International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Diabetic Retinopathy Prediction Using Soft Computing-Based Fuzzy-SVM Integrated Diabetic Retinopathy Prediction Framework (SC-FSIDR-PF)

[1]Mrs. P. Anitha, [2]Dr. P. R. Tamilselvi

Submitted: 15/05/2024 **Revised:** 25/06/2024 **Accepted:** 10/07/2024

Abstract – Diabetic retinopathy (DR) is a severe eye condition caused by diabetes, leading to vision impairment and blindness if not detected early. This study proposes a novel hybrid model integrating soft computing techniques, specifically Fuzzy Logic and Support Vector Machine (SVM), for predicting DR from retinal images. Initially, retinal images from the Indian Diabetic Retinopathy Image (IDRiD) dataset are collected and pre-processed using the Improved Non-local Means (INLM) filter for noise reduction. Feature extraction is performed using the Gray-Level Co-occurrence Matrix (GLCM), followed by dimensionality reduction with Principal Component Analysis (PCA). The fuzzy logic system processes these features, handling the inherent uncertainty and imprecision in medical data and outputs fuzzy values representing the health state. These outputs are then fed into an SVM classifier, which employs a kernel function to handle non-linearity and separates the data into healthy and DR-affected categories. This hybrid approach leverages the interpretability of fuzzy logic and the classification strength of SVM, resulting in a robust and accurate predictive model for Early detection of diabetic retinopathy is possible with soft computing techniques.

Keyword: *Soft computing, fuzzy-logic, Support Vector Machine, Improved Non-local Means filter, Gray-Level Co-occurrence Matrix, Principal Component Analysis.*

1. Introduction

Current medical research and diagnosis relies on medical image processing for efficient detection and treatment. There are diverse imaging modalities like Computed Tomography Scans, fundus imaging, X-ray, ultrasound and Magnetic Resonance Imaging which aids the clinicians in diagnostic procedures. Current advancements in technology have led to the development of digital imaging systems with high-resolution images that are enough for the majority of clinical scenarios. In ophthalmology, retinal digital imaging offers a stable, high quality evidence of the look of the retina with application for diabetic retinopathy test program. Digital fundus or retinal images can be exposed to image analysis, to make an objective quantitative analysis of fundus images and has the potential for automated diagnosis to aid in decision-making. The retinal image analysis is indeed a complicated task particularly because of the color variability of the images. The existence of the retinal anatomical pathological structures and the morphology features in different patients may produce erroneous interpretation. This has led to the expansion of many retinal image examination methods. Retinal pictures are typically processed in an algorithmic fashion, where the results of one step are used to inform the results of subsequent steps. Distinctive sequence may consist of one or more pre-processing measures followed by image feature extraction and classification stages.

Diabetic Retinopathy (DR)

Diabetic Retinopathy is a common complication of diabetes that affects the eyes, specifically the retina. The disease is characterized by damage to the blood vessels of the retina, which can lead to vision impairment and, in severe cases, blindness. DR can be categorized into different types of lesions and severity levels:

[1]Research Scholar, PG and Research Department, Department Computer Science, Periyar University Salem
[2]Research Supervisor & Assistant Professor, Department of Computer Science Govt Arts and Science College, Komarapalayam

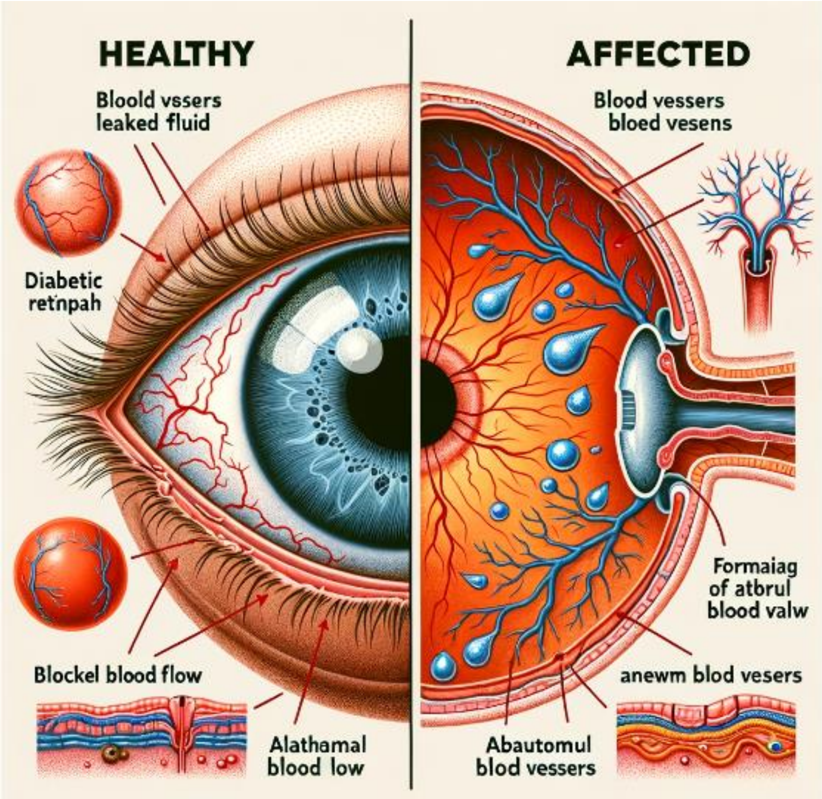

Figure 1.Affected & Non Affected Eye

Types of DR Lesions

1. **Microaneurysms**: Small bulges in the blood vessels of the retina that can leak fluid.
2. **Hemorrhages**: Small blood spots in the retina caused by bleeding from damaged blood vessels.
3. **Exudates**: Lipid or protein deposits that form in the retina due to leaking blood vessels.
4. **Cotton Wool Spots:** A buildup of nerve fiber layer infarcts that results in fluffy white patches on the retina.
5. **Neovascularization**: Formation of new, abnormal blood vessels that are prone to bleeding typically associated with advanced stages of DR.

