Jestr

Journal of Engineering Science and Technology Review 18 (2) (2025) 203- 214

Research Article

JOURNAL OF
Engineering Science and Technology Review

www.jestr.org

Hand Gesture Recognition with Derivative and Summative Feature using Machine Learning Model

Pooja Kataria, Tripti Sharma and Yogendra Narayan*

Electronics and Communication Engineering Chandigarh University, Mohali, India

Received 11 October 2024; Accepted 10 April 2025

Abstract

Hand gesture recognition (HGR) is a crucial aspect of human-computer interaction, enabling intuitive and efficient control in various applications, including assistive technology, robotics, and virtual reality. Surface electromyogram (sEMG) signals have been widely used for gesture recognition due to their ability to capture muscle activity, providing a reliable basis for identifying different hand movements. Despite significant advancements, existing hand gesture recognition systems still face challenges in achieving high accuracy and robustness, particularly in distinguishing complex hand movements. Traditional features extracted from sEMG signals may not fully capture the dynamic and cumulative nature of muscle activity, limiting their effectiveness in real-world applications. There is a need to explore novel feature extraction methods that enhance recognition performance. This study introduces three novel derivative and summative features—Derivative Simple Square Integral (DSSI), Threshold Slope Sign Change (TSSC), and Absolute Value of Summation of the Temporal Moments (AVTM)—to improve hand gesture recognition accuracy. The proposed features aim to better encode both the instantaneous and cumulative dynamics of hand movements, leading to more accurate gesture classification. To evaluate their effectiveness, the performance of Extreme Gradient Boost (XGBoost), Logistic Regression, and Support Vector Machine (SVM) classifiers is analyzed using both existing and proposed features. The experimental results demonstrate that incorporating the proposed features with existing feature sets significantly enhances classification accuracy. The best performance is achieved using the TSSC feature with the SVM classifier, reaching an accuracy of 90.76%, which marks an improvement of approximately 4.98% over traditional feature-based methods. These findings validate the potential of the proposed features in developing more advanced and efficient hand gesture-based interfaces.

Keywords: Hand Gesture Recognition, Time Domain, Time-Frequency Domain, Frequency Domain, DSSI, TSSC, AVTM

1. Introduction

Human hand gesture recognition (HGR) systems have gained popularity due to their multifunctional role in human-computer interaction. The primary reason for this growing interest is their advantages, such as flexibility, ease of use, convenience, and the ability to control machines through gestures [1, 2]. The development of systems that enable command execution via gesture recognition represents a significant technological advancement, offering an alternative to traditional input methods such as joysticks, buttons, or touchscreens. HGR systems are widely applied across various domains, including chatbot-based communication, natural language processing, prosthetics, internet navigation, virtual and augmented reality, human-robot interaction, assistive vehicle command and guidance, medical automation, and video game control. However, developing HGR systems capable of accurately and swiftly recognizing specific gestures remains a significant challenge. This difficulty arises because each user performs gestures slightly differently, leading to variations in the collected data. As a result, even when individuals execute the same gesture, the recorded signals may differ. Furthermore, there is a possibility that measurements from different hand movements may appear similar, potentially leading to misinterpretations of the user's intentions [3, 4]. In the literature, vision-based HGR systems commonly utilize different cameras (RGB, infrared, thermal, stereo, depth) along with various sensors and Leap Motion signal structures. These systems analyze body gestures using multiple sensors. However, challenges such as occlusion, hand-sensor distance, ambient noise, and lighting conditions can adversely affect their performance. Consequently, for many HGR applications, sensor-based approaches, including EMG (electromyography), ultrasonic, or IMU-based methods, are often preferred [5,6]. EMG signals are particularly effective for recognizing static hand gestures, focusing on finger and hand joint movements. Meanwhile, Inertial Measurement Unit (IMU) and Visual-Inertial (VI) signals have proven highly beneficial for identifying dynamic gestures, especially those involving arm movement. EMG signals effectively capture muscle activation, making them a preferred choice for HGR applications. The popularity of EMG-based solutions has led to the development of commercial devices like the G-force and Myo wristbands. EMG data can be recorded through invasive or non-invasive methods. While invasive techniques require direct skin penetration to reach muscles, surface electromyography (sEMG) provides a non-invasive alternative by measuring muscular action potential through the outermost layer of the skin [7,8].

Pattern recognition using machine learning (ML) techniques is a widely adopted approach for gesture

*E-mail address: Narayan.yogendra1986@gmail.com
ISSN: 1791-2377
doi:10.25103/jestr.182.25

recognition leveraging sEMG. A typical machine learning pipeline includes feature extraction, data acquisition, model development, and interpretation. sEMG signals are collected by placing electrodes around the targeted muscle region. Experts typically extract hand-crafted features that capture temporal and frequency aspects of the data. Commonly used features include Root Mean Square (RMS), Variance, Mean Absolute Value (MAV), Zero Crossings (ZC), Slope Sign Variations (SSV), waveform length, and histogram features. These extracted features are then fed into various classification models, such as Random Forest (RF), Support Vector Clustering (SVC), k-Nearest Neighbors (kNN), and Logistic Regression (LR), with optimized parameters for improved classification accuracy. Support Vector Clustering (SVC) is a supervised ML technique that utilizes kernel-based learning [9,10]. It combines regression and classification methods for pattern analysis. The primary objective of SVC is to separate a dataset into two classes by identifying a hyperplane, which serves as a binary classifier. The data points closest to the hyperplane, known as support vectors, play a crucial role in determining its shape, form, and position in an n-dimensional space, where n represents the number of features. The hinge loss function, similar to that of logistic regression, is used to optimize the margin between the hyperplane and data points. The kNN algorithm classifies gestures based on the labels of their nearest neighbors. Euclidean distance is typically used to identify the closest neighbors, with classification accuracy improving as the k value (the number of considered neighbors) increases [11,12]. The Euclidean distance formula computes the distances between data points, aiming to minimize them. In kNN, classification is determined by averaging the labels of the k closest data points. The algorithm's performance is influenced by factors such as the threshold value, similarity measurement, and distance to the nearest neighbors.

Logistic regression aims to Logistic regression aims to utilize regression tools to illustrate the relationship between informative factors and a defined response variable. Unlike standard linear regression, where the response variable Y is continuous, logistic regression deals with a discrete response variable. This distinction is evident in how logistic regression selects thresholds and probabilities, showcasing its predictive nature. Logistic regression employs a linear equation to predict values between 0 and 1, operating based on probability. It converts these probabilities into categorical values using the sigmoid activation function, also known as the logistic function. A random forest classifier utilizes an ensemble of decision trees to label sample data. Multiple decision trees are employed to enhance the algorithm's performance, with each tree designed to optimize prediction accuracy [13,14]. Random forest methods are preferred over other techniques due to several advantages, including their robust performance on large datasets, ability to prevent overfitting, capability to handle both numerical and categorical variables, and ease of use in multi-class environments, all while requiring minimal parameter tuning. The advent of deep learning (DL), a revolutionary approach in machine learning, has marked a new era in data analytics. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are two prominent deep learning models that have demonstrated significant success in speech recognition and image classification, respectively. Unlike conventional machine learning methods that rely on manually engineered features, deep learning models extract and learn features automatically. Although these methods are not entirely new, their rapid advancement can be attributed to the availability of big data and substantial improvements in computational power [15,16], enabling the efficient execution of these resource-intensive models. The structure of the article is as follows: Section 2 explores the related research work. Section 3 describes the materials and methods. Existing time, frequency and time-frequency domain features have been discussed in Section 4 and our proposed research methodology is explained in Section 5. Section 6 comprises the results and discussion; finally, the conclusion and future scope are given in Section 7.

The following are the main contributions of the proposed work:

1. Introduced derivative and summative features to encode both instantaneous and cumulative dynamics of hand gestures, enhancing gesture classification.
2. Proposed three new features — Derivative Simple Square Integral (DSSI), Threshold Slope Sign Change (TSSC), and Absolute Value of Summation of Temporal Moments (AVTM)— to augment the existing features.
3. Applied machine learning algorithms, including Extreme Gradient Boost (XGBoost), Logistic Regression, and Support Vector Machine (SVM), to evaluate the effectiveness of the proposed features.
4. Facilitated the development of more advanced and efficient hand gesture-based interfaces through the incorporation of novel features, improving recognition performance.

