Bol. Soc. Paran. Mat.
©SPM – E-ISSN-2175-1188
SPM: www.spm.uem.br/bspm

(3s.) **v. 2026 (44) 7** : 1–15.
ISSN-0037-8712
doi:10.5269/bspm.81377

Unsupervised Rank-Based Fusion with Rough Fuzzy C-Means (URRFCM) for Diabetic Retinopathy Detection

Vandana Saini, Shalini, Chaman Singh and Krishna Kumar Sharma*

ABSTRACT: Diabetic retinopathy (DR) is the most prevalent type of diabetic eye disease and can affect all individuals at any stage of diabetes. If not detected and treated at an early stage, DR may progress to severe vision impairment or even permanent visual loss. Early diagnosis is therefore crucial for improving treatment outcomes. To facilitate timely detection, the development of automated, computer-aided diagnostic systems offers a fast, cost-effective, and widely accessible solution. This study presents Unsupervised Rank Based Fusion with Rough Fuzzy C-Means (URRFCM) framework for DR detection. The proposed method employs multiple pretrained convolutional neural networks (CNNs) for feature extraction, Rank Based Fusion (RBF) technique for feature fusion and intigrate Rough Fuzzy C-Means approach (RFCM) for DR detection. In RFCM, rough set theory is applied to enhance the extracted features, finally, fuzzy clustering is applied to refine the deep clustering outputs. Extensive experiments conducted on three diabetic retinopathy datasets, that the proposed framework significantly outperforms five state-of-the-art deep clustering approaches in DR detection, as validated through comprehensive analysis.

Key Words: Diabetic Retinopathy, unsupervised deep learning, rough fuzzy C-means, medical imaging, computer-aided diagnosis, retinal image analysis.

Contents

1. Introduction

Diabetic Retinopathy (DR) is among the most common and serious microvascular conditions related to diabetes mellitus. It continues to be a leading cause of vision loss and blindness among working-age adults worldwide [1,28]. Continuous high blood sugar damages the retinal blood vessels, this damage causes structural problems such as microaneurysms, bleeding, hemorrhages and the growth of new blood vessels [31]. These changes progressively impair retinal function and, if left undiagnosed or untreated, may result in irreversible vision loss. The severity of diabetic retinopathy (DR) disease divides into five

* Corresponding author.
2020 *Mathematics Subject Classification:* 68T07, 62H30, 92C50.
Submitted January 16, 2026. Published April 11, 2026

stages: non-DR, Mild Non-Proliferative DR (NPDR), Moderate NPDR, Severe NPDR and Proliferative DR (PDR) [27]. These five stages of DR are illustrated in Figure1. Early identification and appropriate treatments are critical to preventing disease progression and visual loss [9]. However, manual screening and grading of the retinal fundus images by ophthalmologists are time-consuming, observer-dependent, and require expertise. These limitations restrict the deployment of large-scale screening, especially in low-resource areas [11]. In recent years, deep learning (DL) and convolutional neural networks (CNNs) have achieved remarkable success in computer-aided medical imaging, particularly in retinal image segmentation and disease classification [20,19]. Various preprocessing techniques are also employed in DR detection to optimizing model efficiency like non-informative image backgrounds are removed so that the CNNs can focus on important retinal features like microaneurysms, hemorrhages, and exudates that are essential for detection and classification of DR. The investigation also applied data augmentation, image flipping, brightness and rotation techniques addressing dataset imbalance and enhancing model robustness [23]. The performance of Deep CNN based detection method has been enhanced by the use of attention modules is integrated with a bilinear method. These method trained more efficiently and performs better in subtle and fine-grained regions leading to improved classification accuracy [33]. Ensemble learning has also played an significant role in developing robust and powerful artificial intelligence frameworks for DR detection by combining the strengths of multiple models [19,34]. Despite these advances, refining feature extraction strategies remains an active research area aimed at improving the sensitivity and specificity of automated DR detection systems [10]. However, obtaining high-quality annotations for automated DR detection methods is both time-consuming and subjective. Lesion severity often overlaps between adjacent DR stages. These issues lead to label noise and inconsistent grading, which can significantly reduce model performance. In contrast, clustering based approaches offer an unsupervised technique that can identify inherent structures and similarities within retinal images without relying on extensive manual annotations. Such approaches helps in grouping images with similar pathological features, making it useful for automated stage discovery and identifying subtle disease progression. For instance, K-means has been utilized for the segmentation of exudates and the optic disc in the retina, which increases early detection capability of DR [7]. Fuzzy c-means have also been tested, more in situations where the boundaries are ambiguous between the lesion and healthy tissue [21]. We proposed a model illustrated in Figure **??** can achieve more robust, explainable, and generalizable representations, especially when dealing with imbalanced or unlabeled datasets that are common in DR imaging. Primary contributions of this study can be summed up as follows:

- A two-stage cutout image processing technique [30] is employed to enhance feature robustness.
- An unsupervised feature extraction methodology is developed using pretrained CNNs, including: Inception V3 [24], Xception [2], and Resnet50 [5].
- The decision scores of the base learners are processed using two nonlinear functions, followed by rank based fusion [17].
- Feature selection using rough set theory to eliminate irrelevant and noisy components, while fuzzy C-Means module is applied for soft membership assignment in the fused latent space.

The remainder of this manuscript is arranged as follows. Section 2 presented the research background. Section 3 describes the proposed URRFCM clustering method. Section 4 reports the experimental evaluations and in Section 5 provides the concludes with final remarks.

2. Background

In this section provides an overview of techniques which are employed in this research paper. The two state cutout method 2.1 strategy for region extraction and refining. Next, in section 2.2 three pre-trained deep learning models for extract important features from images, Subsequently, in section 2.3 rank based fusion for feature fusion and in section 2.4 rough fuzzy c-means is applied to optimize the extracted features and effective grouping.