DR Severity Grading

DR severity is typically classified into the following stages:

1. **Diabetic Retinopathy with Mild Non-Proliferative Retinopathy (NPDR)**: NPDR is characterized by the presence of micro aneurysms and small retinal hemorrhages. Patients may be asymptomatic at this stage.
2. **Moderate NPDR**: Increased number and severity of hemorrhages, microaneurysms, and exudates. Some blood vessels may become blocked, reducing retinal blood flow.
3. **Severe NPDR**: Significant blockage of retinal blood vessels, leading to extensive hemorrhages and cotton wool spots. higher chance of proliferative diabetic retinopathy developing.
4. **Proliferative Diabetic Retinopathy (PDR)**: Characterized by neovascularization, where new abnormal blood vessels form on the surface of the retina or optic disc. These new vessels are fragile and can bleed, causing vitreous hemorrhage and potentially leading to retinal detachment.

Predicting Diabetic Retinopathy (DR) using machine learning involves developing models that can analyze retinal images and identify features associated with the disease.

2. Literature Survey

1. K. A. Anant (2017) et.al proposed Diabetic retinopathy detection through image mining for type 2 diabetes [1]. Research is needed to diagnose diabetes because it is a leading cause of death in developing nations. Researchers are always trying to identify diabetes using data mining, image processing, and other techniques. Type 1 and type 2 diabetes are the two categories for diabetes.

Diabetes type 2 can be cured if caught in its early stages, whereas type 1 diabetes is fatal and difficult to treat. Diabetic Retinopathy analysis, a retinal examination, is a simple way to identify type 2 diabetes. Many studies have used image processing techniques to find evidence of DR in the retina. We have presented a method in our study that extracts texture and wavelet features for detection through image processing and data mining. The picture of the standard database DIARETDB1 is used to obtain the results, which are then assessed using the sensitivity, specificity, and accuracy metrics. With a high accuracy of 97.75%, our method can aid in the early detection and prevention of diabetes.

2. M. Panwar (2016) et.al proposed K-nearest neighbor based methodology for accurate diagnosis of diabetes mellitus [2]. One of the main causes of death, disability, and loss of income worldwide is diabetes. Ninety to ninety-five percent of people globally have type 2 diabetes. On the other hand, it can be avoided or postponed with the appropriate attention and actions, including an early diagnosis. Particularly in the area of different machine learning algorithms for medical diagnostics, significant progress has been made. However, accuracy frequently declines as a result of incomplete medical data sets, which increases the likelihood of misclassifications that can result in dangerous repercussions. For many researchers, accurately diagnosing and predicting an illness becomes a difficult research topic. As a result, we have developed a new methodology based on cutting-edge preprocessing methods and the K-nearest neighbor classifier in an effort to increase the diagnosis accuracy. Several quantitative measures and a comparative study with previously published studies concentrating on the diagnosis of pimadiabetes condition using the same UCI datasets are employed to confirm the recommended methodology's effectiveness. This is the first work of its sort, demonstrating the superiority of the suggested methodology over current techniques by achieving 100% classification accuracy through feature reduction from eight to two.

3. K. S. Argade (2015) et.al proposed Automatic detection of diabetic retinopathy using image processing and data mining techniques [3]. One of the main causes of blindness in people with diabetes mellitus is a retinal illness called diabetic retinopathy. It's an illness that causes the blood vessels in the retina to enlarge. If diabetes is present at a very high level, this might damage the retina in the eye and cause blindness. Regular screenings for early detection are the most effective form of treatment. The doctors would be able to more accurately and readily assess the patient's condition if these photographs were automatically screened. The identification of retinal pictures through the use of suitable image processing and data mining techniques is emphasized in this. This makes it simple to distinguish between normal and aberrant retinal scans, which will cut down on the amount of reviews that physicians receive.

4. W. Ding (2019) et.al proposed SVM-Based Feature Selection for Differential Space Fusion and Its Application to Diabetic Fundus Image Classification [4]. One of the most significant uses of machine learning is data mining. Complex data categorization in machine learning techniques uses the support vector machine (SVM) algorithm and the fusion kernel principle component analysis (KPCA). A promising classification performance is achieved by the SVM based feature selection approach for differential space fusion (DSF-FS), which addresses the shortcoming of the fusion KPCA and SVM algorithm. Differential space data is obtained by first processing the original data using principle component analysis (PCA). The KPCA algorithm is then applied to the original data and the differential space data, respectively, to obtain the differential space fusion features. Subsequently, the ReliefF algorithm is employed to determine the feature weights, and an initial classification assessment measure is used to choose the best feature combination. Third, the dimensionality reduction data is classified using the SVM method. Lastly, some experimental findings on the five UCI datasets demonstrate that the suggested DSF-FS algorithm can lower the computational complexity of the classification process in addition to increasing classification accuracy. Furthermore, the DSF-FS method may be effectively used for the categorization of diabetic fundus images, and the positive outcomes further highlight its excellent applicability and feasibility.

5. D. K. Prasad (2015) et.al proposed early detection of diabetic retinopathy from digital retinal fundus images [5]. Diabetic retinopathy is the impairment of the retinal blood vessels due to complications of diabetes, which can subsequently lead to loss of vision. The implementation of a retinal screening system, which would identify retinal impairment early on, is the only way to address this issue. This research suggests blood vessel, exudate, and microaneurysm detection using morphological operations and segmentation approaches. There are four subimages within the retinal fundus picture. The retinal fundus image is processed to extract various information. The extracted characteristics undergo application of Haar wavelet transforms. The technique of principal component analysis is then used to improve feature selection. Using back propagation neural networks and one rule classifiers, the images are categorized as either non-diabetic or diabetic. DIARETDB1, a publicly accessible diabetic retinopathy data set, is used for the experiments. Performance is evaluated with metrics like sensitivity, specificity and accuracy, the results obtained are encouraging.

