The Visual Computer
https://doi.org/10.1007/s00371-018-1503-0

ORIGINAL ARTICLE

Image classification by combining local and global features

Leila Kabbai[1] · **Mehrez Abdellaoui**[2] · **Ali Douik**[1]

Abstract
Several techniques have recently been proposed to extract the features of an image. Feature extraction is one of the most important steps in various image processing and computer vision applications such as image retrieval, image classification, matching, object recognition. Relevant feature (global or local) contains discriminating information and is able to distinguish one object from others. Global features describe the entire image, whereas local features describe the image patches (small group of pixels). In this paper, we present a novel descriptor to extract the color-texture features via two information types. Our descriptor named concatenation of local and global color features is based on the fusion of global features using wavelet transform and a modified version of local ternary pattern, whereas, for the local features, speeded-up robust feature descriptor and bag of words model were used. All the features are extracted from the three color planes. To evaluate the effectiveness of our descriptor for image classification, we carried out experiments using the challenging datasets: New-BarkTex, Outex-TC13, Outex-TC14, MIT scene, UIUC sports event, Caltech 101 and MIT indoor scene. Experimental results showed that our descriptor outperforms the existing state-of-the-art methods.

Keywords SURF · BoW · LBP · LTP · Wavelet transform · Image classification

1 Introduction

The main goal of feature extraction is to extract the most relevant information from the image in order to obtain a fine and robust descriptor. Feature extraction is a fundamental step in computer vision and image processing tasks like image retrieval [1], image classification [2], texture classification [3,4], object recognition [5], image matching [6], remote sensing [7], facial classification [8]. Nowadays, feature extraction is considered as a challenging issue as the features vary significantly due to many factors like scale variations, noise, illumination and viewpoint changes.

In the literature, there are three methods to extract the features of an image: low-level methods describe an image with a feature vector using low-level visual attributes such as scale-invariant feature transform (SIFT) [5], speeded-up robust feature (SURF) [9], local binary pattern (LBP) [10]. Mid-level methods attempt to develop a global image representation via the statistical analysis of the extracted local visual attributes [11]. The bag of words (BoW) model [12] is one of the most popular mid-level methods. It is efficient to solve the problem of image classification [13]. High-level methods are widely used for deep learning. In general, deep learning algorithms learn high-level semantic features automatically instead of requiring handcrafted features [14]. Some approaches based on deep convolutional neural networks (CNNs) [15] architecture have achieved success in image classification [16–19]. CNNs and other deep networks are generally trained on a large number of data.

Although, several efforts have recently been made to improve the low-level feature descriptor, it is still remains an open research area. The extraction of the best features is a crucial step to enhance the classification performance leading to a robust system to these challenges [20]. The extraction of an image feature can be classified into two categories: global features which describe the visual content of the entire image by a single vector. They represent the texture, color, shape information which are the most popular for image represen-

✉ Leila Kabbai
kabbai.leila@gmail.com

Mehrez Abdellaoui
mehrez.abdellaoui@enim.rnu.tn

Ali Douik
ali.douik@enim.rnu.tn

[1] National Engineering School of Sousse- ENISo, University of Sousse, Sousse, Tunisia

[2] High Institute of Applied Technologies of Kairouan, University of Kairouan, Kairouan, Tunisia

Published online: 07 April 2018

tation. Differently from global features, local features aim to detect the interest points (IPs) in an image and describe them by a set of vectors. Recently, many researchers have focused on extracting the IPs, texture and color features of an image.

In the literature, a large variety of texture descriptions have been presented: Ojala et al. [10] proposed a new operator named local binary pattern (LBP) for texture classification. Some variants of the LBP descriptor have been presented to enhance its performance. Thus, Guo et al. [21] developed a completed LBP (CLBP) which is based on a combination of three operators, namely CLBP-Sign (CLBP_S), CLBP-Magnitude (CLBP_M) and CLBP-Center (CLBP_C) to attain the rotation invariant for texture classification. Tan et al. [22] presented local ternary pattern (LTP) that quantizes the difference between the central pixel and its eight neighbors into three levels. Nishant et al. [23] developed a robust framework of LBP, named completed local structure pattern (CLSP) and robust local structure pattern (RLSP) for image texture classification. Also, Xiaosheng et al. [24] presented a new feature descriptor called joint-scale LBP. In 2017, Rahman et al. [25] developed a discriminative ternary census transform histogram (DTCTH) for image classification. In the case of illumination variations, Khan et al. [26] proposed a way to normalize the classical texture features based on frequency decomposition. Alvarez et al. [27] presented a descriptor based on the BoW framework to represent color-texture image. Zhu et al. [28] exploited an orthogonal combination of LBP (OC-LBP) descriptor that aims to increase both discriminating power and photometric invariance properties. Porebski et al. [29] presented a descriptor named multi-color space feature selection for evaluating color-texture classification. Cusano et al. [30] proposed a new descriptor by combining LBP and local color contrast for texture classification under varying illumination. Sandid et al. [31] developed a local combination adaptive ternary pattern (LCATP) to encode both color and local feature for texture classification. Ledoux et al. [3] developed an LBP extension that takes the vector information of color into account thanks to a color order for texture classification. Kabbai et al. [32] proposed a new descriptor named upper and lower LBP (UL_{LBP}) that aims to improve the photometric invariant and reduce the size compared to the original LBP. Kabbai et al. [33] also developed a novel descriptor named Histogram of Local and Global features using SURF descriptor (HLF_{SURF}) to apply in image retrieval and classification. Papadopoulos et al. [34] proposed an approach by Combining Global and Local information for knowledge-assisted image analysis and classification. Based on color, texture, shape, and wavelets, Banerji et al. [35] developed a new descriptor named H-fusion. Also, Sinha et al. [36] proposed a descriptor by fused color Gabor-pyramid of histograms of oriented gradients (FC-GPHOG). Moreover, many researchers proposed new descriptors to use for object recognition [37–39] and scene image classification [2,40–42].

However, global features have a few limitations, like sensitivity to noise, illumination variation, scaling, and failure to identify the important features of the image. They are not suitable for some applications. Their drawbacks are resolved by the local features which encode the local information in order to obtain better details of the image like interest points (IPs). Different approaches have been proposed to detect the IPs such as Harris and Stephens [43], Harris and Lapace [44], SIFT [5]. Many extensions of SIFT descriptor were developed, for instance the principal component analysis SIFT (PCA-SIFT) [45], the speeded up robust feature (SURF) algorithm [9]. Over the last few years, IPs have been used for applications like image matching [46] and image classification [47].

Local and global features are two approaches which provide different information of image at the computational level. In the literature, several applications combined local and global features in different domains such as object recognition [48], image retrieval [33,49,50] and image classification [33,51].

This paper presents a novel descriptor based on a combination of the local and global features of three color planes. Thus, SURF algorithm is the most popular local feature which provides good effectiveness and great discriminative power in image classification. This algorithm was proposed to detect the local IPs and generate a descriptor. The number of local features for each image is huge. To overcome this drawback, the model BoW (bag of words) is used to quantize feature descriptors into visual words (codebook). It encodes the local descriptors into a histogram representation. Each image is represented by a k-bins histogram, k being the codebook size. BoW has become popular in the recent years thanks to its effectiveness and its results quality. For global feature, we propose a novel operator, variant of LTP, named elliptical upper and lower local ternary pattern operator (EUL_{LTP}) which is more robust to noise and illumination transformation by replacing the value of each center pixel in 3×3 by its average local gray value of the neighborhood. However, it reduces its dimensionality compared to the original LTP while keeping discriminative power. Also, wavelet transform (WT) is one of the most used texture features. By combining the two approaches of three color planes, we developed a new descriptor named concatenation of local and global color features (CLGC). Experimental results showed that CLGC descriptor achieves higher classification rates than other descriptors.

