



Weber local descriptor for image analysis and recognition: a survey

Arnab Banerjee¹ · Nibaran Das¹ · K. C. Santosh²

Accepted: 4 November 2020 / Published online: 25 November 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

Weber local descriptor (WLD) is applied for addressing the challenges in image/pattern problems, especially in computer vision and pattern recognition domains. In this paper, we review literature on theories and applications of WLD. Using WLD, we address the different challenges of image analysis and recognition features with respect to illumination changes, contrast differences, and geometrical transformations like rotation, scaling, translation, and mirroring. Further, the role of the classifiers and experimental protocols used in the different applications are discussed. Applications include texture classification, medical imaging, agricultural safety, fingerprint analysis, forgery analysis, and face recognition.

Keywords Weber local descriptor · Image analysis and recognition · Computer vision and pattern recognition

1 Introduction

Extensive researches were done on image recognition, where texture analysis-based local and global texture descriptors were employed. Texture provides knowledge about spatial distribution of color and intensities either from the whole image or from a region-of-interest that can be based on the point-based descriptors. Texture analysis also helps in segmenting images into meaningful regions and in classifying them accordingly. Structural approaches reveal that texture is a collection of primitive texels with patterns whereas statistical approaches extract the computable measure of the intensity values in the image. Several texture descriptors were proposed and applied on several different applications viz. computer vision, pattern recognition, material classification, face recognition, emotion analysis, and medical image analysis. Multiple texture models viz. statistical model (co-occurrence matrix, auto-correlation features etc.), geo-

metric models (structural model, Voronoi tessellation etc.), model-based random fields (fractal approach, random field model etc.), and signal processing-based models (Gabor and wavelet method, spatial domain filters, etc.) exist in the literature. Due to robustness and improved execution time, local descriptors are widely used in texture recognition. Further, sparse and dense descriptors are commonly used to classify texture. Primarily, their difference lies in the way they describe the image. Sparse descriptors detect key points to sample a local region and dense descriptors extract local features from all pixels. The feature transformation technique with scale invariance (SIFT) [49] and rotational invariance [44], Histogram of Oriented Gradients (HOG) are the key sparse descriptors. If we see the dense descriptor, Gabor Wavelet [84], LBP [57], LDP [36] played an important role in multiple application domains. In 2010, [13] proposed WLD as an illumination and noise sensitive robust local descriptor. Since then multiple modified versions of WLD came into play. Some of the popular modified WLD versions are Memetically optimized multiscale circular WLD [8], Weber LBP [48], Gabor WLD [69], Log-Gabor WLD [45], deep neural network-based WLD [5], quaternionic Weber Local Descriptor [43], etc. For better understanding, the road map of the development of WLD is presented in Fig. 1.

In the image recognition framework, the WLD has been applied effectively in several areas, like classifying textures, detecting faces, medical image classification, and agricultural domain. Therefore, recognizing the importance of the image recognition via the application of robust local fea-

✉ K. C. Santosh
santosh.kc@ieee.org

Arnab Banerjee
arnab.jdvu@yahoo.in

Nibaran Das
nibaran@gmail.com

¹ Jadavpur University, Kolkata, West Bengal 700032, India

² University of South Dakota, Vermillion, SD 57069, USA



Fig. 1 Roadmap of the development of WLD

ture descriptor: WLD, in this survey, we explain the working principle of WLD, addresses its different applications (plus detailed algorithms) and the result that follows standard datasets.

The remaining paper is arranged in the following manner: the introduction of problem and motivation of this survey is presented in Sect. 2. The principles of WLD, methods to extract the features, and characteristics of WLD with examples are explained in Sect. 3. In Sect. 4, the applications of WLD with the methods and results are presented. The conclusion of the work is given in Sect. 6.

2 Image classification

In image recognition, pixels are associated with the different classes based on the selected features and the used learning technique. For learning/recognition, the system has to be trained with sufficient data, and such techniques are under the scope of supervised and unsupervised learning/classification. Mathematically classification is a mapping from pixels to a class or label $f(I) : I \rightarrow c$, where $C = c_1, c_2, \dots, c_n$, and the function $f(\cdot)$ maps an image to a label c_i where $c_i \in C$ and $1 \leq i \leq n$.

2.1 Problem overview

Figure 2 shows the overview of an image recognition/classification system having three components: preprocessing, feature extraction and classification. In preprocessing, among various techniques, noise removal, gray scaling is very common since not all feature extractors are robust to noise and scaling. Some of the descriptors are having inherent preprocessing capabilities. In feature extraction, distinguished features are extracted. In classification, all the above phases are applied to the unknown image instance and the classifier will give the label of the unknown image, taking the help of the trained image features. The classifier plays a crucial role in image recognition. The performance of the classifiers depends largely on the configuration and, of course, the parameters applied. Therefore, proper selection of features and classifier(s) is required since they vary from one application to another.

2.2 Motivation

After the introduction of WLD by [13], several modifications were made to solve various problems under the scope of pattern recognition. Due to rotation, translation, scaling, and other issues, such as illumination changes, occlusions, and degradation, image recognition is not trivial. Modified WLDs were successfully applied to solve aforementioned inherent problems in image recognition. Though the performances of WLDs are compelling than the other local descriptor, such as LBP, LTP, and SIFT, no survey on WLD has not yet been reported in the literature. The encouraging performances of traditional WLD and its variants in different fields of image/pattern recognition motivate us to write this extensive survey that includes basic theories behind WLD in addition to its modifications. It also covers application domains, where the use of different classifiers (with performance on standard datasets) are addressed. In a nutshell, it is composed of WLD's definition, its modified versions (application-motivated), and its application domains. In 3, the theory, principal, and characteristics of contemporary WLD will be explained.

3 Weber local descriptor: definition, principle and characteristics

3.1 Overview

In the nineteenth century, Ernest Weber explored a constant relationship between incremental threshold and background

intensity. Therefore, one must speak loudly in a crowded environment to be heard while whisper works in peaceful environment. This relationship can be described using the Weber's law, i.e., $\delta I/I = K$, where $\delta I/I$ is called Weber fraction. Since we can simplify it by $\delta I = KI$, a linear relation between δI and I is established by the Weber's law.

3.2 WLD

Using the two principle concepts WLD is formed (a) differential excitation (ψ) and (b) orientation (Θ). The Eqs. 1 to 6 are taken from [13].

(a) *Differential Excitation*: The micro-variation within an image can be computed by taking into account the intensity difference between the neighboring pixels. It can be expressed as:

$$\delta I = \sum_{x=0}^{i-1} \delta I(p_x) = \sum_{x=0}^{i-1} I(p_x) - I(p_c), \tag{1}$$

where x th neighbors of p_c is represented by $p_x(x = 0, 1, \dots, i - 1)$ and i represents total number of neighbors in a region. $I(p_x)$ presents the intensity of the neighbored pixels, and $I(p_c)$ presents the intensity of current pixel. It can be expressed as,

$$\psi(p_c) = \arctan\left(\frac{\delta I}{I}\right) = \arctan\left(\frac{\sum_{x=0}^{i-1} (I(p_x) - I(p_c))}{I(p_c)}\right). \tag{2}$$

If $\psi(p_c)$ is positive, then center pixel is darker respect to the neighbor pixels and if $\psi(p_c)$ is negative, then current pixel is lighter respect to the neighbor pixels.

(b) *Orientation*: It determines the directional property of the pixels. It can be computed as:

$$\Theta(p_c) = \arctan\left(\frac{dI_h}{dI_v}\right), \tag{3}$$

where $dI_h = I(p_7) - I(p_3)$ and $dI_v = I(p_5) - I(p_1)$ is calculated from the two filters as in Fig. 3. The mapping of $f : \Theta \rightarrow \Theta'$ can be expressed as, $\Theta' = \arctan2(dI_h, dI_v) + \pi$, and

$$f(a) = \begin{cases} \Theta, & dI_h > 0 \text{ and } dI_v > 0 \\ \pi - \Theta, & dI_h > 0 \text{ and } dI_v < 0 \\ \Theta - \pi, & dI_h < 0 \text{ and } dI_v < 0 \\ -\Theta, & dI_h < 0 \text{ and } dI_v > 0, \end{cases} \tag{4}$$

Fig. 2 Image recognition overview

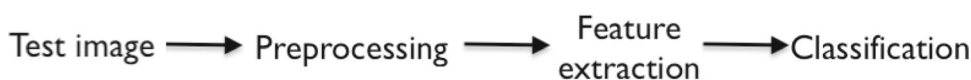
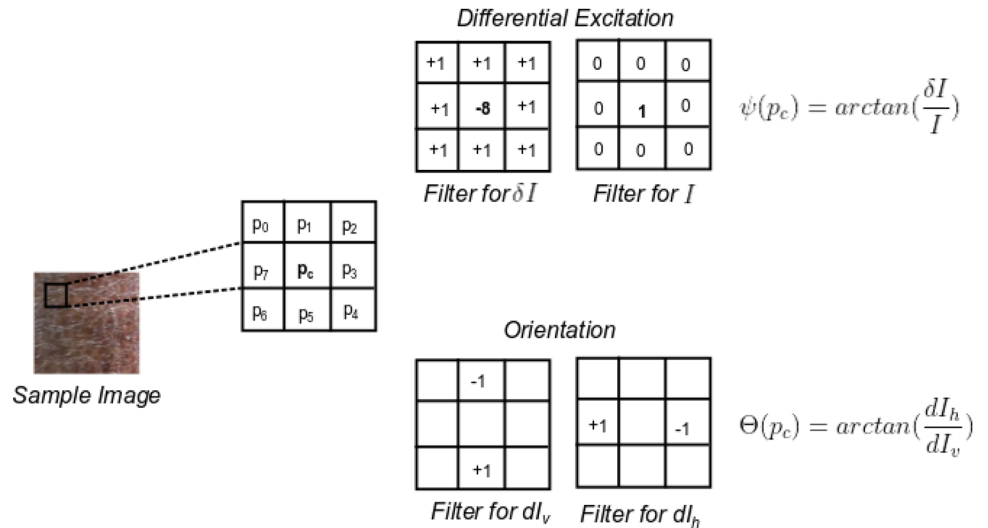


Fig. 3 Calculation of WLD [13] on a sample skin image



where \$\Theta\$ varies from \$-90^\circ\$ to \$90^\circ\$. Further the \$\Theta\$ is quantized into 0 to \$2\pi\$.

3.2.1 Formation of WLD histogram

Using the above two principle components of WLD, we can form the WLD histogram. The WLD histogram can be taken as the image descriptor for several image recognition tasks. Figure 3 represents the calculation of differential excitation (\$\psi_m\$) and orientation (\$\Phi_n\$) of an input image. Next the, two-dimensional histogram is created \$WLD(\psi_m, \Phi_n)\$, where (\$m\$ varies from 0 to \$M - 1\$) and (\$n\$ varies from 0 to \$N - 1\$). Here \$M\$ denotes image dimension and \$N\$ denotes the total number of dominant orientations.

In the 2D histogram, dominant orientation is represented by each column and differential excitation is represented by each row. Next \$N\$ numbers of 1D histograms are formed by decomposing the 2D histogram. Each dominant orientation is referred by each 1D histogram. In each 1D histogram, total \$s\$ segments \$H_s\$, \$s\$ varies from 0 to \$s - 1\$ are present. A subset is formed by combining each row of 1D histogram. The final WLD histogram is formed by concatenating all sub-histograms \$H = H_s, s = 0, 1, \dots, S - 1\$. There are \$B\$ bins in each sub-histogram \$H_{(s,n)}\$ i.e. \$H_{(s,n)} = H_{(s,n,b)}, b = 0, 1, \dots, B - 1\$:

$$H_{s,n,b} = \sum_k \Psi(B_k = B), \left(B_k = \left\lfloor \frac{\psi_k - \gamma_{s,l}}{\frac{\gamma_{s,u} - \gamma_{m,l}}{B}} + \frac{1}{2} \right\rfloor \right), \tag{5}$$

where \$s\$ depends on the interval \$\psi_k\$ belongs and \$n\$ refers to the quantized orientation index and \$\Psi(\cdot)\$ can be expressed

as:

$$\Psi(a) = \begin{cases} 1, & a \text{ is true} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

3.2.2 Characteristics of the WLD

Inspired by the psychological law, the WLD can be computed. It computes the features similar, as a human perceives his/her surroundings. So, such descriptor has some inherent capabilities of human sensing and application of such descriptor has been very successful in several domains of image recognition. This section will address the advantages of using WLD. We basically cover where it is robust to noise, illumination, and rotation.

(a) *Robust to Noise* The derivation of differential excitation and orientation makes the WLD noise redundant. WLD inherently reduces the noises present in the image like the smoothing operation in image processing. While calculating differential excitation \$\psi(p_c)\$, first the intensity difference of \$i - 1\$ neighbors and center pixel is added up and then divided by the current pixel intensity. Thus, it makes WLD robust against multiplicative noise. Thus, the WLD is robust against the multiplicative noise [13]. An example is presented in Fig. 4.

