cross-validation fold, while the validation fold was kept untouched [24]. All data-dependent operations were executed within the 5-fold stratified cross-validation to minimize information leakage, and performance was reported as the mean accuracy and mean macro-F1 across folds [25]. For class-wise interpretation, confusion matrices were analyzed to identify dominant error patterns by contrasting predicted and ground-truth labels, with diagonal cells representing correct classifications [26]. In addition to classification performance, computational cost was evaluated for the proposed handcrafted pipeline and a modern CNN baseline in terms of training time, inference latency, and total cross-validation time. The CNN baseline adopts an end-to-end fine-tuned ResNet18 initialized with ImageNet weights and executed on a CUDA-enabled GPU (NVIDIA RTX 3050), trained under a two-stage protocol (3-epoch classifier-head warm-up with Adam, lr = 1×10^{-3}; followed by 7-epoch light

fine-tuning with AdamW, $lr = 1\times10^{-5}$) using 224×224 ImageNet-normalized inputs within a 5-fold stratified cross-validation framework, whereas the handcrafted pipeline is evaluated entirely on a CPU-based setup.

C. Experimental Setup

Experiments were implemented in Python 3.10.18 (Jupyter Notebook environment) and executed on a Windows 11 workstation equipped with an Intel Core i5 processor and 16 GB RAM.

III. RESULTS AND DISCUSSION

A. Effect of ROI Cropping Strategy

To isolate the effect of macular localization, ROI input strategies were compared under identical downstream settings. Table I summarizes the impact of ROI localization by comparing three input scenarios: Full Fundus (No ROI), Automatic ROI (OD-guided), and the proposed Assisted-ROI. In all scenarios, the same downstream pipeline was applied (LBP-RIU, Haralick/GLCM, and RGB color moments with SVM-RBF) to ensure a controlled comparison. Training on the full fundus yielded the lowest performance (Accuracy = 0.8091; macro-F1 = 0.7729), indicating that non-macular regions may dilute the discriminative texture and color statistics captured by handcrafted descriptors. The Automatic ROI (OD-guided) provided only a modest improvement (Accuracy = 0.8527; macro-F1 = 0.8241), which is consistent with occasional macular mislocalization under heterogeneous illumination and contrast. In contrast, Assisted-ROI achieved the best performance (Accuracy = 0.9063; macro-F1 = 0.8906), indicating that anatomically consistent macular sampling strengthens class separability when the downstream pipeline is held constant.

TABLE I. AMD'S CLASSIFICATION PERFORMANCE ON VARIOUS ROI SCENARIOS

ROI Scenario	Accuracy	F1-score
Full Fundus (No ROI)	0.8091	0.7729
Automatic ROI (OD-guided)	0.8527	0.8241
Assisted ROI (Proposed)	**0.9063**	**0.8906**

The Assisted-ROI strategy is intended as an upper-bound standardization and human-in-the-loop baseline rather than a fully automated screening solution. Enabling anatomically consistent macular centering, it provides a controlled reference for evaluating the impact of ROI quality on downstream classification performance, while highlighting the sensitivity of handcrafted features to ROI placement.

B. Hybrid Features Against Model Accuracy

The ablation results further support the clinical relevance of combining complementary handcrafted descriptors for multi-class AMD detection. From a clinician's perspective, fundus assessment relies on identifying localized granular texture indicative of drusen deposition, as well as broader intensity and chromatic variations associated with exudation, hemorrhage, and neovascular activity. Dry AMD is characterized by fine-grained macular micro-texture effectively captured by LBP-RIU, which achieves the best single-descriptor performance (Accuracy = 0.8309). In contrast, Wet AMD involves exudative and neovascular processes that alter intensity, contrast, and color distributions, making color moment descriptors particularly informative; their fusion with LBP-RIU further improves performance (Accuracy = 0.8929). Although Haralick features alone are less discriminative (Accuracy = 0.6299), they contribute complementary second-order texture information reflecting macular tissue heterogeneity and structural irregularity associated with disease progression, and their integration with LBP-RIU and color descriptors yields the highest overall accuracy (0.9063), as shown in Figure 3.