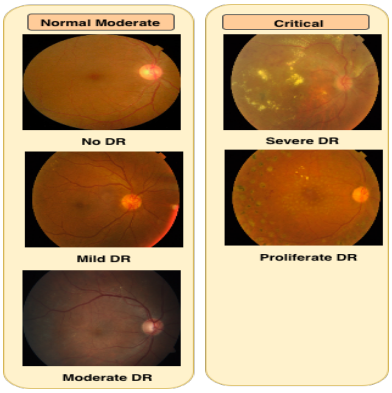

Figure 1: DR Stage

2.1. Two stage cutout method

The cutout technique is employed to address dataset imbalance. Figure 2 shows the two-stage cutout preprocessing framework, including the firt stage and the second stage applied to CIFAR-10 image dataset. In cutout a continuous section of the image is removed rather than individual pixels, dropout units in cutout happens only in the input layer so the masked parts are removed from all subsequent feature maps. In dropout and its variations, a feature may not be removed from one feature map but may be removed from the other feature maps because each feature map is considered independently, which leads to inconsistencies among feature maps and a noisy representation of the images, making the model more robust to noisy inputs. In the first phase of the two-stage cutout, the dataset's dimensions are reduced to 112×112 before a square section of the image is randomly removed and replaced with zero pixels. During the second phase, the dataset size is expanded to 224×224, followed by another random cutout of a larger square area in the image, also filled with zero pixels. The cutout process aims to prompt deep learning networks to concentrate on the broader context of the image, and the two-stage approach enhances this effect. Therefore, after training on images with two different sizes of cutouts, the deep learning networks gain a better understanding of general structures, and the model is able to handle two independent scales of occlusion, which makes the model more robust in distinguishing between images of different dimensions.

Figure 2: An Two stage cutout on CIFAR-10 dataset. (b) First stage mostly masks part-level features of the image, such as heads, legs, or wheels. (b) Second stage of of cutout applied on CIFAR-10 dataset [3]

.

2.2. Deep feature learning

For feature Learning, three pretrained deep CNNs are employed, and they are: The structure of the Inception v3 architecture supports a large number of parallel convolutions. This controls the overfitting issue and enables the generation of deep features with less computation. For feature extraction, three pretrained cnns are employed, and they are: The structure of the Inception v3 architecture supports a large number of parallel convolutions. This controls the overfitting issue and enables the generation of deep features with less computation. The Xception framework, which has an equivalent number of model parameters as the previous one, uses them more effectively. They demonstrated that pointwise convolutions and depthwise separable convolutions exist at opposite ends of a discrete spectrum. ResNet50 has multiple stacked layers; these layers reduced training errors. For unsupervised feature learning.

2.3. Rank based fusion (RBF). method

Figure 3 presents the overall flow diagram flow of RBF where features are extracted using three pretrained CNNs base learners, indexed by $(f = 1, 2, 3)$ and $G = g_1, g_2, .., g_5$ number of classes. Let the base learner f produce the confidence scores are: $H^f = H_1^f, H_2^f, .., H_G^f$
Let $A_1^{f_1}, A_2^{f_1}, \ldots, A_G^{f_1}$ and $A_1^{f_2}, A_2^{f_2}, \ldots, A_G^{f_2}$ be fuzzy ranks generated by two non-linear functions. Output of each pretrained CNNs produce confidence score that is the following:

$$\sum_{g=1}^{G} H_g^f = 1, \forall f = 1, 2, 3.$$

Two non-linear functions for generating fuzzy ranks are calculated in Eq 2.1 and 2.2.

$$A_k^{f_1} = 1 - tanh\left(\frac{(H_g^f - 1)^2}{2}\right) \tag{2.1}$$

$$A_k^{f_2} = 1 - exp\left(-\frac{(H_g^f - 1)^2}{2}\right) \tag{2.2}$$

Then define the fuzzy rank for class k and model f.

$$Q_k^{(f)} = A_k^{f_1} \times A_k^{f_2} \tag{2.3}$$

Fused rank scores for each class is given by Eq 2.4.

$$Q_k^{fused} = \sum_{f=1}^{N} Q_k^{(f)} \tag{2.4}$$

Here in Eq 2.5 N is the number of pretrained CNNs models and k represents the total number of classes. This fused score represents the final score corresponding to each class. We then find the class which has the least fused score considered the winner.

$$F = \min_{\forall k}(Q_k^{fused}) \tag{2.5}$$

2.4. Rough fuzzy Clustering

FRCM [12] integrates the advantage of fuzzy set theory and rough set theory and incorporates fuzzy membership values of each sample to the lower approximation and boundary area of a cluster. Let a set of image data features $F_i(i = 1, 2, 3, ...n)$, each cluster $c_j(j = 1, 2, ...p)$ is regarded as a rough set. It is categorized by $\underline{L}(c_j)$ and $\overline{L}(c_j)$ be the lower and upper bounds of cluster c_j, and let $M(c_j) = \underline{L}(c_j) - \overline{L}(c_j)$ denotes the boundary area of the cluster c_j. The proposed Rfcm algorithm divides a set of n objects into c clusters. The R_{fcm} objective function is defined based on the availability of the lower and boundary

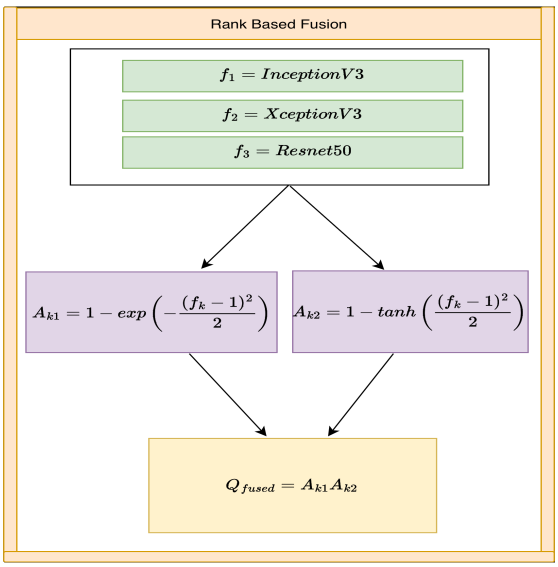

Figure 3: Workflow of Rank Based Fusion

regions of each cluster. F_i represents the features of the input data, while λ_{ij} indicates the degree of membership to the cluster c. The $Rfcm$ algorithm introduces a revised formula for updating fuzzy centroids that considers the effect of fuzzy membership levels. An exponent $m > 1$ is employed to modify the influence of membership values. As outlined in Eq 2.6, the formula for calculating the new fuzzy cluster center is:

