Journal of AI and Data Mining
Vol 8, No 2, 2020, 161-175.
DOI: 10.22044/JADM.2019.7402.1877

Sparse Structured Principal Component Analysis and Model Learning for Classification and Quality Detection of Rice Grains

S. Mavaddati*

Electronic Department, Faculty of Technology and Engineering, University of Mazandaran, Babolsar, Iran.

Received 26 August 2018; Revised 30 November 2019; Accepted 12 December 2019
**Corresponding author: s.mavaddati@umz.ac.ir (S. Mavaddati).*

Abstract
In the scientific and commercial fields associated with modern agriculture, the categorization of different rice types and determination of their quality is very important. In the recent years, various image processing algorithms have been applied to detect different agricultural products. The problem of rice classification and quality detection is presented in this paper based on the model learning concepts including the sparse representation and dictionary learning techniques to yield over-complete models in this processing field. There are color-based, statistical-based, and texture-based features available to represent the structural content of rice varieties. In order to achieve the desired results, different features from the recorded images are extracted and used to learn the representative models of rice samples. Also the sparse principal component analysis and sparse structured principal component analysis are employed to reduce the dimension of classification problem, which leads to an accurate detector with a less computational time. The results of the proposed classifier based on the learned models are compared with the results obtained from neural network and support vector machine. The simulation results along with a meaningful statistical test show that the proposed algorithm based on the learned dictionaries derived from the combinational features can detect the type of rice grain and determine its quality precisely.

Keywords: *Rice Classification, Quality Detection, Sparse Structured Principal Component Analysis, Model Learning, Neural Network, Statistical Test.*

1. Introduction
Rice classification and detection of its quality have been an important field of image processing in the recent years. Rice is the main food product for the people in Iran and the world [1]. The traditional methods used for rice categorization are based on vision and olfaction but they are typically time-consuming and are not particularly reliable for an inexperienced buyer. Also the sensory evaluation in this classification process can be affected by the mental conditions such as fatigue. Therefore, the data mining and image processing techniques are applied in this area to result in an accurate classification with a lower recognition time. The quality of different agricultural products can be specified using these processing techniques and feature extraction of the recorded images. These features contain color-based, morphological-based and texture-based properties [2-6]. This issue is important and a few types of research works have been performed in this field

In [6], the classification results of wheat, oats, and rye grains have been considered using a combination of color, texture, and morphological characteristics. In [7], the color and texture-based features have been used to train a back-propagation neural network and detect the wheat, barley, oats, and rye grains. An appropriate classification accuracy has been obtained using these features. In [8], the bulk and individual samples of wheat, barley, oats, and rye varieties have been categorized using the color- and texture-based characteristics, and proper results have been obtained. A method for categorization of six rice types using the color and morphological features has been presented in [9]. Seven color features and forty morphological properties of each rice grain have been extracted and used in the classification step. In [10], the characteristics related to the length, area, perimeter, maximum length, maximum width, and compression for three rice types have been calculated and classified by the neural network. The appropriate results due to the small number of the considered products can be justifiable. In [11], the color and appearance of rice have been considered using the image processing methods. In [12], the quality of rice varieties has been investigated based on the observed broken rice grains by calculating the length and width of grains based upon the bulk data. In [13], the color-based property and gray level co-occurrence matrix (GLCM) have been extracted from the rice images, and the neural network has been used as a classifier. In [14], the morphological characteristics of four Iranian rice products have been extracted and the classification has been carried out based on the linear discriminant analysis (LDA) and artificial neural network. In [15], the classification of three Iranian rice products from the mixed samples has been performed by the combinational coefficients including the GLCM and local binary pattern (LBP) features. In the following, the combination of two methods of Fisher's coefficient and principal component analysis are used to select the effective features in this categorization. The classifier is the learning vector quantization (LVQ) neural network. In most papers, a rice classification method has been presented using individual recorded samples. Another procedure for rice classification uses the bulk sample [14-15]. In [16], the type of rice grains has been detected using a deep convolutional neural network learning method. This classification procedure can identify the type of rice grains as broken or fine.

Another procedure for rice classification uses the bulk sample [17-20]. A classifier fusion algorithm used to recognize rice variety over bulk samples has been presented in [17]. The length matrix features have been used to earn seventeen superior features in this classification procedure. In [18], a support vector machine has been employed to identify the rice varieties using the features calculated from the co-occurrence matrix. In [19], the texture-based features based on GLCM and back-propagation neural network have been used to determine five Iranian rice products. In [20], different color features extracted from three color channels and a back-propagation neural network have been used to classify rice grains.

The model used in the presented paper is based on bulk samples of rice varieties.

In the commercial trading of rice product, each rice type may be mixed with other poor varieties. In this case, the exact determination of rice purity is very important and the issue of rice quality recognition is discussed.

In this paper, a new method is proposed for rice classification and determination of its quality based on the concept of model learning. The texture-based and statistical-based characteristics captured from the bulk samples of different rice products were used to form a combinational feature vector. Then the sparse principal component analysis (SPCA) and sparse structured principal component analysis (SSPCA) were applied to reduce the dimension of the feature vectors obtained. In this work, the dictionary learning and sparse coding procedures were used to train the over-complete models for different rice types. The proposed dictionary learning algorithm was designed in such a way that the trained model for each rice type had the least coherence parameter value with each other. In the proposed algorithm, a classifier was designed based on the energy of sparse coefficients, and there was no need to have a common classifier such as neural network or support vector machine (SVM). The presented algorithm was not sensitive to the rotation and light of the environment that were challenging problems in this categorization filed since the desired texture-based and statistical-based features were considered. Different rice types considered in this work were Tarom, Shiroodi, Fajr, and Binam, which are used in the north of Iran.

The rest of this paper is organized as what follows. Section 2 is dedicated to the dictionary learning and sparse representation procedures, and then different feature extraction methods followed by definitions of the SPCA and SSPCA algorithms are investigated in Section 3. In Section 4, the proposed rice classification algorithm is proposed. In Section 5, the results of the proposed classifier are reported and compared with other common classifiers. In the last section, the paper is concluded.

2. Sparse representation and dictionary learning

The input image I can be modeled using the dictionary learning technique as:

$$\mathrm{I_m} = \mathrm{DX} \tag{1}$$

where, $\mathrm{I_m}$, D, and X are the input data, dictionary matrix, and coding matrix, respectively. $\mathrm{I_m}$ is a data matrix including different patches of I. The input image is divided into different patches $\mathrm{I_{m \in M}}$, where M involves the dimension of these patches. The data matrix I_m in the sparse representation can be coded by a linear combination of defined atoms in an over-complete dictionary $D \in R^{P \times L}, L > P$, as shown in Eq. 1. This dictionary includes L columns or atoms $\{d_l\}_{l=1}^{L}$ with the unit norm $\|d_{(:,l)}\|_2 = 1, \forall l = 1,...,L$. The over-complete dictionary means that the number of columns or atoms of a dictionary is greater than the number of rows or dimension of the feature space. The ratio of the number of dictionary's column to the dictionary's rows is called the redundancy rate. Also the coding matrix X with K cardinality parameter and $L ? K$ consists of the sparse coefficients of I_m [21-23]. The value of the cardinality parameter K determines how many columns of D matrix can participate in the representation of each input data. Each column of X includes only K non-zero elements. The sparse representation problem with the approximation error and sparsity constraint terms is formulated as [24]:

$$X^* = \underset{X}{\operatorname{argmin}} \|I - DX\|_2^2 \quad \text{s.t.} \quad \|X\|_0 \le K \tag{2}$$

where, X^* is the sparse coding matrix and $\|X\|_0$ denotes the number of non-zero coefficients in each row of X^*. The sparse representation technique makes it possible to show the major information of the input data based on a smaller dimension of the original spatial bases. The over-complete dictionary learning was first presented for image denoising using the definition of the K-SVD technique with proper results in this area [25].

