



Hand Gesture Recognition Using Semi Vectorial Multilevel Segmentation Method with Improved ReliefF Algorithm

Md Shoaibuddin Madni^{1*} C. Vijaya²

¹Department of Electronics and Communication Engineering, Navodaya Institute of Technology, Raichur, India

²Department of Electronics and Communication Engineering,
SDM College of Engineering and Technology, Dharwad, India

* Corresponding author's Email: mdshoaibuddin.madni@navodaya.edu.in

Abstract: Currently, hand gesture recognition becomes a promising system for the digital entertainment field due to the recent development of sensor innovations and machine learning. However, hand gesture recognition is a challenging task in most of the existing models, due to background clutter, motion blur, illumination variations, and occlusions. In this research, a dynamic hand gesture recognition system is proposed to improve the performance of hand gesture recognition. At first, the normalization method is used to enhance the visibility level of the gesture images which are collected from the Indian sign language dataset. Then, semi vectorial multilevel segmentation method is employed to segment the exact gesture regions from the normalized images. Further, improved reliefF algorithm and K-nearest neighbour classifiers are used to select the optimal features and to classify the 16 gesture classes or symbols. In the experimental phase, the proposed improved reliefF- K-nearest neighbour model performance is analysed in light of Matthew's correlation coefficient, accuracy, sensitivity, specificity, and f-score. For overall 16 distinct alphabets; A, B, D, E, F, G, H, K, P, R, T, U, W, X, Y, Z, the proposed improved reliefF-K-nearest neighbour model achieves average accuracy of 98.95%. Hence, the proposed improved reliefF-K-nearest neighbour model showed maximum of 7.97% and minimum of 0.15% improvement in recognition accuracy compared to k means-neural network model, and discriminant correlation analysis based unimodal feature-level fusion model.

Keywords: Hand gesture recognition, Indian sign language, K-nearest neighbour, ReliefF algorithm, Semi vectorial multilevel segmentation method.

1. Introduction

The hand gesture is a non-verbal and natural form of communication, where gestures are meaningful and expressive body motions involving physical movements of the arms, hands, face, and figures to interact and convey meaningful information [1, 2]. The hand gesture finds in applications like understanding social behaviour, human-computer interactions, sign language recognition, intelligent surveillance system, gaming, etc. [3, 4]. Generally, hand gestures are categorized into two types such as dynamic and static. In dynamic hand gesture, the hand trajectory and shape are defined as a gesture, wherein static hand gesture, the hand posture is considered as a gesture [5, 6]. Hence, the dynamic

and static hand gestures are complex and also incur changes in hand shape during gesture performance. The vision-based dynamic hand gesture recognition system is required to detect the hand gesture, tracking the hand, and recognizing the hand shape as it moves [7, 8]. However, hand gesture tracking and detection are challenging tasks, due to the concerns caused by occlusion, motion blur, illumination variations, and background clutter [9, 10]. The new model is proposed in this research article to improve the performance of hand gesture recognition and assist people with speech disabilities.

Initially, the hand gesture images are collected from the Indian sign language dataset, and then the image normalization approach is applied to the collected hand gesture images to improve the visibility level of the gesture images. Then, a semi

vectorial multilevel segmentation method is applied to the normalized images to segment the exact gesture regions. The semi vectorial multilevel segmentation method has an extensive range of applications because it requires a simple process for gesture recognition and also it is computationally stable. In semi vectorial multilevel segmentation method, initially dilation disc operator uses a structural element to probe and expand the hand shapes. Then, the respective output image is combined with an original RGB gesture image and convert the image into Hue, Saturation, and Value (HSV) space. Further, apply multilevel Otsu thresholding algorithm and morphological close operator on the H space, because it eases the process of separating color components from the intensity components. Additionally, the close operator fills the small holes, while preserving the size and shape of the hand gestures in the image. After segmenting the hand gesture region, the segmented image is converted into one-dimensional data for better classification. Then, improved reliefF algorithm is applied to select the optimal features. In improved reliefF algorithm, the Manhattan distance measure is replaced by Minkowski distance measure to find the nearest miss and hit instances. Minkowski distance measure effectively reduces curse of dimensionality problem to achieve better recognition performance. Finally, K-Nearest Neighbour (KNN) classification technique is applied in lower-dimensional data that effectively classifies the gesture classes; A, B, D, E, F, G, H, K, P, R, T, U, W, X, Y and Z. In this research, two-handed signs are considered for experimental investigation, which is drawn from a wide repertoire of words of technical and daily use origins. In the experimental section, the proposed improved reliefF-KNN model performance is analysed through f-score, Matthew's Correlation Coefficient (MCC), accuracy, sensitivity, and specificity.

In section 2, a few recent research papers on the topic "hand gesture recognition" is surveyed. The proposed improved reliefF-KNN model is explained in Section 3. Section 4 details about the experimental analysis of the proposed improved reliefF-KNN model. The conclusion of the research work is given in Section 5.

2. Literature survey

Kane, and Khanna [11] presented a new system for recognizing dynamic gestures, which were characterized by a short-duration posture sequence. In this literature study, a depth-based modification was done for detecting the hand silhouettes, which gives a faster alternative to region-based descriptors.

Also, 1-nearest neighbour and depth matrix were used to recognize the postures, and then a dynamic naive Bayes classifier with an adaptive windowing technique was used to predict the posture sequence labels. In this literature, the experimental validation was conducted on a dynamic gesture dataset to analyse the efficiency of the developed system in light of accuracy. In hand gesture classification, the manual intervention is high in naive Bayes classifier as it works on the assumption of independent predictors.