2. Literature Review

Recently, hand gesture recognition has become a focal point in Human-Computer Interaction (HCI) due to its applications in sign language interpretation, virtual reality, robotics, and more. Interpreting hand movements as meaningful commands or actions is crucial for developing intuitive and effective interfaces. Traditionally, gesture recognition relied on image processing techniques, but with the advent of machine learning, a shift toward more advanced and efficient models has taken place. In recent years, various machine learning models, such as Convolutional Neural Networks (CNNs), Support Vector Machines (SVMs), and deep learning architectures, have been employed for gesture recognition. These models first process visual information from cameras or other sensors and then classify hand gestures with varying degrees of accuracy. However, the selection of features used in training these models is a critical factor influencing their performance. The integration of derivative and summative features in gesture recognition has recently gained attention as a means to enhance model accuracy. Derivative features capture variations in hand movements, such as speed and acceleration, providing dynamic information that complements the static spatial features of hand gestures. In contrast, summative features aggregate information over time, effectively representing complete movement patterns. This literature survey analyzes existing hand gesture recognition methods and explores the role of derivative and summative features in machine learning models. It examines feature extraction techniques, classification models, and their performance across different approaches, emphasizing the positive impact of these feature sets on the accuracy and reliability of gesture recognition systems.

K. S. Prakash centered on utilizing machine learning models to optimize the electrode configuration (number and

placement) for a hand gesture detection system based on wrist-worn electromyography (EMG) [17]. The features were ranked according to their significance scores using the ReliefF and Minimum Redundancy Maximum Relevance (MRMR) feature selection algorithms. By examining the sum of feature ranks and comparing the effectiveness of several machine learning models, including Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Artificial Neural Network (ANN), and Random Forest (RF), across different electrode combinations, the study was able to determine the best electrode configurations. With the SVM classifier in particular, the four-electrode arrangement proved to be the most efficient, providing a balance between high accuracy and low latency. Using SVM, this setup obtained 85.9% accuracy, 85.93% precision, 86.27% recall, and 86.09% F1-score. RF came in second with 85.2% accuracy.

R. Sebbah, sought to construct a sign recognition system combining two descriptors, namely Improved Weber Binary Coding (IWBC) and Low Frequency Descriptor (LFD), in order to distinguish both static and dynamic hand movements [18]. The SVM one-against-all method with RBF Kernel and K-nearest neighbor (KNN) is the foundation of the classification system. Four datasets were utilized to test the recognition system: Jochen Triesch's dataset, the Arabic Sign Language 2018 dataset, the Arabic Sign Language 2001 dataset (Halawani), and Sebastien Marcel's Dynamic dataset. With 97.07% on the Arabic Sign Language 2018 dataset and 95% on Jochen Triesch's dataset, the results showed that IWBC utilizing SVM was able to obtain higher recognition rates. Using SVM, a greater identification rate of 100% was achieved for LFD on Sebastien Marcel's Dynamic Dataset.

H. Feng, et al. suggested a machine learning-enhanced smart sensor system that uses machine learning models like support vector machines (SVM) and long short-term memory networks (LSTM) to process and analyze gesture data. The system combines hardware devices like Flex sensors [19], JY901S, cameras, and smart gloves. Experimental results demonstrate that collaborative inference of multi-modal data considerably enhances the model's classification accuracy when compared to single or bi-modal data in both static and dynamic gesture classification tasks.

In another work sought to evaluate the effectiveness of EMG-based hand gesture identification by combining the most advanced machine and deep learning models with innovative feature extraction techniques, such as fused time-domain descriptors, temporal-spatial descriptors, and wavelet transform-based features [20]. Experiments on the Grabmyo dataset show that the 1D Dilated CNN outperformed the others with a 97% accuracy rate using fused time-domain descriptors such waveform length ratio, power spectral moments, sparsity, and irregularity factor. Similarly, utilizing temporal-spatial descriptors (which contain time domain features together with extra features like coefficient of variation (COV) and Teager-Kaiser energy operator (TKEO)), random forest fared the best on the FORS-EMG dataset, with an accuracy of 94.95%.

S. A. Mousavi, et al. introduced a smart ring that can be worn and has a built-in Bluetooth low-energy (BLE) module [21]. To ensure the viability of the suggested approach, nine preset finger movements were added. To improve the accuracy of gesture recognition, well-distinguished feature vectors were chosen using data preprocessing techniques such as normalization, statistical feature extraction, random forest recursive feature elimination (RF-RFE), and k-nearest neighbors sequential forward floating selection (KNN-SFFS). For gesture categorization, three supervised machine-learning algorithms—support vector machines (SVMs), k-nearest neighbors (KNNs), and naïve Bayes (NB)—were employed. Six one-finger gestures and three two-finger gestures can be correctly identified by the system with 97.1% and 97.0% accuracy, respectively, when the KNN is used as the primary classifier.

K. Challa, et al. presented a new classifier based on electromyography (EMG) for hand gesture identification [22]. To record signals pertaining to hand movements, this method uses separate EMG sensors positioned on different hand portions. Eight healthy participants participated in their studies, performing three different hand gestures, including intricate motions like flexing, lifting, and grasping an object. Eight time-domain features were derived from the four channels of EMG signals that were recorded for this study. These characteristics were then utilized to build classifiers using both logistic regression (LR) and random forest (RF) machine learning techniques for the three hand motions under investigation. The findings show that the mean accuracies of the LR and RF classifiers were 0.94 and 0.966, respectively.

Numerous sensor technologies utilized in the gathering of gesture data, which have been studied to offer information on both static and dynamic motions [23]. Analysis is done on the benefits and drawbacks of both image-based and non-image-based technologies. A neural network-based machine learning technique has been created for high-precision gesture recognition. Testing the new approach on a database accessible to scientific investigation yielded positive results. Therefore, according to the accuracy, precision, recall, and F1-score metrics, the approach was able to identify the letters of the dactyl alphabet with an accuracy of 0.95, 0.92, 0.95, and 0.94, respectively.

In other article the authors decided to use four distinct machine learning models on two distinct datasets—American Sign Language (ASL) and a general gesture set—to accurately recognize hand gestures in order to categorize both static and dynamic gestures for the English language [24]. On both datasets, every classifier we employed demonstrated good accuracy, and after normalization, our results were much better. By adjusting the input, hidden, and output layers, the first Artificial Neural Network (ANN) model was put out, yielding a 99.40% accuracy rate. The accuracy of the second K-Nearest Neighbors (KNN) model was 99.14%. The accuracy of the third Decision Tree (DT) model was 94.52%. Ultimately, they used the Ensemble vote classifier to take the ensemble of these models, which showed overall predictive performance and turned out to be a lot more generic model with an accuracy of 99.45%.

T. -H. Le, et al. presented a fresh dataset of human hand gestures that might be useful for operating household appliances. An inexpensive Internet of Things (IoT) device with gyroscope and accelerometer sensors built in is used to collect the dataset [25]. Next, in order to train several machine learning models, we examine different features that were retrieved from multiple sensor data. In order to increase the identification rate, we also provide a straightforward yet efficient late fusion model from multimodal data. They showed that the suggested late fusion schema significantly increases the accuracy of gesture classification in the initial tests conducted on the gathered dataset. The late fusion procedure has the maximum accuracy of 87.61%.

A comparative study of current hand gesture recognition (HGR) methods is shown in Table 1, which also highlights the methods employed, performance indicators, conclusions, and drawbacks. This comparison offers insightful information about the benefit and drawbacks of different approaches, which forms the basis for comprehending developments and highlighting areas in which HGR systems require more investigation.