The learning of over-complete dictionaries consists of two stages: sparse representation and updating of atoms. Due to the flexibility of these defined steps, each one of these steps can be carried out with any arbitrary procedure. The methods used for sparse coding are different according to how their coding parameters are set. These tuning procedures should be implemented carefully. In addition to the K-SVD algorithm, other dictionary learning methods such as the maximum likelihood (ML), method of optimal direction (MOD), and maximum a posteriori (MAP) exist. In all of these algorithms, the convergence rate decreases with increase in the training data [26-27].

3. Feature extraction and dimension reduction procedure

Similar to all the common classifiers, the feature extraction step is the first step in the field of rice classification [28-30]. Two common types of features including the texture-based and statistical-based feature extractions are utilized in this processing filed. In the following, different methods are investigated in order to extract these properties. Then the dimension reduction algorithms based on SPCA and SSPCA are introduced.

3.1. Gabor filter

One of the important features in the texture-based image analysis is to use the magnitude and phase coefficients generated by applying the Gabor filter to an image. Using this linear filter, the frequency components in different directions of the image are calculated, which is very important to distinguish between the different regions of an image. A filtered image is obtained by applying any Gabor filter in the specified direction of an image. The Gabor filter is yielded by the expansion and rotation of a Gabor function [31-33]. The 2D Gabor function $g(x, y)$ is expressed as:

$$\begin{aligned} g(x, y) = 1/(2\pi\sigma_x\sigma_y)\exp[-1/2(x^2/(\sigma_x^2) \\ + y^2/(\sigma_y^2)) + j\omega(x\mathrm{Cos}\theta + y\mathrm{Sin}\theta)] \end{aligned} \tag{3}$$

where, σ_x and σ_y are the standard deviations along the x and y directions, respectively. Also x and y denote the pixels, and ω and θ are the frequency and desired direction, respectively. In order to eliminate the intensity of the illumination effect in the image brightness, the value of the DC coefficient for each output of the Gabor function is ignored. The output of each Gabor function at the specified angle is two magnitude and phase matrices with the same dimension of the original image. Two final magnitude and phase matrices for the Gabor features are obtained based on averaging the different specified directions [31-33].

3.2. Discrete cosine transform (DCT)

The image is decomposed into various frequency bands to which the human eye system has a high sensitivity using DCT. These bands include low-, middle-, and high-frequency bands. The coefficients in the low bandwidth contain a large portion of the image energy, and the coefficients in the high-frequency bands include the minimum amount of energy. The coefficients of the 2D cosine transform $G(x,y)$ for an input image I with dimensions $N \times M$ can be calculated as follows [31-33]:

$$G(x,y)=\frac{2}{\sqrt{M\times N}}\alpha(x)\alpha(y)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}I(i,j)\times$$
$$\mathrm{Cos}\left(\frac{(2i+1)x\pi}{2M}\right)\times\mathrm{Cos}\left(\frac{(2j+1)y\pi}{2N}\right) \quad (4)$$
$$\alpha(x),\alpha(y)=\begin{cases}\frac{1}{\sqrt{2}} & \text{if } x,y=0\\ 1 & \text{Otherwise}\end{cases}$$

where, I is the input image with the x and y pixels. This transform results in an image in the frequency domain with the same dimension as the initial image.

3.3. Discrete Fourier transform (DFT)

The input image is split into the sinusoidal and cosine components by applying this transform and transfer it into the frequency domain. Therefore, the original gray level scale image is approximated by an unlimited set of these functions. The Fourier transformation function for an input image I with dimensions $N \times M$ can be expressed as [31-33]:

$$F(x,y)=\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}I(i,j)e^{-i(x_{i/M}+y_{j/N})} \quad (5)$$
$$e^{ix}=\cos x+i\sin x$$

In this regard, F represents the input image in the frequency domain with a set of complex numbers and the same dimension as the original image. Therefore, the representation of these Fourier coefficients can be in the form of the magnitude and phase of the frequency content of image. Usually the magnitude of the Fourier coefficients is used in the image processing algorithms since these coefficients involve sufficient information about the content of the input image [31-33].

3.4. Local binary pattern (LBP)

The LBP method is one of the most robust feature extraction procedures that is widely used in the research fields related to image retrieval [34]. This algorithm is a rotation-invariant descriptor used to analyze grayscale images to extract the properties of the adjacent textures. This attribute makes it possible that the change in the position, light or rotation of the sample to have a less adverse effect on a classification algorithm.

There are several procedures available to calculate the local binary pattern coefficients that depend on the choice of neighborhood type. These neighborhoods can be regarded as diagonal or circular with different radii [34].

3.5. Histogram of oriented gradients (HOG)

One of the feature extraction algorithms in machine learning that is very effective in classification purposes is HOG [35, 36]. In this method, the number of gradient occurrences in different directions of the local sections in the image is calculated. This counting is done on different cells that are considered in the overlapped blocks of image. In this case, the image is first divided into the blocks with a 50% overlap, and then each block is divided into four cells. In the following, the magnitude and angle of the gradient in each pixel of the image are calculated. Then a histogram with 9 directions consisting of different gradient angles in this cell is obtained. The values of each column are the sum of gradient magnitudes that have the same angle as the value of this column. In HOG, the gradient angles can be calculated in the range of 0-180° to create a histogram. It should be noted that this descriptor is rotation-variant but the difference in the illumination intensity will have a little effect on the extracted properties [35, 36]. In the experiments carried out in this work, a histogram descriptor with the size of cells 9 × 9 and blocks 2 × 2 was used to extract the HOG features.

3.6. Gray level co-occurrence matrix (GLCM)

Another feature that determines the properties of the image texture is the extracted parameters of GLCM [37]. The textural information of an image can be expressed by means of a matrix with relative abundances $P(i,j)$, where the value of each row with the number of row i and the column number j represent the number of occurrence neighborhoods with the same gray level value in different directions (for example 0°, 45°, 90°, 135°, and 180°).

The coefficients of this matrix are not used directly as a feature for classification, and the statistical parameters that result from it determine the content of the image. These parameters include the mean, variance, energy, range of variations in relative abundance, contrast, homogeneity, uncertainty, maximum relative frequency, correlation, and entropy. The selection of all or part of these second-order statistical properties in the image processing routine will be effective in the texture analysis of image [37].

3.7. Moments

The extracted moments of an image is another statistical feature discussed in the field of image processing, which is very important for its rotation-invariant property [33, 38]. The values of this feature vector, which consists of seven coefficients of the first to seventh moments, do not change with the image rotation in each direction. Therefore, this feature set is very important in the classification algorithms that the rotation problem is a fundamental challenge for them.

3.8. Dimension reduction using SPCA and SSPCA

Principal component analysis (PCA) is a statistical technique used to reduce the dimension of data analysis and find the variables of data with maximum variance [39]. PCA seeks the linear combinations of the original components such that the derived output components are in a k-dimensional sub-space and k is smaller than the initial dimension of data. An approach used to estimate the principal components with sparse constraints is sparse PCA (SPCA). The SPCA technique is designed based on the lasso penalty and regression criterion [40]. SPCA results in a flexible control over the sparse structure of the extracted components, and has different advantages such as computational efficiency, high explained variance, and ability in the identification of important variables [40].

SSPCA is an extension of sparse PCA, which finds the data variance by sparse constraints as well as some a priori structural constraints to model the data content [41]. A non-convex variant of the regularization in [42] is presented for the problem of structured sparse dictionary learning. SSPCA applies an efficient block-coordinate descent algorithm with closed-form updates for a better decomposition of the data. SSPCA is robust to occlusion problem since this technique has a local dictionary learning procedure.

4. Rice classification in proposed method

In the first step of the proposed rice classification problem, an over-complete model is learned for each data category for different rice products. Some samples of this dataset are shown in figure 1. The proposed rice classification method is described in the following sections.

