

Article

Minimum and Maximum Pattern-Based Self-Organized Feature Engineering: Fibromyalgia Detection Using Electrocardiogram Signals

Veysel Yusuf Cambay^{1,2}, Abdul Hafeez Baig³, Emrah Aydemir⁴ , Turker Tuncer¹  and Sengul Dogan^{1,*} 

¹ Department of Digital Forensics Engineering, Technology Faculty, Firat University, Elazig 23119, Turkey; 231144201@firat.edu.tr (V.Y.C.); turkertuncer@firat.edu.tr (T.T.)

² Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Mus Alparslan University, Mus 49250, Turkey

³ School of Management and Enterprise, University of Southern Queensland, Toowoomba, QLD 4350, Australia; abdul.hafeez-baig@unisu.edu.au

⁴ Department of Management Information Systems, Management Faculty, Sakarya University, Sakarya 54050, Turkey; emrahaydemir@sakarya.edu.tr

* Correspondence: sdogan@firat.edu.tr; Tel.: +90-424-2370000

Abstract: Background: The primary objective of this research is to propose a new, simple, and effective feature extraction function and to investigate its classification ability using electrocardiogram (ECG) signals. Methods: In this research, we present a new and simple feature extraction function named the minimum and maximum pattern (MinMaxPat). In the proposed MinMaxPat, the signal is divided into overlapping blocks with a length of 16, and the indexes of the minimum and maximum values are identified. Then, using the computed indices, a feature map is calculated in base 16, and the histogram of the generated map is extracted to obtain the feature vector. The length of the generated feature vector is 256. To evaluate the classification ability of this feature extraction function, we present a new feature engineering model with three main phases: (i) feature extraction using MinMaxPat, (ii) cumulative weight-based iterative neighborhood component analysis (CWINCA)-based feature selection, and (iii) classification using a t-algorithm-based k-nearest neighbors (tkNN) classifier. Results: To obtain results, we applied the proposed MinMaxPat-based feature engineering model to a publicly available ECG fibromyalgia dataset. Using this dataset, three cases were analyzed, and the proposed MinMaxPat-based model achieved over 80% classification accuracy with both leave-one-record-out (LORO) cross-validation (CV) and 10-fold CV. Conclusions: These results clearly demonstrate that this simple model achieved high classification performance. Therefore, this model is surprisingly effective for ECG signal classification.

Keywords: MinMaxPat; feature engineering; ECG fibromyalgia detection; machine learning



Citation: Cambay, V.Y.; Hafeez Baig, A.; Aydemir, E.; Tuncer, T.; Dogan, S. Minimum and Maximum Pattern-Based Self-Organized Feature Engineering: Fibromyalgia Detection Using Electrocardiogram Signals. *Diagnostics* **2024**, *14*, 2708. <https://doi.org/10.3390/diagnostics14232708>

Academic Editor: Mugahed A. Al-antari

Received: 23 October 2024

Revised: 19 November 2024

Accepted: 25 November 2024

Published: 30 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Fibromyalgia is a chronic pain syndrome associated with the central nervous system [1,2] and is characterized by widespread sleep disorders, musculoskeletal pain, cognitive dysfunction, and fatigue [3]. Patients experience chronic widespread pain in different areas of their bodies, and extreme sensitivity can be seen, especially at certain points. These tender points play an important role in the diagnosis of fibromyalgia [4]. Fatigue and sleep disorders are common in fibromyalgia patients and are often severe enough to limit daily activities [5,6]. Cognitive dysfunction, known as “fibro fog”, is also frequent [7]. There is no specific laboratory test for diagnosing fibromyalgia; however, clinical evaluation and physical examination are essential [8]. The American College of Rheumatology (ACR) criteria use the Widespread Pain Index (WPI) and Symptom Severity Scale (SSS) to assess pain and symptom severity [9,10]. Sleep Polygraphy (Polysomnography) may be used to

evaluate sleep disorders, while laboratory tests such as Complete Blood Count (CBC), Thyroid Function Tests, and Erythrocyte Sedimentation Rate (ESR) help exclude inflammatory and autoimmune diseases [11].

Treatment for fibromyalgia includes medication, physical therapy, and psychotherapy [12]. Common medications are Tricyclic Antidepressants (TCAa), Serotonin–Norepinephrine Reuptake Inhibitors (SNRIs), Gabapentinoids, and Nonsteroidal Anti-Inflammatory Drugs (NSAIDs) [13]. Physical therapy and exercise help reduce pain by strengthening muscles. Cognitive–Behavioral Therapy (CBT) and Mindfulness–Based Stress Reduction (MBSR) are used to manage pain [14]. Artificial intelligence (AI) tools like machine learning (ML) and deep learning (DL) also help doctors understand symptoms and create personalized treatments [15].

1.1. Literature Review

Some studies on fibromyalgia in the literature are presented in Table 1.

Table 1. Literature review.

Study	Data	Method	Split Ratio	Results (%)
Sabeti et al. [16]	Physiological signals	Learning using concave and convex kernels	10-fold CV	Acc: 88.38
Paul et al. [17]	Sleep EEG signals	Nonlinear dynamical features	10-fold CV	Acc: 96.15 Sen: 96.88 Spe: 95.65
Santana et al. [18]	rs-fMRI	Dynamic time warping	5-fold CV	Acc: 86.00 ROC: 93.00
Chatterjee et al. [19]	rs-fMRI	CNN	5-fold CV	Acc: 85.20 Pre: 83.00 Rec: 98.00 F1: 90.00
Fukae et al. [20]	Clinical data	CNN	80:20	Acc: 98.00
Gokcay et al. [21]	fNIRS and clinical data	Likelihood-based decision level fusion	10-fold CV 20-fold CV	Spe: 100.0 Sen: 100.0
Alves et al. [22]	Blood plasma samples	Paper spray ionization–mass spectrometry	70:30	Acc: 100.0
Robinson et al. [23]	Structural magnetic resonance imaging and self-report data	J48 decision tree	10-fold CV	Acc: 96.17
Martín-Brufau et al. [24]	Resting-state EEG recordings	Fast Fourier transform, statistical analysis	Unspecified	High discriminative capacity: 100%
Passos et al. [25]	Blood plasma samples	Chemometric analysis	70:15:15	Acc: 84.20 Sen: 89.50 Spe: 79.00
Orrù et al. [26]	Psychometric tests	J48 decision tree	10-fold CV	Acc: 88.16 AUC: 88.00 F1: 88.00
Alves et al. [27]	Blood plasma samples, clinical data	Principal component analysis	10-fold CV	Acc: 88.00 Sen: 100.0 Spe: 75.00 AUC: 87.00

Table 1. Cont.

Study	Data	Method	Split Ratio	Results (%)
Aksalli et al. [28]	Sleep EEG signals	D'hondt pooling, glucose pattern	1. 10-fold CV 2. Leave-one-record-out	1. Acc: 100.0 Pre: 100.0 Rec: 100.0 F1: 100.0 2. Acc: 99.72 Pre: 99.74 Rec: 99.82 F1: 99.78

rs-fMRI: resting-state functional magnetic resonance imaging; CNN: Convolutional Neural Network; fNIRS: functional near-infrared spectroscopy.

1.2. Literature Gaps

The detected literature gaps highlight that the minimum and maximum statistical moments have been utilized in both deep learning (DL) and feature engineering (FE) [29]. In FE models, these moments are typically employed to generate statistical features or normalization [30], whereas in DL models, they are used for pooling functions and normalization [31]. However, no existing descriptor specifically based on these moments has been identified in the literature, as confirmed by our knowledge and review. Additionally, while DL models have been extensively used to achieve high classification performance [32], they are inherently complex and often involve exponential time complexity [33,34]. Consequently, training such models on local devices requires expensive hardware, such as graphical processing units (GPUs) [35].

1.3. Motivation and Our Model

The major motivation of this research is to fill the existing gaps in the literature. Therefore, we have proposed a new feature extraction function based on simple mathematical definitions, specifically using two basic statistical moments: minimum and maximum. By deploying these moments, we developed a new feature extraction function, which serves as a one-dimensional descriptor. This feature extraction method uses a local feature extraction approximation and is named MinMaxPat (minimum and maximum pattern).