3. Research Methodology

A collection of approaches known as "soft computing" offers flexible information processing capabilities. Its goal is to attain tractability, resilience, and low-cost solutions by leveraging the tolerance for imprecision, uncertainty, approximate reasoning, and partial truth. The results of this would aid medical professionals, scientists, and pharmacists in understanding the traits and combinations of traits that cause various illnesses, offering accurate diagnosis techniques, and in the development of novel medications. The below figure 2 depicts a flowchart outlining the proposed framework for predicting diabetic retinopathy (DR) using soft computing techniques. The framework consists of four main stages: Data Collection, Data Pre-Processing, Feature Extraction, and Classification. The Indian Diabetic Retinopathy Image (IDRiD) dataset is utilized for collecting retinal images. This dataset includes images showing various stages of DR. The Improved Non-local Means (INLM) filter is applied to the collected retinal images. This step involves noise reduction to enhance the quality of the images for further analysis. To extract features, the Gray-Level Co-occurrence Matrix (GLCM) and Principal Component Analysis (PCA) are utilized. GLCM captures the texture features of the retinal images, and PCA reduces the dimensionality of the extracted features, retaining the most relevant information. This system handles the uncertainty and imprecision in medical data by processing the extracted features to produce fuzzy values representing the health state of the retina. The fuzzy values are supplied into a Support Vector Machine (SVM) classifier together with the retrieved features. The SVM uses a kernel function to manage non-linearity and classify the images into healthy and DR-affected categories. This hybrid approach leverages the interpretability of fuzzy logic and the classification strength of SVM, resulting in a robust and accurate predictive model for early detection of diabetic retinopathy.

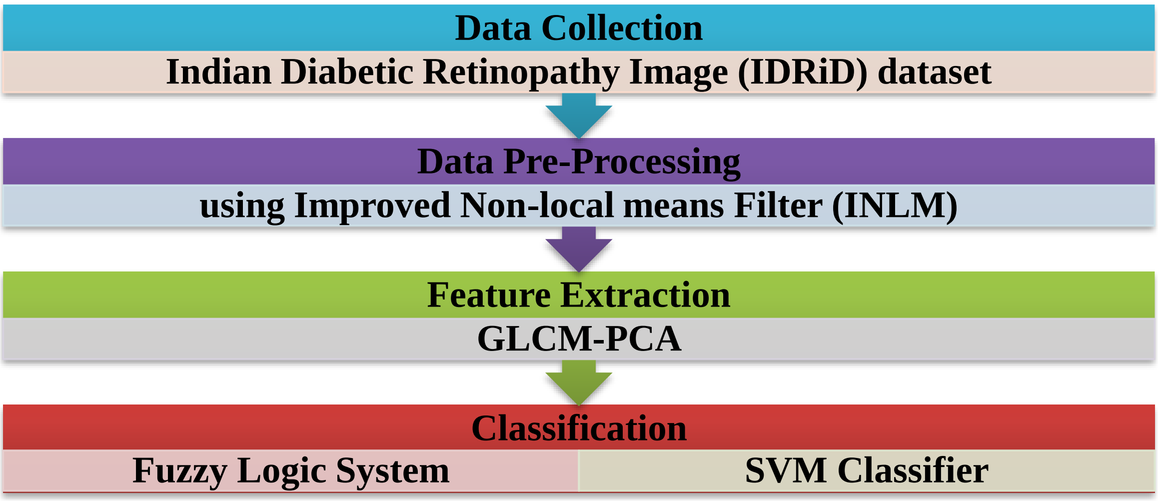

Figure 2.Soft Computing-Based Fuzzy-Svm Integrated Diabetic Retinopathy Prediction Framework

3.1 Data Collection

Gathering a dataset of retinal images—both normal and ones displaying different stages of DR—is the initial step in the process. The retinal fundus images in the Indian Diabetic Retinopathy Image (IDRiD) dataset were obtained from an eye clinic located in Nanded, Maharashtra, India. This collection has 516 photos with a 50° FOV and a resolution of 4288 x 2848 pixels. Of the fundus photos in this dataset, 454 exhibit symptoms indicative of NPDR severity, whereas the remaining 62 fundus images display symptoms indicative of PDR severity. This collection is made up of acceptable quality, therapeutically relevant photos that were gathered from thousands of patient DR and diabetic macular edema examinations.

IDRiD: Indian Diabetic Retinopathy Image Dataset, 2018. [Dataset]. Available: http://dx.doi.org/10.21227/H25W98.

3.2 Data pre-Processing

Data pre-processing using Improved Non-local means Filter (INLM)

Non-local means Filter (NLM)

In this type of filter, the nonlocal assessment of pixels replaces the local comparison of patches. This filter doesn't have any assumptions about where to put the best relevant pixels in order to de-noise an existing pixel. The NLM filter establishes a relationship between the grey values at a specific location and the geometric harmony throughout a neighborhood [17]. In NLM, a pixel's weight is unaffected by differences in intensity or location.

The weighted mean value of all strengths of the pixels (X_a) in picture I serves as the restored intensity $v(i,j)$ of pixel X_a.

$$NLv(i,j) = \sum_{Xb \in I} W(Xa, X\,b)\mathrm{u(Xa)}$$

Here $W(Xa, Xb)$ indicates the average weight value assigned to u(Xa) for restoring the Xa pixel.

Unlike standard filters, the Non-local Means (NLM) filter is a potent denoising approach that evaluates pixels based on similarities across the entire image, not just locally. In contrast to local filters, which solely take into account nearby patches, NLM assesses a pixel's similarity to every other pixel in the image. NLM can successfully reduce noise while maintaining fine details thanks to this method.

Non-local Evaluation: When evaluating a pixel, NLM compares it to every other pixel in the image, not simply those that are close by. Structures and features that could be lost with local filtering are preserved with the aid of this global comparison.

Patch Comparison: NLM compares regions of pixels as opposed to individual pixels. This implies that NLM takes into account similar patches throughout the image rather than simply neighboring ones when denoising a pixel.

Weighting Mechanism: Each pixel's contribution to a target pixel's denoised value depends on how closely it resembles the target pixel's patch. During the averaging process, pixels with more similarity to the target patch are assigned larger weights.

Spatial and Intensity Independence: In contrast to conventional filters, a pixel's weight in NLM is independent of its distance from the target pixel or of the difference in intensity between the two. Because of its independence, NLM can effectively handle intricate noise patterns and preserve image features.

