Received: December 11, 2022. Revised: February 10, 2023.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Handwritten Character Recognition Using Morlet Stacked Sparse Auto Encoder based Feature Dimension Reduction

Keerthi Prasad Ganeshaiah[1]* **Vinay Hegde[2]**

[1]*Department of Artificial Intelligence and Machine Learning, GM Institute of Technology, Davanagere, India*
[2]*Departmentt. of Computer Science & Engineering, Rashtreeya Vidyalaya College of Engineering, Bangalore, India*
* Corresponding author's Email: keerthiprasad@gmit.ac.in.

Abstract: Handwritten character recognition (HCR) is a growing field in the applications of pattern recognition, image processing, communication technologies, and so on. But, the identification of handwritten characters affected because of different styles of writers, or even one writer's style differed according to the conditions. Moreover, the huge amount of features from the handwritten characters also affect the classification. To address the aforementioned issues, the hybrid feature extraction (HFE) with morlet stacked sparse auto-encoder (MSSAE) based feature dimensionality reduction is proposed for improving the classification of HCR. Here, the HFE is the combination of different textural and shape features such as histogram of oriented gradients (HOG), gray-level co-occurrence matrix (GLCM), discrete wavelet transform (DWT), and skeleton features. The morlet wavelet activation function is used in the MSSAE to enhance the dimensionality reduction ability that used for an effective reduction of feature dimensions. The reduction of feature dimension using MSSAE is used to improve the classification using multi-class support vector machine (MSVM). In this research, the real-time indigenous languages i.e., English and Kannada handwritten characters from the Chars74K dataset are used for the analysis i.e., offline recognition. On the other hand, the online recognition of Kannada characters is also done for analyzing the HFE-MSSAE method. The HFE-MSSAE method is analyzed in terms of accuracy, precision, recall, CSI and F-measure. The existing researches namely convolutional neural network (CNN), hybrid feature based long short term memory (HF-LSTM) and elephant herding optimization - long short term memory (EHO-LSTM) are used to evaluate the HFE-MSSAE method. The accuracy of the HFE-MSSAE for Kannada is 96.73% which is higher than the HF-LSTM and EHO-LSTM.

Keywords: Accuracy, Feature dimensionality reduction, Handwritten character recognition, Hybrid feature extraction, Morlet stacked sparse auto-encoder, Multi-class support vector machine.

1. Introduction

HCR approach is used to translate various document types into analyzable, editable, and searchable information. The important goal of HSR is to follow the ability of humans to read in a way that machines can read, edit and communicate with text similar to humans in less time [1]. The HCR's objective is to transfer the handwritten text documents from digital images to encode characters that are readable and editable by using the word processing application [2, 3]. The HCR is used in various applications such as tracking of vehicle number plates, input from paper, comprehensible handwriting, signature verification, signboard reading, and touch screen device [4]. HCR is one of the important applications in the research of pattern recognition which improves the friendliness and socialization of mobile devices and networks. According to the input, the HCR is mainly separated into two types such as online and offline [5, 6].

The characters on the smart devices (e.g., smartphones and tablets) are online whereas the identification of handwritten text on paper is related to the offline [7-9]. The characters written by each person aren't matching and change in both shapes and sizes. The identification process is highly difficult and challenging, because of the several changes in the writing styles of each person's characters [10-12]. Hence, the HCR is difficult because of the huge variations in intra- and inter-intensity and also intra-class differences and inter-class resemblances [13, 14]. In the HCR, the selection of an adequate feature that denotes the input sample is considered one of the important factors which create an impact on the recognition performances. The performance of the learning algorithm may have degraded because of the large number of features with redundant features. Consequently, the feature selection approach is used to search for the optimal features from the entire available features for achieving better recognition [15].

The research contributions are concise as follows:

- The hybrid features of HOG, GLCM, DWT, and skeleton features are extracted during the classification. In this HFE, HOG provides robustness against the edge intensity, shadow, and illumination; GLCM provides the textural information; the combination of low-low and high-low subbands of DWT increases the classification accuracy whereas the skeleton feature provides major curvatures of the object without any noise.
- Further, the MSSAE-based feature dimensionality reduction with enhanced reconstruction ability is proposed for removing the redundant features from the hybrid features. Because the redundant features that exist in the feature set affect the classification. Moreover, the developed HFE-MSSAE provides the real-time English and Kannada HCR.

This research paper is arranged as follows: The existing works related to the HCR are provided in section 2. A clear explanation of the HFE-MSSAE is explained in section 3. Section 4 delivers the outcomes of the HFE-MSSAE whereas the conclusion is presented in section 5.

2. Related work

Albahli [16] implemented the modified faster regional convolutional neural network (Faster-RCNN) to perform an effective recognition of handwritten digits. At first, the annotations were developed for achieving the annotations followed by the deep features were computed by introducing the DenseNet-41 of Faster-RCNN. Further, the digits were classified by using the regressor and classification layer. The developed Faster-RCNN used only a few layers to achieve better classification. The visual similarities among other classes degraded the performance of certain classes.

Asghar Ali chandio [17] presented the multi-scale feature aggregation (MFSA) and multi-level feature fusion (MLFF) for recognizing the separated Urdu characters from natural images. An up-sampling and addition operations were used to aggregate the multi-scale features followed it was integration with the high-level features. Further, the robust features are generated by fusing the outputs of MSFA and MLFF networks. The Softmax classifier was used to accomplish the prediction based on combined features. However, this work was processed only with offline data, it failed to examine with online data.