To investigate the classification performance of the MinMaxPat, we proposed a new feature engineering (FE) model. This model consists of three main phases: (i) MinMaxPat-based feature extraction, (ii) feature selection using cumulative weighted iterative neighborhood component analysis (CWINCA), and (iii) classification using a t-algorithm-based k-nearest neighbors (tkNN) [36] classifier. By employing this strategy, a lightweight FE model was created, as the methods used in the MinMaxPat-based FE model have linear time complexity. Thus, the proposed model is both lightweight and efficient for signal classification.

In this research, we tested a new coding schema using the well-known statistical moments—maximum and minimum—and the proposed MinMaxPat. Our goal was to develop a new signal descriptor similar to the local binary pattern (LBP) [37]. In the testing phase, our proposed MinMaxPat surprisingly achieved high classification performance. Therefore, a new feature engineering model was developed by integrating the CWINCA and tkNN methods. The CWINCA feature selection method identified the most informative features from the 256 generated features, while tkNN, an ensemble classifier, generated the classification results. Both CWINCA and tkNN methods are self-organized, making the proposed FE model a self-organized feature engineering (SOFE) model.

1.4. Innovations and Contributions

A novel feature extraction function, termed MinMaxPat, has been developed. This function serves as the foundation for a new self-organized feature engineering (SOFE) model, which is also introduced in this study.

The main motivation behind this work is to design a simple yet effective descriptor similar to the local binary pattern (LBP). The proposed MinMaxPat leverages minimum and maximum statistical moments along with a base-16 coding method to generate features. Additionally, it employs histogram-based feature extraction, making it a significant contribution to the feature extraction research area.

Using MinMaxPat, a new SOFE model has been presented with a streamlined architecture consisting of three phases: (i) feature extraction, (ii) feature selection, and (iii) classification. Features are extracted using MinMaxPat, while the feature selection and classification phases employ two self-organized methods, ensuring efficiency and adaptability. This integrated approach establishes a new paradigm in SOFE research and highlights the potential of simple architectures for achieving high-performance results.

2. Materials and Methods

2.1. Material

A publicly available ECG fibromyalgia dataset [17] was used in this research. This dataset was collected from participants during sleep stage 2 and sleep stage 3, consisting of 139 records. The collected ECG signals are single-lead signals, and the sampling frequency of these ECG signals is 512 Hz. We divided the ECG signals into segments, each with a length of 15 s. Therefore, each ECG segment contains 7680 ($=15 \times 512$) values, which we used as an array of length 7680. There are two main classes in this dataset, (i) fibromyalgia and (ii) control, as the objective of this dataset is to detect differences between healthy and fibromyalgia ECG signals.

Using these ECG signals, we created three cases, and the features of these cases are tabulated in Table 2.

Table 2. Features of the cases.

Case	Class	Number of Records	Number of Segments
1: Sleep stage 2	Control	42	1810
	Fibromyalgia	32	1012
	Total	74	2822
2: Sleep stage 3	Control	36	1497
	Fibromyalgia	29	771
	Total	65	2268
3: Merged	Control	78	3307
	Fibromyalgia	61	1783
	Total	139	5090

In order to obtain results from these ECG signals, we used both 10-fold cross-validation (CV) and leave-one-record-out (LORO) CV to provide reliable results.

2.2. The Presented MinMaxPat-Based SOFE

In this research, we propose a new SOFE model based on the MinMaxPat method. This model consists of three main phases, two of which are self-organized. The phases are as follows: (1) MinMaxPat-based feature extraction, (2) CWINCA-based feature selection, and (3) tkNN-based classification.

To select the most informative features from these 256 features, the CWINCA feature selector was used. By utilizing the selected features as input for the tkNN classifier, the classification results were generated. To better explain the proposed MinMaxPat-based SOFE model, a graphical depiction is shown in Figure 1.

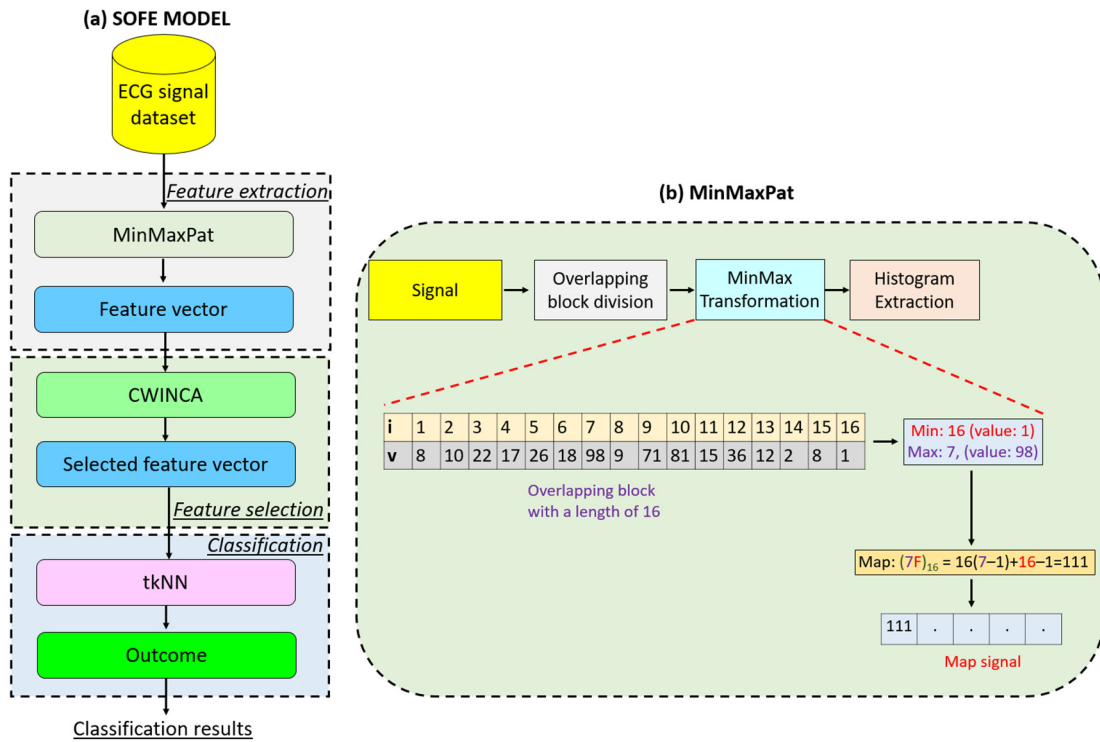


Figure 1. Graphical overview of (a) the proposed MinMaxPat-based SOFE model and (b) the presented MinMaxPat.

The details of the proposed SOFE model are provided step by step below.

Step 1: Apply MinMaxPat to generate 256 features from each ECG signal. The details of the MinMaxPat method are explained below.

In Figure 1b, we demonstrate a block diagram of the proposed MinMaxPat along with a numerical example. The steps of this descriptor are also provided below.

Step 1.1. Divide the signal into overlapping blocks with a length of 16.

$$block^i(j) = ECG(i + j - 1), i \in \{1, 2, \dots, Len - 15\}, j \in \{1, 2, \dots, 16\} \quad (1)$$

where *block*: overlapping block with a length of 16; *ECG*: ECG signal; and *Len*: length of the signal.

Step 1.2. Find indexes of the minimum and maximum values of the block.

$$mini(i) = argmin(block^i) \quad (2)$$

$$maxi(i) = argmax(block^i) \quad (3)$$

Here, *mini*: minimum index and *maxi*: maximum index of the block. Moreover, the *argmin(.)* and *argmax(.)* functions detect indices of the minimum and maximum values.

Step 1.3. Calculate the feature map value.

$$map(i) = 16(maxi(i) - 1) + (mini(i) - 1) \quad (4)$$

where *map*: feature map signal coded as base16.

