

Frost Denoised Regressive Feature Extraction based Relevance Vector Classification for Rheumatoid Arthritis Disease Prediction

G.Hemamalini¹, Dr. V.Maniraj²

Submitted: 03/02/2024 Revised: 11/03/2024 Accepted: 17/03/2024

Abstract Objective: The proposed method aims to predict Rheumatoid Arthritis (RA) disease using Deep Learning Techniques.

Methods: A new technique called Frost Denoised Regressive Feature Extraction-based Relevance Vector Classification (FDRFE-RVC) has been introduced. The FDRFE-RVC technique performs three key processes: image denoising, feature extraction, and classification. It first utilizes an Improved Frost Denoising Filter to enhance image quality and minimize the mean square error. Then, it applies the Michael Index DeFries-Fulker (MIDF) Regression to extract shape, color, and texture features.

Finding: It uses Kernelized Relevance Vector Classification with the extracted features to predict the disease with greater efficiency. An experimental assessment of the FDRFE-RVC technique reveals a significant improvement in accuracy levels, achieving 95% accuracy and a reduced time of 42 milliseconds compared to the existing Novel Gaussian filtering and segmentation algorithm and HGWO-C4.5 methods.

Novelty: The FDRFE-RVC technique offers a promising solution for RA diagnosis with improved accuracy, enhanced image quality, and faster diagnosis times. The proposed FDRFE-RVC technique reduces disease prediction time by 36% and 24%, reduces the false positive rate by 6%, and increases disease prediction accuracy by 11% compared to the existing method.

Keywords: Rheumatoid arthritis, chronic inflammatory, image denoising, improvised frost, denoising filter, kernelized relevance vector classification.

1. Introduction

Rheumatoid Arthritis often involves extra-articular organs, resulting in symptoms such as pain and swelling. In [1], a Novel Gaussian filtering and segmentation algorithm was introduced to eliminate unnecessary disturbances and partition the affected area of an image with greater accuracy. However, while the accuracy level was improved, the time consumption was not reduced. Another method for RA prediction, HGWO-C4.5, which combined C4.5 with the hybridization of GWO and PSO, was introduced in [2], but it also did not reduce time consumption. Demographic and clinical information were combined in [3] to find the intervention. Precision medicine in RA aims at disease prediction and prevention. However, the computational complexity was not reduced.

A study of Human leukocyte antigen (HLA)-DR tissue was conducted in [4] to find the severity of rheumatoid arthritis. HLA-DR recommended serum copper

assessment significance in predicting RA disease. Disease activity score (DAS) and Health Assessment Questionnaire (HAQ) were used in [5] to investigate the relationship between RA and cardiovascular disease (CVD), and the hazard ratios were determined using Cox regression. However, HAQ did not reduce the error rate. The phenotype Knowledge Base (PheKB) algorithm was introduced in [6] to predict RA using computational phenotyping approaches. However, the algorithm had low sensitivity with positive RA cases for phenotyping. The RATING system was designed [7] to improve the clinical assessment of RA from ultrasound images, and it was introduced with deep learning to address real clinical problems. In [8], IPA was carried out to find enriched gene pathways, while DNCM-ICSA was introduced in [9] with an ensemble classifier to identify affected individuals. Additionally, a deep convolutional neural network was introduced in [10] for automated and fast scoring of X-ray images of patients.

Techniques involving computational analysis, including Machine Learning, play a substantial role in genomics and healthcare. Modern methods for assessing gene expression, like microarray technology and RNA sequencing, generate vast volumes of data. Conventionally, statistical methods are employed to compare this gene expression data. Nevertheless, for more intricate tasks like classifying sample observations

¹Research Scholar, Department of Computer Science, A.V.V.M Sri Pushpam College (Autonomous) Poondi, Thanjavur. Affiliated to Bharathidasan University, Tiruchirappalli, Tamil Nadu, India.
E-mail: hemamalini1984@gmail.com

²Associate Professor & Research Supervisor, Department of Computer Science, A.V.V.M Sri Pushpam College (Autonomous) Poondi, Thanjavur. Affiliated to Bharathidasan University, Tiruchirappalli, Tamil Nadu, India.
E-mail: drmanirajv@gmail.com

or identifying significant genes, advanced computational methods are necessary.

Nevertheless, the issues identified from the literature included higher time consumption, lesser accuracy, higher computational cost, higher computational complexity, higher error rates, and so on. To address these issues, a new technique called Frost Denoised Regressive Feature Extraction-based Relevance Vector Classification (FDRFE-RVC) Technique has been introduced.

The main contributions of the FDRFE-RVC Technique are as follows:

It is introduced to improve the prediction performance of rheumatoid arthritis disease in terms of accuracy. The technique utilizes an Improved Frost Denoising Filter to enhance image quality and MIDF Regression to extract shape, color, and texture features. Kernelized Relevance Vector Classification is performed with the extracted features to efficiently predict the disease, resulting in improved disease prediction performance with higher accuracy, lesser time consumption, and reduced error rates. The article is organized into five sections. Section 2 discusses related works, while Section 3 explains the proposed FDRFE-RVC Technique, complete with a diagram. In Section 4, experimental evaluation is presented, including parameters such as accuracy, false positive rate, and time. Finally, Section 5 presents the conclusion.

2. Related Works

In a study aimed at forecasting relapse among rheumatoid arthritis (RA) patients, machine learning was employed, utilizing data from ultrasound (US) examinations and blood tests [11]. The XGBoost classifier was utilized for prediction, and it identified ten crucial factors, including microvascular imaging scores of the wrist and metatarsophalangeal regions, using the recursive feature elimination technique. The performance of this model exceeded that of traditional prognostic markers chosen by researchers. These findings indicate that machine learning can deliver precise predictions of RA relapse, potentially enabling personalized treatment strategies.