Riadh Harizi [18] developed the text recognition system to offer word detection according to the text images cropped from real world images. The detection depends on the joint stepwise character/word modeling based on deep learning. The characters with huge changes that exist in the natural scenes were handled by a CNN which used to perform feature extraction. This work was failed to reduce the dimensions of the features from entire features, therefore the features with redundant information degraded the performance.

Asha Kathigi and Krishnappa Honnamachanahalli Kariputtaiah [19] presented the HF-LSTM approach to perform character recognition. From the image, the individual characters were obtained using the skewed line segmentation approach. Next, the steerable pyramid transform and discrete wavelet transform were used in the hybrid feature extraction method to acquire the differential feature vectors. The hybrid features combined both the higher and lower level features resulted in higher accuracy through LSTM. Here, the LSTM classifier processed all the extracted features, since the features have redundant information it degraded the classification performance.

Zhijun Liu [20] developed the combination of depth neural network (DNN) and probability model to enhance character recognition. The statistical and local gradient features were extracted by using the GLCM, HOG, and grid-level features. The expression of the probability model depends on the posterior possibility value for which a character with a higher value was discovered in the input. Hence, the developed DNN with probability model was used to enhance the accuracy of the system. The classification performance of DNN mainly relies on the amount of hidden layers and neurons that exist in those layers.

Guptha [21] presented the deep learning based approach to perform automatic character recognition.

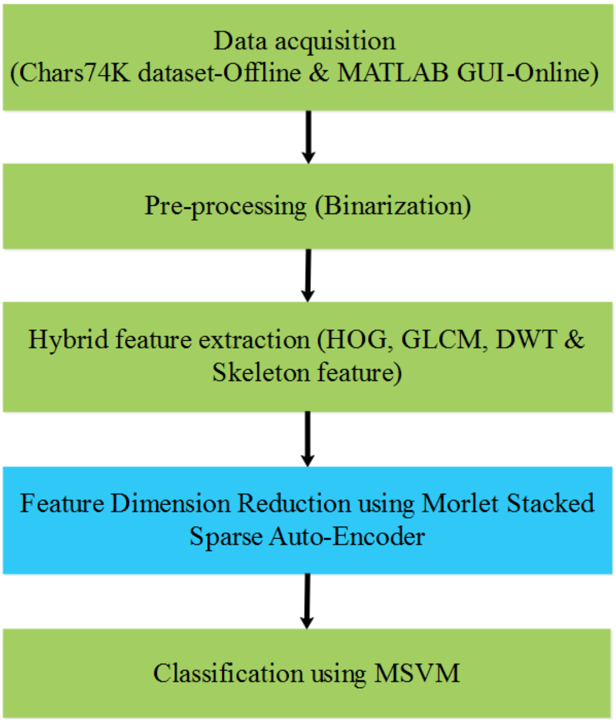

Figure. 1 Block diagram of the HFE-MSSAE method

At first, Gaussian filtering and skew detection were used to perform the pre-processing for the images acquired from Chars74K & MADbase digits datasets. The projection profile and thresholding approaches were used to segment the individual lines and characters. The features from the segmented images were obtained by inverse difference moment normalized (IDMN) and enhanced local binary pattern (ELBP) descriptors. Further, elephant herding optimization (EHO) was used to choose the features followed by the classification was accomplished by long short term memory (LSTM). However, deep learning based feature optimization was required for further improving the classification by selecting the optimal features.

The problems from the existing researches are mentioned as follows: classification using large feature vector, analysis of HCR only using offline data and misclassification because of visual similarities. In this research, an effective classification is performed using the HFE-MSSAE for both the offline and online real-time HCR. The reconstruction ability of MSSAE is used for effective feature dimension reduction. Therefore, this MSSAE is used to enhance the classification performed using MSVM.

3. HFE-MSSAE method

In this research, both the offline and online real-time HCR is performed by using HFE-MSSAE. The performance of the HCR is improved by using the HFE along with feature dimensionality minimization using MSSAE. The MSSAE is developed with effective reconstruction ability helps to improve the dimension reduction performances. The optimal feature reduction using MSSAE from the whole feature set helps to improve the recognition accuracy. Here, the classification of characters is performed by using the MSVM. The block diagram of the HFE-MSSAE method is shown in Fig. 1.

Figure. 2 Sample images of Kannada characters

3.1 Data acquisition

In this HFE-MSSAE method, handwritten character recognition is done both offline and online recognition. In offline recognition, the real-time English and Kannada characters are used from the Chars74K dataset [22]. This Chars74K dataset of Kannada language has 49 basic characters in alpha syllabaries where the consonants and vowels are included for generating the 657 different classes. Moreover, the 64 classes are included in the English which has lower cases (a-z), and upper cases (A-Z) and numbers (0-9). The English characters comprises of 7705 natural, 3410 handwritten and 62,992 synthesized images. The Kannada sample images from the Chars74K dataset are shown in Fig. 2. Moreover, the input for online HCR is given through the MATLAB GUI for classifying the Kannada characters.