Patil and Sinha [12] utilized Scale Invariant Feature Transform (SIFT) key points to extract the distinctive feature vectors from the Indian sign language gestures. Generally, sign languages were developed for hard hearing and deaf people for conveying the message by developing dissimilar sign patterns. SIFT feature descriptor performs reliable matching between the gesture images and extracts a large number of key points from the Indian sign language gesture images. The extracted feature points were rotation, scaling, and invariant to translation which delivers the better matching between the gesture images which were associated with blur and noise. This simulation result shows that the developed SIFT feature descriptor attained better performance in hand gesture recognition. The SIFT feature descriptor utilizes only grayscale information of the gesture images, where a lot of color information are discarded which may degrade the performance of hand gesture recognition.

Wu [13] developed a new gesture recognition system, which comprises three major phases such as image pre-processing, static and dynamic gesture recognition. Initially, background subtraction was employed for excluding the invalid gestures and then detect and track the user's hand. Next adaptive square was applied for extracting the hand region and then the feature vectors were extracted. In this literature study, the Support Vector Machine (SVM) classifier was used to classify the static hand posture. Also, nine commonly utilized dynamic hand gestures were detected by the developed gesture recognition system. In this literature, the number of feature vectors for every data point exceeds the number of training samples, where SVM classifiers underperform in this condition and also it is suitable only for binary classification.

Verma and Choudhary [14] developed a Grassmann manifold based system for hand gesture recognition. In this literature study, depth data was utilized for detecting the fingertip points and then stores it to create the trajectory. At first, hand shapes were detected using the properties like finger orientation angle, finger length, finger width, and

finger thickness. The skeleton information was used to detect the fingertip points in every frame based on the availability of skeleton data. The geometrical features were extracted and then generate a unique gesture sub-spaces using singular value decomposition for each feature matrix of every gesture set. The generated gesture sub-spaces lie on the Grassmann manifold to recognize every test gesture. The experimental result shows that the developed system has achieved better recognition accuracy related to the comparative systems. Due to hand shape variations, detection of exact finger is difficult in Grassmann manifold based system.

Baranwal and Nandi [15] used wavelet descriptor and Mel sec Frequency Cepstral Coefficients (MFCC) for dynamic gesture recognition. Initially, wavelet descriptor was utilized to diminish the dimensionality of the collected data and to extract the moment of invariant points from the hand gestures. Further, MFCC was used to identify the spectral envelope of the image frames. In the complex environment, the spectral envelope was useful in recognizing the hand gestures by removing the darkness present in the images. Finally, SVM and KNN classifiers were utilized for classifying the probe gestures. The experimental result showed that the wavelet descriptor with MFCC provides a higher recognition rate as compared to the existing systems. However, SVM is a binary classifier that supports only two-class classification, where it is not suitable for multi-class gesture recognition.

Ansari and Harit [16] implemented a new model for effective hand gesture recognition, which mainly includes four steps such as image pre-processing, segmentation, feature extraction and classification. Initially, a median filter was employed to remove spiky noise from the collected images and to preserve the image edge information. Then, normalized cuts and k-means clustering algorithms were employed for the exact segmentation of the gesture region. Further, the features were extracted from the segmented regions using SIFT, Speeded up Robust Feature (SURF) and Viewpoint Feature Histogram (VFH). Finally, the classification was accomplished by utilizing the neural network. In this literature study, the developed model performance was investigated on the Indian sign language database through recognition rate. However, the k-means clustering consumes more time to predict the k-value with more number of gestures classes (16 classes).

Joshi [21] used Discriminant Correlation Analysis (DCA) based unimodal feature-level fusion model for hand gesture recognition on Indian sign language dataset. DCA was a feature level fusion approach that minimizes the intra class separability,

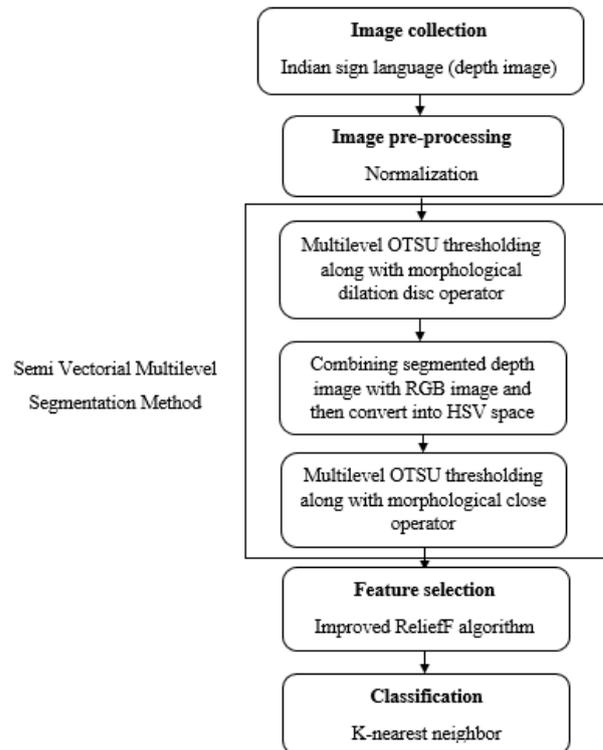


Figure. 1 The workflow of the proposed improved relief F-KNN model

and maximizes the inter class separability of feature sets. The major aim of the developed model was to combine different feature sets into a single feature vector with more discriminative power. Extensive experiment showed that the developed model attained better performance in hand gesture recognition in terms of accuracy. The size of feature vector imposes curse of dimensionality issue in this literature study. The improved reliefF-KNN model is proposed in this research paper to highlight the above-mentioned issues and to improve the performance of hand gesture recognition.