[12] Explored various aspects relevant to gaining insights into the prevention of rheumatoid arthritis (RA). These aspects include discussions about the development of Pre-RA and its pathogenesis, the ability to predict the likelihood and timing of future RA diagnosis, and a comprehensive overview of both completed and ongoing clinical trials related to RA prevention. While new medications and strategies that aim for effective disease management have enhanced control over RA, many

individuals still experience persistent symptoms and do not return to a symptom-free 'pre-RA' state. These challenges, in addition to the high costs associated with RA management, potential medication side effects, and increasing limitations in accessing rheumatology care on a global scale, highlight the potential benefits of preventive approaches for addressing RA.

[13] Suggested a combined approach that incorporates the Complexity parameter, the harmonic fusion vector bispectrum (HFVB), and intrinsic time-scale decomposition (ITD). The outcome suggests that the HFVB method effectively integrates fault information, while the complexity parameter within the Hjorth parameter can act as the key parameter for selecting sensitive characteristic elements of rubbing faults. Utilizing this proposed method, the square demodulation spectrum of reconstructed signals can accurately and efficiently reveal the characteristics of rotor-stator rubbing faults and successfully classify the type of fault.

In a [14] study, a novel approach was introduced for analyzing the medical history of patients with heart disorders. This approach included a new method for selecting and ranking features, specifically targeting highly informative features for early detection of Myocardial Infarction (MI). Since the contribution of various features varied, a combination of varying features (VCF) algorithms was proposed, and probabilistic principal component analysis (PPCA) was applied to enhance feature extraction. The resulting feature vectors underwent covariance analysis, and PPCA identified those vectors with the highest covariance. Consequently, both VCF and PPCA reduced the dataset's dimensionality, effectively addressing the issue of high dimensionality. Subsequently, the most significant features selected were subjected to multi-linear regression (MLR) to identify tightly related combinations.

[15] Gathered a variety of statistical and computational techniques employed in the examination of expression microarray data. Although these methods were deliberated within the scope of expression microarrays, they were also adaptable for scrutinizing RNA sequencing and quantitative proteomics datasets. The data generated by both microarray and RNA sequencing based on NGS underwent several quality checks before analysis. This data was subsequently transformed into a numerical matrix, with genes and samples being represented as rows and columns.

[16] Examined the existing knowledge gaps and encountered technical and ethical challenges in the application of AI. Automated models had been extensively employed to enhance the diagnosis of RA

since the early 2000s, utilizing a diverse range of techniques, such as support vector machines, random forests, and artificial neural networks. AI algorithms had the potential to streamline the screening and identification of at-risk groups, diagnose through omics, imaging, clinical, and sensor data, detect patients within electronic health records (EHR) - a process known as phenotyping, assess treatment responses, monitor the progression of the disease, predict prognosis, discover new drugs, and elevate fundamental scientific research.

Deep neural networks learned from extensive past experiences and had the potential to predict future disease activity, serving as clinical decision support. AdaptiveNet (AN), a newly developed adaptive recurrent neural network, was specifically designed to handle diverse and incomplete clinical data. [17] Examined AN's use in predicting individual disease activity in patients from the rheumatoid arthritis (RA) registry. AN demonstrated greater efficacy in forecasting numeric RA disease activity when compared to traditional machine learning approaches. It's worth noting that all models investigated had limitations in terms of low specificity.

[18] Data from the electronic health record, comprising information on medications, patient characteristics, laboratory results, and previous assessments of disease activity, were organized. A longitudinal deep learning model was employed to forecast the disease activity for rheumatoid arthritis (RA) patients during their subsequent rheumatology clinic appointments. Additionally, the performance of the model and strategies for interhospital compatibility were assessed. Model performance was quantified through the area under the receiver operating characteristic curve (AUROC). RA disease activity was assessed using a composite index score.

The existing method was found to have several limitations, including heightened time consumption, reduced accuracy, greater computational cost, increased computational complexity, and elevated error rates,

3. METHODOLOGY

among others. To overcome these challenges, a proposed technique, Frost Denoised Regressive Feature Extraction-based Relevance Vector Classification (FDRFE-RVC) Technique, was introduced. The FDRFE-RVC Technique is aimed at enhancing the accuracy of RA disease prediction. The technique employs an Improved Frost Denoising Filter to improve image quality, and MIDF Regression to extract shape, color, and texture features. The extracted features are then used in Kernelized Relevance Vector Classification, resulting in improved disease prediction performance marked by heightened accuracy, reduced time consumption, and lower error rates.

A high-performance PLSDA classification model was introduced in [19] for diagnosing RA patients. A personalized approach was introduced in [20] to achieve fast remission in every patient, with a designed approach employed to avoid disability and restore patients without redundant adverse effects in a cost-effective manner. A new framework was introduced in [21] for the evaluation of finger movement patterns in patients with RA. The designed framework employed a 3D hand pose estimation method to visualize 3D skeleton tracking results. A grading method was introduced in [22] for identifying the textural features of bone erosion, and a linear discriminant analysis (LDA) model was designed in [23] to forecast disease activity consistently with validated disease activity indices.

An extracted bone surface with a joint-specific anatomical model was introduced in [24] for the coarse localization of the joint capsule region. A fully automatic method was introduced in [25] for joint location in RA, with the location detection process studied using image features of the joint region. A novel imaging method was introduced in [26] with a classification system for arthritic finger joints, which categorized the finger joints based on their inflammatory status. A two-variable model was introduced in [27] to study the interaction of anti-inflammatory cytokines. A new program that relies on improved user participation was studied in [28] for selecting assistive devices.