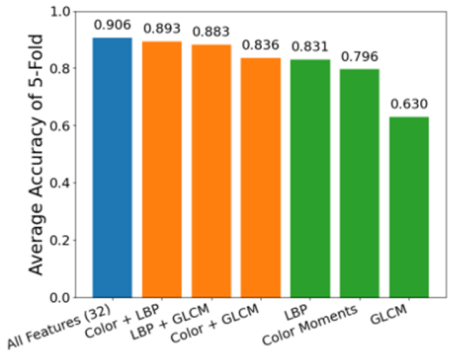

Fig. 3. Accuracy comparison based on feature combination type.

The fold-wise results for the best configuration (all features) further demonstrate a stable generalization, as shown in Table II. The model achieves a mean accuracy of 0.9063 and a mean macro F1-score of 0.8906, indicating robust generalization and balanced class-wise performance.

TABLE II. ACCURACY AND MACRO-AVERAGED F1-SCORE ACROSS 5-FOLD STRATIFIED CROSS-VALIDATION

Fold to	Accuracy	Macro F1-score
1	0.8917	0.8608
2	0.8500	0.8209
3	0.9076	0.8991
4	0.9412	0.9351
5	0.9412	0.9371
Mean	**0.9063**	**0.8906**

C. Comparison with State-of-the-Art Methods

This work reassessed representative handcrafted methods under a unified multi-source and multi-class experimental setting. Feature extraction schemes from previous studies were adopted as originally reported, preserving the original feature definitions and parameter configurations, including image re-annotation by three retinal specialists, preprocessing, Assisted-ROI localization, feature standardization, 5-fold stratified cross-validation, and SVM-RBF classification uniformly applied across all methods to ensure a fair comparison.

As shown in Table III, prior SOTA methods exhibit notable performance degradation when re-evaluated on the same combined multi-source dataset under an identical unified multi-class experimental protocol. Hu Moments [11] show limited robustness under the unified setting (0.5192 accuracy, 0.4936 macro F1-score). PHOG and entropy features [12] achieve moderate performance (0.6300 accuracy, 0.5944 macro F1-

score) but exhibit class imbalance. The hybrid Mean+Range+GLCM+CHKM descriptors [14] provide improved results (0.6918 accuracy, 0.6200 macro F1-score), but remain insufficient for reliable Dry-Wet AMD discrimination. The multi-descriptor approach in [13] performs better (0.8207 accuracy, 0.7743 macro F1-score). In contrast, the proposed approach achieves the highest performance, reaching 0.9063 accuracy and 0.8906 macro F1-score. This margin substantiates the methodological contribution of Assisted-ROI standardization and hybrid handcrafted feature fusion for robust multi-class AMD classification under heterogeneous multi-source conditions.

TABLE III. COMPARISON OF REPRESENTATIVE SOTA METHODS AND THE PROPOSED APPROACH

Feature extraction method	Acc	Macro F1
Hu Moments [11]	0.5192	0.4936
PHOG & Enthropy [12]	0.6300	0.5944
RLBP + HoG + GMH ++ CDH + SDH [13]	0.8207	0.7743
Mean+Range + GLCM + CHKM [14]	0.6918	0.6200
LBP_RIU + GLCM + Color Moments (proposed)	**0.9063**	**0.8906**