In a \$3 \times 3\$ neighborhood, consider any pixel \$p_x\$ where \$0 \le x \le 7\$ and \$p_c\$ represents the current pixel. The multiplicative noise \$\eta\$ is applied to every pixel in the neighborhood. Differential excitation \$\psi_\eta(p_c)\$ and orientation \$\Theta_\eta(p_c)\$ can be derived, respectively, as

Differential excitation:

$$\psi_\eta(p_c) = \arctan \left(\sum_{x=0}^{i-1} \frac{\eta((I(p_x) - I(p_c)))}{\eta(I(p_x))} \right)$$

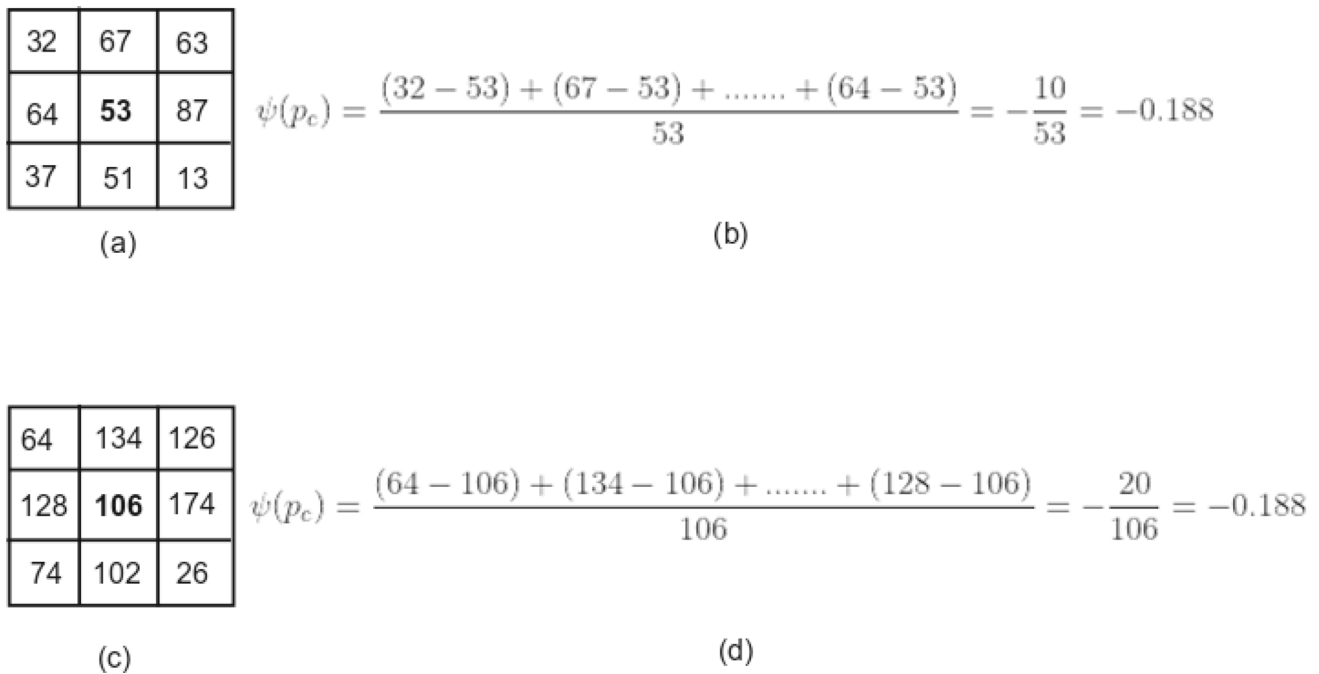


Fig. 4 **a** The gray values of 3 × 3 Neighborhood of an image; **b** the calculation of differential excitation for upper left hand side neighborhood; **c** the noisy gray values of same 3 × 3 neighborhood; and **d** calculation of differential excitation for lower left hand neighborhood

$$= \arctan \left(\sum_{x=0}^{i-1} \frac{((I(p_x) - I(p_c)))}{(I(p_x))} \right)$$

and $\psi_\eta(p_c) = \psi(p_c)$. (7)

Orientation:

$$\Theta_\eta(p_c) = \frac{\eta((I(p_7) - I(p_3)))}{\eta((I(p_5) - I(p_1)))}$$

$$= \frac{((I(p_7) - I(p_3)))}{((I(p_5) - I(p_1)))}$$

and $\Theta_\eta(p_c) = \Theta(p_c)$. (8)

These equations prove that the WLD is robust to multiplicative noise. When WLD is compared with SIFT [49] and LBP [57] for additive noise, it is proved that WLD performs better than these two-texture descriptors. These three-texture descriptors performed same for the added white Gaussian noise with noise strength 5, but WLD outperformed than the SIFT and LBP for the noise strength greater than 5 [13].

(b) Robust to Illumination In WLD, δI is the differences of intensities of neighbored pixels with center pixel intensity. Hence, brightness changes have no effect on the difference values. Therefore, the WLD descriptor is robust to the illumination [13] that can avoid potential use of preprocessing.

(c) Robust to Rotation of Image The differential excitation component of WLD descriptor is having rotational invariance property. Even if a 3 × 3 neighborhood is rotated about some degree, it produced the same difference value of the

intensity with the neighbor pixel and the center pixel. So, rotation of image does not affect the Differential Excitation. [61] proposed WLD with rotational invariance (WLDRI) to make orientation component robust to rotation changes. Here the orientation is computed for all the mutually perpendicular diagonal pairs and the minimum value for orientation is taken. Equation [61] for calculating the orientation component:

$$\Theta_k = \arctan \left(\frac{I(p_{((\frac{i}{2})-k) \bmod i)}) - I(p_i)}{I(p_{((\frac{3i}{4})-k) \bmod i)}) - I(p_{((\frac{i}{4})-k) \bmod i})} \right),$$

and $\theta(p_c) = \min_{x=0}^{i-1} (\theta_x)$. (9)

WLDRI outperformed the conventional WLD in case of skin diseases recognition. Deep neural network (DNN)-based WLDRI [5] also significantly outperformed conventional WLD in case of KTH-TIPS2-a [12] texture dataset, OUTEX-10 [56] for texture recognition.

4 WLD: applications, a quick review

WLD has been successfully applied in verification of faces, facial expressions and facial emotion recognition, fingerprint detection, forgery detection, biometric spoofing detection and texture classification. Face recognition and texture classification are found to be dominant applications. WLD were used in Kinship classification based on self-similarity

representation [41] and used as a problem solver in face recognition in case of aging problem [6]. There were three applications of WLD in medical imaging, one is automatic mass detection in Mammograms [31], other is skin diseases identification using WLDRI [61] and DNN-Based WLDRI for Skin Diseases Detection [5]. There exist several versions of WLD developed by the combination of state-of-the-art descriptor with WLD. These are Memetically optimized multiscale circular WLD [8], Weber LBP [48], Gabor WLD [69], Log-Gabor WLD [45], deep neural network-based WLD [5], quaternionic Weber Local Descriptor [43], etc. These versions of WLD were extensively used in several applications with higher recognition accuracy. The higher-order statistics of WLD was studied to get the highly discriminant image features and applied in texture recognition, and recognition of different food images and pattern of Hep-2 cell [26]. Recently in 2018 [71] applied WLD-based biometric model for fish species classification. A list of works that use WLD are enumerated as follows:

1. Texture classification [1,5,13,15,16]
2. Medical diagnosis [5,31,61]
3. Agriculture safety [19]
4. Fingerprint analysis and detection [23]
5. Forgery detection [33–35,66,74]
6. Face detection and recognition [13,19,22,32,41,46,54,76,79,80,83]
7. Variations of WLD [8,26,45,48,68,69]

4.1 Texture classification

4.1.1 Robust local image descriptor

Authors presented WLD, a local descriptor, and applied in texture classification [13]. WLD processes the texture micropatterns locally. WLD was applied on Brodatz [11] and KTH-TIPS2-a [12] texture dataset. The method of feature extraction is described in the Sect. 3. Brodatz dataset consists 32 texture classes, each with 64 samples in it. Each image is of 256×256 pixels and 256 gray levels. Every image was divided into 64×64 pixels and histogram-equalized for luminance invariance. Additional samples are created using rotation and scaling. The material dataset KTH-TIPS2-a has 4395 images and 11 classes of textures. In this dataset, images are present with different variations like 9 different scales, 4 illumination direction, and 3 different poses. Here size of

each image is 200 by 200 pixels. To avoid over fitting, they tested using cross-validation (tenfold) on the Brodatz dataset. The setup of KTH-TIPS2-a dataset was same as of [12]. For training, 3 samples were applied. The tests were made for four times with four different sets of three images for training. The parameter values of T, M, S were 8, 6, 20 and the Table 1 gives the weight of each sub-histogram. In this experiment, for the classification purpose, K nearest neighbor (KNN) was used. Besides, comparison of multiscale WLD (MWLD) with MLBP was done. WLD or MWLD outperformed the other recently reported techniques for both the datasets. The accuracy of WLD for Brodatz dataset was 97.5 (standard deviation = 0.6) and accuracy of MWLD for KTH-TIPS2-a dataset was 64.7 (better than the normal WLD, accuracy = 58.1) [13]. The accuracy of SIFT for Brodatz dataset and KTH-TIPS2-a was 91.4%, 52.7% and accuracy of LBP for Brodatz dataset and KTH-TIPS2-a was 91.2% and 49.9%. Clearly, WLD outperformed the the SIFT- and LBP-based techniques.

4.1.2 Improved WLD

It considers the orientation component. It blurs the texture of the image and thus responsible for misclassification [16]. Authors proposed an improved WLD and applied in texture classification. The following Eq. [16] express the way of computing gradient at $f(m, n)$,

$$Df = \begin{bmatrix} \text{Gradient}_x \\ \text{Gradient}_y \end{bmatrix} = \begin{bmatrix} \frac{df}{dm} \\ \frac{df}{dn} \end{bmatrix}. \tag{10}$$

The magnitude of the vector is $Df = \text{mag}(Df) = \left[\begin{bmatrix} \text{Gradient}_x^2 \\ \text{Gradient}_y^2 \end{bmatrix} \right]^{\frac{1}{2}}$, and gradient direction is $\beta(x, y) = \arctan \left(\frac{\text{Gradient}_y}{\text{Gradient}_x} \right)$ For each cell of the image, the histogram was computed. The overlapping cells made it powerful for identifying the edges. The image was formed from the histograms of each cell. The differential excitation was considered to keep the neighboring information. This improved version of WLD was applied on Brodatz [11] and KTH-TIPS2-a [12] dataset. The details of the datasets are same as Sect. 4.1.1. Authors compared the performance of improved WLD with LBP, LPQ and WLD. Support vector machine (SVM) was applied for classification purpose. In case of LPQ, the correlation coefficient $\rho = 0.9$ was used. This method achieved 8,

Table 1 Weights for the WLD sub histogram [H_i denotes i th sub histogram.] [13]

	H_0	H_1	H_2	H_3	H_4	H_5
Frequency (%)	0.2519	0.1179	0.1186	0.0965	0.0875	0.3276
Weights (ω_m)	0.2688	0.0852	0.0955	0.1000	0.1018	0.3487

6, and 4% better accuracies than the LBP [57], LTP [58] and WLD [13]. If the textures are blurred, then micro variations in image cannot be achieved by only using the local contrast information. The method computes the features at the cell, and therefore, it helps reduce the blurring of texture.

4.1.3 Integrating the WLD with variance

Authors [15] proposed a technique to combine WLD and variance histogram (WLDV) for better performance than the LBP, LPQ, and WLD. PWM [18] was used for the variance component. The input image was divided into 3×3 window and the variance of each window was calculated. The contrast image was normalized in (0–255) level range. In this approach, WLD and variance histogram was calculated and combined. These two histograms are complementary to one another and exploit the contrast and local spatial pattern. The value of the parameter T , M , S was 8, 6, 20. This combined histogram was applied on Brodatz [11] and KTH-TIPS2-a [12] texture dataset. The details of the two texture datasets are same as Sect. 4.1.1. They used tenfold cross-validation to conduct the experiment. The training and testing data were created in 2:3 ratio. SVM was used for classification and WLDV gained better accuracy than LBP, LPQ and WLD for both datasets. Though accuracy percentage is not reported in their work. The accuracy for the KTH-TIPS2-a dataset is lower than the Brodatz dataset because of diverse nature of the KTH-TIPS2-a dataset (illumination direction, different pose, and rotation).

4.1.4 Texture classification using DNN-based WLDRI

Authors proposed a deep learning-based rotation invariant WLD. OUTEX-10 and KTH-TIPS2-a dataset was used in this experiment [5]. Authors [61] presented WLDRI to make orientation component rotation invariant. In this approach the kernel (or filter) of WLDRI was applied in deep neural network to simulate the behavior of WLDRI into DNN for enhancing the performance of DNN. All the images were grayscale and .bmp file type. In this approach, three-layered DNN was used and 3000 neurons were used in each layer. During DNN training, when the error rate was not decreased after 30 epochs, at max 1000 epochs were used. In OUTEX-10 [56] dataset, total 24 classes of textures are present with 4320 images. At nine different angles (0° , 5° , 10° , 15° , 30° , 45° , 60° , 75° , 90°) the images were rotated. There were no illumination changes for rotating the images. In OUTEX-10 dataset, the accuracies using multiscale WLDRI and DNN based WLDRI were 81.27% and 89.33% accordingly. When KTH-TIPS2-a dataset was used in the experiment, the multiscale WLD achieved 64.7% accuracy and the DNN-based WLDRI achieved 69.73% accuracy. Multiscale WLD was used in KTH-TIPS2-a dataset for easier

comparison with [13]. It is also easy to compare the results with other descriptors such as SIFT and LBP by referring Sect. 4.1.1. DNN-based WLDRI achieved 8.06% better performance in case of OUTEX-10 dataset and 5.03% better performance for the KTH-TIPS2-a dataset. Further the proposed DNN-based WLDRI descriptor proved its statistical significance.

