IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 19 July 2023, accepted 2 August 2023, date of publication 17 August 2023, date of current version 30 August 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3306075*

RESEARCH ARTICLE

# Weighted Constraint Feature Selection of Local Descriptor for Texture Image Classification

**ENTESSAR SAEED GEMEAY[1,2], FARHAN A. ALENIZI[3], ADIL HUSSEIN MOHAMMED[4], MOHAMMAD HOSSEIN SHAKOOR[5], AND REZA BOOSTANI[6]**

[1]Department of Computer Engineering, Computer and Information Technology College, Taif University, Taif 21944, Saudi Arabia
[2]Department of Electronics and Communication Engineering, College of Engineering, Tanta University, Tanta 31527, Egypt
[3]Electrical Engineering Department, College of Engineering, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[4]Department of Communication and Computer Engineering, Faculty of Engineering, Cihan University-Erbil, Kurdistan Region, Erbil 44001, Iraq
[5]Department of Computer Engineering, Faculty of Engineering, Arak University, Arak 38156-88349, Iran
[6]Computer Engineering Department, School of Electrical and Computer Engineering, Shiraz University, Shiraz 71348-14336, Iran

Corresponding author: Mohammad Hossein Shakoor (mh-shakoor@araku.ac.ir)

**ABSTRACT** There are several statistical descriptors for feature extraction from texture images. Local binary pattern is one of the most popular descriptors for revealing the underlying structure of a texture. Recently several variants of local binary descriptors have been proposed. The completed local binary pattern is an efficient version that can provide discriminant features and consequently provide a high classification rate. It finely characterizes a texture by fusing three histograms of features. Fusing histograms is applied by jointing the histograms and it increases the feature number significantly; therefore, in this paper, a weighted constraint feature selection approach is proposed to select a very small number of features without any degradation in classification accuracy. It significantly enhances the classification rate by using a very low number of informative features. The proposed feature selection approach is a filter-based feature selection. It employed a weighted constraint score for each feature. After ranking the features, a threshold estimation method is proposed to select the most discriminant features. For a better comparison, a wide range of different datasets is used as a benchmark to assess the compared methods. Implementations on Outex, UIUC, CUReT, MeasTex, Brodatz, Virus, Coral Reef, and ORL face datasets indicate that the proposed method can provide high classification accuracy without any learning step just by selecting a few features of the descriptor.

**INDEX TERMS** Local binary pattern, weighted constraint feature selection, texture image classification.

## I. INTRODUCTION

Texture images usually include material and objects that present texture information. They can be regarded as any visual non-regular or irregular patterns. Texture image processing is a popular area in many applications of image processing recently. One of the important processing of textures is related to textures classification. It is an important part of image processing that is very useful in many applications. Some applications such as the detection of defect fabric [1], [2], medical image processing [3], remote sensing [4], face recognition [5], and image retrieval [6] are related to texture classification.

The key point of texture processing is related to extracting a set of discriminative features. In the last two decades, many feature extraction approaches [7], [8], [9] have been introduced that are used for texture images for classification [10], [11], [12], [13], [14] or segmentation [15], [16]. Most of these techniques employed statistical local descriptors for texture analysis. Statistical models are some types of techniques that extract features for all kinds of textures. Furthermore, it is possible to combine these descriptors with some other local descriptors to provide more discriminative features for classification. Haralick methods are some of the statistical feature extraction techniques based on matrix processing. For

The associate editor coordinating the review of this manuscript and approving it for publication was Senthil Kumar.

example, GLCM [17], GLDM, GLRLM, and SGLDM [18], [19] were proposed for the extraction of features from texture. These methods are simple however, they are sensitive to changes in illumination and rotation.

One of the popular and simple techniques for texture feature descriptors is Local Binary Patterns (LBP). Ojala et al. proposed LBP [20] for the first time. It is a basic operator to describe local patterns and achieved a high classification rate for texture images [21]. The main goal of proposing LBP is related to texture classification. However, LBP can be used in different applications, such as shape localization [22], face recognition [23], texture recognition [24], and medical image analysis [25]. LBP is a local descriptor that provides noise-sensitive features. It extracts too many features that are sensitive to rotation. So Ojala et al. proposed some robust and rotation-invariant LBP versions (rotation-invariant, robust to noise) [26], [27].

After proposing the first version of LBP, some extended LBP variants [28], [29] were proposed. They could provide more discriminant features for textures. Furthermore, these variants can extract more features from each image. The number of features was raised exponentially by increasing the neighborhoods of LBP. Many other methods are proposed based on LBP. Ojala et al. combined the LBP and variance operator and LBP/VAR [16] to prepare more discriminant features. Some researchers [30] proposed Local Ternary Pattern (LTP). It can quantize the difference between each center and its neighborhood values into three levels. Some versions of LBP, such as derivative-based LBP [22] used derivative patterns for feature extraction. Liao et al. proposed dominant LBP (DLBP) [31]. The DLBP method used the most frequently occurring patterns. It includes uniform and non-uniform local patterns that describe textural information. It can select the most frequent patterns of texture. Center-symmetric LBP [32] is another type of LBP. It was proposed to decline the number of LBP features. Local Derivative Patterns (LDP) [33] is another type of this group of descriptors that used high order form difference of LBP. It could capture more detailed features than first-order LBP, but it was sensitive to noise. Despite the development of LBP in texture classification and pattern recognition, its fundamental working mechanism requires further study and investigation [34].

In this paper, the CLBP [34] method is used because it achieves high accuracy in texture classification. CLBP provides high features number, especially for large neighborhoods. In this research, the feature number is decreased in two parts, the first part is related to mapping or merging the features into a smaller number of features and the second part includes the feature selection. Some mapping approaches [35] such as riu2 decrease the features by using only a uniform pattern. Although CLBP achieves a high classification rate, it joints the LBP_S, LBP_M, and LBP_C histograms by the 3D jointing method, therefore, the feature number grows significantly. Therefore, a dimensionality reduction method is needed to address this problem. In other words, it is necessary to apply a feature selection approach to decrease the irrelevant features of CLBP. Feature selection is an important step in computer vision and image processing applications which deletes irrelevant and redundant features and improves learning performance [36], [37], [38], [39]. It can be used to reduce the size of big data [40] and remove irrelevant features [41]. Some approaches were proposed to select features for the histogram of LBP [42], [43]. Applying feature selection not only declines the number of features but also raises the classification rate significantly. Feature selection is an approach to finding a set of the most discriminative features [44].

There are three groups of feature selection, Filter, Wrapper, and Embedded techniques. The wrapper method selects features based on the result of classification accuracy. Therefore, it optimizes the classification rate and includes the train step. Filter-type methods select variables regardless of the model. They are based only on general features like the correlation with the variable to predict. The third class, embedded methods, is similar to wrapper methods. The difference to wrapper methods is that an intrinsic model-building metric is used during learning. They have been proposed to combine the advantages of both filter and wrapper methods [45]. Feature selection can be divided into supervised and unsupervised approaches. Supervised feature selection techniques use labeled data for feature selection and evaluate feature relevance by measuring the feature correlation or information gain [46], [47], [48]. Unsupervised feature selection methods estimate feature relevance by the capability of keeping particular attributes of the data, such as the variance or the locality-preserving ability [49], [50].

Some methods provided more discriminative and robust features of texture images by decreasing the extracted features. One type of these methods identified the most discriminative local pattern elements based on some information about textures. These approaches prepare mapping or merging techniques for local binary features [26], [27]. The second group of these methods decreases the feature number by selecting a discriminative set of features from the entire of them. Some researchers used feature selection for color textures [42]. They combined some color channels and selected features from a set of color textures. Porebski et al. proposed a novel approach to select the most discriminant features from the histogram of local patterns [51]. Some papers proposed a supervised feature selection based on the Laplacian score [43], [52]. The Laplacian score is used for feature selection for histogram bins as features. In another approach, feature selection used a block-based Local Binary Pattern for face recognition [53]

In this paper, a score-based feature selection method is proposed. The main goal of this paper is related to the proposed framework to achieve a high classification rate only by a very small number of features. It is because of some limitations of memory in some applications. According to the

implementation results of the proposed method, it is possible to reach a high accuracy only by using a small number of best features. The results indicate that by using 3, 5,or 10 features from the entire of all, it is possible to record a classification rate that is only 5 or 10% lower than the best rate with all features.

The contribution of this paper is related proposing a novel feature selection method for selecting the completed local binary pattern features. It combines the constraint and fisher feature selection methods and provides a new method that named weighted constraint feature selection (WCFS). The implementation shows that the features that selected by the proposed method are more discriminant than Entropy, Fisher, and Constraint approaches. The constraint method is used as a semi-supervised technique for feature selection [42]. However, in this paper, it is used as a supervised method, and by using, the Fisher score as a weight of constraint scores it can be possible to increase the classification rate by selecting a few CLBP features. Furthermore, a new method is proposed to estimate the threshold value. This value is used for selecting the ranked features. In other words, it indicates the best cut of value for ranked features and the best selection of them. In addition, an enhancement part is proposed at the end of the feature selection part to remove or append some features for better selection.

The paper is organized as follows: in the next section LBP and CLBP are explained. Section three presents the proposed method. Experimental results and conclusion part are reported in sections four and five respectively.

## II. RELATED WORK

In this part, a summary of some local descriptors is discussed. In addition, a review of feature selection approaches that are related to the proposed method is presented.

### A. LOCAL BINARY PATTERN (LBP)

The LBP [20] is a powerful tool to describe texture characteristics and has been used frequently in many applications of machine vision [54]. It provides a binary string to compare P neighbor values with the center point. It generates a binary value of 1 if the value of the neighborhood pixel is greater or equal to the center pixel; otherwise, it is set to 0. Next, each binary bit is multiplied by a weight value, and the LBP code is generated. This value is calculated as follows in Eq. (1). If the center point of a patch is $g_c$ and $g_i$ is the point value of the i-th neighborhood, P is the number of neighbor points and R is the radius of the local patch.

$$LBP_{P,R(x,y)} = \sum_{i=0}^{P-1} s\left(g_i - g_c\right) 2^i \tag{1}$$

$$s\left(g_i - g_c\right) = \begin{cases} 1 \; g_i \geq g_c \\ 0 \; g_i < g_c \end{cases} \tag{2}$$

Figure 1 illustrates the calculation steps of LBP. Occurring to this figure, the square shape neighborhood is employed. The

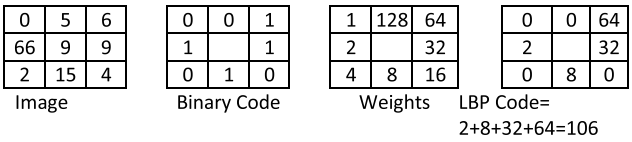


**FIGURE 1.** Example of the LBP code.

square shape patch is not rotation invariant. There are several versions of LBP extract the textural features of images. Most of these methods are not rotation invariant. To make a rotation invariant descriptor, the original LBP extended to a circular symmetric neighbor set of P values on a circular patch. It has a radius of R and uses uniform patterns for exploiting the local information [39]. For circular neighbor point values, an interpolation operation for each neighbor of the square patch must be applied. The uniform LBP with rotation invariant specification can be obtained as it is mentioned in (Eq. (3)–(5)): In these relations, riu2 illustrates the rotation invariant uniform patterns. They have a U value of at most 2. U is estimated as the uniformity of texture. It is equal to the number of transitions, i.e., bitwise 0/1 changes between successive bits in the neighborhood of each patch.

$$LBP^{riu2}_{P,R}(x, y) = \begin{cases} \sum_{i=0}^{P-1} s\left(g_i - g_c\right) \text{ if } U\left(LBP_{P,R}\right) \leq 2 \\ P + 1 \text{ otherwise} \end{cases} \tag{3}$$

$$s\left(g_i - g_c\right) = \begin{cases} 1 \; g_i \geq g_c \\ 0 \; g_i < g_c \end{cases} \tag{4}$$

$$U\left(LBP_{P,R}\right) = \left|s\left(g_{P-1} - g_c\right) - s\left(g_0 - g_c\right)\right| + \sum_{i=1}^{P} \left|s\left(g_i - g_c\right) - s(g_{i-1} - g_c)\right| \tag{5}$$

### B. COMPLETED LOCAL BINARY PATTERN (CLBP)

The Complete LBP (CLBP) [34] is proposed as a variant of LBP. It utilizes local difference sign and magnitude as LBP_S and LBP_M. LBP_S is shown in Eq. 3. LBP_M is similar to LBP_S however, it compares the absolute difference magnitude of the center value and each neighbor value to a threshold value. In addition, for each local patch, the center value is compared to the entire image average and LBP_C is made. The LBP utilizes only the sign vector to represent the local patterns. This may lead to some incorrect classification results. The sign components give higher accuracy than the magnitude components. Besides, using sign and magnitude components together and combining them gives much better results. The sign components preserve plenty of local details. However, the magnitude components may provide extra discriminant patterns. CLBP is prepared by jointing three histograms of LBP_S, LBP_M, and LBP_C as a 3D joint histogram. It is possible to combine these features (i.e. CLBP_S, CLBP_M, and CLBP_C) by

concatenation (i.e. CLBP_S_M_C) or joint (CLBP_S/M/C) of these histograms. In addition, it is possible to provide two or three combinations of them [34], the best features are prepared by CLBP_S/M/C. Therefore, in this research and other papers [34], [35] CLBP_S/M/C illustrates CLBP. The below relations indicate the CLBP method. In this relation, LBP_S is the same as LBP and LBP_M calculates the magnitude of LBP. $m_i$ is the magnitude of each neighborhood and c is the average value of all $m_i$ Also the Avg value in LBP_C is the average value of the entire of each image.

$$\mathrm{LBP\_S}^{\mathrm{riu2}}_{\mathrm{P,R}}(x,y) = \begin{cases} \sum_{i=0}^{P-1} s(g_i - g_c) \text{if } U\left(\mathrm{LBP\_S_{P,R}}\right) \le 2 \\ P+1 \text{ otherwise} \end{cases} \quad (6)$$

$$s(g_i - g_c) = \begin{cases} 1 g_i \ge g_c \\ 0 g_i < g_c \end{cases} \quad (7)$$

$$\mathrm{LBP}^{\mathrm{riu2}}_{\mathrm{M_{P,R}}} = \begin{cases} \sum_{i=0}^{P-1} t(m_i - c) \text{if } U\left(\mathrm{LBP\_M_{P,R}}\right) \le 2 \\ P+1 \text{ otherwise} \end{cases} \quad (8)$$

$$t(x-c) = \begin{cases} 1 x \ge c \\ 0 x < c \end{cases}$$

$$\mathrm{LBP\_C}(x,y) = t(g_c, \mathrm{Avg}) \quad (9)$$

### C. SCORE-BASED AND FILTER METHOD FOR FEATURE SELECTION

In this part, several supervised or semi-supervised score functions are explained that are used for feature selection as a filter method, including constraint score [37], Fisher score [55], information gain, or entropy-based score [56].