3. Methodology

In hand gesture recognition, the proposed improved reliefF-KNN model contains five phases like image collection: Indian sign language, pre-processing: image normalization, segmentation: semi vectorial multilevel segmentation method, feature selection: improved reliefF algorithm, and classification: k nearest neighbour. The workflow of the proposed improved reliefF-KNN model is presented in Fig. 1.

3.1 Dataset description and pre-processing

In the proposed improved reliefF-KNN model, the input gesture images are collected from the Indian

sign language dataset [16]. The undertaken dataset comprises 140 symbols, which are collected from 18 individuals with a total of 5041 gesture images. Mostly, the symbols consist of two-handed signs, which are drawn from daily use origins. According to [16], a total of 16 symbols (A, B, D, E, F, G, H, K, P, R, T, U, W, X, Y, and Z) are considered for experimental investigation in this research article with the total of 576 depth and color images. After the collection of depth gesture images, a normalization method is undertaken for improving the visibility level of the depth gesture images by varying the range of pixel values [17]. Image normalization method is mathematically indicated Eq. (1).

$$I_{norm} = (I - Min) \times \frac{newMax - newMin}{Max - Min} + newMin \quad (1)$$

where, I is represented as depth gesture images, I_{norm} is denoted as normalized gesture images, and $Max - Min$ is stated as the maximum and minimum image pixel value that ranges between 0 to 255. The sample original depth and pre-processed gesture image of Indian sign language dataset is stated in Fig. 2.

3.2 Semi vectorial multilevel segmentation method

After pre-processing the original depth gesture images, a semi vectorial multilevel segmentation method is used to segment the exact gesture regions. The semi vectorial multilevel segmentation method is simple and effective for segmentation that utilizes only the maximum variance values of the classes. Compared to other thresholding algorithms, the semi vectorial multilevel segmentation method is ease in the implementation, fast processing speed, and smaller storage space. At first, the intensity level L of

the normalized image I_{norm} is calculated by using Eq. (2).

$$PH_i^e = \frac{H_i^e}{M}, \sum_{i=1}^M PH_i^e = 1, e = \begin{cases} 1,2,3 & \text{if } RGB \\ 1 & \text{if } grayscale \end{cases} \quad (2)$$

where, PH_i^e is denoted as distribution probability, H_i^e is stated as pixel value, which corresponds to the intensity level from i until e , M is stated as the number of pixels in the normalized image I_{norm} , IN is denoted as intensity level ($0 \leq i \leq L - 1$) and E is indicated as image components (RGB or grayscale). In probability distribution, the histogram value is normalized using Eq. (3).

$$w_o^E(th) = \sum_{i=1}^{th} PH_i^e, w_1^E(th) = \sum_{i=th+1}^L PH_i^e \quad (3)$$

$$E_1 = \frac{PH_1^e}{w_o^e(th)}, \dots, \frac{PH_{th}^e}{w_o^e(th)}, C_2 = \frac{PH_{th+1}^e}{w_1^e(th)}, \dots, \frac{PH_L^e}{w_1^e(th)}$$

$w_0(th)$ and $w_1(th)$ are indicated as the distribution probability from E_1 and E_2 . Then, compute the average levels and variants between the classes C using the Eq. (4) and (5).

$$\mu_0^e = \sum_{i=1}^{th} \frac{iPH_i^e}{w_o^e(th)}, \mu_1^e = \sum_{i=th+1}^L \frac{iPH_i^e}{w_1^e(th)} \quad (4)$$

$$\sigma^{2e} = \sigma_1^e + \sigma_2^e, \quad (5)$$

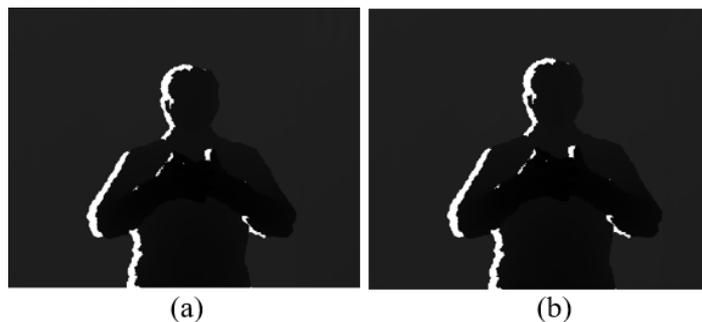


Figure. 2: (a) original depth gesture image and (b) pre-processed image



Figure. 3 Output image: (a) multilevel Otsu thresholding algorithm and (b) dilation disc operator