4.1. Procedure of rice data recording

In order to prepare the bulk samples for different types of rice grains, a box with dimensions of 30 cm × 40 cm × 60 cm was prepared, as shown in figure 2. In order to create a consistent light in the interior of the box, 4 LED strips with length 25 cm were used in the inner and upper sides of the box to prevent shadow over samples. A circle with a diameter of 5 cm was created at the center of the upper level of the box to record the images. In order to prepare the samples, a rectangular container with dimensions 10 cm × 10 cm was used. The rice sample was placed in this container, and the surface was flattened so that no shadow could be created in the sample. A Sony Imaging Camera with a 300imx sensor and 19-Megapixel resolution was used to record the images.

4.2. Sparse representation in proposed method

A structured set of each data class can be used as a comprehensive model to properly categorize the rice varieties using the dictionary learning technique. The first step in the dictionary learning process is the sparse coding of the training data, where each patch of the image is represented by several atoms. The concept of sparsity in this representation means that each data patch will be coded only by a linear combination of different atoms specified based on the cardinality value. The next step in the dictionary learning algorithms is to update the dictionary atoms according to the input data patches. Since the dictionary is over-complete and the dimensions of the problem in Eq. 1 is high, it leads to an under-determined solving procedure, where the number of linear equations is much lower than the parameters.

Therefore, the atom learning process is carried out in the mentioned steps including sparse representation and update of the dictionary atoms. In the first step, an initial dictionary is selected randomly from the training data. The sparsity coefficients are obtained in the sparse coding procedure.

In the second step, the dictionary atoms are updated according to the obtained sparsity coefficient matrix. The fundamental difference between the dictionary learning algorithms is in the coding algorithm or the employed dictionary learning method. The sparse coding method and the dictionary learning algorithm employed in this work are introduced in this section and the next section, respectively. In the proposed rice classification algorithm, the least angle regression with coherence criterion (LARC) sparse representation method is used, which is an extension of the LARS algorithm [43]. In the LARC coding, a stop condition based on the coherence value between the atoms and training data is set [43-44].

Figure 1. Training data used in the rice classification problem a) Tarom rice b) Shiroodi rice c) Fajr rice d) Binam rice.

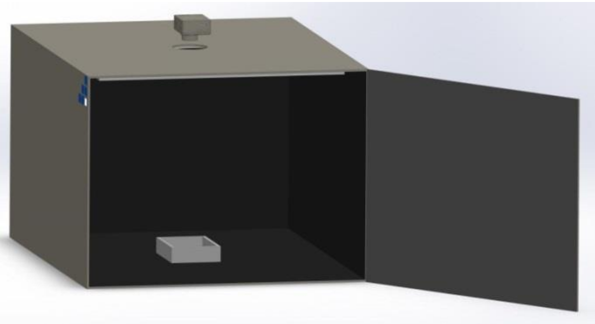


Figure 2. Box designed to record image of the rice grains.

Using this algorithm, only the atoms whose coherence values with the data patches are more than a certain value Coh, named as the residual coherence, are involved in the dictionary. In this case, the atoms must be coherent with the data patches in order to better represent the content of the training data. Another characteristic of this algorithm is that it uses the variable cardinality parameter instead of a constant rate used in most sparse coding algorithms. In this method, the upper bound of the cardinality parameter K is determined. It means that each data patch can be represented at most with K atoms. The LARC technique was first introduced in order to represent the speech signal [43]. The sparse representation based on this algorithm can be expressed as follows:

$$X^* = \mathrm{LARC}(D, X, K, Coh) \tag{6}$$

where, K indicates the variable cardinality parameter and Coh is the value of residual coherence. The Coh parameter, which expresses the coherence between the atom and the training data, should be set precisely. If the value of this parameter is high, it will only be possible to include a few number of atoms that have a coherence value higher than the Coh parameter. Then a problem such as a high approximation error may occur in this learning procedure. Also if the value of this parameter is low, this coherence parameter will not affect the sparse coding procedure.

4.3. Dictionary learning procedure in proposed method

The K-SVD algorithm is a procedure for atom training based on a set of input data [34]. In this algorithm, each patch of the input image is coded by the linear combination of the trained atoms. The coherence parameter in the learning procedure should be set precisely. The coherence parameter between the training data and atoms expressed in the previous section should have high a value. Also a low value for the coherence parameter between the atoms leads to a lower approximation error according to (1). In the case of the coherence between the dictionary atoms, it should be mentioned that there is a least coherence value between the dictionary atoms when the spatial bases are independent from each other as much as possible and the best representation for the content of the data patches is achieved. These parameters should be considered with more precision when the training data belonging to different classes has a similar structure. The higher coherence value between data/atoms and also lower mutual coherence between atoms result in the minimum reconstruction error. Therefore, the learned atoms can better represent the structural characteristics of the training data. The mutual coherence between the dictionary atoms is obtained based on the maximum value of the correlation for different atoms:

$$\mu(D) = \max_{1 \le i, j \le L, i \ne j} \left| d_i . d_j \right| \tag{7}$$

where, d is the dictionary atoms or the columns of dictionary matrix. The maximum absolute value of the correlation parameter between the atoms should be obtained as small as possible to result in a trained model with the independent spatial bases [43].

It is usually difficult to find a dictionary with this characteristic when the dictionary dimension is high. Therefore, the approximate methods are used to access the dictionary with the incoherent atoms. The approximation procedure used to solve this problem leads to the definition of Gram matrix $G = D^T D$. It should be noted that if the Gram matrix of a dictionary is in the form of the unitary matrix, then the atoms of this dictionary will be as independent as possible [44, 45]. In order to achieve this Gram matrix for any desired dictionary dimension, a different approximate solution is considered. One of these solutions is the post-processing algorithms in the dictionary learning process. In [46], an iterative projection followed by the rotation (IPR) technique has been introduced to obtain this desired Gram matrix.

In the first step, the non-diagonal coefficients of this matrix are bounded by a constraint set called the structural constraints, and then the number of Eigenvalues is limited. In this thresholding procedure, the non-diagonal coefficients that must be zero in the ideal form of Gram matrix are set to the pre-defined small coherent value μ_0 to reduce the Frobenius norm between the Gramm matrix and the unitary matrix as $\|G - I\|_F$. Then the Eigenvalues of the Gramm matrix are limited. This procedure is performed by maintaining only the N largest Eigenvalues.

In the second step, the obtained atoms are rotated with an orthogonal matrix W since the approximation error may increase in the thresholding procedure of the first step. The rotation step tries to decrease the approximation error $\|Y - WDX\|_F$. This technique was first used to improve the reconstruction error for the music signal. This technique was used in this work to model the training data in the rice classification problem, and resulted in a better classification rate [46].

4.3.1. Correction of dictionary atoms

It is important that the designed dictionaries that are related to each data class do not have any similarity with each other, and a distinction exists between different categories. Therefore, the learned dictionary must have the least coherence value with the atoms of other dictionary classes. In this section, this issue is checked whether the atoms with the same structure exist in the representation of the training data in a fixed dictionary related to each rice class. If this problem is confirmed, then a routine is designed to reduce this similarity.