#### 1) FISHER SCORE

Fisher score [55] is a supervised feature selection method that seeks the features with the best discriminant ability. The basic idea of Fisher's score is to maximize distances between data samples in the different classes and minimizes the distances between data samples in the same class. It is illustrated in Eq. (10). In this equation $\mu_i^r$ and $\sigma_i^r$ show the average and variance values of the r-th feature from the i-th class of data. $\mu^r$ indicates the average value of the r-th feature for all of the samples and $n_i$ shows the number of samples of the i-th class.

$$Fisher^r = \frac{\sum_{i=1}^{n} n_i(\mu_i^r - \mu^r)^2}{\sum_{i=1}^{n} n_i(\sigma_i^r)^2} \quad (10)$$

#### 2) CONSTRAINT SCORE

The constraint method is a supervised method that estimates the score of features such as Eq. 11. In this equation $X_i$ and $X_j$ are two sample values. If they belong to the same class they are members of M sets, otherwise, they belong to C. $F_i^r$ indicates the r-th feature value of the i-th sample. The sum of the magnitude difference of each feature such as $X_i$ and $X_j$ is estimated and used in this equation. Some papers proposed the score based on the absolute difference of each pair of features as a constraint method [49]. However, in this research, this relation is used as a part of the proposed supervised technique to determine the rank of the feature. In the feature selection step, each feature with a low score is removed. Eq. 11 shows the score estimation of the r-th feature.

$$Constraint^r = \frac{\sum_{(X_i,X_j)\in C} |F_i^r - F_j^r|}{1 + \sum_{(X_i,X_j)\in M} |F_i^r - F_j^r|} \quad (11)$$

#### 3) ENTROPY SCORE (INFORMATION GAIN)

Information Gain (IG) is an entropy-based feature evaluation method [56]. As information gain is used in feature selection, it is defined as the gain of each feature according to the label of data. It is illustrated by the entropy of the r feature in Eq. (12). The higher information gain the better feature can be selected for classification. In this relation, $F^r$ is related to values of the r-th feature and $y_i$ shows the label of the i-th class.

$$Entropy^r = Entropy(F^r) - \sum_{i=1}^{n} P(y_i).Entropy(F^r|y_i) \quad (12)$$

## III. PROPOSED METHOD

The proposed method is a score-based approach that selects features from the whole features of CLBP. It belongs to the filter selection methods. Such as some previous well-known papers [20], [34] the riu2 method is used for mapping the extracted features of CLBP into rotation-invariant and uniform patterns. However, some papers employed different mapping techniques for feature extraction [48].

As it is mentioned in the previous section, the filter approach is faster than the wrapper method. It selects a discriminative subset of features without a learning step. It includes two parts. First, a criterion score sorts the features. The criterion score may be based on the label of the sample or not. They are named as supervised and unsupervised respectively. Some popular supervised criteria were proposed before, such as variance [55], correlation [44], mutual information [57], [58], and consistency [59]. In the second part, high-rank features are selected and the remaining features are removed. If the rank criterion uses label information of samples for selection, it is named an unsupervised method. Some well-known unsupervised methods such as Variance [55], unsupervised Laplacian [60], and unsupervised sparsity score [61]. The techniques estimate the feature's relation by using some criteria such as information gain, distance, correlation, and dependency. In supervised approaches, class labels play a main role in feature selection [62].

The proposed approach is a filter-based method that uses label information for feature selection. Therefore, it belongs to the supervised and score-based techniques. It is explained here.

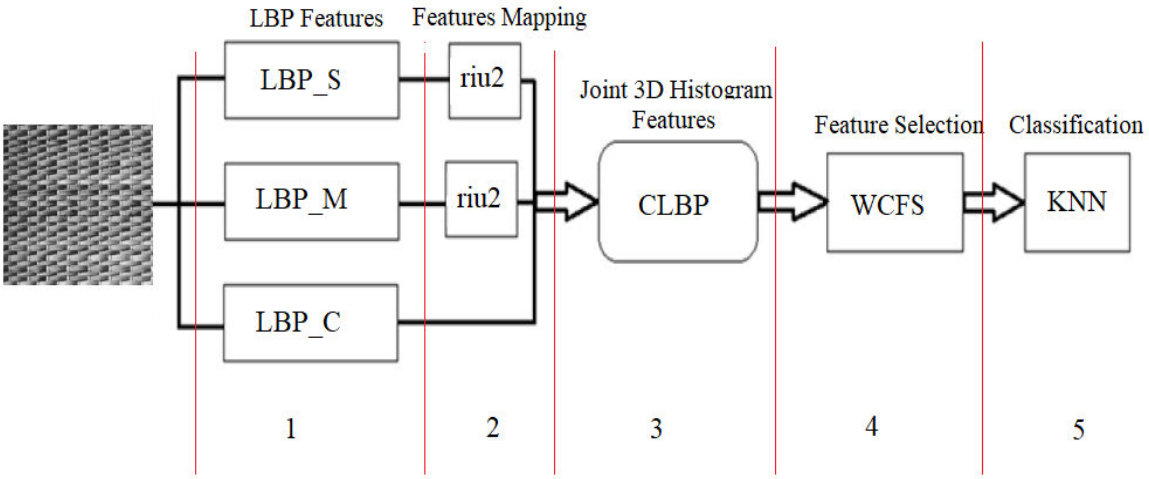


**FIGURE 2.** The block diagram of the proposed framework.

**TABLE 1.** The parameters of the proposed method.

| Parameter | Comment |
|---|---|
| R | The radius of each local patch of LBP |
| P | Number of points for each local patch of LBP |
| m | Number of original features |
| $F_i^r$ | r-th feature value of i-th sample |
| $F_j^r$ | r-th feature value of j-th sample |
| $n_C$ | Number of samples in class C |
| $n_M$ | Number of samples in class M |
| $\mu_C^r$ | The average value of r-th feature for data in class C |
| $\mu^r$ | The average value of r-th feature |
| $\sigma_M^r$ | The standard deviation of r-th feature for data in class M |
| $B_C^r$ | Between class variance of r-th feature for data in class C |
| $W_M^r$ | Within-class variance of r-th feature for data in class M |
| L | Fix coefficient of within class variance |
| K | Fix coefficient for threshold value |
| $IG_i^2$ | Information gain of i-th feature |
| $\overline{IG}$ | Average of information gain for all features |
| Tr | Threshold value to add or remove features |
| D(S, M) | The distance value of two samples S, and M of data for KNN classification |

### A. PROPOSED METHOD (WEIGHTED CONSTRAINT FEATURE SELECTION)

Here a supervised method is proposed. It proposed a weight of dissimilarity of each pair of features as a selection criterion.

Fig. 2 shows the block diagram of the entire framework. This diagram has some steps. In the first step, the LBP features are extracted by LBP_S, LBP_M,and LBP_C. As it is mentioned before LBP_S and LBP_M extracted 2^P features in each texture if each local patch has P points. Also, LBP_C extracted two features. In the next step (i.e. step 2) the feature number is reduced by the mapping method. As it is shown in relation 3 all of the non-uniform patterns are merged into one feature therefore by using riu2 mapping method 2^P features map into P+2 features. It means LBP_S provides P+2 features (also LBP_M). In the third step by the 3D joint of LBP_S, LBP_M,and LBP_C features, the CLBP features are obtained. By 3D jointing, the number of features increases into 2(P+2)(P+2) features.

The main contribution of this framework is related to step 4. In this step the best features are selected by the proposed method (WCFS) then these selected features are used for classification by KNN in step 5. Table 1 shows the parameters of proposed method. Also, all details and relations of the proposed method (WCFS) are illustrated in pseudo-codes of Table 2.

The proposed method estimates the score of the feature as it is shown in Eq. 13. It estimates the score of each feature based on maximizing the inter-class value and minimizing the intra-class value. In this equation, $X_i$ and $X_j$ are two samples. These samples include features such as $X_i = [F_i^1, F_i^2, \ldots F_i^m]$ and $X_j = \left[F_j^1, F_j^2, \ldots F_j^m\right]$.If they belong to the same class, it is indicated by M, else it is mentioned by C. Sum of the

**TABLE 2.** The pseudo-code of the proposed method for feature selection (WCFS).

1. For each feature such as r:
   Calculate the average of each feature ($\mu^r$)
   Calculate the average of each feature in each class such as C ($\mu_C^r$)
   Calculate the standard seviation of each feature in each class such as M ($\sigma_M^r$)
   End

2. For each Feature such as r (r from 1 to m):
   Set $BC^r$ =0 and $WC^r$=0
   For each pair samples such as $X_i = [F_i^1, F_i^2, \ldots F_i^m]$ and $X_j = [F_j^1, F_j^2, \ldots F_j^m]$
   If $X_i$ and $X_j$ belong to the different class such as C
   $BC^r = BC^r + (\mu_C^r - \mu^r)^2 \, |F_i^r - F_j^r|$
   Else ($X_i$ and $X_j$ belong to the same class such as M)
   $WC^r = WC^r + {\sigma_M^r}^2 \, |F_i^r - F_j^r|$
   End
   End
   $WC^r = BC^r - L.WC^r$
   End

3. Rank all of the features by WC score.
4. Calculate μ-σ of WC scores.
5. For all features (r=1:m)
   If $WC^r$ > μ-σ Then
   Add $F^r$ to FS1 subset.
   End
   End
6. Calculate Threshold T:
$$Tr = K \sqrt{\frac{m \sum_{i=1}^{m} IG_i^2 - (\sum_{i=1}^{m} \overline{IG})^2}{m(m-1)}}$$
7. For all features in FS1 subset (r=1:m1)
   If Information Gain of $F^r$ < Tr Then
   Remove $F^r$ from the FS1 subset.
   End
   End
8. For all features in FS - FS1 subset (i=1:m-m1)
   If Information Gain of $F^r \geq$ Tr Then
   Add $F^r$ to FS2 subset.
   End
   End
9. FS1+FS2 is the final subset of selected features for classification.

absolute difference for each pair of features is calculated and used for both similarity and dissimilarity. For each pair of samples, this relation estimates the score value for each feature such as r.

Here the numerator and denominator values of the Fisher method (i.e. Eq. 10) are employed as coefficients of the proposed method. In Eq. 13 $B_C^r$ is used as a weight of dissimilarity. It is a difference between the average of feature r in class C and a whole average of feature r. Furthermore, $W_M^r$ shows the similarity of features. It determines the within-class variance of class M for the r-th feature. Values $n_C$ and $n_M$ indicate the sample number of class C

**TABLE 3.** Specifications of some texture datasets that are used in this paper.

| Data | Subset of Data | Type | Number of classes | Sample per class | Size of image | Comments |
|---|---|---|---|---|---|---|
| Outex | TC3 | General textures | 24 | 20 | 128×128 | inca illumination |
| Outex | TC10 | General textures | 24 | 180 | 128×128 | inca illumination and 9 rotation angles |
| Outex | TC12(t) | General textures | 24 | 380 | 128×128 | t184 illumination |
| Outex | TC12(h) | General textures | 24 | 380 | 128×128 | horizon illumination |
| Outex | TC13 | General textures | 68 | 20 | 128×128 | 9 rotation angles |
| CUReT | - | General textures | 61 | 92 | 200×200 | small and similar pattern details |
| UIUC | - | General textures | 25 | 40 | 640×480 | large pattern details |
| Virus | - | Virus | 15 | 100 | 41×41 | low size and low contrast |
| MeasTex | - | General textures | 69 | 16 | 128×128 | high details textures |
| Brodatz | - | General textures | 112 | 16 | 128×128 | Sharp edges textures |
| ORL | - | Face | 40 | 10 | 128×128 | face texture |
| Coral Reef | RSMAS | Coral Reef | 8 | 48 | 256×256 | coral reef textures |

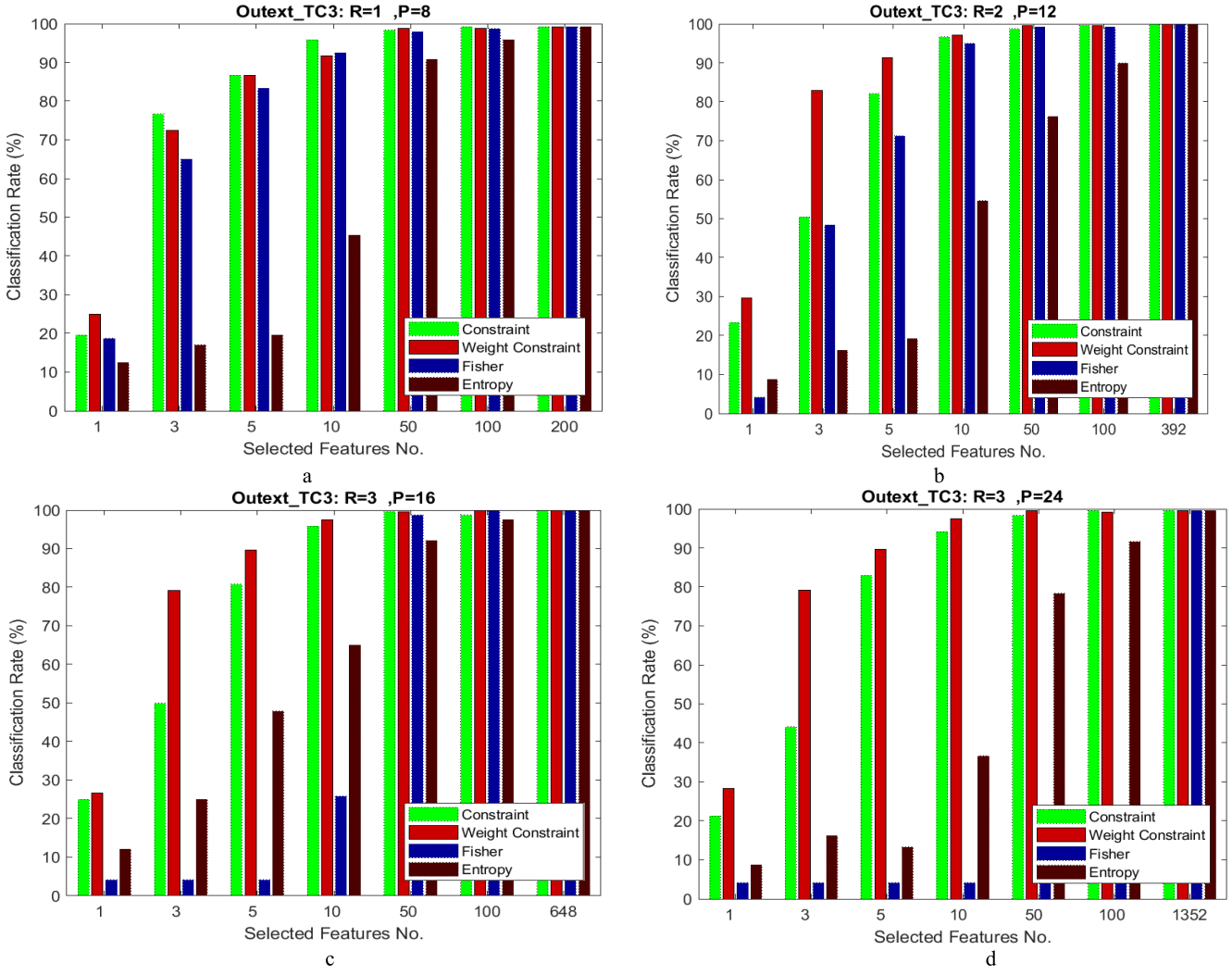


**FIGURE 3.** The results for Outex_TC3 texture for some local size.

and M respectively. In Eq. 13 the L, is a coefficient value that changes the percentage of similarity and dissimilarity of features for the estimation of the score. Here it is set by one. However, it can be trained and set for different datasets. The suitable value of L can be estimated by dividing the average values of $B^r_C.|F^r_i - F^r_j|$ by the average value