Step 1.4. Extract the histogram of the feature map signal to obtain the feature vector.

$$feat(d, 1 : 256) = \zeta(map), d \in \{1, 2, \dots, NoS\} \quad (5)$$

Here, *feat*: features; $\zeta(.)$: histogram extraction function; *NoS*: number of sample.

The given Steps 1.1–1.4 are defined as the proposed MinMaxPat feature extraction function.

Step 2. Select the best features by deploying the CWINCA feature selector.

The CWINCA function used is a self-organized feature selector. By deploying cumulative weights, we determined the start and stop indexes for the loop of the INCA [38]. Subsequently, iterative feature selection was applied using the INCA feature selector, and the best feature vectors were selected automatically. In this regard, CWINCA functions as a self-organized feature selector. The steps of this feature selector are defined below.

$$[w, id] = NCA(feats, y) \quad (6)$$

where w : weight of the features; id : indices of the sorted features obtained by computing feature weights; $NCA(\cdot)$: NCA feature selection function; and y : actual output.

Step 2.2. Compute start and stop values by deploying cumulative weight computation. Herein, in order to compute the start and stop values, we have used 0.80 and 0.99 threshold values.

$$start = CW(w, id, 0.80) \quad (7)$$

$$stop = CW(w, id, 0.99) \quad (8)$$

Herein, $start$: start value of the loop; $stop$: stop value of the loop; and $CW(\cdot)$: cumulative weight computation function.

Step 2.3. Select features iteratively and compute the loss values of the selected features.

$$f_s^{r-start+1}(d, i) = feat(d, ind(i)), i \in \{1, 2, \dots, r\}, r \in \{start, start + 1, \dots, stop\} \quad (9)$$

$$loss(r - start + 1) = \mathbb{C}(f_s^{r-start+1}, y) \quad (10)$$

where f_s : selected feature vector; $loss$: the misclassification rate; and \mathbb{C} : classifier.

Step 2.4. Choose the final feature vector by deploying the greedy algorithm.

$$ca(i) = \varphi(f_s^{r-start+1}, y) \quad (11)$$

$$id = \operatorname{argmax}(ca) \quad (12)$$

$$SF = f_s^{id} \quad (13)$$

Here, ca : classification accuracy; $\varphi(\cdot)$: classification accuracy calculation function; id : indices of the maximum classification accuracy; and SF : the final selected feature vector.

Step 3: Classify the selected features by deploying the tkNN classifier. The tkNN classifier is an ensemble and self-organized classifier. In this classifier, iterative parameter changes are applied, and parameter-based outcomes are generated. Subsequently, iterative majority voting (IMV) [39] is applied to these parameter-based outcomes, and voted outcomes are produced. In the final phase of the tkNN algorithm, the best outcome (the one with the maximum classification accuracy) is selected using a greedy algorithm. In this sense, the tkNN classifier is a self-organized classifier. The steps of this classifier are as follows:

Step 3.1. Generate parameter-based outcomes by deploying iterative parameter changes and the kNN classifier.

$$pot_a = kNN(SF, y, D_i, K_j, W_k), a \in \{1, 2, \dots, 90\} \quad (14)$$

$$D \in \{Cityblock, Euclidean, Cosine\} \quad (15)$$

$$K \in \{1, 2, \dots, 10\} \quad (16)$$

$$D \in \{Equal, Inverse, SquaredInverse\} \quad (17)$$

where *pot*: parameter-based outcome; *kNN*(.): kNN classifier [40]; *D*: distance; *K*: k-values; and *W*: weights. We used 3 distances, 10 k values, and 3 weights. Therefore, the tkNN classifier generated 90 ($=3 \times 3 \times 10$) classifier-based outcomes.

Step 3.2. Deploy the IMV algorithm and generate 88 more voted outcomes.

$$ca(a) = \varphi(pot_a, y) \quad (18)$$

$$ix = \text{argsort}(-ca) \quad (19)$$

$$vot_b = \omega(pot_{ix(1)}, pot_{ix(2)}, \dots, pot_{ix(b+2)}), b \in \{1, 2, \dots, 88\} \quad (20)$$

Here, *ix*: the qualified indices; *vot*: voted outcomes; and $\omega(\cdot)$: mode function.

Step 3.3. Choose the best outcome by deploying the greedy algorithm to the generated 178 outcomes.

$$ca(90 + b) = \varphi(vot_b, y) \quad (21)$$

$$x = \text{argmax}(ca) \quad (22)$$

$$FinalOut = \begin{cases} pot_x, & x \leq 90 \\ vot_{x-90}, & x > 90 \end{cases} \quad (23)$$

Here, *x*: index of the maximum classification accuracy and *FinalOut*: the final/ultimate outcome.

These three steps are defined as the proposed MinMaxPat-based SOFE model.

3. Results

In order to investigate the classification performance of the proposed MinMaxPat-based SOFE model, it was applied to an ECG fibromyalgia dataset. The dataset was downloaded, and cases were created for analysis. Subsequently, the MinMaxPat-based SOFE model was developed using MATLAB (version 2024a), employing functional programming with four key functions: (i) main, (ii) MinMaxPat, (iii) CWINCA, and (iv) tkNN.

The used feature extraction function is a basic function and the time complexity of this function is computed as $O(n)$. Herein, *n*: length of the signal and $O(\cdot)$: big O notation. In the feature selection phase, we used an iterative feature selector called CWINCA. The time complexity of the CWINCA feature selector is equal to $O(N + iK)$. Herein, *N*: time complexity coefficient of the NCA [41]; *i*: number of iterations; and *K*: time complexity of the used classifier to compute misclassification rates. The last phase is the classification phase. The tkNN classifier is an iterative classifier. Therefore, the time complexity of this classifier is equal to $O(iK)$. In this aspect, the total time complexity of the recommended model is equal to $O(n + N + iK + iK) \cong O(n + N + iK)$. This computation clearly demonstrated the linear time complexity of the recommended MinMaxPat-based SOFE model. Therefore, we used a central processing unit (CPU) mode to implement the presented MinMaxPat-based model.

The recommended MinMaxPat-based model is a parametric model, and to provide a repetition of this model, the initial parameters used in this model are given in Table 3.

By deploying the parameters listed in Table 3, the proposed MinMaxPat-based model was created.

To obtain the classification results, we used five commonly known performance evaluation metrics: (i) classification accuracy, (ii) geometric mean, (iii) precision, (iv) recall, and (v) F1-score.

Moreover, we applied three cases and used two validation techniques: 10-fold CV and LORO CV. To compute the results, we used confusion matrices for these cases based on the validation technique. The computed confusion matrices are shown in Figure 2.

Table 3. The initial parameters of the introduced MinMaxPat-based SOFE.

Phase	Method	Parameters
Feature extraction	MinMaxPat	Input: ECG signal Number of leads: 1 Length of each overlapping blocks: 16 Base: 16 Output: 256 features
Feature selection	CWINCA	Input: Generated feature vector Loop range detector: Cumulative weights Threshold values: 0.85 for start value and 0.99 for stop value Output: Selected feature vector
Classification	tkNN	Input: Selected feature vector Distances: City block, Cosine, Euclidean k value: from 1 to 10 Weight: Squared inverse, equal, inverse Number of the parameter-based outcome: 90 Number of voted outcome: 88 Selection criteria: Maximum classification accuracy Output: tkNN-based outcomes Parameters of the IMV: Loop range: from 3 to 90