Table 1. Comparison of some existing HGR Approaches

Author & Year	Technique Used	Performance Metric	Findings	Limitations
K. S. Prakash, et al. [17]	EMG, SVM, KNN, ANN, RF, MRMR, ReliefF	Accuracy: 85.9%, Precision: 85.93%, Recall: 86.27%, F1-score: 86.09%	Optimal 4-electrode configuration achieved high accuracy and low latency with SVM classifier.	Limited to wrist-wearable EMG systems; further investigation needed for other electrode placements.
R. Sebbah, et al. [18]	Low Frequency Descriptor (LFD), Improved Weber Binary Coding (IWBC), SVM (one-vs-all), KNN	IWBC: 97.07% (ArSL 2018), 95% (Jochen Triesch's), LFD: 100% (Sebastien Marcel's)	IWBC achieved higher recognition rates with SVM, especially on Arabic and Jochen Triesch's datasets.	Limited to static and dynamic gesture datasets; could benefit from real-time recognition testing.
H. Feng, et al. [19]	Multi-modal data, SVM, LSTM, Flex sensor, JY901S, Camera, Smart gloves	Improved accuracy with multi-modal data over single/bi-modal data	Multi-modal data fusion significantly improved classification accuracy for static and dynamic gestures.	Need for better hardware integration and scalability for real-world use.
P. N. Aarotale, et al. [20]	1D Dilated CNN, RF, Fused time-domain descriptors, Temporal-spatial descriptors	97% (1D Dilated CNN on Grabmyo dataset), 94.95% (RF on FORS-EMG dataset)	Best performance observed with 1D Dilated CNN using fused time-domain features on Grabmyo dataset.	Limited to specific datasets; other gesture datasets could be tested for generalization.
S. A. Mousavi, et al. [21]	Smart ring, Bluetooth, Accelerometer, Gyroscope, RF-RFE, KNN-SFFS, SVM, KNN, NB	Accuracy: 97.1% (1-finger gestures), 97% (2-finger gestures)	Proposed method achieved high accuracy with KNN using KNN-SFFS features, enhancing feature selection efficiency.	Focuses on finger gestures; more complex gestures could be explored.
K. Challa, et al. [22]	EMG, Random Forest, Logistic Regression, Time-domain features	Accuracy: 96.6% (RF), 94% (LR)	Both RF and LR achieved high accuracy in hand gesture recognition using EMG signals from different hand parts.	Limited to simple hand gestures; other gestures and sensor types could improve generalization.
F. J. Abdullayeva, et al. [23]	Neural networks, Static and dynamic gesture recognition	Accuracy: 0.95, Precision: 0.92, Recall: 0.95, F1-score: 0.94	Achieved high precision in identifying gestures from the dactyl alphabet.	Relatively small scope (dactyl alphabet); could be extended to other gesture types.
M. Mahzabin, et al. [24]	ANN, KNN, Decision Tree, Ensemble voting classifier, ASL, General gesture set	Accuracy: 99.4% (ANN), 99.14% (KNN), 99.45% (Ensemble vote classifier)	Ensemble method outperformed individual classifiers, with high accuracy on ASL and general gesture datasets.	Limited to ASL and English gestures; might not generalize to other languages.
T.-H. Le, et al. [25]	IoT, Accelerometer, Gyroscope, Late fusion model	Accuracy: 87.61%	Late fusion of multimodal data enhanced gesture recognition accuracy for home appliance control.	Low-cost IoT device may limit real-world applications due to sensor precision and range.

This research work is based on the hand gesture recognition based on EMG signals. The proposed schemes are having three phases which are data pre-processing, feature extraction and classification. The feature extraction

based are broadly classified into time, frequency and time-frequency based schemes for the hand gesture recognition. Three new features are introduced which are derivative simple square integral, threshold slope sign change, absolute value of Summation of the Temporal Moments. These three new introduced features are the main contribution and motivation to increase accuracy for the hand gesture recognition. The proposed features are tested on different classifiers for hand gesture recognition. Machine learning models such as XGBoost, logistic regression, and SVM were then applied to both the existing features and the newly proposed features. The proposed features resulted in an improvement of up to 4.98% in classification accuracy over the existing features.

3. Materials and methods

The sEMG signal is a one-dimensional time series influenced by various factors, including the motor unit (MU) model, the number of MUs, and metabolic status. As a result, sEMG signals provide valuable insights into neuromuscular system activity. Essentially, sEMG signals are bioelectrical signals captured from the surface of human skeletal muscles using surface electromyography electrodes. These signals reflect neuromuscular activity and contain extensive information related to limb motion. Different limb movements exhibit distinct muscle contraction patterns, leading to variations in myoelectric signal features. By analyzing these features, it is possible to distinguish between various limb movement patterns. This research involved six gestures: elbow flexion, elbow extension, wrist extension, wrist flexion, wrist supination, and wrist pronation, as illustrated in Fig. 1. The study included ten healthy male participants, aged 20 to 37 years, all of whom were right-handed. To collect sEMG signals from the forearm muscles, two MyoTrace 400 devices were used in the biomedical laboratory of the National Institute of Technical Teachers' Training & Research (NITTTR), Chandigarh, India. Typically, the dominant energy of sEMG signals is concentrated in the frequency range of 20–500 Hz, with amplitudes limited to 0–10 mV. Consequently, the sampling frequency was set to 1000 Hz, and a bandpass filter with a bandwidth of 20–500 Hz was applied [26]. Each subject performed six hand motions: wrist flexion, wrist extension, elbow flexion, elbow extension, wrist pronation, and wrist supination. The tests were repeated ten times, with each repetition lasting 5 seconds, resulting in a total of 50 seconds of EMG signals per subject for each motion. None of the subjects had neuromuscular disorders. They were well-trained in all the movements before performing the hand motions and received a demonstration of the research procedure beforehand.

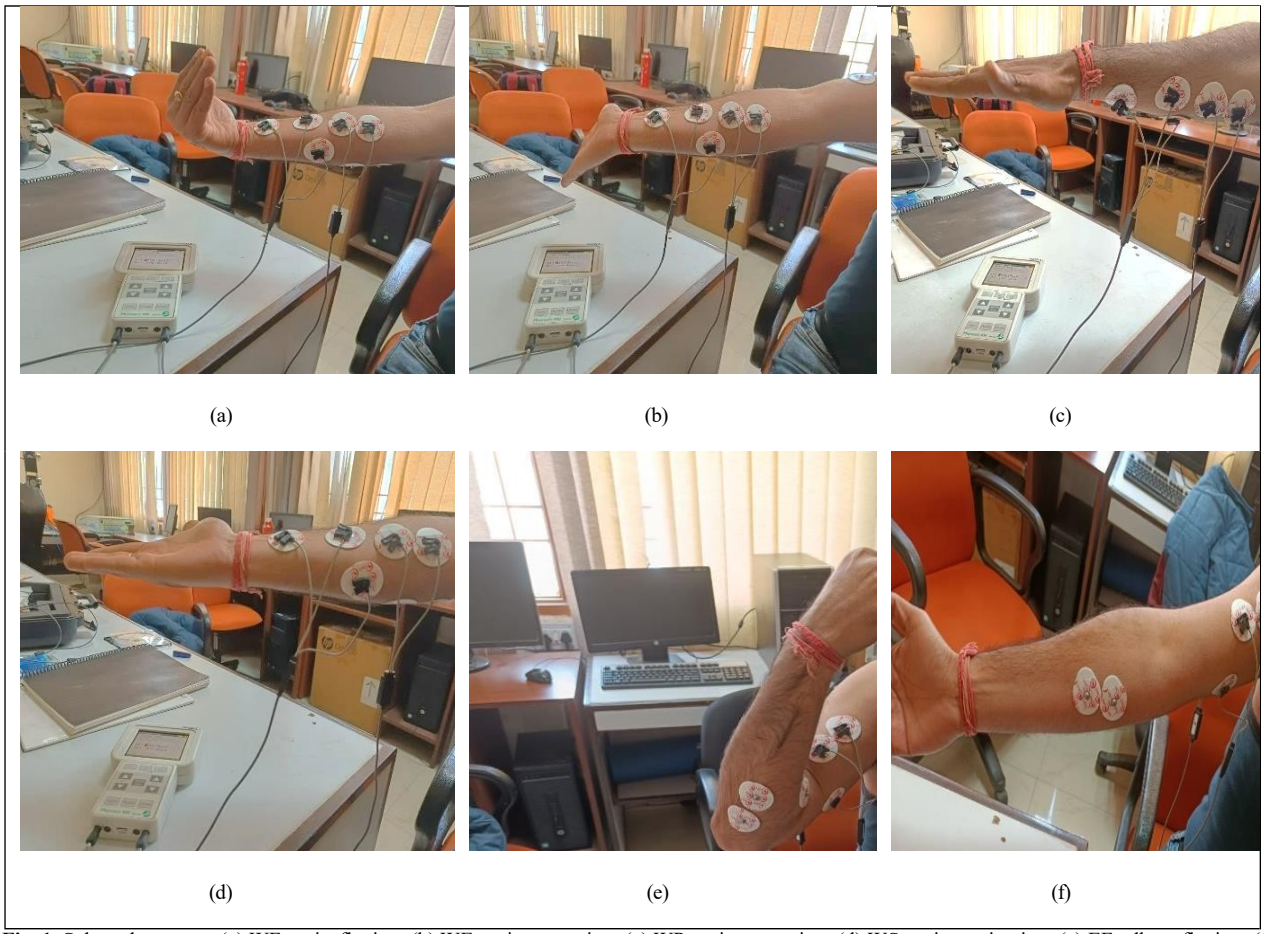

Fig. 1. Selected gestures. (a) WF: wrist flexion; (b) WE: wrist extension; (c) WP: wrist pronation; (d) WS: wrist supination; (e) EF: elbow flexion; (f) EE: elbow extension;