3.2 Pre-processing using binarization and segmentation

The images from the dataset are pre-processed using binarization for improving the ability of classification. Otsu's thresholding approach is used to identify the appropriate threshold for binarization. In each block, the pixel with a lesser grey level than the local threshold is considered as foreground; Otherwise, it is considered as the background which is expressed in Eq. (1).

$$I'(p,q) = \begin{cases} 0, & I(p,q) \geq T_0 \\ 1, & I(p,q) < T_0 \end{cases} \tag{1}$$

Where, I is the input image; p and q defines the

Received: December 11, 2022. Revised: February 10, 2023.

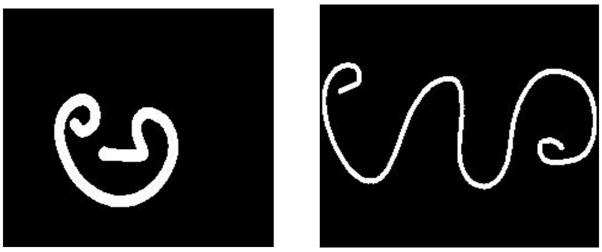

Figure. 3 Pre-processed images of Kannada characters

locations of row and column for image; I' is the preprocessed image and the threshold identified from Otsu's thresholding value is represented as T_0. The examples of preprocessed images are shown in Fig. 2. Further, the segmentation is done when the given input of online Kannada HCR is in a text format. Here, skewed line segmentation [19] is used to segment the individual characters from the text. In that skewed line segmentation, initially, line segmentation takes place followed by character segmentation done to obtain individual characters.

3.3 Hybrid feature extraction

In this HFE-MSSAE method, four different feature extraction approaches such as HOG, GLCM, DWT and skeleton features are utilized to obtain the relevant features from the preprocessed image. After extracting the features from the preprocessed image, the features from HFE are concatenated together to generate the whole feature set. The information about the feature extraction using HFE is given as follows:

3.3.1. Histogram of oriented gradients (HOG)

The HOG is generally a shape descriptor that gets the appearance and shape of the local character. This HOG defines the intensity gradient's dissemination through the region which is utilized for identifying the character. Generally, the HOG depends on the gradient direction gathering through the pixels of a small spatial region namely a cell. The gradient direction's local 1-D histogram is gathered by the HOG descriptor in each cell. The aforementioned process is completed by gathering the local histogram through the huge spatial region namely a block that is composed of cells and utilizing the results for normalizing all cells of the block. Here, the detection window is given over the overlapping grid.

The horizontal and vertical gradient of the pixel is represented in Eq. (2).

$$\begin{cases} G_p = I'(p+1,q) - I'(p-1,q) \\ G_q = I'(p,q+1) - I'(p,q-1) \end{cases} \quad (2)$$

Eqs. (3) and (4) shows the magnitude and orientation of the gradient for HOG respectively.

$$G(p,q) = \sqrt{G_p^2 + G_q^2} \quad (3)$$

$$\theta(p,q) = \arg\tan\left(\frac{G_q}{G_p}\right) \quad (4)$$

The gradient magnitude's weight which is in the same orientation are gathered to form the histogram vector of the cell. Further, the output vector of HOG is normalized cells of all components and blocks in the detection window.

3.3.2. Gray-level co-occurrence matrix (GLCM)

The GLCM [23] is a powerful feature descriptor that is used to examine the textural features of the image by evaluating the spatial connection between two pixels. Since, the distance and orientation among the pixels are varied to identify the mutual existence of pixel pairs. There are different textural features obtained from GLCM such as angular second moment, contrast, correlation, the sum of squares, sum average, inverse difference moment, entropy, sum variance, sum entropy, difference entropy, difference variance, maximum correlation coefficient and information measures of correlation. This textural information obtained from this GLCM is used to perform effective recognition.

3.3.3. DWT

The 2D DWT [24] is applied over the preprocessed image which returns four subbands such as low–low (LL), high–low (HL), low–high (LH), and high–high (HH). This 2D-DWT returns the spatial frequency components obtained from the LL subband. Eq. (5) expresses the various frequency components and each component with a related resolution to its scale.

$$DWT\left(a(I')\right) = \begin{cases} d_{WS,TF} = \sum a(I')h \times WS(I' - 2 \times WS \times TF) \\ d_{WS,TF} = \sum a(I')g \times WS(I' - 2 \times WS \times TF) \end{cases} \quad (5)$$

Where the component attribute to the preprocessed image is denoted as $a(I')$; wavelet scale is denoted as WS; translation factor is denoted as TF; high pass and low pass filter coefficients are denoted as $h(I')$ and $g(I')$ respectively.

3.3.4. Skeleton feature

The computation of smooth medial function with enhanced quality is essential in skeleton extraction

[25] because it offers effective skeleton features without noise. The curvature maxima with the level-sets of the medial function are important in skeleton point identification. The probability of each pixel being a skeleton point is referred to as the skeleton strength map (SSM) which is the mean curvature of the level-sets as shown in Eq. (6).

$$SSM = \nabla.\left(\frac{\nabla I\prime(SV,t_s)}{|\nabla I\prime(SV,t_s)|}\right) \tag{6}$$

Where the space vector is denoted as SV and diffusion time is denoted as t_s. Here, $SSM < 0$ defines the brighter regions and $SSM > 0$ defines the darker regions.