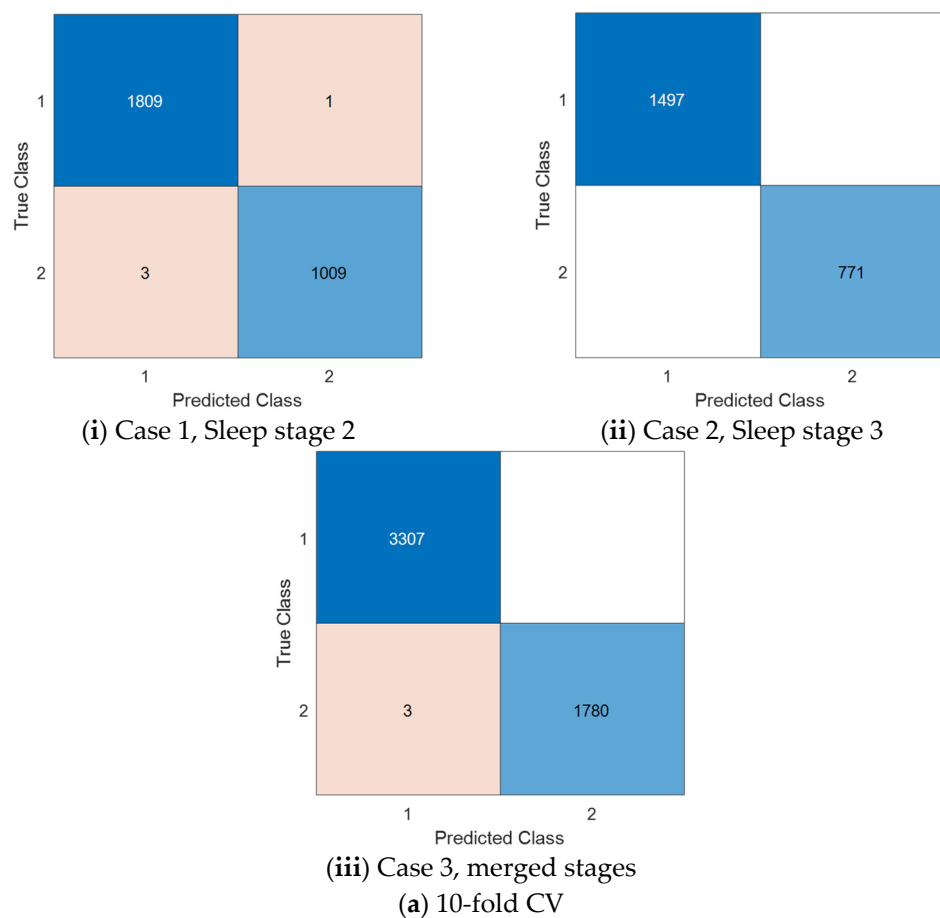


Figure 2. Cont.

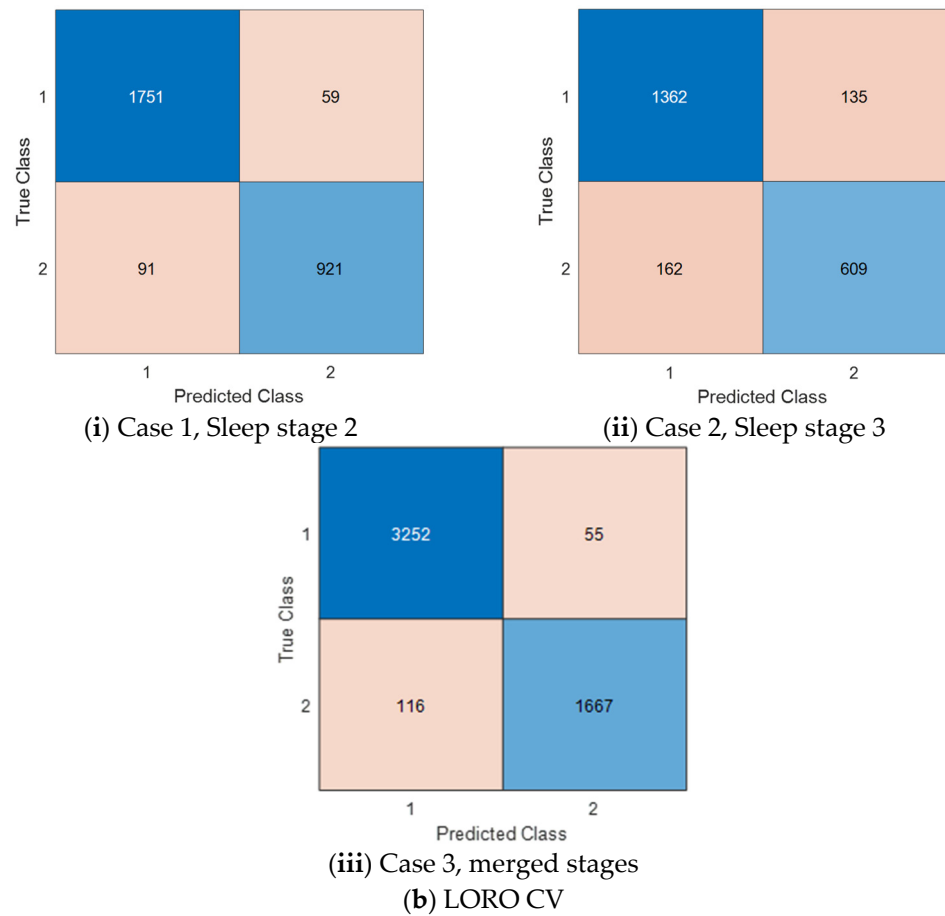


Figure 2. The computed confusion matrices per the cases and the used validation techniques. Herein, 1: control; 2: fibromyalgia. (a) 10-fold CV; (b) LORO CV.

By using the given confusion matrices in Figure 2, the results of this model were computed and are listed in Table 4.

Table 4. The computed classification results.

Validation	Performance Evaluation Metric	Case 1: Sleep Stage 2	Case 2: Sleep Stage 3	Case 3: Merged Sleep Stage
10-fold CV	Accuracy	99.86	100	99.94
	Geometric mean	99.82	100	99.92
	Precision	99.87	100	99.95
	Recall	99.82	100	99.92
	F1-score	99.85	100	99.94
LORO CV	Accuracy	94.68	86.90	96.64
	Geometric mean	93.83	84.77	95.88
	Precision	94.52	85.61	96.68
	Recall	93.87	84.99	95.92
	F1-score	94.18	85.28	96.28

Table 4 shows that the highest classification accuracy and geometric mean were 100% for the ECG signals collected during sleep stage 3. For LORO CV, the best classification accuracy and geometric mean were 96.64% and 95.88%, respectively, in Case 3. Additionally, the proposed MinMaxPat-based model achieved over 84% classification performance for all cases using both validation techniques.

Moreover, the receiver operating characteristic (ROC) curves of the introduced model for the defined Cases 1, 2, and 3 are given in Figure 3.