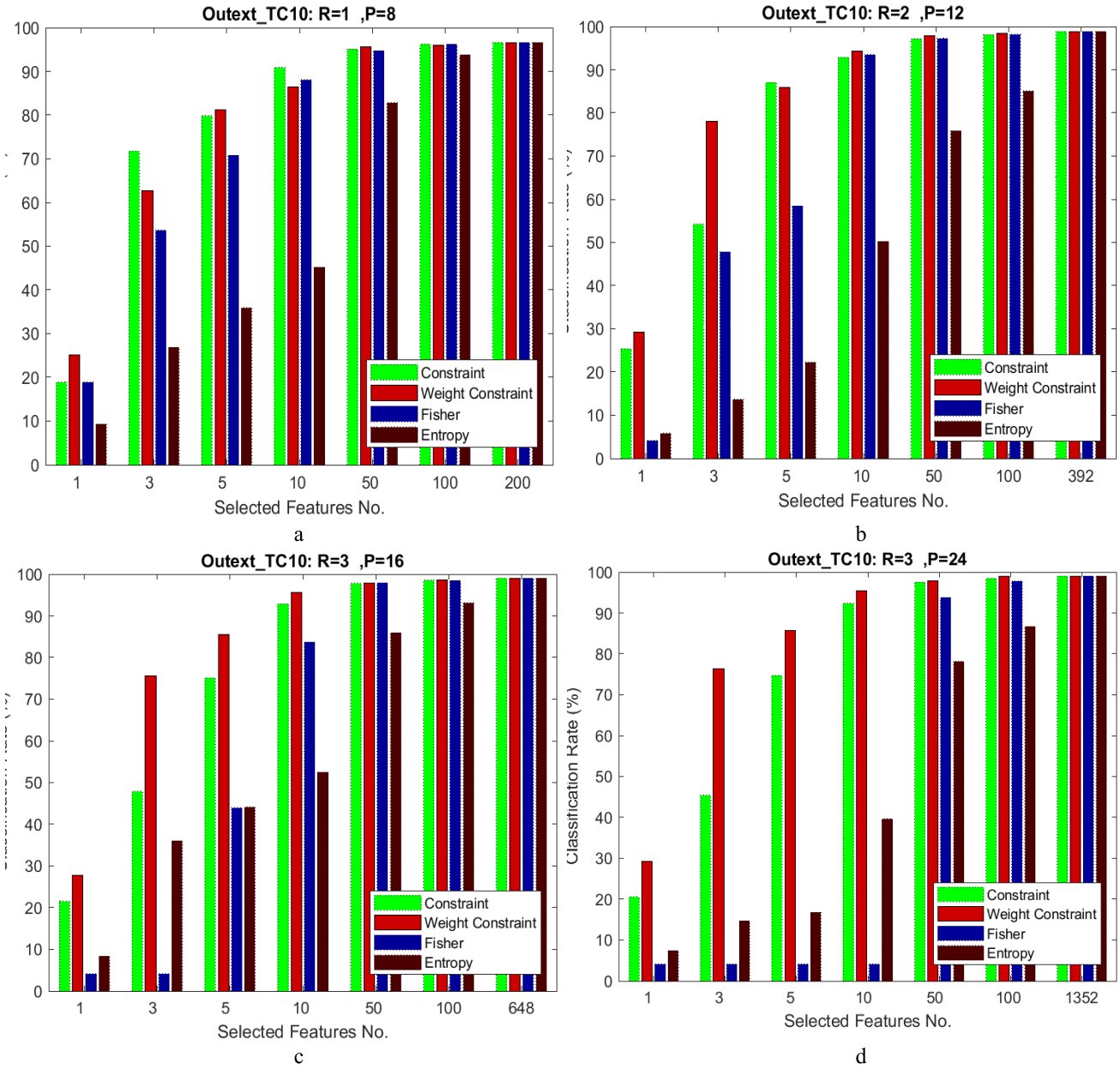


**FIGURE 4.** The results of Outex_TC10 texture for some local size.

of $W_M^r \cdot \left|F_i^r - F_j^r\right|$.

$$\begin{aligned} B_C^r &= n_C \left(\mu_C^r - \mu^r\right)^2 \\ W_M^r &= n_M \left(\sigma_M^r\right)^2 \\ &\text{Weighted\_Constraint}^r \\ &= \sum_{(X_i,X_j)\in C} B_C^r . |F_i^r - F_j^r| \\ &-L. \sum_{(X_i,X_j)\in M} W_M^r . |F_i^r - F_j^r| \end{aligned} \tag{13}$$

Table 2 shows the pseudo-code of the proposed framework for feature score, the threshold for selecting,and feature selection.

### B. PROPOSED METHOD (SET THRESHOLD FOR FEATURE SELECTION)

In the previous part a weighted constrain score is proposed to score each feature. According to this score, the features are sorted and selected. In this section, a threshold value is required to select high-score features and remove remain of them that their scores are lower than this threshold. In other words, it is necessary to determine the cutting value of the score to remove low-score features. Some papers [63] proposed an automatic threshold method for feature selection. It combines some data complexity measures to estimate the threshold value. In some research [48] the threshold value is set by using normal distribution values. Standard deviation

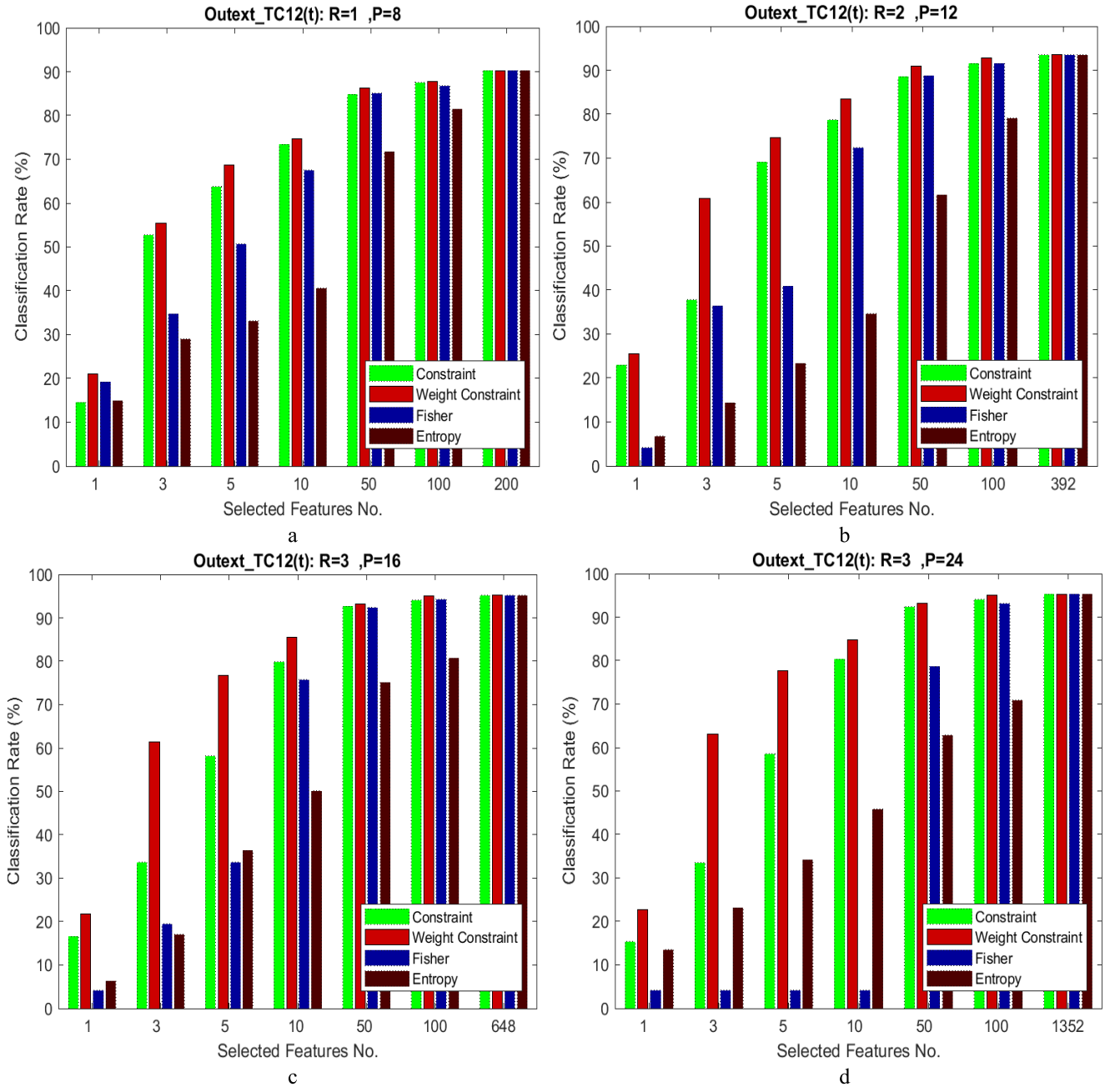


**FIGURE 5.** The results for Outex_TC12(t) texture for some local size.

($\sigma$) and average value ($\mu$) are two parameters of normal or Gaussian distribution.

In this section, a method for the threshold of feature selection is proposed. Some papers [64] introduced a technique that estimates the threshold based on information gain [56]. Information gain is a score of each feature that estimates the correlation between each feature and the label of the class based on entropy. It is shown in Eq. 8. Prasetiyowati et al. proposed [64] a threshold estimation method based on information gain. It is illustrated in Eq. 10. In this relation, m is the number of original features. The information gain for each feature i is calculated (i.e. Eq. 8) as a $IG_i$. Relation 10 estimates the threshold value Tr as a cutoff value for feature selection. $\bar{IG}$ is the average of all $IG_i$. In our paper a coefficient K ($0.5<K<1.5$) is multiplied by the threshold of the Prasetiyowati et al. [64] method. This value is depending on the dataset.

$$Tr = K\sqrt{\frac{m\sum_{i=1}^{m} IG_i^2 - \left(\sum_{i=1}^{m} \bar{IG}\right)^2}{m(m-1)}} \tag{14}$$

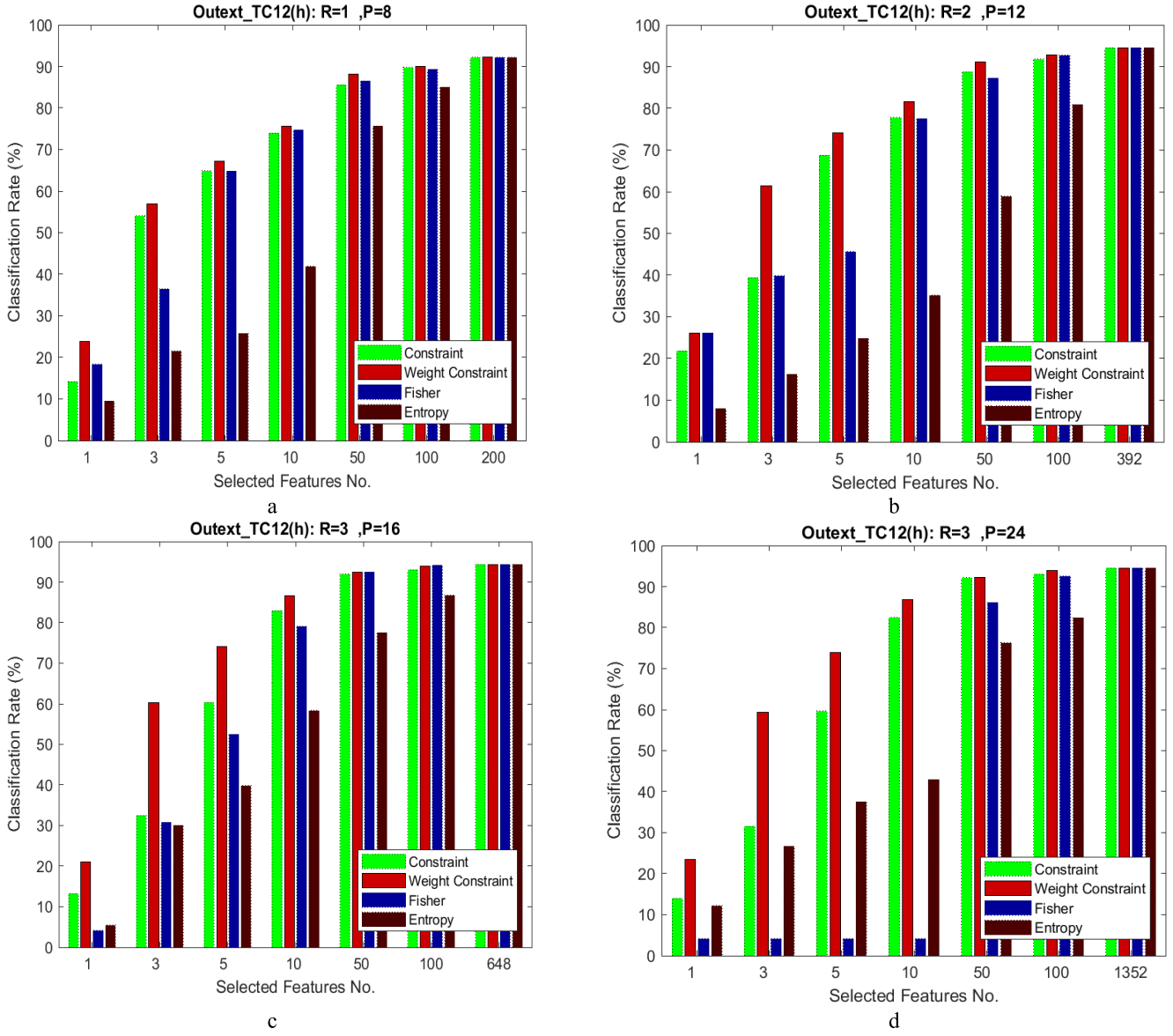


**FIGURE 6.** The results for Outex_TC12(h) texture for some local size.

### C. PROPOSED METHOD (SELECT MORE DISCRIMINATIVE FEATURES)

As it is mentioned in before two sections, first the features are extracted by using the CLBP descriptor. This descriptor prepares features by jointing 3 histograms of LB_S, LBP_M, and LBP_C and provides a high number of features, especially for large neighborhoods. In the second part of the proposed framework, a weighted constraint feature selection (WCFS) method is proposed that selects some of the features based on the proposed score. After scoring the features, an estimation of the threshold value is proposed. It selects features that have scored more than the threshold value and removes the remain of them. The third part of our proposed method is introduced here. In this part, some discriminative features that are removed by threshold value are selected and some irrelevant features that are selected in the previous part are deleted from the set of selected features.