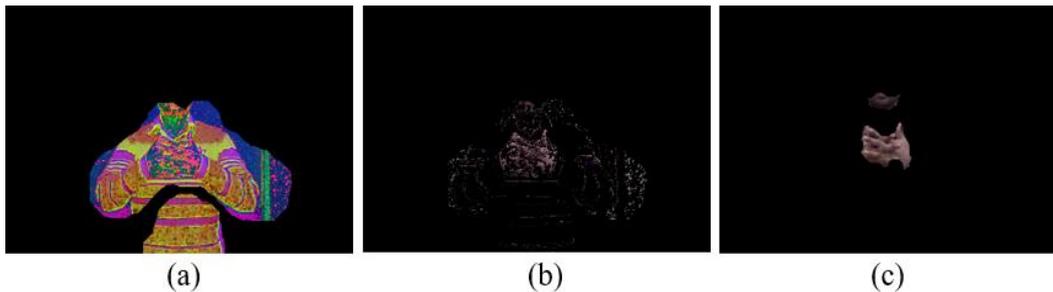


Figure. 4 Output image: (a) HSV space, (b) multilevel Otsu thresholding algorithm and (c) close operator

where,

$\sigma_1^e = w_0^e (\mu_0^e + \mu_T^e)^2, \sigma_2^e = w_1^e (\mu_1^e + \mu_T^e)^2, \mu_0^e$ and μ_1^e are represented as average rate for class variants 1 and 2, σ^{2e} is indicated as variants between the classes C , σ_1^e and σ_2^e are indicated as class variants 1 and 2 [18]. Hence, the objective function is estimated using Eq. (6).

$$J(th) = \max(\sigma^{2e}(th)), 0 \leq th_i \leq L - 1, i = 1,2,3, \dots K \quad (6)$$

where, $th = th_1, th_2, \dots, th_{K-1}$ is indicated as a vector that consists of multiple thresholds. By increasing the objective function, the Otsu between class variance function is maximized to obtain an optimal threshold level of depth gesture image for better segmentation. Then, the morphological dilation disc operator (with the size of 25) is applied to the output of the multilevel Otsu thresholding algorithm. Hence, the dilation disc operator uses a structural element to probe and expand the shapes that contain in the output image of the multilevel Otsu thresholding algorithm. The output image of the multilevel Otsu thresholding algorithm and dilation disc operator is denoted in Fig. 3.

Further, combine the output image of the multilevel Otsu thresholding algorithm with the original RGB gesture image and convert the image into HSV space. Next, apply multilevel Otsu thresholding algorithm and morphological close operator on the H space, where hue space represents the angle of the color on the RGB circle, where 0°

result in red, 120° result in green, and 240° result in blue. Additionally, the morphological close operator dilates the image and then erodes the image by utilizing a structuring element. The morphological close operator fills the small holes in the image while preserving the size and shape of the hand gesture in the image. The output image of HSV space, multilevel Otsu thresholding algorithm, and close operator are stated in Fig. 4.

For better hand gesture recognition, convert the segmented gesture image 32×32 into one-dimensional values A , where the final segmented image is presented in Fig. 4(c). After image conversion, feature selection is accomplished using Particle swarm optimization (PSO), reliefF, infinite algorithm, and improved reliefF algorithm to reduce the dimensionality of feature values. Among the available techniques, improved reliefF algorithm showed better performance in feature selection, because it effectively estimates the quality of features in dimensionality problems with strong dependencies between features. A detailed explanation about the improved reliefF algorithm is given in section 3.3.

3.3 Improved ReliefF Algorithm

ReliefF is an effective feature selection algorithm, which is robust and effectively deals with noisy and incomplete data. In reliefF algorithm, the instances f_i are selected randomly and then the k nearest neighbour values are searched in similar classes which are called as nearest hit h_j instances and the instances in the different classes is called as nearest

miss m_j instances. In reliefF algorithm, Manhattan distance is used to determine both h_j and m_j instances, but it consumes more time to calculate the distance D between the instances f_i . In order to highlight this concern, Minkowski distance measure is used instead of Manhattan distance to identify the nearest miss and hit instances.

Minkowski distance measure uses a low number of instances to represent the data to achieve precise gesture recognition, also it effectively decreases the curse of dimensionality issue. For all the attributes A , improved reliefF algorithm updates the quality estimation $W[A]$ based on the f_i , h_j and m_j instances. If the instances f_i and h_j have the same values of A , it is separated into two instances with the same classes that help in minimizing the $W[A]$ quality estimation. Correspondingly, if the instances f_i and h_j have different values of A , it is separated into two instances with different classes that help to maximize the $W[A]$ quality estimation [19]. The whole process is repeated for n times, where n is indicated as user-determined parameter value. In this research study, the user-determined parameter value is fixed as 25, because the selection process is saturated after 25. In reliefF algorithm, the quality estimation $W[A]$ is updated using Eq. (7), (8), (9), and (10).

$$W[A] = (\bar{h} + \bar{m})/25 \quad (7)$$

Where,

$$\bar{h} = - \sum_{j=1}^k D(A, f_i, h_j)/k \quad (8)$$

$$\begin{aligned} & \bar{m} \\ &= \sum_{C \neq cl(f_i)} \left[\frac{P(C)}{1 - P(cl(f_i))} \right] \sum_{j=1}^k D(A, f_i, h_j(C)) \\ & /k \end{aligned} \quad (9)$$

Where,

$$D = |f_i - h_j| + |f_i - m_j| \quad (10)$$

where, $P(C)$ is indicated as previous class, f_i is denoted as instances, D is represented as Manhattan distance between the selected instances f_i , $W[A]$ is indicated as quality estimation, $cl(f_i)$ is represented as class of i^{th} sample, h_j is specified as nearest hit

Table 1. Features selected after applying reliefF and improved reliefF algorithm

Algorithm	Original feature values A	Selected feature values $W[A]$
ReliefF	1024	307
Improved reliefF	1024	272

as class of i^{th} sample, h_j is specified as nearest hit instances, A is represented as attributes, m_j is stated as nearest miss instances and C is stated as a total number of classes ($C = 16$). In improved reliefF algorithm, Manhattan distance D is replaced by Minkowski distance measure, and updated in the Eq. (9) and (10). The Minkowski distance measure is mathematically denoted in Eq. (11).

$$D = \left(\sum_{i=1, j=1}^N |f_i - h_j| + |f_i - m_j|^s \right)^{1/s} \quad (11)$$

where, $s \geq 1$ states that the Minkowski distance measure is a metric as a result of the Minkowski inequality, and N is indicated as number of samples $W[A]$. In table 1, the original and selected feature vectors are mentioned.