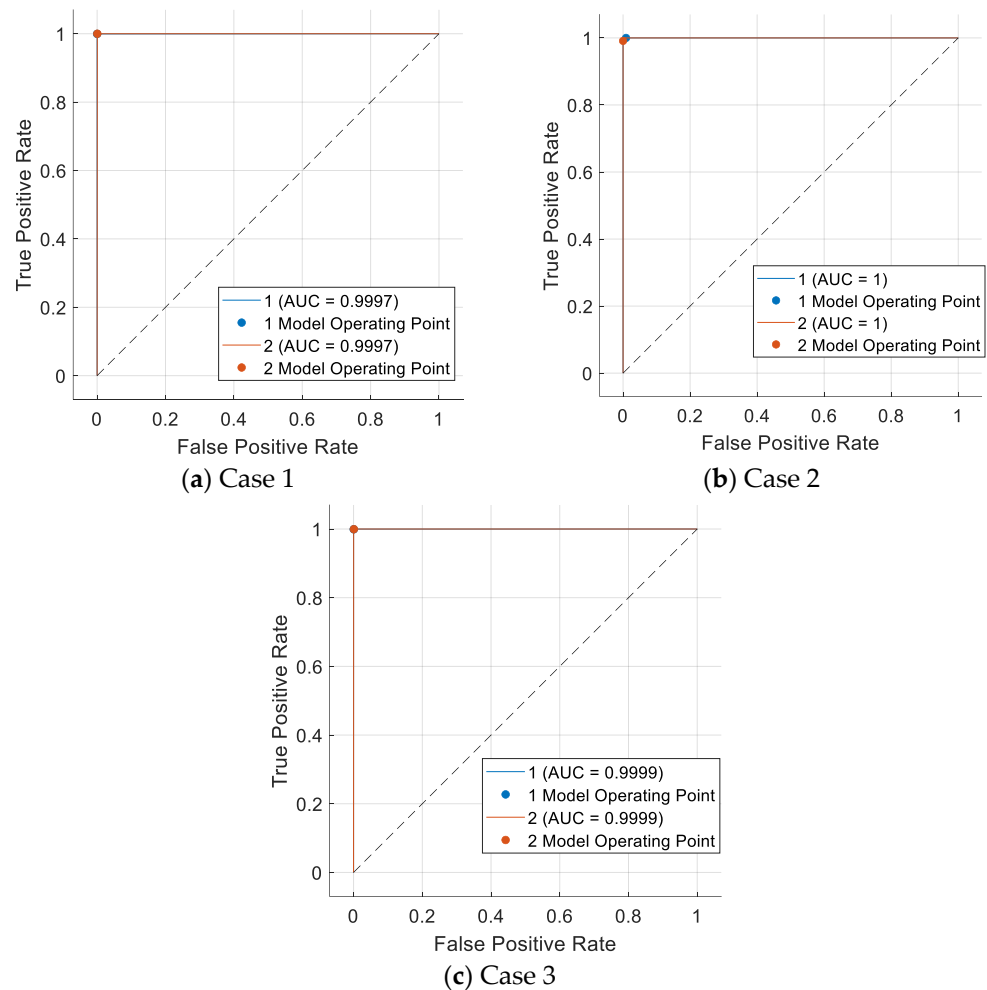


Figure 3. The computed ROC curves. (a) Case 1; (b) Case 2; (c) Case 3.

Per the given ROC curves, all recommended area under curve (AUC) values are above 0.999.

The performance of the MinMaxPat-based SOFE model lies in its ability to maintain high accuracy even with few features. By selecting the most informative features (38 for Case 1, 35 for Case 2, and 31 for Case 3) using CWINCA, the model demonstrates that it can achieve high classification accuracy with minimal computational load. Furthermore, the highest accuracy rates for 10-fold CV were obtained in sleep stage 3 (Case 2), showcasing the model's ability to effectively handle specific and challenging cases. These results clearly demonstrate the effectiveness of the MinMaxPat-based SOFE model. A discussion of the model's results is provided in Section 4, along with a more in-depth analysis.

4. Discussion

In this research, we proposed a new MinMaxPat-based SOFE model to investigate the classification performance of the MinMaxPat feature extraction. We presented a simple structure for the SOFE model, which consists of three main phases. In the feature selection and classification phases, we used self-organized CWINCA and tkNN methods, making this feature engineering model self-organized.

In the feature extraction phase, we extracted 256 features from each ECG signal, and CWINCA selected the most informative 38, 35, and 31 features as the final feature vectors

for Case 1, Case 2, and Case 3, respectively. At this point, the proposed model achieved high classification performance with fewer features.

We also used the 10-fold CV and LORO CV techniques in the MinMaxPat-based SOFE model, and the results are demonstrated in Table 4. According to Table 4, the worst case is Case 2, with a classification accuracy of 86.90% using LORO CV. However, Case 2 performed the best using 10-fold CV, achieving 100% classification accuracy. To better understand the performance of each case, the comparative results are showcased in Figure 4.

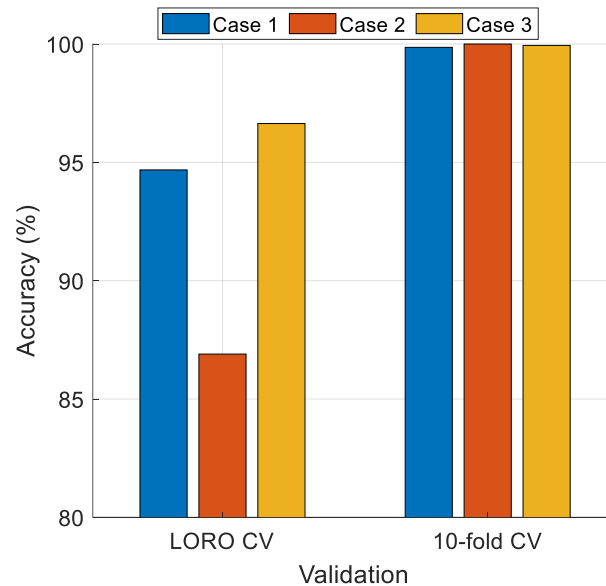


Figure 4. Performance comparison of the proposed model according to cases.

Figure 4 shows that the best cases for LORO CV and 10-fold CV are Case 3 and Case 2, respectively. Moreover, for 10-fold CV, all cases achieved over 99% classification accuracy (see Table 4). Additionally, both Case 1 and Case 3 attained over 90% classification accuracy using both validation techniques.

The tkNN classifier used is an ensemble classifier. Therefore, we compared the classification performance of the tkNN classifier to other ensemble classifiers, namely (i) subspace discriminant (SD), (ii) subspace kNN (SkNN), (iii) bagged tree (BaT), and (iv) boosted tree (BoT). The comparative results of tkNN with these classifiers are displayed in Figure 5.

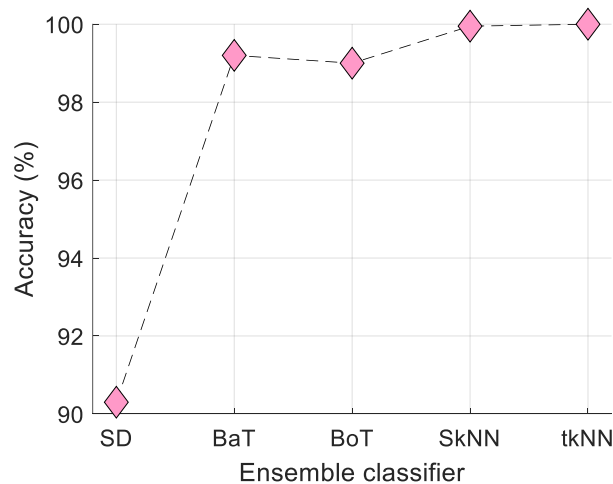


Figure 5. Comparative results of the ensemble classifiers on Case 2 with 10-fold CV.

Figure 5 shows that the closest classifier to tkNN is SkNN, which achieved 99.95% classification accuracy, while tkNN achieved 100% classification accuracy. Therefore, we selected the tkNN classifier for this model.

The proposed MinMaxPat is a simple feature extractor (signal descriptor), inspired by the local binary pattern (LBP). We aimed to create a more effective and simpler feature extraction function than LBP, which led to the development of MinMaxPat. To ensure a fair comparison, we used both LBP and MinMaxPat as feature extractors and conducted ablation studies. The classification results computed using 10-fold CV are also demonstrated in Figure 6.

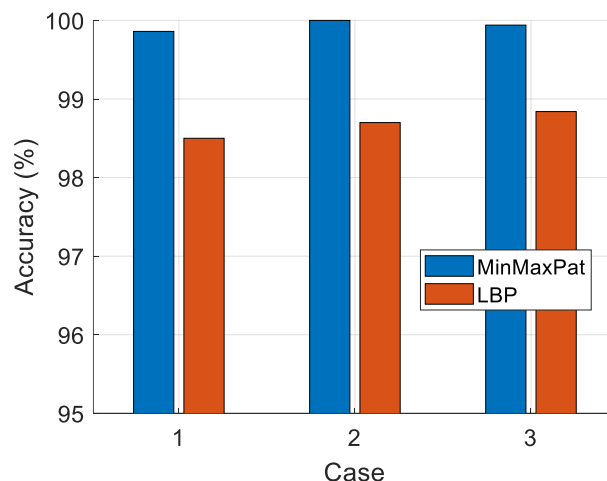


Figure 6. MinMaxPat versus one-dimensional LBP.

Figure 6 clearly demonstrates that the proposed MinMaxPat achieved over 99% classification accuracy for all cases, while the LBP feature extraction function could not reach 99% accuracy in all cases.

To highlight the significance of the proposed model in the literature, we present comparative results in Table 5.

Table 5. Comparative results deploying LORO CV.