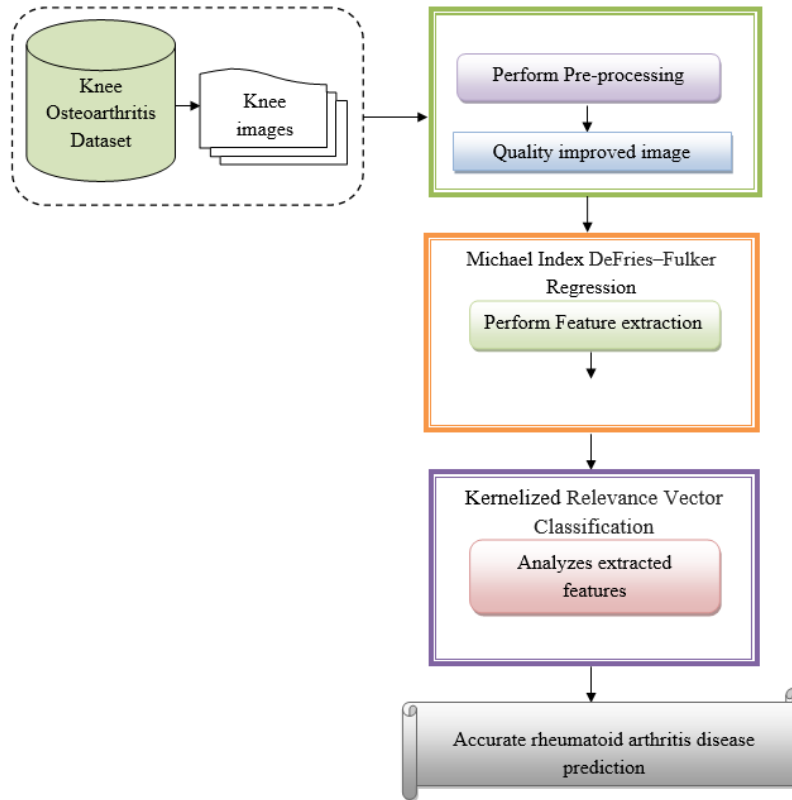


Figure 1. Structural Diagram of the Proposed FDRFE-RVC Technique

Rheumatoid arthritis (RA) is a persistent autoimmune disorder characterized by joint inflammation." Drawing inspiration from this, a new method known as FDRFE-RVC has been presented. This approach is designed to enhance the accuracy of diagnosing and managing rheumatoid arthritis.

Figure 1 illustrates the structural diagram of the FDRFE-RVC technique with three processes: preprocessing, feature extraction, and classification. Let us consider the knee image database 'D' and collect the number of knee images $ki_1, ki_2, ki_3, \dots, ki_n$ from the input database. Image denoising is performed using an improvised Frost

denoising filter. Then, MIDF Regression is used to extract features from the input knee images. Finally, the classification process is carried out using a kernelized relevance vector machine.

2.1 Improved Frost Denoising Filter

The image denoising process is carried out to improve the raw image quality. The image denoising process is employed for noise removal and contrast enhancement. The proposed FDRFE-RVC technique employs an Improved Frost Denoising Filtering process to maintain the image's sharpness after eliminating the noise.

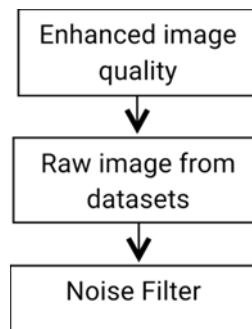


Figure 2. Structure of Improved Frost Denoising Filter

Figure 2 illustrates the structure of the improvised Frost-Denoising Filter in the FDRFE-RVC Technique. Let us consider the number of knee images

$ki_1, ki_2, ki_3, \dots, ki_n$ and their pixels $pi_0, pi_1, pi_2, \dots, pi_m$ are placed in a window in the form of rows and columns.

pi_0	pi_1	pi_2	pi_3	pi_4
pi_5	pi_6	pi_7	pi_8	pi_9
pi_{10}	pi_{11}	pi_{ij}	pi_{13}	pi_{14}
pi_{15}	pi_{16}	pi_{17}	pi_{18}	pi_{19}
pi_{20}	pi_{21}	pi_{22}	pi_{23}	pi_{24}

Figure 3. 5*5 Filtering Window

Figure 3 depicts the pixel arrangement in a 5x5 matrix. The center value is obtained by arranging the pixels in ascending order. The center pixel is then replaced with the weighted sum of the neighboring pixel values. The image-denoising process is determined by the following equation:

$$DO = |p_{ij} - pi_j| \quad (1)$$

In equation (1), ‘DO’ represents the denoised output, ‘ pi_{ij} ’ represents the center pixel, and ‘ pi_j ’ represents the neighboring pixel. The pixel value that differs from the center value is considered a noisy pixel and is replaced by taking the mean of all pixels in the filtering window. By following these steps, noise from the images can be eliminated and the image quality can be improved. The algorithmic steps for the image-denoising process are given below.

Algorithm 1: Improvised Frost Denoising Filter
Input: Knee image database, knee images ‘ $ki_1, ki_2, ki_3, \dots, ki_n$ ’
Output: Quality-enhanced knee image
Begin
1: For each knee image ‘ ki ’
2: Organize pixels $pi_0, pi_1, pi_2, \dots, pi_m$ in the row and column
3: Identify the center pixels from the filtering window
4: Identify and replace the noise pixels with the average of all pixels
5: Obtain quality enhanced knee image
6: end for
End for

Algorithm 1 provides the algorithmic steps of knee image denoising in FDRFE-RVC Technique to attain quality enhanced image. The input knee images are gathered from the input database. Image pixels are placed in the filtering window. The pixel deviates from the center value termed noisy pixels and is replaced by the weighted mean of neighboring pixels. Finally, the quality-enhanced knee image is obtained.