3.4 K-nearest neighbour classifier

After feature selection, classification is carried out by using KNN to classify 16 symbols (A, B, D, E, F, G, H, K, P, R, T, U, W, X, Y, and Z). The concept of KNN classifier depends on the calculation of the distance between the training and testing samples $W[A]$ to identify the nearest neighbour. In the KNN classifier, the nearest neighbours are considered based on the training and testing samples $W[A]$. Several distance measures are utilized to calculate the distance between training and testing samples $W[A]$ like Chebyshev distance, city-block, Euclidean, Minkowsky, etc. [20]. Among these available distance measures, Euclidean distance is used to calculate the distance between training and testing samples $W[A]$ in KNN classifier. The general formula of Euclidean distance measure is mathematically defined in Eq. (12).

$$Euclidean\ distance = \sqrt{\sum_{i=1}^N (q_i - p_i)^2} \quad (12)$$

where, N is indicated as the number of samples $W[A]$, q_i and p_i are indicated as testing and training

samples. The steps involved in the KNN classifier are listed as follows:

- Estimate the number of nearest neighbours (k value).
- Calculate the distance between the training and testing samples.
- Based on the k^{th} minimum distance, sort the distance, and find the nearest neighbours.
- Arrange the categories of nearest neighbours.
- Use the simple majority of nearest neighbours categories as the prediction value of new data.

4. Experimental Result and Discussion

In this research article, the proposed improved reliefF-KNN model is simulated using MATLAB (2019a) environment with windows 10 operating system, 8GB RAM, Intel i5 processor, and 4 TB hard disk. In this study, the performance of the proposed improved reliefF-KNN model is compared with K means-neural network model [16], and DCA based unimodal feature-level fusion model [21] for validating the effectiveness of the proposed model. In this scenario, Indian sign language database is used for validating the effectiveness of the proposed and existing models using accuracy, sensitivity, specificity, MCC and f-score. The mathematical expressions of accuracy, MCC, f-score, specificity and sensitivity are represented in Eq. (13-17).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (13)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \times 100 \quad (14)$$

$$F - score = \frac{2TP}{2TP + FP + FN} \times 100 \quad (15)$$

$$Specificity = \frac{TN}{TN + FP} \times 100 \quad (16)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \quad (17)$$

where, TP is denoted as true positive, TN is stated as true negative, FP is stated as false positive, and FN is represented as false negative.

4.1 Quantitative investigation

In this research, the Indian sign language database is undertaken for validating the performance of the proposed improved relief F-KNN model. The Indian sign language database comprises 140 symbols that are collected from 18 individuals and has a total of 5041 images. In this study, 16 symbols (A, B, D, E, F, G, H, K, P, R, T, U, W, X, Y, and Z) are considered for experimental analysis with the total of 576 images. Here, the undertaken symbols contain two-handed signs that are drawn from daily use origins. In this research, the performance is investigated with dissimilar classifiers like Earth Mover Distance (EMD), random forest, Deep Neural Network (DNN), Multi Support Vector Machine (MSVM) and KNN classifier, and feature selection algorithms like Infinite, reliefF, Particle Swarm Optimization (PSO), and improved reliefF algorithm. In table 2, the performance is analysed with dissimilar classifiers without using feature selection algorithm. By investigating table 2, the KNN classifier attained better performance in hand gesture recognition compared to other classifiers in light of f- score, accuracy, sensitivity, MCC, and specificity. Hence, EMD attained minimum without feature selection algorithm recognition accuracy, because it completely relies on discriminative distance between the feature vectors. The MSVM is not suitable for

Table 2. Performance analysis of dissimilar classifiers without feature selection algorithm

Without feature selection					
Performance metrics	EMD	DNN	MSVM	Random forest	KNN
Accuracy (%)	66.84	33.45	91.32	92.56	96.33
Sensitivity (%)	64.04	31.12	93.52	93.30	94.81
Specificity (%)	62.30	31.13	95.78	93.67	93.33
MCC (%)	69.30	67.54	92.13	93.62	93.69
F-score (%)	65.32	55.32	95.39	93.87	95.83