### D. PROPOSED METHOD (EXPLANATION ABOUT PARAMETERS)

Table 1 shows the parameters of the proposed method. These parameters are used in Figure 2 and Table 2. They illustrate the diagram and pseudo code of the proposed method respectively. R is the radius of each local patch and P is the number of neighbor points of LBP. Other values are related to the feature selection part. All of these parameters are explained in the proposed method part. K is a coefficient for normalization and its value depends on the average value of

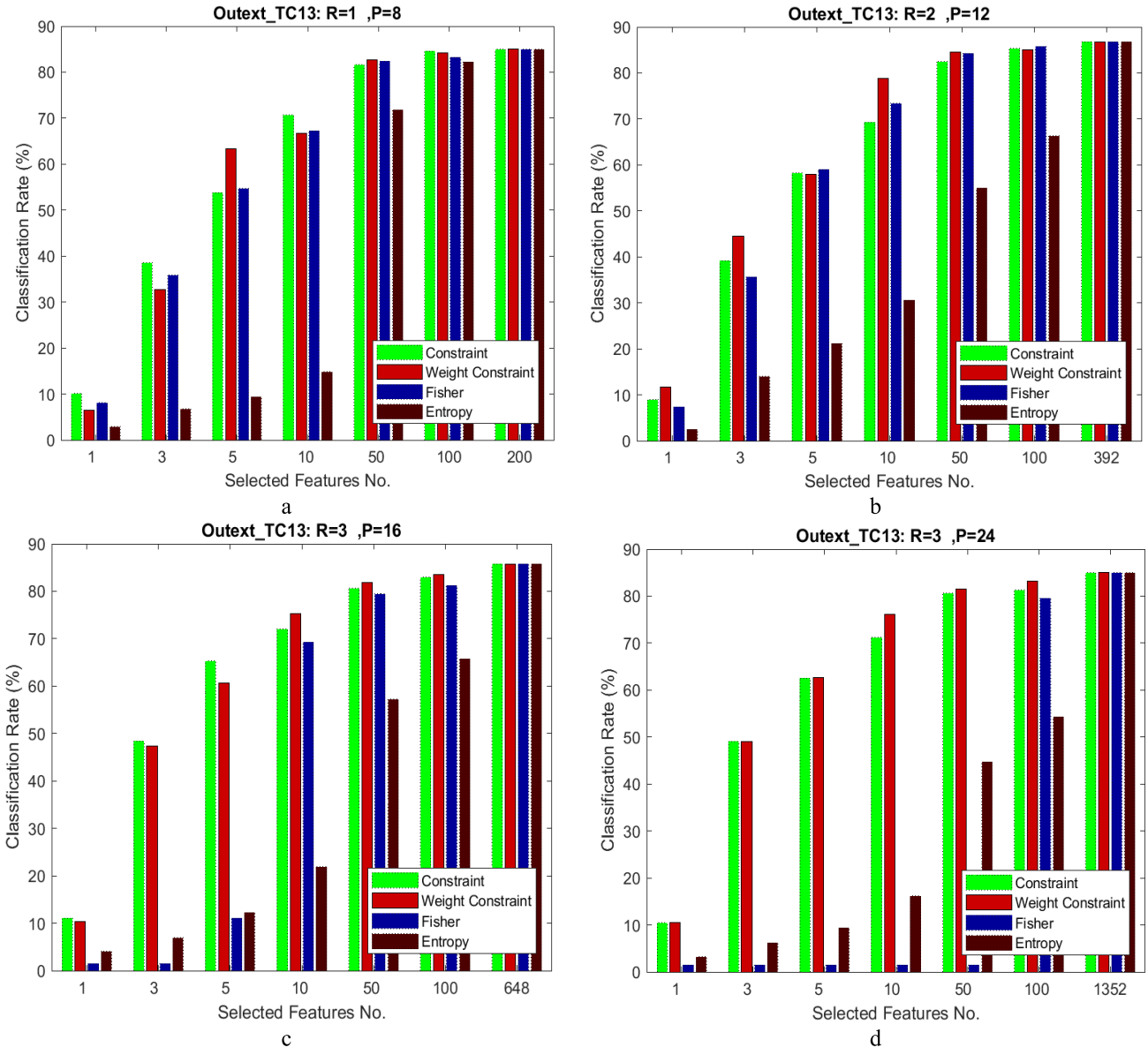


**FIGURE 7.** The results for Outex_TC13 texture for some local size.

information gain. Tr is a threshold value to select the best score features and D(S,M) is a distance value that is used in KNN classification.

## IV. EXPERIMENTAL RESULTS

Here some different textures are used for implementation and comparing the performance of the proposed approach and some novel methods: the Outex dataset [21], includes 24 classes. It includes textures under nine viewpoint angles and three illuminations. Five groups of Outex are used in this paper (TC3, TC10, TC12(t), TC12(h), and TC13). In addition, Columbia-Utrecht Reflection (CUReT) dataset [57] includes 5612 images in 61 classes and 92 samples per class. These textures are prepared in different illumination and viewpoint. UIUC dataset is another dataset that has 25 classes and each class has 40 images [66]. Another texture dataset is MeasTex [67] with 69 classes and 16 samples per class. The virus dataset [68] includes some low-quality and small-size textures of virus images. It contains 1500 images in 15 classes. ORL [69] is a face dataset with 400 faces in 40 classes. RSMAS [70] is one of the special textures. it is a collection of coral reefs in 8 classes and 48 images per class. Also, Brodatz [71] textures include 112 classes of sharp edges textures. It contains 16 images in each class. Table 3 indicates these textures and their specifications.

### A. DISSIMILARITY METRIC METHOD

Such as some scientific papers [33] K-NN (K=1) is used for classification. To compare a pair of textures, the CLBP

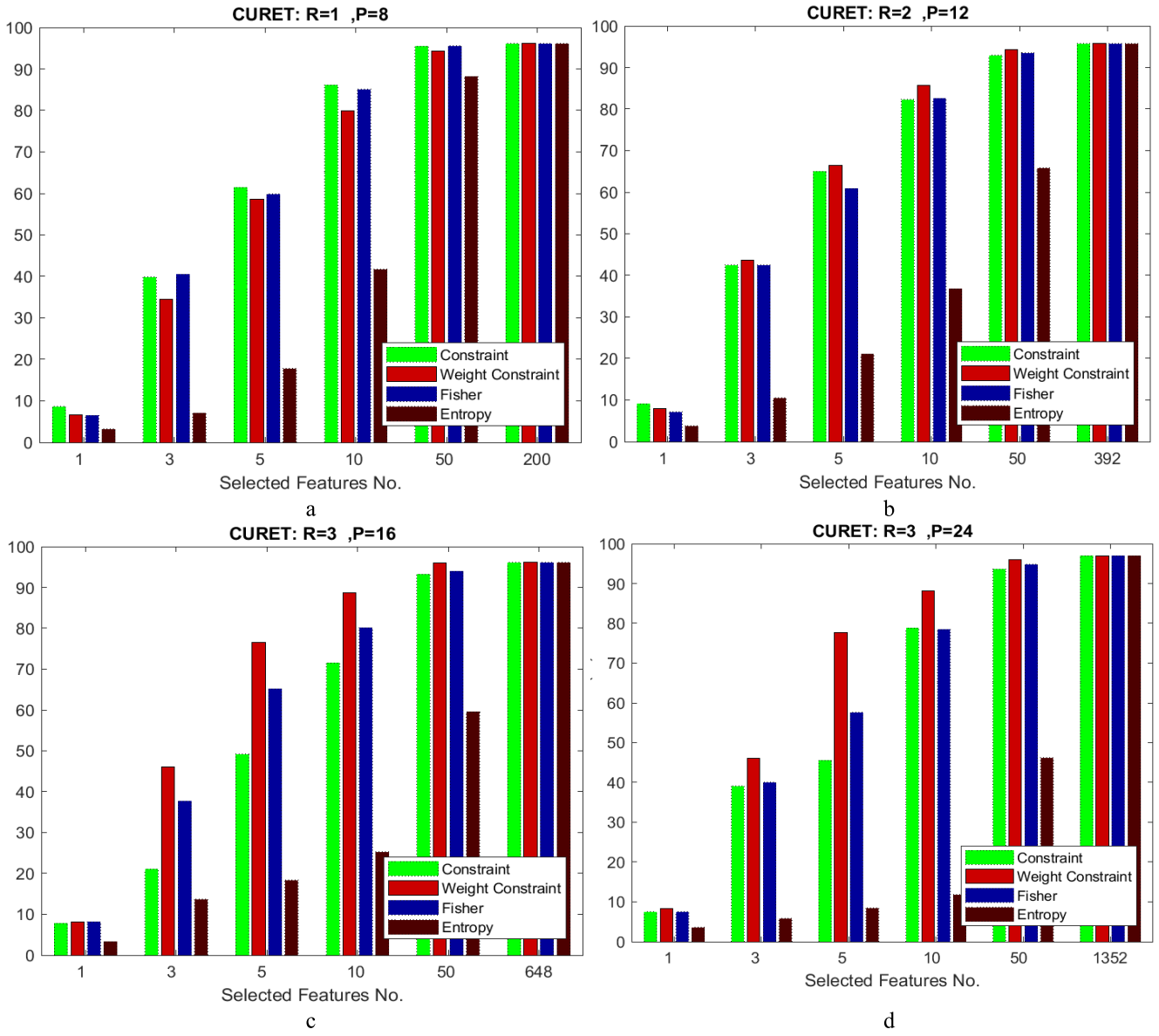


**FIGURE 8.** The results for CUReT texture for some local size.

histograms or CLBP feature vectors of them must be compared. For comparing two histograms some criteria were proposed [33]. Some of the popular methods such as chi-square metric, log-likelihood ratio, and histogram intersection [16] are used for this purpose. Such as some scientific papers [22], [25] in this paper, chi-square is employed for the K-NN distance metric. Relation (15) shows the chi-square method. S is a test sample that is assigned to the class of model M to minimize the chi-square value. In Eq. 15, N is the number of features, and Si and Mi are two values for the sample and model images of the i-th feature.

$$D(S, M) = \sum_{i=1}^{N} \frac{(S_i - M_i)^2}{S_i + M_i} \tag{15}$$

The same as other previous papers [22], [25], [33], the aim of this paper is related achieving the highest classification rate. The classification rate is obtained by dividing the number of textures that are classified correctly by all of the textures.

### B. EXPERIMENTAL RESULTS

Here the implementation of results for the proposed method is compared to some state-of-the-art and advanced approaches. The proposed method is a filter-based method. Therefore, the results of implementations are illustrated without the learning phase. In addition, it is a supervised approach. Therefore the implementations of the proposed method are compared to some supervised and filter-based methods (i.e. Constraint, Fisher, and Entropy). Figs. 3 to 13 compare these