2.2 Michael Index DeFries–Fulker Regressed Feature Extraction

The feature extraction process is performed in the FDRFE-RVC Technique for knee disease prediction. An input knee image comprises local and global features. Consequently, the FDRFE-RVC Technique uses MIDF Regression to extract features for knee disease prediction. DeFries–Fulker Regression is a regression analysis used for determining environmental effects.

MIDF Regression is employed to extract features such as color, shape, and texture from the denoised knee image based on similarity. It is given as follows:

$$RO = B_1P + B_2R + K \quad (2)$$

From (2) ‘RO’, represents the regression output, ‘P’ represents the proband score, ‘K’ and ‘R’ denote the regression coefficients, ‘ B_1 ’ denotes co-twin relatedness while ‘ B_2 ’ represents the extreme scores of the knee image. Based on the regression analysis, color, shape, and texture features are extracted from the input knee image. The texture feature is determined based on the pixel intensity correlation, which depends on the mean and standard deviation, and is computed as follows:

$$Tex = \sum_i \sum_j \frac{1}{deviation^2} [(p_i - mean)(p_j - mean)] \quad (3)$$

From (3), Tex represents the texture correlation between pixel p_i and neighboring pixels p_j depends on the mean and deviation. The color features are extracted from the input image after converting into the HSV (hue, saturation, value).

$$Color = \frac{1}{m} * P \quad (4)$$

From (4), ' P ' symbolizes the pixel intensity while ' m ' denotes the total number of pixels in the knee image. The shape features are determined based on the contour representation. The knee image is converted into

2-D space and the center is represented as the origin(0,0). The distance between the origin and boundary is symbolized as,

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (5)$$

From (5), ' D ' represents the distance, (x_1, y_1) denotes an origin. The point (x_2, y_2) indicates the perimeter of an object. Then, the shape of the boundary is extracted. In this manner, all image features are extracted from preprocessed images. The algorithmic steps of the feature extraction process are given as,

Algorithm 2: Michael Index DeFries–Fulker Regressed Feature Extraction	
Input:	Preprocessed knee image
Output:	Extract knee features
Begin	
Step 1:	For each preprocessed knee image ' ki '
Step 2:	Apply similarity-based regression to identify the best features
Step 3:	Extract texture, color, and shape features
end for	

Algorithm 2 illustrates the feature extraction process. The regression function is used to identify the shape, color, and texture among the multiple features in the region. The feature extraction process of the proposed FDRFE-RVC Technique reduces the disease prediction time.

2.3 Kernelized Relevance Vector Classification

The classification process is the third step in knee disease diagnosis. The kernelized relevance vector is a machine learning method with two possible classification results, i.e., $Y_i \in \{+1, -1\}$. The kernelized relevance vector uses two separating hyperplanes to categorize the knee images as normal or diseased images. The hyperplane acts as the decision boundary between the two classes.

$$DB \rightarrow \alpha \cdot ki_i + b = 0 \quad (6)$$

From equation (6), ' H_d ' symbolizes the decision boundary. ' α ' denotes the normal weight vector to

training images, and ' b ' denotes the bias. Two marginal hyperplanes are created on either side of the decision boundary.

$$H_1 \rightarrow \alpha \cdot ki_i + b > 0 \quad (7)$$

$$H_2 \rightarrow \alpha \cdot ki_i + b < 0 \quad (8)$$

From (7) and (8), ' H_1 ' and ' H_2 ' symbolize the upper and lower marginal hyperplanes to classify the knee images on either side of the decision boundary. The output of the relevance vector machine with the help of the kernel function is calculated as:

$$CR = sign \sum we k_s(ki_i, ki_t) \quad (9)$$

From (9), ' CR ' symbolizes the classification results. ' we ' represent the weight value. ' k_s ' symbolizes the kernel function between the training knee image (ki_i) and testing knee images (ki_t), ' $sign$ ' indicates the classification output as either positive (+1) or negative (-1).

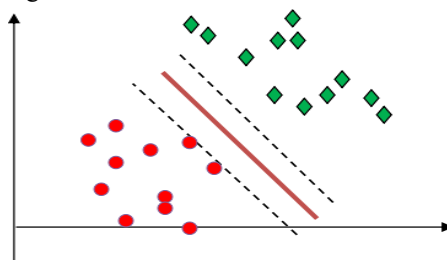


Figure 4. Kernelized Relevance Vector Machine

Figure 4 illustrates the kernelized relevance vector machine used to categorize knee images. As shown in the figure, knee images on the upper side of the hyperplane are classified as diseased, whereas those on the lower side are classified as normal. This leads to improved accuracy in knee disease prediction with a minimal false positive rate.

4. Results and Discussion

3.1 Experimental setup

The experimental analysis of the FDRFE-RVC technique and two existing methods namely the Novel Gaussian filtering and segmentation algorithm [1] and C4.5 with hybridization of Grey Wolf Optimization (GWO) and Particle Swarm Optimization (PSO) termed HGWO-C4.5 [2] were implemented using MATLAB simulator with Knee Osteoarthritis Datasets with KL Grading - 2018. The URL of the datasets is given as <https://www.kaggle.com/datasets/tommyngx/kneeo>. The dataset includes 4796 participants with ages ranging from 45 to 79 comprising 4130 X-ray images with 8260 knee joints. The experimentation analysis is carried out with different performance metrics like disease prediction accuracy, false-positive rate, and disease prediction time to several images.