Research	Method	Accuracy (%)
Barua et al. [42]	Multiple filter-based discrete wavelet transform, quantum inspired local binary pattern (3LBP), Chi2- and NCA-based feature selection, kNN and support vector machine classifiers, IMV-based information fusion	Case 1: 93.97 Case 2: 92.02
Our model	MinMaxPat, CWINCA, tkNN	Case 1: 94.68 Case 2: 86.90

The results in Table 5 indicate that the proposed model outperforms Barua et al.'s method for Case 1, achieving an accuracy of 94.68% compared to 93.97%. This clearly highlights the effectiveness of the presented MinMaxPat feature extraction approach. While Barua et al.'s method achieves higher accuracy in Case 2, their approach is significantly more complex, requiring multiple feature selectors, classifiers, and a knowledge fusion stage. In contrast, the proposed MinMaxPat-based SOFE model employs a simple three-stage process that includes MinMaxPat for feature extraction, CWINCA for feature selection, and tkNN for classification. Moreover, the comparative results are shown in Figure 7 [42].

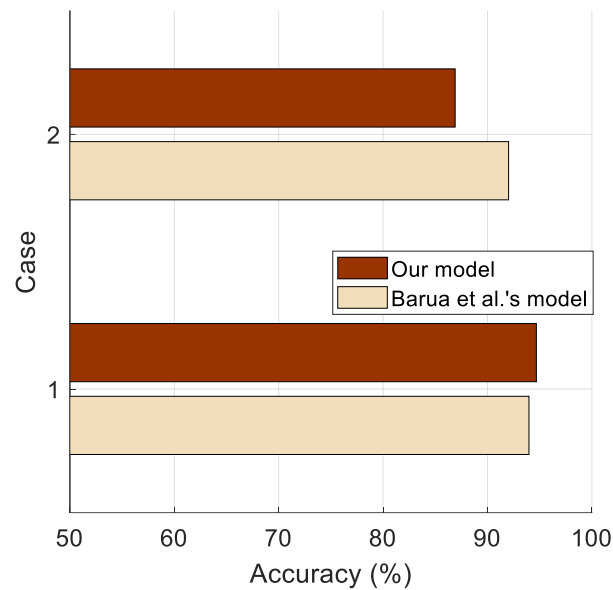


Figure 7. Comparative results with LORO CV (Our model and Barua et al. [42]).

Table 5, Figure 7, and these models' structures demonstrate that the simplicity of the MinMaxPat-based SOFE model does not compromise its effectiveness. The proposed model offers competitive performance in Case 1, presenting an accessible and efficient approach to feature extraction and classification. Furthermore, by focusing on lightweight computation, the MinMaxPat-based SOFE model provides a practical solution for real-time fibromyalgia detection systems, advancing the field of feature engineering by balancing simplicity and performance. The introduced model's straightforward structure allows it to be implemented in various environments, such as FPGA, and it effectively distinguishes ECG signals. Additionally, its adaptability makes it suitable for other signal structures with minimal modifications.

Furthermore, we used an LSTM classifier to obtain the classification results, and the selected features achieved accuracies of 99.81% and 92.30% using LSTM with 10-fold CV and LORO CV, respectively, while the tkNN classifier attained 99.84% and 96.64% classification accuracy with 10-fold CV and LORO CV, respectively. The hyperparameters for the LSTM classifier were as follows: maximum epochs: 50, number of hidden units: 100, mini-batch size: 64, initial learning rate: 0.01, gradient threshold: 1, and validation using 10-fold CV. For Case 3 (the largest case, encompassing Case 1 and Case 2), the classification results of the LSTM and tkNN classifiers are depicted in Figure 8.

Figure 8 clearly shows that the tkNN classifier achieved higher classification accuracy compared to the LSTM classifier for Case 3.

The findings of this research study are as follows: The MinMaxPat-based SOFE model achieved high classification performance across all three defined cases. This success can be attributed to the feature extraction phase, where 256 meaningful features are extracted from each ECG signal using MinMaxPat, and CWINCA selects the most informative features. The model maintained high classification accuracy with a reduced number of features, selecting 38 for Case 1, 35 for Case 2, and 31 for Case 3. The performance of the proposed model was evaluated using both 10-fold cross-validation (CV) and leave-one-record-out (LORO) CV, producing robust and reliable results with both techniques. Notably, Case 2 achieved an accuracy rate of 86.90% with LORO CV and 100% accuracy with 10-fold CV. The proposed SOFE model also demonstrated strong performance in Cases 1 and 3, with Case 3 identified as the best-performing case in both validation techniques.

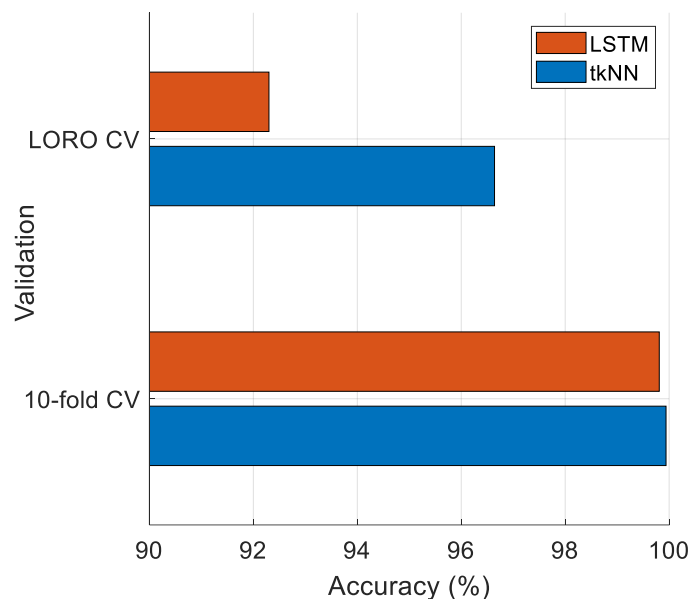


Figure 8. Comparison of tkNN and LSTM for Case 3 deploying 10-fold CV.

The tkNN classifier used in the model was compared against ensemble classifiers such as SD, SkNN, BaT, and BoT to validate its superiority. Using 10-fold CV, tkNN achieved 100% accuracy in Case 2, outperforming its closest competitor, SkNN, which achieved 99.95% accuracy. Furthermore, the MinMaxPat feature extraction method outperformed the classification performance of the LBP feature extractor, proving that it is a simple yet effective method. When compared with existing methods in the literature, the proposed method consistently outperformed its competitors, further validating its effectiveness.

The advantages of the recommended model are given as follows. The proposed model achieved high classification performance, notably reaching 100% accuracy in Case 2 with 10-fold CV. By employing both leave-one-record-out (LORO) CV and 10-fold CV, the model produced robust and reliable results. Despite its simplicity, the MinMaxPat feature extraction function demonstrated remarkable effectiveness in feature extraction and classification tasks. With its linear time complexity, the MinMaxPat-based model is lightweight and computationally efficient. The integration of self-organized methods, specifically CWINCA for feature selection and tkNN for classification, further enhanced the model's performance, establishing it as a self-organized feature engineering (SOFE) approach. Consistently strong performance across different cases highlights the model's reliability and versatility.

The limitations of this research are as follows: The classification performance of the proposed MinMaxPat-based SOFE model dropped to 86.90% accuracy for sleep stage 3 when evaluated using LORO CV. To further validate and enhance the model's performance, more diverse and larger datasets could be utilized for testing.

The future works of this research are as follows: The proposed MinMaxPat-based SOFE model can be extended to larger and more diverse datasets, including ECG data from different states (e.g., awake or exercising), to enhance its generalizability. To address the performance degradation observed in Case 2, more advanced feature selection and classification techniques, such as deep learning, can be integrated into the model, and the performance of these deep models can also be evaluated as benchmarks. Despite its simplicity, the MinMaxPat-based SOFE model is an effective FE approach with manageable time complexity, making it highly suitable for real-time fibromyalgia detection systems and a viable option for practical applications. Additionally, future models could incorporate other physiological signals, such as electromyography (EMG) or functional near-infrared spectroscopy (fNIRS), alongside ECG data to improve detection accuracy and address more

complex cases. Furthermore, research efforts can focus on developing next-generation MinMaxPat-based deep learning models to advance biomedical signal classification further.