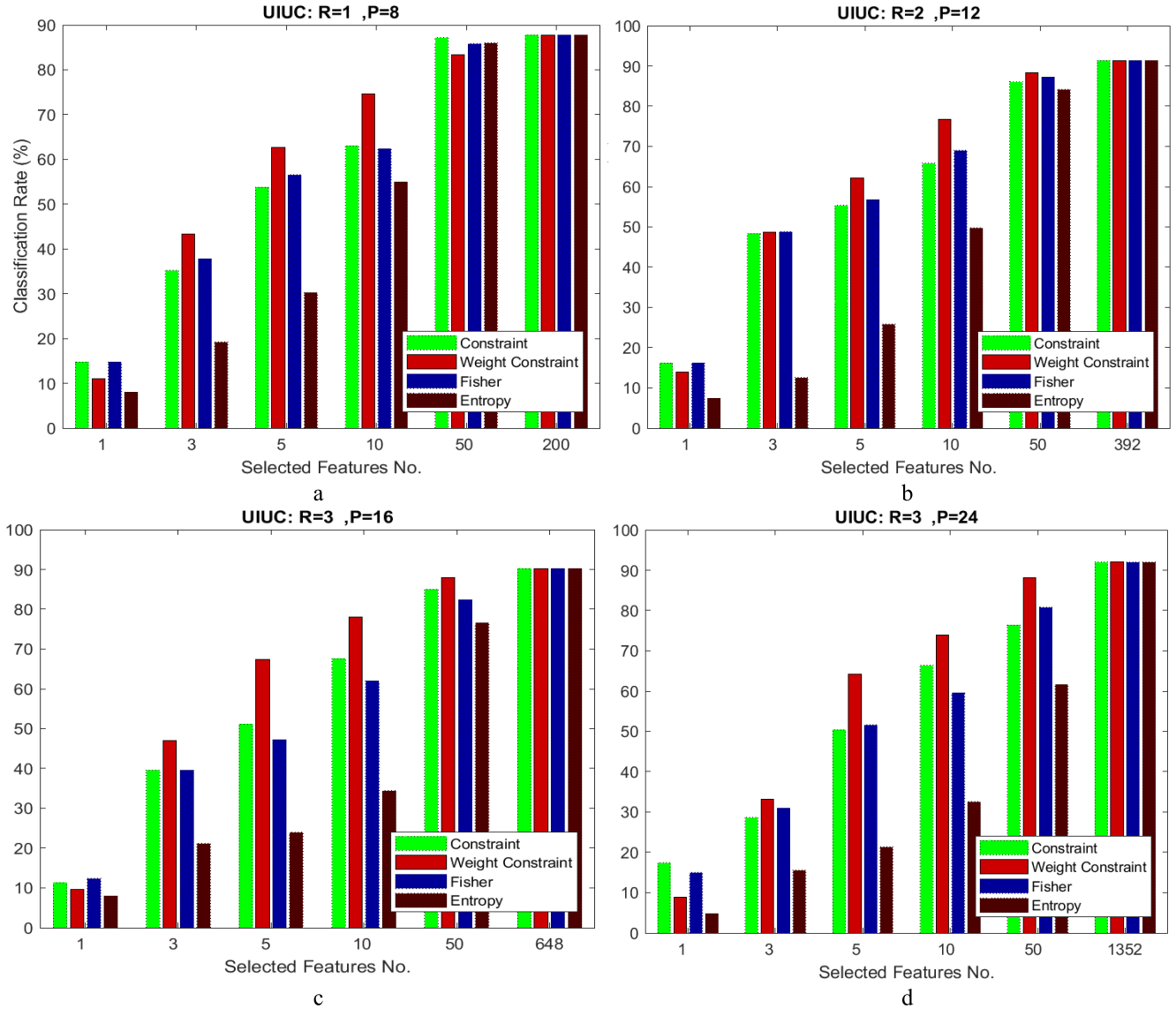


**FIGURE 9.** The results for UIUC texture for some local size.

results. In each figure, four plots are illustrated for different neighborhoods. (i.e. R=1,P=8, R=2,P=12, R=3,P=16 and R=3,P=24). In each plot, the classification rate is compared for 1, 3, 5, 10, 50, and 100 selected features. In addition, the result for whole features is illustrated. The Outex [21], CUReT [65], UIUC [66], MeasTex [67], Virus [68], ORL [69], RSMAS [70],and Brodatz [71] datasets are used in this part.

### C. EXPERIMENTAL RESULTS FOR OUTEX DATASET

The Outex texture includes many test suites [21]. They were collected in different viewpoints, light, and scaling size. In this paper TC3, TC10, TC12(t), TC12(h) and TC13 are employed. The comparison of results for Outex textures is shown in plots of Figs 3 to 7.

Plots of Fig. 3 show the results of TC3 of the Outex dataset. It includes 480 textures in 24 classes. This figure compares the results of the proposed weighted constraint method and the other three feature selection methods for different R, P, and features number. As can be seen from this figure, for all cases, the proposed method outweighs the other approaches. Only for R=1 and P=8 the constraint technique is slightly better than weighted constraint. Because in Outex textures the small regions do not have enough distinctive information and their weighting cannot provide more distinctive features. In each plot, the classification rates of four methods are compared for different numbers of features. The max number of features (i.e. 200, 392, 648, and 1352) indicates the whole of features (without feature selection). These plots show that the classification rate of the proposed method only for 50

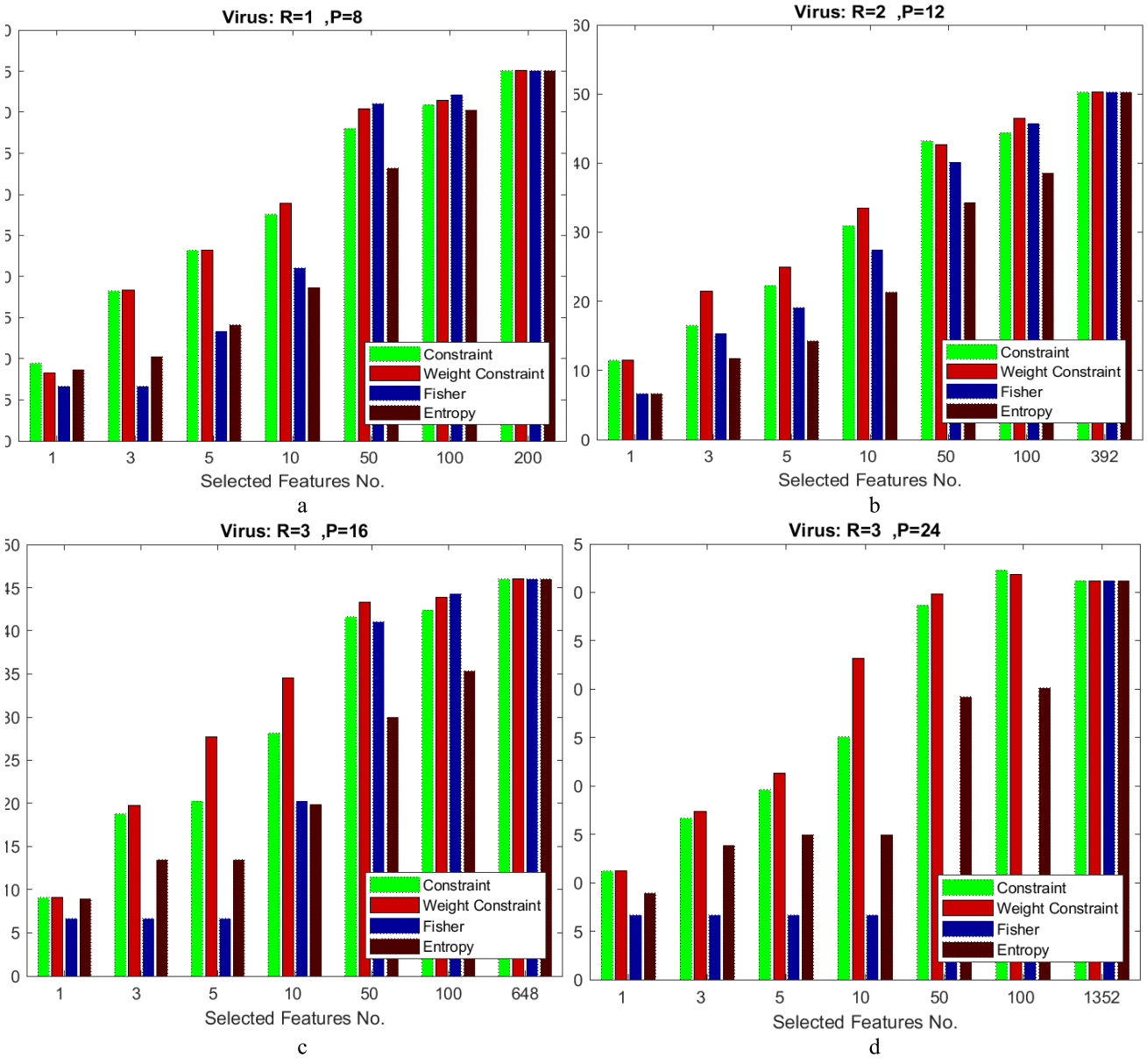


**FIGURE 10.** The results for Virus texture for some local size.

(or even for 10) selected features is the same as using whole features. In addition, for 3 or 5 selected features the proposed method provides a rate that is around 10% to 80% better than Fisher and Entropy method. Considering Fig3.c or Fig3.d they indicate that for the 10 best features that are selected by the proposed method, they can reach the rate of around 97% which is slightly lower than the rate of using whole features.

Fig. 4 shows the classification accuracy of TC10 texture for some local size and features number. TC10 includes 4320 textures in 24 classes. So it includes more textures than TC3. These textures have nine rotations and Inca illumination. Fig. 4 indicates that for larger neighborhoods by using the weighting proposed method the classification rate is increased. By considering these results they indicate that only 10 features selected by the proposed method can provide more than 90% of accuracy for all plots. In addition, for all cases, the weighted constraint is better than the constraint method. Unless for R=1, P=8 when 3 or 10 features are selected, the constraint is slightly better than the proposed method.

Fig. 5 and 6 illustrate the results of Outex_TC12(t) and Outex_TC12(h). These datasets contain 9120 textures in 24 classes. The difference between these two datasets is related to illumination of them. The trend of these plots is the same as in previous figures.

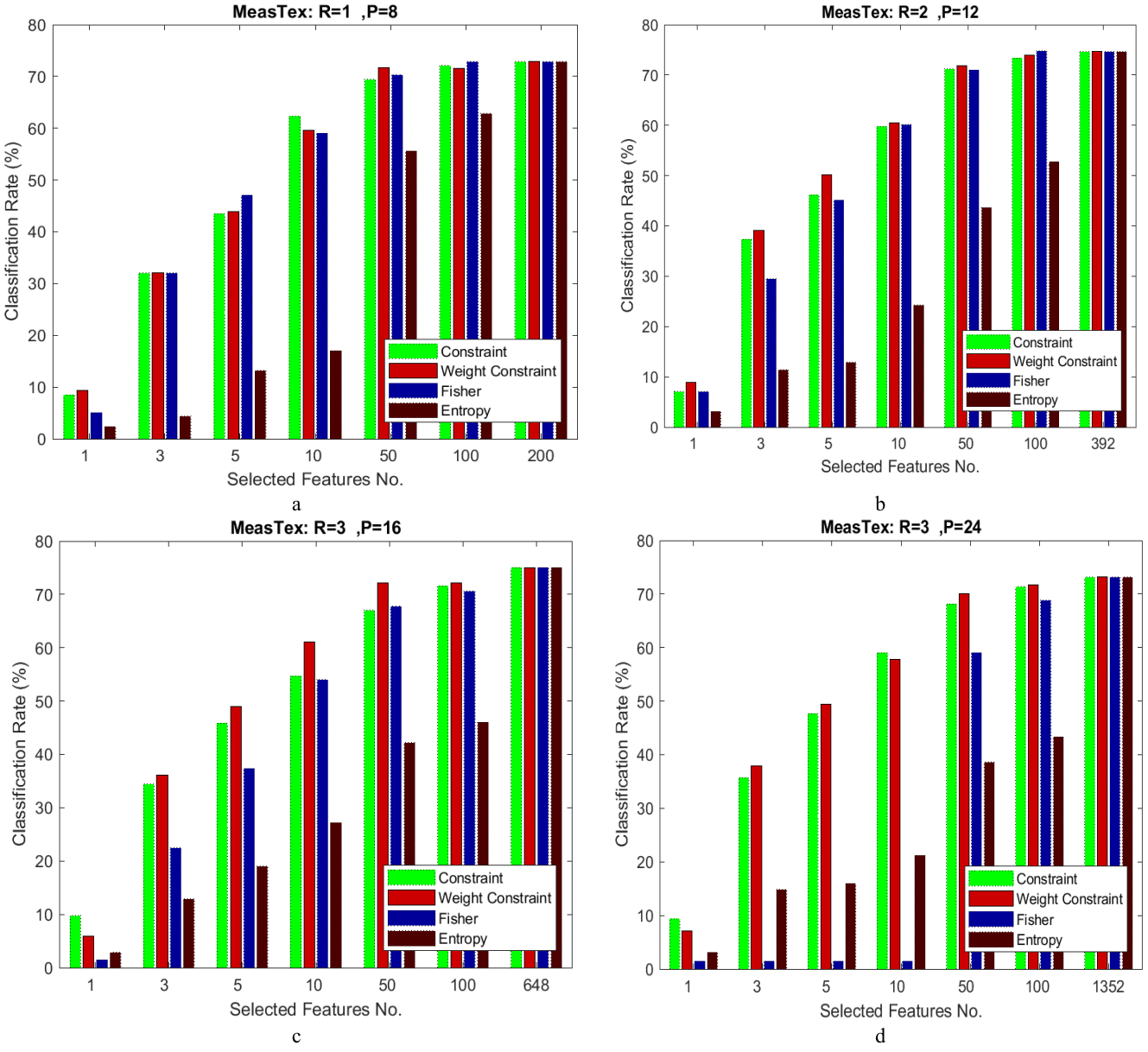


**FIGURE 11.** The results for MeasTex texture for some local size.

Fig. 7 illustrates the results of Outex_TC13. It contains a high number of texture classes. This texture set includes color images. However here, only grayscale information is employed and they are converted into grayscale for feature extraction. The difference between this suit and the previous suits of Outex is related to the number of classes. TC13 includes 68 classes of textures. This figure indicates that the proposed selection method can select discriminative features for a high number of classes. However, it slightly increases the accuracy of the constraint method for this suit of data.