In the proposed correction step, a composite dictionary $D = [D_T \, D_{SH} \, D_F \, D_B]$ including the dictionaries related to the Tarom rice D_T, Shiroodi rice D_{SH}, Fajr rice D_F, and Binam rice D_B is regarded. Then the data related to the Tarom rice is coded over this composite dictionary:

$$\boldsymbol{X}_T^*, \boldsymbol{X}_{SH}^*, \boldsymbol{X}_F^*, \boldsymbol{X}_B^* = \mathrm{LARC}\left(Y_T, [D_T \; D_{SH} \; D_F \; D_B], \mathrm{coh}\right) \rightarrow \underset{\boldsymbol{X}_T, \boldsymbol{X}_{SH}, \boldsymbol{X}_F, \boldsymbol{X}_B}{\arg\min} \left\| Y - [D_T \; D_{SH} \; D_F \; D_B] \begin{bmatrix} X_T \\ X_{SH} \\ X_F \\ X_B \end{bmatrix} \right\|_F^2 \tag{8}$$

where, X_T, X_{SH}, X_F, and X_B show the sparse matrix related to the Tarom rice, Shiroodi rice, Fajr rice, and Binam rice, respectively. In the following, the energy of the sparse coefficients is calculated for different rice varieties:

$$E_T = \frac{1}{L}\sum_{l=1}^{L} X^{*2}_{T,l}, \; E_{SH} = \frac{1}{L}\sum_{l=1}^{L} X^{*2}_{SH,l}, \quad E_F = \frac{1}{L}\sum_{l=1}^{L} X^{*2}_{F,l}, \; E_B = \frac{1}{L}\sum_{l=1}^{L} X^{*2}_{B,l} \tag{9}$$

In this equation, E_T, E_{SH}, E_F, and E_B include the energy of sparse matrix calculated for the Tarom rice, Shiroodi rice, Fajr rice, and Binam rice, respectively. Since the data related to the Tarom rice should not be represented on the other dictionaries, the atoms of other dictionaries such as D_{SH}, D_F, and D_B that have the largest energy in this representation are replaced by the dictionary atoms that have the least energy in this sparse coding. Also the sparse coding of the other data class is performed on this composite dictionary and the energy of this representation related to the rice data is calculated. This routine is performed four times for each data of different rice category and the atoms are corrected. In this case, the reconstruction error in the dictionary learning process of each data class is reduced as much as possible.

4.4. Proposed procedure for rice classification

It is important in dictionary learning that the captured atoms have the highest degree of coherence with the training data. Also these atoms should have the least degree of coherence with each other and with the atoms of other data classes. In this section, the proposed procedure for identification of the rice grains based on the incoherent dictionaries is proposed. In order to classify the input rice data in this work, the conventional classifiers such as neural network and support vector machine are not used but it is suggested that a dictionary-based classifier based on the extracted features of the sparse representation algorithm introduced in Section 3 is designed and applied.

In the proposed procedure, at first, the sparse representation is performed using the LARC sparse coding algorithm over the composite dictionary introduced in the previous section. This process is carried out with the same amount of cardinality value for each dictionary in the training step. Then the energy of the coefficient matrix in this representation for each dictionary is calculated. The energy value for this representation is used to classify the input data into the desired class since this energy will be more for the dictionary related to each data class. In this case, there will be no need to use other classifiers, and the label estimation of the input data will only be possible using the sparse representation technique. A block diagram of the proposed method including the training and test steps for classification of rice varieties based on learning the incoherent dictionary and dimension reduction technique is shown in figure 3.

5. Simulation results

In order to evaluate the performance of the proposed classifier, a collection of rice images is recorded according to Section 4.1. This collection contains 500 color image data for each rice type with size 350 × 600. In the first step of simulations, the three color channels of the input image are obtained, and the gray levels of each channel are calculated. Then the extracted features of each channel are used to make the final feature vector.

In the proposed method, the learned dictionaries are applied to determine the rice variety. The various features introduced in Section 3 are used as the training data extracted from the input images in the simulations. The cardinality parameter required by the LARC algorithm in the sparse coding step depends on the dimension of training data, and will be different for each calculated feature. The Coh parameter expressed in (6) and (8) in all simulations is adjusted to 0.2.

The redundancy rate of over-complete dictionaries for all features and rice categories is set to 4. This means that the learned models are over-complete with a redundancy rate of 4. In order to learn the dictionary in the training step, 400 images of each rice type are used and also 100 images are considered in the test step to evaluate the performance of the proposed method.

The performance evaluation of different methods is carried out by the classification accuracy rate calculated by the percentage of true categorized image divided by the total number of test images.

5.1. Feature extraction based on HOG coefficients

The extracted feature in this experiment is based on the HOG coefficients with 9 × 9 cell size, 50% overlap, and 2 × 2 block size. At the first step of simulations, the input image is resized to 70 × 70. The dimension of the HOG-based features including 36 blocks with 9 × 9 cell size, 9 bins, and 4 directions is 1296. The convergence plot of the proposed dictionary learning for the extracted HOG features captured from the Tarom rice grains is shown in figure 4. As it can be seen, a suitable convergence with a low approximation error is obtained based on the calculated root mean square error (RMSE). This plot for learning over-complete dictionaries with a redundancy rate of 4 is derived from training of a data matrix with 1296 × 400 dimension for each rice product.

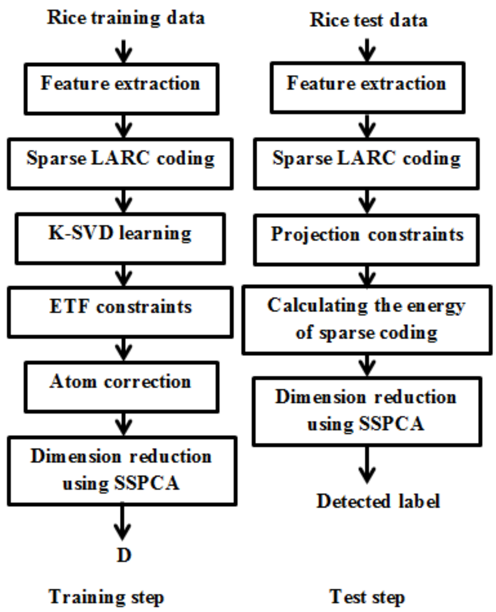
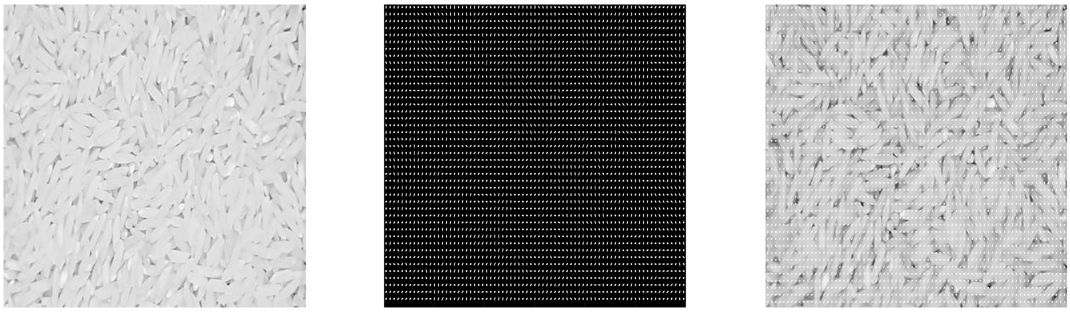
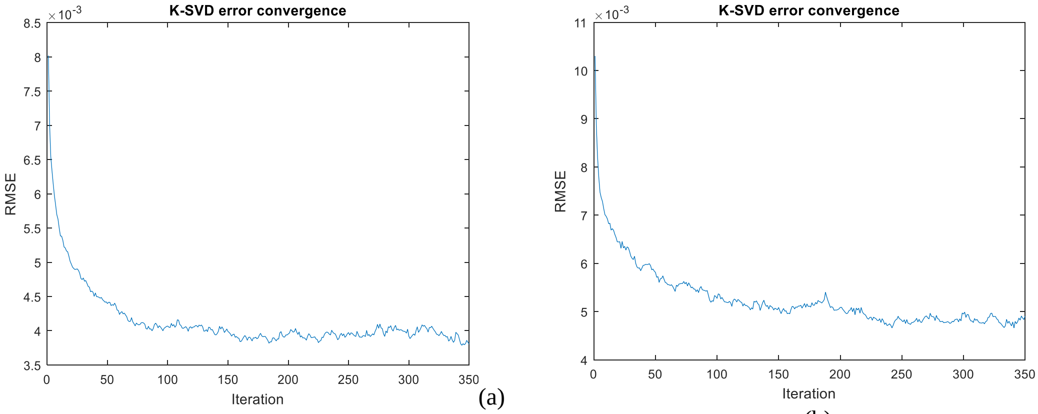


Figure 3. Block diagram of the proposed rice classification method based on the learned incoherent models and the dimension reduction technique.