Denoising Effectiveness: Because of its global approach and patch-based comparison, NLM performs particularly well in scenarios where noise is present consistently throughout the image or where standard local filters are unable to keep fine features.

The Non-local Means filter (NLM), which employs patch-based comparisons and global image information, significantly improves image denoising and yields excellent results. It is a helpful tool in many image processing applications because of its capacity to reduce noise while maintaining image characteristics.

Improved Non-local Means (INLM) Filter Algorithm

Step 1: *I_I denotes the noisy input image and define the patch size $P \times P$ for comparing patches.*

Step 2: *Initialize parameters such as search window size $S \times S$, similarity threshold h, and filtering parameter σ.*

Step 3: *For each pixel (x,y) in the image I_I:*

Step 4: *Define a search window $W(x,y)$ centered at (x,y) with size $S \times S$.*

Step 5: *For each pixel (x',y') within the search window $W(x,y)$:*

5.1 Extract patches $P(x,y)$ and $P(x',y')$ centered at (x,y) and (x',y'), respectively.

5.2 Compute the similarity between patches using an improved metric to better capture complex textures and patterns:

$$Sim\big(P(x,y),P(x',y')\big) = exp\left(-\frac{\| P(x,y)-P(x',y') \|_2^2}{h^2}\right)$$

Here, h is the similarity threshold.

Step 6: *Calculate the weight $w(x',y')$ for each pixel (x',y') based on the patch similarity:*

$$w(x',y') = \frac{Sim(P(x,y),P(x',y'))}{\sum_{x'',y''\in W(x,y)} Sim(P(x,y),P(x'',y''))}$$

Step 7: *Compute the denoised pixel value $I_{denoised}(x,y)$:*

$$I_{denoised}(x,y) - \sum_{(x'.y')\in W(x,y)} w(x'.y').I(x'.y')$$

Step 8: *Output the denoised image $I_{denoised}$.*

Improvements in INLM Filter

The proposed method enhances patch similarity calculation by utilizing a more advanced similarity metric, specifically the exponential decay of patch differences, which significantly improves the discrimination between relevant and irrelevant patches. Additionally, the Parameters h, σ and S are tuned empirically or adaptively to optimize denoising performance across various noise levels and image characteristics. Improves weight calculation by normalizing based on the sum of similarities within the search window, ensuring that contributions from more similar patches are prioritized. It implements optimizations to handle large search windows efficiently, possibly through data structures or algorithmic improvements, maintaining computational feasibility. These enhancements collectively contribute to the Improved Non-local Means (INLM) filter's ability to effectively denoised images while preserving finer details, making it suitable for a wide range of image processing applications.

3.3 Feature extraction using GLCM

The spatial relationship between adjacent pixels in an image is quantified by the GLCM. It contains complete grayscale image information on the following: direction, surrounding interval, and range capability. To put it simply, every element $X_{d,\theta}$ (i,j) in the co-occurrence matrix denotes the likelihood that the grayscale j and i will appear at a certain d-spatial distance and θ-orientation. Typically, orientation is selected from the following four directions: 0°, 45°, 90°, and 135° for the horizontal, vertical, and left diagonal, respectively. The spatial distance, d, is typically one and is a member of the set of positive numbers. GLCM is a matrix whose elements X (i,j) are thus obtained for $(i,j) \in (Nc, Nc)$, where No is the number of grayscale image. In other words, for (d,θ) fixed. Table 1 lists the GLCM features that Haralick established.

No.	Feature Name	Notation used	Formulation
1	Contrast	CONTRA	$\sum_{i=1}^{N_G}\sum_{j=1}^{N_G}(i-j)^2.P(i,j)$
2	Correlation	CORRE	$\frac{\sum_{i=1}^{N_G}\sum_{j=1}^{N_G}(i-\mu_x)(j-\mu_y).P(i,j)}{\sigma_x\sigma_y}$
3	Energy	ENERG	$\sum_{i=1}^{N_G}\sum_{j=1}^{N_G}[P(i,j)]^2$
4	Homogeneity	HOMOG	$\sum_{i=1}^{N_G}\sum_{j=1}^{N_G}\frac{P(i,j)}{1+(i-j)^2}$
5	Sum of square: variance	SUMOF	$\sum_{i=1}^{N_G}\sum_{j=1}^{N_G}(i-\mu)^2.P(i,j)$
6	Entropy	ENTRO	$-\sum_{i=1}^{N_G}\sum_{j=1}^{N_G}P(i,j).\log[P(i,j)]$

Since the high-dimensional feature vector can deteriorate classification performance, PCA is used to reduce the dimensionality of the high-dimensional feature vector, including the extracted texture characteristics. Only the pertinent principle components linear transformations of the initial, uncorrelated features are chosen by the PCA technique. Only the most important of these main components are used to identify the GLCM features that have a higher correlation. These features are used as the classifier's inputs and are identified as pertinent GLCM features.

3.4 Classification

Fuzzy Logic System:

Fuzzy logic is utilized to handle the uncertainty and imprecision inherent in medical data. Fuzzy sets and membership functions are defined for each input feature, capturing the gradual transition between healthy and diseased states. Rules are formulated based on expert knowledge, facilitating the inference process that mimics human reasoning. Use membership functions to transform input data from sharp values into fuzzy ones. For example, if x is a feature vector, the membership function $\mu(x)$ defines the degree to which x belongs to a fuzzy set.

$$\mu_i(x) = \frac{1}{1 + (\frac{x \;\; c_i}{\sigma_i})^2}$$

Where $\mu_i(x)$ is the membership value of x for the i-th fuzzy set, c_i is the center, and σ_i is the spread of the fuzzy set.