As it can be seen from the results of OUTEX datasets, the proposed method can extract more discriminative features, especially for the higher local patches. In all plots that are shown in Figs. 3 to 7, the weighted constraint and Fisher methods obtained the best and the worst rate respectively. However for some cases especially for small neighborhoods, the constraint method may record a slightly better rate than the proposed method. It is because of this fact that the LBP features of large neighborhoods contain more uniform and more discriminative patterns than small neighborhoods [34] and if the local patch is not large enough the local information may not provide discriminative details and weighting this data does not strengthen them. For some suits of OUTEX dataset such as TC12(t) and TC12(h) which include a high number of textures the proposed method is better than other methods even for the small local patches

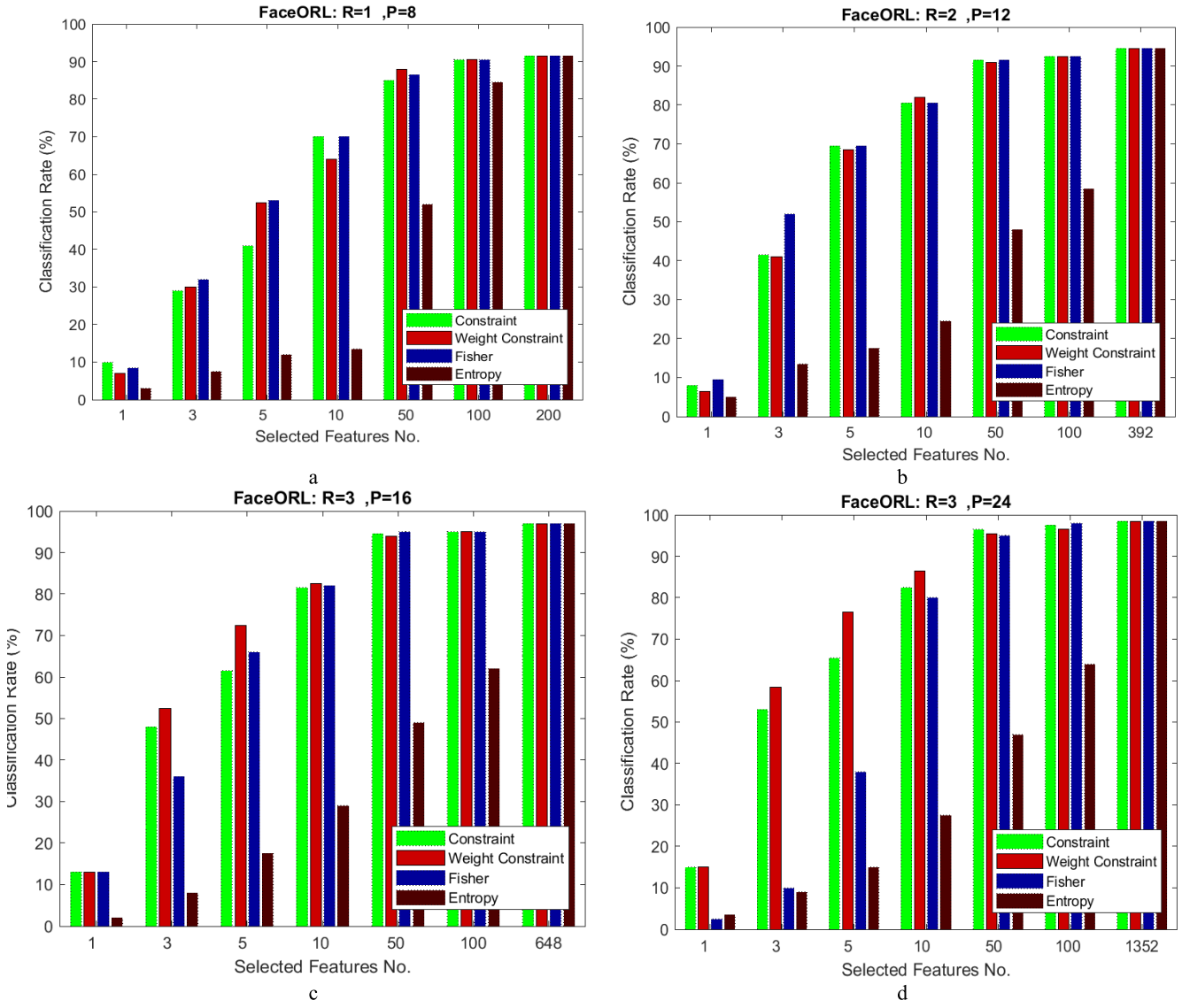


**FIGURE 12.** The results for ORL Face texture for some local size.

(R=1, P=8). In other words, the higher number of samples the better performance of WCFS is obtained. It is because of more complexity of features for a large number of samples and the proposed method can elicit better features by using the weight technique.

### D. EXPERIMENTAL RESULTS FOR CURET DATASET

The CUReT data [65] includes 92 classes of texture and each class has 61 images. This dataset includes different viewpoints and illumination textures. For each class, 92 images are selected from the images that have a viewing angle of less than 60°. These textures contain some classes of very similar images that include very small details. In each training step, 46 images (i.e. 50% of images) are randomly chosen for the train, and remain half is used for the test. The average classification rate over 20 random tests is used as the result of each test. Fig. 8 demonstrates the results of implementation for CUReT textures. This figure illustrates that for all neighborhoods the proposed method provides the best rate. Unless in Fig. 8.a for R=1, P=8 the constraint method is better than the weighted constraint. As it is mentioned for OUTEX dataset in CUReT textures, the small local area does not include enough information therefore the extracted features are not discriminative such as a large local patch. The best rate is obtained in Fig 8.d and it is around 97.8% by selecting the best 50 features.

The plots of Fig 8 especially b, c,and d show that the classification rate of the proposed method declines slightly when a small number of selected features are used for classification. Whereas the classification accuracy decreases significantly for Fisher and Entropy methods. The same as the Outex dataset in CUReT textures the small local patch does not

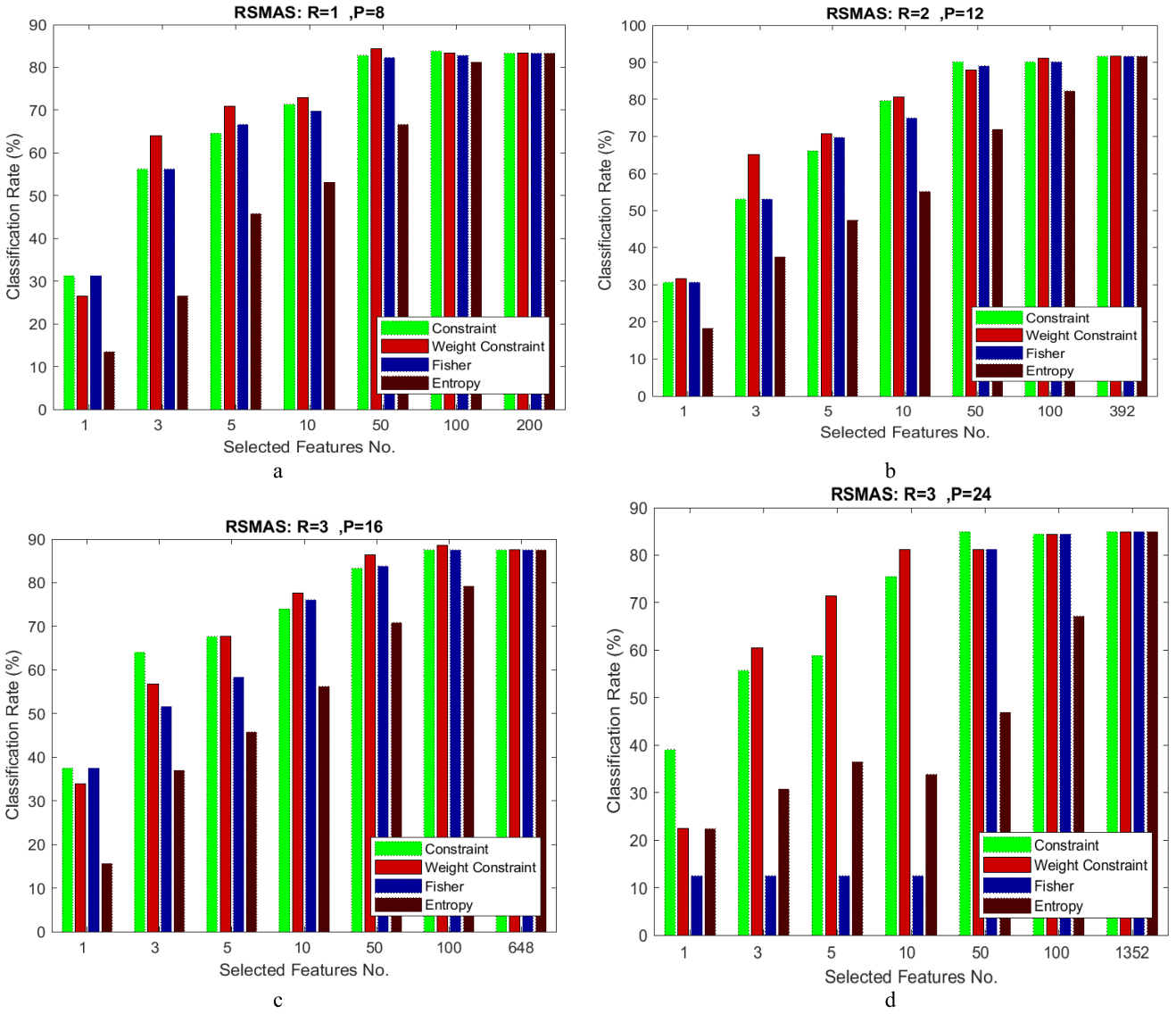


**FIGURE 13.** The results for RSMAS Coral Reef texture for some local size.

include enough discriminative details and the weighting process does not enhance the classification rate.

### E. EXPERIMENTAL RESULTS FOR UIUC DATASET

The UIUC textures [66] include 25 classes of data and each class has 40 images. The size of the texture is 640∗480. Half of the texture of each class is used for train and remains for the test. The test operation is repeated 50 times and the average of the test results is shown in Fig. 9. Despite OUTEX and CUReT textures, this texture includes large textures with sharp edges and high-contrast local details. Because of using textures with high contrast and sharp edges, the LBP features provide high discriminative features even for small patches such as Fig9.a. Therefore, in spite of previous textures the proposed method has better performance than constraint even for small local patches. Fig. 9 indicates that for all neighborhoods the proposed method outweighs other methods. This figure shows that using weight for the constraint method increases the classification rate by around 1 to 20%. Only in Fig9.a for R=1, P=8 when 50 features are selected the constraint method is slightly better than the proposed approach. These results indicate that the proposed method provides better results for sharp-edge textures. However, for one best feature, the proposed method is slightly worse than other methods. Nevertheless, one feature is not enough for a fair comparison.

### F. EXPERIMENTAL RESULTS FOR OTHER DATASETS

For better comparison not only, the general textures are used for implementation, but also some special textures are employed. Fig. 10 shows the performance of classification for the Virus [68] dataset. The Virus dataset includes 1500 small

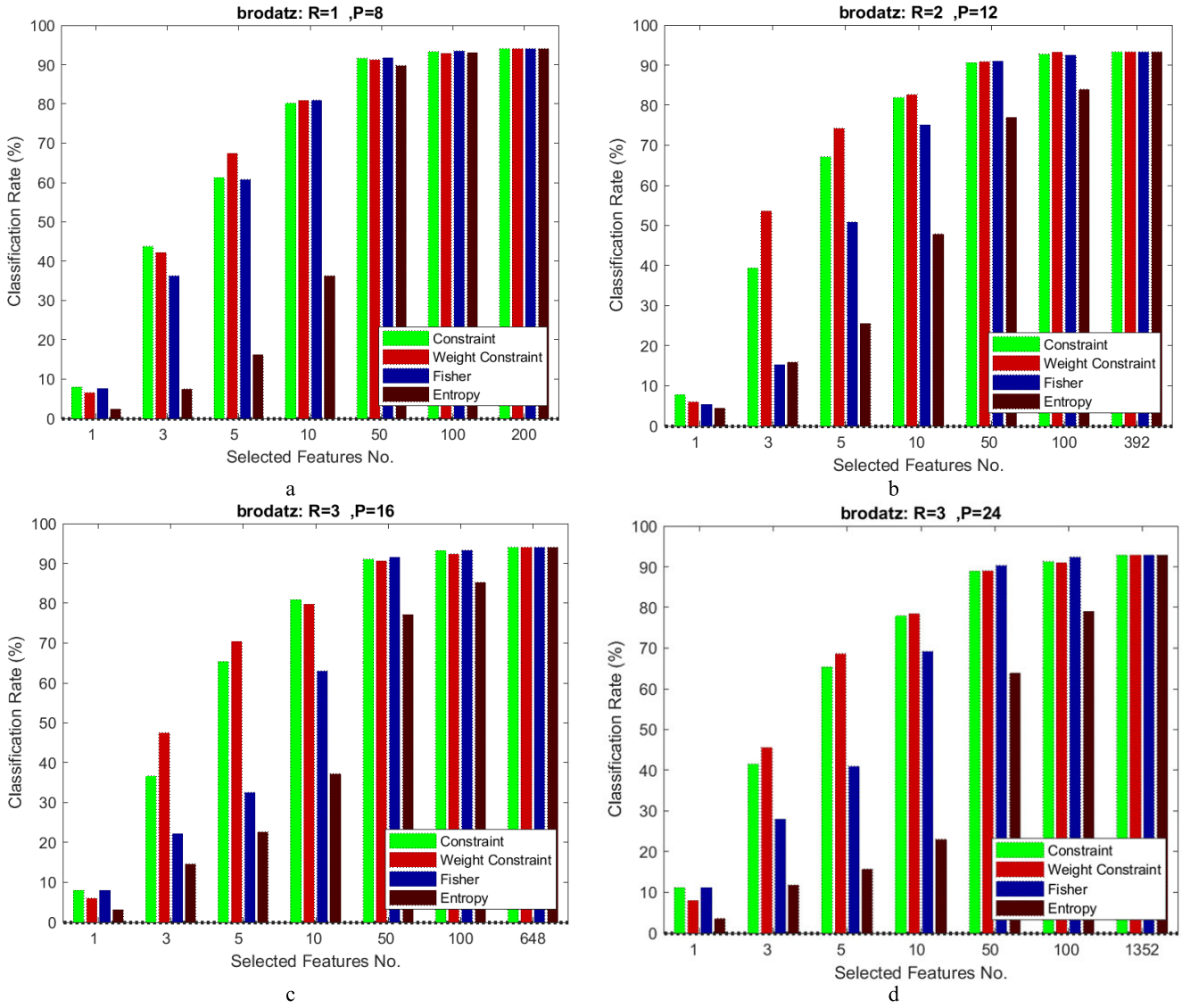


**FIGURE 14.** The results for Brodatz texture for some local size.

and low-contrast images in 15 classes. As can be seen from the figure, the classification rate is almost lower than 50% in all plots. It is because of the low quality and low size of textures of virus images. This figure indicates that the proposed method provides a slightly better rate than other methods, especially for the large neighborhood. Only n Fig 10.a for R=1, P=8 the Fisher method is negligibly better than the proposed method for 50 and 100 selected features. The larger local patch the better features are obtained from each local area and the higher classification rate is prepared. It can be seen in Figures 10.b, 10.c and 10.d. From these four plots, the best rate is obtained in Fig.10.b. In this plot, the classification rate is around 52% for whole features. Whereas it is around 45% and 47% for 50 and 100 selected features. The results of these histograms indicate that for low-quality images the weighted constraint method can select better features than other approaches.

Fig. 11 shows the results of implementation on Meas-Tex [67] textures. This dataset is the same as the Outex textures, therefore the results are the same as Outex results. Fig 11.a has a slightly lower rate than constraint whereas Fig 11.b, 11.c and 11.d indicate the highest rate for weighted features. The change of classification accuracy for this texture is the same as previous general datasets such as Outex.