3.2 Impact of Disease Prediction Accuracy

Disease prediction accuracy is described as the number of knee images that are correctly classified into different classes to the total number of knee images from the input dataset. The disease prediction accuracy is computed as,

$$DPA = \left[\frac{\text{Number of knee images correctly classified}}{\text{Total number of knee images}} \right] * 100 \quad (10)$$

From (10), 'DPA' symbolizes the prediction accuracy, and it is measured in percentage (%). There are three different methods are used for computing the prediction

accuracy results. Let us consider 800 knee images for performing disease prediction. The proposed FDRFE-RVC technique attained 89% disease prediction accuracy whereas the Novel Gaussian filtering and segmentation algorithm [1] and HGWO-C4.5 [2] attained 83% and 85% respectively. Likewise, ten different disease prediction accuracy results are observed for each method.

Figure 5 illustrates the experimental results of disease prediction accuracy using three methods against the number of knee images. Among these methods, the proposed FDRFE-RVC technique achieves better results than the other methods. This improvement is due to the application of MIDF Regression for extracting shape, color, and texture features, as well as Kernelized Relevance Vector Classification for efficient knee disease prediction with better accuracy. The average comparison outcomes show that the suggested FDRFE-RVC technique increases the disease prediction accuracy by 11% and 7% when compared to the Novel Gaussian filtering and segmentation algorithm [1] and HGWO-C4.5 [2], respectively.

3.3 Impact of False Positive Rate

The false positive rate (FPR) refers to the proportion of knee images in the input dataset that are inaccurately categorized into various classes, expressed as a ratio between the misclassified knee images and the total number of knee images. The FPR is calculated as,

$$FPR = \left[\frac{\text{Number of knee images incorrectly classified}}{\text{Total number of knee images}} \right] * 100 \quad (11)$$

From (11), 'FPR' symbolizes the false positive rate. The false positive rate is measured in percentage (%).

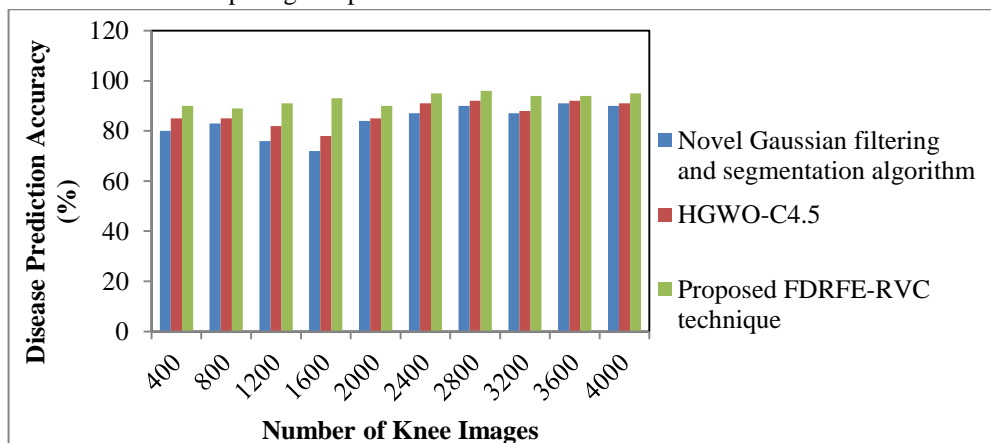


Figure 5 Measurement of Disease Prediction Accuracy

Let us consider that 3600 knee images for performing disease prediction. The proposed FDRFERVC technique attained a 6% false positive rate whereas the Novel Gaussian filtering and segmentation algorithm [1] and HGWO-C4.5 [2] attained 9% and 8% respectively.

Figure 6 shows the experimental outcomes of the FPR of three methods versus the number of knee images.

Among three different methods, the proposed FDRFE-RVC technique reduces the false positive rate more than the existing methods. This improvement is due to the application of MIDFRegression and Kernelized Relevance Vector Classification results in more efficient detection of Rheumatoid arthritis with a lower false positive rate.

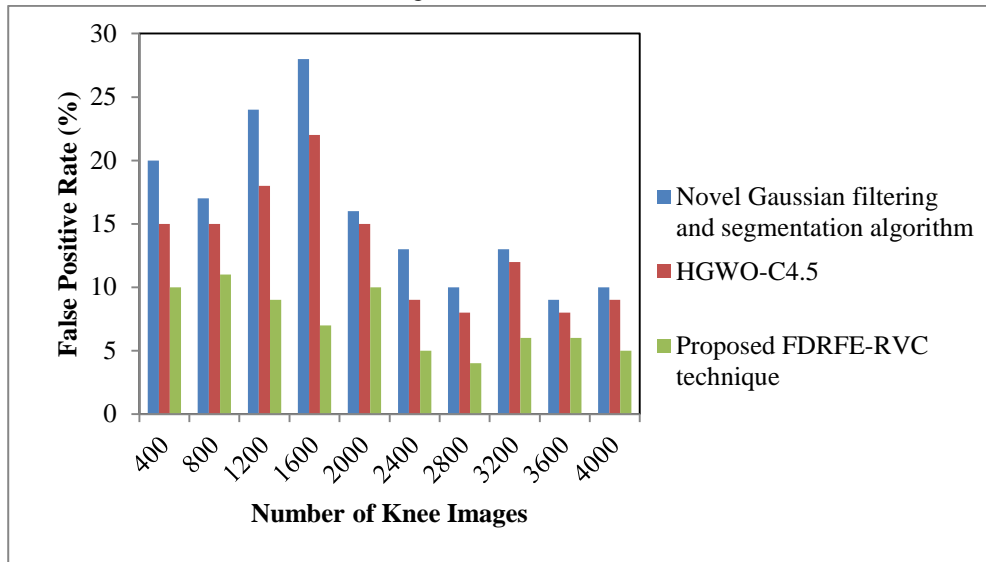


Figure 6 Measurement of False Positive Rate

As a result, the FDRFE-RVC technique reduces the false positive rate. The average results of the comparison show that the proposed FDRFE-RVC technique reduces the false positive rate by 52% and 43% compared to the existing Novel Gaussian filtering and segmentation algorithm [1] and HGWO-C4.5 [2], respectively.