4.2 Medical diagnosis

4.2.1 Rotation invariant WLD for skin disease recognition

Authors proposed a rotation invariant WLD to recognize skin diseases: Leprosy, Tinea Versicolor, and Vitiligo [61]. They proposed that the orientation component of WLD [13] is easily disrupted by rotation. The proposed technique help make rotation invariant, since the orientation is computed for all the mutually perpendicular diagonal pairs and the minimum value for orientation is taken. The formula is shown in Eq. (9). To build WLD histogram, 8 and 6 bins were used for differential excitation and orientation. In the experiment, 4 regions of the image were taken using the center of gravity (CG) of the image. The experiment was done on the skin diseases dataset [61] with the images of Leprosy, Tinea Versicolor, and Vitiligo affected skin. The normal skin images were also taken for classification purpose. Altogether, from 876 randomly collected images 702 images were taken for training and 174 images were used for testing purpose. Two experiments were done using WLD and WLDRI—center-of-gravity-based partition method and without center-based partition-based partition method. Using SVM classifier, WLDRI achieved 4.79% better accuracy than the WLD (without using center-of-gravity-based partition). When center-of-gravity-based partition was used, multiscale WLDRI achieved an accuracy of 89.08%, whereas multiscale center-of-gravity-based WLD achieved an accuracy of 86.78%.

4.2.2 Multiscale spatial WLD for mass detection in mammograms

Authors proposed multiscale spatial WLD to detect mass in the mammograms [31]. The method was proposed to encode the local region patterns and spatial structure in the masses. The application of this method includes the detecting masses, including suspicious parenchymal regions. Traditional WLD does not incorporate the spatial location factor. In spatial WLD, the image was partitioned into different regions and multiscale WLD was applied in each region. The histograms associated with each region was concatenated and final histogram was built. The method was applied on Digital Database for Screening Mammography (DDSM) [27] dataset which consists of images labelled as either mass or normal by

the experienced radiologists. 256 ROIs were extracted with the size ranging from 240×240 to 1180×1180 depending on the size of the masses. 256 ROIs of normal but suspicious region images were also present in the dataset. For the classification purpose, SVM (with RBF kernel) was applied and performance was tested by the Area under curve (AUC). The appropriate value of the parameter T , M , S was (4, 4, 5) and 4×4 optimal number of blocks were used per image. The AUC of MSWLD was 0.988 ± 0.006 , whereas the AUC of LBP and basic WLD was 0.922 ± 0.016 and 0.697 ± 0.032 . This method outperformed LBP and basic WLD with this DDSM dataset.

4.2.3 DNN-based WLDRI for skin disease recognition

Authors proposed a deep neural network-based system based on WLDRI [5]. The description of the method and experimental setup are described in Sect. 4.1.4. CMATER skin dataset [61] was used for the experiment. This dataset consists of the texture images of skin diseases—Leprosy, Vitiligo and Tinea Versicolor and also normal skin images. In the experiment, 702 images were taken as the training data and 174 images were taken for the testing. Center of gravity (CG) of the image was used for the partition of the images into four regions. The experiments were done using partition based on the CG and without using the partition, and results were compared between them. The results were compared with the method by [61]. The proposed method without CG based partition achieved 2.88% better accuracy than the WLDRI using the concatenated features of multiscale version, whereas in case of CG-based partition, the proposed method achieved 5.74% improved accuracy over WLDRI. The precision, recall, and F -score of DNN-based WLDRI was also greater than WLDRI. The FPR of DNN-based WLDRI was less than WLDRI. The observation is that the use of combined multiscale features from the CG-based partitioned images is more discriminative for making the differences between the skin diseases affected images and normal skin images.

4.3 Agricultural safety

4.3.1 WLD for biometric cattle identification

Authors proposed a system based on WLD in biometric cattle identification from the cattle muzzle print images [19]. Local discriminant analysis (LDA) [64] was used for the reduction of the feature dimension to discriminate between different classes. From each of 31 heads of cattle, 7 images were taken and this approach was evaluated on total 217 muzzle print images. Training images were increased from one to six for each head of cattle and rest were used for testing. AdaBoost classifier was used for the recognition of unknown muz-

zle image. There were different parameters involved in the experiment-WLD parameters (patch size), AdaBoost parameters (type of weak learner, learning rate and number of weak learners). The patch size of WLD was determined by various experiments done with different patch sizes. Using tree and discriminant learner number of experiments were done to get the appropriate type of weak learner. In the experiment, number of weak learner was 200 and learning rate was 0.1. The discriminant learner had the less error rate than he Tree learner. KNN and Fuzzy k -NN classifiers were used for comparing performance with the AdaBoost classifier. Several experiments were done for different odd values of K . This method used AdaBoost classifier and achieved 98.9% accuracy using 4 training images, and KNN and Fuzzy k -NN gained, respectively, accuracies of 96.8% and 97.9%. The system is very robust with rotation and occlusion even though no preprocessing was applied. The statistical measures [sensitivity, specificity, AUC, equal error rate (ERR)] of this approach using AdaBoost is relatively better than the KNN and Fuzzy k -NN.

4.4 Fingerprint analysis and detection

4.4.1 Local features via WLD for fingerprint detection

Authors studied the fingerprint liveness recognition using local discriminative feature space, which is considered as a very challenging problem to track the fake fingerprints in several sources, such as silicon, gelatin or clay [23]. The spatial WLD histogram and the combination of WLD and local phase quantization (LPQ) [59] were used for fingerprint liveness recognition. The LivDet 2009 [53] and LivDet2011 [82] dataset was used for the experiment. The works by [52] and [20] were also considered for comparative analysis of the results. The performance of WLD was better than LBP for all sensors. The performance of WLD was better than LPQ in LivDet2009 dataset and comparable with LPQ for LivDet2011 dataset. Performance of the combination of WLD and LPQ on LivDet2009 dataset was better and statistically significant ($p < 0.05$) than the other approaches for all the sensors, but in case of LivDet2011 dataset performance of the combination was better than the other approaches except Biometrika sensor. Combination of LBP and LPQ achieved the best for Biometrika sensor in LivDet2011 dataset, but the result is very close with the combination of WLD and LPQ. The combination of WLD and LPQ outperformed the other methods because two approaches are complementary to one other. The authors have not identified the reason of getting better result with the combination of WLD and LPQ for this problem. A through experiment may reveal the discriminant features, which are playing the lead role and the insight behind it also will be clear.

4.5 Forgery detection

4.5.1 Multiresolution WLD for image forgery analysis

Authors proposed the multiresolution-based WLD in image forgery detection [33]. This work addressed the copy–move, splice, and deformation forgery with the images. The human eye is less sensitive with the chrominance than the luminance component of color. Basically, the image forgery is done with RGB color space without no trace of tampering. In this approach, the YCbCr color space was considered and the WLD histogram was generated from the chrominance component. The method was applied on the CASIA TIDEV1.0 dataset [17] with large number of authenticate and fake images with 384×256 pixels size and JPEG file type. The T , M and S parameter values were optimally used as (4, 4, 20). Seven different scales ($C1$ – $C7$) were considered for the experiments, where $C1$, $C2$, $C3$ scale was, respectively, (8, 1), (16, 2), and (24, 3). $C4$ scale was the combination of $C1$ and $C2$ scale, $C5$ scale was the combination of $C1$ and $C3$ scale, $C6$ scale was the combination of $C2$ and $C3$ scale and $C7$ scale was the combination of $C1$, $C2$ and $C3$ scale. In splicing detection, the Cr component outperformed the Cb component for all the scales. 91.54% was found for the Cr component with $C7$ scale, whereas for Cb component, the accuracy was 89.88%. Size of feature vector for $C7$ was 960. Joint histogram of Cr and Cb with $C7$ scale achieved 93.33% accuracy and AUC of 0.93 in splice forgery detection. When detecting copy–move forgery, accuracy of Cr channel with $C7$ scale was 90.69% and Cb channel was 87.77%. Joint histogram of Cr and Cb with $C7$ scale achieved accuracy of 91.52% and AUC 0.88. This method is compared with [75] on CASIA TIDEV1.0 dataset, that used chrominance channels and achieved 79.90% accuracy for splice forgery detection and 76.30% accuracy for copy–move forgery detection. Different types of transformation were applied on the images like deform, rotation, resize. The experiment was done with Cr, Cb and Cr + Cb channel with $C7$ scale. The accuracy of rotation transformation was less than the other transformations. Another experiment was done on different shapes of copied regions in forgery (Arbitrary, Circular, Rectangular and Triangular). In case of splicing detection Cr + Cb component gained the better accuracy but for copy–move detection Cr component gained better accuracy except the circular tampered region. The Cr + Cb component gained better accuracy for circular tampered region in copy–move detection.

4.5.2 Analysis of non-intrusive image forgery using multiscale WLD and LBP

Authors presented a comparative analysis between multiscale WLD and multiscale LBP for the detection of image

forgery [35]. The texture micropatterns are distorted due to the image forgery made by different image processing tools and applications. Two texture descriptors (Multiscale-WLD and Multiscale-LBP) were used to detect the distortion in the texture micropatterns and their performances were compared for different experiments. In this approach to form the multiscale WLD histogram, histogram of three different neighborhoods was combined. The neighborhoods were (radius = 1, pixels = 8), (radius = 2, pixels = 16) and (radius = 3, pixels = 24). The T , M , and S parameters were optimally used as (4, 4, 20) for the experiment and in multiscale scenario the size of the WLD histogram feature vector was 960. Three different versions of LBP were used, LBP with rotational invariance ($LBP_{P,R}^{ri}$), uniform LBP ($LBP_{P,R}^{u2}$) and uniform LBP with rotational invariance ($LBP_{P,R}^{riu2}$) for comparison with the Multiscale-WLD. To avoid the redundant features Local Learning Based (LLB) [70] feature subset selection technique was used. In this approach, for the experiment CASIA TIDE V1.0 dataset [17] was applied. In this dataset, the number of original and tampered images were 800 and 921. Using Adobe Photoshop, the images were tampered. Different transformations and cut and paste method were used for tampering. A total of 459 images were forged by copy–move forgery and rest were forged using splicing. For the classification, SVM (polynomial kernel) was used. They used performance accuracy and AUC for performance evaluation. Three experimental cases were considered in this work—detecting splicing, detecting copy–move, and detecting copy–move and splicing forgery combinedly. They extracted the features from the Cr component, Cb component and from the combination of both the component using feature level fusion (FLF). 94.29% accuracy was obtained for the splicing detection using Multiscale-WLD using FLF (Combination of Cr and Cb) and the AUC was 0.938 ± 0.024 . AUC for the Cr component 0.94 ± 0.02 was slightly better than the FLF. When FLF was applied on copy–move detection experiment, 90.97% accuracy and AUC of 0.90 was obtained. The third experiment for the combination of splice and copy–move dataset was done and achieved 94.19% accuracy for the FLF. In case of the combined dataset the performance of FLF was far better than the Cr. Multiscale LBP was applied for the Cr and Cb channel. ($LBP_{P,R}^{ri}$) using Cr channel achieved the accuracy with standard deviation of 90.48 ± 4.20 and AUC of 0.90 ± 0.05 for splicing detection, whereas $LBP_{P,R}^{u2}$ achieved the accuracy and AUC (90.36 ± 2.94 and 0.90 ± 0.04). The standard deviation for the $LBP_{P,R}^{u2}$ was less than the $LBP_{P,R}^{ri}$ for both accuracy and AUC. The same case occurred with Cb component and $LBP_{P,R}^{u2}$ achieved the best performance (86.55 ± 2.81 accuracy and AUC of 0.86 ± 0.04) than the $LBP_{P,R}^{ri}$ in respect of the standard deviation. In case of copy–move detection $LBP_{P,R}^{ri}$ outperformed the other two variants for Cr and Cb channel. The accuracy and AUC for

copy–move detection with Cr channel was 85.56 ± 4.91 and 0.83 ± 0.06 respectively and 85.83 ± 5.31 and for Cb channel 0.83 ± 0.08 , respectively. But in case of combined dataset, $LBP(P, R)^{u2}$ with Cr channel outperformed the other two variants with accuracy of 85.93% and AUC of 0.86 ± 0.04 . This study showed that in intrusive forgery detection Multiscale WLD achieved better performance than the multiscale LBP and Cr component is best channel from where the features should be extracted to track the distortion due to the image forgery. This method was also compared with [75] using CASIA TIDE V1.0 dataset. Multi-WLD gained better accuracy than Multi-LBP.