Fig. 12 shows the comparison between the proposed method and other supervised feature selection methods for the ORL dataset [69]. This dataset is different from other datasets. It includes 400 face images in 40 classes. The highest classification rate is recorded by using completely 1352 features in Fig. 12.d. and it is around 98%. Despite

**TABLE 4.** Comparison of the best classification rate for some textures.

| Method | TC10 | TC12(t) | TC12(h) | CUReT | Average | Published |
|---|---|---|---|---|---|---|
| LBP [16] | 97.70 | 87.30 | 86.40 | 87.22 | 89.66 | TPAMI 2002 |
| LBP/VAR [16] | 98.15 | 87.13 | 87.08 | 91.12 | 90.87 | TPAMI 2002 |
| LTP [30] | 98.64 | 92.59 | 91.52 | 95.46 | 94.55 | TIP 2010 |
| DLBP [31] | 98.10 | 91.60 | 87.40 | 84.93 | 90.51 | TIP 2009 |
| CLBP [34] | 98.93 | 95.32 | 94.53 | 95.85 | 96.16 | TIP 2010 |
| EMCLBP [70] | *99.61* | 96.37 | 95.32 | 95.17 | 96.62 | MMTAP 2018 |
| SLGP [72] | 97.79 | 84.17 | 83.82 | 94.83 | 90.15 | SPL 2018 |
| NTLNP [73] | 98.65 | 92.15 | 92.35 | 91.64 | 93.70 | PRL 2012 |
| BRINT [10] | 97.76 | 95.95 | 96.92 | 96.85 | 96.87 | TIP 2014 |
| DRLBP [74] | 99.19 | **97.15** | 95.37 | 90.77 | 95.62 | PRL 2016 |
| NRLBP [75] | 93.44 | 86.13 | 87.38 | 90.95 | 89.47 | TIP 2013 |
| LDDP [76] | 98.64 | 95.99 | 94.16 | 94.49 | 95.82 | NCAA 2012 |
| CLBC [11] | 98.96 | 95.37 | 94.72 | 95.39 | 96.11 | TIP 2012 |
| DNS+LBP [77] | 99.27 | 94.40 | 93.85 | 95.00 | 95.63 | TIP 2011 |
| LBPV [78] | 97.63 | 95.06 | 93.88 | 96.04 | 95.65 | PR 2010 |
| VZ-Patch [79] | 92.00 | 91.41 | 92.06 | 96.59 | 93.02 | TPAMI 2009 |
| CRLBP [80] | 99.43 | 96.46 | 96.44 | 95.39 | 96.93 | NERO COMP 2012 |
| CJLBP [81] | 98.85 | 95.21 | 95.56 | 97.60 | 96.80 | TVC 2017 |
| CRDP [82] | 97.99 | 94.07 | 91.67 | 89.88 | 93.40 | PR 2017 |
| CLBP_CPS [83] | 99.38 | 95.63 | 95.53 | 97.64 | 97.04 | ESWA 2019 |
| CLBP_ACPS [84] | 99.22 | 96.64 | 96.04 | 96.96 | 97.21 | ESWA 2021 |
| CLBP_SFB_OCPS [85] | 99.22 | 95.79 | 94.58 | 96.70 | 96.57 | ESWA 2023 |
| CLBP_WCFS | **99.63** | *97.12* | **97.22** | **97.81** | **97.95** | Proposed |

other datasets, in this dataset, the Fisher method provides high accuracy. It is because face images have low within-class variance and a high difference between the whole average and the average of each face class. However, for features selecting the proposed and constraint methods prepare better performance than Fisher and entropy methods. The lower number of selected features the more accuracy can be obtained for constraint and weighted constraint methods.

Fig. 13 illustrates the results of RSMAS [70] images. This dataset includes eight classes of coral reefs of 48 images. For this, image the rate of classification is slightly better than the constraint method.

Fig. 14 shows the comparing of the results of weighted constraint and other methods for the Brodatz dataset [71]. These plots indicate that for a small number of best features such as 3, 5,and 10 features the proposed method outweighs the other methods. Only for some cases such as small patches, the constraint method is slightly better than the proposed method. Similar to previous datasets the proposed selected features are better than Fisher and Entropy features. The smaller number of selected features the higher rate is obtained for proposed method compared to Fisher and Entropy.

### G. COMPARING TO OTHER DESCRIPTORS AND LBP VERSIONS

In this section, for some datasets, the classification rate is compared to some advanced and state-of-the-art methods [78-102]. Most of these methods in this part use R=3, and P=24 for each local patch.

Table 4 compares the results of the proposed method and some state-of-the-art and advanced local descriptors of textures. The proposed method is CLBP_WCFS (i.e. Weighted Constraint Feature Selection of CLBP). This table illustrates the best results of each textures descriptors for three types of Outext and CUReT datasets. As can be seen from the table, WCFS achieves the best rate for TC10 and TC12(h) and they are 99.63% and 97.22% respectively. For TC12(t) textures the best result belongs to DRLBP and it is only 0.03% better than the proposed method. For the CUReT dataset, the proposed method achieves the highest rate and it is 97.81%. For better comparison, the average accuracy of these four datasets is shown in the last column of data. The best average rate is recorded by the proposed method and it is 97.95%.

### H. THE COMPARISON OF FEATURES NUMBER VS. CLASSIFICATION RATE

Table 5 shows the comparison of the feature number of the proposed method and some state-of-the-art methods. This table indicates the average classification rate for some datasets that provide by each extracted feature. As it can be seen from the table, the proposed method records the best accuracy whereas the feature number is lower than most of the other methods.

The features of the proposed method are selected from 1352 features. Therefore the feature number for the best rate is lower than 1352. In some cases, it is very low such as 10 features to achieve the best classification rate.

### I. THE COMPARISON WITH DEEP FEAURE

Table 6 indicates the comparison of deep features [86] and proposed method features for Outex and CUReT textures. Table 6 illustrates that for Outext10 and Outex12(t) the selected features have higher accuracy than some deep features [87], [88]. The deep models of this table include

**TABLE 5. Comparison of the best classification rate for some textures.**

| Method | Average Classification Rate | Features No. |
|---|---|---|
| LBP [16] | 89.66 | 26 |
| LBP/VAR [16] | 90.87 | 2400 |
| LTP [30] | 94.55 | 676 |
| DLBP [31] | 90.51 | 52428 |
| CLBP [34] | 96.16 | 1352 |
| EMCLBP [70] | 96.62 | 8712 |
| SLGP [72] | 90.15 | 256 |
| BRINT [10] | 96.87 | 1296 |
| CLBC [11] | 96.11 | 4490 |
| LBPV [78] | 95.65 | 555 |
| VZ-Patch [79] | 93.02 | 240 |
| CRLBP [80] | 96.93 | 1352 |
| CLBP_CPS [83] | 97.04 | 1352 |
| CLBP_ACPS [84] | 97.21 | 1352 |
| CLBP_SFB_OCPS [85] | 96.57 | 1352 |
| CLBP_WCFS | **97.95** | <1352 |

**TABLE 6. Comparing the results of selected CLBP_WCFS features and deep features for datasets.**

| Features | Outex10 | Outex12(t) | CUReT |
|---|---|---|---|
| AlexNet [87] | 67.30 | 73.30 | 98.40 |
| FV+ALexNet [88] | 67.30 | 72.30 | 98.40 |
| FV+CNN [88] | 80.00 | 82.16 | - |
| FV-VGGM [87] | 72.80 | 77.50 | 98.70 |
| FV-VGG-VD [88] | 80.00 | 82.30 | **99.00** |
| CLBP_WCFS | **99.63** | **97.12** | 97.81 |

AlexNet and AlexNet with Fisher Vector (FV). VGG net and VGG-VD (Very Deep VGG net) with Fisher vector, Convolutional Neural Network with Fisher vector. The results illustrate that for CUReT textures all of the deep features provide higher accuracy than the proposed method. It is because of very similar texture classes and very small details of these types of textures. Very Deep VGG [87] achieves the best rate for CUReT textures. CLBP_WCFS provides the rotation and illumination invariant features whereas the deep are not rotation and illumination invariant. Therefore the proposed method prepares better results for rotated and different illumination textures such as Outex10 and Outex12(t).

## V. CONCLUSION

In this paper, a framework for texture classification was proposed. The main part of this framework belongs to the selection of extracted features. The weighted constraint feature selection method was proposed. In this technique, the Fisher score is used as the weight of the constraint score. In other words, this method uses the numerator and denominator of the Fisher equation as weights of the constraint method to select better features than both of these methods. This score is based-on minimizing the within-class variance and maximizing the between class variance. By selecting the features, it is possible to use small memory for keeping features without decreasing the classification rate. The improvement of the proposed methods is significant for large values of R and P. It is because of LBP features. The LBP features include more uniform patterns for large local patches. The more uniform patterns the higher discriminant attributes can be extracted from the textures. If LBP is used for large neighborhoods the number of features increases significantly. However, if these features are filtered by the proposed method the selected features of large neighborhoods provide better results than all features of small neighborhoods. Furthermore, the implementation proved this statement that for textures with sharp edges and high contrast details the proposed method provides better results. It is because of more discriminative features of LBP for high contrast and high details textures.

After the selection of features by the proposed score, an enhancement technique is proposed that by employing the threshold value some enhancement is applied to the features set. The threshold value is estimated based on information gain of features. The discriminant features that are removed because of their low score, are added to selected features. Furthermore, some removed features with higher scores than threshold values are added to the features set. The proposed threshold can select the discriminative features properly and decline the feature number drastically. Therefore, feature selection can remove redundant features and increase the classification rate significantly.

The proposed framework can be used for some applications with a lack of storage. Because of the small size of memory, a small number of features can be used for classification. The implementation results of almost all of the datasets of this paper indicate that by selection of a very small number of best features (i.e. 3 or 5 or 10 feature) we can achieve the classification rate that is only 5 to 10% lower than the highest rate. The maximum rate almost can be obtained by using all of the features such as 1352 features for a large local patches in CLBP.

## DECLARATIONS

Conflict of interest Authors declare that they have no conflict of interest.

## REFERENCES


[1] K. S. Kumar and M. R. Bai, "LSTM based texture classification and defect detection in a fabric," *Meas., Sensors*, vol. 26, Apr. 2023, Art. no. 100603, doi: 10.1016/j.measen.2022.100603.

[2] F. Tajeripour, E. Kabir, and A. Sheikhi, "Fabric defect detection using modified local binary patterns," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 1, pp. 1–12, Dec. 2007.

[3] H. Ali, M. Sharif, M. Yasmin, and M. H. Rehmani, "A shallow extraction of texture features for classification of abnormal video endoscopy frames," *Biomed. Signal Process. Control*, vol. 35, no. 1, pp. 405–415, 2023.

[4] H. Anys and D.-C. He, "Evaluation of textural and multipolarization radar features for crop classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 5, pp. 1170–1181, Sep. 1995.

[5] N. Daniel and A. Anitha, "Texture and quality analysis for face spoofing detection," *Comput. Electr. Eng.*, vol. 94, Sep. 2021, Art. no. 107293.

[6] S. Murala, R. P. Maheshwari, and R. Balasubramanian, "Local tetra patterns: A new feature descriptor for content-based image retrieval," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2874–2886, May 2012.

[7] M. H. Shakoor and R. Boostani, "Radial mean local binary pattern for noisy texture classification," *Multimedia Tools Appl.*, vol. 77, pp. 21481–21508, Aug. 2018.

[8] M. H. Shakoor and F. Tajeripour, "Repeating average filter for noisy texture classification," *Scientia Iranica*, vol. 24, no. 3, pp. 1419–1436, Jun. 2017.

[9] M. H. Shakoor and F. Tajeripour, "Noise robust and rotation invariant entropy features for texture classification," *Multimedia Tools Appl.*, vol. 76, pp. 8031–8066, Mar. 2016.

[10] L. Liu, Y. Long, P. W. Fieguth, S. Lao, and G. Zhao, "BRINT: Binary rotation invariant and noise tolerant texture classification," *IEEE Trans. Image Process.*, vol. 23, no. 7, pp. 3071–3084, Jul. 2014.

[11] Y. Zhao, D.-S. Huang, and W. Jia, "Completed local binary count for rotation invariant texture classification," *IEEE Trans. Image Process.*, vol. 21, no. 10, pp. 4492–4497, Oct. 2012.

[12] T. Lacombe, H. Favreliere, and M. Pillet, "Modal features for image texture classification," *Pattern Recognit. Lett.*, vol. 135, pp. 249–255, Jul. 2020.

[13] N. Shrivastava, "A novel texture classification scheme based on completed multiple adaptive threshold patterns," *Proc. Comput. Sci.*, vol. 171, pp. 331–340, Jan. 2020.

[14] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[15] H. Shah, N. Badshah, F. Ullah, and A. Ullah, "A new selective segmentation model for texture images and applications to medical images," *Biomed. Signal Process. Control*, vol. 48, pp. 234–247, Feb. 2019.

[16] A. Chaudhari and J. Kulkarni, "Cerebral edema segmentation using textural feature," *Biocybernetics Biomed. Eng.*, vol. 39, no. 3, pp. 599–612, Jul. 2019.

[17] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. 3, no. 6, pp. 610–621, 1973.

[18] M. M. Galloway, "Texture analysis using gray level run lengths," *Comput. Graph. Image Process.*, vol. 4, no. 2, pp. 172–179, Jun. 1975.

[19] A. H. Mir, M. Hanmandlu, and S. N. Tandon, "Texture analysis of CT images," *IEEE Eng. Med. Biol. Mag.*, vol. 14, no. 6, pp. 781–786, Nov./Dec. 1995.

[20] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, Jan. 1996.

[21] T. Ojala, T. Maenpaa, M. Pietikainen, J. Viertola, J. Kyllonen, and S. Huovinen, "Outex—New framework for empirical evaluation of texture analysis algorithms," in *Proc. Int. Conf. Pattern Recognit.*, 2002, pp. 701–706.

[22] X. Huang, S. Z. Li, and Y. Wang, "Shape localization based on statistical method using extended local binary pattern," in *Proc. 3rd Int. Conf. Image Graph. (ICIG)*, 2004, pp. 184–187.

[23] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.

[24] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 915–928, Jun. 2007.

[25] M. H. Shakoor, "Lung tumour detection by fusing extended local binary patterns and weighted orientation of difference from computed tomography," *IET Image Process.*, vol. 13, no. 6, pp. 877–884, May 2019.

[26] T. Ojala, "Nonparametric texture analysis using simple spatial operators, with applications in visual inspection," Ph.D. thesis, Dept. Elect. Eng., Univ. Oulu, Finland, 1997.

[27] M. Pietikäinen, T. Ojala, and Z. Xu, "Rotation-invariant texture classification using feature distributions," *Pattern Recognit.*, vol. 33, no. 1, pp. 43–52, Jan. 2000.

[28] D. Huang, Y. Wang, and Y. Wang, "A robust method for near infrared face recognition based on extended local binary pattern," in *Proc. Int. Symp. Vis. Comput.*, 2007, pp. 437–446.

[29] Y. Huang, Y. Wang, and T. Tan, "Combining statistics of geometrical and correlative features for 3D face recognition," in *Proc. Brit. Mach. Vis. Conf.*, 2006, pp. 879–888.