(a) (b) (c)

Figure 4. Extracted HOG features for a Tarom rice image a) Tarom rice data B) HOG features c) Extracted HOG features.



(b)

Figure 5. Convergence plot of the proposed dictionary learning algorithm using the HOG features: a) for Tarom rice data B) for Shiroodi rice data.

The convergence plot of the proposed learning procedure for the HOG features extracted from the training data of the Tarom and Shiroodi rice products are shown in figure 5. As shown, a low RMSE value is obtained in this learning process.
The cardinality rate in the training and test steps of this simulation is set according to the smallest calculated approximation error in the experimental results. The classification results of the rice grains based on HOG feature for the proposed dictionary-based classifiers using SPCA and SSPCA with different percentages of dimension reductions are shown in table 1.
In this simulation, 10% and 20% of the dimension in the training data matrix are reduced to provide a classification process using the main feature coefficients.
The results of the classification rate using the HOG feature for different classifiers including the feed-forward neural network with 30 hidden layers, support vector machine, and incoherent dictionary learning technique with dimension reduction algorithms are reported in table 2. As it can be seen, the classification accuracy rate for the proposed method is greater than the results for the other algorithms.

5.2. Feature extraction based on Gabor coefficients

According to Section 3.1, the magnitude and phase of the Gabor coefficients are the main features in the texture analysis since they show the frequency components of the input image in different directions. The magnitude and phase matrices of the Gabor coefficients calculated for the Tarom rice image are shown in figure 6, which have the same dimensions as the initial image. The Gabor coefficients in this simulation are calculated in four directions: 0°, 45°, 90°, and 135°. The magnitude matrix of the Gabor coefficients in the specified directions for the rice image is shown in figure 7, which indicates the content of the frequency components in different directions. Then the sum of the row and column elements of the magnitude Gabor matrix is considered as the final feature for dictionary learning. As a result, the 560×400 training data is used for each rice category. Using these features, the cardinality parameter is set to 30 in the training and test steps according to the data dimension of 560, which is estimated from the experimental tests to result in the lowest approximation error. The results of the classification rate using the Gabor features calculated for different classifiers including a feed-forward neural network with 30 hidden layers, support vector machine, and incoherent dictionary learning algorithm with the SPCA and SSPCA techniques are presented in table 3. As it can be seen, the proposed classifier achieves the desired results using the magnitude feature of the Gabor coefficients. Therefore, this category of features can provide an appropriate distinction between the different data categories. Also along with the magnitude of the Gabor coefficients, the phase of these features is also used, and the final feature vector is set based on the column mean of the magnitude and phase matrices to learn the dictionaries. Therefore, for the rice image data, the 140 × 400 training data is used. The cardinality rate for these features in the training and test steps is adjusted to 20 according to the data dimension. It means that each training data is maximally coded with 20 atoms. The results of the classification rate using these features for the mentioned classifiers are reported in table 4. As it can be observed, the classification results of the proposed classifier are lower than those obtained for other defined features introduced so far. Therefore, the combination of the magnitude and phase of the Gabor coefficients cannot adequately distinguish the different rice types.

5.3. Feature extraction based on Fourier transform and discrete cosine transform

The main goal of this work is to consider the extracted features from the rice images and determine their efficiency to solve the rice classification problem. Therefore, the results of these features are analyzed, and a desirable feature category with the ability to make more distinctions between the rice classes is chosen. The results of rice classification using the extracted DFT and DCT features are shown in table 5. Also the results of these coefficients in combination with the HOG feature vector are reported in this table. These results show that the combination of these features has not succeeded to solve this problem.

Table 1. Results of the classification rate for rice images based on the HOG feature for the proposed dictionary-based classifiers using SPCA and SSPCA with different percentages of dimension reduction.

	Tarom rice	Shiroodi rice	Fajr rice	Binam rice
Proposed	90	92.5	91	93.5
Proposed with SPCA (20%)	86	88	90	89.5
Proposed with SPCA (10%)	92	93	92.5	94
Proposed with SSPCA (20%)	88	90.5	89.5	91
Proposed with SSPCA (10%)	**94**	**94.5**	**93.5**	**95.5**

Table 2. Results of the classification rate for rice images based on the HOG feature for neural network, support vector machine, and proposed dictionary-based classifier with the SPCA and SSPCA techniques.

	Tarom rice	Shiroodi rice	Fajr rice	Binam rice
Support vector machine [18]	82.5	83	88.5	85
Feed-forward neural network [19]	83	84.5	89	86
Dictionary-based classifier (Proposed)	90	92.5	91	93.5
Proposed with SPCA (10%)	92	93	92.5	94
Proposed with SSPCA (10%)	**94**	**94.5**	**93.5**	**95.5**

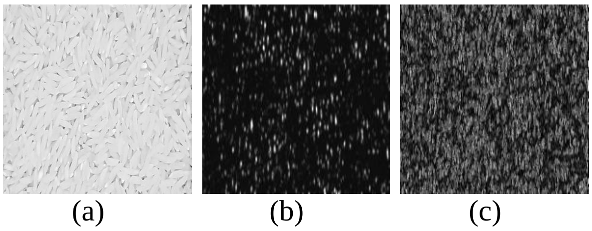


Figure 6. a) Tarom rice image b) Magnitude of the Gabor coefficient matrix c) Phase of the Gabor coefficient matrix.

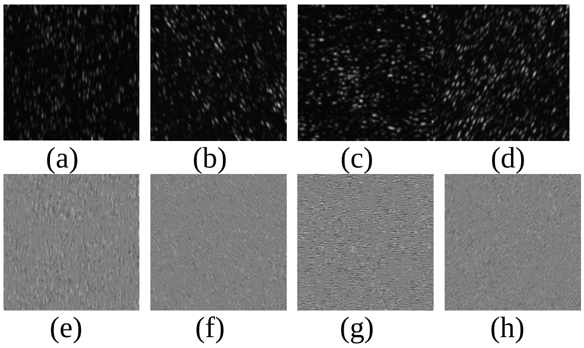


Figure 7. A representation of the magnitude of the Gabor coefficient in the directions a) 0° b) 45° c) 90° d) 135°, and the phase of the Gabor coefficient in the directions e) 0° f) 45° g) 90° h) 135°.

Table 3. Results of the classification rate based on the magnitude of the Gabor features for the neural network, support vector machine, and proposed learning-based classifier with the SPCA and SSPCA techniques.

	Tarom rice	Shiroodi rice	Fajr rice	Binam rice
Support vector machine [18]	92.5	93	92.5	94
Feed-forward neural network [19]	94	95	94.5	93.5
Dictionary-based classifier (Proposed)	96	97.5	96	96
Proposed with SPCA (10%)	96.5	97	96.5	96.5
Proposed with SSPCA (10%)	**98**	**98.5**	**97**	**97.5**

Table 4. Results of the classification rate based on the magnitude and phase of the Gabor coefficients for different classifiers and the proposed learning-based classifier with the SPCA and SSPCA techniques.

	Tarom rice	Shiroodi rice	Fajr rice	Binam rice
Support vector machine [18]	76	72	74.5	73.5
Feed-forward neural network [19]	81.5	80	82.5	80.5
Dictionary-based classifier (Proposed)	87	86.5	89	86.5
Proposed with SPCA (10%)	89	88.5	87.5	88.5
Proposed with SSPCA (10%)	**90**	**91**	**91.5**	**90.5**

5.4. Feature extraction based on combinational feature coefficients

In this section, the feature extraction process is investigated using the GLCM coefficients and its captured parameters. The selected parameters from the GLCM feature vector are nine characteristics including the mean, variance, energy, range of relative abundances, contrast, homogeneity, maximum relative frequency, correlation, and entropy. The GLCM coefficients are calculated in the four directions of 0˚, 45˚, 90˚, and 135˚. Then the mentioned parameters in these directions are calculated. As a result, a feature vector with 36 coefficients is obtained for the input image. The simulation results show that this dimension is not sufficient for classification and does not obtain a satisfactory classification rate. Therefore, this extracted feature vector should be used in combination with the other features. Among these features, the coefficients derived from HOG with a dimension of 1296 for each image data, a local binary pattern LBP with a dimension of 10 for each image data, and moments with a dimension of 7 for each image data introduced in Section 3 is applied. The results of this combination to solve the rice classification problem are presented in table 6.