4.5.3 WLD for watermark authentication

Authors proposed a watermark authentication technique using WLD descriptor [74]. The system can authenticate the watermark that has been applied or affected by the noise corruption, compression or cropping of the image. The illumination, rotation and scale variance of WLD played the key role for achieving the result. DCT coefficient modification was used for embedding the watermark bits into the image. The WLD histogram was calculated from the image. WLD histogram was stored in register file with the key and encrypted with AES. At the receiver end the WLD histogram was calculated for the distorted image. The register file was decrypted using AES the normalized coefficient of correlation (η) and Euclidean distance (ρ) was calculated as [74],

$$\eta = \sum_{x=0}^{k-1} \frac{M(x) - M'(x)}{\sqrt{\sum_{x=0}^{k-1} M^2(x)} \sqrt{\sum_{x=0}^{k-1} M'^2(x)}} \quad \text{and}$$

$$\rho = \sqrt{\sum_{x=0}^k (M'(x) - M(x))^2}. \quad (11)$$

In this approach, an image was considered as watermarked image when value of η and ρ was beyond some threshold. This method was applied on the different gray level nature images. Total 256 watermark bits were used. The host image and the watermark image is having PSNR ratio of 40.17 db. This work achieved the PSNR close to 40db as of [81] except for 1024-bit watermark. The threshold value of T1 for η is 0.7 and T2 for ρ is 107. Several geometrical transformations were applied on the images like rotation, flipping, cropping, scaling, and translation. The images were rotated using bilinear interpolation using MATLAB7.0. For rotated image, the descriptor does not change significantly with respect to the original image. The variation was for the incorrectness of the rotation technique. In case of flipping, the images were flipped in both horizontal and vertical direction and the approach was robust against the flipping attack in both direction. Vertical and random cropping was done

with the images and system was robust with the cropping also. The images were scaled with the factor of 2 and 4 and normalized to 32×32 pixel. The descriptor was tested with the original watermarked images and the system was robust to the scaling. The images were translated by 16×16 about the centroid and the system proved the robustness against the translation transformation. The approach was tested by applying the Gaussian noise. The experiment was done with the changes of mean and variance. The value of η is near to 0.7 when low contrast image was used, which means the low contrast image was sensitive to noise. Their approach is robust to sharpness, contrast stretching and JPEG compression. 50 unwatermarked images were tested and only one was identified as the watermarked image (host image). The approach reported no false negative rate. The results of the system were compared with SIFT and LBP. SIFT is not robust to the vertical and horizontal flipping. SIFT was also not able to identify the blurred image. For SIFT, it holds the same for degradation and rotation. Also, the complexity of SIFT was very high compared to WLD and WLD takes less time to get executed. LBP has failed to authenticate the blurred images, but performance of LBP was good for other enhancement techniques and the geometric transformation.

4.5.4 Integrating WLD with statistical features for copy–move detection

Authors proposed a copy–move detection method affected by the geometrical transformation in the small copied area using the point descriptor derived from the integration of WLD histogram and some statistical features [66]. In this approach, the key points were extracted using SIFT and WLD components were calculated from every key point and all the pixels around it in the circular area. The WLD components were extracted for 3×3 , 5×5 and 7×7 neighborhood and histograms H1, H2 and H3 were concatenated to make the final histogram. The approach also makes the WLD histogram rotation invariant using the dominant orientation. The input was blurred by the Gaussian blur filter and from every key point WLD components were extracted. The histogram of T equal-sized bins was formed with the width of $\frac{360}{T}$ using the orientation of the pixels proportional to the differential excitation gradient of the pixels. The dominant orientation of the key point is represented by the maximum peak in the histogram. To make the rotational invariance each pixel orientation in the region was subtracted from the dominant orientation of the key point. Several statistical features were calculated such as, mean intensity, mean color channel and the color channels (HR, HG, HB) histogram around each key point with the radius r . Key point similarities were found using the normalized histogram intersection measure. The weighted Euclidean distance (WED) was computed between the two-candidate matched point with some

weights and those points that had the smaller distance value than the T_{WED} (1.8 for this approach) were identified as the matched points. The values of (T, M, S) for calculating WLD components was (8, 6, 20). The spatial clustering was used to group together the matched key points to recognize the copied areas. MICC_F220 [3] dataset was used for the experiment without preprocessing and using geometrical transformations like scaling, rotation, and combination of them. The dataset consists with 110 original and 110 tampered images with resolution varied from 722×480 to 800×600 pixels. MICC_NBCM was created with 121 tampered and 121 original images after application of some postprocessing techniques with the original and tampered images. MICC-SMALL dataset consists of images tampered in a small copy area, created using copying the 48×48 square image area. This dataset consists with 30 original and 30 tampered images. The results were compared with the SIFT-based method [3]. TPR and FPR was used for the result comparison. In case of copy-move-tampered images without any postprocessing, both SIFT-based method and above approach performed equally well. Both the methods were applied on the MICC-SMALL dataset and the multi-scale WLD-based method achieved best result because the WLD features and statistical measures were extracted from the different granularities. In case of rotation postprocessing the SIFT-based method worked better for the MICC_F220 dataset because the dominant orientation calculation in SIFT is very robust than this method. The SIFT had better result for the scaling postprocessing on the MICC_F220 dataset because features were extracted from each key point using different scale space, whereas the WLD components were extracted using only three-scale variation. SIFT also had better result in case of the combination of rotation and scaling postprocessing. WLD had significantly good result for added white Gaussian noise as WLD is robust against the noise. The WLD also had the better results for the blurring and JPEG compression than the SIFT. The mean color channel Y played the key role here. In case of mirror postprocessing WLD had the significantly better result than the SIFT-based method.

4.5.5 Multiscale WLD and its influence in image forgery detection

Authors [34] extended their previous work and used CASIA v1.0 dataset, CASIA v2.0 dataset [17] and Columbia color dataset [55]. The method was same as of the approach described in Sect. 4.5.1. The grid search method was used for the finding the optimal parameter for the kernels used in SVM. This approach used tenfold cross-validation for the experiments. The TPR, TNR, ACC, and AUC was used for the evaluation purpose. The values for WLD (T, M, S) was same as of the previous work (4, 4, 20). In splicing detection, the Cr component achieved the accuracy of 94.29%,

whereas the FLF(Cr + Cb) achieved 94.52%. Though the AUC of Cr and FLF was same but for TPR, TNR, and accuracy the performance of FLF was better than the Cr channel. In copy-move detection, the accuracy using the Cr channel was 91.11%, it was better than the accuracy of the other channel but for TPR, TNR and AUC, the FLF performed better than the Cr channel. The accuracy using FLF for copy-move detection was 90.97%. In copy-move detection, from the image small part is taken and introduced in the original image, so the hidden noise patterns are present in the image. This is the reason for which the performance of copy-move detection is less than the splicing detection. They also created the combined dataset using the splicing and copy-move. The approach was tested using C7 scale on the combined dataset. They used SVM (polynomial kernel) for the classification purpose. LLB was used for dimensionality reduction from 960 to 770. Accuracy of the approach using C7 scale and 770 features with FLF(Cr + Cb) was 94.19%. When detecting splicing and copy-move, multiscale WLD was tested for different transformations, shape and size of the tampered regions. The deformation transformation for splicing detection achieved the best result of 95%, whereas the rotation transformation achieved 88.57% because of small dataset. The results of copy-move detection for different transformations were not reported due to the lack of images in the dataset for this attack. In case of arbitrary shape of tampered region, the accuracy was 94.33%, but for the rectangular and circular shape the approach achieved 85% and 90% accordingly. In splicing detection, using Cr channel multi-WLD achieved best result, even if tampered region size changes. This method achieved 93% accuracy when tampered region size was large, but for medium and small tampered region the accuracy was 91%. In copy-move detection, this method using FLF achieved the best result of 86.67% with large tampered region, which was 86.67%. The above approach was tested on CASIA v2.0 dataset where the number of images was large than CASIA v1.0. After feature selection, the feature vector size for Cr, Cb and FLF were 185, 379, and 359, respectively. The multi-WLD achieved the best accuracy of 96.52% with FLF after feature selection. FLF with feature selection (size of feature vector is 316) achieved 94.17% accuracy for the Columbia color dataset, which consists with the TIFF images. This result was the best, compared to the results shown by the method of [86]. It proves the invariance with the different file format and size of the images. The results of multi-WLD approach were tested with the performance of three variants of LBP for the CASIA v1.0 dataset. LBP^{u2} achieved the best performance (compared to other variants) of 90.36% and 86.81% for splicing and copy-move detection accordingly. In splicing detection LBP^{u2} achieved the performance with the feature selection done by LLB and size of the feature vector was 256, whereas in copy-move detection LBP^{u2} achieved the best accuracy without any fea-

ture selection and size of feature vector was 437. This proved that for the copy–move detection the system needs more number of features to learn because of the hidden noise pattern remaining the same.

4.6 Face detection and recognition

4.6.1 Robust local image descriptor via WLD for face texture analysis

Authors proposed WLD—an enriched local descriptor and tested its performance in recognition of textures and detection of faces. The details of the approach and the formation of the WLD histogram is described in Sect. 3. This subsection describes the application of SVM-WLD in face detection. In case of face detection, each face image was divided into 9 overlapping regions of 32×32 dimension. Each region of the images was of 16×16 dimension and overlapping region of size 8 by 8 was used. The WLD histogram was generated for each overlapped block with the parameter value $M = 6$, $T = 4$ and $S = 3$. For each block, SVM with polynomial kernel was trained using the histogram feature and then used to detect valid face block. When total number of valid face block was above the threshold value, then the image was detected as a face. A positive set with 1,00,000 images of different pose, illumination and lighting conditions and negative set with 31,805 images were used for the training samples. Testing was done on three different datasets—MIT + CMU frontal face dataset [65] with 507 upright faces, AR face dataset [51] with 1512 frontal view face images and CMU profile testing set [67] with 441 multiview face. Due to large training data, resampling method was used to train the SVM classifier. The threshold value for determining face block was set to 8, 7, 6 for MIT + CMU, AR and CMU accordingly. In case of MIT + CMU face dataset without false alarm, SVM-WLD recognized 89.3% faces. They did a comparative study with the other reported methods [10,21,25,29,47,72]. And SVM-WLD had comparable performance with the other methods. In case of AR dataset, SVM-WLD detected 99.3% faces without any false signal. SVM-WLD detected all faces with three false signals. In case of CMU face dataset, 85.7% faces was recognized by SVM-WLD without any false signal.

4.6.2 Integrating WLD-based human perception and LBP for face analysis

Authors proposed an combined technique for recognizing faces based on the human perception using Weber's law and LBP [80]. In this approach on the face image, LBP was applied at first and then the image was partitioned into non-overlapping regions. The WLBPH was computed from each non-overlapping regions of LBP image according to the

perception of the local micro-patterns as weights. Only the uniform patterns (59 uniform patterns) of the LBP were used to form the WLBH. All the weighted histograms were combined to make the feature vector. Nearest neighborhood based on Chi-square metric was used for the comparison between the two global histograms. The approach was applied on the ORL [60], Yale and extended Yale-B dataset [38]. The (P, R) value in this experiment was set to (8, 1). Results was compared with the LBP histograms for different partitioning modes—without partition, 2×2 , 2×4 , 4×2 and 4×4 partition. The best average result was found for the 4×2 partition of the face image for all the datasets. In case of ORL dataset the WLBPH achieved 98.0% accuracy, whereas LBP achieved 96.5% accuracy for 4×2 partition of the face images. WLBPH also outperformed the LBP for the Yale dataset. But for the extended Yale-B dataset LBP achieved 99% accuracy and outperformed the WLBPH due to the variable illumination in the dataset.

4.6.3 Multiscale and spatially enhanced WLD for face analysis

Authors presented a technique of face recognition using the spatially enhanced multiscale WLD [32]. The spatial information played the key role for extracting local micro patterns for better description and to increase the discriminative power. In this approach, WLD components were calculated from different regions of the image with different neighborhood. The components of different neighborhood were concatenated to make the resulting 1D WLD histogram. Fisher score was used for the selection of the key features. Key features were always having larger F -score. The approach was tested on FERET [63] and AT&T face database [60]. FERET database consists of large number of images with $60 * 48$ -pixel size, collected during different photo sessions. The training was done with 1204 images (fa set) and testing was done with 1196 images (fb set). The testing images were taken in different illumination, facial expression and poses. The WLD parameter values (T, M, S) for the experiment was set to (16, 4, 5) and the number of block in a face image was set to 16 (4×4 with $n = 4$). The approach achieved 96.15% accuracy with the above configuration and using the combination of three scale [(8, 1) + (16, 2) + (24, 3)] without the feature selection process. The application of feature selection using F -score reduced the feature vector from 15,360 to 910 and the approach achieved 98.07% accuracy for the above configuration of the parameters. Thus, F -score-based feature selection reduced the amount of the redundant features, which improves performance and takes less time. The AT&T face database consists of 40 subject (10 images for each subject) acquired with some pose variation. The approach was tested with two-experiment protocol. The first experiment

done with 3 training images and 7 testing images per subject and achieved 96.89% accuracy. The second experiment was done with 5 training images and 5 testing images. This method achieved 99.37% accuracy in this experiment. The performance of LBP and Eigenface using FERET and *At&T* face databases was (94.67%, 74.41%) and (97.96%, 95.37%). Clearly, this approach outperformed the Eigenface and LBP-based approach for the above two datasets.

4.6.4 WLD for race identification via face images

Authors proposed a WLD-based race recognition technique from the face images. In this approach, face images were normalized and from them WLD features were extracted [54]. They extracted most discriminative features using the Kruskal–Wallis method. The FERET dataset was used for the experiment. They had taken the images from five major race groups, such as Black, White, Asian, Middle and Hispanic. All the race groups were with more than 50 subjects. Training was done using the *fa* (1204 images) set and testing was done using *fb* (1195 images) set. There are total of 1180 images of 60×48 -pixel size of five major race groups. Different values of the *T*, *M* and *S* were used (*T* = 6 or 8, *M* = 4 or 6 and *S* = 10 and 15). The changes of these parameters did not affect the result of the application. In the experiment the (8, 4, 5) configuration was used for *T*, *M* and *S*. For the comparison purpose PCA features were used with 200 principal components and accuracy achieved 79.17%. The proposed approach with the full image size obtained 74.09% accuracy with the above configuration of the parameters. Clearly the result is worse than the PCA because the local features were used to express the global features. Different blocks of the image should be used to extract the features so that the global and local information could be captured. Out of the several block sizes, 10×16 size blocks achieved the average performance 96.88% using the city block minimum distance classifier. The performance of Euclidean distance classifier is comparable with the city block classifier but the performance of the chi-square method was worse than the others. The KW techniques were applied with different significance values to get a threshold above which the features are discarded. The significance value 0.16 with 1632 features achieved the same performance as of the full length WLD. This method achieved 97.74% accuracy for Asian race group, 96.89% accuracy for Black race group, 92.06% for Hispanic race group, 98.33% for Middle race group and 99.53% for White race group.