5. Conclusions

The presented MinMaxPat-based SOFE model has demonstrated its effectiveness in detecting fibromyalgia by analyzing ECG signals. The model utilizes MinMaxPat for feature extraction, generating 256 features from each signal, with the most informative features being selected using the CWINCA method. The tkNN classifier, integrated within the model, enabled high classification performance across multiple cases, reaching 100% classification accuracy for Case 2 with 10-fold cross-validation and 99.94% accuracy for the merged sleep stages.

When validated with LORO cross-validation, the model achieved 94.68% accuracy for Case 1 and 96.64% accuracy for the merged sleep stages. The geometric mean values were similarly high, ensuring reliable results across different validation methods. Compared to other classifiers, such as SkNN, the tkNN-based classifier outperformed its counterparts, achieving superior results in the fibromyalgia detection task.

This work also shows that the MinMaxPat feature extraction method is more effective than traditional approaches like LBP, as evidenced by its consistent classification accuracy exceeding 99%. Additionally, the model outperformed Barua et al.'s approach in Case 1, indicating the robustness of MinMaxPat as a feature extraction method in biomedical signal analysis [42].

These findings and the computed results showcase that the MinMaxPat-based SOFE model is a reliable tool for fibromyalgia detection, offering a balance between simplicity and accuracy in the classification of ECG signals.

Author Contributions: Conceptualization, V.Y.C., A.H.B., E.A., T.T. and S.D.; methodology, V.Y.C., A.H.B., E.A., T.T. and S.D.; software, T.T. and S.D.; validation, V.Y.C., A.H.B., E.A., T.T. and S.D.; formal analysis, V.Y.C., A.H.B., E.A., T.T. and S.D.; investigation, V.Y.C., A.H.B., E.A., T.T. and S.D.; resources, V.Y.C., A.H.B., E.A., T.T. and S.D.; data curation, V.Y.C., A.H.B., E.A., T.T. and S.D.; writing—original draft preparation, V.Y.C., A.H.B., E.A., T.T. and S.D.; writing—review and editing, V.Y.C., A.H.B., E.A., T.T. and S.D.; visualization, V.Y.C., A.H.B., E.A., T.T. and S.D.; supervision, S.D.; project administration, S.D.; All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Scientific Research Projects Coordination Unit of Firat University. Project number TEKF.24.49.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The authors are committed to making the data available if requested by the Journal.

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

References

1. Marino, Y.; Inferrera, F.; D'Amico, R.; Impellizzeri, D.; Cordaro, M.; Siracusa, R.; Gugliandolo, E.; Fusco, R.; Cuzzocrea, S.; Di Paola, R. Role of mitochondrial dysfunction and biogenesis in fibromyalgia syndrome: Molecular mechanism in central nervous system. *Biochim. Biophys. Acta (BBA)-Mol. Basis Dis.* **2024**, *1870*, 167301. [[CrossRef](#)]
2. Favretti, M.; Iannuccelli, C.; Di Franco, M. Pain biomarkers in fibromyalgia syndrome: Current understanding and future directions. *Int. J. Mol. Sci.* **2023**, *24*, 10443. [[CrossRef](#)] [[PubMed](#)]
3. Younes, A. Fibromyalgia: An Inflammatory Disease Characterized by Widespread Pain, Sleep Disturbances, Fatigue, and Memory Problems. *Eur. J. Neurodegener. Dis.* **2023**, *12*, 77–81.
4. Jurado-Priego, L.N.; Cueto-Ureña, C.; Ramírez-Expósito, M.J.; Martínez-Martos, J.M. Fibromyalgia: A Review of the Pathophysiological Mechanisms and Multidisciplinary Treatment Strategies. *Biomedicines* **2024**, *12*, 1543. [[CrossRef](#)]
5. Ruschak, I.; Montesó-Curto, P.; Rosselló, L.; Aguilar Martín, C.; Sánchez-Montesó, L.; Toussaint, L. Fibromyalgia syndrome pain in men and women: A scoping review. *Healthcare* **2023**, *11*, 223. [[CrossRef](#)] [[PubMed](#)]