As shown in the confusion matrix (Figure 4), the proposed pipeline (Assisted-ROI + hybrid handcrafted features + SVM-RBF) achieves 90.63% accuracy, consistent with the cross-validation summary. Correct predictions dominate the diagonal for all classes, with Normal showing the highest separability (281/287 correctly classified), which is clinically plausible because non-AMD images typically lack macula-centered abnormalities that drive discriminative texture and color-intensity statistics. The remaining errors are mainly concentrated in Dry-Wet confusion (Dry→Wet: 12; Wet→Dry: 17). This pattern is clinically reasonable, as Dry and Wet AMD can share overlapping macular appearance in color fundus images (e.g., drusen-related changes and pigmentary alterations), while neovascular activity may be subtle in a single image and further affected by cross-dataset variability in illumination and field-of-view. From a methodological perspective, Assisted-ROI reduces background-driven feature dilution by standardizing macular sampling, and the hybrid descriptors LBP-RIU (local texture), RGB color moments (color-intensity statistics), and Haralick/GLCM (second-order texture) jointly enhance separability; however, borderline presentations and acquisition heterogeneity can still lead to residual Dry-Wet ambiguity.

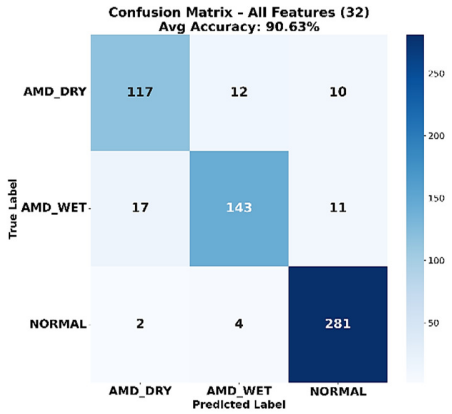

Fig. 4. Confusion matrix of the proposed method for multi-class AMD classification.

These results support a multi-component representation (local texture + supporting texture + color) for improving fine-grained AMD classification on aggregated multi-source fundus data. The proposed method demonstrates competitive performance against state-of-the-art approaches and exhibits consistent class-wise behavior as reflected in the confusion matrix analysis, but practical applicability also requires careful consideration of computational efficiency. To address this aspect, a quantitative comparison of computational cost between the proposed handcrafted pipeline and a modern CNN-based baseline is summarized in Table IV.

As summarized in Table IV, the proposed handcrafted pipeline achieves an accuracy of 0.9063 and a macro F1-score of 0.8906, which are comparable to the end-to-end ResNet18 baseline (0.9012 accuracy and 0.8750 macro F1-score). However, despite GPU acceleration, the ResNet18 baseline incurs substantially higher computational cost, requiring 682.54 s for full cross-validation and 50.81 ms per-image inference. In contrast, the CPU-based handcrafted pipeline completes the entire end-to-end evaluation, including feature extraction and 5-fold cross-validation, in 96.57 s, corresponding to an approximately ×7 reduction in total cross-validation time and over ×1000 faster per-image inference. These results indicate that, even when compared with a GPU-accelerated end-to-end CNN, the proposed handcrafted framework maintains comparable classification performance with a markedly reduced computational footprint, supporting its suitability for deployment in resource-constrained screening settings.

TABLE IV. PERFORMANCE AND COMPUTATIONAL COST COMPARISON BETWEEN THE PROPOSED METHOD AND A RESNET18 BASED CNN BASELINE EVALUATED UNDER THE SAME 5-FOLD CROSS-VALIDATION PROTOCOL

Method	Hardware	Accuracy	Macro F1	Training fold (s)	Inference time/ Image (ms)	Total CV time (s)
Handcrafted (Proposed)	CPU	0.9063	0.8906	0.02	0.044	96.57
CNN (ResNet18)	GPU (CUDA)	0.9012	0.8750	130.02	50.81	682.54