The dimension of the training matrix using the GLCM/HOG, GLCM/HOG/LBP, and GLCM/HOG/LBP/Moment combinational properties for each rice class are 1305 × 400, 1315 × 400, and 1322 × 400, respectively. The cardinality rate for these features in the training and test steps is set to 30, 35, and 35, respectively. These results show that using the statistical characteristics such as the extracted parameters from the GLCM coefficients along with other rotation-invariant features such as LBP and moments leads to the desired results in the rice classification problem. The result of calculating the relative confusion matrix using the combinational feature is reported in table 7. In order to consider the performance of the proposed algorithm in comparison with other mentioned classifiers, a statistical significance test has been employed. The Friedman test with the Holms post hoc test is used to compare the results of more than two methods [47-48]. This test investigates the results of all classifiers (neural network, support machine, and learning-based technique with the SPCA and SSPCA dimension reduction algorithms) for the five methods considered in table 2, five methods in table 3, five methods considered in table 4, sixteen methods in table 5, and twelve methods in table 6.

Table 5. Results of the classification rate for rice images based on the DFT and DCT coefficients as well as these coefficients in combination with the PCA algorithm for the neural network classifications, support vector machine, and proposed dictionary-based classifier with the SPCA and SSPCA techniques.

		Support vector machine [18]	Feed-forward neural network [19]	Dictionary-based classifier (Proposed)	Proposed with SPCA (10%)	Proposed with SSPCA (10%)
DCT feature	Tarom rice	48	53	56	58	**59.5**
	Shiroodi rice	49.5	54	60	62	**64**
	Fajr rice	52	56	61.5	62.5	**63.5**
	Binam rice	56.5	55.5	63.5	64	**66**
DCT/HOG feature	Tarom rice	58.5	62	68.5	68	**69.5**
	Shiroodi rice	54.5	57	61.5	64.5	**66.5**
	Fajr rice	57.5	59	62	63.5	**65**
	Binam rice	63.5	61.5	65.5	66	**68.5**
DFT feature	Tarom rice	57	53.5	62.5	65	**68.5**
	Shiroodi rice	58.5	58.5	61.5	65.5	**68**
	Fajr rice	61	60.5	66	67.5	**70**
	Binam rice	61.5	62	64.5	65.5	**69**
DFT/HOG feature	Tarom rice	63	66.5	68.5	69.5	**70.5**
	Shiroodi rice	61.5	63	67	68.5	**71.5**
	Fajr rice	63	61.5	68.5	69.5	**72.5**
	Binam rice	64	62.5	70	71.5	**73**

Table 6. Results of the classification rate for rice images based on the combinational feature coefficients for the neural network classifications, support vector machine, and proposed dictionary-based classifier with the SPCA and SSPCA techniques.

		Support vector machine [18]	Feed-forward neural network [19]	Dictionary-based classifier (Proposed)	Proposed with SPCA (10%)	Proposed with SSPCA (10%)
GLCM/HOG	Tarom rice	90.5	90	94.5	95.5	**96.5**
	Shiroodi rice	91.5	92.5	96.5	96.5	**97**
	Fajr rice	92	91	95	96	**97.5**
	Binam rice	93.5	89.5	96.5	97	**98.5**
GLCM/HOG/LBP	Tarom rice	93.5	93	**100**	97	98.5
	Shiroodi rice	95	95.5	97.5	98.5	**100**
	Fajr rice	93.5	94.5	98.5	98	**99.5**
	Binam rice	94	92	98	98.5	**99**
GLCM/HOG/LBP /Moment	Tarom rice	95	94.5	**100**	99	99.5
	Shiroodi rice	94.5	93	98.5	99	**100**
	Fajr rice	96	96	100	100	**100**
	Binam rice	94.5	96.5	100	99.5	**100**

Table 7. Results of the relative confusion matrix based on the HOG feature for the proposed dictionary-based classifier with the SSPCA technique.

	Tarom rice	Shiroodi rice	Fajr rice	Binam rice
Tarom rice	**99.5**	0.5	0	0
Shiroodi rice	0	**100**	0	0
Fajr rice	0	0	**100**	0
Binam rice	0	0	0	**100**

The numbers of different methods and conditions in this test are $J = 5$ and $I = 43$, respectively. Also better results are obtained from this test if the number of conditions is higher than the number of the examined algorithms. This test is one of the best procedures to compare several methods in different conditions without the need to have initial assumptions.

In the first step of this test, the average performance rating of R_j for the $j-th$ method from the J method in I different conditions is calculated as follows:

$$R_j = \frac{1}{I}\sum_{i=1}^{I} r_{ij} \tag{10}$$

where, r_{ij} is the performance rank of the $j-th$ method in the $i-th$ test state. The lowest rate value in this statistical test obtains the method with the best performance. This significance test starts with a null-hypothesis where all the compared methods have the same performance, and then it should be proved that this assumption is wrong, and then the rank of different methods according to their efficiency is calculated [49]. This test starts with the definition of the critical value:

$$\chi_F^2 = \frac{12I}{J(J+1)}\left[\sum_{j=1}^{J} R_j^2 - \frac{J(J+1)^2}{4}\right] \tag{11}$$

where, J and I show the number of methods and different conditions, respectively. Also the modified statistical value of the Friedman test, which is based on the F distribution, is defined by $(J-1)$ and $(I-1) \times (J-1)$ degrees of freedom [46]:

$$F_F = (I-1)\chi_F^2 / \left(I(J-1)-\chi_F^2\right) \tag{12}$$

In this statistical test, the null-hypothesis is rejected if the F_F value is greater than the critical value χ_F^2. It means that the mentioned algorithms do not have the same performance. In the next step, the algorithm with a better performance is determined. This decision is made using the post-hoc test. In this case, at first, the Z_j parameter for each compared algorithm in this test is calculated [50]:

$$Z_j = (R_0 - R_j)/\sqrt{J(J+1)/6I} \tag{13}$$

where, R_0 is the method with the lowest performance level and the highest average rate according to (10). The values for the Z and ρ parameters are calculated at a statistical significance level of $\alpha = 0.05$. The value of ρ in the probability density of the standard normal distribution, corresponding to the underlying surface of this distribution, is outside the range ($-Z$, Z) [51]:

$$\rho = 1 - \int_{-Z}^{Z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \tag{14}$$

The simulation results of this statistical test using the Holm measure for the mentioned classification methods are reported in table 8 [49]. These results are calculated to determine the averaging rank for different classifiers based on the various feature vectors. The value of χ_F^2 for the results of this table according to (11) is 29.58. Also the value of F_F based on (12) is 25.46. The critical F value with the degree of freedom (5-1) and (5-1) × (43-1) is 2.4254 [50].

The reported results show that all the F_F values achieved in this statistical test have a value more than the F parameter, and therefore, the initial null hypothesis using the performance similarity between all the mentioned algorithms is rejected. Therefore, the post-hoc test can be employed to determine the performance rank of different algorithms. As mentioned, the rank R_0 belongs to the method with the lowest performance level and the highest average rank between the mentioned methods. This rank in the statistical test corresponds to the support vector machine. In order to calculate the Z parameter to calculate the ρ-value at a statistical significance level $\alpha = 0.05$, (13) is used. Also the value for the γ parameter in $\left(\alpha/\left(J\text{-}\gamma\right)\right)$ of table 8 is related to the row number and start from 1, which includes a method with a better performance and the lowest average rank up to $J-1$ that means the method with the lowest performance and higher average rank value. The results of the last row in table 8 are not reported since the comparisons are based on the method with a lower performance according to (13). As it can be seen, the ρ-values in this table are lower than the corresponding Holm critical values for all methods. Therefore, it can be concluded that the algorithms are listed in table 8 from the best performance of the method with the worst performance in the rice classification problem. The results of the statistical test for the proposed algorithm based on SSPCA with a 10% dimension reduction rate have a rank value that is close to 1. It means that this algorithm has a better performance than the other methods. Since the support vector machine has an average rate R_0 and the value of Z parameter for this classifier is set to zero, the ρ parameter cannot be calculated for it.