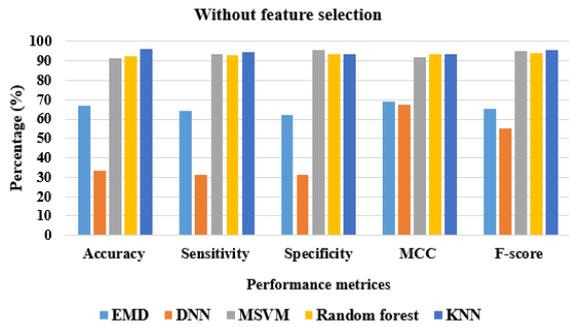


Figure. 5 Graphical analysis of dissimilar classifiers

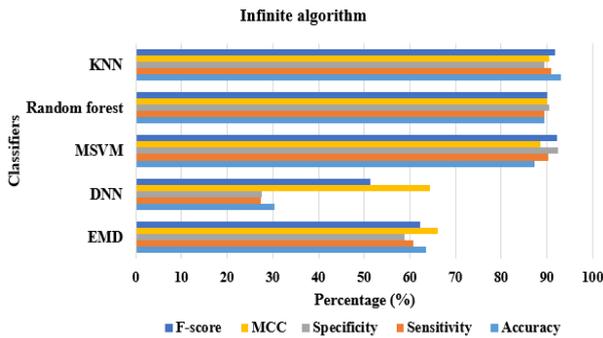


Figure. 6 Graphical analysis of dissimilar classifiers with an infinite algorithm

Table 3. Performance analysis of dissimilar classifiers with an infinite algorithm

Infinite algorithm					
Performance metrics	EMD	DNN	MSVM	Random forest	KNN
Accuracy (%)	63.49	30.30	87.34	89.48	93.16
Sensitivity (%)	60.68	27.34	90.36	89.35	91.01
Specificity (%)	58.88	27.57	92.45	90.60	89.36
MCC (%)	66.17	64.43	88.64	90.07	90.58
F-score (%)	62.30	51.44	92.29	90.02	91.85

Table 4. Performance analysis of dissimilar classifiers with the PSO algorithm

PSO algorithm					
Performance metrics	EMD	DNN	MSVM	Random forest	KNN
Accuracy (%)	60.57	26.53	84.56	85.62	89.37
Sensitivity (%)	57.53	24.19	86.74	86.64	88.72
Specificity (%)	56.05	25.06	88.80	86.71	86.75
MCC (%)	62.96	60.56	85.63	87.47	87.62
F-score (%)	59.08	49.25	88.60	87.13	89.06

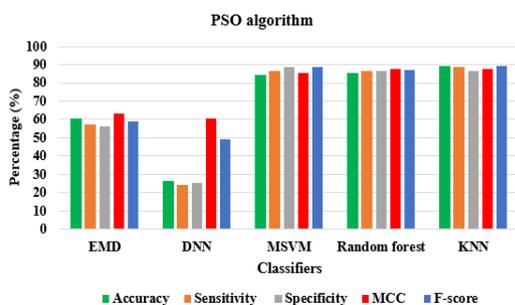


Figure. 7 Graphical analysis of dissimilar classifiers with PSO algorithm

large databases and also it does not perform very well when the target classes are overlapped. The graphical analysis of dissimilar classifiers without feature selection algorithm is denoted in Fig. 5.

However, the infinite algorithm works based on linear regression, where it has a major concern of high computational complexity that is costly for high dimensional data. The graphical analysis of dissimilar classifiers with infinite algorithms is presented in Fig. 6. In Table 4, the performance is analysed with dissimilar classifiers and the PSO algorithm. Table 4 shows that the combination (PSO with KNN classifier) attained average classification accuracy of 89.37%, which is limited compared to the combination of infinite with KNN classifier. The major concern in the PSO algorithm is it easily falls into the local optimum of higher dimensional space and also has a lower convergence rate in the iterative process. The graphical analysis of dissimilar classifiers with the PSO algorithm is denoted in Fig. 7.

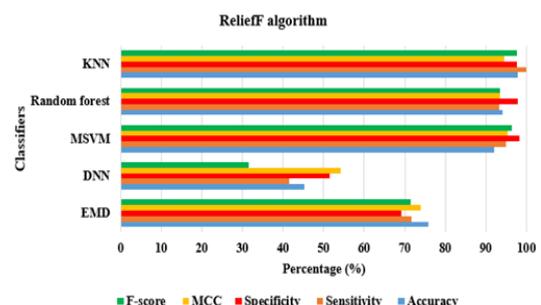


Figure. 8 Graphical analysis of dissimilar classifiers with relief algorithm

Table 5. Performance analysis of dissimilar classifiers with reliefF algorithm

ReliefF algorithm					
Performance metrics	EMD	DNN	MSVM	Random forest	KNN
Accuracy (%)	75.89	45.24	92.05	94.16	97.85
Sensitivity (%)	71.62	41.54	94.91	93.19	99.84
Specificity (%)	69.26	51.46	98.20	97.58	97.60
MCC (%)	73.85	54.23	95.32	93.46	94.46
F-score (%)	71.39	31.44	96.30	93.48	97.68

Table 6. Performance analysis of dissimilar classifiers with improved reliefF algorithm

Improved reliefF algorithm					
Performance metrics	EMD	DNN	MSVM	Random forest	KNN
Accuracy (%)	76.92	46.02	94	95.64	98.95
Sensitivity (%)	72.10	42	95.97	94.84	99.88
Specificity (%)	70.23	54.12	98.92	98	98.10
MCC (%)	74.20	58.80	96.30	94.26	95.50
F-score (%)	74.94	34.42	97.45	95.80	98.26