4.6.5 Weighted LBP based on WLD for face analysis

Authors proposed an infrared-based technique for face recognition using weighted LBP [79]. The intensity of pixels in local regions (IOL) is calculated using the equation,

$$\text{IOL} = \frac{1}{p} \sum_{i=0}^{p-1} |g_i - g_c|$$

. Here g_i represent intensities of neighbored pixels, and g_c represent intensity of current pixel. The mouth, nose and eye region play crucial rule in infrared face recognition. In formation of normal LBP histogram same weight 1 is assigned to each micro-pattern. This paper addressed this issue and assigned adaptive weight. The uniform weighted LBP histogram was extracted from each of the non-overlapped regions of the image and combined to make the final feature representation. The feature vector is of length $59 \times$ total patches under consideration. The chi-square statistics was used as the nearest-neighbor classifier. The training set consists 500 images of 50 subjects captured in a controlled air-conditioned environment. The test set is portioned into two groups: one is same session data which consists with 500 thermal images of 50 subjects captured with same environment setting as the training set and other is time-elapsd data where for each individual 165 thermal images were present. The resolution of the images for the experiment was 80×60 . The approach achieved 98.6% accuracy for same session data using WLBPH with 2×2 partitioning method and (8, 1) scale whereas LBPH achieved 97.2% accuracy. The difference of accuracy between LBP and WLBP was very small in this experiment. But in case of time-elapsd data, the WLBPH achieved 91.2% accuracy, whereas LBP achieved 87.4% accuracy. The results were compared with PCA and LDA. PCA and LDA combination achieved the accuracy of 92.4% and 33.6% accuracy for same session and time-elapsd data, respectively.

4.6.6 Weber faces for self-similarity representation for kinship classification

Authors proposed a kinship classification technique using the self-similarity of the Weber faces [41]. Due to the non-availability of the dataset and the inherent variation among the kins, kinship classification is a less explored application. In this approach at first using the Adaboost face detector they detected the faces from the image. The face images were normalized by WLD, and hence illumination factor was removed. The images were represented by the reflectance only. The key points were identified by local extrema of Gaussian differences (DOG). The discriminatory key points were extracted using the threshold centering and gradient detection. The facial features and similarity measure of textures were encoded using the self-similarity descriptor (SSD). The approach was applied on the IITD Kinship Database [41] and UB Kinship dataset [78]. IITD Kinship database consists with 544 images of 272 pairs and 272 non-kin pairs were added with this database. There were four different societies: Afro-American, American, Indian and Asian except Indian. The kinship relationship had seven relation categories. The UB dataset consists with 200 kin-pairs. SVM classifier was

used for the binary classification (kin or non-kin). They used RBF kernel on IIITD kinship dataset and achieved 75.2% accuracy. The method of [87] achieved 57.5% for the IITD kinship database. The high accuracy was observed for the Indian and American ethnicity because 85% of the images belong to these groups. For all the kinship classes and ethnicity, the approach of [41] outperformed the method in [87]. In case of the UB dataset only 175 groups were considered for the experiment. In case of child versus older parents group, this approach achieved 52.5% accuracy and when using child versus young parents group 55.3% of accuracies was achieved. The accuracy for this dataset was less because of the failure of the detection of key points. In case of UB dataset also, the above approach outperformed the [87] by at least 4.1%.

4.6.7 Region-based WLD for face analysis

Authors used WLD features for face recognition purpose. In this approach, the face images were smoothed using the Gaussian filter [22]. The smoothed face image was partitioned into sub-regions and WLD components were calculated for each sub-region. Sobel operator was used to extract the gradient orientation to avoid the noise disruption in the orientation component. All the sub-regions of the test and gallery images were considered and the Euclidean distances in the feature space were computed. Voting based decision fusion was used for the improvement of the performance. ORL and Yale dataset were used for testing this approach. In ORL dataset total forty subjects are there and each having ten different images (112×92 – pixelsize) with small occlusions, orientations, different scales and various expressions. In Yale dataset, all 15 individuals have 11 different images (100×100 – pixelsize) and images expresses the facial expressions in different lighting conditions. For evaluation purpose, leave one out strategy was used. The recognition accuracy of the approach for ORL dataset was 99.25% ($T = 10$ and $N = 5$) and for Yale dataset accuracy was 96.97% ($T = 20$ and $N = 10$). They done experiments by changing number of sub-regions of image. The accuracy was increased if number of blocks increased up to a certain limit, and then the recognition accuracy decreased. The observation is that there should be tradeoff between the number of regions or blocks of an image. The results of this approach were compared with two popular texture descriptors, LBP and LTP (for threshold value 1 to 4). The recognition accuracy for ORL dataset using LBP was 96% whereas best result for LTP was 99% for threshold 2 and for Yale dataset LBP achieved 90.30% accuracy and LTP achieved 91.52% for threshold 3. Clearly the WLD approach outperformed LBP and LTP. The also done comparative study with some other reported methods, viz. ICA, Eigenfaces,

Kernel Eigenfaces, and 2DPCA. The above method outperformed all the methods in terms of recognition accuracy.

4.6.8 Integrating HOG and the WLD for recognition of facial expression

Authors proposed a ensemble technique for recognizing facial expressions using HOG and WLD [76]. The local information of image was encoded using the gradient and orientation density distribution. WLD was used to encode the lack of information and shape distribution. The approach ensemble the HOG and WLD to get the advantages of both the descriptor. In this approach, the faces were detected using the AdaBoost face detector. The scaled images were normalized into 128×128 size. The images were divided into several blocks and assigned different weight to the blocks to encode different behavior in facial expression recognition. The optimal parameters of T , M , S was (8, 3, 5). The approach was tested on the JAFFE [50] and Cohn–Kanade [37] face dataset. JAFFE dataset contains total 70 facial expressions taken from ten individuals. Every expression has 3 or 4 images in it and total of 213 images are there. In Cohn–Kanade dataset 100 university student images are there with age group from 18 to 30 years. They used chi-square and nearest neighbor for the classification purpose. Experiment was done for three times on JAFFE dataset, where each time 1–2 images were taken for training and another 1–2 images were used for testing purpose. There were 15 images in both training and testing dataset. The results were compared with LBP, AAM [28] and Gabor Wavelet. AAM has the better performance than LBP. This approach achieved 93.97% accuracy and outperformed the other methods. This result was the average performance of all the classes. In case of Cohn–Kanade dataset for every subject, there were six images present per expression. They used three images of each expression of different people for training and other images were used in testing. The experiment was done for four times. The approach achieved 95.86% accuracy and outperformed the Gabor, LBP, and AAM. The proposed approach had taken little more time than the LBP, but time complexity is less than the Gabor and AAM approach. Therefore, the fusion of WLD and HOG achieved the best for facial expression recognition.

4.6.9 Nonlinear quantization-based multiscale WLD for face analysis

Authors proposed a face recognition technique using a nonlinear quantization-based multiscale WLD [46]. In this approach, the nonlinear quantization was applied to compute differential excitation and orientation. The face image was divided into some non-overlapped regions (internal sub-image). Each of the internal sub-image was taken as the

center and several sub-images with different sizes were extracted. WLD components were extracted from each scaled sub-image and fused together to make the feature vector for each internal sub-image. The chi-square-based nearest neighborhood was used to calculate the similarity of two sub-regions. They used a voting function on individual result of the sub-regions of the image. They used Yale, AR and FERET datasets for testing this approach. In Yale dataset, total 165 facial images (100×100 -pixel size) of 15 individuals were present. The images were taken in different illumination conditions and details (glass or without glass). In the experiment, facial images of 50 men and 50 women were taken into consideration. There were 13 images (100×100 -pixel size) per person per session (two session separated by 2 weeks). A Small portion FERET dataset of 1400 images from 200 individuals was considered for the experiment. Several experiments were done to set the number of internal sub-regions and number of sub-images per internal sub-image. For the AR and FERET dataset, if the number of internal sub-images increased the recognition accuracy was also increased. The average best performance for AR and FERET dataset achieved with 81 (9×9) sub-images. The performances of the system were increased at first with the size of the sub-images increased and then keeps stable or degrade. The number of sub-images was 4 for both the dataset. A comparative study was done with the other reported methods in face recognition. In case of AR dataset, the proposed method achieved 96%, 95.33% and 96.67% accuracy for the facial expression, sunglass, and scarf occlusion probe set. The above method got the best result of 89.83% accuracy with the FERET dataset, when training was done with 4 images and testing done with rest 4 images. Using the leave-one-out strategy this method achieved 98.18% accuracy for the Yale dataset and outperformed the ICA, Eigen faces and 2DPCA approach. The results of the approach on AR and FERET dataset also outperformed the other state-of-the-art approach. The most important fact is that the result using the nonlinear quantization of WLD is far better than the linear quantization method. The approach also performed better for the random occlusion than the SRC and the partitioned SRC algorithm.

4.6.10 Realistic facial expression learning from web images

Authors presented a search-based framework for collecting web search engine-based images of facial expressions [83]. This approach was based on the active learning approach (using SVM) to select the relevant images from the noisy result given by the search engine. A novel histogram contextualization-based WLD was also proposed for the handling of such a challenging dataset. However, there are some popular datasets (CK and JAFFE) available for the recognition of facial expression, but the number of samples are not enough to capture the task reliably. This approach can col-

lect the large number of samples and thus enhanced the area of facial expression recognition research. In this approach, Viola–Jones [73] approach was applied for removal of noisy images (low quality or lack of frontal face). The Average of Synthetic Exact Filters (ASEF) [9] was used for the localization of eye in the face and different face image were aligned in a common coordinate. For further improvement of the dataset the semantically relevant images were selected using the binary SVM, which was learned from active learning-based training set. Multiscale-WLD was used for the recognition of the images of the dataset prepared. Each face image was downsampled and divided into some non-overlapped regions. From every sub-region of downsampled images, WLD components were computed and fused together for the construction of the 1D feature vector. To encode the spatial contextual information of the image, contextual information histogram was constructed. The dataset consists with seven categories of expressions with 2000 to 2500 images in each category. For the experiment, validation set G_v (for determination of stopping criteria of active learning) with 350 images (50 images of each category), seed training set (20 images for each category) and active learning pools with rest images were prepared. In the experiment number of regions of an image was $25(5 \times 5)$, number of downsampled version was 3 at scale 0.6, and WLD parameter T, M, S was (6, 2, 4). The size of feature vector for contextual multiscale WLD histogram was 10,800 and PCA was used for the reduction of feature vector to 400. $LBP_{8,2}^{U2}$ with 59 bin was used for the comparison purpose. In case of the proposed dataset and for CK and JAFFE, the results were good for the happiness and neutral category when fivefold cross validation was used. The misclassification occurred due to some same appearances between the categories. The approach was confused with the anger, fear and sadness category. Another point is that for the disgust and surprise category the result was not promising for the proposed dataset but results were good for the CK and JAFFE dataset. The reason may be the two expressions are over exaggerated for the CK and JAFFE dataset which is not applicable for the real-world environment. The proposed multiscale WLD achieved 59.9%, 85.7%, 95.7% accuracy for proposed dataset, JAFFE and CK dataset whereas multiscale LBP achieved 48.8%, 84.9% and 92.5% accuracy's accordingly. It shows that multiscale WLD clearly outperformed multiscale LBP. Another experiment was also done for the cross-dataset. Training using the proposed dataset and testing using the proposed dataset achieved 48.8% accuracy, and testing using CK and JAFFE achieved 49.3% and 45.1% accuracy. When the system was trained using CK dataset and testing using CK, JAFFE and proposed dataset achieved 95.6%, 35.3% and 26.4% accordingly. In case of training using JAFFE and testing using JAFFE, CK and proposed dataset achieved 85.7%, 35.4% and 24.2% accordingly. Their approach is robust to cross-dataset as well. The approach

achieved the benchmark result 58.2% when trained with the proposed dataset and tested on the BU-3DFE dataset.

4.6.11 Region-based multiscale WLD in e-healthcare for facial emotion recognition

Authors proposed facial emotion recognition approach based on multiscale WLD for the initial assessment of the patient in e-Healthcare system [2]. In this approach, face image was cropped from the full image (via mobile app) and sent to the cloud server to extract feature. Multiscale-WLD was used for extraction of WLD components from each sub-regions of the image. Only two neighborhoods were used—one with (8, 1) and another with the (16, 2) scale. The features of the different subregions were combined to get the final feature histogram. They used Fisher discrimination ratio (FDR) for section of significant bins from the feature set. After capturing the emotion, the information is sent to the e-Healthcare professionals. The approach was applied on CK and JAFFE dataset. In JAFFE dataset from 10 female Japanese actresses total 213 face images were taken. In CK dataset, 100 university students (96 students were selected finally) from different ethnicities. Here for experiment purpose, three most expressive emotional frames were considered from total 408 image sequences. In case of neutral, the first frame of 408 image sequences were selected. In case of JAFFE dataset, an eye localization method was applied and using rectangular approach cropped face images were created. In case of CK dataset eye labeling was already provided. The images of both the dataset were in grayscale and with 150×110 pixels size. SVM with RBF kernel using single scale WLD achieved the best accuracy of 82.34% and 76.28% for the CK and JAFFE dataset, respectively. The face images were subdivided using four variations—two horizontal blocks, two vertical blocks, three block and four block. The best accuracies of 99.28% (for JAFFE) and 97.44% (for CK) were achieved with the four blocks with SVM (via RBF kernel and WLD parameter, $T = 6$, $M = 8$, and $S = 20$). The FDR ratio was used to detect the significant features and achieved 98.82% accuracy with CK dataset and 97% accuracy with JAFFE dataset, respectively. Their method outperformed the LDP, LFDA and the combination of LBP with isomap. Best results achieved for surprise emotion, and in case of anger emotion the accuracy was lowest.