IV. CONCLUSION

This study demonstrates that anatomically consistent macular sampling is critical for robust handcrafted-feature learning on heterogeneous multi-source fundus data. An Assisted-ROI protocol was introduced to standardize the macular region, and complementary descriptors (LBP-RIU, selected Haralick/GLCM features, and RGB color moments) were fused at the feature level and classified using an SVM-RBF. Under 5-fold stratified cross-validation, the proposed pipeline achieved an accuracy of 0.9063 and a macro-F1 of 0.8906, outperforming both full-fundus (No ROI) and automatic OD-guided ROI baselines under identical downstream settings. Fold-wise stability and confusion matrix

analysis indicate strong separability for Normal cases, while residual errors are primarily concentrated in Dry-Wet confusion, which is clinically plausible given overlapping macular appearance in borderline presentations and acquisition variability. Although the evaluation was conducted on a merged multi-source dataset with re-annotation by three retinal specialists to ensure consistent Dry and Wet AMD labeling, the absence of external validation on an independent cohort remains a limitation. Consequently, future work will prioritize validation on fully independent datasets acquired under different clinical conditions to further assess generalizability. Additionally, compact deep embeddings can be integrated with handcrafted features and feature selection strategies as a potential future extension to further enhance class-wise separability in multi-class AMD classification.

ACKNOWLEDGMENT

This work was supported by the Lembaga Pengelola Dana Pendidikan (LPDP), Ministry of Finance, Republic of Indonesia. The authors gratefully acknowledge LPDP for the financial support that enabled this research and publication.

REFERENCES

[1] W. L. Wong *et al.*, "Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis," *The Lancet Global Health*, vol. 2, no. 2, pp. e106–e116, Feb. 2014, https://doi.org/10.1016/S2214-109X(13)70145-1.

[2] A. R. Fernandes *et al.*, "Exudative versus Nonexudative Age-Related Macular Degeneration: Physiopathology and Treatment Options," *International Journal of Molecular Sciences*, vol. 23, no. 5, Feb. 2022, Art. no. 2592, https://doi.org/10.3390/ijms23052592.

[3] U. Schmidt-Erfurth *et al.*, "Guidelines for the management of neovascular age-related macular degeneration by the European Society of Retina Specialists (EURETINA)," *British Journal of Ophthalmology*, vol. 98, no. 9, pp. 1144–1167, Sept. 2014, https://doi.org/10.1136/bjophthalmol-2014-305702.

[4] H. P. N. Scholl *et al.*, "Inter- and intra-observer variability in grading lesions of age-related maculopathy and macular degeneration," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 241, no. 1, pp. 39–47, Jan. 2003, https://doi.org/10.1007/s00417-002-0602-8.

[5] S. T. Berkowitz, A. P. Finn, R. Parikh, A. E. Kuriyan, and S. Patel, "Ophthalmology Workforce Projections in the United States, 2020 to 2035," *Ophthalmology*, vol. 131, no. 2, pp. 133–139, Feb. 2024, https://doi.org/10.1016/j.ophtha.2023.09.018.

[6] M. Tounsi, E. Aram, A. T. Azar, A. Al-Khayyat, and I. K. Ibraheem, "A Comprehensive Review on Biomedical Image Classification using Deep Learning Models," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19538–19545, Feb. 2025, https://doi.org/10.48084/etasr.8728.

[7] F. Li, H. Chen, Z. Liu, X. Zhang, and Z. Wu, "Fully automated detection of retinal disorders by image-based deep learning," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 257, no. 3, pp. 495–505, Mar. 2019, https://doi.org/10.1007/s00417-018-04224-8.

[8] P. M. Burlina, N. Joshi, M. Pekala, K. D. Pacheco, D. E. Freund, and N. M. Bressler, "Automated Grading of Age-Related Macular Degeneration From Color Fundus Images Using Deep Convolutional Neural Networks," *JAMA Ophthalmology*, vol. 135, no. 11, pp. 1170–1176, Nov. 2017, https://doi.org/10.1001/jamaophthalmol.2017.3782.

[9] S. Kadry, V. Rajinikanth, R. González Crespo, and E. Verdú, "Automated detection of age-related macular degeneration using a pre-trained deep-learning scheme," *The Journal of Supercomputing*, vol. 78, no. 5, pp. 7321–7340, Apr. 2022, https://doi.org/10.1007/s11227-021-04181-w.