$$v_j = \frac{\sum_{i=1}^{n} \lambda_{ij}^m F_i}{\sum_{i=1}^{n} \lambda_{ij}^m} \tag{2.6}$$

$$Rfcm = \sum_{i=1}^{n} \sum_{j=1}^{c} \lambda_{ij}^m \left\| F_i - v_j \right\|^2 \tag{2.7}$$

$$\lambda_{ij} = \frac{1}{\sum_{k=1}^{p} \left(\frac{\|F_i - v_j\|}{\|F_i - v_k\|} \right)^{\frac{2}{m-1}}} \tag{2.8}$$

$$i = 1, 2, 3, ...n,$$

$$F_i \in \underline{L}(c_j), \tag{2.9}$$

$$F_i \in M(c_j), \tag{2.10}$$

$$k = 1, 2, 3, ...p.$$

$$Loss(Rfcm) = \min_{\theta_t} Rfcm \tag{2.11}$$

The loss function of R_{fcm} in Equation 2.11. Here θ_t are considered as parameters an initial condition in the clustering phase.

3. Proposed URRFCM Clustering Method

Input: DR image dataset $\mathcal{I}$; pretrained cnns; cutout parameters;
RBF fusion parameters; Rfcm parameters (m, η).
Output: cluster assignments c; membership matrix λ; cluster centers v.
Step 1: Two-stage cutout Augmentation
foreach *image* $x \in \mathcal{I}$ **do**
  Apply stage-1 cutout to remove a contiguous region from x
  Apply stage-2 cutout at a different position
  store the augmented sample x_{cut} in $\mathcal{I}_{aug}$
end
Step 2: Unsupervised feature extraction
foreach *Cutout image* $x_{cut} \in \mathcal{I}_{aug}$ **do**
  Extracted features using pretrained cnns:
    $f_1 = \text{cnn1}(x_{cut})$,
    $f_2 = \text{cnn2}(x_{cut})$,
    $f_3 = \text{cnn3}(x_{cut})$
end
Step 3: rank based fusion (RBF)
foreach *extracted feature vectors* (f_1, f_2, f_3) **do**
  compute fused feature: $F = Q_k^{(fused)}$
end
Step 4: Rough fuzzy c-means (Rfcm)
Assign cluster centers v and membership matrix λ
repeat
  Compute fuzzy memberships λ_{ij} for each data point
  Evaluate rough set regions for each cluster c_j
    Lower approximation: $\underline{L}(c_j)$
    Boundary approximation: $M(c_j)$
  foreach *cluster* c_j **do**
    Compute centroid contribution S using objects in $\underline{L}(c_j)$
    Compute centroid contribution T using objects in $M(c_j)$
    if $\underline{L}(c_j) \neq \Phi$ ***and*** $M(c_j) \neq \Phi$ **then**
      $v_i = \eta S + (1 - \eta)T$
    else
      if $\underline{L}(c_j) \neq \Phi$ ***and*** $M(c_j) = \Phi$ **then**
        $v_i = S$
      else
        $v_i = T$
      end
    end
  end
  Update membership matrix λ using the Rfcm objective function
until *convergence*
return c, λ, v

Figure 4 and figure **??** shows the proposed method URRFCM for finding DR detection. We employ two stage cutout image processing technique with unsupervised feature learning technique where the cutout regularization technique in place of a dropout layer to avoid overfitting. Cutout is a straightforward regularization method for cnns that adds partially occluded versions of existing samples to the dataset by removing contiguous portions of input images. Similar to dropout, this method improves model generalization performance by working directly on the image space. We integrate feature extraction from DR images using pretrained cnns where we remove the final classification layer and include a fully connected layer with 1024, 1028, and 256 neurons. The layer employs a Rectified Linear Unit (ReLU)

nonlinear mapping function to overcome the problem of vanishing gradients and for faster learning. Extracted features from these pretrained cnns are fused using RBF feature fusion technique. F $= Q_k^{(fused)}$ is the output from RBF and apply fuzzy clustering with rough sets, describe a new fuzzy c-means, referred to as Rfcm. The proposed $Rfcm$ integrates fuzzy membership from fuzzy sets, as well as the lower and upper approximations of rough sets, into c-means clustering process.