4.6.12 Pose-invariant face recognition using WLD and facial landmarks

Authors presented a face recognition technique which is invariant to pose using combination of WLD and facial landmarks [85]. Due to the facial rotation, the intra-class variation increased and face recognition performance is degraded. In this approach, from the inner face N (31) landmarks were

used. Multiscale patches were extracted at each landmark with the window size same as the minimum of the width of eyes and mouth from the training set. Thus, total N (31) local features were extracted from the 31 landmarks of the inner face. According to the different semantic components, six groups were formed dividing the N (31) landmarks. From each of the six group randomly, one local feature using local-random strategy was selected and formed a fusion feature of length $6 \times L \times T1 \times T2$, where L was the number of patches, $T1$ was the number of differential excitations and $T2$ was the number of orientations). Total M number of fusion features were extracted recursively. Every face image had $N + M$ number of feature vectors. The cosine angle-based KNN was used to draw the similarity between the feature vectors. To establish the significance of the fusion features an experiment was done with the subset of the FERET dataset. During training, 1 to 5 samples were used. The proposed approach achieved over 90% accuracy, when training was done with 2–4 images from every subject, whereas the local feature (N) achieved about 85% accuracy when for each subject 4 training samples were used. The value of $C (= M/N)$ was 0.5 for the applications of this approach on the FERET, ORL, GT [4] and the LFW dataset [30]. A comparative study was done with the other reported methods in face recognition. In FERET dataset total 1400 grayscale images of 80×80 dimension were taken from the 200 subjects. The value of M and K (KNN with 3 neighbor) was set to 10 and 3 respectively. In the experiment, 1 to 5 training samples were used. This method achieved the best performance 92.5% with 4 samples for training. In case of three training images per class this method achieved 91.9% accuracy whereas the traditional WLD achieved 75.5% accuracy. In case of ORL dataset, 400 face images were taken from the 40 subjects with the dimension 112×92 after cropping. LDA achieved 90.8% accuracy when six training images were used whereas this method achieved 97.5% with only 5 training images. Performance of the system is very effective even the number of samples are less. In case of GT dataset, from 50 people 750 color images were collected. The images were then grayscaled and cropped and resized to 120×100 . The results proved this approach has 13% less classification error than the WLD. This method achieved the best performance of 85.1% when training was done with 7 images per person. The LFW dataset is an unconstrained dataset with 13,000 face images and two or more distinct face images were taken from 1680 people. The experiment was done with the face image of 158 subjects with no less than 10 photographs and no more than 20 photographs. This method achieved improved performance than the other reported methods with the same experiment protocol and achieved the best performance 44.8% when number of training images per person was 7. Another experiment was done on the ORL dataset with random occlusion. The block size was set to 20×20 and 30×30 . In this experiment also,

this approach outperformed the other reported methods. The final observation is that it is very robust to pose and occlusion.

4.7 Variations of the WLD

4.7.1 Gabor wavelet WLD

Authors presented the WLD based on Gabor wavelet (GWLD) [69]. In this approach, every pixel of Gabor magnitude map was taken for the calculation of differential excitation and orientation. GWLD was applied for the Bovine Iris recognition. In this approach, active contour model helped to find the edge curves at inner and outer iris. The elliptical boundaries were obtained by using the edge curves. Further histogram quantization was applied for the enhancing of bovine iris region. To compute the GWLD, multiple Gabor magnitude maps (40 Gabor magnitude maps for the experiment) were extracted in the frequency domain using the multiscale and multiorientation Gabor filters. Next in the Gabor magnitude maps, from every pixel differential excitation and orientation were calculated to form the GWLD histogram. Finally, the 1D GWLD feature vector was generated. Histogram intersection method was used to find the similarity between the two GWLD histograms. The GWLD descriptor was used on the SEU bovine iris dataset. 18 subjects with 90 original images were selected for the experiment. The size of the bovine iris image was 253×61 . Since 40 Gabor magnitude maps, the 1D GWLD histogram was generated using the 40×61 pairs of differential excitation and orientation. Several experiments were done for the parameter (mask size of the Gabor filter and length of the histogram) selection. The best performance was achieved with the 5×5 mask size. The feature vector length was 32000. The proposed GWLD achieved 98.87% accuracy on the SEU bovine iris dataset. On the same dataset, LBP and center-epsilon LBP achieved 93.1% and 95.79% accuracy. The WLD and modified WLD achieved 96.20% and 98.73% accuracy accordingly and Gabor filter alone achieved 98.30% accuracy. GWLD outperformed other methods with no significant values except LBP. The feature vector is large and it suffers from heavy computationally cost.

4.7.2 Log-Gabor WLD

Authors proposed WLD based on Log-Gabor (LGWD) [45] for recognizing faces. The Log-Gabor representation of the image and Weber LBP was used to form the LGWD. In this approach, log-Gabor transform was applied on each face image and WLD based on Log-Gabor magnitude (LGMWD) [45] and WLD based on Log-Gabor phase [45] were extracted from it. LGMWD encodes the variation of center pixel with its neighbors, whereas the phase feature is encoded by the LGPWD. The LGMWD and LGPWD feature histogram was concatenated to make the final vector.

This approach was applied on the ORL, Yale and UMIST face dataset [24]. ORL face dataset consists with 10 different images of 40 different subjects with 92×112 dimension. In Yale dataset 165 gray scale images of 15 individuals with 100×100 dimension. UMIST dataset contains 565 images from twenty people with changes in pose, race, sex, appearance. The images are of 92×112 pixels and gray-scaled. For the experiment the images of each dataset was divided into K subsets ($K = 10, 8, 5$). Training was done using one subset. They used average result of K iterations to report the performance of the approach. In the first experiment the results were compared with the WLBP and LBP. The performance of individual components of the approach (i.e. LGMWD and LGPWD) was analyzed. The performance of LBP and WLBP was not satisfactory than the individual components of the approach. LGMWD contributed more than the LGPWD for all the datasets. With nearest neighbor (chi-square distance metric) classifier, the performance of approach using combined histogram (i.e., combination of LGMWD and LGPWD—LGWD) was 89.88%, 81.52% and 93.48% for ORL, Yale and UMIST face dataset for 5 subsets used as training. With the same settings the performance of LBP, WLBP, and Gabor-WLBP for ORL, Yale and UMIST dataset was (55.44%, 47.88%, 77.61%), (69.50%, 56.97%, 86.17%) and (83.88%, 67.42%, 89.87%). For smaller value of K , the LGWD performed better. This approach outperformed the Gabor-LBP, Log-Gabor statistic, Log-Gabor magnitude PCA, Log-Gabor phase and MBC method also. This approach proves the discrimination power of WLBP and Log-Gabor transform.

4.7.3 Memetically optimized MCWLD

Authors proposed an evolutionary memetically optimized multiscale circular WLD for crime investigation [8]. The discriminative information was extracted from digital and sketch face images. In this approach, 6×7 non-overlapping portions of images were used and multiscale circular WLD components were extracted. Three different scales were used such as- (number of pixels = 8, radius = 1), (number of pixels = 16, radius = 2) and (number of pixels = 24, radius = 3). The Memetic algorithm [42], was applied to find the optimized weights for different facial regions. For the comparison between MCWLD histograms, chi-square distance was applied. This approach was applied on viewed sketch dataset (combination of the CUHK [77] and IIIT-Delhi sketch dataset [7]) with 549 pairs of sketch-digital images, IIIT-Delhi semi-forensic sketch dataset with 140 sketch images by an expert from the digital images as per memory and Forensic sketch dataset with 190 forensic sketches from different source. Three experiments were done with the viewed sketch dataset, using digital face images as gallery and as probe sketch images were used. Training was done using

40% data and 60% was used for the testing. The results were compared with WLD, MWLD, MCWLD, SIFT [39], EUCLBP+GA [7], LFDA [40] and two commercially package named COTS-1 and COTS-2. In case of CUHK dataset this approach achieved 97.28% rank-I accuracy and at least 2% effective accuracy than the WLD, EUCLBP+GA, SIFT MWLD, and MCWLD. There was a slight performance hike using this approach than the LFDA. The proposed approach achieved at least 5% better performance than the two commercial products. MWLD achieved 1% better accuracy than the WLD for all the experiments and the multiscale circular WLD improved 1%, 2.8% and 2.9% accuracy on CUHK, IIIT-Delhi and the combined dataset. Memetic optimization achieved 2.2% improved accuracy on CUHK dataset, 5.7% improved accuracy when IIIT-Delhi dataset was used and 4.9% improved accuracy when both the dataset was used. This proved the effectiveness of the memetic optimization. In case of the combined dataset, the proposed approach achieved at least 2% better accuracy than the other methods and 13% better result than the two commercial products. In case of semi-forensic sketches memetically optimized MCWLD achieved rank-I accuracy of 63.24% and outperformed the other methods by 2–5%. This approach achieved at least 9% better result than the two commercial systems. In total four experiments were done for forensic sketch images matching. For conducting experiment 1, for training 140 sketch-digital pairs were taken from IIIT-Delhi viewed sketch dataset [7] and testing was done using 190 forensic images. In experiment 2, for training 140 pairs of sketch-digital pairs were taken from IIIT-Delhi semi-forensic sketch dataset [8] and 190 forensic sketches and 599 digital face images were taken for testing. Experiment 4 was done with 140 pairs of sketch-digital images for training and testing was done with the rest of data with preprocessing and without preprocessing. In case of experiment 1 the propose algorithm achieved about 2% better accuracy than the existing algorithms and at least 3% improvement than the two commercial systems. When the system was trained with the semi-forensic sketch images (experiment 2) about 7% improved performance achieved than the experiment 1 with the proposed algorithm and 4% improved accuracy in case of the other methods. In case of experiment 3, all the algorithms achieved an improvement of 2–3%. For experiment 4 when system was trained with viewed sketch dataset and without preprocessing of forensic image data the approach achieved 23.94% accuracy and it achieved 3% improved accuracy than all the reported techniques. The proposed approach achieved 28.52% accuracy when the system was trained with the semi-forensic sketch images without preprocessing. When compared to other reported methods, 4% better accuracy was achieved using this approach and compared to the two commercial methods 15% better performance was achieved. The results claimed that for the matching of the forensic images with

the digital images, training using the semi-forensic sketch images was very effective than the viewed sketch dataset.

4.7.4 Weber LBP

Authors proposed weber local binary pattern (WLBP) which was formed by using the differential excitation of WLD and LBP [48]. The differential excitation extracted the perception features and LBP extracted the local feature. The Laplacian of Gaussian (LOG) was used to enhance to reduce noise level. The interval of the differential excitation component was divided into low $(-K, K)$ and high perception pattern $[(-\frac{\pi}{2}, -K)$ and $(K, \frac{\pi}{2})]$ where K is a constant. The interval of differential excitation is further divided. Uniform LBP was computed using $(8, 2)$ scale to extract local features. The WLBP 2D histogram was generated using the S number of intervals of differential excitation of WLD and T number of pattern in LBP. The 2D histogram of size $T \times S$ is further coveted into 1D histogram for the better discriminative features. WLBP was applied in face recognition (FERET and AR dataset) and texture classification (Brodatz and KTH-TIPS2-a dataset). They used 5 intervals (represented by S) in differential excitation and (P, R) value was $(8, 2)$. Each face image was divided into 4×4 regions to compute the spatial features. Chi-square distance was used for the similarity measurement and NN classifier was used for the classification purpose. The result of WLBP was compared with PCA, KPCA, 2DPCA, LBP and WLD. In case of face recognition on the FERET dataset the WLBP achieved accuracies of 91.07%, which is 3.49% better than the WLD and 8.24% better than LBP. WLBP had at least 21.34% better accuracy than PCA, KPCA, and 2PCA. In case of AR dataset, the experiment was done for different time span, lighting condition and different expression of face. WLBP achieved at least 5.60% better accuracy than the other methods in case of different time span but with the changes of lighting condition and expression, though the performance of WLBP was better than the other methods but the performance of WLBP was not significantly better than the second position holder method (LBP for lighting and WLD for expression difference). Thus, WLBP extracted the similar set of features even if the face images were taken in different time. Another experiment was done on the FERET dataset in addition to white Gaussian noise. WLBP performed significantly better. In case of texture classification, WLBP achieved 95.68% accuracy on the Brodatz dataset which is 0.63% better than the second position holder multiscale LBP (MLBP). Again, in case KTH-TIPS2-a dataset MLBP stood second position and WLBP achieved 64.42% accuracy (4.74% better than MLBP). WLBP descriptor is robust to different lighting condition, expression, and time span in case of face recognition and robust to rotation, scaling, illumination and pose. But the major drawback is it uses large size of the feature vector.