The rest of this paper is organized as follows: Sect. 2 gives a brief account of the feature descriptors. The proposed method is described in Sect. 3. Section 4 illustrates the exper-

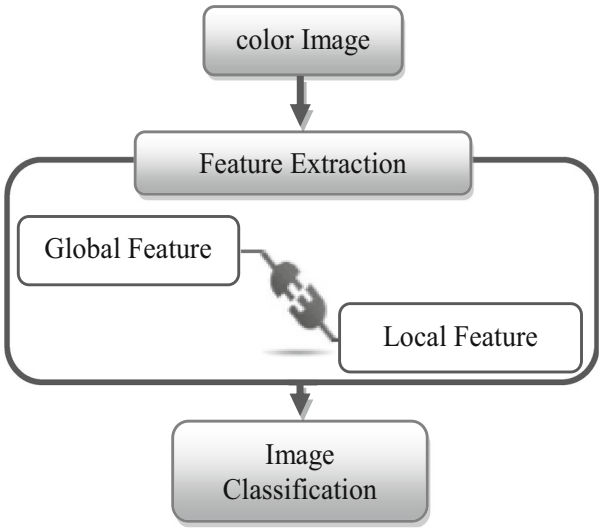

Fig. 1 Synoptic schematics for image classification

imental results on texture and image classification. Finally, a conclusion is presented.

2 Feature descriptors

In recent years, many algorithms are used to extract the local or global features of an image. Global features describe the visual content of the whole image which represents an image by one vector, whereas the local features extract the IPs of image and describe them as a set of vectors. Figure 1 presents the synoptic schematics for image classification.

The speeded-up robust features (SURF) descriptor is inspired by the SIFT descriptor to compute distinctive invariant local features proposed by Bay et al. [9]. This algorithm is used to detect the IPs and describe the features that presented some invariance to image noise, scaling, rotation and change of illumination and viewing direction. SURF is more robust and faster than SIFT. In addition, wavelet transform (WT) is one of the most used descriptors to extract the texture feature from the entire image. The computation of the WT in a two-dimensional space involves analyzing an image in multi-scale and multi-resolution of both space and frequency domains [52]. At each level, the image is decomposed into four frequency sub-bands, LL, LH, HL, and HH, where L and H are, respectively, low, and H denotes high frequency. Smith et al. [53] used the mean and variance of the wavelet coefficients as texture features. Moreover, local binary pattern (LBP) operator was developed by Ojala et al. [10] to extract the texture feature from an image. This operator is robust against illumination changes and presents a computational simplicity and ability to encode texture. The LBP code is computed by comparing the center pixel with its 8 neighbors in a 3-by-3 kernel. If the value of the neighborhood is bigger than the value of the central pixel, the result is set to '1'. Otherwise, it is set to '0'. The final LBP code is obtained by multiplying the results by weights given by powers of two and summing the different computed values, as defined in Eq. (1).

$$\begin{aligned} \mathrm{LBP}_{P,r}\,(x_\mathrm{c}, y_\mathrm{c}) &= \sum_{i=0}^{P-1} S\,(g_\mathrm{i} - g_\mathrm{c}) \times 2^\mathrm{i} \\ S(d) &= \begin{cases} 1\,; & \text{if } d \geq 0 \\ 0\,; & \text{otherwise} \end{cases} \end{aligned} \tag{1}$$

where S stands for the function sign, P and r are the total number and the radius of neighboring pixels. g_c and g_i are, respectively, the gray level of central pixel and the neighboring one. $(x_\mathrm{c}, y_\mathrm{c})$ is the coordinate of the central pixel. d is the difference between the neighboring and the center pixel. The LBP code tends to produce a histogram of the size 256 bins.

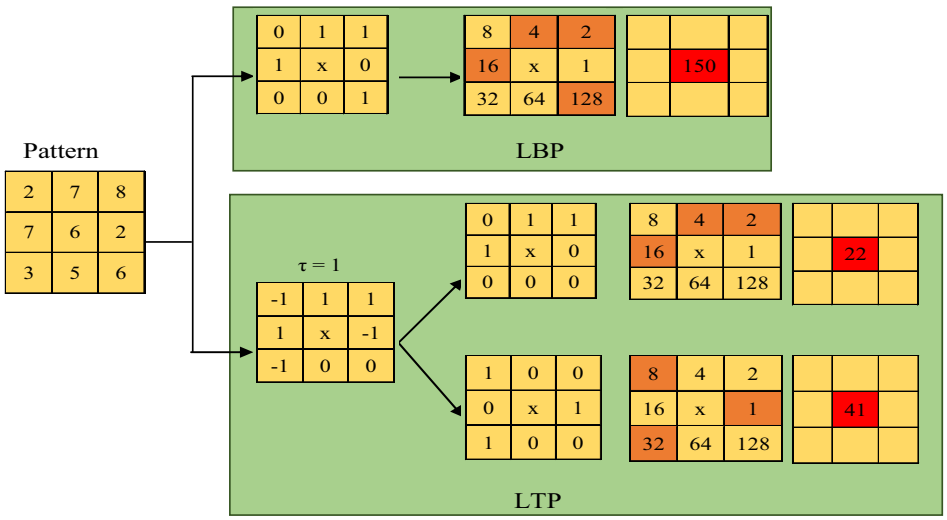

Fig. 2 Example to calculate LBP and LTP for 3 × 3 pattern

Inspired by LBP, Tan and Triggs [22] proposed the local ternary pattern (LTP) operator. This operator allows to solve some problems in the LBP by creating three patterns instead of two and using a threshold τ to tackle intensity fluctuation. LTP operator is less sensitive to noise in uniform regions and has more discriminative power. Thus, the LTP code is computed as follows:

$$\mathrm{LTP}_{P,r}(x_c, y_c) = \sum_{i=0}^{P-1} \mathrm{S}(g_i - g_c) \times 3^i$$
$$S(d) = \begin{cases} 1; & \text{if } d \geq \tau \\ -1; & \text{if } d \leq -\tau \\ 0; & \text{otherwise} \end{cases} \quad (2)$$

In order to reduce the feature size, an LTP code is split into two binary codes (upper and lower pattern). In addition, these two histograms are concatenated to build a feature vector of an image with a size of 512 bins. Figure 2 illustrates an example of the LBP and LTP encoding procedure.

3 Proposed approach

3.1 Elliptical upper and lower local ternary pattern

In this paper, a new variant of LTP so-called elliptical upper and lower local ternary pattern operator ($\mathrm{EUL_{LTP}}$) is proposed. This operator allows to extract the texture feature of an image using the vertical and horizontal motif ellipse patterns. The idea of the proposed method is to use the pixel value of the different directions. Firstly, we take a 5×5 local window around a center pixel as shown in Fig. 3. Secondly, each pixel around the center pixel is computed by the average of the pixel's neighborhood. For instance, the pixels which are found in the horizontal or vertical motif around the center pixel (x0) are calculated by the average of the three neighborhoods [before, after and right or (left)]. The pixels which are found in the diagonals are computed by the average with the two neighborhoods (before and after).

Finally, we get a new motif noted elliptical local pattern (ELP) of the size (3×3) as illustrated in Fig. 3. The parameters of the new pattern are presented as follows:

$$\begin{aligned} w1 &= \frac{1}{4}(x1 + x2 + x8 + x9) \\ w2 &= \frac{1}{3}(x1 + x2 + x3) \\ w3 &= \frac{1}{4}(x2 + x3 + x4 + x13) \\ w4 &= \frac{1}{3}(x3 + x4 + x5) \\ w5 &= \frac{1}{4}(x4 + x5 + x6 + x17) \\ w6 &= \frac{1}{3}(x5 + x6 + x7) \\ w7 &= \frac{1}{4}(x6 + x7 + x8 + x21) \\ w8 &= \frac{1}{3}(x7 + x8 + x1) \end{aligned} \quad (3)$$

where R1 and R2 represent, respectively, the distance between central and neighboring pixels horizontal and vertical directions. The direction of the Ellipse depends on the