Another problem discussed in this paper is the detection of the rice quality. The last step in the rice classification is consideration of the rice impurity percentage. The Tarom rice has a higher quality and price between the rice varieties mentioned in this paper. Therefore, mainly in the rice trade market, only this product may be combined with other rice types that have a lower quality and price. Therefore, in the proposed procedure in this paper, at first, different sets of the Tarom rice were combined with 5%, 10%, 15%, and 20% of a mixture of other rice varieties. Obviously, the combination of more than 20% with the main rice is not common due to the fundamental changes in the appearance of rice that is easily visible by the eye. After classification of the rice type, the purity or impurity of the rice type is considered. This simulation is carried out based on the extracted combinational feature GLCM/HOG/ LBP/Moment that obtains the proper results in the previous investigation.

Table 8. Results of the statistical test to compare the performance of the neural network, support vector machine, and proposed dictionary-based classifier to categorize the rice images based on the different features.

	R_j	Z	ρ -value	Holm $\left(\alpha/\left(J-\gamma\right)\right)$
Proposed with SSPCA (10%)	1.04	5.1236	0	0.0125
Dictionary-based classifier (Proposed)	1.23	4.8969	0	0.0167
Proposed with SPCA (10%)	1.83	2.7432	0.0061	0.0250
Feed-forward neural network [18]	2.06	2.4985	0.0125	0.0500
Support vector machine [19]	2.56	0	-	-

Table 9. Results of the accuracy rate for the Tarom rice classification and detection of its purity or impurity based on the combinational features for the neural network, support vector machine, and proposed dictionary-based classifier with SSPCA (10%).

	Support vector machine [18]	**Feed-forward neural network [19]**	**Proposed**	**Proposed with SSPCA (10%)**
Pure Tarom rice	94	94.5	**100**	99.5
Impure Tarom rice 5%	93.5	94.5	98.5	**99**
Impure Tarom rice 10%	92	93.5	98	**98.5**
Impure Tarom rice 15%	95	92.5	98.5	**100**
Impure Tarom rice 20%	94.5	93	**99**	**99**

The results of the proposed algorithm for detection of the rice quality are shown in table 9, which emphasize on the accurate performance of the proposed dictionary-based classifier with the SSPCA dimension reduction algorithm.

6. Conclusion

One of the most widely used areas in the field of rice trade, especially in the north of Iran, is rice classification and detection of its quality. This paper presents a novel algorithm to solve this problem. In the proposed method, the sparse coding and dictionary learning algorithms were used to train the over-complete models. These techniques were used in the proposed classifier procedure. Various statistical and texture-based features were applied in the classification procedure, and the calculated results were compared with each other. Also due to the high dimension of the problem, the dimension reduction technique was investigated to result in a better classification rate. These techniques involve the SPCA and SSPCA algorithms. The simulations show that the proposed dictionary-based classifier with the SSPCA algorithm precisely detects the rice type and achieves the appropriate classification rate.

Reference

[1] Fao. (2012). Rice Market Monitor. vol. XV.

[2] Neelamma, K. P., Virendra, S. M., & Ravi, M. Y. (2011). Color and texture based identification and classification of food grains using different color models and haralick features. International Journal on Computer Science and Engineering (IJCSE), vol. 3, pp. 3669-3680.

[3] Silva, C. S., & Sonnadara, U. (2013). Classification of rice grains using neural networks. Proceedings of Technical Sessions; Sri Lanka, Institute of Physics, pp. 9-14.

[4] Philip, M. F., Mary, T., & Anita, H. B. (2017). Rice Grain Classification using Fourier Transform and Morphological Features. Indian Journal of Science and Technology, vol. 10, no. 14, pp. 1-6.

[5] Wah, T. N., San, P. E., & Hlaing, T. (2018). Analysis on Feature Extraction and Classification of Rice kernels for Myanmar Rice Using Image Processing Techniques. International Journal of Scientific and Research Publications (IJSRP), vol. 8, no. 8, pp. 603-606.

[6] Majumdar, S., & Jayas, D. S. (2000). Classification of cereal grains using machine vision: I. Morphology models. Transactions of the ASAE, vol. 43, no. 6, pp. 1669-1675.

[7] Visen, N. S., Jayas, D. S., Paliwal, J., & White, N. D. G. (2004). Comparison of two neural network architectures for classification of singulated cereal grains. Journal of Canadian Biosystem Engineering, vol. 46, pp. 7-14.

[8] Paliwal, J., Borhan, M. S., & Jayas, D. S. (2004). Classification of cereal grains using a flatbed scanner. Canadian Biosystems Engineering, vol. 46, no. 3, pp. 1-35.

[9] Liu, Z. y., Cheng, F., Ying, Y. b., & Rao, X. Q. (2005). Identification of rice seed varieties using neural network. Journal of Zhejiang University. Science.B, vol. 6, no. 11, pp. 1095-1100.

[10] Verma, B. (2010). Image processing techniques for grading & classification of rice. In International conferences on computer and communication technology, India, pp. 220-223.

[11] Jinorose, M., Prachayawarakorn, S., & Soponronnarit, S. (2010). Development of a computer vision system and novel evaluation criteria to characterize color and appearance of rice. Drying Technology, vol. 28, pp. 1118-1124.

[12] Dalen, G. V. (2004). Determination of the size distribution and percentage of broken kernels of rice using flatbed scanning and image analysis. Food research international, vol. 37, pp. 51-58.

[13] Pabamalie, L., & Premaratne, H. (2011). An intelligent rice quality classifier. International Journal of Internet Technology and Secured Transactions, vol. 3, pp. 386-406.

[14] Hatami, M., Rahmanididar, A., & Khazaee, J. (2010). Identification of common Iranian rice varieties using machine vision techniques. Sixth National Congress of Agricultural Machinery Engineering and Mechanization, pp. 1-9.

[15] Faiazi, S., Abaspourfard, M. H., Monajemi, S. A., Sadrnia, H., & Rouhani, A. (2013) .Identification and separation of three Iranian rice products in mixed samples using texture based feature and LVQ. Journal of Agricultural Mechanization, vol. 1, pp. 35-44.

[16] Patel, V. A., & Joshi, M. V. (2018). Convolutional neural network with transfer learning for rice type classification. In 10th International Conference on Machine Vision (ICMV), Vienna, Austria, pp. 1-8.

[17] Mousavirad, F. A., & Mollazade, K. (2012). Real time identification of rice variaties using texture features based on classifier fusion method. In the 16th CSI International Symposium on Artificial Intelligence and Signal Processing, Shiraz University, Iran.

[18] Mousavirad, F. A., & Mollazade, K. (2012). Identification of rice variaties using co-occurence matrix features of bulk samples and support vector machine. Presented at the 20th Iranian Conference on Electrical Engineering, University of Tehran, Teharn, Iran.

[19] Golpour, I., Amiripanah, J., Amirichaichan, R., & Khazaee, J. (2015). Detection of rice varieties, brown rice and white rice based on image and artificial neural network. Journal of Agricultural Machinery, vol. 1, no. 5, pp. 73-81.

[20] Golpour, I., Parian, J. A., & Chayjan, R. A. (2014). Identification and classification of bulk paddy, brown, and white rice cultivars with colour features extraction using image analysis and neural network. Czech J. Food Sci., vol. 32, pp. 280-287.

[21] Davis, G., Mallat, S., & Zhang, Z. (1994). Adaptive time-frequency decompositions. Optical-Engineering, vol. 33, pp. 218-391.