4. Feature Extraction

In hand gesture recognition, effective feature extraction plays a pivotal role in improving the accuracy and performance of machine learning models. This paper focuses on utilizing both derivative and summative features to enhance gesture recognition systems. Derivative features capture dynamic aspects of gestures, such as changes in

velocity, acceleration, and directional shifts, which are crucial for distinguishing gestures that involve movement. On the other hand, summative features aggregate the overall motion patterns across the gesture sequence, providing a holistic representation of the gesture. By combining these two feature types, the proposed method ensures that both the fine-grained details of hand motion and the general trajectory of the gesture are captured, leading to more robust recognition, particularly for complex or similar gestures. This approach leverages key spatial and temporal features, such as hand shape, orientation, and motion over time, enhancing the machine learning model's ability to classify both static and dynamic gestures with higher precision. EMG signal analysis features are typically categorized into three primary domains: time, frequency, and time- frequency or time-scale representation.

4.1 Time-based Features

Time domain features are computed directly from the unprocessed EMG time series, making them easy to use without requiring additional modifications. Their strong classification capability has led to widespread application in both scientific and clinical domains, particularly in low-noise environments. Additionally, they are preferred due to their lower computational complexity compared to frequency and time-scale domain features. The key time domain features are described below:

4.1.1. Integrated electromyography (IEMG)

In medical settings and for detecting onsets in EMG non-pattern recognition, IEMG is commonly utilized. It calculates the total absolute values of EMG signal amplitudes and is expressed as [27]:

$$IEMG = \sum_{i=1}^{N} |x_i| \tag{1}$$

In this context, x_i is used to denote the signal of EMG in a part i. In this, variable N represents the EMG signal's length.

4.1.2. Variance (VAR)

This formula, as presented in [27], is used for calculating the power density of the sEMG signal, which is one of its key applications.

$$\text{VAR} = \frac{1}{N-1} \sum_{N-1}^{N} x_i^2 \tag{2}$$

4.1.3. Zero crossing (ZC)

The frequency characteristics of the EMG signal are derived from this time-domain metric. It measures the rate at which the signal's amplitude values deviate from zero. A threshold condition is implemented to mitigate low-voltage variations and background noise. The computation is given as follows [27]:

$$ZC = \sum_{i=1}^{N-1} [sgn(x_i \times x_{i+1}) \cap |x_i - x_{i+1}| \geq threshold]; \tag{3}$$

$$sgn(x) = \begin{cases} 1, if\ x \geq threshold \\ 0, \quad otherwise \end{cases} \tag{4}$$

4.1.4. RMS (Root mean square)

For the study of EMG signals, this characteristic is often used. It is presented as a Gaussian random process with amplitude modulation, corresponding to contraction conditions with a constant force and no fatigue. The formula for the RMS feature is the following [27]:

$$RMS = \sqrt{\frac{1}{N}} \sum_{i=1}^{N} x_i^2 \tag{5}$$

4.1.5. MYOP (Myopulse percentage rate)

This characteristic pertains to the average myopulse output [27], which is assigned a value of 1 if the absolute value of the EMG signal exceeds a pre-set threshold. Using a specific numerical setup, the computation is shown as

$$Myop = \frac{1}{N} \sum_{i=1}^{N} [f\ (x_i)] \tag{6}$$

$$f(x) = \begin{cases} 1, if\ x \geq threshold \\ 0, \quad otherwise \end{cases} \tag{7}$$

4.1.6. MAV (Mean absolute value)

Just like the IEMG feature used as an onset index, MAV is a feature commonly used in EMG signal analysis, especially for detecting sEMG prosthetic limb control signals. The amplitude of signal within a section is represented as an average by this feature, and its mathematical expression is as follows:

$$MAV = \frac{1}{N} \sum_{i=1}^{N} |x_i| \tag{8}$$

4.1.7. WL (Waveform length)

WL computes the complexity of the signal, representing the increasing magnitude of the waveform over a given duration of time. Some works refer to this feature as wavelength (WAVE) [27]. Its mathematical expression is as follows:

$$\text{WL} = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{9}$$

4.1.8. ACC (Average amplitude change)

It corresponds to WL, with the distinction that the waveform length is averaged [27]. Its formula is given below:

$$AAC = \frac{1}{N} \sum_{i=1}^{N} |x_{i+1} - x_i| \tag{10}$$

4.1.9. LOG

An estimate of the force used to contract muscles is also provided by this feature. However, the characterization of the non-linear detector is modified to depend on logarithms, giving rise to the LOG feature [27], which can be formulated as follows

$$LOG = e^{\frac{1}{N} \sum_{i=1}^{N} log(|x_i|)} \tag{11}$$

4.2 Frequency-based Features

Studies on motor unit (MU) recruitment and muscle fatigue primarily rely on frequency or spectral domain characteristics. Power Spectral Density (PSD) analysis is particularly important in this domain. Derived from the autocorrelation function of the EMG signal's Fourier transform, PSD is analyzed using various statistical features. Some examples of frequency-based features include

4.2.1. MNF (Mean frequency)

The Mean Frequency (MNF) is obtained by summing the EMG power spectrum and its corresponding frequencies. This sum is then divided by the total power of the spectrum. Its computation formula is [27]:

$$MNF = \sum_{j=1}^{M} f_j P_j / \sum_{j=1}^{M} P_j) \tag{12}$$

In this formula, f_j and P_j respectively represent frequency spectrum and power spectrum at frequency bin j. Also, M in the indicator of the length of this bin.

4.2.2. Mean power (MNP):
MNP represents the mean power of the signal's power spectrum and can be calculated as follows: [27]:

$$MNP = \sum_{j=1}^{M} P_j / M \tag{13}$$

4.2.3. Total power (TP)
It represents the summation of the EMG power spectrum and can be defined as follow [27]:

$$TP = \sum_{j=1}^{M} P_j = SMO \tag{14}$$

4.2.4. Median frequency (MDF)
For the study of EMG signals, it is a frequently utilized characteristic. It is modeled as a Gaussian random process with amplitude modulation, relevant to contraction conditions involving a steady force and no fatigue. The mathematical expression for the RMS feature i [27]:

$$\sum_{j=1}^{MDF} P_j = \sum_{j=MDF}^{M} P_j = \frac{1}{2}\sum_{j=1}^{M} P_j \tag{15}$$

4.2.5. Peak frequency (PKF)
PKF represents the frequency at which maximum power occurs. It is calculated as follows: [27]:

$$PKF = \max(P_j),\ j = 1,2,\dots,M \tag{16}$$

4.2.6. FR (Frequency ratio)
FR is recommended for distinguishing between muscle contraction and relaxation. To achieve this, the percentage ratio between the EMG signal's high and low frequencies must be determined [27].

$$FR = \sum_{j=LLC}^{ULC} P_j / \sum_{j=LHC}^{UHC} P_j \tag{17}$$

In the low-frequency band, ULC represents the upper cutoff frequency, while LLC denotes the lower cutoff frequency. Similarly, in the high-frequency band, LHC and UHC represent the lower and upper cutoff frequencies, respectively

4.3 Time-Frequency-based Features
These features are implemented using Fast Discrete Wavelet Transform (DWT), incorporating both temporal and spectral information to capture various aspects such as trends, segmentations, discontinuities in higher derivatives, and self-similarity, which cannot be identified using time and frequency features alone.σ

4.3.1. DASDV (Difference absolute standard deviation value)
This feature is similar to the RMS feature and is evaluated as the standard deviation (SD) of the wavelength [27]:

$$DASDV = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_{i+1} - x_i)^2} \tag{18}$$

In this, x_i is utilized to denote EMG signal within section i and the signal's length is illustrated with N.