Apply fuzzy rules to generate output fuzzy sets. If R_i represents a fuzzy rule, the rule can be expressed as:

$$R_i : \text{ IF } x_1 \; is \; A_1 \text{ AND } x_2 \text{ is } A_2 \text{ THEN } y \text{ is } B_i$$

Where $A_1, A_2, \ldots$ are fuzzy sets for input features and B_i is the output fuzzy set. Combine the results of all fuzzy rules to produce a fuzzy output. The fuzzy output $\bar{y}$ is obtained by aggregating the contributions of all rules.

$$\bar{y} = \sum_i w_i . B_i$$

Where w_i is the weight of the i-th rule, typically the minimum or product of the membership values of the inputs. Convert the fuzzy output $\bar{y}$ to a crisp value y_f using defuzzification methods like the centroid method.

$$y_f = \frac{\sum_i \mu_{\bar{y}}(y_i) . y_i}{\sum_i \mu_{\bar{y}}(y_i)}$$

Where $\mu_{\bar{y}}$ the membership is function of the fuzzy output, and y_i are the discrete values of the output variable.

Support Vector Machine:

SVM is a powerful machine learning algorithm known for its effectiveness in classification tasks. In this study, SVM is employed to classify the extracted features into healthy and DR-affected categories. The kernel trick is used to handle non-linearity, mapping input features into higher-dimensional spaces where they become more separable.

Construct a feature vector X that includes both original features and the output of the fuzzy logic system y_f.

$$X = [\, x_1, x_2\, , \ldots, x_n, y_f]$$

Train the SVM model using the constructed feature vector X. The SVM aims to find the optimal hyperplane that separates the classes. The decision function for SVM is given by:

$$f(X) = \sum_{i=1}^{N} \alpha_i y_i K(X_i, X) + b$$

Where α_i, are the Lagrange multipliers, y_i are the class labels, X_i are the support vectors, K is the kernel function, and b is the bias term.

To make the feature vector X more separable, map it into a higher-dimensional space using a kernel function $K(X_i, X)$. Common kernel functions include the radial basis function (RBF):

$$K(X_i, X) = \; exp(-\gamma \parallel X_i \; - \; X \, 2 \parallel^2)$$

Where γ is a parameter that determines the width of the Gaussian kernel.

Hybrid Model Integration:

The fuzzy logic system's output is fed into the SVM as an additional feature, enhancing the model's ability to handle complex, nonlinear relationships. This hybrid approach leverages the interpretability of fuzzy logic and the classification power of SVM, resulting in a robust predictive model.

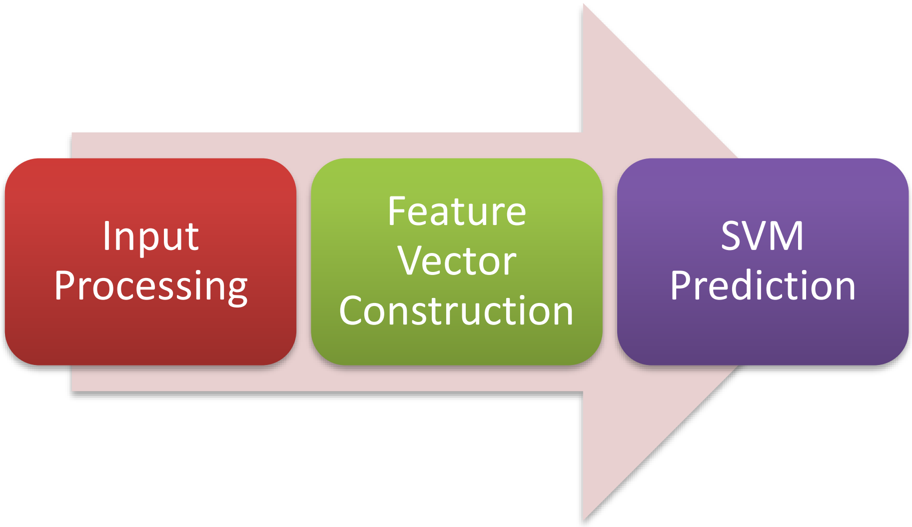

Figure 3.Hybrid Model Integration

For a given input, first process it through the fuzzy logic system to obtain y_f. Construct the feature vector X using the original input features and y_f.Use the trained SVM model to predict the class label $\hat{y}$ for the constructed feature vector X.

$$\hat{y} = sign\left(\sum_{i=1}^{N} \alpha_i y_i K(X_i, X) + b\right)$$

This hybrid approach leverages the interpretability of fuzzy logic and the classification power of SVM, resulting in a robust predictive model capable of handling complex, nonlinear relationships.

Step 1.Input: Pre-processed and feature-selected data D.

Step 2.Fuzzy Logic System:

- Define membership functions $\mu_i(x)$ for each feature x in D.
- Formulate fuzzy rules R_i based on expert knowledge.
- For each input x in D:
- Calculate membership values $\mu_i(x)$.
- Apply fuzzy rules to generate output fuzzy sets B_i.
- Aggregate output fuzzy sets to produce fuzzy output $\overline{y}$.
- Defuzzify $\bar{y}$ to obtain crisp output y_f.

Step 3.Feature Vector Construction:

- Construct feature vector $X = [x_1, x_2, \dots, x_n, y_f]$for each input x in D.

Step 4.SVM Training:

- Train SVM model on constructed feature vectors X:
- Find optimal hyper plane using:

$$f(X) = \sum_{i=1}^{N} \alpha_i y_i K(X_i, X) + b$$

- Choose kernel function $K(X_i, X)$, e.g., radial basis function (RBF):

$$K(X_i, X) = exp(-\gamma \parallel X_i - X\ 2 \parallel^2)$$

Step 5.Hybrid Model Prediction:

- For new input 2:
- Process x through fuzzy logic system to obtain y
- Construct feature vector $X = [x_1, x_2, \dots, x_n, y_f]$.
- Predict class label y using trained SVM model:

$$\overline{y} = sign\left(\sum_{i=1}^{N} \alpha_i y_i K(X_i, X) + b\right)$$

Step 6: Output: Predicted class label $\bar{y}$ for each input data point.

The predicted class label $\bar{y}$ for each input data point, indicating whether the input is classified as healthy or diabetic retinopathy-affected.