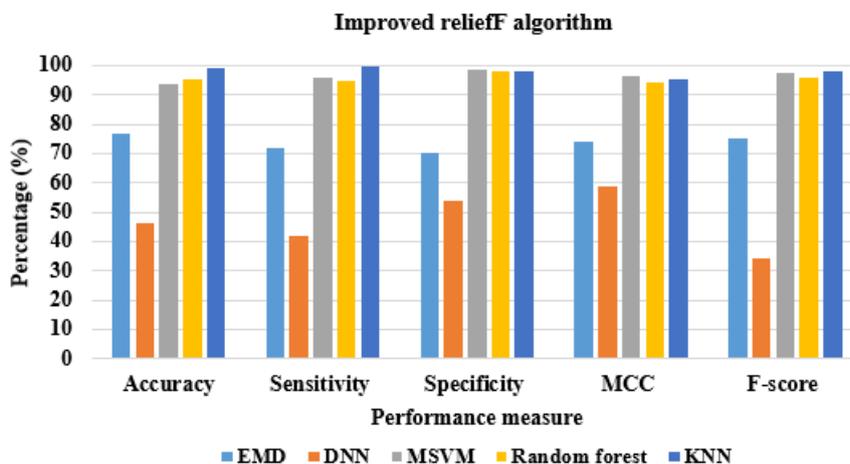


Figure. 9 Graphical analysis of dissimilar classifiers with improved reliefF algorithm

In Table 5, the performance is investigated with dissimilar classifiers and reliefF algorithm. Table 5 shows that the reliefF with KNN combination obtained average accuracy of 97.85%, f-score of 97.68%, sensitivity of 99.84%, the specificity of 97.6% and MCC of 94.46%, which is better compared to other combinations in hand gesture recognition.

For individual class, reliefF with KNN combination showed minimum classification accuracy of 93% in E and F symbols, because both the gestures are similar in manner, where reliefF with KNN achieved the maximum recognition accuracy of 100% in P and T symbols. The graphical analysis of dissimilar classifiers with reliefF algorithm is presented in Fig. 8.

In Table 6, the performance is analyzed with dissimilar classifiers and improved reliefF algorithm in terms of accuracy, sensitivity, specificity, MCC, and f-score. Table 6 shows that the KNN classifier with improved reliefF algorithm obtained average

classification accuracy of 98.95%, sensitivity of 99.88%, specificity of 98.10%, MCC of 95.50%, and f-score of 98.26%, which are better compared to other classifiers in hand gesture recognition. The KNN classifier is very easy to implement for multi class problems and works well in higher-dimensional data. Additionally, the improved reliefF algorithm reduces the training time of KNN classifier by reducing the dimensionality of the data. The graphical analysis of dissimilar classifiers with improved reliefF algorithm is presented in Fig. 9.

4.2 Comparative Investigation

The comparative investigation between the proposed and existing model is represented in table 7. Ansari and Harit, [16] implemented a new model for dynamic hand gesture recognition. The gesture images were collected from the Indian sign language dataset and then a median filter was applied to remove spiky noise and to preserve the image edge

Table 7. A comparative investigation between proposed and existing models

Methodology	Recognition accuracy (%)	Running time in seconds
K means-neural network [16]	90.68	6.0847
DCA based unimodal feature-level fusion model [21]	98.80	-
Improved ReliefF-KNN	98.95	5.984

information. Then, the k-means clustering algorithm and normalized cuts were applied to segment the exact gesture region from the denoised image. Further, SIFT, SURF, and VFH feature descriptors were used to extract the feature vectors from the segmented region. At last, the neural network classification technique is used for classifying 16 gesture classes such as A, B, D, E, F, G, H, K, P, R, T, U, W, X, Y, and Z. In the experimental segment, the developed model attained 90.68% of accuracy in hand gesture recognition and consumed running time of 6.0847 seconds. Additionally, Joshi [21] developed DCA based unimodal feature-level fusion model for hand gesture recognition on Indian sign language dataset. Extensive experiment showed that the developed model achieved 98.80% accuracy in hand gesture recognition. Compared to the existing research papers, the proposed improved reliefF-KNN model achieved 98.95% of recognition accuracy, and consumed limited running time of 5.984 seconds in hand gesture recognition.

Several features are obtained by converting the two-dimensional segmented image into one dimensional numerical values, so in this paper improved reliefF algorithm is applied to reduce the dimensionality of the features, where the effect of feature selection is represented in table 6. Hence, the recognition accuracy of with improved reliefF algorithm is 2.62% better than the without feature selection algorithm in hand gesture recognition. The proposed algorithm effectively overcomes the problems mentioned in the literatures [13, 15, 16], where the proposed improved reliefF algorithm consumes less running time, and it achieved better performance in multi class classification.

5. Conclusion

In this research, improved reliefF-KNN model is proposed to enhance the performance of hand gesture recognition. Initially, the normalization method is undertaken to enhance the visual level of the images, which are collected from the Indian sign language database. Then, the semi vectorial multilevel segmentation method is used to segment the gesture regions and further improved reliefF algorithm is

applied to reduce the dimension of the data that completely resolves the problem of “curse of dimensionality”. Additionally, KNN classifier is applied to classify the 16 gesture symbols. In the experimental phase, the improved reliefF-KNN model achieved effective performance in hand gesture recognition in terms of accuracy, sensitivity, specificity, MCC and f-score. Compared to the existing k means-neural network model and DCA based unimodal feature-level fusion model, the proposed improved reliefF-KNN model showed maximum of 7.97% and minimum of 0.15% improvement in recognition accuracy. In future work, a new optimization-based clustering algorithm can be included in the proposed improved reliefF-KNN model to further improve the performance of hand gesture recognition.