4.7.5 High-order information of the WLD

Authors proposed higher order statistical measures of WLD for improving discriminant representation power of image representation [26]. This approach extracted highly discriminative features from an image using the robust WLD. They used Weber's law to represent original image in the domain of differential excitation and local patch was generated from the transformed image. A generative probability model was employed to adaptively characterize the WLD space and learn parameters. The higher order statistics of WLD was applied on three image classification problems- texture classification using KTH-TIPS2-a dataset, food images classification using PFID dataset [14] and HEp-2 cell recognition using HEp-2 cell dataset [62]. Linear SVM was used for all the experiments to gain the promising recognition performance. In case of KTH-TIPS2-a dataset results were compared with GMM and other local descriptors. The higher order statistics of WLD achieved the best performance 75.35% for GMM with 128 components whereas the SIFT and microstructure base approach achieved 73.59% and 71.46% accordingly. The proposed approach outperformed the other two local descriptors for all the component numbers of GMM. The higher order statistics of WLD also achieved better performance than the popularly used local descriptor LBP. LBP achieved 58.1% accuracy whereas this approach achieved 75.35% and 75.58% for 3×3 and 5×5 local structures. Another important observation is that the proposed approach achieved best performance for low, first and second order statistics of WLD. In case of PFID food dataset the proposed approach achieved 36.9763% accuracy and outperformed the color histogram (11.2%), SIFT with lower order statistics (9.3%), WLD (28.05%), SPLF (28.2%). The task for HEp-2 cell dataset is to recognize the intermediate and positive intensity image. Their approach achieved the best performance of 95.97% for positive intensity image and 85.14% for intermediate intensity image using 3×3 local structure. It outperformed the SIFT, LBP, WLD, and microstructure for both the cases.

5 Authors comments on this review

WLD has been applied in several different domains ranging from biometrics, medical image analysis to agriculture. Regardless of the applications, WLD's performance remains promising and statistically significant. The reason behind this is due to the fact that human perception is converted into the descriptor, where illumination and rotational invariance, robustness to noise and scaling are considered. Even though we have several variants of WLD, we observe that the Original WLD performs significantly better in common tasks, such as texture classification and face recognition as

compared to other local descriptor, such as LBP and GLCM. Further, understood that if WLD is combined with variance histogram, performance can be augmented. Recently, deep learning emerges as key tool in almost every domain of computer science. We find that when WLDRI kernel in deep neural network model performs better than conventional one. Skin disease identification is one of the examples to state fact that the DNN-based WLDRI is better [61]. The multiscale and multiresolution-based WLD performed well in agriculture safety and face recognition domain. In other cases, WLD is combined with many well-known computer vision image descriptors and in each combined version, the performance is better than when they are separately used. This means that WLD is found be complementing other descriptors. Therefore, WLD has become an obvious choice in both emerging and existing image recognition problems. With WLD, this review work groups together potential related methods, experiment protocols and performance. Like other descriptors in the domain, we need to tune parameters for WLD. We can consider it as an open research area to study parameter optimization of WLD. Also, we find it interesting that performance vary when different classifiers are varied. In the view of the recent developments of deep neural network, it may be a new research area to explore the difference between what we achieved and what we can. Looking at the performances of the methods using WLD, we find it interesting to apply WLD in the other codomains, such as image processing and pattern recognition.

6 Conclusion

We have presented a comprehensive survey on the different approaches applied in various image classification problems that are based on the robust and rich local feature descriptor WLD. This survey illustrates all the applications that are based on WLD. Indicatively WLD feature descriptor has been applied on texture classification, medical image classification, face detection, and recognition and in several agricultural applications. Texture classification problem itself resolves many other pattern recognition problems. In the Sect. 3 we illustrated the theory, principle and characteristics of WLD. The WLD is robust to noise, rotation, changes of illumination and changes in scale. This robustness is the key factor in gaining the best results in different applications. The Sect. 4.1 described the approaches used in texture classification. Maximum of the approaches were applied on the KTH-TIPS2-a and Brodatz dataset. Recently proposed deep neural network-based WLDRI [5] achieved the best result than the other WLD-based approaches. Next, the Sect. 4.2 explains the methods of the application of WLD on medical diagnosis. [61] proposed WLDRI and applied for the skin diseases recognition. A single work was done by [31] for the

detection of mass in the mammograms. [5] have proposed deep neural network-based WLDRI for the three popular skin diseases recognition and achieved the improved performance than the method of [61]. The Sect. 4.3 illustrates a biometric approach to identify cattle from muzzle print images. Combination of WLD and LPQ was used for fingerprint liveness detection [53] on LivDet2009 and LivDet2011 dataset and achieved the best performance than the other state-of-art methods. WLD has been efficiently utilized in image forgery detection and achieved significant better results. Combination of WLD with other local descriptors was proposed to increase the recognition performance. In recent years WLD has been extensively used in face detection and recognition on some standard datasets like ORL, Yale, AR, CMU, MIT+CMU. The approaches in face analysis domain properly used the local and spatial information from the faces using WLD. Some of approaches combined the WLD with some global descriptors for gaining the better discriminative features for the recognition. Gabor-Weber local descriptor [69] was proposed and used in bovine iris recognition for the security in the agricultural field. Log-Gabor weber local descriptor LGWD [45] was proposed for face recognition purpose. Memetically optimized WLD [8], a variant of WLD was used in forensic sciences for the matching of digital face images and sketches. Another popular WLD variation is WLBP which takes the benefit of both WLD and LBP. In recent times, higher order statistics was studied and applied on texture classification using KTH-TIPS2-a dataset, food images classification using PFID dataset and HEp-2 cell recognition using HEp-2 cell dataset. This approach achieved improved performance than the other reported methods. It has been seen that due to robustness and effectiveness in extracting the local micropatterns of the image WLD resulted always the best in the above application domains. But the problem in these filed is for different or same type of problem several classifiers were trained with different hyperparameters. The future work may be to develop some framework such that it would be possible to apply the same for all the same type problems. Another approach would be to apply WLD in other unexplored research fields and to analyze the results with the other reported methods.

Acknowledgements Authors are thankful to CMATER LAB, CSE Dept., Jadavpur university for getting the well and good infrastructure during this work. The first author is also grateful to Dr. B. C. Roy Polytechnic, Durgapur for the overall support during the progress of this work. The work is partially supported by the SERB, GOI via project no SB/S3/EECE/054/2016. Also, the authors would like to acknowledge Sk Md Obaidullah for his initial discussion.

Author contributions A.B., N.D. and K.C.S.: Conceptualization and Methodology. A.B. and N.D.: Writing- Original draft preparation. N.D. and K.C.S.: Supervision. K.C.S.: Writing- Reviewing and Editing

Compliance with ethical standards

Conflict of interest Authors declared no conflict of interest.

References

1. Agrawal, D.G., Jangale, P.M.: Dynamic texture feature extraction using weber local descriptor. *Int. J. Eng. Res. Appl.* **4**(3), 502–506 (2014)
2. Alhussein, M.: Automatic facial emotion recognition using weber local descriptor for e-healthcare system. *Cluster Comput.* **19**(1), 99–108 (2016). <https://doi.org/10.1007/s10586-016-0535-3>
3. Amerini, I., Ballan, L., Caldelli, R., Bimbo, A.D., Serra, G.: A sift-based forensic method for copy-move attack detection and transformation recovery. *IEEE Trans. Inf. Forensics Secur.* **6**(3), 1099–1110 (2011). <https://doi.org/10.1109/TIFS.2011.2129512>
4. Nefian, A.V.: Georgia tech face database (1999). http://www.anefian.com/research/face_reco.htm. Accessed 21 July 2020
5. Banerjee, A., Das, N., Nasipuri, M.: Texture classification using deep neural network based on rotation invariant weber local descriptor. In: Santosh, K., Hangarge, M., Bevilacqua, V., Negi, A. (eds.) *Recent trends in image processing and pattern recognition*, pp. 277–292. Springer, Singapore (2017)
6. Bereta, M., Karczmarek, P., Pedrycz, W., Reformat, M.: Local descriptors in application to the aging problem in face recognition. *Pattern Recogn.* **46**(10), 2634–2646 (2013). <https://doi.org/10.1016/j.patcog.2013.03.010>
7. Bhatt, H.S., Bharadwaj, S., Singh, R., Vatsa, M.: On matching sketches with digital face images. In: 2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 1–7 (2010). <https://doi.org/10.1109/BTAS.2010.5634507>
8. Bhatt, H.S., Bharadwaj, S., Singh, R., Vatsa, M.: Memetically optimized MCWLD for matching sketches with digital face images. *IEEE Trans. Inf. Forensics Secur.* **7**(5), 1522–1535 (2012). <https://doi.org/10.1109/TIFS.2012.2204252>
9. Bolme, D.S., Draper, B.A., Beveridge, J.R.: Average of synthetic exact filters. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2105–2112 (2009). <https://doi.org/10.1109/CVPR.2009.5206701>
10. Bourdev, L., Brandt, J.: Robust object detection via soft cascade. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 2, pp. 236–243 (2005). <https://doi.org/10.1109/CVPR.2005.310>
11. Brodatz, P.: *Textures: A Photographic Album for Artists and Designers*. Dover Photography Collections. Dover Publications (1999). <https://books.google.co.in/books?id=XbrIJQAACAAJ>
12. Caputo, B., Hayman, E., Mallikarjuna, P.: Class-specific material categorisation. In: Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, vol. 2, pp. 1597–1604 (2005). <https://doi.org/10.1109/ICCV.2005.54>
13. Chen, J., Shan, S., He, C., Zhao, G., Pietikainen, M., Chen, X., Gao, W.: WLD: a robust local image descriptor. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(9), 1705–1720 (2010). <https://doi.org/10.1109/TPAMI.2009.155>
14. Chen, M., Dhingra, K., Wu, W., Yang, L., Sukthankar, R., Yang, J.: Pfid: pittsburgh fast-food image dataset. In: 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 289–292 (2009). <https://doi.org/10.1109/ICIP.2009.5413511>
15. Dawood, H., Dawood, H., Guo, P.: Combining the contrast information with WLD for texture classification. In: 2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE), vol. 3, pp. 203–207 (2012). <https://doi.org/10.1109/CSAE.2012.6272939>