4. Result Analysis

1. Accuracy

The degree of agreement between a measurement and its actual value is known as accuracy. The formula for accuracy is:

$$Accuracy = \frac{(true\ value - measured\ value)}{true\ value} * 100$$

Dataset	KNN	SVM	Proposed SC-FSIDR-PF
100	70	63	88
200	73	65	90
300	77	68	93
400	80	71	96
500	82	73	98

Table 1.Comparison table of Accuracy

The Comparison table 1 of Accuracy demonstrates the different values of existing KNN, SVM and Proposed SC-FSIDR-PF. While comparing the Existing algorithm and Proposed SC-FSIDR-PF, provides the better results. The existing algorithm ranges from 70 to 82 and 63 to 73, while the proposed SC-FSIDR-PF method achieves values ranging from 88 to 98, demonstrating significantly improved results.

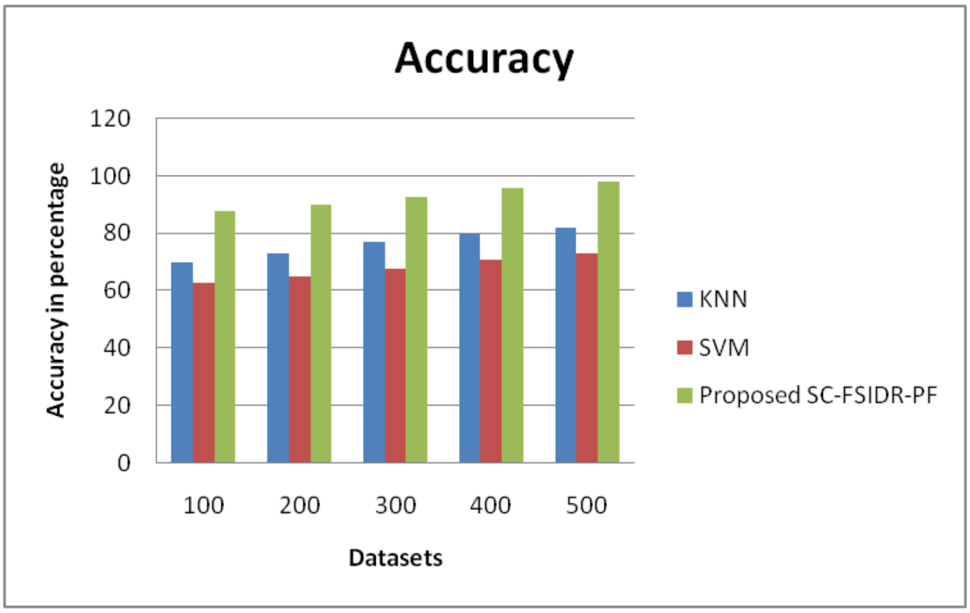

Figure 4.Comparison chart of Accuracy

The Figure 4 Shows the comparison chart of Accuracy demonstrates the existing KNN, SVM and Proposed SC-FSIDR-PF. X axis denote the Dataset and y axis denotes the Accuracy ratio. The Proposed SC-FSIDR-PF values are better than the existing algorithm. The existing algorithm ranges from 70 to 82 and 63 to 73, whereas the Proposed SC-FSIDR-PF method achieves values ranging from 88 to 98, demonstrating superior performance.

2. Precision

Precision is a measure of how well a model can predict a value based on a given input. The precision of a model is the ratio of true positive predictions to all positive predictions.

$$Precision = \frac{true\ positive}{(true\ positive + false\ positive)}$$

Dataset	KNN	SVM	Proposed SC-FSIDR-PF
100	83.23	86.73	98.02
200	81.91	84.05	95.81

300	79.82	82.18	93.28
400	77.32	79.86	91.63
500	74.14	77.32	89.30

Table 2.Comparison table of Precision

The Comparison table 2 of Precision demonstrates the different values of existing KNN, SVM and Proposed SC-FSIDR-PF. When comparing the Existing algorithm and Proposed SC-FSIDR-PF, the latter yields superior results. The existing algorithm ranges from 74.14 to 83.23 and 77.32 to 86.73 and Proposed SC-FSIDR-PF values starts from 89.30 to 98.02. The proposed method provides the great results.

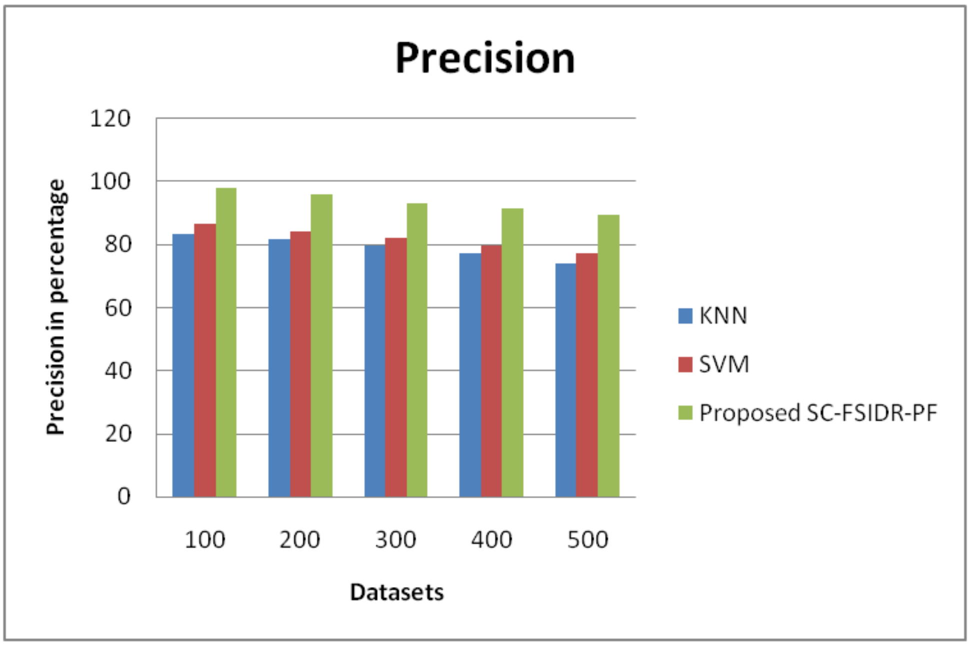

Figure 5.Comparison chart of Precision

The Figure 5 Shows the comparison chart of Precision demonstrates the existing KNN, SVM and Proposed SC-FSIDR-PF. X axis denote the Dataset and y axis denotes the Precision ratio. The Proposed SC-FSIDR-PF values are better than the existing algorithm. The existing algorithm values start from 74.14 to 83.23, 77.32 to 86.73 and Proposed SC-FSIDR-PF values starts from 89.30 to 98.02.The proposed method provides the great results.