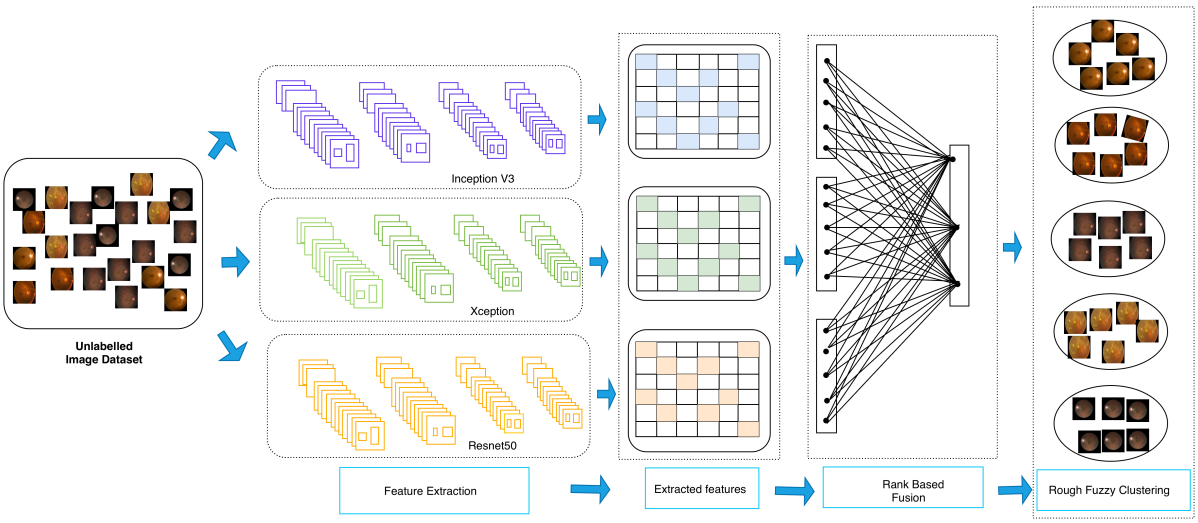

Figure 4: Proposed framework URRFCM

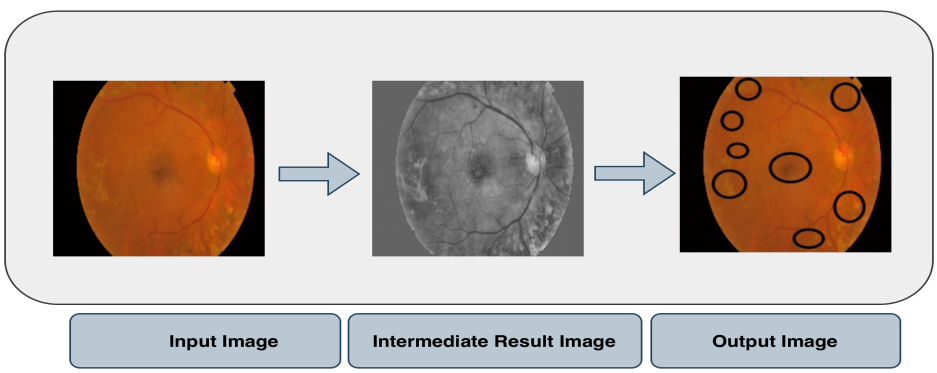

Figure 5: Diabetic Retinopathy Detection using the Proposed URRFCM Method

4. Experiments

Proposed model is evaluated by conducting experiments on three publicly available benchmark DR datasets commonly used for diabetic retinopathy diagnosis [14,8,6]- IDRiD [18], DDR [16], and DiaRetDB1 [15].

4.1. Datasets

The IDRiD dataset comprises 516 high-resolution fundus images collected from an ophthalmology clinic in India. The DDR dataset consists of 12,522 images gathered from 147 hospitals across China. DiaRetDB1 contains 89 fundus images captured in a controlled environment for DR analysis.

4.2. Unsupervised Learning Training

Table 1 presents the hyperparameters used during the unsupervised learning training phase. The learning rate is fixed at 0.0005, which ensures a stable and gradual convergence of the model by controlling the magnitude of weight updates during optimization. The stochastic gradient descent optimizer is chosen for its robustness and effectiveness for handling large scale and high dimensional data. In particular, it is suitable for deep learning tasks. To prevent overfitting and improve model generalization, L2 regularization is incorporated, which discourages large weight values and encourages smoother parameter distributions. The model is trained for 100 epochs, allowing sufficient iterations to minimize the loss and achieve stable convergence.To balance between computational efficiency and convergence stability the batch size is set to 32. The fuzzy loss function is utilized to guide the model's learning in an unsupervised manner. Unlike conventional loss functions based on crisp cluster assignments, fuzzy loss incor porates the concept of membership degrees, enabling the model to capture uncertainty and partial belonging of data samples to multiple clusters. This enhances the clustering quality and provides a more robust representation of the data in the latent space.

Table 1: Hyperparameters Used for Unsupervised Learning Training

Hyperparameter	**Value**
Learning rate	0.0005
Optimizer	Stochastic Gradient Descent
Regularization	L2
Epochs	100
Batch size	32
Loss function	Fuzzy Loss

4.3. Baselines

To evaluate the performance of proposed method, we compare it with existing deep learning models with clustering methods. We assess the efficacy of our suggested approach by contrasting it with a number of representative baselines that are selected to emphasize various model features:

- **AE + KMeans** [4]: A simple deep clustering method that uses KMeans after dimensionality reduction with an autoencoder. This baseline compares the advantages of autoencoder and probabilistic encoding to straightforward deterministic representation learning.

- **DEC** [29]: A deep embedded clustering model that jointly learns feature representations and cluster assignments through a deep neural network. Although DEC demonstrates good clustering performance, it lacks a generative modeling component—similar to KMeans and therefore cannot model the underlying data distribution or generate new samples.

- **VaDE** [13]:VaDE combines a gaussian mixture model within a variational autoencoder network to learn probabilistic latent representations for clustering. It is trained by maximizing the evidence lower bound on the data log likelihood using the stochastic gradient variational bayes estimator((SGVB) and the reparameterization trick.

- **DAFC** [25]: DAFC consists of a deep feature quality verification module and a fuzzy clustering module. It jointly learns deep feature representation learning and integrated fuzzy clustering through a weighted adaptive entropy based loss function.

- **DSCFS** [32]: DSCFS uses a deep neural network to map remote sensing images and further refines it through regularized embedding. In addition, an adaptive feature projection mechanism is leveraged to extract the most discriminative features from the remote sensing images.

4.4. Evaluation Metrics

To evaluate the clustering performance, three standard metrics are used: Accuracy, ARI, NMI [22,26]. These metrics compare the predicted cluster assignments with the ground-truth labels to quantitatively assess the quality of clustering results..

Accuracy. Accuracy indicates the how well the correctly predicted cluster labels with the ground truth labels. It is defined as:

$$\text{Accuracy} = \frac{\sum_{i=1}^{n} \delta(t_i, \text{map}(c_i))}{n} \tag{4.1}$$

where t_i and c_i denote the correctly predicted cluster labels and ground truth labels of the sample i, $\delta(\cdot)$ is the Kronecker *delta*, and map$(\cdot)$ is the optimal permutation mapping of the correctly predicted cluster labels and ground truth labels.

NMI. NMI is a reliable and easy-to-understand metric for assessing how similar two clustering are [26].

$$\text{NMI}(U, V) = \frac{2 \cdot I(U; V)}{H(U) + H(V)}$$

ARI. ARI assesses clustering performance by comparing how pairs of data points are grouped in both the predicted and true labels, accounting for random agreement.

$$ARI = \frac{\sum_{n,m} \binom{q_{nm}}{2} - \left[\sum_n \binom{a_i}{2} \sum_m \binom{b_j}{2} \Big/ \binom{q}{2}\right]}{\frac{1}{2}\left[\sum_n \binom{a_n}{2} + \sum_m \binom{b_m}{2}\right] - \left[\sum_n \binom{a_n}{2} \sum_m \binom{b_m}{2} \Big/ \binom{q}{2}\right]}$$

where q_{ij} represent each element of $A \times B$ matrix.