4.3.2. WAMP (Wilson amplitude)
The purpose of this feature is to specify the signal's frequency data. This measure is evaluated as the number of times the difference in signal amplitude between two consecutive segments exceeds a predefined threshold. WAMP is expressed as [27]:

$$WAMP = \sum_{i=1}^{N-1}[f|x_n - x_{n+1}|]; \tag{19}$$

$$f(x) = \begin{cases} 1, if\ x \geq threshold \\ 0, \quad otherwise \end{cases} \tag{20}$$

4.3.3. Wavelet entropy (WENT)
It is utilized to compute the degree of order/disorder in sEMG signals and to generate information related to the primary dynamical process of sEMG signals. This feature is defined a [27]:

$$E(c) = -\sum_i c_i^2 \log(c_i^2) \tag{21}$$

In this, c_i is used to illustrate the i^{th} coefficient of an sEMG signal within an orthonormal base.

5. Research Methodology

In this research work, three new features have been added for hand gesture recognition from EMG signals. Machine learning models such as XGBoost, logistic regression, and SVM were then applied to both the existing features and the newly proposed features. The proposed features resulted in an improvement of up to 4.98% in classification accuracy over the existing features. The complete flowchart of the proposed work is depicted in Figure 2. The first feature is the derivative of the simple square integral, which is the second-order derivative of the simple square integral. The second feature is the threshold slope sign change, and the last feature is the absolute value of the summation of the temporal moment, which is derived from the temporal moment. The descriptions of these three features are given below:

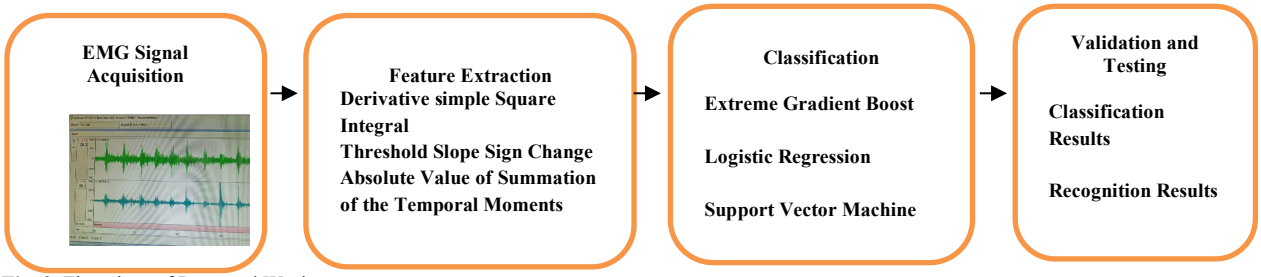

Fig. 2. Flowchart of Proposed Work

5.1 DSSI (Derivative Simple Square Integral)

An EMG signal's energy can be represented by the SSI as:

$$SSI = \sum_{i=1}^{N} x_i^2 \quad (22)$$

where N is the number of specimens in each section and xi is the ith specimen. DSSI is the second derivative of simple square integral and it is defined as:

y=$f(x)$

$\frac{dy}{dx} = f'(x)$

$f'(x)$=0 where x_1 and x_2

$\frac{d}{dx} f'(x) \cdot = f''(x)$

Where $x_1 = f''(x) \rightarrow f''(x_1) < 0$

x_1 is $local$ $minimum$

Where $x_2 = f''(x) \rightarrow f''(x_1) > 0$

x_2 is $local$ $miximum$

$$\text{DSSI} = \sum_{i=1}^{N} f''(x_i) \quad (23)$$

By measuring abrupt fluctuations in energy in any direction, the Absolute Value of DSSI is able to identify important acceleration patterns in EMG data. This improves feature discrimination for intricate hand movements, reduces noise effects, and increases classification accuracy by identifying unique transition points between gestures.

5.2 Threshold Slope Sign Change (TSSC)

The threshold slope sign change parameter is derived from the slope sign change parameter. Utilizing the threshold function which eliminates noise-induced slope changes, TSSC shows how frequently the EMG waveform's slope alters sign across an analysis window. This characteristic defines the frequency parameter of the EMG signal. The mathematical representation is given below:

$$\text{TSSC}=\sum_{i=2}^{N-1} f(x_1 - x_{i-1}) * (x_i - x_{i+1}) \quad (24)$$

$$f(x) = \begin{pmatrix} 1, if\ x > threshold \\ 0, otherwise \end{pmatrix} \quad (25)$$

Where x_i is the current scope change, x_{i+1} is future change and x_{i-1} is the previous change in the scope relative to frequency

5.3 Absolute value of Summation of the Temporal Moments (AVTM)

The moments and probability density function moments are comparable, with multiple applications for the square of the temporal history. Negative amplitude issues are circumvented. When employing the squared time period, there are useful properties that establish a relationship between time-domain moments and frequency-domain moments

$$\text{Temporal Moment (TM)} = \frac{1}{N} \sum_{i=1}^{N} x_i^{order(1\ to\ N)} \quad (26)$$

As represented in the equation, the temporal moment is described as the number of samples in the signal, denoted by N. The variable x represents the current sample of the signal, and the order defines the analysis window number

$$\text{Analysis Window}=\sum_{n=1}^{k} x_n^{\frac{1}{2}} \quad (27)$$

The analysis window is the current segment of the signal that needs to be analyzed, and its value always ranges between 1 and 5.

$$\text{AVTM}=\left| \frac{\frac{1}{N} \sum_{i=1}^{N} x_i^k}{k} \right| \quad (28)$$

AVTM represents the absolute value of the summation, and the window of AVTM, defined within brackets, calculates the absolute summation value Temporal signal patterns are captured by the AVTM feature, which uses absolute values to prevent negative amplitude issues. It examines signal dynamics across analysis windows, which makes it useful for identifying subtle variations between analogous gestures. Normalization by window size (k) assures uniform feature scaling despite window length.

6. Results & Discussion

This research focuses on hand gesture recognition from EMG signals using machine learning models. To predict hand movements, various types of information are extracted from EMG signals, categorized into time-domain, frequency-domain, and time-frequency-domain features. In this study, three new features are introduced: Derivative Simple Square Integral, Threshold Slope Sign Change, and Absolute Value of Summation of the Temporal Moments. Machine learning models such as XGBoost, Logistic Regression (LR), and Support Vector Machine (SVM) are employed for hand gesture recognition. The results are presented by comparing the performance of existing features and the newly proposed features across different classifiers. Cross-validation is performed using different fold sizes to assess the robustness of the features. The features used for hand gesture recognition include time-domain, frequency-domain, and time-frequency-domain features. The classifiers XGBoost, Logistic Regression, and SVM—are applied to various features such as iEMG, VAR, ZC, RMS, MYOP, MAV, WL, ACC, LOG, MNF, MNP, TP, MDF, PKF, FR, DASDV, WAMP, and WENT. As shown in Table 2, the PKF feature yields the highest accuracy across all classifiers, including XGBoost, Logistic Regression, and SVM.

Table 2. Accuracy of machine learning models with existing features

Feature	XGboost	Logistic Regression	SVM
iEMG	63.45	59.89	67.89
VAR	72.45	70.89	74.67
ZC	46.78	52.34	56.78
RMS	67.89	66.78	67
MYOP	69.75	70.67	68.12
MAV	66.13	43.54	60.19
WL	61.90	60.45	59.18
ACC	76.89	79.65	72.34
LOG	61.23	60.74	67.34
MNF	81.52	80.81	76.89
MNP	77.67	75.74	76.84
TP	65.67	63.45	63.76
MDF	78.90	75.89	81.12

PKF	85.78	82.34	86.78
FR	65.78	66.65	71.90
DASDV	56.78	56.12	59.45
WAMP	72.34	73.76	74.32
WENT	72.56	70.56	74.56

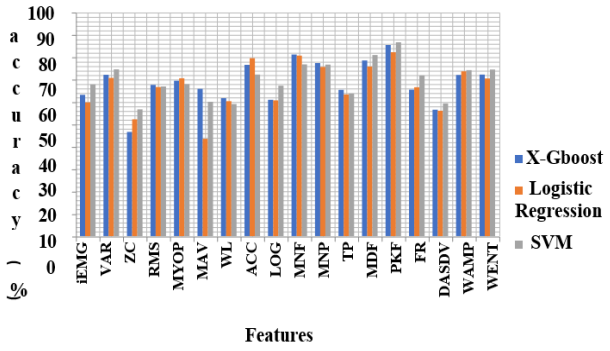

Fig. 3. Performance of Machine Learning Models with Existing Features

Figure 3 illustrates the application of three machine learning algorithms—XGBoost, Logistic Regression (LR), and Support Vector Machine (SVM)—for hand gesture recognition. The models utilize features extracted from time, frequency, and time-frequency domains. The extracted features include iEMG, VAR, ZC, RMS, MYOP, MAV, WL, ACC, LOG, MNF, MNP, TP, MDF, PKF, FR, DASDV, WAMP, and WENT. Among these, PKF emerges as the most effective feature, achieving the highest accuracy across all classifiers, including XGBoost, Logistic Regression, and SVM.