[30] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *Proc. Int. Workshop Anal. Model. Faces Gestures*, 2007, pp. 168–182.

[31] S. Liao, M. W. K. Law, and A. C. S. Chung, "Dominant local binary patterns for texture classification," *IEEE Trans. Image Process.*, vol. 18, no. 5, pp. 1107–1118, May 2009.

[32] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognit.*, vol. 42, no. 3, pp. 425–436, Mar. 2009.

[33] B. Zhang, Y. Gao, S. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 533–544, Feb. 2010.

[34] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1657–1663, Jun. 2010.

[35] M. H. Shakoor and R. Boostani, "Extended mapping local binary pattern operator for texture classification," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 31, no. 6, Jun. 2017, Art. no. 1750019.

[36] P. Wang, B. Xue, J. Liang, and M. Zhang, "Feature selection using diversity-based multi-objective binary differential evolution," *Inf. Sci.*, vol. 626, pp. 586–606, May 2023.

[37] P. Zhu, X. Hou, K. Tang, Y. Liu, Y.-P. Zhao, and Z. Wang, "Unsupervised feature selection through combining graph learning and $\ell_{2,0}$-norm constraint," *Inf. Sci.*, vol. 622, pp. 68–82, Apr. 2023.

[38] A. Hashemi, M. B. Dowlatshahi, and H. Nezamabadi-pour, "VMFS: A VIKOR-based multi-target feature selection," *Exp. Syst. Appl.*, vol. 182, Nov. 2021, Art. no. 115224, doi: 10.1016/j.eswa.2021.115224.

[39] M. S. Jahani, G. Aghamollaei, M. Eftekhari, and F. Saberi-Movahed, "Unsupervised feature selection guided by orthogonal representation of feature space," *Neurocomputing*, vol. 516, pp. 61–76, Jan. 2023.

[40] F. BenSaid and A. M. Alimi, "Online feature selection system for big data classification based on multi-objective automated negotiation," *Pattern Recognit.*, vol. 110, Feb. 2021, Art. no. 107629.

[41] S. Solorio-Fernández, J. F. Martínez-Trinidad, and J. A. Carrasco-Ochoa, "A supervised filter feature selection method for mixed data based on spectral feature selection and information-theory redundancy analysis," *Pattern Recognit. Lett.*, vol. 138, pp. 321–328, Oct. 2020.

[42] A. Porebski, V. T. Hoang, N. Vandenbroucke, and D. Hamad, "Multi-color space local binary pattern-based feature selection for texture classification (Erratum)," *J. Electron. Imag.*, vol. 27, no. 3, Jun. 2018, Art. no. 039801, doi: 10.1117/1.JEI.27.3.039801.

[43] M. Kalakech, A. Porebski, N. Vandenbroucke, and D. Hamad, "A new LBP histogram selection score for color texture classification," in *Proc. Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Nov. 2015, pp. 242–247.

[44] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Jan. 2003.

[45] E. Saghapour, S. Kermani, and M. Sehhati, "A novel feature ranking method for prediction of cancer stages using proteomics data," *PLoS ONE*, vol. 12, no. 9, Sep. 2017, Art. no. e0184203, doi: 10.1371/journal.pone.0184203.

[46] X. Song, J. Zhang, Y. Han, and J. Jiang, "Semi-supervised feature selection via hierarchical regression for web image classification," *Multimedia Syst.*, vol. 22, no. 1, pp. 41–49, Feb. 2016, doi: 10.1007/s00530-014-0390-0.

[47] Y. Han, Y. Yang, Y. Yan, Z. Ma, N. Sebe, and X. Zhou, "Semisupervised feature selection via spline regression for video semantic recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 2, pp. 252–264, Feb. 2015.

[48] M. H. Shakoor, R. Boostani, M. Sabeti, and M. Mohammadi, "Feature selection and mapping of local binary pattern for texture classification," *Multimedia Tools Appl.*, vol. 82, no. 5, pp. 7639–7676, Feb. 2023.

[49] D. Zhang, S. Chen, and Z.-H. Zhou, "Constraint score: A new filter method for feature selection with pairwise constraints," *Pattern Recognit.*, vol. 41, no. 5, pp. 1440–1451, May 2008, doi: 10.1016/j.patcog.2007.10.009.

[50] K. Benabdeslem and M. Hindawi, "Constrained Laplacian score for semi-supervised feature selection," in *Machine Learning and Knowledge Discovery in Databases*. Berlin, Germany: Springer, 2011, pp. 204–218.

[51] A. Porebski, N. Vandenbroucke, and D. Hamad, "LBP histogram selection for supervised color texture classification," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 3239–3243.

[52] V. T. Hoang, A. Porebski, N. Vandenbroucke, and D. Hamad, "LBP histogram selection based on sparse representation for color texture classification," in *Proc. 12th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, 2017, pp. 476–483.

[53] A. Moujahid, A. Abanda, and F. Dornaika, "Feature extraction using block-based local binary pattern for face recognition," *Electron. Imag.*, vol. 28, no. 10, pp. 1–6, Feb. 2016.

[54] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[55] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford Univ. Press, 1995.

[56] B. Azhagusundari and A. S. Thanamani, "Feature selection based on information gain," *Int. J. Innov. Technol. Exploring Eng.*, vol. 2, no. 2, pp. 18–21., 2013.

[57] L. Yu and H. Liu, "Feature selection for high-dimensional data: A fast correlation-based filter solution," in *Proc. ICML*, vol. 3, 2003, pp. 856–863.

[58] Z. Zhao and H. Liu, "Spectral feature selection for supervised and unsupervised learning," in *Proc. 24th Int. Conf. Mach. Learn.*, Jun. 2007, pp. 1151–1157.

[59] M. Dash and H. Liu, "Feature selection for classification," *Intell. Data Anal.*, vol. 1, nos. 1–4, pp. 131–156, 1997.

[60] X. He, D. Cai, and P. Niyogi, "The Laplacian score for feature selection," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 18, 2005, pp. 1–8.

[61] M. Liu and D. Zhang, "Sparsity score: A novel graph-preserving feature selection method," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 28, no. 4, Jun. 2014, Art. no. 1450009.

[62] Q. Gu, Z. Li, and J. Han, "Generalized Fisher score for feature selection," in *Proc. 27th Conf. Uncertainty Artif. Intell.*, 2011, pp. 266–273.

[63] B. Seijo-Pardo, V. Bolón-Canedo, and A. Alonso-Betanzos, "On developing an automatic threshold applied to feature selection ensembles," *Inf. Fusion*, vol. 45, pp. 227–245, Jan. 2019.

[64] M. I. Prasetiyowati, N. U. Maulidevi, and K. Surendro, "Determining threshold value on information gain feature selection to increase speed and prediction accuracy of random forest," *J. Big Data*, vol. 8, no. 1, p. 84, Dec. 2021, doi: 10.1186/s40537-021-00472-4.

[65] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real world surfaces," *ACM Trans. Graph.*, vol. 18, no. 1, pp. 1–34, 1999.

[66] S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine regions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1265–1278, Aug. 2005.

[67] G. Smith, "MeasTex image texture database and test suite centre for sensor signal and information processing," Univ. Queensland, Australia, Tech. Rep., Jan. 1998.

[68] G Kylberg. (2012). *Virus Texture Dataset V. 1.0*. [Online]. Available: http://www.cb.uu.se/~gustaf/virustexture

[69] (2002). *AT&T Laboratories, Database of Faces*. [Online]. Available: http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html

[70] M. H. Shakoor and R. Boostani, "A novel advanced local binary pattern for image-based coral reef classification," *Multimedia Tools Appl.*, vol. 77, no. 2, pp. 2561–2591, Jan. 2018, doi: 10.1007/s11042-017-4394-6.

[71] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. New York, NY, USA: Dover, 1966.

[72] T. Song, L. Xin, C. Gao, G. Zhang, and T. Zhang, "Grayscale-inversion and rotation invariant texture description using sorted local gradient pattern," *IEEE Signal Process. Lett.*, vol. 25, no. 5, pp. 625–629, May 2018, doi: 10.1109/LSP.2018.2809607.

[73] A. Fathi and A. R. Naghsh-Nilchi, "Noise tolerant local binary pattern operator for efficient texture analysis," *Pattern Recognit. Lett.*, vol. 33, no. 9, pp. 1093–1100, Jul. 2012.

[74] R. Mehta and K. Egiazarian, "Dominant rotated local binary patterns (DRLBP) for texture classification," *Pattern Recognit. Lett.*, vol. 71, pp. 16–22, Feb. 2016, doi: 10.1016/j.patrec.2015.11.019.

[75] J. Ren, X. Jiang, and J. Yuan, "Noise-resistant local binary pattern with an embedded error-correction mechanism," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 4049–4060, Oct. 2013, doi: 10.1109/TIP.2013.2268976.

[76] Z. Guo, Q. Li, J. You, D. Zhang, and W. Liu, "Local directional derivative pattern for rotation invariant texture classification," *Neural Comput. Appl.*, vol. 21, no. 8, pp. 1893–1904, Nov. 2012, doi: 10.1007/s00521-011-0586-6.

[77] F. M. Khellah, "Texture classification using dominant neighborhood structure," *IEEE Trans. Image Process.*, vol. 20, no. 11, pp. 3270–3279, Nov. 2011, doi: 10.1109/TIP.2011.2143422.

[78] Z. Guo, L. Zhang, and D. Zhang, "Rotation invariant texture classification using LBP variance (LBPV) with global matching," *Pattern Recognit.*, vol. 43, no. 3, pp. 706–719, 2010, doi: 10.1016/j.patcog.2009.08.017.

[79] M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 2032–2047, Nov. 2009, doi: 10.1109/TPAMI.2008.182.

[80] Y. Zhao, W. Jia, R. X. Hu, and H. Min, "Completed robust local binary pattern for texture classification," *Neurocomputing*, vol. 106, no. 6, pp. 68–76, 2012, doi: 10.1016/j.neucom.2012.10.017.

[81] X. Wu and J. Sun, "Joint-scale LBP: A new feature descriptor for texture classification," *Vis. Comput.*, vol. 33, no. 3, pp. 317–329, Mar. 2017.

[82] K. Wang, C.-E. Bichot, Y. Li, and B. Li, "Local binary circumferential and radial derivative pattern for texture classification," *Pattern Recognit.*, vol. 67, pp. 213–229, Jul. 2017.

[83] Z. Pan, X. Wu, and Z. Li, "Central pixel selection strategy based on local gray-value distribution by using gradient information to enhance LBP for texture classification," *Exp. Syst. Appl.*, vol. 120, pp. 319–334, Apr. 2019, doi: 10.1016/j.eswa.2018.11.041.

[84] Z. Pan, S. Hu, X. Wu, and P. Wang, "Adaptive center pixel selection strategy in local binary pattern for texture classification," *Exp. Syst. Appl.*, vol. 180, Oct. 2021, Art. no. 115123.

[85] S. Lan, H. Fan, S. Hu, X. Ren, X. Liao, and Z. Pan, "An edge-located uniform pattern recovery mechanism using statistical feature-based optimal center pixel selection strategy for local binary pattern," *Exp. Syst. Appl.*, vol. 221, Jul. 2023, Art. no. 119763, doi: 10.1016/j.eswa.2023.119763.

[86] L. Liu, P. Fieguth, X. Wang, M. Pietikäinen, and D. Hu, "Evaluation of LBP and deep texture descriptors with a new robustness benchmark," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 69–86, doi: 10.1007/978-3-319-46487-9_5.

[87] M. Cimpoi, S. Maji, and A. Vedaldi, "Deep filter banks for texture recognition and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3828–3836.

[88] M. Cimpoi, S. Maji, I. Kokkinos, and A. Vedaldi, "Deep filter banks for texture recognition, description, and segmentation," *Int. J. Comput. Vis.*, vol. 118, no. 1, pp. 65–94, May 2016, doi: 10.1007/s11263-015-0872-3.

**ENTESSAR SAEED GEMEAY** received the B.Sc. and M.Sc. degrees in electronics and communications engineering and the Ph.D. degree in wireless communications from Tanta University, Tanta, Egypt, in 1996, 2002, and 2010, respectively. She is currently an Assistant Professor with the Department of Electronics and Communications Engineering, Tanta University. She is also with the College of Computer and Information Technology, Taif University, Saudi Arabia. Her research interests include the applications of IoT, mobile generations, cognitive radio, and signal processing.

**FARHAN A. ALENIZI** received the B.Sc. and M.Sc. degrees in electrical engineering from King Saud University, Saudi Arabia, in 1999 and 2006, respectively, and the Ph.D. degree in electrical engineering and computer science from the University of California at Irvine, Irvine, CA, USA, in 2017. Since 2010, he has been a Faculty Member with the Department of Electrical Engineering, Prince Sattam bin Abdulaziz University, Saudi Arabia, where he is currently an Assistant Professor. Besides his academic experience, he worked for ten years with Saudi Telecom Company (STC), a leading telecommunication company in the Middle East as a Designer and a Consultant in the satellite and optical fiber transmission networks. Moreover, he is also involved in drones spoofing and jamming projects. His research interests include images and video processing, signal processing, discrete signal processing (DSP), images and video watermarking, 3D-mesh objects watermarking, and secure multimedia exchanges.

**ADIL HUSSEIN MOHAMMED** received the B.Sc. degree in electrical and electronic engineering and the M.Sc. degree in communication engineering and EW.

He is currently an Assistant Lecturer (Full) with the Department of Communication and Computer, Cihan University-Erbil, Kurdistan Region, Iraq. His research interests include communication engineering, electro magnetic waves and antenna, the IoT, intelligent systems, solar energy, air pollution, neural networks, and medical devices.

Mr. Mohammed is a member of the Iraqi Engineering Union.

**MOHAMMAD HOSSEIN SHAKOOR** received the B.Sc. degree in computer engineering from Shiraz University, Shiraz, Iran, in 1998, the M.S. degree in computer architecture from Isfahan University, Isfahan, Iran, in 2003, and the Ph.D. degree in artificial intelligent of computer engineering from Shiraz University, in 2016. His research interests include texture classification, pattern recognition, and computer vision.

**REZA BOOSTANI** received the B.S. degree in electronic engineering from Shiraz University, Shiraz, Iran, in 1996, and the M.S. and Ph.D. degrees in biomedical engineering from Amir Kabir University, Tehran, Iran, in 1999 and 2005, respectively. He is currently an Associate Professor with Shiraz University. His research interests include statistical pattern recognition, signal processing, and biomedical engineering.