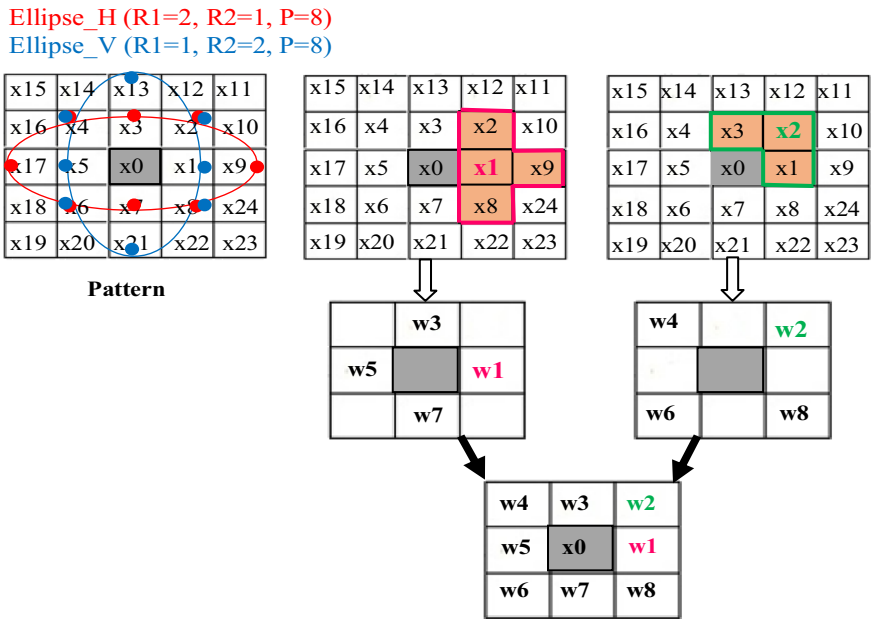

Fig. 3 New elliptical local pattern (ELP) motif

R1 and R2 values. If R1 > R2, we have a horizontal ellipse; otherwise, we obtain a vertical ellipse.

The value of the center pixel for the new motif is the average local gray level (ALG) of all pixels in the motif. The ALG is insensitive to noise and more robust to illumination variants when compared to gray level value. Equation (4) illustrates the calculation of ALG that is simple but effective.

$$\mathrm{ALG}\,(x_c, y_c) = M_c = \frac{\sum_{i=1}^{P}(w_i) + x_0}{9} \tag{4}$$

In this paper, we propose a new operator inspired from the traditional LTP named elliptical upper and lower LTP ($\mathrm{EUL_{LTP}}$) based on the new elliptical local pattern (ELP) motif. Our operator can reduce the dimensionality of the LTP operator while keeping robustness against noise interference and strong discriminative ability for describing texture analysis.

In this paper, we divide the ternary pattern into two binary patterns: positive binary pattern (PBP) and negative binary pattern (NBP) presented as follows:

$$\mathrm{PBP}\,(x_c, y_c) = \mathrm{S}\,(w_i - M_c) \quad \text{where} \quad S\,(d) \begin{cases} 1 & \text{if } d \geq +\tau_1 \\ 0 & \text{if otherwise} \end{cases} \tag{5}$$

$$\mathrm{NBP}\,(x_c, y_c) = \mathrm{S}\,(w_i - M_c) \quad \text{where} \quad S\,(d) \begin{cases} 1 & \text{if } d \leq -\tau_1 \\ 0 & \text{if otherwise} \end{cases} \tag{6}$$

Thus, we compute only the difference between the neighboring and the new center pixel in order to obtain a binary code, with τ_1 being the threshold.

After that, we divide the 8 neighboring binaries of the positive binary pattern (PBP) into two groups and compute the code separately for each one. Group 1 contains the four upper neighboring binaries (UNB) (w1, w2, w3, w4) and group 2 consists of the four lower neighboring binaries (LNB) (w8, w7, w6, w5) as illustrated in Fig. 4. To encode each group, it is necessary to multiply the four UNB and the four LNB with the weights given by powers of two to give, respectively, $\mathrm{UNB_{PLTP}}$ (positive upper neighboring binaries of LTP) and $\mathrm{LNB_{PLTP}}$ (positive lower neighboring binaries of LTP), as presented below. Each group has a size of 16 bins.

Positive part

$$\mathrm{UNB_{PLTP}} = \sum_{i=1}^{\left(\frac{P}{2}\right)} \mathrm{PBP}_i \times 2^{i-1} \tag{7}$$

$$\mathrm{LNB_{PLTP}} = \sum_{i=1}^{\left(\frac{P}{2}\right)} \mathrm{PBP}_{(P-i)+1} \times 2^{i-1} \tag{8}$$

The same step is applied for the negative binary pattern (NBP) to obtain, respectively, $\mathrm{UNB_{NLTP}}$ (negative upper neighboring binaries of LTP) and $\mathrm{LNB_{NLTP}}$ (negative lower neighboring binaries of LTP) as follows:

Negative part

$$\mathrm{UNB_{NLTP}} = \sum_{i=1}^{\left(\frac{P}{2}\right)} \mathrm{NBP}_i \times 2^{i-1} \tag{9}$$

$$\mathrm{LNB_{NLTP}} = \sum_{i=1}^{\left(\frac{P}{2}\right)} \mathrm{NBP}_{(P-i)+1} \times 2^{i-1} \tag{10}$$

Equation (11) presents the concatenation of the histograms of $\mathrm{UNB_{PLTP}}$, $\mathrm{LNB_{PLTP}}$, $\mathrm{UNB_{NLTP}}$, and $\mathrm{LNB_{NLTP}}$ to obtain a novel descriptor named elliptical upper and lower LTP ($\mathrm{EUL_{LTP}}$) of a 64-bin size, which is more compact than the original LTP.

$$\mathrm{EUL_{LTP}} = [\mathrm{UNB_{PLTP}}; \mathrm{LNB_{PLTP}}; \mathrm{UNB_{NLTP}}; \mathrm{LNB_{NLTP}}] \tag{11}$$

Figure 4 illustrates the steps to compute the $\mathrm{EUL_{LTP}}$ feature.

3.2 Concatenation of local and global color features

The main objective from developing a significant descriptor is to increase system performance. In this paper, we propose a new descriptor by concatenating local and global features based on the color-texture information for image classification. To extract the local features of an image, we use SURF algorithm which is the most popular local feature representation. It was used to detect the IPs and generate many descriptors per image. The number of IPs for each image is huge, to resolve this weakness, the BoW model was used to quantize descriptors feature into visual words (codebook). It encodes the local descriptors into a histogram representation using k-means algorithm to cluster these descriptors into a codebook. Each image is represented by a k-bins histogram, k being the codebook size. In this work, we concatenate the best local features information extracted from each color plane as illustrated in Fig. 5.

For global features, we take the best information of each color plane using wavelet transform (WT) with four levels in order to extract the standard deviation and the mean wavelet transform coefficients to build a 40-bin sized vector. By fusing the three histograms of the three planes (R, G, B),we obtained a vector called Color WT (CWT). Then, the process is repeated using the new operator ($\mathrm{EUL_{LTP}}$) to obtain a 64-bin vector for each plane color. By concatenating the three histograms, a vector noted Color $\mathrm{EUL_{LTP}}$ ($\mathrm{CEUL_{LTP}}$) is generated.

By concatenating the two vectors, we generate a novel histogram named $\mathrm{CWEUL_{LTP}}$ which represents the global features. Figure 5 presents the different steps for generating the $\mathrm{CWEUL_{LTP}}$ descriptor with a size of 312 bins. Finally,