3.4 Impact of Disease Prediction Time

Disease prediction time is computed as the amount of time consumed to perform the disease prediction through the classification process. Consequently, the disease prediction time is determined as,

$$DPT = n * t [kip] \quad (12)$$

From (12), 'DPT' symbolizes the disease prediction time. 'n' denotes the number of knee images. 't[kip]' represents the time consumed for single knee image prediction. It is measured in terms of milliseconds (ms).

Let us consider using 2400 knee images for performing disease prediction. The proposed FDRFE-RVC

technique took 32ms to perform disease prediction, whereas, the Novel Gaussian filtering and segmentation algorithm [1] and HGWO-C4.5 [2] consumed 48ms and 41ms, respectively.

Figure 7 shows the experimental results of disease prediction time for three methods versus the number of knee images. Among three different methods, the proposed FDRFE-RVC technique reduces the disease prediction time than the existing methods. This is due to the use of MIDF Regression and Kernelized Relevance Vector Classification which are more efficient predictions of Rheumatoid arthritis disease with reduced time consumption. The average comparison results show that the proposed FDRFE-RVC technique reduces the disease prediction time by 36% and 24% when compared to the existing Novel Gaussian filtering and segmentation algorithm [1] and HGWO-C4.5 [2] respectively.

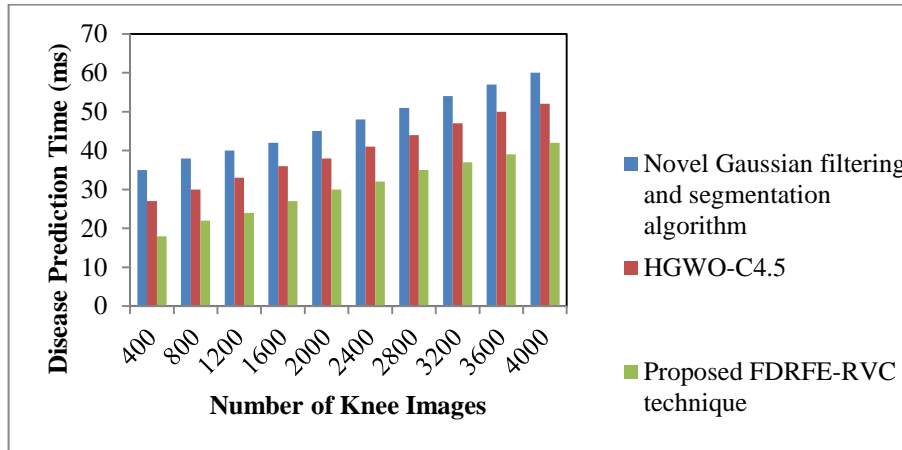


Figure 7 Measurement of Disease Prediction Time

The utilization of the kernelized relevance vector machine has been suggested as a means to classify knee images. The categorization involves designating knee images in the proposed work to show that the upper side of the hyperplane as displaying pathological characteristics, while those located on the lower side are deemed to be within normal parameters. This approach yields a significant improvement in the accuracy of predicting knee ailments, accompanied by a negligible incidence of false positives. The classification model employed is intended to enhance the precision of knee disease prediction while minimizing the false positive rate. In addition, the FDRFE-RVC Technique has been developed to streamline the feature extraction process and reduce the time needed for disease prediction. The computation of disease prediction time is based on the duration of the classification process. As such, our methodology surpasses existing methods.

Results show that the proposed FDRFE-RVC technique achieved a high disease prediction accuracy of 89% and a low false positive rate of 6%, with a disease prediction time of 32ms using Kernelized Relevance Vector Classification and MIDF Regression. In comparison, the Novel Gaussian filtering and segmentation algorithm [1] and HGWO-C4.5 [2] achieved disease prediction accuracy and false positive rate of 83% and 85%, respectively. Furthermore, the Novel Gaussian Filtering and segmentation algorithm [1] and HGWO-C4.5 [2] required 48ms and 41ms, respectively, to perform the disease prediction process.

5. Conclusion

A new technique called the FDRFE-RVC Technique improves the prediction performance of RA disease. The FDRFE-RVC Technique performs image denoising, feature extraction, and classification. Image denoising removes noise and improves image quality using an

Improved Frost Denoising Filter. After that, MIDF Regression in the FDRFE-RVC Technique extracts the shape, color, and texture. Finally, the Kernelized Relevance Vector Classification (KRVC) categorizes the images using the extracted features to predict RA disease more efficiently.

The Frost denoising algorithm enhances the quality of the input data. This denoising step helps to reduce the impact of noise and improve the signal-to-noise ratio, which can lead to more accurate predictions. Next, a regressive feature extraction process is employed to identify the most informative and relevant features from the denoised data. This feature extraction step helps to reduce the dimensionality of the data and retain only the most discriminative features, which can improve the performance of the subsequent classification algorithm.

In addition, an experiment was conducted to assess the performance of the proposed FDRFE-RVC Technique alongside existing methods, using three metrics related to the number of knee images. The results indicate that the FDRFE-RVC Technique achieved higher accuracy in predicting disease while requiring less time for prediction and demonstrating a lower false positive rate compared to existing methods. Specifically, the FDRFE-RVC technique reduced disease prediction time by 36% and 24% and decreased the false positive rate by 6%. Additionally, it improved disease prediction accuracy by 11% and 7% when compared to the existing method. It's worth noting that prompt treatment is essential for addressing RA-induced shoulder affliction. Delaying the treatment of RA in the shoulder can lead to joint damage and loss of function. Therefore, it is advisable to use advanced technologies for early diagnosis of RA in the shoulder region to achieve higher accuracy and faster management of this condition.