[10] O. Mamyrbayev *et al.*, "Hybrid Neural Architectures Combining Convolutional and Recurrent Networks for the Early Detection of Retinal Pathologies," *Engineering, Technology & Applied Science Research*, vol. 15, no. 4, pp. 25150–25157, Aug. 2025, https://doi.org/10.48084/etasr.11521.

[11] A. García-Floriano, Á. Ferreira-Santiago, O. Camacho-Nieto, and C. Yáñez-Márquez, "A machine learning approach to medical image classification: Detecting age-related macular degeneration in fundus images," *Computers & Electrical Engineering*, vol. 75, pp. 218–229, May 2019, https://doi.org/10.1016/j.compeleceng.2017.11.008.

[12] U. R. Acharya *et al.*, "Automated screening tool for dry and wet age-related macular degeneration (ARMD) using pyramid of histogram of oriented gradients (PHOG) and nonlinear features," *Journal of Computational Science*, vol. 20, pp. 41–51, May 2017, https://doi.org/10.1016/j.jocs.2017.03.005.

[13] S. Khalid *et al.*, "Automated diagnosis system for age related macular degeneration using hybrid features set from fundus images," *International Journal of Imaging Systems and Technology*, vol. 31, no. 1, pp. 236–252, Mar. 2021, https://doi.org/10.1002/ima.22456.

[14] I. F. Kallel and S. Kammoun, "Hybrid human-artificial intelligence system for early detection and classification of AMD from fundus image," *Signal, Image and Video Processing*, vol. 18, no. 5, pp. 4779–4796, July 2024, https://doi.org/10.1007/s11760-024-03115-2.

[15] N. Li, T. Li, C. Hu, K. Wang, and H. Kang, "A Benchmark of Ocular Disease Intelligent Recognition: One Shot for Multi-disease Detection," in *Benchmarking, Measuring, and Optimizing*, 2021, pp. 177–193, https://doi.org/10.1007/978-3-030-71058-3_11.

[16] H. Fang *et al.*, "ADAM Challenge: Detecting Age-Related Macular Degeneration From Fundus Images," *IEEE Transactions on Medical Imaging*, vol. 41, no. 10, pp. 2828–2847, July 2022, https://doi.org/10.1109/TMI.2022.3172773.

[17] K. Jin *et al.*, "FIVES: A Fundus Image Dataset for Artificial Intelligence based Vessel Segmentation," *Scientific Data*, vol. 9, no. 1, Aug. 2022, Art. no. 475, https://doi.org/10.1038/s41597-022-01564-3.

[18] S. Pachade *et al.*, "Retinal Fundus Multi-Disease Image Dataset (RFMiD): A Dataset for Multi-Disease Detection Research," *Data*, vol. 6, no. 2, Feb. 2021, https://doi.org/10.3390/data6020014.

[19] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*, Academic Press Professional, Inc., 1994, pp. 474–485.

[20] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, July 2002, https://doi.org/10.1109/TPAMI.2002.1017623.

[21] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973, https://doi.org/10.1109/TSMC.1973.4309314.

[22] M. A. Stricker and M. Orengo, "Similarity of color images," in *Storage and Retrieval for Image and Video Databases III*, Mar. 1995, vol. 2420, pp. 381–392, https://doi.org/10.1117/12.205308.

[23] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines*, Regularization, Optimization, and Beyond. MIT Press, 2002.

[24] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, June 2002, https://doi.org/10.1613/jair.953.

[25] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, May 1995, pp. 1137–1143.

[26] G. Zeng, "On the confusion matrix in credit scoring and its analytical properties," *Communications in Statistics - Theory and Methods*, vol. 49, no. 9, pp. 2080–2093, May 2020, https://doi.org/10.1080/03610926.2019.1568485.