Table 3 presents the results of applying existing features with the SVM classifier under different fold cross-validations (FCV) settings, including 2, 4, 6, 8, and 10 folds. The analysis reveals that 8-fold and 10-fold cross-validation yield the most stable results for hand gesture recognition

Table 3. Accuracy of SVM classifier on existing features

Feature	2 FCV	4 FCV	6 FCV	8 FCV	10 FCV
iEMG	67.89	69.90	70.45	70.13	70.13
VAR	74.67	65.56	66.23	66.37	66.37
ZC	56.78	56.10	57.89	57.34	57.34
RMS	67	65.67	63.67	63.61	63.61
MYOP	68.12	66.89	65.46	65.43	65.43
MAV	60.19	61.21	60.89	60.83	60.83
WL	59.18	56.67	55.12	55.10	55.10
ACC	72.34	70.34	68.45	68.43	68.43
LOG	67.34	68.89	67.34	67.34	67.34
MNF	76.89	78.78	75.34	75.30	75.30
MNP	76.84	79.10	78.45	78.42	78.42
TP	63.76	64.56	64.34	64.30	64.30
MDF	81.12	80.34	80.12	80.10	80.10
PKF	86.78	87.34	87.78	87.72	87.72
FR	71.90	75.34	75.42	75.40	75.40
DASDV	59.45	58.45	58.32	58.30	58.30
WAMP	74.32	73.34	73.90	73.91	73.91
WENT	74.56	72.89	72.90	72.86	72.86

As shown in Table 4, Logistic Regression is applied to classify existing features, including iEMG, VAR, ZC, RMS, MYOP, MAV, WL, ACC, LOG, MNF, MNP, TP, MDF, PKF, FR, DASDV, WAMP, and WENT. The model is evaluated using cross-validation with different fold settings. The results indicate that the highest accuracy for hand gesture recognition is achieved at the 10th fold.

Table 4. Average Accuracy of Logistic Regression classifier on existing features

Feature	2 FCV	4 FCV	6 FCV	8 FCV	10 FCV
iEMG	60.18	61.11	60.09	60.09	61.11
VAR	70.21	70.31	69.32	69.32	70.31
ZC	52.34	51.78	50.17	50.17	51.78
RMS	66.78	65.23	66.12	66.12	65.23
MYOP	70.67	69.34	68.54	68.54	69.34
MAV	43.54	44.60	43.43	43.43	44.60
WL	60.45	61.20	62.10	62.10	61.20
ACC	79.65	80.10	79.05	79.05	80.10
LOG	60.74	61.56	59.51	59.51	61.56
MNF	80.81	79.45	78.42	78.42	79.45
MNP	75.74	77.20	78.10	78.10	77.20
TP	63.45	64.14	65.11	65.11	64.14
MDF	75.89	77.11	76.23	76.23	77.11
PKF	82.34	83.10	82.56	82.56	83.10
FR	66.65	65.45	64.36	64.36	65.45
DASDV	56.12	56.11	55.34	55.34	56.11
WAMP	73.76	74.89	75.12	75.12	74.89
WENT	70.56	73.52	75.78	75.78	73.52

As shown in Table 5, the XGBoost classifier is applied with five cross-validation settings: 2, 4, 6, 8, and 10 folds. These cross-validations are tested using the existing feature set, including iEMG, VAR, ZC, RMS, MYOP, MAV, WL, ACC, LOG, MNF, MNP, TP, MDF, PKF, FR, DASDV, WAMP, and WENT. The results indicate that the 8-fold and 10-fold methods provide the most stable performance, achieving an accuracy of approximately 62%.

Table 5. Average Accuracy of XGBoost classifier on existing features

Feature	2 FCV	4 FCV	6 FCV	8 FCV	10 FCV
iEMG	63.45	62.41	62.40	62.40	62.40
VAR	72.45	70.23	70.65	70.65	70.65
ZC	46.78	44.35	44.30	44.30	44.30
RMS	67.89	64.10	64.20	64.20	64.20
MYOP	69.75	68.17	68.12	68.12	68.12
MAV	66.13	65.12	65.64	65.64	65.64
WL	61.90	58.18	58.21	58.21	58.21
ACC	76.89	74.34	74.30	74.30	74.30
LOG	61.23	59.10	59.20	59.20	59.20
MNF	81.52	78.32	78.33	78.33	78.33
MNP	77.67	76.16	76.75	76.75	76.75
TP	65.67	64.21	64.31	64.31	64.31
MDF	78.90	75.14	75.19	75.19	75.19
PKF	85.78	84.10	84.96	84.96	84.96
FR	65.78	64.64	64.84	64.84	64.84
DASDV	56.78	55.23	55.28	55.28	55.28
WAMP	72.34	70.21	70.38	70.38	70.38
WENT	72.56	70.54	70.75	70.75	70.75

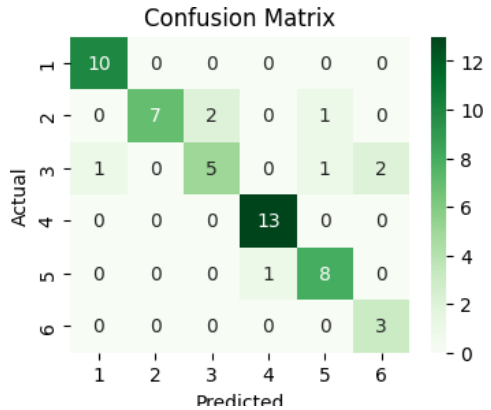

Fig. 4. Confusion of SVM with PKF Feature

As shown in Figure 4, a confusion matrix is generated for the PKF feature with the SVM classifier, corresponding to the number of classes. With an F1-score of 1.0, Wrist Flexion (class 1) attains ideal precision (10/10 = 100%) and recall (10/10 = 100%) while examining assessment metrics by class. Class 3 wrist pronation performs the worst, with a precision of 71.4% (5/7) and recall of 55.6% (5/9). As a whole, the model's accuracy is 85.78%. According to a normalized confusion matrix, 11% of samples with wrist pronation are incorrectly classified as elbow flexion, 11% as wrist flexion, and 22% as elbow extension. Wrist pronation has a true positive rate of just 0.56. Physiological proximity associations are shown through assessing erroneous patterns; gestures that share comparable muscle activation patterns (wrist-elbow movements) exhibit a greater degree of uncertainty. Particularly between Wrist Extension and Wrist Pronation, the matrix shows asymmetric misinterpretation (2 instances in one direction, 0 in the other). Clear identification for classes 1, 4, and 6 can be observed through analysis of diagonal supremacy indicating that these gesture patterns have more unique PKF traits than others, thereby indicating that the PKF feature, when used with the SVM classifier, outperforms all other feature-classifier combinations.

Furthermore, as illustrated in Table 6, different machine learning models, including XGBoost, Logistic Regression, and SVM, are evaluated on the proposed features. The analysis reveals that the SVM model achieves the highest accuracy compared to XGBoost and Logistic Regression for hand gesture recognition

Table 6. Accuracy of Machine learning model with proposed features

Features	XGboost	Logistic Regression	SVM
DSSI	89.78	88.78	90.34
TSSC	90.76	86.78	91.23
AVTM	75.78	74.78	79.56

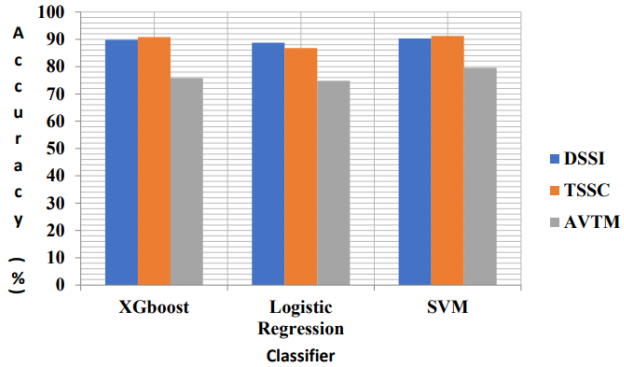

Fig. 5. Performance of Machine Learning Models with proposed Features

As shown in Figure 5, three new features—DSSI, TSSC, and AVTM—are introduced in this research. When XGBoost, Logistic Regression, and SVM algorithms are applied to these features, the highest accuracy of 91% is achieved using the TSSC feature.