Further, the features of the HOG, GLCM, DWT and skeleton are concatenated together for generating the feature set (x) as shown in Eq. (7). Therefore, a total 20 features i.e., 1 feature from HOG, 14 features from GLCM, 4 features from DWT, and 1 feature from the skeleton are extracted from HFE.

$$x = \{HOG, GLCM, DWT, SSM\} \tag{7}$$

3.4 Feature dimension reduction using morlet stacked sparse auto-encoder

In this phase, the morlet stacked sparse auto-encoder (MSSAE) is used for performing the dimensionality reduction for the extracted features (x). Since, the sparse auto-encoder is the base for the MSSAE that is designed from auto encoder (AE). The conventional stacked sparse auto-encoder (SSAE) was required huge amount of weights, because of higher width and the depth of the model that makes the SSAE is difficult to train. Even weight decay value is incorporated in cost function for eliminating the over-fitting and numerous non-zero connection weights that affects the reconstruction quality and decreases the sparsity. Theerefore, the morlet wavelet activation function is used in the stacked sparse auto-encoder (SSAE) for avoiding the issue of sparsity reduction and enhancing the reconstruction quality. The sparsity constraints are accomplished in the hidden layer units by incorporating the sparsity constraints in AE. The kullback-leibler (KL) divergence is used by MSSAE for making it near to the specified sparse value ρ by limiting the average activation value $\hat{\rho}$ of the hidden layer neuron output. This KL divergence is used for realizing the prevention effects and it is incorporated with the cost function as the penalty parameter. Eq. (8) is used to update the cost function of MSSAE (τ).

$$\tau(x, z, \lambda) = \psi(x, z) + \lambda.\Omega_{weights} \tag{8}$$

Where, z is the output layer; the parameter utilized for controlling the regularization strength is denoted as λ; the error among the inputs and reconstructed ones is represented as $\psi(x, z)$ and the weight attenuation term is denoted as $\Omega_{weights}$ which is expressed in Eq. (9).

$$\Omega_{weights} = \frac{1}{2}\sum_{l=1}^{u}\sum_{i=1}^{s}\sum_{j=1}^{v}\left(W_{ij}^{(l)}\right)^2 \tag{9}$$

Where the number of the hidden layers is represented as u; input layer units are represented as v; hidden layer units are represented as s and the weight matrix is denoted as W.

The function of Morlet wavelet activation used in this MSSAE is expressed in Eq. (10).

$$\alpha(T) = k.\exp(-T^2/f_b)^2.\cos(2\pi f_c T) \tag{10}$$

Where, the waveform parameter set is denoted as (k, f_b, f_c); amplitude is denoted as k; bandwidth is denoted as f_b; central frequency is denoted as f_c and time is denoted as T.

The penalty term of MSSAE is also referred to as sparse regularization term which is expressed in Eq. (11).

$$E(x, \lambda, \beta, \rho) = \frac{1}{2n}\sum_{n=1}^{n}\|x_i - z_i\|^2 + \lambda.\Omega_{weights} + \beta.\Omega_{sparsity} \tag{11}$$

$$\Omega_{sparsity} = \sum_{i=1}^{D} KL(\rho||\hat{\rho}_i) = \sum_{i=1}^{D}\rho\log\left(\frac{\rho}{\hat{\rho}_i}\right) + (1-\rho)\log\left(\frac{1-\rho}{1-\hat{\rho}_i}\right) \tag{12}$$

Where, sparsity parameter is denoted as ρ; the amount of training samples is represented as n and the parameter utilized for controlling the sparse regularization is denoted as β. Eq. (13) shows the average activation values for all training samples of hidden layer's neurons i.

$$\hat{\rho}_i = \frac{1}{n}\sum_{j=1}^{n} f(w_i^{(1)}x_j + b_i^{(1)}) \tag{13}$$

Where, the training samples of j is x_j; the weight matrix row at 1st layer is $w_i^{(1)}$ and the bias vector entry is denoted as $b_i^{(1)}$.

In this MSSAE, the KL is used due to its capacity for computing the variance among two distributions. The $KL(\hat{\rho}_i||\rho_i) = 0$, when $\hat{\rho}_i = \rho_i$. The divergence of KL makes ρ and $\hat{\rho}$ close, when the difference between the ρ and $\hat{\rho}$ is large. The series integration of numerous MSSAE is used to generate the deep learning architecture of MSSAE. In this MSSAE, the output of each layer is given as input to the successive layer. The greedy training layer by layer is accomplished to train the MSSAE. The MSSAE with the Softmax regression neural network is trained and it adjusts the perimeters which are referred to as fine-tuning. Moreover, the output vector's dimension is equal to the number of labels. Therefore, the MSSAE with Morlet wavelet activation is finely tuned with respect to the labels for reducing the dimension of the features from 20 to 16.

3.5 Classification using MSVM

After reducing the dimension of the features, the features chosen by MSSAE are given to the MSVM to perform the accurate recognition of handwritten characters. A group of binary support vector machines (SVMs) are integrated for creating the MSVM while performing the classification [26]. In that, the one versus all approach is utilized for solving the problems related according to the multi-class. This one versus all establishes N-SVM models for N classes.