4.5. Quantitative Results

Table 2 shows proposed model consistently outperforms all baseline methods across the IDRiD, DDR, and DiaRetDB1 datasets. Specifically, compared to the best performing baselines (AE + KMeans, DEC, VaDE, DAFC andDSCFS), proposed mode achieves up to **7.0%** improvement in Accuracy, **11.2%** improvement in NMI, and **3.3%** improvement in ARI on the IDRiD dataset in Figure 7a, Similar trends are observed on the DDR dataset, where proposed model yields **7.9%**, **7.6%**, and **8.4%** gains in Accuracy, NMI, and ARI, respectively in Figure 7b. On the DDR dataset, our model improves Accuracy by **7.3%**, NMI by **3.3%**, and ARI by **5.0%** in Figure 7c.
Table 3 presents a comparison of different feature fusion techniques. When the Rank-Based Fusion technique is applied to the IDRiD dataset, the values of Accuracy, NMI, and ARI show a notable improvement compared to the Concatenation and Weighted Sum fusion methods. The visualization of these results is illustrated in Figure 6.

Table 2: Clustering performance (%) on benchmark datasets: IDRiD, DDR and DiaRetDB1

Method	**IDRiD Dataset**			**DDR Dataset**			**DiaRetDB1 Dataset**		
	Accuracy	NMI	ARI	Accuracy	NMI	ARI	Accuracy	NMI	ARI
AE + KMeans	76.8	67.4	62.7	73.9	64.9	60.8	72.1	63.0	58.4
DEC	81.3	72.5	68.9	79.6	70.3	66.5	77.4	68.2	64.3
VaDE	83.5	75.1	70.6	81.2	72.8	69.0	79.3	71.1	67.5
DAFC	85.4	78.3	73.1	83.0	75.6	71.2	82.1	74.0	70.3
DSCFS	87.6	79.5	80.1	85.3	77.5	72.4	84.2	79.4	76.2
Proposed method	**94.6**	**90.7**	**83.4**	**93.2**	**85.1**	**80.9**	**91.5**	**82.7**	**81.2**

4.6. Ablation Study Settings

An ablation study is performed by selectively disabling in order to assess the distinct contributions of each component in proposed model framework. The baseline is the **full model**, which has all the components: RBF for fusion, Rough fuzzy clustering for soft cluster assignment. The ablation study's

Table 3: Percentage Improvement of rank based fusion method over concatenation and weighted sum fusion methods on IDRiD Dataset

Feature fusion	**Accuracy (%)**	**NMI (%)**	**ARI (%)**
Concatenation	+4.8	+6.8	+1.1
Weighted Sum	+5.3	+8.6	+1.6
Rank Based Fusion	+7.0	+11.2	+3.3

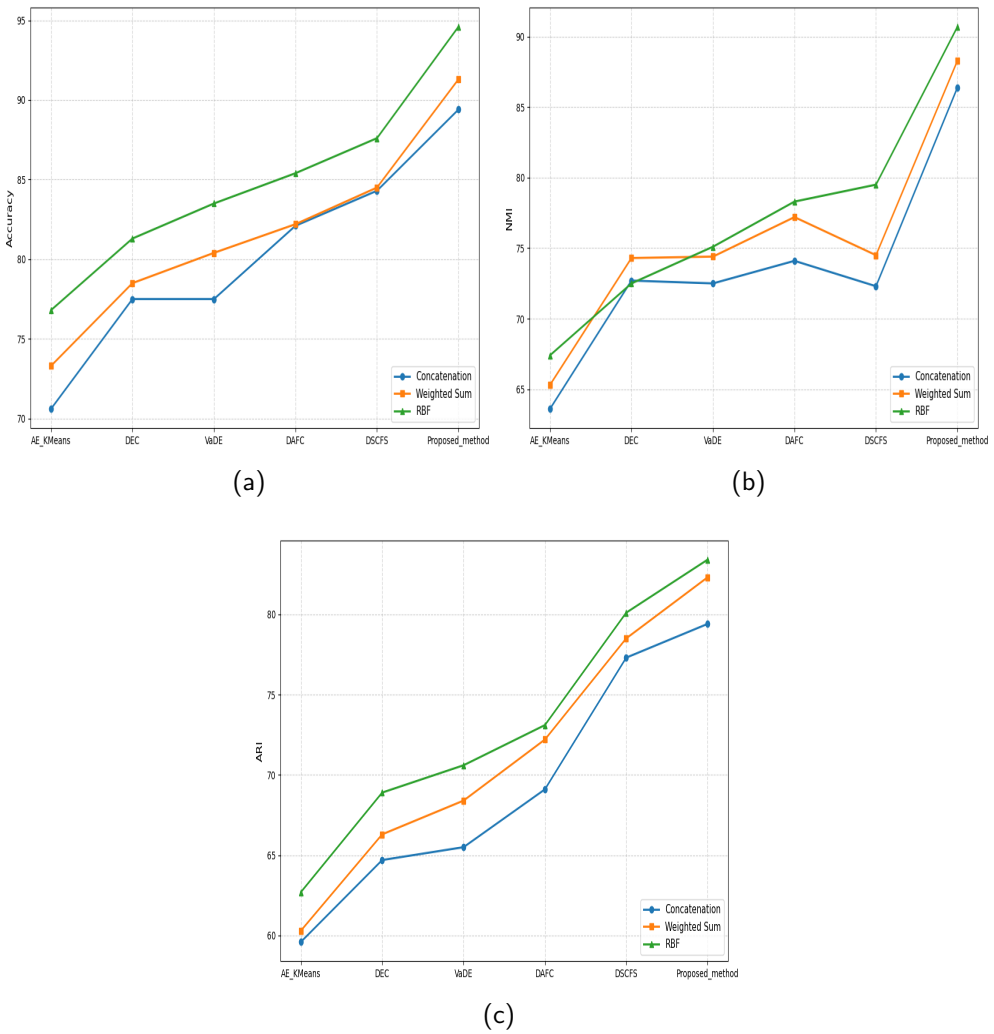

Figure 6: Visualization of Accuracy, NMI and ARI on IDRiD dataset using different fusion techniques