Additionally, as shown in Table 7, the SVM classifier is implemented on the proposed features across different cross-validation folds, including 2, 4, 6, 8, and 10 FCV. The proposed features (DSSI, TSSC, AVTM) are evaluated using the SVM classifier, with results indicating that the 10-fold cross-validation provides the most stable accuracy for hand gesture recognition.

Table 7. Average Accuracy of SVM classifier on Proposed features

Features	2 FCV	4 FCV	6 FCV	8 FCV	10 FCV
DSSI	90.34	90.89	90.65	90.54	92.01
TSSC	91.23	91.34	91.30	92.10	91.67
AVTM	79.56	79.52	79.50	78.89	80.21

As shown in Table 8, the Logistic Regression classifier is applied to the proposed features—DSSI, TSSC, and AVTM—using different cross-validation folds (2, 4, 6, 8, and 10 FCV). The model is tested across these folds, and the results indicate that the 10-fold cross-validation provides the most stable accuracy for hand gesture recognition

Table 8. Average Accuracy of Logistic Regression classifier on Proposed features

Features	2 FCV	4 FCV	6 FCV	8 FCV	10 FCV
DSSI	88.78	88.12	87.10	88.20	89.34
TSSC	86.78	86.70	84.56	83.32	84.30
AVTM	74.78	74.54	74.50	74.54	73.50

As shown in Table 9, the XGBoost classifier is applied to the proposed features—DSSI, TSSC, and AVTM—using different cross-validation folds (2, 4, 6, 8, and 10 FCV). The model is tested across these folds, and the results indicate that the 10-fold cross-validation provides the most stable accuracy for hand gesture recognition.

Table 9. Average Accuracy of XGBoost classifier on Proposed features

Features	2 FCV	4 FCV	6 FCV	8 FCV	10 FCV
DSSI	89.78	89.10	88.20	87.10	88.13
TSSC	90.76	90.10	90.05	90.32	90.10
AVTM	75.78	73.53	73.40	74.54	76.31

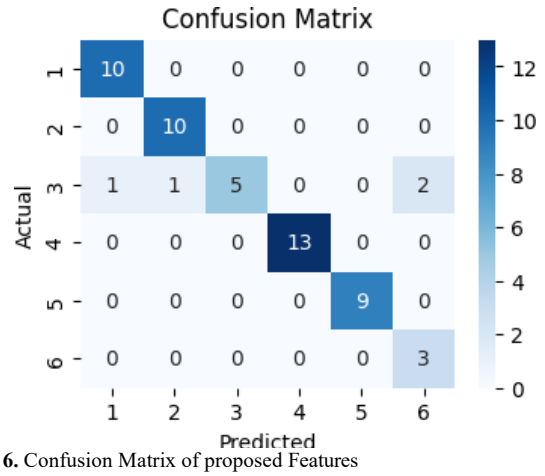

Fig. 6. Confusion Matrix of proposed Features

As shown in Figure 6, the confusion matrix for the SVM classifier using the TSSC feature is presented with respect to the number of classes. The x-axis represents the predicted values, while the y-axis represents the actual values. The matrix exhibits a remarkable 90.76% overall accuracy, with a particularly significant diagonal dominance that suggests viable class separation. For the majority of gestures, class-wise metrics of performance demonstrate superior results: Classes 1 (WF), 2 (WE), 4 (WS), and 5 (EF) attain precision and recall levels at or close to 100%.Class 1 (one instance), Class 2 (one instance), and Class 6 (EE) (two instances) all exhibit misclassifications, with Class 3 (wrist pronation) having the lowest recall (5/9 accurate) and the most

challenging gesture. This implies that categorization performance is impacted by anatomical proximity, with pattern overlap occurring due to identical muscle activation characteristics. While Class 6(EE) exhibits lower precision at 60% (3/5), the majority of classes have noticeably higher accuracy scores: Class 3 (100%), Class 4 (100%), and Class 5 (100%). The majority of the model's perplexity is caused by wrist pronation, as evidenced by the skewed pattern of the error distribution. The assertion that TSSC provides better hand gesture identification ability with the SVM classifier is evidenced by the TSSC feature's increased performance for Class 2 (WE) and Class 5 (EF) in contrast to the PKF feature confusion matrix in Figure 4. The analysis indicates that the TSSC feature, when used with the SVM classifier, delivers the best performance for hand gesture recognition.

Table 10. Precision Analysis

Model	2 FCV	4 FCV	6 FCV	8 FCV	10 FCV
XGBoost	72..78	72.67	74.89	76.23	75.90
Logistic Regression	81.23	80.89	80.12	82.89	82.45
SVM	85.78	86.12	86.90	87.67	88.90
Proposed	91.78	90.16	90.16	90.89	90

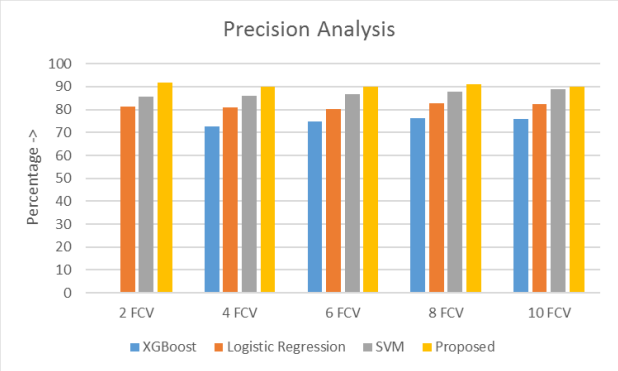

Fig. 7. Precision Analysis

As shown in table 10 and figure 7, the precision value of proposed model is compared with other machine learning models like XGBoost, Logistic regression and SVM on different cross validation values. It is analyzed that proposed model has maximum precision value for the hand gesture recognition

Table 11. Recall Analysis

Model	2 FCV	4 FCV	6 FCV	8 FCV	10 FCV
XGBoost	72..78	72.67	74.89	76.23	75.90
Logistic Regression	81.23	80.89	80.12	82.89	82.45
SVM	85.78	86.12	86.90	87.67	88.90
Proposed	91.78	90.16	90.16	90.89	90

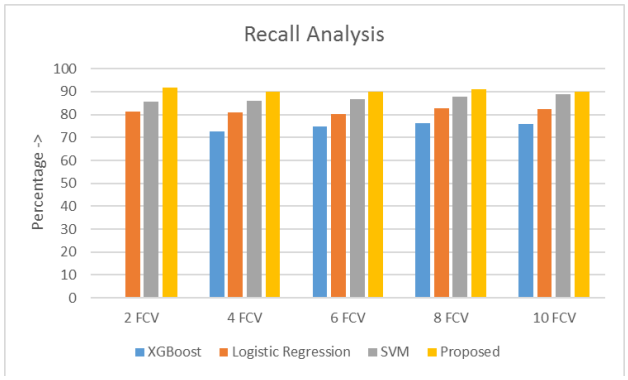

Fig. 8. Recall Analysis

As shown in table 11 and figure 8, the recall value of proposed model is compared with other machine learning models like XGBoost, Logistic regression and SVM on different cross validation values. It is analyzed that proposed model has maximum precision value for the hand gesture recognition

7. Conclusion and Future Scope

In this study, a sEMG dataset from ten healthy subjects was utilized to recognize six different hand motions. After acquiring the sEMG signals, preprocessing, feature extraction, and classification were performed. To improve classification accuracy, three new feature sets were introduced and critically compared with existing features. Machine learning classifiers, namely XGBoost, Logistic Regression, and SVM, were applied to assess the accuracy of both existing and proposed features. Among the existing features, PKF achieved the highest accuracy of 85.78% for hand gesture recognition. However, the newly proposed TSSC feature demonstrated superior performance, achieving a maximum accuracy of 90.76% with the SVM classifier—an improvement of approximately 4.98% over existing features.

For future work, integrating advanced sensor technologies, such as infrared depth cameras or wearable motion sensors, could enhance the robustness of gesture recognition systems. Additionally, accuracy can be further optimized using genetic algorithms (GA), particle swarm optimization (PSO), or a parallel-series combination of different optimization techniques.