4. Results and discussion

The outcomes of the proposed HFE-MSSAE method-based HCR are given in this section. The implementation and simulation of the HFE-MSSAE method are done using MATLAB R2020a software. The system configuration used for this research is an i5 processor with 8GB of RAM. The dataset used to analyze the HFE-MSSAE method is the Chars74k dataset where 80% of data was taken for training and 20% of data is taken for testing. In that, the data is randomly taken for training and testing according to the iterations. The HFE-MSSAE method is analyzed by means of accuracy, precision, recall, CSI, and F-measure which are expressed in Eqs. (14)-(18).

$$Accuracy = \frac{TP+TN}{TN+TP+FN+FP} \times 100 \tag{14}$$

$$Precision = \frac{TP}{TP+FP} \times 100 \tag{15}$$

$$Recall = \frac{TP}{TP+FN} \times 100 \tag{16}$$

$$CSI = \frac{TP}{TP+FN+FP} \times 100 \tag{17}$$

$$F-measure = \frac{2TP}{2TP+FN+FP} \times 100 \tag{18}$$

Where, TP is true positive; TN is true negative; FP is false positive and FN is false negative.

4.1 Performance analysis of the HFE-MSSAE method

This section shows the performance analysis of offline HCR using the Kannada characters of Chars74k dataset. In the Chars74k dataset, the HFE without MSSAE-based feature dimension reduction achieves the recognition accuracy of 94.01% whereas the individual features such as HOG, GLCM, DWT and skeleton achieves 89.44%, 91.38%, 90.76% and 88.76%. From the results, it is known that the combined set of features provides better classification accuracy than the individual features. Additionally, the performance of the HFE-MSSAE method is improved by using the MSSAE based feature dimension reduction. The MSVM is used by the HFE-MSSAE method for classifying the real-time Kannada characters.

In this section, the HFE-MSSAE method is analyzed with different classifiers and with different feature dimension reduction approaches. The different classifiers used to evaluate the HFE-MSSAE method are neural network (NN), k-nearest neighbour (KNN), and random forest classifier (RFC). Moreover, the performance evaluation of the proposed MSVM with different classifiers is analyzed with and without MSSAE which is shown in Table 1. Figs. 4 and 5 show the graphical comparison of the MSVM with MSSAE with different classifiers where Fig. 3 is for classifiers without MSSAE and Fig. 4 is for classifiers with MSSAE. From the analysis, it is concluded that the MSVM with MSSAE achieves higher classification accuracy of 96.73% than the MSVM without MSSAE. Moreover, the MSVM with MSSAE provides better performance than the other classifiers such as NN, KNN, and RFC with and without MSSAE. For example, the MSVM with MSSAE achieves 96.73%, whereas the MSSAE with NN obtains 89.05%, KNN obtains 93.76% and RFC obtains 95.22%. The MSVM with the capacity of handling high dimension spaces and optimal feature dimension reduction using MSSAE is used to improve the classification of real-time kannada character.

The performance analysis of HFE-MSSAE with different feature dimension reduction approaches is provided in Table 2. Fig. 6 shows a graphical comparison of MSSAE with different feature dimension reduction approaches. The dimension reduction approaches considered for the evaluation are principal component analysis (PCA),

Table 1. Performance analysis of HFE-MSSAE with different classifiers

Feature dimension reduction	Classifiers	Accuracy (%)	Precision (%)	Recall (%)	CSI (%)	F-measure (%)
Without MSSAE	NN	87.56	85.63	88.37	84.25	86.42
	KNN	90.08	88.11	91.07	89.29	90.76
	RFC	92.64	90.39	93.18	92.11	92.00
	MSVM	94.01	91.17	93.44	92.16	92.55
With MSSAE	NN	89.05	90.17	91.34	92.33	88.37
	KNN	93.76	92.47	92.72	94.11	90.18
	RFC	95.22	94.05	94.11	95.34	92.90
	MSVM	96.73	95.18	97.30	96.97	94.08

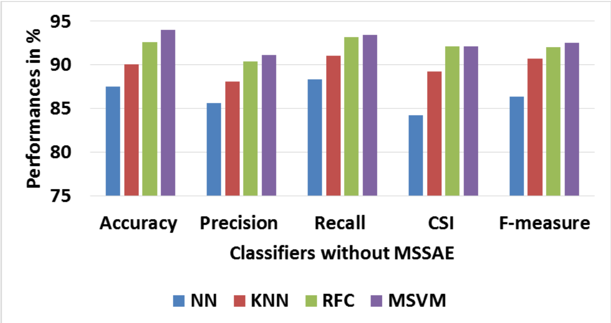

Figure. 4 Graphical comparison of the classifiers without MSSAE

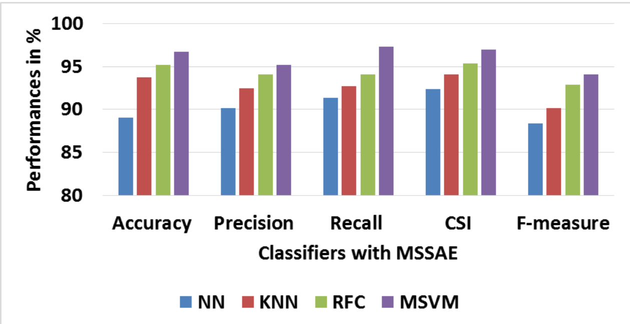

Figure. 5 Graphical comparison of the classifiers with MSSAE

independent component analysis (ICA) and reconstruction independent component analysis (RICA). From Table 2 and Fig. 6, it is decided that the proposed MSSAE offers advanced classification accuracy of 96.73% to the PCA, ICA and RICA approaches. The HFE-MSSAE provides higher accuracy because of the enhanced reconstruction ability obtained using Morlet wavelet activation function.