findings are summed up in Table 4 and visually depicted in Figure 8a. corresponding to the FCM method, many of the clusters are not well separated. RFCM extracts features and optimized using rough set which provide some separation between clusters, but the representation lacks clear boundaries in Figure 8b, indicating its poor clustering performance. RBF + FCM is used for cluster performance improvement Figure 8c. On the other hand, Figure 8d clearly shows improved clustering behavior with the proposed method. The latent representations are more compact, and well-defined cluster boundaries. The model effectively integrates information.

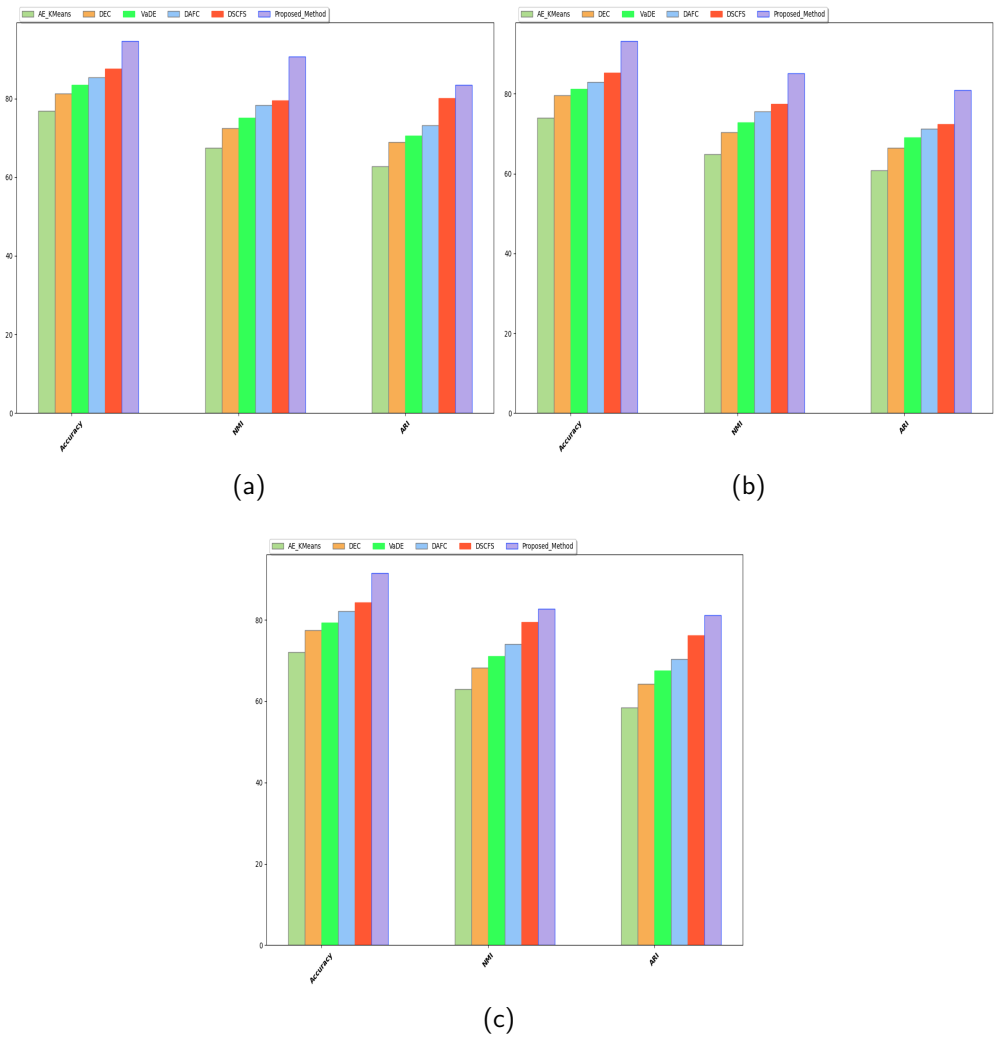

Figure 7: Visualization of Accuracy, NMI and ARI on (a).IDRiD, (b).DDR, (c).DiaRetDB1 datasets

Table 4: Ablation study on the proposed model.

Configuration	**NMI**	**ARI**	**Accuracy**
Proposed Model (RBF + RFCM)	**0.89**	**0.87**	**0.91**
RBF + FCM	0.84	0.81	0.86
RFCM	0.80	0.77	0.83
FCM	0.82	0.79	0.85

4.7. Complexity Analysis

From Algoritim 1, the computational complexity of our proposed URRFCM consists mainly of four parts, which correspond the initialization and updates of variables respectively. To be more specific the update of F where n be the number of images, f the number of pretrained CNNs, and d the extracted feature dimension. Feature extraction has a complexity of $O(fnd)$. RBF operates on confidence scores produced by the base learners. Ranking across G classes and f models for n samples incurs a relatively low complexity of $O(nfGlogG)$, making this step computationally lightweight compared to deep feature extraction. The RFCM clustering step dominates the online computation. For c clusters, fuzzification