References

- [1] D. Uma Nandhini, Janani Naga Sai Pravallika, Beulah Jabaseeli and S. Udhayakumar, "Prediction of Rheumatoid Arthritis disease using improved segmentation algorithm", *Materials Today Proceedings*, Elsevier, Volume 62, Part 7, 2022, Pages 4940-4945. <https://doi.org/10.1016/j.matpr.2022.03.684>
- [2] Shanmugam Sundaramurthy and Preethi Jayavel, "A hybrid Grey Wolf Optimization and Particle Swarm Optimization with C4.5 approach for prediction of Rheumatoid Arthritis", *Applied Soft Computing*, Elsevier, Volume 94, September 2020, Pages 1-15. <https://doi.org/10.1016/j.asoc.2020.106500>
- [3] Michael Mahler, Laura Martinez-Prat, Jeffrey A. Sparks and Kevin D. Deane, "Precision medicine in the care of rheumatoid arthritis: Focus on prediction and prevention of future clinically-apparent disease", *Autoimmunity Reviews*, Elsevier, Volume 19, Issue 5, May 2020, Pages 1-15. <https://doi.org/10.1016/j.autrev.2020.102506>
- [4] Khalid Ahmad Omer Aldabbagh and Dashty Abbas Al-Bustany, "Relationship of serum copper and HLADR4 tissue typing to disease activity and severity in patients with rheumatoid arthritis: A cross sectional study", *Annals of Medicine and Surgery*, Elsevier, Volume 73, January 2022, Pages 1-15. doi: 10.1016/j.amsu.2021.103193
- [5] Hong Ki Min, Hae-Rim Kim, Sang-Heon Lee, Kwi Young Kang, Sung-Hwan Park, Dae-Seung Ki, Kwok "Time-averaged DAS28 and HAQ predict cardiovascular disease in patients with rheumatoid arthritis: Data from KORONA registry", *Joint Bone Spine*, Elsevier, Volume 89, Issue 5, October 2022, Pages 1-14. DOI: 10.1016/j.jbspin.2022.105401
- [6] Henry W. Zheng, Veena K. Ranganath, Lucas C. Perry, David A. Chetrit, Karla M. Criner, Angela Q. Pham, Richard Seto, Sitaram Vangala, David A. Elashoff and Alex A. T. Bui, "Evaluation of an automated phenotyping algorithm for rheumatoid arthritis", *Journal of Biomedical Informatics*, Elsevier, Volume 135, November 2022, Pages 1-12. doi: <https://doi.org/10.1016/j.jbi.2022.104214>.
- [7] Zhanping Zhou, Chenyang Zhao, Hui Qiao, Ming Wang, Yuchen Guo, Qian Wang, Rui Zhang, Huaiyu Wu, Fajin Dong, Zhenhong Qi, Jianchu Li, Xinping Tian, Xiaofeng Zeng, Yuxin Jiang, Feng Xu, Qionghai Dai and Meng Yang, "RATING: Medical knowledge-guided rheumatoid arthritis assessment from multimodal ultrasound images via deep learning", *Patterns*, Elsevier, Volume 3, Issue 10, 14 October 2022, Pages 1-15. DOI: 10.1016/j.patter.2022.100592
- [8] Xinqiang Song, Yu Zhang, Erqin Dai, Lei Wang and Hongtao Du, "Prediction of triptolide targets in rheumatoid arthritis using network pharmacology and molecular docking", *International Immunopharmacology*, Elsevier, Volume 80, March 2020, Pages 1-12. <https://doi.org/10.1016/j.intimp.2019.106179>
- [9] B. Chithra and R. Nedunchezian, "Dynamic neutrosophic cognitive map with improved cuckoo search algorithm (DNM-ICSA) and ensemble classifier for rheumatoid arthritis (RA) disease", *Journal of King Saud University - Computer and Information Sciences*, Elsevier, Volume 34, Issue 6, Part B, June 2022, Pages 3236-3246. <https://doi.org/10.1016/j.jksuci.2020.06.011>
- [10] Janick Rohrbach, Tobias Reinhard, Beate Sick and Oliver Dürr, "Bone erosion scoring for rheumatoid arthritis with deep convolutional neural networks", *Computers & Electrical Engineering*, Elsevier, Volume 78, September 2019, Pages 472-481. <https://doi.org/10.1016/j.compeleceng.2019.08.003>
- [11] Matsuo, H., Kamada, M., Imamura, A. et al. Machine learning-based prediction of relapse in rheumatoid arthritis patients using data on ultrasound examination and blood test. *Sci Rep* 12, 7224(2022). <https://doi.org/10.1038/s41598-022-11361-y>
- [12] Deane KD, Holers VM. Rheumatoid Arthritis Pathogenesis, Prediction, and Prevention: An Emerging Paradigm Shift. *Arthritis Rheumatol*. 2021 Feb; 73(2):181-19. doi: <https://doi.org/10.1002/art.41417>
- [13] Yu M, Cong H, Chen W, Fang M. Feature extraction of rotor-stator rubbing faults based on harmonic fusion vector bispectrum, ITD and Hjorth parameters. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*. 2023; 237(3):600-613. doi: <https://doi.org/10.1177/09544100221107256>
- [14] Sulthana, A.R., Jaithunbi, A.K. Varying combination of feature extraction and modified support vector machines based prediction of my

ocardial infarction. *Evolving Systems* 13, 777–794 (2022).<https://doi.org/10.1007/s12530-021-09410-4>