6. Denche-Zamorano, Á.; Pastor-Cisneros, R.; Tomas-Carus, P.; Adsuar-Sala, J.C.; Salas-Gómez, D.; Parraca, J.A. Relationship of Pain, Depression, Fatigue, and Sleep Problems with Functional Capacity, Balance, and Fear of Falling in Women with Fibromyalgia: Cross-Sectional Study. *Nurs. Rep.* **2024**, *14*, 2819–2836. [[CrossRef](#)] [[PubMed](#)]
7. Varrassi, G.; Rekatsina, M.; Perrot, S.; Bouajina, E.; Paladini, A.; Coaccioli, S.; Tamayo, M.A.N.; Puttini, P.S. Is Fibromyalgia a Fashionable Diagnosis or a Medical Mystery? *Cureus* **2023**, *15*, e44852. [[CrossRef](#)]
8. Galvez-Sánchez, C.M.; Reyes del Paso, G.A. Diagnostic criteria for fibromyalgia: Critical review and future perspectives. *J. Clin. Med.* **2020**, *9*, 1219. [[CrossRef](#)] [[PubMed](#)]
9. Galvez-Sánchez, C.M.; De la Coba, P.; Duschek, S.; Reyes del Paso, G.A. Reliability, factor structure and predictive validity of the Widespread Pain Index and Symptom Severity scales of the 2010 American College of Rheumatology criteria of fibromyalgia. *J. Clin. Med.* **2020**, *9*, 2460. [[CrossRef](#)]
10. Pontes-Silva, A.; de Sousa, A.P.; Dibai-Filho, A.V.; de Souza, M.C.; DeSantana, J.M.; Avila, M.A. Do the instruments used to assess fibromyalgia symptoms according to American College of Rheumatology criteria generate similar scores in other chronic musculoskeletal pain? *BMC Musculoskelet. Disord.* **2023**, *24*, 467. [[CrossRef](#)]
11. Siracusa, R.; Paola, R.D.; Cuzzocrea, S.; Impellizzeri, D. Fibromyalgia: Pathogenesis, mechanisms, diagnosis and treatment options update. *Int. J. Mol. Sci.* **2021**, *22*, 3891. [[CrossRef](#)] [[PubMed](#)]
12. Lladser, A.-N.; Monteso-Curto, P.; Lopez, C.; Rossello, L.; Sydney, L.; Toussaint, L.; Casado-Martin, L.C. Multidisciplinary rehabilitation treatments for patients with fibromyalgia: A systematic review. *Eur. J. Phys. Rehabil. MedicinE* **2022**, *5*, 76. [[CrossRef](#)]
13. Giorgi, V.; Sarzi-Puttini, P.; Pellegrino, G.; Sirotti, S.; Atzeni, F.; Alciati, A.; Torta, R.; Varrassi, G.; Fornasari, D.; Coaccioli, S. Pharmacological Treatment of Fibromyalgia Syndrome: A Practice-Based Review. *Curr. Pain Headache Rep.* **2024**, 1–15. [[CrossRef](#)] [[PubMed](#)]
14. Jahanshahi Hesari, N.; Asgari, P.; Naderi, F.; Heidari, A. Effectiveness of positive mindfulness therapy and emotion-focused cognitive-behavioral therapy in improving pain perception and sleep quality in fibromyalgia patients. *Int. J. Musculoskelet. Pain Prev.* **2021**, *6*, 564–572. [[CrossRef](#)]
15. Jenssen, M.D.K.; Bakkevoll, P.A.; Ngo, P.D.; Budrionis, A.; Fagerlund, A.J.; Tayefi, M.; Bellika, J.G.; Godtlielsen, F. Machine learning in chronic pain research: A scoping review. *Appl. Sci.* **2021**, *11*, 3205. [[CrossRef](#)]
16. Sabeti, E.; Gryak, J.; Derksen, H.; Biwer, C.; Ansari, S.; Isenstein, H.; Kratz, A.; Najarian, K. Learning using concave and convex kernels: Applications in predicting quality of sleep and level of fatigue in fibromyalgia. *Entropy* **2019**, *21*, 442. [[CrossRef](#)] [[PubMed](#)]
17. Paul, J.K.; Iype, T.; Dileep, R.; Hagiwara, Y.; Koh, J.W.; Acharya, U.R. Characterization of fibromyalgia using sleep EEG signals with nonlinear dynamical features. *Comput. Biol. Med.* **2019**, *111*, 103331. [[CrossRef](#)]
18. Santana, A.N.; Cifre, I.; De Santana, C.N.; Montoya, P. Using deep learning and resting-state fMRI to classify chronic pain conditions. *Front. Neurosci.* **2019**, *13*, 1313. [[CrossRef](#)]
19. Chatterjee, I.; Baumgartner, L.; Cho, M. Detection of brain regions responsible for chronic pain in osteoarthritis: An fMRI-based neuroimaging study using deep learning. *Front. Neurol.* **2023**, *14*, 1195923. [[CrossRef](#)] [[PubMed](#)]
20. Fukae, J.; Isobe, M.; Hattori, T.; Fujieda, Y.; Kono, M.; Abe, N.; Kitano, A.; Narita, A.; Henmi, M.; Sakamoto, F. Convolutional neural network for classification of two-dimensional array images generated from clinical information may support diagnosis of rheumatoid arthritis. *Sci. Rep.* **2020**, *10*, 5648. [[CrossRef](#)]
21. Gökçay, D.; Eken, A.; Baltacı, S. Binary classification using neural and clinical features: An application in fibromyalgia with likelihood-based decision level fusion. *IEEE J. Biomed. Health Inform.* **2018**, *23*, 1490–1498. [[CrossRef](#)] [[PubMed](#)]
22. Alves, M.V.; Maciel, L.I.; Ramalho, R.R.; Lima, L.A.; Vaz, B.G.; Morais, C.L.; Passos, J.O.; Pegado, R.; Lima, K.M. Multivariate classification techniques and mass spectrometry as a tool in the screening of patients with fibromyalgia. *Sci. Rep.* **2021**, *11*, 22625. [[CrossRef](#)] [[PubMed](#)]
23. Robinson, M.E.; O’Shea, A.M.; Craggs, J.G.; Price, D.D.; Letzen, J.E.; Staud, R. Comparison of machine classification algorithms for fibromyalgia: Neuroimages versus self-report. *J. Pain* **2015**, *16*, 472–477. [[CrossRef](#)] [[PubMed](#)]
24. Martín-Brufau, R.; Gómez, M.N.; Sanchez-Sanchez-Rojas, L.; Nombela, C. Fibromyalgia detection based on EEG connectivity patterns. *J. Clin. Med.* **2021**, *10*, 3277. [[CrossRef](#)]
25. Passos, J.O.S.; dos Santos Alves, M.V.; Morais, C.L.; Martin, F.L.; Cavalcante, A.F.; Lemos, T.M.A.M.; Moura, S.; Freitas, D.L.; Mariz, J.V.M.; Carvalho, J.L. Spectrochemical analysis in blood plasma combined with subsequent chemometrics for fibromyalgia detection. *Sci. Rep.* **2020**, *10*, 11769. [[CrossRef](#)]
26. Orrù, G.; Gemignani, A.; Ciacchini, R.; Bazzichi, L.; Conversano, C. Machine learning increases diagnosticity in psychometric evaluation of alexithymia in fibromyalgia. *Front. Med.* **2020**, *6*, 319. [[CrossRef](#)]
27. Alves, M.V.; Maciel, L.I.; Passos, J.O.; Morais, C.L.; Dos Santos, M.C.; Lima, L.A.; Vaz, B.G.; Pegado, R.; Lima, K.M. Spectrochemical approach combined with symptoms data to diagnose fibromyalgia through paper spray ionization mass spectrometry (PSI-MS) and multivariate classification. *Sci. Rep.* **2023**, *13*, 4658. [[CrossRef](#)]
28. Karabey Aksalli, I.; Baygin, N.; Hagiwara, Y.; Paul, J.K.; Iype, T.; Barua, P.D.; Koh, J.E.; Baygin, M.; Dogan, S.; Tuncer, T. Automated characterization and detection of fibromyalgia using slow wave sleep EEG signals with glucose pattern and D’hondt pooling technique. *Cogn. Neurodynamics* **2024**, *18*, 383–404. [[CrossRef](#)]

29. Long, W.; Lu, Z.; Cui, L. Deep learning-based feature engineering for stock price movement prediction. *Knowl.-Based Syst.* **2019**, *164*, 163–173. [[CrossRef](#)]
30. Gutmann, M.U.; Hyvärinen, A. Noise-contrastive estimation of unnormalized statistical models, with applications to natural image statistics. *J. Mach. Learn. Res.* **2012**, *13*, 307–361.
31. Garbin, C.; Zhu, X.; Marques, O. Dropout vs. batch normalization: An empirical study of their impact to deep learning. *Multimed. Tools Appl.* **2020**, *79*, 12777–12815. [[CrossRef](#)]
32. Sarker, I.H. Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Comput. Sci.* **2021**, *2*, 420. [[CrossRef](#)]
33. Smyl, S.; Dudek, G.; Peřka, P. ES-dRNN: A hybrid exponential smoothing and dilated recurrent neural network model for short-term load forecasting. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *35*, 11346–11358. [[CrossRef](#)] [[PubMed](#)]
34. Shrestha, A.; Mahmood, A. Review of deep learning algorithms and architectures. *IEEE Access* **2019**, *7*, 53040–53065. [[CrossRef](#)]
35. Xu, D.; Li, T.; Li, Y.; Su, X.; Tarkoma, S.; Jiang, T.; Crowcroft, J.; Hui, P. Edge intelligence: Empowering intelligence to the edge of network. *Proc. IEEE* **2021**, *109*, 1778–1837. [[CrossRef](#)]
36. Tuncer, T.; Dogan, S.; Tasci, I.; Baygin, M.; Barua, P.D.; Acharya, U.R. Lobish: Symbolic Language for Interpreting Electroencephalogram Signals in Language Detection Using Channel-Based Transformation and Pattern. *Diagnostics* **2024**, *14*, 1987. [[CrossRef](#)]
37. Ojala, T.; Pietikainen, M.; Maenpää, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987. [[CrossRef](#)]
38. Tuncer, T.; Dogan, S.; Özyurt, F.; Belhaouari, S.B.; Bensmail, H. Novel Multi Center and Threshold Ternary Pattern Based Method for Disease Detection Method Using Voice. *IEEE Access* **2020**, *8*, 84532–84540. [[CrossRef](#)]
39. Dogan, A.; Akay, M.; Barua, P.D.; Baygin, M.; Dogan, S.; Tuncer, T.; Dogru, A.H.; Acharya, U.R. PrimePatNet87: Prime pattern and tunable q-factor wavelet transform techniques for automated accurate EEG emotion recognition. *Comput. Biol. Med.* **2021**, *138*, 104867. [[CrossRef](#)] [[PubMed](#)]
40. Maillo, J.; Ramírez, S.; Triguero, I.; Herrera, F. kNN-IS: An Iterative Spark-based design of the k-Nearest Neighbors classifier for big data. *Knowl.-Based Syst.* **2017**, *117*, 3–15. [[CrossRef](#)]
41. Goldberger, J.; Hinton, G.E.; Roweis, S.; Salakhutdinov, R.R. Neighbourhood components analysis. *Adv. Neural Inf. Process. Syst.* **2004**, *17*, 513–520.
42. Barua, P.D.; Kobayashi, M.; Tanabe, M.; Baygin, M.; Paul, J.K.; Iype, T.; Dogan, S.; Tuncer, T.; Tan, R.-S.; Acharya, U.R. Innovative Fibromyalgia Detection Approach Based on Quantum-Inspired 3LBP Feature Extractor Using ECG Signal. *IEEE Access* **2023**, *11*, 101359–101372. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.