3. Recall

Recall is a measure of a model's ability to correctly identify positive examples from the test set:

$$\boldsymbol{Recall} = \frac{\boldsymbol{True\ Positives}}{(\boldsymbol{True\ Positives} + \boldsymbol{False\ Negatives})}$$

Dataset	KNN	SVM	Proposed SC-FSIDR-PF
100	67	74	89
200	70	76	91
300	72	79	93
400	75	82	96
500	80	84	97

Table 3.Comparison table of Recall

The Comparison table 3 of Recall demonstrates the different values of existing KNN, SVM and Proposed SC-FSIDR-PF. While comparing the Existing algorithm and Proposed SC-FSIDR-PF, provides the better results. The existing algorithm ranges from 67 to 80 and 74 to 84, while the Proposed SC-FSIDR-PF method achieves values between 89 and 97, indicating superior results from the proposed approach.

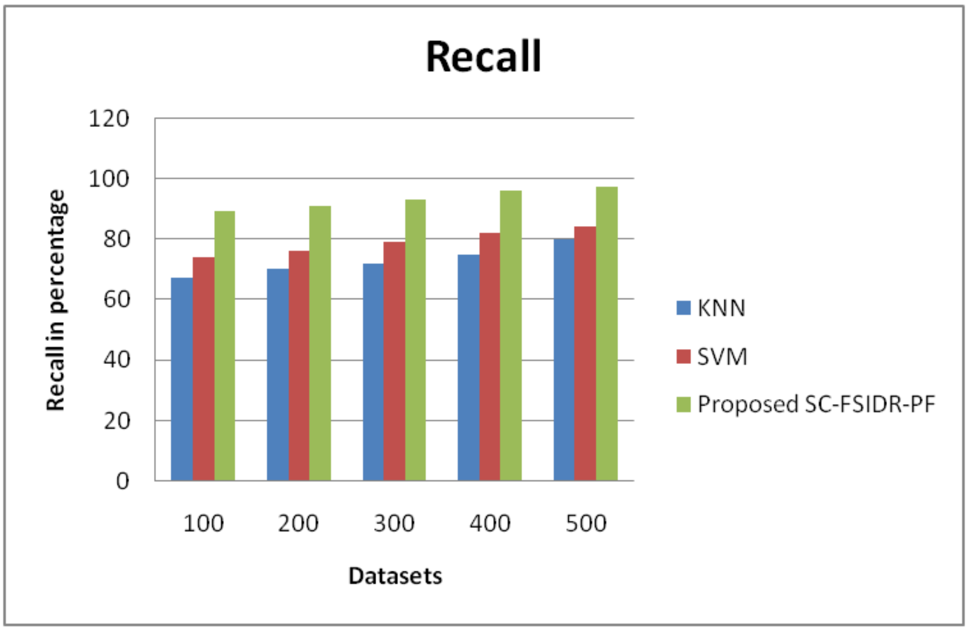

Figure 6.Comparison chart of Recall

The Figure 6 Shows the comparison chart of Recall demonstrates the existing KNN, SVM and Proposed SC-FSIDR-PF. X axis denote the Dataset and y axis denotes the Recall ratio. The Proposed SC-FSIDR-PF values are better than the existing algorithm. The existing algorithm ranges from 67 to 80 and 74 to 84, while the Proposed SC-FSIDR-PF method achieves values between 89 and 97, showcasing significant improvements in results.

4. F -Measure

F1-measure is a test's accuracy that combines precision and recall. It is calculated by taking the harmonic mean of precision and recall.

$$\boldsymbol{F1 - Measure = \frac{(2 * Precision * Recall)}{(Precision + Recall)}}$$

Dataset	KNN	SVM	Proposed SC-FSIDR-PF
100	86	79	98
200	84	74	96
300	83	70	95
400	80	69	93
500	78	65	91

Table 4.Comparison table of F -Measure

The Comparison table 4 of F -Measure Values explains the different values of existing KNN, SVM and Proposed SC-FSIDR-PF. While comparing the Existing algorithm and Proposed SC-FSIDR-PF, provides the better results. The existing algorithm ranges from 78 to 86 and 65 to 79, whereas the Proposed SC-FSIDR-PF method achieves values ranging from 91 to 98, demonstrating significant improvements in results.