4.2. Comparative analysis

This section shows the comparative analysis of the HFE-MSSAE for real-time English and Kannada character recognition. The dataset used for the English and Kannada HCR is the Chars74k dataset. The comparative analysis between the HF-LSTM [19], EHO-LSTM [21] and HFE-MSSAE for Kannada characters is shown in Table 3 and Fig. 7,

Table 2. Performance analysis of HFE-MSSAE with different feature dimension reduction approaches

Feature dimension reduction	Accuracy (%)	Precision (%)	Recall (%)	CSI (%)	F-measure (%)
PCA	91.25	90.33	92.37	91.51	89.07
ICA	93.22	91.17	93.45	92.18	90.26
RICA	95.02	94.11	95.38	94.13	92.67
MSSAE	96.73	95.18	97.30	96.97	94.08

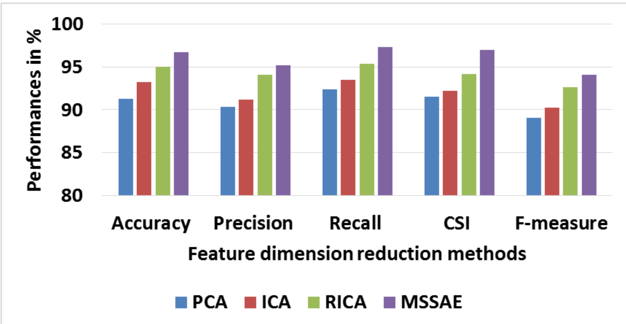

Figure. 6 Graphical comparison of MSSAE with different feature dimension reduction approaches

where NA refers the values which are not available. Similarly, the comparative analysis between the CNN [18], HF-LSTM [19], EHO-LSTM [21] and HFE-MSSAE for English characters is shown in Table 4 and Fig. 8. From the analysis, it is known that the HFE-MSSAE provides better performance than the CNN [18], HF-LSTM [19] and EHO-LSTM [21]. For example, the accuracy of the HFE-MSSAE for Kannada is 96.73% whereas the accuracy of HF-LSTM [19] is 95.06% and EHO-LSTM [21] is 96.67%. The CNN [18] and HF-LSTM [19] obtains less accuracy because it processes all the extracted features during the classification. The EHO-LSTM [21] obtains slightly lesser accuracy, because of inappropriate fitness function considered during

Table 3. Comparative analysis of HFE-MSSAE for Kannada characters

Performance measures	Methods		
	HF-LSTM [19]	EHO-LSTM [21]	HFE-MSSAE
Accuracy (%)	95.06	96.67	96.73
Precision (%)	92.14	94.92	95.18
Recall (%)	95.73	95.93	97.30
CSI (%)	93.05	NA	96.97
F-measure (%)	93.09	96.73	94.08

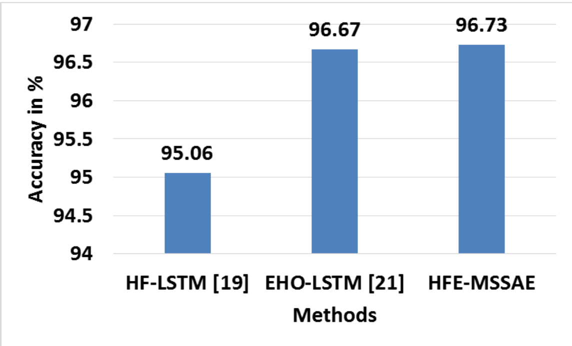

Figure. 7 Accuracy comparison of Kannada characters for HFE-MSSAE

Table 4. Comparative analysis of HFE-MSSAE for English characters

Performance measures	Methods			
	CNN [18]	HF-LSTM [19]	EHO-LSTM [21]	HFE-MSSAE
Accuracy (%)	76.30	98.78	96.66	99.18
Precision (%)	NA	96.05	94.38	98.55
Recall (%)	NA	95.32	96.90	99.71
CSI (%)	NA	95.89	NA	99.09
F-measure (%)	NA	97.03	96.75	99.16

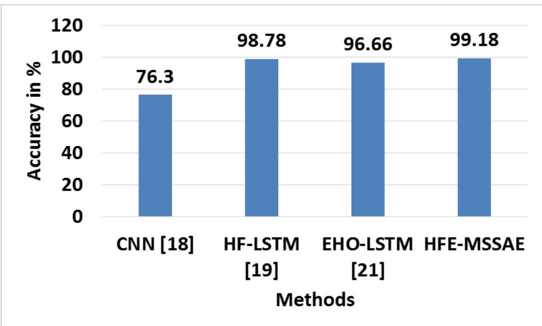

Figure. 8 Accuracy comparison of English characters for HFE-MSSAE

feature selection. Therefore, an enhanced reconstruction ability of MSSAE is used for an effective feature dimension reduction that used to improve the recognition accuracy of both the English and Kannada characters.

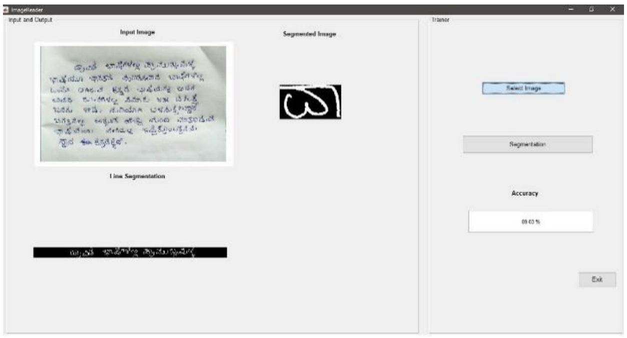

Figure 9. Screenshot for online Kannada HCR

Table 5. Case study for online Kannada HCR

Performance measures	HFE-MSSAE
Accuracy (%)	89.03
Precision (%)	90.70
Recall (%)	90.37
CSI (%)	88.59
F-measure (%)	91.52

4.3 Case study

In this case study, an input of Kannada text is given by using the MATLAB GUI. Since, the input is given on-spot for classification it is termed an online HCR. For verification, the screenshot of MATLAB GUI while giving the input is shown in Fig. 9. The case study for online Kannada HCR is given in Table 5. The proposed HFE-MSSAE provides better performance for the online Kannada HCR whereas the accuracy for the Kannada text given through the MATLAB GUI is 89.03%.

5. Conclusion

This paper proposed the HFE along with the MSSAE based feature dimensionality reduction for improving the classification of English and Kannada characters. The hybrid features include the features from HOG, GLCM, DWT, and skeleton features are used to provide an effective shape and textual features. Reaching higher accuracy in HCR using a huge amount of features is challenging, so the MSSAE with morlet wavelet activation function is used to minimize the feature dimension from the overall feature set. The morlet wavelet activation function is used to achieve improved reconstruction ability and minimize the sparsity. Therefore, the optimal features obtained from the MSSAE are used to improve the classification from the MSVM. From the performance analysis, it is concluded that the HFE-MSSAE provides better performance in both the offline and online Kannada HCR as well as in

Received: December 11, 2022. Revised: February 10, 2023.

English HCR. Specifically, the HFE-MSSAE outperforms well in English and Kannada HCR when compared to the CNN, HF-LSTM and EHO-LSTM. The classification accuracy of the HFE-MSSAE method for Kannada is 96.73% which is high than the HF-LSTM and EHO-LSTM.

Notations

Parameter	Description
I	Input image
I'	Preprocessed image
T_0	Otsu's thresholding value
p and q	Locations of row and column for image
G_p	Horizontal gradient of the pixel
G_q	Vertical gradient of the pixel
$G(p,q)$	Magnitude of the gradient
$\theta(p,q)$	Orientation of the gradient
$a(I')$	Component attribute to the preprocessed image
WS	Wavelet scale
TF	Translation factor
$h(I')$	High pass filter coefficients
$g(I')$	Low pass filter coefficients
SSM	Skeleton strength map
SV	Space vector
t_s	Diffusion time
x	Feature set
HOG	Features extracted from HOG
$GLCM$	Features extracted from GLCM
DWT	Features extracted from DWT
$\hat{\rho}$	Average activation value
z	Output layer
λ	Parameter utilized for controlling the regularization strength
$\psi(x,z)$	Error among the inputs and reconstructed ones
$\Omega_{weights}$	Weight attenuation term
u	Number of the hidden layers
v	Input layer units
s	Hidden layer units
W	Weight matrix
α	Function of Morlet wavelet activation
k	Amplitude
f_b	Bandwidth
f_c	Central frequency
T	Time
ρ	Sparsity parameter
n	Amount of training samples
β	Parameter utilized for controlling the sparse regularization
E	Sparse regularization term
KL	Kullback-Leibler divergence
$w_i^{(1)}$	Weight matrix row at 1st layer
$b_i^{(1)}$	Bias vector entry
TP	True positive
TN	True negative
FP	False positive
FN	False negative

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

The paper background work, conceptualization, methodology, dataset collection, implementation, result analysis and comparison, preparing and editing draft, visualization have been done by first. The supervision, review of work and project administration, have been done by second author.

References

[1] I. Khandokar, M. Hasan, F. Ernawan, S. Islam, and M. N. Kabir, "Handwritten character recognition using convolutional neural network", *In Journal of Physics: Conference Series*, Vol. 1918, No. 4, pp. 042152, 2021.

[2] M. Eltay, A. Zidouri, and I. Ahmad, "Exploring deep learning approaches to recognize handwritten arabic texts", *IEEE Access*, Vol. 8, pp. 89882-89898, 2020.

[3] S. Ahlawat, A. Choudhary, A. Nayyar, S. Singh, and B. Yoon, "Improved handwritten digit recognition using convolutional neural networks (CNN)", *Sensors*, Vol. 20, No. 12, pp. 3344, 2020.

[4] R. Geetha, T. Thilagam, and T. Padmavathy, "Effective offline handwritten text recognition model based on a sequence-to-sequence approach with CNN–RNN networks", *Neural Computing and Applications*, Vol. 33, No. 17, pp. 10923-10934, 2021.

[5] W. Simayi, M. Ibrayim, and A. Hamdulla, "Character type based online handwritten Uyghur word recognition using recurrent neural network", *Wireless Networks*, pp. 1-11, 2021.

[6] N. Altwaijry and I. A. Turaiki, "Arabic handwriting recognition system using convolutional neural network", *Neural Computing and Applications*, Vol. 33l, No. 7, pp. 2249-2261, 2021.

[7] F. J. J. Joseph, "Effect of supervised learning methodologies in offline handwritten Thai character recognition", *International Journal of Information Technology*, Vol. 12, No. 1, pp. 57-64, 2020.

[8] B. R. Kavitha and C. Srimathi, "Benchmarking on offline Handwritten Tamil Character Recognition using convolutional neural

networks", *Journal of King Saud University-Computer and Information Sciences*, 2019.
[9] R. Ptucha, F. P. Such, S. Pillai, F. Brockler, V. Singh, and P. Hutkowski, "Intelligent character recognition using fully convolutional neural networks", *Pattern Recognition*, Vol. 88, pp. 604-613, 2019.
[10] H. M. Balaha, H. A. Ali, E. K. Youssef, A. E. Elsayed, R. A. Samak, M. S. Abdelhaleem, M. M. Tolba, M. R. Shehata, M. R. A. Mahmoud, M. M. Abdelhameed, and M. M. Mohammed, "Recognizing arabic handwritten characters using deep learning and genetic algorithms", *Multimedia Tools and Applications*, Vol. 80, No 21, pp. 32473-32509, 2021.
[11] C. Boufenar, A. Kerboua, and M. Batouche, "Investigation on deep learning for off-line handwritten Arabic character recognition", *Cognitive Systems Research*, Vol. 50, pp. 180-195, 2018.
[12] N. Modhej, A. Bastanfard, M. Teshnehlab, and S. Raiesdana, "Pattern separation network based on the hippocampus activity for handwritten recognition", *IEEE Access*, Vol. 8, pp. 212803-212817, 2020.
[13] H. Kusetogullari, A. Yavariabdi, J. Hall, and N. Lavesson, "DIGITNET: a deep handwritten digit detection and recognition method using a new historical handwritten digit dataset", *Big Data Research*, Vol. 23, p. 100182, 2021.
[14] P. Sahare and S. B. Dhok, "Multilingual character segmentation and recognition schemes for Indian document images", *IEEE Access*, Vol. 6, pp. 10603-10617, 2018.
[15] N. D. Cilia, C. D. Stefano, F. Fontanella, and A. S. D. Freca, "A ranking-based feature selection approach for handwritten character recognition", *Pattern Recognition Letters*, Vol. 121, pp. 77-86, 2019.
[16] S. Albahli, M. Nawaz, A. Javed, and A. Irtaza, "An improved faster-RCNN model for handwritten character recognition", *Arabian Journal for Science and Engineering*, Vol. 46, No 9, pp. 8509-8523, 2021.
[17] A. A. Chandio, M. Asikuzzaman, and M. R. Pickering, "Cursive character recognition in natural scene images using a multilevel convolutional neural network fusion", *IEEE Access*, Vol. 8, pp. 109054-109070, 2020.
[18] R. Harizi, R. Walha, F. Drira, and M. Zaied, "Convolutional neural network with joint stepwise character/word modeling based system for scene text recognition", *Multimedia Tools and Applications*, Vol. 81, No. 3, pp. 3091-3106, 2022.
[19] A. Kathigi and K. H. Kariputtaiah, "Handwritten character recognition using skewed line segmentation method and long short term memory network", *International Journal of System Assurance Engineering and Management*, pp. 1-13, 2021.
[20] Z. Liu, X. Pan, and Y. Peng, "Character recognition algorithm based on fusion probability model and deep learning", *The Computer Journal*, Vol. 64, No. 11, pp. 1705-1714, 2021.
[21] N. S. Guptha, V. Balamurugan, G. Megharaj, K. N. A. Sattar, and J. D. Rose, "Cross lingual handwritten character recognition using long short term memory network with aid of elephant herding optimization algorithm", *Pattern Recognition Letters*, Vol. 159, pp. 16-22, 2022.
[22] T. E. D. Campos, B. R. Babu, and M. Varma "Character recognition in natural images", In: *Proc of the 4th International Conference on Computer Vision Theory and Applications, VISAPP*, pp. 273–280, 2009.
[23] J. S. Wang and X. D. Ren, "GLCM based extraction of flame image texture features and KPCA-GLVQ recognition method for rotary kiln combustion working conditions", *International Journal of Automation and Computing*, Vol. 11, No. 1, pp. 72-77, 2014.
[24] N. V. Shree, and T. N. R. Kumar, "Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network", *Brain Informatics*, Vol. 5, No. 1, pp. 23-30, 2018.
[25] C. Direkoglu, R. Dahyot, and M. Manzke, "On using anisotropic diffusion for skeleton extraction", *International Journal of Computer Vision*, Vol. 100, No. 2, pp. 170-189, 2012.
[26] P. Gangsar and R. Tiwari, "Taxonomy of induction-motor mechanical-fault based on time-domain vibration signals by multiclass SVM classifiers", *Intelligent Industrial Systems*, Vol. 2, No. 3, pp. 269-281, 2016.