[22] Agarwal, A., Anandkumar, A., Jain, P., Netrapalli, P., & Tandon, R. (2014). Learning Sparsely Used Overcomplete Dictionaries. JMLR: Workshop and Conference Proceedings, vol. 35, pp. 1-15.

[23] Lee, H., Battle, A., Raina, R., & Ng, A. Y. (2006). Efficient sparse coding algorithms. Advances in Neural Information Processing Systems.

[24] Mavaddati, S. (2019). A Novel Face Detection Method Based on Over-complete Incoherent Dictionary Learning. Journal of AI and Data Mining, vol. 7, no. 2, pp. 263-278.

[25] Aharon, M., Elad, M., & Bruckstein, A. (2006). K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans. Signal Process, vol. 54, no. 11, pp. 4311-4322.

[26] Engan, K., Aase, S. O., & Hakon-Husoy, J. H. (1999). Method of optimal directions for frame design. IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, pp. 2443-2446.

[27] Delgado, K., Murray, J. F., Rao, B. D., Engan, K., Lee, T., & Sejnowski, T. J. (2003). Dictionary learning algorithms for sparse representation. Neural Computation, vol. 15, no. 2, pp. 349-396.

[28] Pradeep, J., Srinivasan, E., & Himavanthi, S. (2012). Neural network based recognition system integrating feature extraction and classification for English handwritten. International Journal of Engineering Transactions B: Applications, vol. 25, no.2, pp. 99-106.

[29] Saleh, R., & Farsi, H. (2018). Optimum ensemble classification for fully polarimetric SAR data using global-local classification approach. IJE TRANSACTIONS B: Applications, vol. 31, no. 2, pp. 331-338.

[30] Singh, M. P. (2009). Handwritten character recognition using modified gradient descent technique of neural networks and representation of conjugate descent for training patterns. IJE TRANSACTIONS A: Basics, vol. 22, no. 2, pp.145-158.

[31] Nixon, M., & Aguado, A. (2012). Feature Extraction & Image Processing for Computer Vision. Third Edition, Academic Press.

[32] Mathur, N. P., & Gupta, S. K. (2017). Performance analysis of feature extraction techniques for facial expression recognition. International Journal of Computer Applications, vol. 166, no. 1, pp. 1-3.

[33] Nixon, M., & Aguado, A. (2008). Feature extraction & image processing. 2nd Edition, Academic Press, Cambridge.

[34] Ojala, T., & Pietikäinen, M. (1999). Unsupervised texture segmentation using feature distributions. Pattern recognition, vol. 32, no. 3, pp. 477-486.

[35] William, T., & Roth, M. (1994). Orientation histograms for hand gesture recognition. Tech. Rep. TR94-03, Mitsubishi Electric Research Laboratories, Cambridge, MA.

[36] Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 1, 886-893.

[37] Majumdar, S., & Jayas, D. S. (2000). Classification of cereal grains using machine vision: I. Morphology models. Transactions of the ASAE, vol. 43, no. 6, pp. 1669-1675.

[38] Flusser, J., Suk, T., & Zitová, B. (2009). Moments and moment invariants in pattern recognition. Hichester, West Sussex, U.K: J. Wiley.

[39] Jolliffe, I.T. (2002). Principal Component Analysis, Springer, 2nd edition.

[40] Zou, H., Hastie, T., & Tibshirani, R. (2006). Sparse principal component analysis. Journal of Computational and Graphical Statistics, vol. 15, pp. 265-286.

[41] R. Jenatton, Obozinski, G. & Bach, F. (2010). Structured Sparse Principal Component Analysis. In Proceedings of the 13th International Conference on Artificial Intelligence and Statistics, Sardinia, Italy.

[42] Jenatton, R., Audibert, J. Y., & Bach, F. (2009). Structured variable selection with sparsity-inducing norms. Technical report, arXiv: 0904.3523,

[43] Sigg, C. D., Dikk, T., & Buhmann, J. M. (2012). Speech enhancement using generative dictionary learning. IEEE Transactions on Audio, Speech and Language Processing, vol. 20, no. 6, pp. 1698-1712.

[44] Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. Ann. Stat, vol. 32, pp. 407-499.

[45] Tropp, J., Dhillon, I., Heath, R. J. & Strohmer, T. (2005). Designing structural tight frames via an alternating projection method. IEEE Trans. on Information Theory, vol. 51, no.1, pp. 188-209.

[46] Barchiesi, D. & Plumbley, M. D. (2013). Learning incoherent dictionaries for sparse approximation using iterative projections and rotations. IEEE Transactions on Signal Processing, vol. 61, no. 8, pp. 2055-2065.

[47] Benesty, J. (2008). Springer handbook of speech processing. Springer's publication, pp. 843-871.

[48] Demsar, J. (2006). Statistical comparisons of classifiers over multiple data set. The Journal of Machine Learning Research, vol. 7, pp. 1-30.

[49] Sheskin, D. J. (2000). Handbook of Parametric and Nonparametric Statistical Procedures. 4nd ed. Boca Raton, FL: Chapman & Hall/CRC.

[50]https://www.danielsoper.com/statcalc/calculator.aspx?id=4

[51]http://onlinestatbook.com/2/calculators/normal_dist.html

نشریه هوش مصنوعی و داده کاوی

ارائه یک روش جدید تشخیص نوع و کیفیت برنج به کمک تحلیل مولفه‌های اساسی ساختاریافته تنک و یادگیری مدل

سمیرا مودّتی *

دانشکده فنی و مهندسی، دانشگاه مازندران، بابلسر، ایران.

ارسال ۲۰۱۸/۰۸/۲۶؛ بازنگری ۲۰۱۹/۱۱/۰۹؛ پذیرش ۲۰۱۹/۱۲/۱۲

چکیده:

در علوم و حوزه‌های تجاری مربوط با کشاورزی مدرن، دسته‌بندی انواع دانه‌های برنج و تشخیص کیفیت آن‌ها بسیار حائز اهمیت است. در سال‌های اخیر، الگوریتم‌های مختلف پردازش تصویر به منظور تشخیص انواع محصولات کشاورزی بکار گرفته شده‌اند. در این مقاله، مسئله دسته‌بندی دانه برنج و تشخیص کیفیت دانه‌ها براساس مفاهیم یادگیری مدل شامل بازنمایی تنک و شیوه‌های آموزش واژه‌نامه به منظور دستیابی مدل‌های فراکامل صورت می‌گیرد. به این منظور ویژگی‌های تصویری مبتنی‌بر رنگ، آمار و یا بافت برای تعیین محتوای ساختاری انواع برنج مورد استفاده قرار می‌گیرند. برای دستیابی به نتایج مطلوب، ویژگی‌های مختلفی از تصاویر ضبط شده استخراج و برای یادگیری مدل‌های بازنمایی کننده از نمونه‌های برنج استفاده می‌گردد. همچنین از الگوریتم‌های تحلیل مولفه‌های اساسی تنک و تحلیل مولفه‌های اساسی ساختاریافته تنک به منظور کاهش بعد مسئله دسته‌بندی بهره گرفته می‌شود که موجب تشخیص دقیق‌تر و کاهش زمان محاسباتی می‌شود. نتایج دسته‌بند پیشنهادی براساس آموزش مدل با نتایج حاصل از دسته‌بندهای مبتنی‌بر شبکه عصبی و ماشین بردار پشتیبان مقایسه گردیده است. نتایج حاصل از طبقه‌بند مبتنی‌بر یادگیری مدل پیشنهادی به همراه تست معناداری آماری نشان می‌دهد که روش پیشنهادی با دقت بالا قادر به تشخیص نوع و تعیین کیفیت ارقام مختلف برنج می‌باشد.

کلمات کلیدی: دسته‌بندی برنج، تشخیص کیفیت، تحلیل مولفه‌های اساسی ساختاریافته تنک، یادگیری مدل، شبکه عصبی، تست آماری.