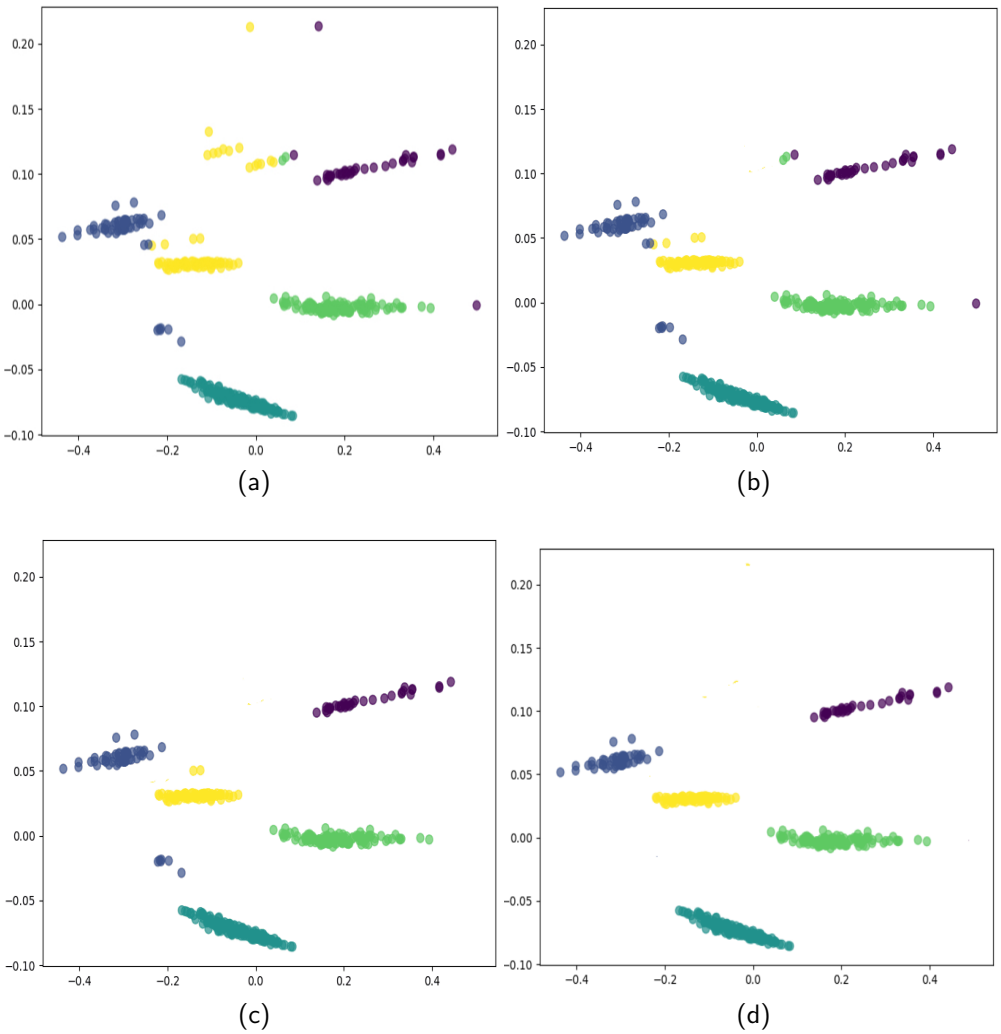

Figure 8: Visualization of clusters on on IDRiD Dataset using different methods (a) FCM, (b) RFCM, (c) RBF + FCM, (d) RBF + RFCM

parameter m and t iterations, the complexity is approximately $O(tncd)$. which is comparable to classical FCM. The incorporation of rough set boundaries introduces only minor overhead through upper and lower approximation updates. Overall, the proposed URRFCM method remains computationally feasible for large-scale DR datasets. Most intensive computations occur during offline feature extraction, while the fusion and clustering stages scale linearly with the number of samples. Experimental runtimes demonstrate stable convergence within a small number of iterations, confirming the practical applicability of the method for unsupervised DR detection and screening scenarios.

5. Conclusion

In this work, we proposed a novel unsupervised deep learning framework for DR detection that integrates two stage cutout for image processing, RBF for feature fusion and rough fuzzy c-means. The model leverages multiple pretrain CNNs to extract robust and discriminative features from retinal images. Additionally, we embedding rough set-based feature selection with deep fuzzy clustering. This strategy enables the model to minimize clustering loss in a unified end-to-end pipeline, leading to more interpretable and structure-aware representations. Extensive empirical validation across three DR bench-

marks demonstrates the superiority of the proposed method over existing unsupervised and deep learning baselines. The experimental results confirm the effectiveness of combining rough and fuzzy paradigms in unsupervised representation learning for medical image analysis. The results suggest that a tool to help ophthalmologists in DR screening at large level can be developed. Moreover, It can be further enhanced for incomplete and and imprecise diagnostic environments. Despite its promising results, proposed UR-RFCM has certain limitations. The ongoing assessment is restricted to single-center benchmark datasets, and the model's robustness in multi-center, highly noisy, or significantly imbalanced clinical data remains to be explored. Additionally, the computational expenses linked to various pretrained CNNs could restrict real-time implementation. Future efforts will aim to expand the framework to accommodate multi-center and diverse datasets, enhance resilience to noise and incomplete data, and investigate efficient or self-supervised feature extractors to improve scalability and real-world clinical utility.