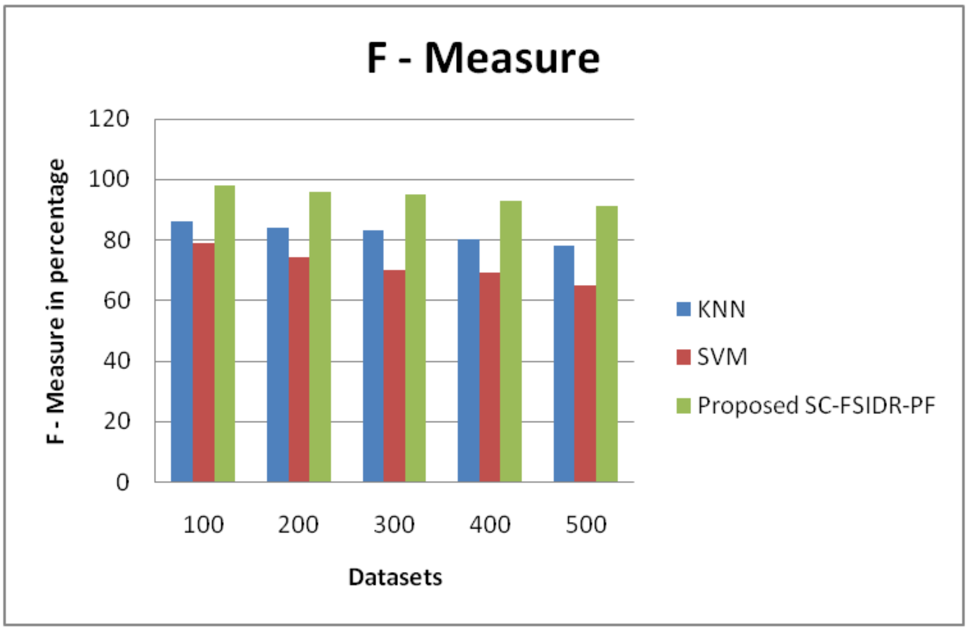

Figure 7.Comparison chart of F -Measure

The Figure 7 Shows the comparison chart of F -Measure demonstrates the existing KNN, SVM and Proposed SC-FSIDR-PF. X axis denote the Dataset and y axis denotes the F -Measure ratio. The Proposed SC-FSIDR-PF values are better than the existing algorithm. The existing algorithm ranges from 78 to 86 and 65 to 79, whereas the Proposed SC-FSIDR-PF method achieves values ranging from 91 to 98, highlighting its superior performance.

5. Conclusion

This study demonstrates the efficacy of a hybrid model combining soft computing techniques, specifically Fuzzy Logic and Support Vector Machine (SVM), is using retinal scans to anticipate the onset of diabetic retinopathy (DR). By integrating these two methods, the model effectively addresses the challenges of uncertainty and non-linearity in medical data. The use of the Improved Non-local Means (INLM) filter for pre-processing and Gray-Level Co-occurrence Matrix (GLCM) for feature extraction, followed by Principal Component Analysis (PCA) for dimensionality reduction, ensures high-quality input for the fuzzy logic system and SVM classifier. The hybrid approach enhances the model's ability to handle complex patterns and improves classification accuracy. The results highlight the potential of this model for early detection of DR, which is crucial for timely intervention and preventing vision loss in diabetic patients. Future work can explore the application of this hybrid soft computing model to other medical imaging tasks and further optimize the parameters for even better performance.

References

[1] K. A. Anant, T. Ghorpade and V. Jethani, "Diabetic retinopathy detection through image mining for type 2 diabetes," 2017 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2017, pp. 1-5, doi: 10.1109/ICCCI.2017.8117738.

[2] M. Panwar, A. Acharyya, R. A. Shafik and D. Biswas, "K-nearest neighbor based methodology for accurate diagnosis of diabetes mellitus," 2016 Sixth International Symposium on Embedded Computing and System Design (ISED), Patna, India, 2016, pp. 132-136, doi: 10.1109/ISED.2016.7977069.

[3] K. S. Argade, K. A. Deshmukh, M. M. Narkhede, N. N. Sonawane and S. Jore, "Automatic detection of diabetic retinopathy using image processing and data mining techniques," 2015 International Conference on Green Computing and Internet of Things (ICGCIoT), Greater Noida, India, 2015, pp. 517-521, doi: 10.1109/ICGCIoT.2015.7380519.

[4] W. Ding, "SVM-Based Feature Selection for Differential Space Fusion and Its Application to Diabetic Fundus Image Classification," in IEEE Access, vol. 7, pp. 149493-149502, 2019, doi: 10.1109/ACCESS.2019.2944899.

[5] D. K. Prasad, L. Vibha and K. R. Venugopal, "Early detection of diabetic retinopathy from digital retinal fundus images," 2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS), Trivandrum, India, 2015, pp. 240-245, doi: 10.1109/RAICS.2015.7488421.

[6] Kishor A, Chakraborty C. Early and accurate prediction of diabetics based on FCBF feature selection and SMOTE. International Journal of System Assurance Engineering and Management. 2021; 3(12):1–9.

[7] Sosale B, Aravind SR, Murthy H, Narayana S, Sharma U, Gowda SGV, Naveenam M. Simple Mobile-based Artificial intelligence algorithm in the detection of diabetic Retinopathy (SMART) study. BMJ Open Diabetes Research & Care. 2020; 8(1):1–6.

[8] Bhatti UA, Huang M, Wu D, Zhang Y, Mehmood A, Han H. Recommendation system using feature extraction and pattern recognition in clinical care systems. Enterprise Information Systems. 2018; 13(3): 329–351.

[9] Leena Nesamani S, Nirmala S, Josphine M, S Jacinth Salome J. Deep Learning-Based Mammogram Classification for Breast Cancer Diagnosis Using Multi-Level Support Vector Machine. Springer Lecture Notes in Electrical Engineering.2021; 700(2): 371-383.

[10] Goh JKH, Cheung CY, Sim SS, Tan PC, Tan GSW, Wong TY. Retinal imaging techniques for diabetic retinopathy screening. Journal of Diabetes Science and Technology. 2016; 10(2): 282–294.

[11] R.J. Winder, P.J. Morrow, I.N. McRitchie, J.R. Bailie, P.M. Hart,Algorithms for digital image processing in diabetic retinopathy,Computerized Medical Imaging and Graphics,Volume 33, Issue 8,2009,Pages 608-622,ISSN 0895-6111.

[12] Singh R, Ramasamy K, Abraham C, Gupta V, Gupta A. Diabetic retinopathy: an update. Indian J Ophthalmol. 2008 May-Jun;56(3):178-88. PMID: 18417817; PMCID: PMC2636123.

[13] Dai, L., Wu, L., Li, H. *et al.* A deep learning system for detecting diabetic retinopathy across the disease spectrum. Nat Commun 12, 3242 (2021).

[14] Rao, K.K., SVP Raju, G. (2011). An Overview on Soft Computing Techniques. In: Mantri, A., Nandi, S., Kumar, G., Kumar, S. (eds) High Performance Architecture and Grid Computing. HPAGC 2011. Communications in Computer and Information Science, vol 169. Springer, Berlin, Heidelberg.