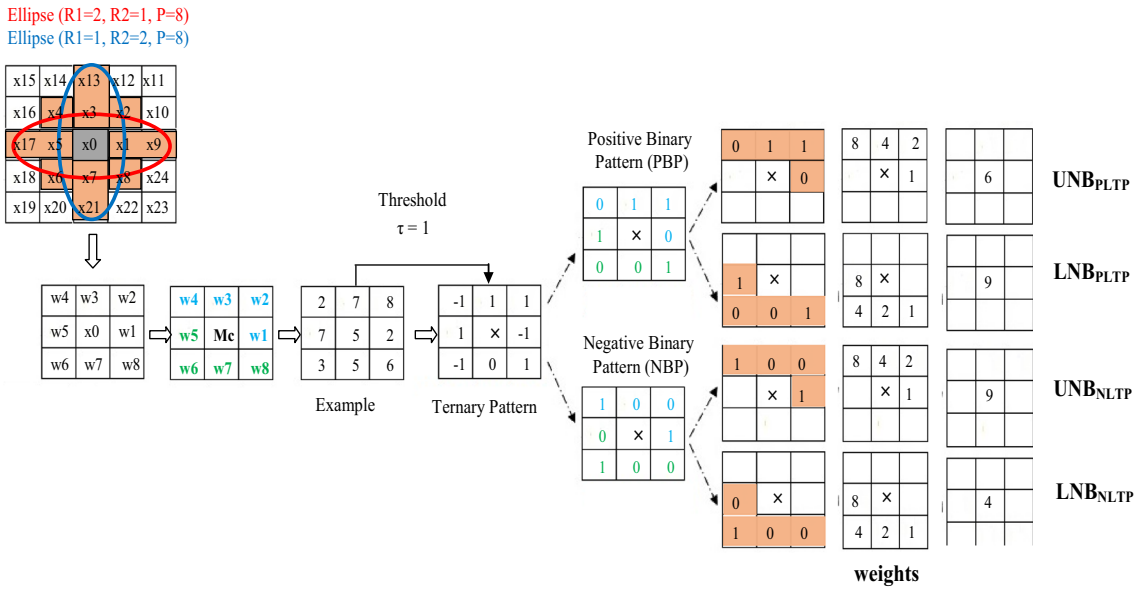

Fig. 4 The flowchart of computing EUL_{LTP}

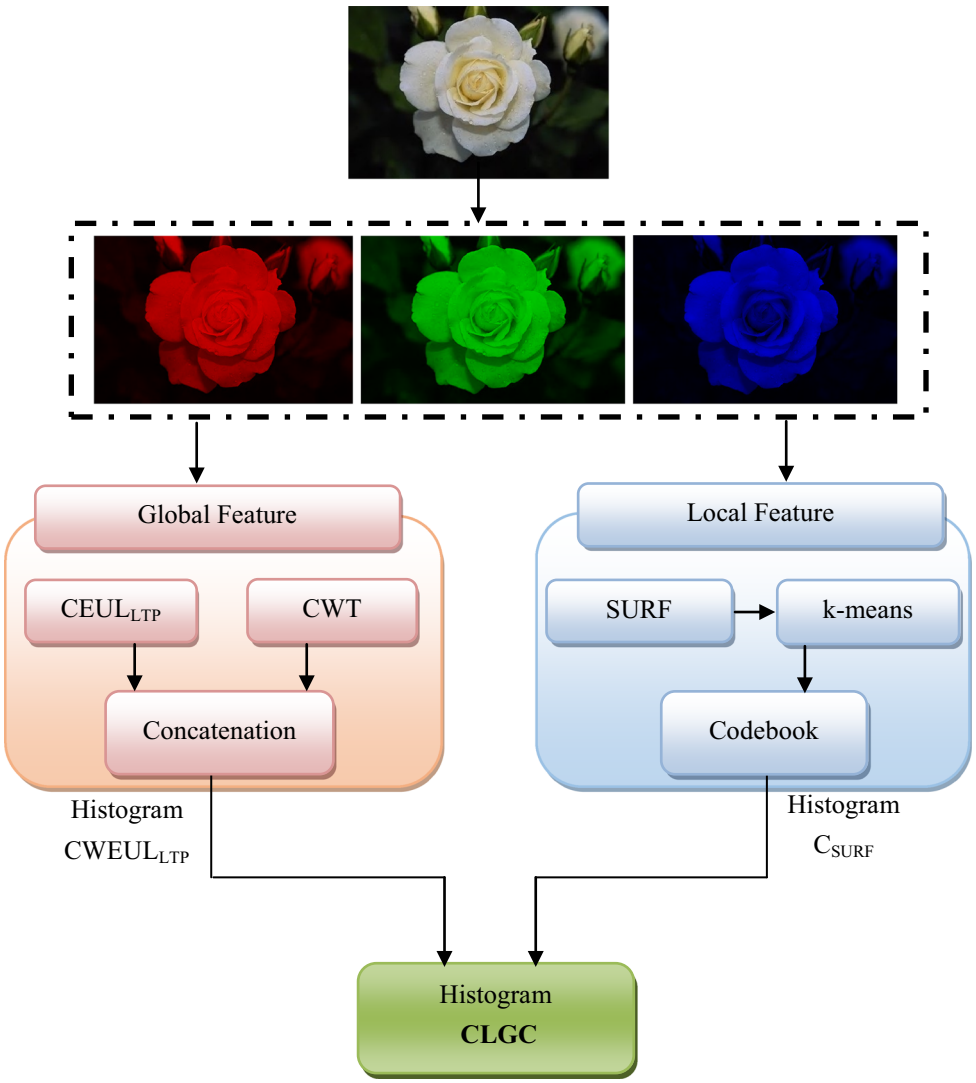

Fig. 5 Overview of the proposed descriptor

by combining the two vectors local and global features, we create a final descriptor called concatenation of local and global color features (CLGC). Figure 5 presents an overview of the proposed descriptor.

4 Experimental results

In this section, we assess the performance of our proposed descriptors $CWEUL_{LTP}$ and CLGC for image classification. Extensive experiments were carried out on some datasets to evaluate the robustness of our approach. We start by presenting the different datasets used. Then, we discuss the obtained results.

4.1 Datasets

4.1.1 New-BarkTex dataset

The New-BarkTex dataset[1] [29] contains a total of 1632 color images classified into six tree bark classes. Each class contains 272 images, of 64 × 64 pixels. Half of the images are used for training while the rest for testing [29]. The images of the dataset characterize the natural color textures acquired under natural illumination conditions. All the images of the BarkTex dataset represent a single type of object, as bark of trees. Figure 6 illustrates some sample images from New-BarkTex dataset.

4.1.2 Outex-TC13 dataset

The Outex-TC13 dataset[2] [54] contains 1360 images, of a size of 128 × 128 pixels divided into 68 classes of different color textures. Half of the images are used for training while the rest for testing. Figure 7a illustrates some classes of the TC13 dataset. The specificity of this database is that it contains several categories with similar color and texture, resulting in a high inter-class similarity. The Outex-TC13 dataset base is a publicly accessible dataset representing a wide variety of textural material acquired under the same controlled conditions.

4.1.3 Outex-TC14 dataset

The Outex-TC14 dataset[3] includes 4080 images in 68 classes of different color textures collected under three different illuminants: Inca, Horizon and TL84. Each image is split into 20 non-overlapping sub-images having a size of 128 × 128 pixels; then, each illuminant has 1360 images. Half of the images under the Inca illuminant are selected as the training set, and half of the images under the two other illuminants (TL84 and Horizon) are selected as the testing set. The total numbers of training and test images are, respectively, 680 and 1360. Figure 7b illustrates the Outex-TC14 dataset of five textures acquired under three different illuminants.

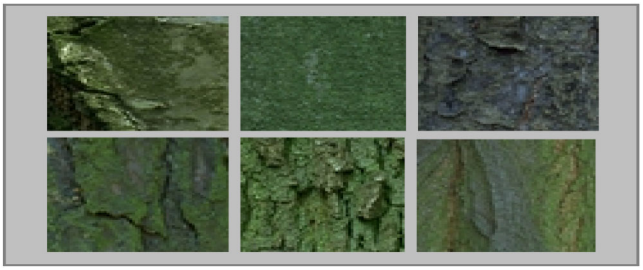

Fig. 6 Class samples from the New-BarkTex dataset

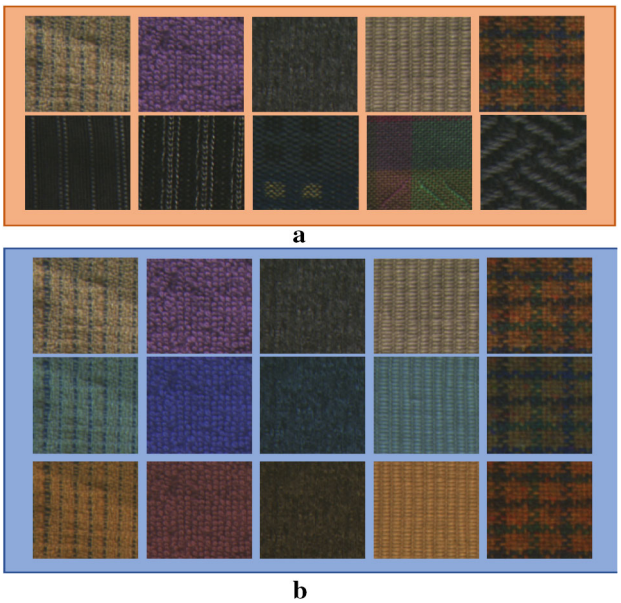

Fig. 7 Class samples from: **a** the Outex-TC13 dataset, **b** Outex-TC14 Dataset of five textures acquired under three different illuminants

4.1.4 MIT scene dataset

The MIT scene dataset[4] [55] holds 2688 color images divided into 8 scene categories: coast, forest, highway, inside cities, mountain, open country, streets and tall buildings. Each category contains a minimum of 260 images and a maximum of 410 images. These images are JPEG format with a size of 256 × 256 pixels. There is a large variation in light, pose and angles, along with a high intra-class variation. Some images of this dataset are given in Fig. 8.

[1] https://www-lisic.univ-littoral.fr/~porebski/BarkTex_image_test_suite.html.

[2] http://www.outex.oulu.fi/index.php?page=classification.

[3] http://www.outex.oulu.fi/index.php?page=classification.

[4] http://people.csail.mit.edu/torralba/code/spatialenvelope/spatial_envelope_256x256_static_8outdoorcategories.zip.

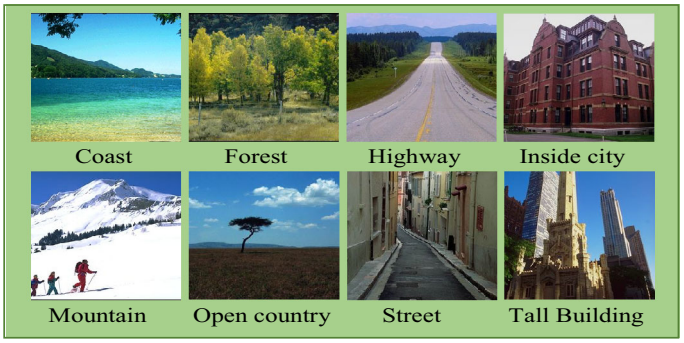

Fig. 8 Some images of the MIT scene dataset

Fig. 9 Some images of the UIUC sports event dataset

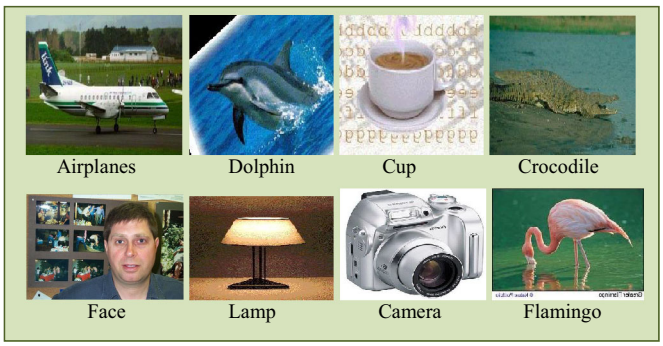

Fig. 10 Some images of the Caltech-101 dataset

4.1.5 UIUC sports event dataset

The UIUC sports event dataset[5] [41] includes a total of 1579 color images classified into eight Sports Event categories: rowing (250 images), badminton (200 images), polo (182 images), bocce (137 images), snowboarding (190 images), croquet (236 images), sailing (190 images) and rock climbing (194 images). Figure 9 illustrates some sample images from UIUC Sports Event dataset.

4.1.6 Caltech-101 dataset

Caltech101 dataset[6] contains 9144 images divided into 101 object categories (Airplane, Crocodile, Faces, Camera, Flamingo, etc) and one background category. The image number per category ranges from 31 to 800. Following the standard experimental setting, we randomly select 30 images for training and test on up to 50 images. Some images of this dataset are presented in Fig. 10.

4.1.7 MIT indoor scene dataset

The MIT indoor scene dataset[7] holds a total of 15620 images of 67 indoor scene classes. The number of images varies from 101 to 734 per category like classroom (113 images), mall (176 images), closet (135 images), etc. All images are in JPEG format. Following the standard experimental setting, we randomly select 80 images for training and 20 images for testing. Some images of this dataset are given in Fig. 11.

[5] http://vision.stanford.edu/lijiali/event_dataset/.

[6] http://www.vision.caltech.edu/Image_Datasets/Caltech101/101_ObjectCategories.tar.gz .

[7] http://groups.csail.mit.edu/vision/LabelMe/NewImages/indoorCVPR_09.tar.

Fig. 11 Some images of the MIT indoor scene dataset

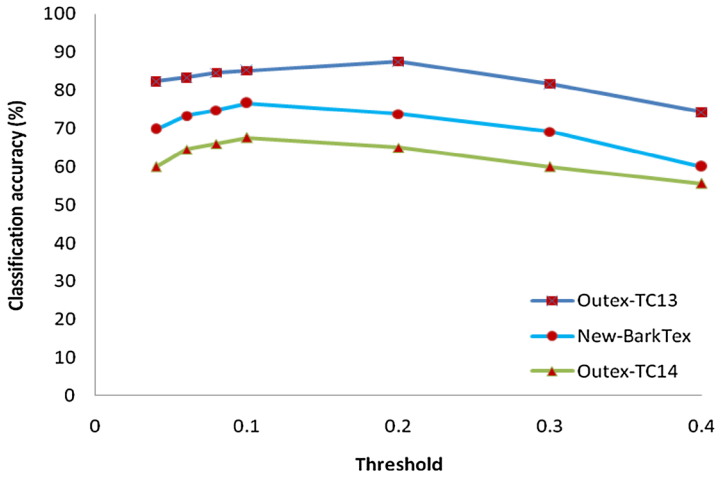

Fig. 12 Classification accuracy of $CWEUL_{LTP}$ with different thresholds

4.2 Evaluation of the experimental results

To validate the best adopted classifier, many experiments were achieved using different datasets. For global features, we tried the descriptors $CEUL_{LTP}$ and CWT separately then their concatenation using image from New-BarkTex, Outex-TC13 and Outex-TC14 datasets. Also, we tried the descriptors $CWEUL_{LTP}$ and CLGC to find the best descriptors using MIT scene, UIUC sports event, Caltech 101 and MIT indoor scene datasets.

The classification rate is adopted as the performance measure in the experiments. To obtain the best reliable result, we run the experiments with several randomly chosen images.

4.2.1 Evaluation of the experimental results on image texture classification

To evaluate both the discriminating power and robustness of color-texture against illumination changes, we used Outex-TC13, Outex-TC14 and New-BarkTex datasets, in which the training and test images are acquired under the same acquisition conditions or under different illuminants. The different trials are conducted using the nearest neighbor classifier based on the Chi-square distance for texture classification.

Table 1 Classification rates (%) on the Outex-TC13, New-BarkTex and Outex-TC14 datasets

Methods	Outex-TC13	New-BarkTex	Outex-TC14
C_{LTP}	83.5	73	63.5
CWT	64.4	58.3	50.2
$CEUL_{LTP}$	85.1	75.1	65.9
$CWEUL_{LTP}$	87.4	76.6	67.5

In order to get the best classification results, It is necessary to find the best threshold. Thus, we conducted a series of trials using different thresholds $\tau = \{0.04, 0.06, 0.08, 0.1, 0.2, 0.3, 0.4\}$ on several datasets using the fusion descriptor $CWEUL_{LTP}$ as illustrated in Fig.12.

In this figure, we observe that the best accuracy results are found with threshold 0.2 for Outex-TC13 around 87.4%. Thus, for New-BarkTex and Outex-TC14, the best threshold is found with value 0.1. In order to assess the performance and the efficiency of the descriptor, we compare the classification rate between the different descriptors C_{LTP}, CWT, $CEUL_{LTP}$ and $CWEUL_{LTP}$ as defined in Table 1. In this table, we see that $CWEUL_{LTP}$ reached a higher success rate than other descriptors for the three texture datasets.

In addition, we compare the classification results of our descriptor $CWEUL_{LTP}$ with other descriptors in the state of

Table 2 Classification rates (%) achieved on the Outex-TC13, New-BarkTex and Outex-TC14 datasets

Methods	Outex-TC13	New-BarkTex	Outex-TC14
[3]	83.4	72.5	64.8
[26]	84.8	–	64.9
[27]	86.7	–	–
[28]	–	–	65.4
[29]	–	75.9	–
[30]	85.3	–	–
$CWEUL_{LTP}$	87.4	76.6	67.5

the art. Detailed comparison results are shown in Table 2. From these results, we show that our proposed descriptor $CWEUL_{LTP}$ outperforms the results of the state of the art for the three datasets. Moreover, the $CWEUL_{LTP}$ descriptor reached a success rate of 87.4% which is superior to that of [3,26,27] and [30] by using Outex-TC13. For New-BarkTex dataset, the proposed descriptor $CWEUL_{LTP}$ achieved a success rate of 76.6% which is higher than the results provided by [29] and [3] which have 75.9 and 72.5%, respectively. As illustrated in Table 2, $CWEUL_{LTP}$ descriptor gives the best classification rate, around 67.5% which is superior to [3,26] and [28] with values around 64.8, 64.9 and 65.4%, respectively. These results show that the $CWEUL_{LTP}$ operator is the best descriptor for texture discrimination.

4.2.2 Evaluation of the experimental results on image classification

In this section, we assess the performance and robustness of our proposed CLGC descriptor for image classification using the nature scene image. Many experiments were carried out on MIT scene, UIUC sport event, Caltech 101 and MIT indoor scene datasets to evaluate the effectiveness of our descriptor. As evaluation criteria, the classification accuracy is the easiest classification metric.

In this work, we take the average results after running the experiments with several randomly selected images for training and testing in order to obtain the best reliable result. The different trials are conducted using the least squares support vector machines (LS-SVMs) classifier [56]. We train the LS-SVM classifier using the training set and then we settle the parameters in the validation phase and obtained the classification results.

The size of CLGC descriptor varies with the size of the codebook. To assess the performance of codebook, we chose seven increasing sizes: 100, 256, 512, 700, 800, 1000 and 1024.

Figure 13 illustrates the impact of codebook for classification accuracy obtained with CLGC descriptor for two datasets. We observe that the classification accuracy improved with the increase in codebook size. Considering both efficiency and accuracy, 1024 is a preferred choice. Finally, the size of CLGC descriptor is equal to 1336 bins.

a. Image scene classification on the MIT scene dataset

In this section, we present two ways to extract features using our descriptor CLGC. First, we extract the local and global features using RGB color space to obtain CLGC(RGB-RGB). Second, we extract the local features using RGB color space and global features using HSV color space to achieve CLGC(RGB-HSV).

In the first experiment on the MIT scene dataset, we randomly choose 100 images per class for the training phase and the rest for testing. Then, we carry out a comparative evaluation of our proposed CLGC descriptor with other descriptors in the state of the art. Table 3 depicts the results obtained with different descriptors.

The classification rate of $CEUL_{LTP}$(RGB) descriptor is around 72.7%, which is superior to CWT(RGB) and C_{LTP}(RGB). By fusing the two descriptors CWT(RGB) and $CEUL_{LTP}$(RGB), the $CWEUL_{LTP}$(RGB) descriptor obtains the best results compared to each descriptor separately.

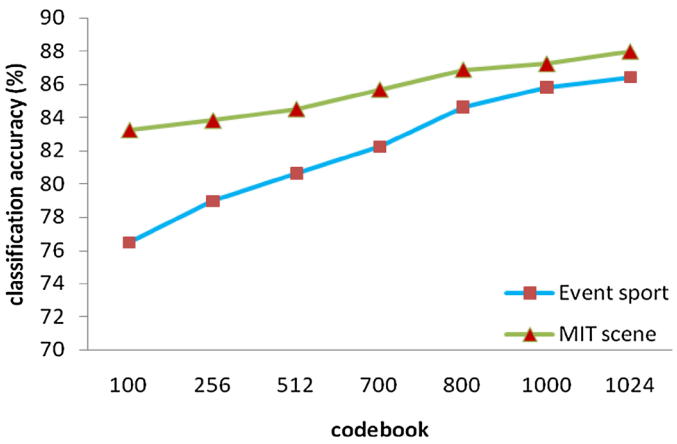

Fig. 13 Classification rate of CLGC with various sizes of the codebook

Table 3 Classification rates (%) per class on the MIT scene dataset

Class	C_{LTP} (RGB)	CWT (RGB)	$CEUL_{LTP}$ (RGB)	$CWEUL_{LTP}$ (RGB)	$CWEUL_{LTP}$ (HSV)	C_{SURF} (RGB)	CLGC (RGB-RGB)	CLGC (RGB-HSV)
Coast	75	65	79	75.4	76.5	78.	87.7	91.5
Forest	80.5	75.5	84.2	87.3	88	94.7	92.9	95.7
Highway	70	61.4	76	78.1	80.6	85.6	89.7	93.8
Inside city	73	64	75	76.9	78.4	81.3	86.6	89.9
Mountain	67.4	57	65.4	68.6	70.8	75.6	89.9	94.3
Open country	64.8	55	60.8	63.5	63.9	70.9	79.6	82.4
Street	70	64	74	78.1	81	87.5	86.5	89.2
Tall building	63	50.8	67	71.5	75.6	89.9	91.5	93.8
Average	70.5	61.6	72.7	74.8	76.8	82.9	88	91.3

Table 4 Comparison of classification rates (%) on the MIT scene dataset

#Train = 100, # test = rest		
Methods		Accuracy (%)
CLGC(RGB-RGB)	Our descriptor	88.0
CLGC(RGB-HSV)	Our descriptor	91.30
DTCTH(LSVM)	[25]	87.88
DTCTH(HI)	[25]	89.18
H-fusion	[35]	87.70
HLG_{SURF}	[33]	86.90
FC-GPHOG	[36]	86.0
CGLF + PHOG	[2]	84.30

Moreover, by combining $CWEUL_{LTP}$(RGB) and C_{SURF} (RGB) descriptors, we observe that CLGC(RGB-RGB) descriptor has the best classification accuracy, around 88% compared to the other descriptors. We also see that the average classification rate $CWEUL_{LTP}$(HSV) is superior to $CWEUL_{LTP}$(RGB). Their combination leads to CLGC(RGB-HSV) that has better results than CLGC(RGB-RGB).

Table 4 illustrates the classification rates obtained by our descriptor and other descriptors of the state of the art. CLGC(RGB-RGB) and CLGC(RGB-HSV) descriptors have the best classification rates around 88 and 91.3%, respectively, compared to other descriptors.

b. Image event classification on the UIUC sports event dataset

In the second experiment on Sports Event dataset, we further made a comparative evaluation of our descriptor CLGC with some popular descriptors used by other researcher for image classification. In the experiment, we used randomly chosen 70 images per class for the training phase and 60 images for testing.

Table 5 depicts the average classification rates per class. In this table, we see that CLGC(RGB-RGB) descriptor reached a better success rate around 86.4% than that of $CWEUL_{LTP}$(RGB) and C_{SURF} (RGB) which achieved 77.7 and 81.7%, respectively. The table also shows that $CWEUL_{LTP}$(RGB) has the best results compared to C_{LTP} (RGB), CWT(RGB) and $CEUL_{LTP}$(RGB). Moreover, $CWEUL_{LTP}$(HSV) has better classification accuracy, around 80.5%, than $CWEUL_{LTP}$(RGB). We also observe that CLGC(RGB-HSV) has a better success rate, around 90.85%, than that of CLGC(RGB-RGB).

Table 6 presents the classification rates obtained by our descriptor and some popular descriptors used by other researchers. As can be seen in this table, the best classification rates were reached by CLGC(RGB-RGB) and CLGC(RGB-HSV) descriptors with a value of around 86.4 and 90.85% which outperform other descriptors in the state the art [25,35,40–42]. Also, in this table, we see that Places-CNN

Table 5 Classification rates (%) per class on the UIUC sports event dataset

Class	C_{LTP} (RGB)	CWT (RGB)	$CEUL_{LTP}$ (RGB)	$CWEUL_{LTP}$ (RGB)	$CWEUL_{LTP}$ (HSV)	C_{SURF} (RGB)	CLGC (RGB-RGB)	CLGC (RGB-HSV)
Rock climbing	82	66	80.4	84	87.6	93	95	98
Badminton	84.5	70.2	85.6	86	88.5	81.9	86.5	93.9
Bocce	79.4	59.4	80.6	81.7	83	70.6	84	87.9
Croquet	74.5	61	76.5	75	79.8	80.5	78.8	82.6
Polo	67	56.2	72.9	73.5	76	85	90	95.9
Rowing	55	40	61.9	69	73.5	79.8	83.3	85.8
Sailing	73.9	57	75.5	78.7	79.4	81.7	87	91.9
Snowboarding	70	55	71.5	73.7	75.9	80.9	86.7	90.8
Average	73.3	58.1	75.6	77.7	80.5	81.7	86.4	90,85

Table 6 Comparison of classification rates (%) on the UIUC sports event dataset

#Train = 70, # test = 60		
Methods		Accuracy (%)
CLGC(RGB-RGB)	Our descriptor	86.40
CLGC(RGB-HSV)	Our descriptor	90.85
DTCTH(LSVM)	[25]	85.16
DTCTH(HI)	[25]	88.18
H-fusion	[35]	86.20
HMP	[42]	85.70
SIFT+SC	[42]	82.70
OB	[40]	76.30
SIFT+GGM	[41]	73.40
Places-CNN	[17]	94.12
ImageNet-CNN	[17]	94.42
Hybrid-CNN	[17]	94.22
TPN-FS	[19]	95.20

Table 7 Comparison of classification rates (%) on the Caltech 101 dataset

#Train = 30 # test = rest		
Methods		Accuracy (%)
CLGC(RGB-RGB)	Our descriptor	72.63
CLGC(RGB-HSV)	Our descriptor	79.85
DTCTH(LSVM)	[25]	72.26
DTCTH(HI)	[25]	78.56
SPS	[37]	68.40
DRLBP	[38]	66.95
IKSVM	[39]	56.59
CDBN	[18]	65.40
Places-CNN	[17]	65.18
ImageNet-CNN	[17]	87.22
Hybrid-CNN	[17]	84.79
CNN M 2048	[16]	86.64

[17], ImageNet-CNN [17], Hybrid-CNN [17] and TPN-FS [19] have better results compared to our descriptor because deep CNN features are more discriminating but still needs several training conditions (GPU specifications, computing time,…).

c. Image object classification on the Caltech 101 dataset

In the third experiment on Caltech 101 dataset, we compare the performance of our descriptor CLGC against other descriptors used in the state the art for image classification. In this experiment, we randomly select 30 images per class for training and up to 50 images per class for testing.

Table 7 summarizes the classification accuracies of our method with several methods of the state-of-the-art including deep learning. We observe that CLGC(RGB- RGB) and

Table 8 Comparison of classification rates (%) on the MIT indoor scene dataset

# Train = 80 # test = 20		
Methods		Accuracy (%)
CLGC(RGB-RGB)	Our descriptor	60.95
CLGC(RGB-HSV)	Our descriptor	64.26
DTCTH(LSVM)	[25]	43.33
DTCTH(HI)	[25]	46.22
OB	[40]	37.60
Places-CNN	[17]	68.24
ImageNet-CNN	[17]	56.79
Hybrid-CNN	[17]	70.80
TPN-FS	[19]	71.94

CLGC(RGB- HSV) achieve the best accuracy rates around 72.63 and 79.85%, respectively, compared to all the low-level descriptors as DTCTH [25], SPS [37], DRLBP [38] and IKSVM [39] which achieved 78.56, 68.4, 66.59 and 56.59%, respectively. As can be seen from the table, our descriptors outperform the methods CDBN [18] that is based on deep networks. Also, our proposed method reached a better success rate than the Places-CNN [17], which is famous as a deep learning method, with a value of around 65.18%. In this table, we also see that Hybrid-CNN [17] and CNN M 2048 [16] have better results than our descriptors because deep CNN features are more discriminating.

d. Image scene classification on the MIT indoor scene dataset

In the ford experiment on MIT indoor scene dataset, we evaluate the performance of our descriptor CLGC with other descriptors used in the state the art for image classification. In this experiment, we randomly select 80 images per class for training and 20 images per class for testing. Table 8 summarizes the classification accuracies of our method with several methods of the state of the art including deep learning. We show that CLGC(RGB-RGB) and CLGC(RGB-HSV) attain the best accuracy rates around 60.95, 64.26%, respectively, compared to ImageNet-CNN [17], DTCTH [25] and OB [40]. Also, we observe that Places-CNN [17], Hybrid-CNN [17] and TPN-FS [19] methods have slightly higher than our descriptors.

The results of the experiments demonstrate that the CLGC descriptor is the best descriptor for image classification discrimination.

5 Conclusion

In this paper, we propose a new descriptor named concatenation of local and global color features (CLGC) based on the concatenation of the local features using speeded-up robust feature (SURF) descriptor and global feature by fusing the features of wavelet transform (WT) with a modified version of LTP (elliptical upper and lower local ternary pattern operator (EUL_{LTP})). It takes into account the information of the three color planes (RGB, HSV). The new EUL_{LTP} operator reduces the size of the original LTP to obtain a vector of a size of 64 bins instead of 512 bins and keeps its discriminative power and computational efficiency.

To assess the proposed method, we performed experiments for image classification on six popular and challenging datasets: New-BarkTex, Outex-TC13, MIT scene, UIUC sports event, Caltech 101 and MIT indoor scene datasets. The experimental results prove that the proposed descriptor outperforms the existing state-of-the-art low-level descriptors and give close results to others using deep learning considered as high-level techniques.

References

1. Yu, J., Qin, Z., Wan, T., Zhang, X.: Feature integration analysis of bag of features model for image retrieval. Neurocomputing **120**, 355–364 (2013)
2. Banerji, S., Verma, A., Liu, C.: Cross disciplinary biometric systems. LBP and Color Descriptors for Image Classification, pp. 205–225. Springer, Berlin (2012)
3. Ledoux, A., Losson, O., Macaire, L.: Color local binary patterns: compact descriptors for texture classification. J Electron Imaging **25**(6), 061404 (2016)
4. Yongsheng, D., et al.: Multi-scale counting and difference representation for texture classification. Vis. Comput. (2017). https://doi.org/10.1007/s00371-017-1415-4
5. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision **60**(2), 91–110 (2004)
6. Xuemei, H., Yan, D.: Image matching with an improved descriptor based on SIFT. In: Proceedings Volume 10322, Seventh International Conference on Electronics and Information Engineering, pp. 1-7 (2017). https://doi.org/10.1117/12.2265595
7. Romero, A., Gatta, C., Camps-Valls, G.: Unsupervised deep feature extraction for remote sensing image classification. IEEE Trans. Geosci. Remote Sens. **54**(3), 1349–1362 (2016)
8. Berbar, M.: Three robust features extraction approaches for facial gender classification. Vis. Comput. **30**(1), 19–31 (2014)
9. Bay, H., et al.: Speeded-up robust features (SURF). Comput. Vis. Image Underst. **110**(3), 346–359 (2008)
10. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. Pattern Recognit. **29**(1), 51–59 (1996)
11. Li, Y., Ruixi, Z., Nan, M., Yi, L.: Improved class-specific codebook with two-step classification for scene-level classification of high resolution remote sensing images. Remote Sens. **9**(3), 1–24 (2017)
12. Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: Proceedings of the International Workshop on Statistical Learning in Computer Vision, pp. 1–16 (2004)
13. Xiaoyong, B., Chen, C., Long, T., Qian, D.: Fusing local and global features for high-resolution scene classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **10**(6), 2889–2901 (2017)
14. Daoyu, L., et al.: MARTA GANS: unsupervised representation learning for remote sensing image classification. IEEE Geosci. Remote Sens. Lett. **14**(11), 1–5 (2017)

15. Alex, K., Ilya, S., Geoffrey, E.H.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
16. Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A.: Return of the devil in the details: delving deep into convolutional nets. CoRR, arXiv:1405.3531 (2014)
17. Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., Oliva, A.: Learning deep features for scene recognition using places database. In: Advances in Neural Information Processing Systems, pp. 487–495 (2014)
18. Lee, H., Grosse, R., Ranganath, R., Ng, A.: Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In: Proceedings Annual International Conference on Machine Learning, pp. 609–616 (2009)
19. Shuang, B., Zhaohong, L., Jianjun, H.: Learning two-pathway convolutional neural networks for categorizing scene images. Multimed. Tools Appl. **76**(15), 16145–16162 (2017)
20. Sandid, F., Douik, A.: Robust color texture descriptor for material recognition. Pattern Recognit. Lett. **80**, 15–23 (2016)
21. Guo, Z., Zhang, L., Zhang, D.: A completed modeling of local binary pattern operator for texture classification. IEEE Trans. Image Process. **19**(6), 1657–1663 (2010)
22. Tan, X., Triggs, B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. IEEE Trans. Image Process. **19**(6), 1635–1650 (2010)
23. Nishant, S., Vipin, T.: An effective scheme for image texture classification based on binary local structure pattern. Vis. Comput. **30**(11), 1223–1232 (2014)
24. Xiaosheng, W., Junding, S.: Joint-scale LBP: a new feature descriptor for texture classification. Vis. Comput. **33**(3), 317–329 (2017)
25. Rahman, M.M., et al.: DTCTH: a discriminative local pattern descriptor for image classification. EURASIP J. Image. Video Process. **2017**, 1–24 (2017)
26. Khan, R., Muselet, D., Trémeau, A.: Texture classification across illumination color variations. Int. J. Comput. Theory Eng. **5**, 65–70 (2013)
27. Alvarez, S., Vanrell, M.: Texton theory revisited: a bag-of-words approach to combine textons. Pattern Recognit. **45**(12), 4312–4325 (2012)
28. Zhu, C., et al.: Image region description using orthogonal combination of local binary patterns enhanced with color information. Pattern Recognit. **46**(7), 1949–1963 (2013)
29. Porebski, A., et al.: A new benchmark image test suite for evaluating colour texture classification schemes. Multimed. Tools Appl. **70**(1), 543–556 (2014)
30. Cusano, C., Napoletano, P., Schettini, R.: Combining local binary patterns and local color contrast for texture classification under varying illumination. J. Opt. Soc. Am. A **31**(7), 1453–1461 (2014)
31. Sandid, F., Douik, A.: Texture descriptor based on local combination adaptive ternary pattern. IET Image Process. **9**(8), 634–642 (2015)
32. Kabbai, L., Abdellaoui, M., Douik, A.: Content based image retrieval using local and global features descriptor. In: IEEE International Conference on Advanced Technologies for Signal and Image Processing, pp. 151–154 (2016)
33. Kabbai, L., Abdellaoui, M., Douik, A.: Hybrid local and global descriptor enhanced with colour information. IET Image Process. **11**(2), 109–117 (2016)
34. Papadopoulos, G.Th, Mezaris, V., Kompatsiaris, I., Strintzis, M.G.: Combining global and local information for knowledge-assisted image analysis and classification. EURASIP J. Image Video Process. **2007**, 1–15 (2007)
35. Banerji, S., Sinha, A., Chengjun, L.: New image descriptors based on color, texture, shape, and wavelets for object and scene image classification. Neurocomputing **117**, 173–185 (2013)
36. Sinha, A., Banerji, S., Liu, C.: New color GPHOG descriptors for object and scene image classification. Mach. Vis. Appl. **25**(2), 361–375 (2014)
37. Khan, R., Barat, C., Muselet, D., Ducottet, C.: Spatial histograms of soft pairwise similar patches to improve the bag-of-visual words model. Comput. Vis. Image Underst. **132**, 102–112 (2015)
38. Amit, S., Xudong, J., How, L.E.: LBP-based edge-texture features for object recognition. IEEE Trans. Image Process. **23**(5), 1953–1964 (2014)
39. Maji, S., Berg, A.C., Malik, J.: Efficient classification for additive kernel SVMs. IEEE Trans. Pattern Anal. Mach. Intell. **35**(1), 66–77 (2013)
40. Li, L.J., et al.: Object bank: A high-level image representation for scene classification & semantic feature sparsification. In: Advances in Neural Information Processing Systems, pp. 1378–1386 (2010)
41. Li, L.J., Li, F.F.: What, where and who? Classifying events by scene and object recognition. In: IEEE International Conference on Computer Vision, pp. 1–8 (2007)
42. Bo, L., Ren, X., Fox, D.: Hierarchical matching pursuit for image classification: architecture and fast algorithms. In: Advances in Neural Information Processing Systems, pp. 2115–2123 (2011)
43. Harris, C., Stephens, M.J.: A combined corner and edge detector. In: Proceedings of Fourth Alvey Vision Conference, pp. 147–151 (1988)
44. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Trans. Pattern Anal. Mach. Intell. **27**(10), 1615–1630 (2005)
45. Ke, Y., Sukthankar, R.: PCA-SIFT: a more distinctive representation for local image descriptors. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 506–513 (2004)
46. Kabbai, L., Abdellaoui, M., Douik, A.: New robust descriptor for image matching. J. Theor. Appl. Inf. Technol. **87**(3), 451–460 (2016)
47. Ai, D.N., et al.: Color independent components based SIFT descriptors for object/scene classification. IEICE Trans. Inf. Syst. **93**(9), 2577–2586 (2010)
48. Muralidharan, R., Chandrasekar, C.: Combining local and global feature for object recognition using SVM-KNN. In: IEEE International Conference on Pattern Recognition, Informatics and Medical Engineering, pp. 1–7 (2012)
49. Chaudhary, M.D., Upadhyay, A.B.: Fusion of local and global features using stationary wavelet transform for efficient content based image retrieval. In: IEEE Students' Conference on Electrical, Electronics and Computer Science, pp. 1–6 (2014)
50. Li, L., et al.: Fusion framework for color image retrieval based on bag-of-words model and color local haar binary patterns. J. Electron. Imaging **25**(2), 023022 (2016)
51. Zou, J., et al.: Scene classification using local and global features with collaborative representation fusion. Inf. Sci. **348**, 209–226 (2016)
52. Mallat, S.G.: theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans. Pattern Anal. Mach. Intell. **11**(7), 674–693 (1989)
53. Smith, J. R., Chang, S. F.: Automated binary texture feature sets for image retrieval. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 2239–2242 (1996)
54. Ojala, T., et al.: Outex new framework for empirical evaluation of texture analysis algorithms. In: Proceedings of IEEE International Conference on Pattern Recognition, pp. 701–706 (2002)
55. Oliva, A., Torralba, A.: Modeling the shape of the scene: a holistic representation of the spatial envelope. Int. J. Comput. Vision **42**(3), 145–175 (2001)
56. De Brabanter, K., et al.: LS-SVMlab toolbox user's guide version 1.8. Internal Report, ESAT-SISTA, K.U. Leuven, Leuven, Belgium, pp. 10–14 (2010)

Leila Kabbai was born in Monastir. In 2016, she received the doctorate in Electrical Engineering from the National Engineering School of Monastir, University of Monastir, Tunisia. She is a researcher at the Networked Objects, Control and Communication Systems (NOCCS) Research Laboratory, National School of Engineering of Sousse, Tunisia. Her main areas of interest are image retrieval, image classification, image matching and hardware implementation. She has published several research papers in international conferences and journals.

Mehrez Abdellaoui obtained his engineer's, M.S. and Ph.D. degrees from National School of Engineering of Monastir, Tunisia, in 2003, 2005 and 2012, respectively. He is currently Associate Professor in signal and image processing at the High Institute of Applied Technologies, University of Kairouan. He is also a researcher at the Networked Objects, Control and Communication Systems (NOCCS) research Laboratory, National School of Engineering of Sousse, Tunisia. His research interests include Image and video processing, computer vision, machine learning.

Ali Douik was born in Tunis, Tunisia. He received his B.S., M.S. and Ph.D. degrees in electrical engineering from the ENSET and ESSTT of Tunis, TUNISIA, in 1988, 1990 and 1996, respectively, and HDR in electrical engineering from the University of Monastir, Monastir, Tunisia, in 2010. He was in National Engineering School of Monastir from September 1991 to September 2014 and actually he is Full Professor in Industrial Computing Department of National Engineering School of Sousse. His research interests include digital image processing, automatic control, optimization and evolutionary algorithms.