Notation	Parameter
PH_i^e	Distribution probability
H_i^e	Pixel value
E	Image components
IN	Intensity level
$P(C)$	Previous class
D	Distance between the selected instances
f_i	Instances
$W[A]$	Quality estimation
h_j	Nearest hit instances
A	Attributes
m_j	Nearest miss instances
C	Total number of classes
N	Number of samples

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st author. The supervision and project administration, have been done by 2nd author.

References

- [1] H. Tang, H. Liu, W. Xiao, and N. Sebe, “Fast and robust dynamic hand gesture recognition via key frames extraction and feature fusion”, *Neurocomputing*, Vol. 331, pp. 424-433, 2019.

- [2] A. Sharma, A. Mittal, S. Singh, and V. Awatramani, "Hand Gesture Recognition using Image Processing and Feature Extraction Techniques", *Procedia Computer Science*, Vol. 173, pp. 181-190, 2020.
- [3] S. Ameur and A.B. Khalifa, "A novel Hybrid Bidirectional Unidirectional LSTM Network for Dynamic Hand Gesture Recognition with Leap Motion", *Entertainment Computing*, pp. 100373, 2020.
- [4] J. C. Nunez, R. Cabido, J. J. Pantrigo, A. S. Montemayor, and J. F. Velez, "Convolutional neural networks and long short-term memory for skeleton-based human activity and hand gesture recognition", *Pattern Recognition*, Vol. 76, pp. 80-94, 2018.
- [5] Y. Yao and Y. Fu, "Contour model-based hand-gesture recognition using the Kinect sensor", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 24, No. 11, pp. 1935-1944, 2014.
- [6] G. A. Rao and P. V. V. Kishore, "Selfie video based continuous Indian sign language recognition system", *Ain Shams Engineering Journal*, Vol. 9, No. 4, pp. 1929-1939, 2018.
- [7] J. Jeong and Y. Jang, "Max-min hand cropping method for robust hand region extraction in the image-based hand gesture recognition", *Soft Computing*, Vol. 19, No. 4, pp. 815-818, 2015.
- [8] P. Bao, A. I. Maqueda, C. R. del-Blanco, and N. García, "Tiny hand gesture recognition without localization via a deep convolutional network", *IEEE Transactions on Consumer Electronics*, Vol. 63, No. 3, pp. 251-257, 2017.
- [9] M. F. Wahid, R. Tafreshi, M. Al-Sowaidi, and R. Langari, "Subject-independent hand gesture recognition using normalization and machine learning algorithms", *Journal of computational science*, Vol. 27, pp. 69-76, 2018.
- [10] O. K. Oyedotun and A. Khashman, "Deep learning in vision-based static hand gesture recognition", *Neural Computing and Applications*, Vol. 28, No. 12, pp. 3941-3951, 2017.
- [11] L. Kane and P. Khanna, "Depth matrix and adaptive Bayes classifier based dynamic hand gesture recognition", *Pattern Recognition Letters*, Vol. 120, pp. 24-30, 2019.
- [12] S. B. Patil and G. R. Sinha, "Distinctive feature extraction for Indian Sign Language (ISL) gesture using scale invariant feature Transform (SIFT)", *Journal of The Institution of Engineers (India): Series B*, Vol. 98, No. 1, pp. 19-26, 2017.
- [13] C. H. Wu, W. L. Chen, and C. H. Lin, "Depth-based hand gesture recognition", *Multimedia Tools and Applications*, Vol. 75, No. 12, pp. 7065-7086, 2016.
- [14] B. Verma and A. Choudhary, "Grassmann manifold based dynamic hand gesture recognition using depth data", *Multimedia Tools and Applications*, Vol. 79, No. 3, pp. 1-25, 2019.
- [15] N. Baranwal and G. C. Nandi, "An efficient gesture based humanoid learning using wavelet descriptor and MFCC techniques", *International Journal of Machine Learning and Cybernetics*, Vol. 8, No. 4, pp. 1369-1388, 2017.
- [16] Z. A. Ansari and G. Harit, "Nearest neighbour classification of Indian sign language gestures using kinect camera", *Sadhana*, Vol. 41, No. 2, pp. 161-182, 2016.
- [17] K. M. Koo and E. Y. Cha, "Image recognition performance enhancements using image normalization", *Human-centric Computing and Information Sciences*, Vol. 7, No. 1, pp. 1-11, 2017.
- [18] M. Fachrurrozi, N. R. Dela, Y. Mahyudin, and H. K. Putra, "Tongue Image Segmentation using Hybrid Multilevel Otsu Thresholding and Harmony Search Algorithm", *Journal of Physics: Conference Series, IOP Publishing*, Vol. 1196, No. 1, pp. 012072, 2019.
- [19] L. Sun, X. Kong, J. Xu, R. Zhai, and S. Zhang, "A hybrid gene selection method based on ReliefF and ant colony optimization algorithm for tumor classification", *Scientific Reports*, Vol. 9, No. 1, pp. 1-14, 2019.
- [20] N. Ali, D. Neagu, and P. Trundle, "Evaluation of k-nearest neighbour classifier performance for heterogeneous data sets", *SN Applied Sciences*, Vol. 1, No. 12, pp. 1559, 2019.
- [21] G. Joshi, R. Vig, and S. Singh, "DCA-based unimodal feature-level fusion of orthogonal moments for Indian sign language dataset", *IET Computer Vision*, Vol. 12, No. 5, pp. 570-577, 2018.