- [15] Bhandari Nikita, WalambeRahee, Kotecha Ketan, Khare Satyajeet P. (2022). A comprehensive survey on computational learning methods for analysis of gene expression data. *Frontiers in Molecular Biosciences*.9.<https://doi.org/10.3389/fmolb.2022.907150>
- [16] Momtazmanesh S, Nowroozi A, Rezaei N. (2022). Artificial Intelligence in Rheumatoid Arthritis: Current Status and Future Perspectives: A State-of-the-Art Review. *Rheumatology and Therapy*.9(5):1249-1304.<https://doi.org/10.1007/s40744-022-00475-4>.
- [17] Kalweit M, Walker UA, Finckh A, Müller R, Kalweit G, Scherer A, Boedecker J, Hügle T. (2021). Personalized prediction of disease activity in patients with rheumatoid arthritis using an adaptive deep neural network. *PLoS One*. 2021.16(6):e0252289.<https://doi.org/10.1371/journal.pone.0252289>
- [18] Norgeot B, GlicksbergBS, Trupin L, et al. (2019). Assessment of a Deep Learning Model Based on Electronic Health Record Data to Forecast Clinical Outcomes in Patients With Rheumatoid Arthritis. *JAMA Network Open*. 2019;2(3):e190606.<https://doi.org/10.1001/jamanetworkopen.2019.0606>
- [19] Neha Chaudhary, Thi Nguyet Que Nguyen, Muddassar Ahmad, Robert Harrington, Caroline A. Jefferies, Grainne Kearns, Aidan D. Meade and Claire Wynne, “A spectroscopic diagnostic for rheumatoid arthritis using liquid biopsies”, *Clinical Spectroscopy*, Elsevier, Volume 3, December 2021, Pages 1-18. <https://doi.org/10.1016/j.clispe.2021.100009>
- [20] C. A. Wijbrandts and P. P. Tak, “Prediction of Response to Targeted Treatment in Rheumatoid Arthritis”, *Mayo Clinic Proceedings*, Elsevier, Volume 92, Issue 7, July 2017, Pages 1129-1143. <https://doi.org/10.1016/j.mayocp.2017.05.009>
- [21] Luciano Walenty Xavier Cejnog, Teofilode Campos, Valeria Meirelles Carril Elui, Roberto Marcondes Cesar, “A framework for automatic hand range of motion evaluation of rheumatoid arthritis patients”, *Informatics in Medicine Unlocked*, Elsevier, Volume 23, 2021, Pages 1-12 DOI:10.1016/j.imu.2021.100544
- [22] Tengfei Yang, Haijiang Zhu, Xiaoyu Gao, Yiyuan Zhang, Yanan Hui and Fangfang Wang, “Grading of Metacarpophalangeal Rheumatoid Arthritis on Ultrasound Images Using Machine Learning Algorithms”, *IEEE Access*, Volume 8, March 2020, Pages 67137 – 67146. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9043487>
- [23] Kristine L. Richardson, Caitlin N. Teague, Samer Mabrouk, Brandi N. Nevius, Goktug C. Ozmen, Rachel S. Graham, Daniel P. Zachs, Adam Tuma, Erik J. Peterson, Hubert H. Lim and Omer T. Inan, “Quantifying Rheumatoid Arthritis Disease Activity Using a Multimodal Sensing Knee Brace”, *IEEE Transactions on Biomedical Engineering*, Volume 69, Issue 12, December 2022, Pages 3772 – 3783. <https://doi.org/10.1109/TBME.2022.3177074>
- [24] Kunlin Cao, David M. Mills, Ralf G. Thiele and Kedar A. Patwardhan, “Toward Quantitative Assessment of Rheumatoid Arthritis Using Volumetric Ultrasound”, *IEEE Transactions on Biomedical Engineering*, Volume 63, Issue 2, February 2016, Pages 449 – 458
- [25] Yinghe Huo, Koen L. Vincken, Desiree van der Heijde, Maria J. H. De Hair, Floris P. Lafeber and Max A. Viergever, “Automatic Quantification of Radiographic Finger Joint Space Width of Patients with Early Rheumatoid Arthritis”, *IEEE Transactions on Biomedical Engineering*, Volume 63, Issue 10, October 2016, Pages 2177 – 2186 DOI: <https://doi.org/10.1109/TBME.2015.2512941>
- [26] A. Schwaighofe, V. Tresp, P. Mayer, A. Krause, J. Beuthan, H. Rost, G. Metzger, G.A. Muller and A.K. Scheel, “Classification of rheumatoid joint inflammation based on laser imaging”, *IEEE Transactions on Biomedical Engineering*, Volume 50, Issue 3, March 2003, Pages 375 – 382 DOI: 10.1109/TBME.2003.808827
- [27] Michelle Baker, Sarah Denman-Johnson, Bindi S. Brook, Ian Gaywood and Markus R. Owen, “Mathematical modeling of cytokine-mediated inflammation in rheumatoid arthritis”, *Mathematical Medicine and Biology: A Journal of the IMA*, Volume 30, Issue 4, December 2013, Pages 311 – 337. <https://doi.org/10.1093/imammb/dqs026>

[28] U. Hass, H. Brodin, A. Andersson and J. Persson, “Assistive technology selection: a study of participation of users with rheumatoid arthritis”, IEEE Transactions on Rehabilitation Engineering, Volume 5, Issue 3, Pages 263 – 275 DOI: 10.1109/86.623018

<https://www.sciencedirect.com/search?qs=rheumatoid%20arthritis%20classification>

<https://www.kaggle.com/datasets/tommyngx/kneea>

<https://www.sciencedirect.com/science/article/pii/S0306456522002182>

Referred links: