Arab J Sci Eng (2018) 43:4315–4326
https://doi.org/10.1007/s13369-017-2945-9

RESEARCH ARTICLE - SPECIAL ISSUE - COMPUTER ENGINEERING AND COMPUTER SCIENCE

Reduced Scattering Representation for Malayalam Character Recognition

K. Manjusha[1] · M. Anand Kumar[1] · K. P. Soman[1]

Received: 9 May 2017 / Accepted: 24 October 2017 / Published online: 23 November 2017

Abstract Scattering convolution network generates stable feature representation by applying a sequence wavelet decomposition operation on input signals. The feature representation in higher layers of the network builds a large-dimensional feature vector, which is often undesirable in most of the applications. Dimension reduction techniques can be applied on these higher-dimensional feature descriptors to produce an informative representation. In this paper, singular value decomposition is applied to the higher-layer scattering representation to generate informative feature descriptors. The effectiveness of the reduced scattering representation is evaluated on Malayalam printed and handwritten character recognition using support vector machine classifier. The reduced scattering representation improves the recognition performance when combining with lower-layer scattering network features.

Keywords Scattering convolutional network · Singular value decomposition · Malayalam language · Character recognition · Feature extraction · Dimensionality reduction

1 Introduction

Variability within classes is the major challenge in most of the pattern recognition applications. In character recognition applications, transformations and deformations are the main reasons for variability inside participating character classes. Often these variabilities greatly affect the recognition performance. Feature descriptors extracted from input patterns for classification usually affect the recognition performance [1]. If the feature descriptor representing character images are invariant to transformations and are able to linearize small deformations, then these variabilities can be effectively handled during the recognition process.

Feature extraction algorithms which have the ability to produce invariants that are stable to deformations can be utilized in building robust recognition systems. Deep convolutional neural network (CNN)-based architectures are the state of the art in character recognition applications due to their capability to build invariant and stable representations of the character images despite of the large-scale variabilities present in the images [2]. CNN makes use of trainable filter banks, nonlinear activations and pooling operators to build invariant image representations. Inside CNN, self-learned features are generated through trainable filter banks. Scattering convolutional network is based on scattering transform which computes invariant image representation by cascading wavelet transform with modulus and averaging functions [3,4]. Wavelet transforms are localized feature descriptors which are stable to deformations, but are covariant with translation. With the support of nonlinear modulus operator and averaging functions, wavelet transform can be transformed to robust feature descriptors. Compared to CNN, a scattering convolutional network makes use of predefined wavelets to create filter banks. A series of scattering transforms are applied on input image with the help of wavelet filter banks, to build a scattering convolutional network, and each node in the network provides invariant feature representation. On the higher layers of scattering convolution network, the invariant represen-

✉ K. Manjusha
k_manjusha@cb.amrita.edu

M. Anand Kumar
m_anandkumar@cb.amrita.edu

K. P. Soman
kp_soman@amrita.edu

1 Center for Computational Engineering & Networking, Amrita University, Coimbatore, India

tations obtained in different scales and angles are highly correlated. Scattering coefficients extracted from scattering network nodes having same path length have very strong correlation [3]. Discrete cosine transform (DCT) across scales and orientation can de-correlate the scattering coefficients [3]. In this paper, singular value decomposition (SVD) technique utilized for de-correlating scattering coefficients.

Malayalam is an official Indian language script widely used in Kerala, the southwestern state of India [5]. The research on Malayalam character recognition dates behind the early 1990s. Like other Indian languages, Malayalam language lipi (script) is syllabic in nature. The basic graphical unit used to represent different syllables in the Malayalam language is usually termed as *Akshara*. The basic *Akshara* (character) set of Malayalam language contains 16 vowels and 36 consonants. Besides this basic character set, the Malayalam script contains large number of compound characters and special symbols [6]. Usually, the recognition module inside optical character recognition (OCR) system relies either on the individual character or on the whole word extracted from the document image. In the Malayalam language script, both character- and word-based approaches can be applied as the script is written non-cursively using characters inside the valid *Akshara* set. The present work focuses on the character-based recognition of Malayalam documents. In terms of character recognition of Malayalam language documents, the main challenge is the strong structural similarity among different character shapes (glyph) and the presence of a large number of character classes. In addition to these challenges, the concurrent usage of both old and new scripts and the non-availability of the standard dictionary make the document recognition in the Malayalam language script further complicated [6]. In order to cope up with all these problems, robust character recognizers have to be utilized for the Malayalam language script. Various feature descriptors have applied for character recognition in other Indian language scripts [7–11]. The earlier works in Malayalam character recognition [12–20] mainly concentrated on wavelet-based or structural features classified with the help of support vector machines (SVM), neural network classifiers or quadratic discrimination functions. Features created from wavelet transform of projection histogram [15], zero crossings in wavelet transform [17], wavelet features at different resolution scales [13] and wavelet energy features [12] could excel in Malayalam character recognition as wavelets have the capability to generate robust localized feature descriptors. Multistage classification approaches have employed in Indian scripts to handle similarly shaped character classes [21–23] by making use of script specific features. In this paper, the translation invariant features obtained from cascaded wavelet decomposition are utilized for Malayalam machine-printed and hand-printed character recognition.

Section 2.1 describes about scattering convolutional network. Section 2.2 explains the SVD-based proposed approach for feature extraction. Finally, Sect. 4 describes about the experiments and results obtained on Malayalam character recognition using proposed scattering-based feature descriptor.

2 Feature Extraction and Classification Methods

This section describes about the algorithms that are utilized in implementing the character recognition system considered in the experiments.

2.1 Scattering Network

Scattering network [24] is similar to a deep convolutional network in which invariant image descriptors are computed sequentially over multiple layers. Instead of trainable filters inside CNN, scattering network contains complex wavelet filters over spatial and angular variables. Consider a signal $x(u)$ where $u \in \mathbb{R}^2$ and an integer J, the spatial scale of scattering transform. If r belongs to the rotation group of $\mathbb{R}^2$ and $j < J$, then the 2D directional wavelets with $\lambda = 2^j r$ can be defined as

$$\psi_\lambda(u) = 2^{-2j}\psi\left(2^{-j}r^{-1}u\right) \tag{1}$$

Convolving $x(u)$ with these directional wavelets produces wavelet coefficients $\{x * \psi_\lambda\}$ which are not translation invariant. Applying point-wise modulus operator on these wavelet coefficients and then applying averaging functions help to build invariants. The resulting translation invariant coefficients are represented in Eq. 2:

$$\|x * \psi_\lambda\|_1 = \int |x * \psi_\lambda(u)|\mathrm{d}u \tag{2}$$

The integration in Eq. 2 removes all nonzero frequencies, and the information loss happening through this integration can be recovered by applying second wavelet transform on $|x * \psi_\lambda|$ using ψ_{λ_i}. The translation invariants created from those coefficients can be represented as shown in Eq. 3

$$\left\||x * \psi_\lambda(u)| * \psi_{\lambda_i}\right\|_1 = \int \left||x * \psi_\lambda(u)| * \psi_{\lambda_i}\right|\mathrm{d}u \tag{3}$$

A large set of translation invariants are computed by further iterating on the wavelet transform and modulus operators. Thus, the scattering coefficients computed through a cascade of convolution and modulus operators through different paths

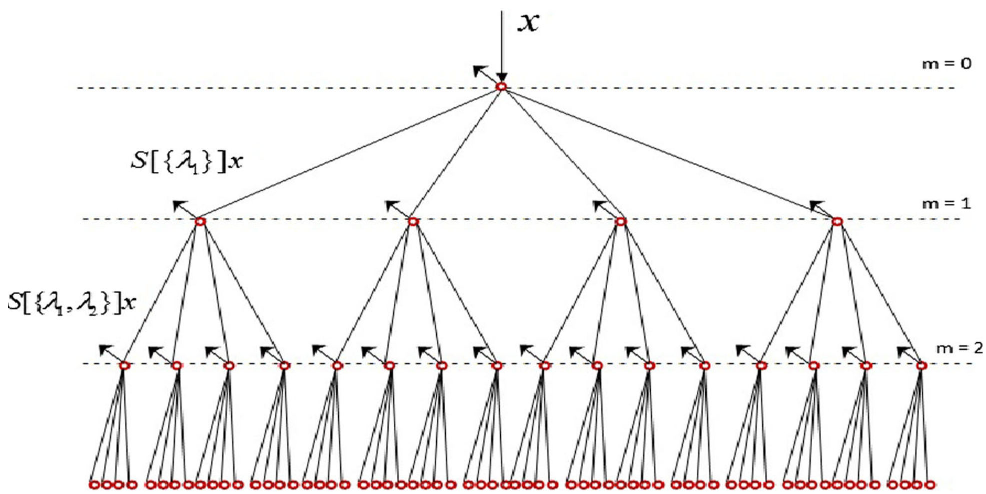


Fig. 1 Scattering convolution network [26]

possible give rise to a convolutional network. In classification tasks, calculating invariants smaller than a predefined scale 2^J is preferred. This is achieved through windowed scattering transform in which the localized invariant feature descriptors are computed through a low pass average at scale 2^J using $\phi_{2^J}(u) = 2^{-2J}\phi(2^{-J}u)$. The scattering coefficient S, calculated through the path $p = \{\lambda_1, \lambda_2 \ldots \lambda_m\}$, is shown in Eq. 4 with $S[\emptyset]x = x * \phi_{2^J}$.

$$S[p]x = \left|\left|\left|\left|x * \psi_{\lambda_1}\right| * \psi_{\lambda_2}\right| \cdots\right| * \psi_{\lambda_m}\right| * \phi_{2^J}(u) \tag{4}$$

$S[p]x$ is of order m and computed at the layer m of scattering convolutional network. In experiments, ψ is Morlet wavelet, and the ϕ is chosen to be Gaussian. The rotation group is in between angles $[0,\pi)$. In contrast to CNN, the scattering coefficients are produced at each layer rather than the last layer. The first-order coefficients in the scattering network is equivalent to SIFT [25] coefficients, if the wavelet functions are chosen appropriately [3]. The visualization of scattering network with two layers is shown in Fig. 1. The scattering coefficients are calculated only among path $\{\lambda_1, \lambda_2\}$, where $2^{j_2} < 2^{j_1}$. $\left|x * \psi_{\lambda_1}\right| * \psi_{\lambda_2}$ is negligible on cases where $2^{j_2} \geqslant 2^{j_1}$ [24].

For an image of N pixels, the scattering network produces $2^{-2J}NK^q\binom{J}{q}$ scattering coefficients in layer q, where K is the number of angles in rotation group of directional wavelets. If the number of layers in the scattering network is m, then the total number of scattering coefficients in scattering vector is $2^{-2J}N\sum_{q=0}^{m}K^q\binom{J}{q}$. The total number of scattering coefficients increases drastically in higher layers. Besides this, there is a great correlation among higher-layer scattering coefficients calculated at paths having the same length. Dimensionality reduction techniques are capable of removing the correlation among higher-layer scattering coefficients and capturing informative feature descriptors.

2.2 Singular Value Decomposition (SVD)-Based Dimensionality Reduction

SVD is a matrix factorization technique widely used in most of the data processing applications for dimensionality reduction. SVD provides relevant information contained in the complete matrix in decreased dimensions. SVD orders the information contained in the matrix so that the principal components capture most of the variation among data while retaining the total variation same. SVD provides the best low-rank approximation of a matrix [27]. Let A be a $p \times q$ matrix. Through SVD, A is factorized as shown in Eq. 5. U represents $p \times p$ orthogonal matrix with its columns the eigen vectors of AA^{T} and V represents $q \times q$ orthogonal matrix and its columns represent eigen vectors of $A^{\mathrm{T}}A$. Σ is a $p \times q$ diagonal matrix with decreasingly ordered singular values in its diagonal.

$$A = U\Sigma V^{\mathrm{T}} \tag{5}$$

Applying SVD to the large-dimensional higher-layer scattering coefficients removes correlation and captures informative feature descriptors. In this paper, we consider features from layer $m \geq 2$, as higher-layer features and dimensionality reduction is applied only on those features. Figure 2 shows the process flow of the proposed feature extraction method on scattering network having two layers. For each character image, the scattering representation on each layer extracted and on the second-layer feature ($m = 2$), SVD-based dimension reduction applied. Extracting the low-level representation of the higher dimension second-layer scattering features can be performed utilizing very few columns inside

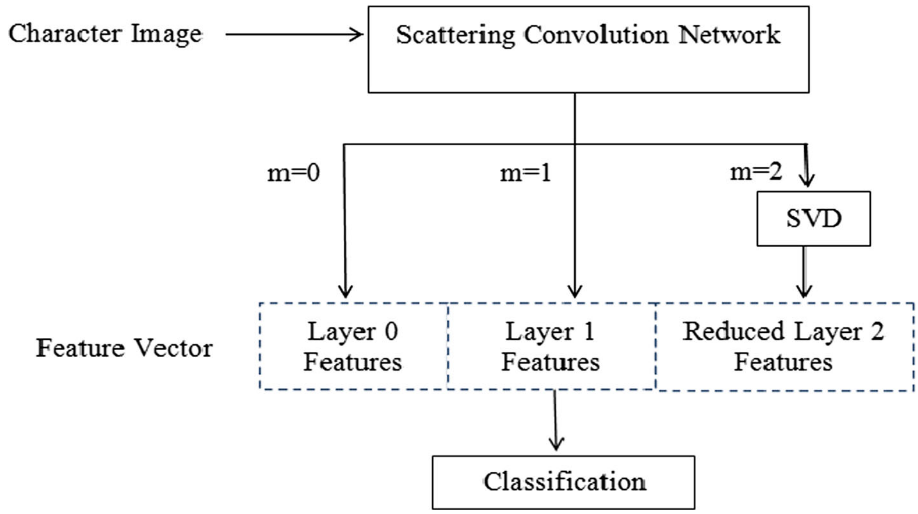


Fig. 2 Feature extraction from scattering convolution network

SVD orthogonal matrices. The proposed feature descriptor obtained through appending zeroth- and first-layer features with dimension reduced second-layer features.

In this paper, SVD approach described in [28] is utilized for dimensionality reduction of higher-layer scattering coefficients. The proposed feature extraction method on SVM classifier can be summarized algorithmically as shown in Table 1. The training phase needs two parameters, minimum and maximum number of SVD features ($minNs$ and $maxNs$) desired for feature creation. $Scat_{image}$ in Table 1 represents the scattering features obtained from the scattering convolution network for input image, and $Scat_{image}(i)$ represents vectorized form of the features obtained from nodes in the ith layer. The algorithm in Table 1 is for scattering network having two layers but can be extended to scattering network having any number of layers. The training feature matrix $Fmat$ is the vectorized second-layer scattering descriptors of training images appended column-wise. Applying SVD on this training feature matrix captures the latent relations in between data present in the matrix in lower dimensions. k-fold cross-validation is applied on SVD features to decide optimal size of SVD features, $optSVD$. In this paper, fivefold cross-validation is performed to estimate $optSVD$. The proposed feature vector, $ProScat$, is obtained by appending lower-layer features $Scat_{image}(0)$, $Scat_{image}(1)$ with the reduced second-layer feature representation.

SVD requires O(N^3) operations to calculate the singular bases for a square matrix of dimension N. Scattering feature representation over scattering network can be calculated with O($NlogN$) [3]. Even though SVD is computationally expensive, in the proposed approach SVD calculation required only in training phase. During testing, the reduced scattering features computation requires only dot product calculation between already calculated SVD bases and the second-layer scattering features extracted from test character images. The test time complexity of the proposed approach is O($NlogN$). The heavy computation required in training phase due to SVD calculation can be accelerated by using parallel processing architectures. As training is only a one time job, what really matters is the test phase time complexity.

Table 1 Algorithmic description of proposed approach with scattering network having two layers

Character Recognition based on Reduced Scattering Features

Training Phase:
Input : Training character image dataset, $train$; Minimum SVD feature vector size $minNs$; Maximum SVD feature vector size $maxNs$
Algorithm :
$Fmat = [\,]$
foreach $image \in train$
  Calculate scattering features $Scat_{image}$, for $m = 0, 1, 2$
  $Fmat = [Fmat \;\; Scat_{image}(2)]$
Compute U, Σ and V by applying SVD on $Fmat$
Iterate $ns = minNs$ till $maxNs$
  Calculate training SVD features by taking ns rows from Σ and V.
  Compute cross-validation accuracy $cValAcc$ for ns features on SVM classifier
Estimate $optSVD$ as ns having maximum $cValAcc$
foreach $image \in train$,
  Compute proposed feature representation,
  $ProScat = [Scat_{image}(0); Scat_{image}(1); <U(1 : optSVD), Scat_{image}(2)>]$
Build SVM Classifier Model, $ModelSVM$ on $ProScat$

Testing Phase:
Input : Testing character image dataset, $test$.
Algorithm :
foreach $image \in test$
  Calculate scattering features $Scat_{image}$ for $m = 0, 1, 2$
  Compute proposed feature representation as, $ProScat = [Scat_{image}(0); Scat_{image}(1); <U(1 : optSVD), Scat_{image}(2)>]$
  Evaluate $ProScat$ on $ModelSVM$ to get target character class

2.3 Support Vector Machine (SVM)

In order to classify the scattering features, SVM [29] classification algorithm utilized. SVM is a kernel-based discriminative classification algorithm for machine learning. Basic SVM is a binary classifier which searches for an optimal hyperplane to maximally separate the training instances such that instances of two classes lie on the two opposite sides of the hyperplane. In order to achieve the good generalization ability among test instances, the hyperplane has to be chosen such that the maximum margin value is achieved. Consider a binary classification task with M training instances, $\{(x_i, y_i), x_i \in \mathbb{R}^d, y_i \in \{+1, -1\}, i = 1, 2, \ldots M\}$. If the training instances are linearly separable in $\mathbb{R}^d$, then there exists a hyperplane such that,

$$\left\{\begin{array}{l} \forall y_i = +1,\ \omega.x_i + b > 0 \\ \forall y_i = -1,\ \omega.x_i + b < 0 \end{array}\right\}, \quad \omega \in \mathbb{R}^d, b \in \mathbb{R} \tag{6}$$

During SVM training, the hyperplane parameters are estimated by minimizing the optimization function in Eq. 7

$$\left\{\begin{array}{l} \min\limits_{\omega,b} \frac{1}{2}\|\omega\|^2 \\ \text{s.t. } \forall i,\ \ y_i(\omega.x_i + b) \geqslant 1 \end{array}\right\} \tag{7}$$

On introducing Lagrange multipliers, an equivalent dual program of Eq. 7 can be formulated and is shown in Eq. 8.

$$\left\{\begin{array}{l} \max\limits_{\alpha_i} \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i . x_j \\ \text{s.t. } \forall i,\ \alpha_i \geqslant 0,\ \sum_i \alpha_i y_i = 0 \end{array}\right\} \tag{8}$$

For nonlinearly separable instances, SVM classifier generalized by applying kernel trick [29]. The scalar product $x_i.x_j$ in Eq. 8, replaced with a positive-definite kernel function $K(x_i, x_j)$ for creating nonlinear decision surface. Through kernel functions, the separating hyperplane in higher-dimensional space is estimated using SVM. In this paper, the Gaussian radial basis function (RBF), $K(x, y) = \exp(-\|x - y\|^2/2\sigma^2)$ utilized for creating the nonlinear SVM classifier model.

The basic SVM classifier concept is extended to the multi-class classification context with the support of several binary SVM classifiers. Single multi-class classification problem is reduced to multiple binary classification problems [30]. Two common methods for combining binary SVM classifiers for large classification problems are one-versus-all and one-versus-one. In one-versus-all strategy, binary SVM classifiers trained with instances from one class as positive and other classes as negative. In one-versus-one strategy, binary classifiers are trained between each pair of distinct classes. For experiments, LibSVM [31] toolkit employed for implementing multi-class SVM classifier. LibSVM makes use of one-versus-one approach and the final class decision is taken through voting strategy.

3 Character Image Databases

This section describes the image databases employed for experimenting character recognition systems.

3.1 Malayalam Character Image Databases

Even though many isolated works have been reported in the literature toward Malayalam character recognition, no standard character image database is available publicly for the Malayalam language script. Two character image databases, printed (MAL_PrintedDB[1]) and handwritten (MAL_HandwrittenDB[2]), are created for implementing Malayalam character recognizer.

3.1.1 Malayalam Printed Character Image Database (MAL_PrintedDB)

For printed Malayalam character recognition, the character images are collected by scanning Malayalam literature books and by creating synthetic character images from different Malayalam font styles. For segmenting the characters from scanned images, the active contour model [32] utilized. The Malayalam printed character image database, MAL_PrintedDB, contains 24,553 character images of 130 different character classes with each character image of dimension 32×32. Figure 3 shows sample character images taken from MAL_PrintedDB for representing five different Malayalam character classes. The first column in Fig. 3 represents the Malayalam vowel character '*a*,' the second column represents Malayalam consonant *'ka,'* and the third column represents Malayalam consonant *'tha.'* The fourth and fifth columns in Fig. 3 represent Malayalam combined characters *'kka'* and *'ddha,'* respectively. For the training process, 75% of images from each character class is considered and the rest is considered for the validation process. The test dataset of MAL_PrintedDB collected from 67 real-world document images taken from Malayalam magazines, literature books and newspaper articles. Three sample document images taken from those images are shown in Fig. 5. Figure 5a shows text paragraph segmented from scanned document image of a Malayalam literature work, and Fig. 5b represents

[1] Available from URL http://nlp.amrita.edu:8080/OCR/dataset/Printed.zip.

[2] Available from URL http://nlp.amrita.edu:8080/OCR/dataset/Handwritten.zip.

Fig. 3 Sample images from MAL_PrintedDB

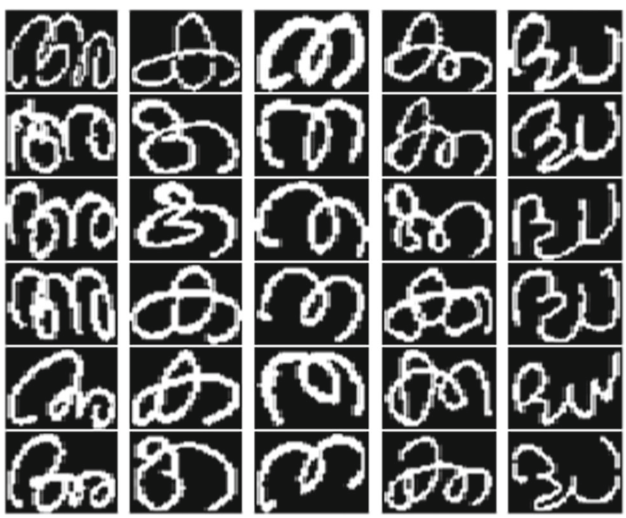


Fig. 4 Sample images from MAL_HandwrittenDB

part of a story in one of the Malayalam magazine. Figure 5c is taken from one of the Malayalam newspapers article. Among the segmented 22,712 character images in test dataset, 1316 images are labeled as ERROR, as they are representing character classes that are not present in MAL_PrintedDB and segmentation errors happened during dataset creation. Test dataset for MAL_PrintedDB created from real-world document images and has a varying count distribution among considered Malayalam character classes.

3.1.2 Malayalam Handwritten Character Image Database (MAL_HandwrittenDB)

In handwritten character recognition context, 85 Malayalam character classes which are most frequently used in writing are considered. MAL_HandwrittenDB created by collecting handwritten documents from 59 different persons. The character images from 46 persons considered for training and remaining considered for testing the recognition performance. Broken and largely distorted images are excluded from the experiments. Each image is binarized and resized to dimension 32×32. Training dataset of MAL_HandwrittenDB contains 18,874 character images, while the evaluation test dataset have 4095 character images. Handwritten character images representing the same five Malayalam character classes as that of Fig. 3 are shown in Fig. 4.

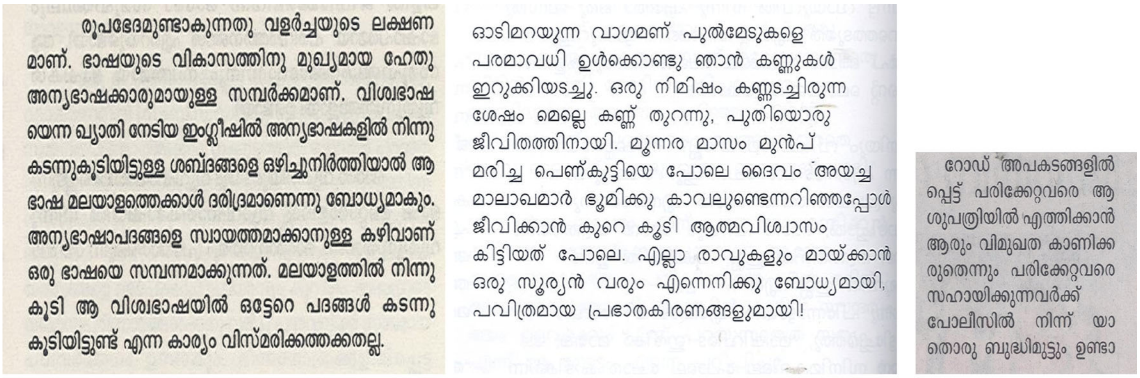
3.2 MNIST Handwritten Digit Database (MNIST)

MNIST [2] handwritten digit database is the most widely used benchmark database for character recognition experiments. MNIST is employed in experiments to confirm the effectiveness of reduced scattering representation. MNIST is a subset of NIST dataset. The handwritten digit images in MNIST dataset are in grayscale format with 28×28 dimension. Some sample handwritten digit images taken from MNIST are shown in Fig. 6. Different handwritten image samples of digit 1 to digit 5 are shown column-wise in Fig. 6. The dataset contains 60,000 training and 10,000 testing images. For conducting experiments, the grayscale images in MNIST dataset are binarized using Otsu [33] algorithm.

3.3 Telugu Handwritten Character Database (HP_Telugu)

The other handwritten image database used in this paper represents Telugu language script. Like Malayalam language, Telugu is also a south Indian language descended from the ancient Brahmi script. The Telugu script is derived from Bhattiporulu, while Malayalam is derived from Grantha alphabets. Both Malayalam and Telugu have alpha-syllabic nature, in which consonant–vowel sequences written as a single unit. Because of this nature, both scripts have a large number of character classes. The character classes in the respective scripts have strong structural similarity. It is very difficult for a person knowing Malayalam script to understand the Telugu script and vice versa as the character glyphs representing basic graphemes have entirely different structure. In character recognition scenario, the recognition problem in both scripts defined in the same way as challenging large-class classification problem. The Telugu dataset used in experiments is part of HP Labs India Indic Handwriting Datasets. It is the offline version of online isolated handwritten Telugu dataset[3] [34]. The HP Labs Telugu dataset contains

[3] Available from URL http://lipitk.sourceforge.net/datasets/teluguchardata.htm.



രൂപഭേദമുണ്ടാകുന്നതു വളർച്ചയുടെ ലക്ഷണ
മാണ്. ഭാഷയുടെ വികാസത്തിനു മുഖ്യമായ ഹേതു
അന്യഭാഷക്കാരുമായുള്ള സമ്പർക്കമാണ്. വിശ്വഭാഷ
യെന്ന ഖ്യാതി നേടിയ ഇംഗ്ലീഷിൽ അന്യഭാഷകളിൽ നിന്നു
കടന്നുകൂടിയിട്ടുള്ള ശബ്ദങ്ങളെ ഒഴിച്ചുനിർത്തിയാൽ ആ
ഭാഷ മലയാളത്തെക്കാൾ ദരിദ്രമാണെന്നു ബോധ്യമാകും.
അന്യഭാഷാപദങ്ങളെ സ്വായത്തമാക്കാനുള്ള കഴിവാണ്
ഒരു ഭാഷയെ സമ്പന്നമാക്കുന്നത്. മലയാളത്തിൽ നിന്നു
കൂടി ആ വിശ്വഭാഷയിൽ ഒട്ടേറെ പദങ്ങൾ കടന്നു
കൂടിയിട്ടുണ്ട് എന്ന കാര്യം വിസ്മരിക്കത്തക്കതല്ല.

(a)

ഓടിമറയുന്ന വാഗമണ് പുൽമേടുകളെ
പരമാവധി ഉൾക്കൊണ്ടു ഞാൻ കണ്ണുകൾ
ഇറുക്കിയടച്ചു. ഒരു നിമിഷം കണ്ണടച്ചിരുന്ന
ശേഷം മെല്ലെ കണ്ണ് തുറന്നു, പുതിയൊരു
ജീവിതത്തിനായി. മൂന്നര മാസം മുൻപ്
മരിച്ച പെണ്കുട്ടിയെ പോലെ ദൈവം അയച്ച
മാലാഖമാർ ഭൂമിക്കു കാവലുണ്ടെന്നറിഞ്ഞപ്പോൾ
ജീവിക്കാൻ കുറെ കൂടി ആത്മവിശ്വാസം
കിട്ടിയത് പോലെ. എല്ലാ രാവുകളും മായ്ക്കാൻ
ഒരു സൂര്യൻ വരും എന്നെനിക്കു ബോധ്യമായി,
പവിത്രമായ പ്രഭാതകിരണങ്ങളുമായി!

(b)

റോഡ് അപകടങ്ങളിൽ
പ്പെട്ട് പരിക്കേറ്റവരെ ആ
ശുപത്രിയിൽ എത്തിക്കാൻ
ആരും വിമുഖത കാണിക്ക
രുതെന്നും പരിക്കേറ്റവരെ
സഹായിക്കുന്നവർക്ക്
പോലീസിൽ നിന്ന് യാ
തൊരു ബുദ്ധിമുട്ടും ഉണ്ടാ

(c)

Fig. 5 Sample images taken from document images employed for creating testing dataset of MAL_PrintedDB

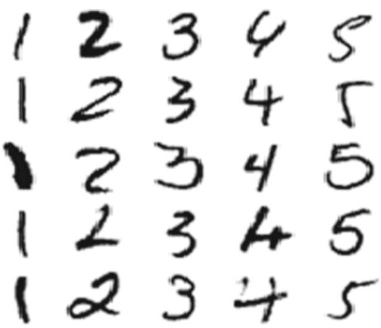


Fig. 6 Sample images from MNIST handwritten digit database

166 character classes collected from different native Telugu writers using pen–paper-based device to make data collection more natural. The dataset includes independent vowels, dependent vowel diacritics, consonants, consonant modifiers and some combined consonant–vowel symbols which are hard to segment. The Telugu database has 34,325 training and 10,892 testing images. Figure 7 shows handwritten character images taken from the Telugu dataset representing five different Telugu character classes. The first column in Fig. 7 shows Telugu vowel representing *'a,'* and the second column is Telugu consonant *'ka.'* The third column in Fig. 7 represents Telugu consonant *'da'*, and columns four and five show compound characters formed by combining Telugu consonant *'da'* with Telugu vowels *'i'* and *'ii,'* respectively. The images in the Telugu dataset are in binary format and have unequal size dimensions. In this paper, the images are resized to 150 × 150 dimension for conducting experiments.

4 Experimental Results and Discussion

This section describes the experiments conducted for evaluating scattering feature representations on character recognition process.

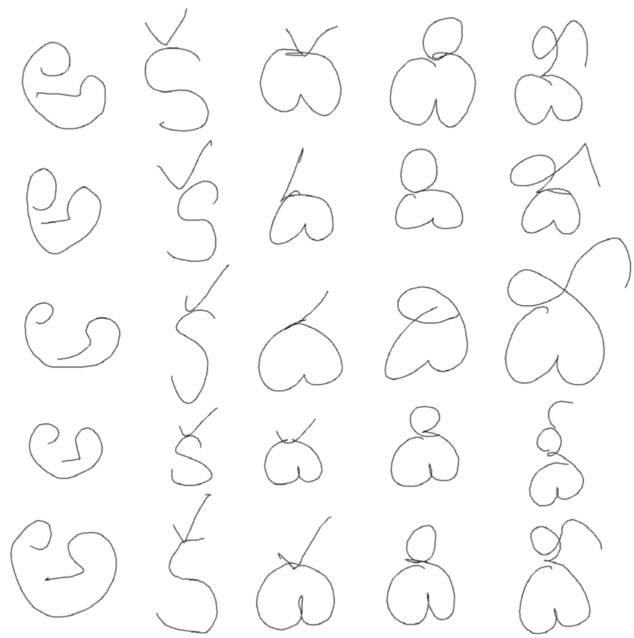


Fig. 7 Sample images from Telugu handwritten character database (HP_Telugu)

4.1 Creating Scattering Network

Each node in the scattering network produces invariant image representation, through wavelet decomposition and averaging functions. In order to implement a scattering network in our experiments, we have utilized ScatNet toolbox [35]. The scattering coefficients generated on each node in the scattering network can be visualized by plotting a disk covering image frequency support divided into sectors [3]. Figure 8 shows arrays of windowed scattering transform plotted on the disk covering image frequency support for layers $m = 1, 2$ of the scattering network using two different character images representing Malayalam letters *'i'* and *'ii,'* respectively.

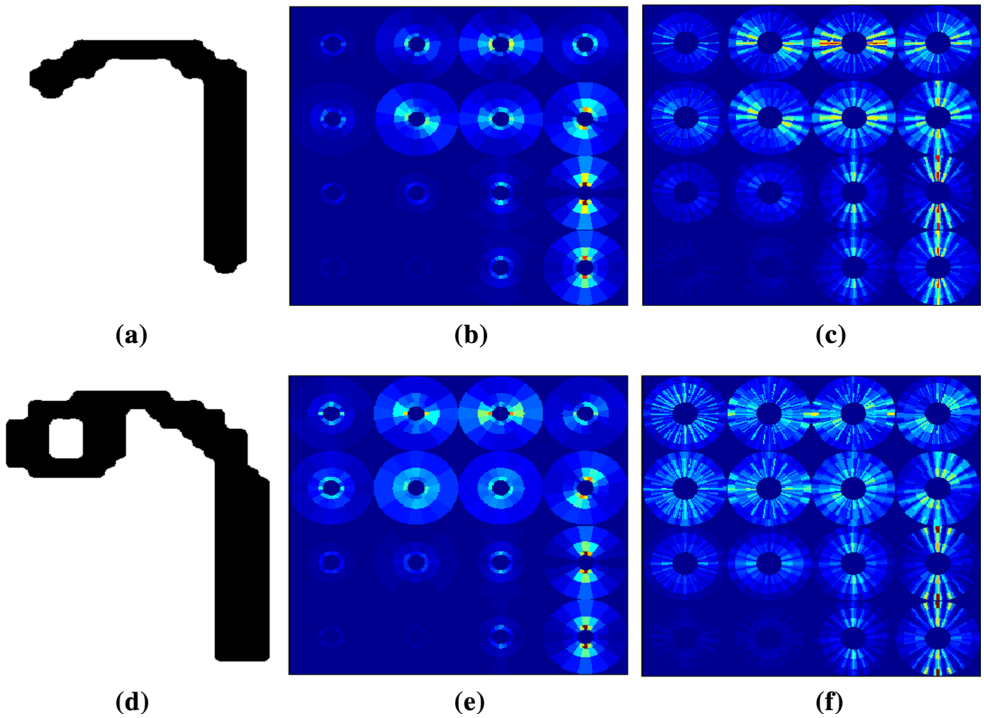


Fig. 8 Visualization of scattering coefficients with scale $J = 3$, extracted from Malayalam character images representing Malayalam vowel signs *i* and *ii*. **a** Malayalam letter *i*, **b** scattering coefficients in $m = 1$ of *i*, **c** scattering coefficients in $m = 2$ of *i*, **d** Malayalam letter *ii*, **e** scattering coefficients in $m = 1$ of *ii* and **f** scattering coefficients in $m = 2$ of *ii*

Even though the two Malayalam letters have very strong structural similarity, the first two disks in the first column representing scattering coefficients in $m = 1$ and $m = 2$ layers could capture discriminating features. On higher layers of scattering network, more discriminating features are generated. Incorporating all those higher-layer information for classification is difficult as it produces very large-dimensional feature vector. As the number of layers increases, the number of nodes in the resulting scattering network explodes and this give rise to a large number of image invariant representations. Table 2 shows the number of features generated at different layers of the network with $J = 3$ and $K = 8$. Bruna et al. [3] demonstrated that, on higher layers the invariant representations in nodes residing in the same layer have strong correlations. Direct application of higher-layer features in classifier reduces the recognition performance [26]. Dimension reduction techniques can be applied to the image scattering coefficients in higher layers to extract informative lower dimension features.

Table 2 Number of scattering coefficients in different layers of scattering network

Layer	Number of features
$m = 0$	$2^{-2J} N \times K^0 \binom{J}{0} = 16 \times 1$
$m = 1$	$2^{-2J} N \times K^1 \binom{J}{1} = 16 \times 24$
$m = 2$	$2^{-2J} N \times K^2 \binom{J}{2} = 16 \times 192$

4.2 Reduced Scattering Representation on Malayalam Character Recognition

For dimension reduction, SVD approach described in Sect. 2.2 applied on higher-layer scattering coefficients. The maximum scale J of wavelet function employed for scattering transform is set to 3, and the number of angular orientations, K, is set as 8 for Malayalam character image databases. The maximum number of layers employed in our experiments is 2, as Bruna et al. [3] reported that the energy content captured after the second layer of scattering network is very small.

On the higher-layer features generated for training images, SVD is applied to calculate orthogonal bases vectors. By employing SVD orthogonal bases, SVD features are extracted for each training image. SVD features obtained through considering the different number of orthogonal bases are evaluated using SVM classifier for finding the minimum number of bases $optSVD$, which acquires maximum cross-validation accuracy among trained features. In experiments, the maximum number of orthogonal bases $maxNs$ is fixed as

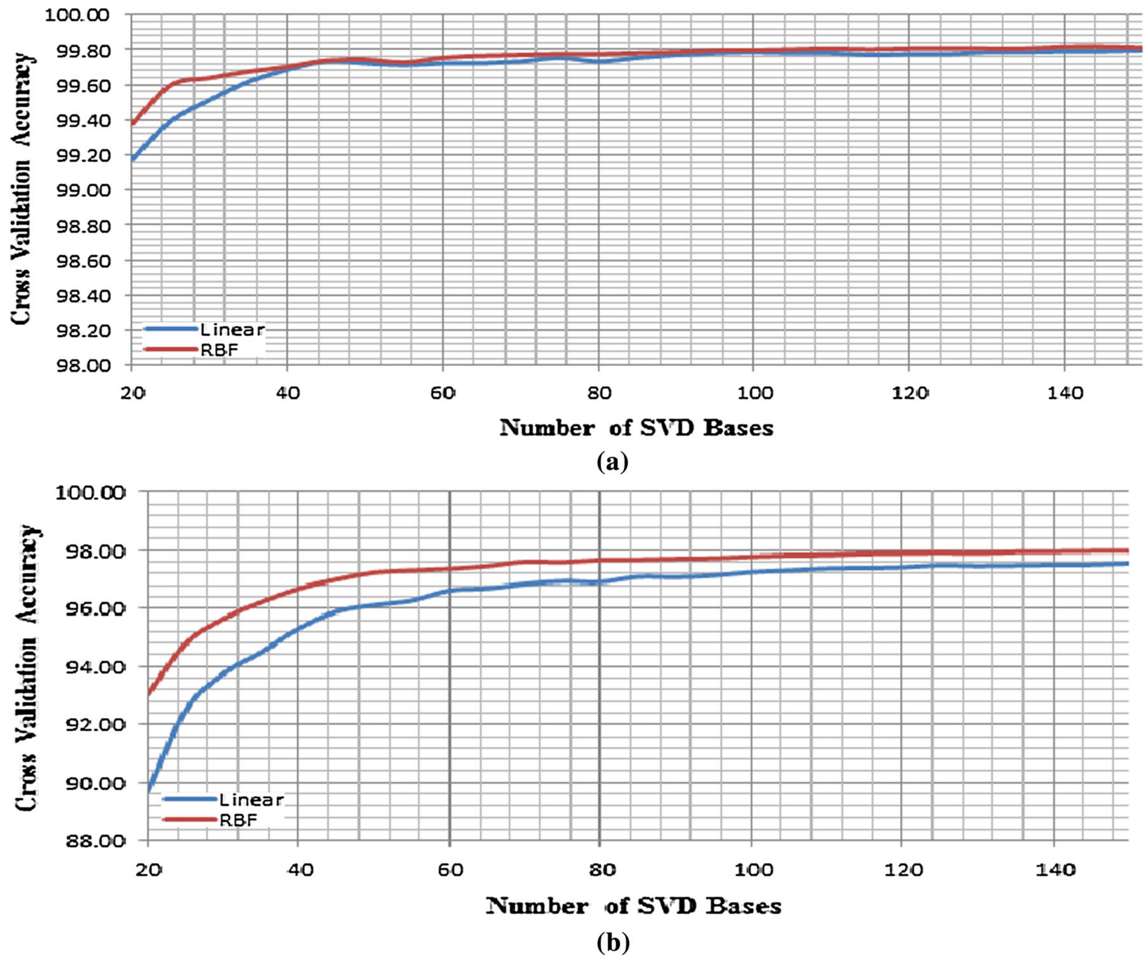


Fig. 9 Analysis on recognition accuracy with the increase in SVD bases using Linear and RBF SVM classifier on MAL_PrintedDB and MAL_HandwrittenDB. **a** MAL_PrintedDB and **b** MAL_HandwrittenDB

150 and $minNs$ as 20. Figure 9 shows the variation in cross-validation accuracy for the different number of SVD bases on Linear and RBF SVM classifier for MAL_PrintedDB and MAL_HandwrittenDB. From Fig. 9, it is clear that the recognition accuracy improves initially with the increase in the number of SVD orthogonal bases. The initial increase in recognition accuracy is because of the discriminating information concentrated on the first few SVD bases. The number of SVD orthogonal vectors which acquires maximum cross-validation accuracy is estimated and used further in creating reduced scattering feature representation. The SVD features are concatenated with the lower-layer scattering feature vectors to get proposed scattering representations. In Fig. 9, RBF SVM classifier has always better performance than Linear SVM. It is because of the fact that Linear SVM is a degenerate form of RBF SVM classifiers and properly tuned RBF SVM can build better decision surfaces compared to Linear SVM. During testing, the trend may or may not change based on the nature of testing dataset.

The recognition performance of scattering features compared with IMG and SVD_IMG features. The raw pixel value features (IMG) are directly classified with Linear and RBF SVM classifiers. SVD_IMG feature follows the feature extraction method outlined in Table 1, but on IMG features instead of scattering second-layer features. SVD_IMG features able to capture informative representation of IMG feature in less dimension with the help of SVD bases. SVD_IMG obtains improved performance than IMG features, and the recognition accuracy obtained on Malayalam character databases is listed in Table 3. Features obtained from the scattering network obtain better recognition performance compared to IMG and SVD_IMG features. Scattering ($m = 0$) represents zeroth-layer scattering network features. Scattering ($m = 0, 1$) represents features combined from zeroth and the first layer of scattering network. Layer2SVD represents the reduced scattering representation of the second layer of scattering network using SVD. Layer2SVD features improves the recognition performance when combined with zeroth and first-layer scattering features. SVD

Table 3 Recognition accuracy on Malayalam character image databases

Feature descriptor	Classifier	Recognition accuracy (%)	
		MAL_PrintedDB	MAL_HandWrittenDB
IMG	Linear SVM	87.21	83.10
IMG	RBF SVM	87.59	86.33
SVD_IMG	Linear SVM	87.70	84.10
SVD_IMG	RBF SVM	87.97	88.33
Scattering ($m = 0$)	Linear SVM	85.99	82.78
Scattering ($m = 0$)	RBF SVM	86.59	85.06
Scattering ($m = 0, 1$)	Linear SVM	89.56	95.07
Scattering ($m = 0, 1$)	RBF SVM	89.51	95.48
Scattering ($m = 0, 1$)+Layer2SVD	Linear SVM	89.70	95.58
Scattering ($m = 0, 1$)+Layer2SVD	RBF SVM	89.61	95.73

effectively managed to capture informative features from the large-dimensional second-layer scattering coefficients. Reduced scattering representation along with the lower-layer features achieve 89.70% on MAL_PrintedDB using Linear SVM classifier. In MAL_HandwrittenDB, the maximum accuracy obtained is 95.73% through the same feature in RBF SVM classifier. Scattering representation in MAL_HandwrittenDB obtains better recognition accuracy compared to that of IMG and SVD_IMG features.

In Table 3, MAL_HandwrittenDB obtains better recognition accuracy than MAL_PrintedDB. Compared with printed character recognition, handwritten character recognition is more complex and difficult. Even though the inter-class and intra-class variability is high in handwritten character images compared to printed character images, the contrariety in recognition accuracy listed in Table 3 happened due to the difference in nature of evaluation test dataset employed for printed and handwritten character recognition. The evaluation dataset for MAL_PrintedDB contains ERROR character images (mentioned in Sect. 3.1.1). As we are not applying rejection methods on recognition outcome to identify those ERROR, the maximum achievable accuracy in case of MAL_PrintedDB is 94.21% because of these ERROR character images. In effect this implies that 5.79% of misclassification in MAL_PrintedDB is due to these ERROR images. Beside this, the evaluation dataset for MAL_PrintedDB follows more natural count distribution of character classes in documents as it is created from real-world document images representing Malayalam language works. The number of test instances belonging to each class is very much varying in MAL_PrintedDB evaluation dataset compared to MAL_HandwrittenDB. 3.61% of misclassification in MAL_PrintedDB for proposed feature ($Scattering(m = 0, 1) + Layer2SVD$) happened for Malayalam dependent character sign *'anusvara'* (◌ം) and dependent vowel sign *'aa'* (◌ാ). The test character image belonging to these two classes is 10.66% of test dataset. Both the characters have a very strong structural resemblance with Malayalam consonant *'ttha'* (ഠ). In MAL_PrintedDB, only 0.89% of misclassification happened among other character classes. MAL_HandwrittenDB follows almost equal count distribution among character classes in test dataset and the misclassification happened among these three similarly shaped classes could not raise the overall misclassification rate much.

4.3 Comparison of the Proposed Technique with Other Feature Descriptors

The performance of proposed scattering feature descriptor (Scattering(m=0,1)+Layer2SVD), compared with other feature extraction methods, reported in the literature of Malayalam character recognition. Table 4 lists the recognition accuracy of different feature descriptors on MAL_PrintedDB and MAL_HandwrittenDB. Features based on projection profile, Gabor filters, wavelet transform and curvelets compared with proposed scattering feature descriptors. The proposed feature descriptor obtains better recognition performance compared to all the considered feature descriptors. Wavelet transform-based features achieve better performance compared to other considered literature techniques.

4.4 Evaluation of the Proposed Technique on Other Character Databases

In order to confirm the effectiveness of scattering representation, the recognition experiments extended to two other standard handwritten databases. MNIST handwritten digit database and Telugu handwritten database (HP_Telugu) are evaluated with different feature descriptors in RBF SVM classifier. From both handwritten databases (MNIST and HP_Telugu), only 10,000 handwritten images are considered during training as our main focus is on evaluating the effectiveness of reduced scattering representation rather

Table 4 Comparison of reduced scattering features with other Malayalam character recognition techniques

Feature descriptor	Recognition accuracy (%)	
	MAL_PrintedDB	MAL_HandwrittenDB
Wavelet transform of projection profile in MLP [15]	44.37	32.89
Gabor features in SVM [36]	82.43	74.90
Haar wavelet transform in SVM [13]	88.54	86.96
Curvelet in MLP [16]	81.03	72.23
Proposed scattering features	89.70	95.73

Table 5 Recognition accuracy of scattering features in MNIST and Telugu handwritten databases

Feature descriptor	Recognition accuracy (%)	
	MNIST	HP_Telugu
IMG	91.63	–
SVD_IMG	96.40	58.66
Scattering ($m = 0, 1$)	98.20	77.93
Scattering ($m = 0, 1$)+Layer2SVD	98.82	79.74

than acquiring state of art results on those databases. All the handwritten images in the test dataset are evaluated on the SVM classifier model built from different feature descriptors. Table 5 lists the recognition accuracy obtained on both the databases for IMG, SVD_IMG and scattering feature descriptors. Extracting IMG features from HP_Telugu database resulted in a very large dimension feature vector ($\gg$10K) which is almost undesirable in pattern recognition applications, so we omitted experimenting IMG features in HP_Telugu database. From Table 5, it is pretty much clear that the proposed feature descriptor obtains improved recognition performance compared with other feature descriptors on both databases. The results follow the same trend obtained on Malayalam character image databases.

The improvement in recognition performance for proposed feature descriptor on MAL_HandwrittenDB, MNIST and HP_Telugu indicates its ability in producing invariant feature descriptors even in the presence of greater degrees of variability in input images. Effectively reduced higher-layer features in the scattering convolution network could acquire better recognition accuracy when combined with the lower-layer features, which implies that the higher-layer scattering representations can produce even more discriminating feature descriptors. The SVD technique could effectively capture those discriminating information in lower-dimensional feature representation, which in turn helped to improve the recognition accuracy.

5 Conclusion and Future Work

Scattering convolutional network has the same concept as that of convolutional neural network. Instead of self-trained filters in CNN, the scattering network uses predefined wavelet filter coefficients. The scattering network generates feature descriptors through a cascade of wavelet decompositions followed by modulus and averaging operations. In this paper, the scattering representations are employed for Malayalam printed and handwritten character recognition. The Malayalam language printed (MAL_PrintedDB) and handwritten (MAL_HandwrittenDB) character databases considered in this paper contain 130 and 85 different character classes respectively. Scattering network-based feature descriptors could achieve far better recognition accuracy compared to IMG and SVD_IMG features. The explosion of scattering coefficients in the second layer of scattering network makes higher-layer features not suitable for applying directly inside classifier. By applying SVD-based dimension reduction technique, large-dimensional higher-layer features are transformed to lower-dimensional informative feature descriptors. The reduced representation of second-layer feature when combined with zeroth- and first-layer scattering coefficients improves the recognition accuracy further. The proposed scattering feature descriptor achieves recognition accuracy of 89.70% in MAL_PrintedDB and 95.73% in MAL_HandwrittenDB, respectively. The reduced scattering representation could capture the discriminating features in higher-layer scattering coefficients and that in turn lead to improve recognition performance. As scattering coefficients are capable to generate powerful feature descriptors, the predefined wavelet filters in scattering network can act as good initial values for receptive filter fields in convolutional neural networks. The future works are on integrating scattering features with self-learned features in CNN to improve the recognition accuracy further.

References

1. Due Trier, Ø.; Jain, A.K.; Taxt, T.: Feature extraction methods for character recognition—a survey. Pattern Recognit. **29**(4), 641–662 (1996)
2. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P.: Gradient based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998)
3. Bruna, J.; Mallat, S.: Invariant scattering convolution networks. IEEE Trans. Pattern Anal. Mach. Intell. **35**(8), 1872–1886 (2013)
4. Mallat, S.: Group invariant scattering. Commun. Pure Appl. Math. **65**(10), 1331–1398 (2012)
5. Obaidullah, S.M.; Halder, C.; Santosh, K.; Das, N.; Roy, K.: Phdindic_11: page-level handwritten document image dataset of 11 official indic scripts for script identification. Multimed. Tools Appl. 1–36 (2017) (**in Press**)
6. Govindaraju, V.; Setlur, S. (eds.): Guide to OCR for Indic Scripts. Springer, Berlin (2009)
7. Bag, S.; Harit, G.: A survey on optical character recognition for bangla and devanagari scripts. Sadhana **38**(1), 133–168 (2013)
8. Kumar, M.; Sharma, R.; Jindal, M.: Efficient feature extraction techniques for offline handwritten gurmukhi character recognition. Natl. Acad. Sci. Lett. **37**(4), 381–391 (2014)
9. Majhi, B.; Pujari, P.: On development and performance evaluation of novel odia handwritten digit recognition methods. Arab. J. Sci. Eng. 1–15 (2017) (**in Press**)
10. Mishra, T.K.; Majhi, B.; Sa, P.K.; Panda, S.: Model based odia numeral recognition using fuzzy aggregated features. Front. Comput. Sci. **8**(6), 916–922 (2014)
11. Shanthi, N.; Duraiswamy, K.: A novel SVM-based handwritten tamil character recognition system. Pattern Anal. Appl. **13**(2), 173–180 (2010)
12. Chacko, B.P.; Vimal Krishnan, V.R.; Raju, G.; Babu Anto, P.: Handwritten character recognition using wavelet energy and extreme learning machine. Int. J. Mach. Learn. Cybern. **3**(2), 149–161 (2012)
13. John, J.; Pramod, K.; Balakrishnan, K.: Unconstrained handwritten Malayalam character recognition using wavelet transform and support vector machine classifier. Procedia Eng. **30**, 598–605 (2012)
14. John, J.; Pramod, K.V.; Balakrishnan, K.: Offline handwritten Malayalam character recognition based on chain code histogram. In: International Conference on Emerging Trends in Electrical and Computer Technology, ICETECT 2011, pp. 736–741 (2011)
15. John, R.; Raju, G.; Guru, D.S.: 1D wavelet transform of projection profiles for isolated handwritten Malayalam character recognition. In: Proceedings—International Conference on Computational Intelligence and Multimedia Applications, ICCIMA 2007, vol. 2, pp. 481–485 (2008)
16. Manuel, M.; Saidas, S.: Handwritten Malayalam character recognition using Curvelet transform and ANN. Int. J. Comput. Appl. **121**(6) 24–31 (2015)
17. Raju, G.: Recognition of unconstrained handwritten Malayalam characters using zero-crossing of wavelet coefficients. In: Proceedings—14th International Conference on Advanced Computing and Communications, ADCOM 2006, pp. 217–221 (2006)
18. Raju, G.; Moni, B.; Nair, M.: A novel handwritten character recognition system using gradient based features and run length count. Sadhana **39**(6), 1333–1355 (2014)
19. Moni, B.S.; Raju, G.: Modified quadratic classifier for handwritten Malayalam character recognition using run length count. In: International Conference on Emerging Trends in Electrical and Computer Technology, ICETECT 2011, pp. 600–604 (2011)
20. Rahiman, M.A.; Rajasree, M.: Printed Malayalam character recognition using back-propagation neural networks. In: IEEE International Advance Computing Conference. IACC 2009, pp. 197–201. IEEE (2009)
21. Bhattacharya, U.; Shridhar, M.; Parui, S.K.; Sen, P.; Chaudhuri, B.: Offline recognition of handwritten bangla characters: an efficient two-stage approach. Pattern Anal. Appl. **15**(4), 445–458 (2012)
22. Bhowmik, T.K.; Ghanty, P.; Roy, A.; Parui, S.K.: Svm-based hierarchical architectures for handwritten bangla character recognition. Int. J. Doc. Anal. Recognit. (IJDAR) **12**(2), 97–108 (2009)
23. Kumar, M.; Jindal, M.; Sharma, R.: A novel hierarchical technique for offline handwritten gurmukhi character recognition. Natl. Acad. Sci. Lett. **37**(6), 567–572 (2014)
24. Bruna, J.; Mallat, S.: Classification with scattering operators. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1561–1566 (2011)
25. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
26. Manjusha, K.; Anand Kumar, M.; Soman, K.P.: Scattering representation in Malayalam character recognition. In: Twenty Third National Conference on Communication (NCC). IEEE (2017)
27. Kalman, D.: A singularly valuable decomposition: the svd of a matrix. Coll. Math. J. **27**(1), 2–23 (1996)
28. Kumar, S.S.; Manjusha, K.; Soman, K.P.: Novel SVD based character recognition approach for Malayalam language script. Adv. Intell. Syst. Comput. **235**, 435–442 (2014)
29. Soman, K.P.; Loganathan, R.; Ajay, V.: Machine learning with SVM and other kernel methods. PHI Learning Private Limited, New Delhi (2009)
30. Aly, M.: Survey on multiclass classification methods. Technical report, Caltech, USA (2005)
31. Chang, C.C.; Lin, C.J.: Libsvm: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST) **2**(3), 27 (2011)
32. Bresson, X.; Esedoglu, S.; Vandergheynst, P.; Thiran, J.P.; Osher, S.: Fast global minimization of the active contour/snake model. J. Math. Imaging Vis. **28**(2), 151–167 (2007)
33. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979)
34. Prasanth, L.; Babu, V.; Sharma, R.; Rao, G.; Dinesh, M.: Elastic matching of online handwritten Tamil and Telugu scripts using local features. In: Ninth International Conference on Document Analysis and Recognition, ICDAR 2007, vol. 2, pp. 1028–1032. IEEE (2007)
35. Sifre, L.; Kapoko, M.; Oyallon, E.; Lostanlen, V.: Scatnet: A matlab toolbox for scattering networks (2013)
36. Ramanathan, R.; Nair, A.; Thaneshwaran, L.; Ponmathavan, S.; Valliappan, N.; Soman, K.: Robust feature extraction technique for optical character recognition. In: ACT 2009—International Conference on Advances in Computing, Control and Telecommunication Technologies, pp. 573–575 (2009)