References

1. World health organization: Vision impairment and blindness. *World Health Organization*, 2023.
2. Basem S Abunasser, Mohammed Rasheed J AL-Hiealy, Ihab S Zaqout, and Samy S Abu-Naser. Breast cancer detection and classification using deep learning xception algorithm. *International Journal of Advanced Computer Science and Applications*, 13(7), 2022.
3. Nikoo Aghaei. A model-agnostic derivative of cutout for image data augmentation. 2022.
4. Elie Aljalbout, Vladimir Golkov, Yawar Siddiqui, Maximilian Strobel, and Daniel Cremers. Clustering with deep learning: Taxonomy and new methods. *arXiv preprint arXiv:1801.07648*, 2018.
5. Md Zahangir Alom, Tarek M Taha, Christopher Yakopcic, Stefan Westberg, Paheding Sidike, Mst Shamima Nasrin, Brian C Van Esesn, Abdul A S Awwal, and Vijayan K Asari. The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv preprint arXiv:1803.01164*, 2018.
6. Asia Pacific Tele-Ophthalmology Society (APTOS). Aptos 2019 blindness detection dataset. https://www.kaggle.com/c/aptos2019-blindness-detection, 2019. Acces.
7. RS Biyani and BM Patre. A clustering approach for exudates detection in screening of diabetic retinopathy. In *2016 International Conference on Signal and Information Processing (IConSIP)*, pages 1–5. IEEE, 2016.
8. Etienne Decencière et al. Messidor: Methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology. In *Proceedings of the 22nd International Conference on Pattern Recognition (ICPR)*, page 200–203, 2014.
9. Li Ding et al. A review on automated diabetic retinopathy detection using deep learning. *Computers in Biology and Medicine*, 2021.
10. Shradha Dubey and Manish Dixit. Recent developments on computer aided systems for diagnosis of diabetic retinopathy: a review. *Multimedia Tools and Applications*, 82(10):14471–14525, 2023.
11. Varun Gulshan et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*, 2016.
12. Qinghua Hu and Daren Yu. An improved clustering algorithm for information granulation. In *International Conference on Fuzzy Systems and Knowledge Discovery*, pages 494–504. Springer, 2005.
13. Zhuxi Jiang, Yin Zheng, Huachun Tan, Bangsheng Tang, and Hanning Zhou. Variational deep embedding: An unsupervised and generative approach to clustering. *arXiv preprint arXiv:1611.05148*, 2016.
14. Kaggle. Eyepacs diabetic retinopathy dataset. https://www.kaggle.com/c/diabetic-retinopathy-detection, 2015. Accessed: 2025-11-10.
15. RVJPH Kälviäinen and H Uusitalo. Diaretdb1 diabetic retinopathy database and evaluation protocol. In *Medical image understanding and analysis*, volume 2007, page 61. Citeseer, 2007.
16. Tao Li, Yingqi Gao, Kai Wang, Song Guo, Hanruo Liu, and Hong Kang. Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. *Information Sciences*, 501:511–522, 2019.
17. Ankur Manna, Rohit Kundu, Dmitrii Kaplun, Aleksandr Sinitca, and Ram Sarkar. A fuzzy rank-based ensemble of cnn models for classification of cervical cytology. *Scientific Reports*, 11(1):14538, 2021.
18. Prasanna Porwal, Samiksha Pachade, Ravi Kamble, Manesh Kokare, Girish Deshmukh, Vivek Sahasrabuddhe, and Fabrice Meriaudeau. Indian diabetic retinopathy image dataset (idrid): a database for diabetic retinopathy screening research. *Data*, 3(3):25, 2018.
19. Harry Pratt et al. Convolutional neural networks for diabetic retinopathy. *Procedia Computer Science*, 90:200–205, 2016.
20. Guillaume Quellec et al. Deep image mining for diabetic retinopathy screening. *Medical Image Analysis*, 2017.

21. RS Rajkumar and A Grace Selvarani. Diabetic retinopathy diagnosis using resnet with fuzzy rough c-means clustering. *Computer Systems Science & Engineering*, 42(2), 2022.

22. Alexander Strehl and Joydeep Ghosh. Cluster ensembles—a knowledge reuse framework for combining multiple partitions. In *Journal of Machine Learning Research*, volume 3, pages 583–617, 2002.

23. Chaichana Suedumrong, Suriya Phongmoo, Tachanat Akarajaka, and Komgrit Leksakul. Diabetic retinopathy detection using convolutional neural networks with background removal, and data augmentation. *Applied Sciences*, 14(19):8823, 2024.

24. Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2818–2826, 2016.

25. Dayu Tan, Zheng Huang, Xin Peng, Weimin Zhong, and Vladimir Mahalec. Deep adaptive fuzzy clustering for evolutionary unsupervised representation learning. *IEEE Transactions on Neural Networks and Learning Systems*, 35(5):6103–6117, 2023.

26. Nguyen Xuan Vinh, Julien Epps, and James Bailey. Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *Journal of Machine Learning Research*, 11:2837–2854, 2010.

27. Charles P Wilkinson, Frederick L Ferris, Ronald E Klein, Peter P Lee, Carl-Daniel Agardh, Matthew Davis, Don Dills, Anselm Kampik, R Pararajasegaram, and John T Verdaguer. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology*, 110(9):1677–1682, 2003.

28. Tien Yin Wong et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: Systematic review and meta-analysis. *The Lancet Diabetes & Endocrinology*, 2016.

29. Junyuan Xie, Ross Girshick, and Ali Farhadi. Unsupervised deep embedding for clustering analysis. In *International conference on machine learning*, pages 478–487. PMLR, 2016.

30. Zhaomin Yao, Wenxin Mao, Yizhe Yuan, Zhenning Shi, Gancheng Zhu, Wenwen Zhang, Zhiguo Wang, and Guoxu Zhang. Fuzzy-vgg: A fast deep learning method for predicting the staging of alzheimer's disease based on brain mri. *Information Sciences*, 642:119129, 2023.

31. Joanne W. Yau et al. Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care*, 35(3):556–564, 2012.

32. Yang Zhao, Zixuan Bi, Peican Zhu, Aihong Yuan, and Xuelong Li. Deep spectral clustering with projected adaptive feature selection. *IEEE Transactions on Geoscience and Remote Sensing*, 2025.

33. Ziyuan Zhao, Kerui Zhang, Xuejie Hao, Jing Tian, Matthew Chin Heng Chua, Li Chen, and Xin Xu. Bira-net: Bilinear attention net for diabetic retinopathy grading. In *2019 IEEE international conference on image processing (ICIP)*, pages 1385–1389. IEEE, 2019.

34. Yi Zhou and Hao Chen. Ensemblenet: Improving convolutional neural networks performance for diabetic retinopathy grading. *Biomedical Signal Processing and Control*, 55:101632, 2020.

Vandana Saini,
Department of Computer Science and Informatics,
University of Kota,
Kota, Rajasthan-324005, India.
E-mail address: vandana.saini@uok.ac.in

and

Shalini,
Department of Computer Science and Informatics,
University of Kota
Kota, Rajasthan-324005, India.
E-mail address: shalini.sharma@uok.ac.in

and

Chaman Singh,
Department of Mathematics,
Acharya Narendra Dev College
Delhi, India,
Country.

E-mail address: chamansingh@andc.du.ac.in

and

Krishna Kumar Sharma,
Department of Computer Science and Informatics,
University of Kota
Kota, Rajasthan-324005, India.
E-mail address: krisshna.sharma@uok.ac.in