References

[1] E. Ceolini *et al.*, "Hand-Gesture Recognition Based on EMG and Event-Based Camera Sensor Fusion: A Benchmark in Neuromorphic Computing," *Front. Neurosci.*, vol. 14, Aug. 2020, Art. no. 637, doi: 10.3389/fnins.2020.00637.

[2] M. Yasen and S. Jusoh, "A systematic review on hand gesture recognition techniques, challenges and applications," *PeerJ Computer Science*, vol. 5, Sep. 2019, Art. no. e218, doi: 10.7717/peerj-cs.218.

[3] M. J. Cheok, Z. Omar, and M. H. Jaward, "A review of hand gesture and sign language recognition techniques," *Int. J. Mach. Learn. & Cyber.*, vol. 10, no. 1, pp. 131–153, Jan. 2019, doi: 10.1007/s13042-017-0705-5.

[4] O. Zahra and D. Navarro-Alarcon, "A Self-organizing Network with Varying Density Structure for Characterizing Sensorimotor Transformations in Robotic Systems," in *Towards Auton. Robotic Sys.*, vol. 11650, K. Althoefer, J. Konstantinova, and K. Zhang, Eds., in Lecture Notes in Computer Science, vol. 11650, Cham: Springer International Publishing, 2019, pp. 167–178. doi: 10.1007/978-3-030-25332-5_15.

[5] Y. Zhang, Z.-R. Wang, and J. Du, "Deep Fusion: An Attention

Guided Factorized Bilinear Pooling for Audio-video Emotion Recognition," in *2019 Int. Joint Conf. Neural Netw. (IJCNN)*, Budapest, Hungary: IEEE, Jul. 2019, pp. 1–8. doi: 10.1109/IJCNN.2019.8851942.

[6] L. Chen, J. Fu, Y. Wu, H. Li, and B. Zheng, "Hand Gesture Recognition Using Compact CNN via Surface Electromyography Signals," *Sensors*, vol. 20, no. 3, Jan. 2020, Art. no. 672, doi: 10.3390/s20030672.

[7] K. Koizumi, K. Ueda, and M. Nakao, "Development of a Cognitive Brain-Machine Interface Based on a Visual Imagery Method," in *2018 40th Annual Int. Conf. IEEE Eng. Medic. Biology Soc. (EMBC)*, Honolulu, HI: IEEE, Jul. 2018, pp. 1062–1065. doi: 10.1109/EMBC.2018.8512520.

[8] A. Toro-Ossaba, J. Jaramillo-Tigreros, J. C. Tejada, A. Peña, A. López-González, and R. A. Castanho, "LSTM Recurrent Neural Network for Hand Gesture Recognition Using EMG Signals," *Appl. Sci.*, vol. 12, no. 19, Sep. 2022, Art. no. 9700, doi: 10.3390/app12199700.

[9] Y.-U. Jo and D.-C. Oh, "Real-Time Hand Gesture Classification using CRNN with Scale Average Wavelet Transform," *J. Mech. Med. Biol.*, vol. 20, no. 10, Dec. 2020, Art. no. 2040028, doi: 10.1142/S021951942040028X.

[10] N. Salankar, D. Koundal, C. Chakraborty, and L. Garg, "Automated attention deficit classification system from multimodal physiological signals," *Multimed. Tools Appl.*, vol. 82, no. 4, pp. 4897–4912, Feb. 2023, doi: 10.1007/s11042-022-12170-1.

[11] A. Jaramillo-Yánez, M. E. Benalcázar, and E. Mena-Maldonado, "Real-Time Hand Gesture Recognition Using Surface Electromyography and Machine Learning: A Systematic Literature Review," *Sensors*, vol. 20, no. 9, Apr. 2020, Art. no. 2467, doi: 10.3390/s20092467.

[12] S. Bhat and D. Koundal, "Multi-focus Image Fusion using Neutrosophic based Wavelet Transform," *Appl. Soft Comp.*, vol. 106, Jul. 2021, Art. no. 107307, doi: 10.1016/j.asoc.2021.107307.

[13] A. Prasad, A. K. Tyagi, M. M. Althobaiti, A. Almulihi, R. F. Mansour, and A. M. Mahmoud, "Human Activity Recognition Using Cell Phone-Based Accelerometer and Convolutional Neural Network," *Appl. Sci.*, vol. 11, no. 24, Dec. 2021, Art. no. 12099, doi: 10.3390/app112412099.

[14] J. G. Colli Alfaro and A. L. Trejos, "User-Independent Hand Gesture Recognition Classification Models Using Sensor Fusion," *Sensors*, vol. 22, no. 4, p. 1321, Feb. 2022, doi: 10.3390/s22041321.

[15] Q. Li and R. Langari, "EMG-based HCI Using CNN-LSTM Neural Network for Dynamic Hand Gestures Recognition," *IFAC-PapersOnLine*, vol. 55, no. 37, pp. 426–431, 2022, doi: 10.1016/j.ifacol.2022.11.220.

[16] Z. Zhang and E. C. Kan, "Novel Muscle Monitoring by Radiomyography(RMG) and Application to Hand Gesture Recognition," Nov. 07, 2022, *arXiv*: arXiv:2211.03767. doi: 10.48550/arXiv.2211.03767.

[17] K. S. Prakash and N. Kunju, "An optimized electrode configuration for wrist wearable EMG-based hand gesture recognition using machine learning," *Expert Sys. Applic.*, vol. 274, May 2025, Art. no. 127040, doi: 10.1016/j.eswa.2025.127040.

[18] R. Sebbah and F. Z. Chelali, "IWBC and LFD for Static and Dynamic Hand Gesture Recognition," in *2024 8th Int. Conf. Image Sign. Proces. Applic. (ISPA)*, Biskra, Algeria: IEEE, Apr. 2024, pp. 1–7. doi: 10.1109/ISPA59904.2024.10536841.

[19] H. Feng, "Machine Learning Enhanced Smart Sensors for Hand Gesture Recognition," in *2024 3rd Int. Symp. Sensor Techn. Control (ISSTC)*, Zhuhai, China: IEEE, Oct. 2024, pp. 70–73. doi: 10.1109/ISSTC63573.2024.10824205.

[20] P. N. Aarotale and A. Rattani, "Machine Learning-based sEMG Signal Classification for Hand Gesture Recognition," in *2024 IEEE Int. Conf. Bioinform. Biomed. (BIBM)*, Lisbon, Portugal: IEEE, Dec. 2024, pp. 6319–6326. doi: 10.1109/BIBM62325.2024.10822133.

[21] S. A. Mousavi and R. Selmic, "Wearable Smart Rings for Multifinger Gesture Recognition Using Supervised Learning," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–12, 2023, doi: 10.1109/TIM.2023.3304703.

[22] K. Challa, I. W. AlHmoud, A. K. M. Kamrul Islam, and B. Gokaraju, "EMG-Based Hand Gesture Recognition Using Individual Sensors on Different Muscle Groups," in *2023 IEEE Appl. Imag. Pattern Recogn. Worksh. (AIPR)*, St. Louis, MO, USA: IEEE, Sep. 2023, pp. 1–4. doi: 10.1109/AIPR60534.2023.10440702.

[23] F. J. Abdullayeva and K. Sh. Gurbanova, "Sign Language Hand Gesture Recognition Method based on Machine Learning," in *2022 IEEE 16th Int. Conf. Applic. Inform. Commun. Techn. (AICT)*, Washington DC, DC, USA: IEEE, Oct. 2022, pp. 1–5. doi: 10.1109/AICT55583.2022.10013627.

[24] M. Mahzabin, M. Hasan, S. Nahar, and M. U. Ahmed, "Automated Hand Gesture Recognition Using Machine Learning," in *2021 24th Int. Conf. Comp. Inform. Techn. (ICCIT)*, Dhaka, Bangladesh: IEEE, Dec. 2021, pp. 1–6. doi: 10.1109/ICCIT54785.2021.9689817.

[25] T.-H. Le, T.-H. Tran, and C. Pham, "The Internet-of-Things based hand gestures using wearable sensors for human machine interaction," in *2019 Int. Conf. Multim. Anal. Pattern Recogn. (MAPR)*, Ho Chi Minh City, Vietnam: IEEE, May 2019, pp. 1–6. doi: 10.1109/MAPR.2019.8743542.