16. Dawood, H., Dawood, H., Guo, P.: Texture image classification with improved weber local descriptor. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) *Artificial intelligence and soft computing*, pp. 684–692. Springer, Cham (2014)
17. Dong, J., Wang, W., Tan, T.: Casia image tampering detection evaluation database. In: 2013 IEEE China Summit and International Conference on Signal and Information Processing, pp. 422–426 (2013). <https://doi.org/10.1109/ChinaSIP.2013.6625374>
18. Downton, F.: Linear estimates with polynomial coefficients. *Biometrika* **53**(1/2), 129–141 (1966)
19. Gaber, T., Tharwat, A., Hassanien, A.E., Snaes, V.: Biometric cattle identification approach based on weber's local descriptor and adaboost classifier. *Comput. Electron. Agric.* **122**(C), 55–66 (2016). <https://doi.org/10.1016/j.compag.2015.12.022>
20. Galbally, J., Alonso-Fernandez, F., Fierrez, J., Ortega-Garcia, J.: A high performance fingerprint liveness detection method based on quality related features. *Future Gener. Comput. Syst.* **28**(1), 311–321 (2012). <https://doi.org/10.1016/j.future.2010.11.024>
21. Garcia, C., Delakis, M.: Convolutional face finder: a neural architecture for fast and robust face detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(11), 1408–1423 (2004). <https://doi.org/10.1109/TPAMI.2004.97>
22. Gong, D., Li, S., Xiang, Y.: Face recognition using the weber local descriptor. In: *The First Asian Conference on Pattern Recognition*, pp. 589–592 (2011). <https://doi.org/10.1109/ACPR.2011.6166675>
23. Gragnaniello, D., Poggi, G., Sansone, C., Verdoliva, L.: Fingerprint liveness detection based on weber local image descriptor. In: 2013 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications, pp. 46–50 (2013). <https://doi.org/10.1109/BIOBMS.2013.6656148>
24. Graham, D.B., Allinson, N.M.: *Characterising Virtual Eigensignatures for General Purpose Face Recognition*, pp. 446–456. Springer, Berlin (1998). https://doi.org/10.1007/978-3-642-72201-1_25
25. Hadid, A., Pietikainen, M., Ahonen, T.: A discriminative feature space for detecting and recognizing faces. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 2, pp. II–II (2004). <https://doi.org/10.1109/CVPR.2004.1315246>
26. Han, X., Chen, Y., Xu, G.: High-order statistics of weber local descriptors for image representation. *IEEE Trans. Cybern.* **45**(6), 1180–1193 (2015). <https://doi.org/10.1109/TCYB.2014.2346793>
27. Heath, M., Bowyer, K., Kopans, D., Kegelmeyer, P., Moore, R., Chang, K., Munishkumar, S.: Current status of the digital database for screening mammography. In: *Computational Imaging and Vision*, pp. 457–460. Springer, Netherlands (1998). https://doi.org/10.1007/978-94-011-5318-8_75
28. Hommel, S., Handmann, U.: AAM based continuous facial expression recognition for face image sequences. In: 2011 IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI), pp. 189–194 (2011). <https://doi.org/10.1109/CINTI.2011.6108497>
29. Huang, C., Ai, H., Yamashita, T., Lao, S., Kawade, M.: Incremental learning of boosted face detector. In: 2007 IEEE 11th International Conference on Computer Vision, pp. 1–8 (2007). <https://doi.org/10.1109/ICCV.2007.4408850>
30. Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst (2007)
31. Hussain, M., Khan, N.: Automatic mass detection in mammograms using multiscale spatial weber local descriptor. In: 2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP), pp. 288–291 (2012)
32. Hussain, M., Muhammad, G., Bebis, G.: Face recognition using multiscale and spatially enhanced weber law descriptor. In: 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems, pp. 85–89 (2012). <https://doi.org/10.1109/SITIS.2012.24>
33. Hussain, M., Muhammad, G., Saleh, S.Q., Mirza, A.M., Bebis, G.: Image forgery detection using multi-resolution weber local descriptors. *Eurocon* **2013**, 1570–1577 (2013). <https://doi.org/10.1109/EUROCON.2013.6625186>
34. Hussain, M., Qasem, S., Bebis, G., Muhammad, G., Aboalsamh, H., Mathkour, H.: Evaluation of image forgery detection using multi-scale weber local descriptors. *Int. J. Artif. Intell. Tools* **24**, 500 (2015). <https://doi.org/10.1142/s0218213015400163>
35. Hussain, M., Saleh, S.Q., Aboalsamh, H., Muhammad, G., Bebis, G.: Comparison between wld and lbp descriptors for non-intrusive image forgery detection. In: 2014 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) Proceedings, pp. 197–204 (2014). <https://doi.org/10.1109/INISTA.2014.6873618>
36. Jabid, T., Kabir, M.H., Chae, O.: Local directional pattern (LDP) for face recognition. In: 2010 Digest of Technical Papers International Conference on Consumer Electronics (ICCE), pp. 329–330 (2010). <https://doi.org/10.1109/ICCE.2010.5418801>
37. Kanade, T., Cohn, J.F., Tian, Y.: Comprehensive database for facial expression analysis. In: *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, pp. 46–53 (2000). <https://doi.org/10.1109/AFGR.2000.840611>
38. Kao, W.C., Hsu, M.C., Yang, Y.Y.: Local contrast enhancement and adaptive feature extraction for illumination-invariant face recognition. *Pattern Recogn.* **43**(5), 1736–1747 (2010). <https://doi.org/10.1016/j.patcog.2009.11.016>
39. Klare, B., Jain, A.K.: Sketch to photo matching : A feature-based approach. In: *Proceedings Volume 7667, Biometric Technology for Human Identification VII* (2010)
40. Klare, B., Li, Z., Jain, A.K.: Matching forensic sketches to mug shot photos. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(3), 639–646 (2011). <https://doi.org/10.1109/TPAMI.2010.180>
41. Kohli, N., Singh, R., Vatsa, M.: Self-similarity representation of weber faces for kinship classification. In: 2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 245–250 (2012). <https://doi.org/10.1109/BTAS.2012.6374584>
42. Krasnogor, N., Smith, J.: A tutorial for competent memetic algorithms: model, taxonomy, and design issues. *IEEE Trans. Evol. Comput.* **9**(5), 474–488 (2005). <https://doi.org/10.1109/TEVC.2005.850260>
43. Lan, R., Zhou, Y., Tang, Y.Y.: Quaternionic weber local descriptor of color images. *IEEE Trans. Circuits Syst. Video Technol.* **27**(2), 261–274 (2017). <https://doi.org/10.1109/TCSVT.2015.2492839>
44. Lazebnik, S., Schmid, C., Ponce, J.: A sparse texture representation using affine-invariant regions. In: 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., vol. 2, pp. II–II (2003). <https://doi.org/10.1109/CVPR.2003.1211486>
45. Li, J., Sang, N., Gao, C.: Log-Gabor weber descriptor for face recognition. In: Jawahar, C., Shan, S. (eds.) *Computer Vision—ACCV 2014 Workshops*, pp. 541–553. Springer, Cham (2015)
46. Li, S., Gong, D., Yuan, Y.: Face recognition using weber local descriptors. *Neurocomputing* **122**, 272–283 (2013). <https://doi.org/10.1016/j.neucom.2013.05.038>. Advances in cognitive and ubiquitous computing
47. Lin, Y.Y., Liu, T.L., Fuh, C.S.: Fast object detection with occlusions. In: Pajdla, T., Matas, J. (eds.) *Computer Vision—ECCV 2004*, pp. 402–413. Springer, Berlin (2004)
48. Liu, F., Tang, Z., Tang, J.: WLBP: Weber local binary pattern for local image description. *Neurocomputing* **120**, 325–335 (2013). <https://doi.org/10.1016/j.neucom.2012.06.061>

49. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004). <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
50. Lyons, M.J., Budynek, J., Akamatsu, S.: Automatic classification of single facial images. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**(12), 1357–1362 (1999). <https://doi.org/10.1109/34.817413>
51. M. Martinez, A., Benavente, R.: The AR face database. Technical Report 24 CVC Technical Report (1998)
52. Marasco, E., Sansone, C.: Combining perspiration- and morphology-based static features for fingerprint liveness detection. *Pattern Recogn. Lett.* **33**(9), 1148–1156 (2012). <https://doi.org/10.1016/j.patrec.2012.01.009>
53. Marcialis, G.L., Lewicke, A., Tan, B., Coli, P., Grimberg, D., Congiu, A., Tidu, A., Roli, F., Schuckers, S.: First international fingerprint liveness detection competition–livdet 2009. In: Proceedings of the 15th International Conference on Image Analysis and Processing, ICIAP '09, pp. 12–23. Springer-Verlag, Berlin (2009). https://doi.org/10.1007/978-3-642-04146-4_4
54. Muhammad, G., Hussain, M., Alenezzy, F., Bebis, G., Mirza, A.M., Aboalsamh, H.: Race recognition from face images using weber local descriptor. In: 2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP), pp. 421–424 (2012)
55. Ng, T.T., Chang, S.F.: A data set of authentic and spliced image blocks. Technical report Columbia University (2004)
56. Ojala, T., Maenpaa, T., Pietikainen, M., Viertola, J., Kyllonen, J., Huovinen, S.: Outex—new framework for empirical evaluation of texture analysis algorithms. In: Object Recognition Supported by User Interaction for Service Robots, vol. 1, pp. 701–706 (2002). <https://doi.org/10.1109/ICPR.2002.1044854>
57. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002). <https://doi.org/10.1109/TPAMI.2002.1017623>
58. Ojansivu, V., Heikkilä, J.: Blur insensitive texture classification using local phase quantization. In: Proceedings of the 3rd International Conference on Image and Signal Processing, ICISP '08, pp. 236–243. Springer, Berlin (2008). https://doi.org/10.1007/978-3-540-69905-7_27
59. Ojansivu, V., Rahtu, E., Heikkilä, J.: Rotation invariant local phase quantization for blur insensitive texture analysis. In: 2008 19th International Conference on Pattern Recognition, pp. 1–4 (2008). <https://doi.org/10.1109/ICPR.2008.4761377>
60. ORL Dataset: Olivetti research laboratory (ORL) face database (1992–1994). <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>
61. Pal, A., Das, N., Sarkar, S., Gangopadhyay, D., Nasipuri, M.: A new rotation invariant weber local descriptor for recognition of skin diseases. In: Maji, P., Ghosh, A., Murty, M.N., Ghosh, K., Pal, S.K. (eds.) *Pattern Recognit. Mach. Intell.*, pp. 355–360. Springer, Berlin (2013)
62. Perner, P.: Image analysis and classification of hep-2 cells in fluorescent images. In: Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No.98EX170), vol. 2, pp. 1677–1679 (1998). <https://doi.org/10.1109/ICPR.1998.712043>
63. Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.J.: The FERET evaluation methodology for face-recognition algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(10), 1090–1104 (2000). <https://doi.org/10.1109/34.879790>
64. Roth, V., Steinhage, V.: Nonlinear discriminant analysis using kernel functions. In: *Advances in Neural Information Processing Systems*, pp. 568–574. MIT Press (1999)
65. Rowley, H.A., Baluja, S., Kanade, T.: Neural network-based face detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(1), 23–38 (1998). <https://doi.org/10.1109/34.655647>
66. Saadat, S., Moghaddam, M.E., Mohammadi, M.: A new approach for copy-move detection based on improved weber local descriptor. *J. Forensic Sci.* **60**(6), 1451–1460 (2015). <https://doi.org/10.1111/1556-4029.12853>
67. Schneiderman, H., Kanade, T.: A statistical method for 3D object detection applied to faces and cars. In: Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662), vol. 1, pp. 746–751 (2000). <https://doi.org/10.1109/CVPR.2000.855895>
68. Shekar, B.H., Smitha, M.L.: Text localization in video using multiscale weber's local descriptor. In: 2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES) (2015). <https://doi.org/10.1109/spices.2015.7091559>
69. Sun, S., Zhao, L., Yang, S.: Gabor weber local descriptor for bovine iris recognition. *Math. Probl. Eng.* **10**, 15 (2013). <https://doi.org/10.1155/2013/920597>
70. Sun, Y., Todorovic, S., Goodison, S.: Local-learning-based feature selection for high-dimensional data analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(9), 1610–1626 (2010). <https://doi.org/10.1109/TPAMI.2009.190>
71. Tharwat, A., Hemedan, A.A., Hassanien, A.E., Gabel, T.: A biometric-based model for fish species classification. *Fish. Res.* **204**, 324–336 (2018). <https://doi.org/10.1016/j.fishres.2018.03.008>
72. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, vol. 1, pp. I–I (2001). <https://doi.org/10.1109/CVPR.2001.990517>
73. Viola, P., Jones, M.J.: Robust real-time face detection. *Int. J. Comput. Vis.* **57**(2), 137–154 (2004). <https://doi.org/10.1023/B:VISI.0000013087.49260.fb>
74. Walia, E., Suneja, A.: A robust watermark authentication technique based on weber's descriptor. *SIVIP* **8**(5), 859–872 (2014). <https://doi.org/10.1007/s11760-012-0312-6>
75. Wang, W., Dong, J., Tan, T.: Image tampering detection based on stationary distribution of Markov chain. In: 2010 IEEE International Conference on Image Processing, pp. 2101–2104 (2010). <https://doi.org/10.1109/ICIP.2010.5652660>
76. Wang, X., Jin, C., Liu, W., Hu, M., Xu, L., Ren, F.: Feature fusion of HOG and WLD for facial expression recognition. In: Proceedings of the 2013 IEEE/SICE International Symposium on System Integration, pp. 227–232 (2013). <https://doi.org/10.1109/SII.2013.6776664>
77. Wang, X., Tang, X.: Face photo-sketch synthesis and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(11), 1955–1967 (2009). <https://doi.org/10.1109/TPAMI.2008.222>
78. Xia, S., Shao, M., Fu, Y.: Kinship verification through transfer learning. In: Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence—Volume Volume Three, IJCAI'11, pp. 2539–2544. AAAI Press (2011). <https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-422>
79. Xie, Z., Liu, G.: Weighted local binary pattern infrared face recognition based on weber's law. In: 2011 Sixth International Conference on Image and Graphics, pp. 429–433 (2011). <https://doi.org/10.1109/ICIG.2011.51>
80. Xie, Z., Liu, G., Fang, Z.: Face recognition based on combination of human perception and local binary pattern. In: Zhang, Y., Zhou, Z.H., Zhang, C., Li, Y. (eds.) *Intell. Sci. Intell. Data Eng.*, pp. 365–373. Springer, Berlin (2012)
81. Xin, Y., Liao, S., Pawlak, M.: Circularly orthogonal moments for geometrically robust image watermarking. *Pattern Recogn.* **40**(12), 3740–3752 (2007). <https://doi.org/10.1016/j.patcog.2007.05.004>
82. Yambay, D., Ghiani, L., Denti, P., Marcialis, G.L., Roli, F., Schuckers, S.: Livdet 2011—fingerprint liveness detection competition 2011. In: 2012 5th IAPR International Conference on Biomet-

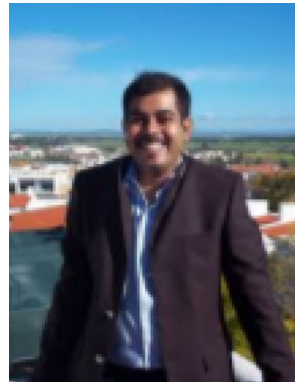
- rics (ICB), pp. 208–215 (2012). <https://doi.org/10.1109/ICB.2012.6199810>
83. Yu, K., Wang, Z., Zhuo, L., Wang, J., Chi, Z., Feng, D.: Learning realistic facial expressions from web images. *Pattern Recogn.* **46**(8), 2144–2155 (2013). <https://doi.org/10.1016/j.patcog.2013.01.032>
84. Zhang, W., Shan, S., Gao, W., Chen, X., Zhang, H.: Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition. In: Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, vol. 1, pp. 786–791 (2005). <https://doi.org/10.1109/ICCV.2005.147>
85. Zhang, Z., Wang, L., Zhu, Q., Chen, S.K., Chen, Y.: Pose-invariant face recognition using facial landmarks and weber local descriptor. *Knowl. Based Syst.* **84**(C), 78–88 (2015). <https://doi.org/10.1016/j.knosys.2015.04.003>
86. Zhao, X., Li, J., Li, S., Wang, S.: Detecting digital image splicing in chroma spaces. In: Kim, H.J., Shi, Y.Q., Barni, M. (eds.) *Digital Watermarking*, pp. 12–22. Springer, Berlin (2011)
87. Zhou, X., Hu, J., Lu, J., Shang, Y., Guan, Y.: Kinship verification from facial images under uncontrolled conditions. In: Proceedings of the 19th ACM International Conference on Multimedia, MM '11, pp. 953–956. ACM, New York (2011). <https://doi.org/10.1145/2072298.2071911>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Arnab Banerjee received his B.Tech. degree in Computer Science and Engineering from Bengal Engineering College under West Bengal University of Technology in 2011. He received his M.C.S.E. degree from Jadavpur University in 2015. He is registered in the PhD program in the Department of Computer Science and Engineering at Jadavpur University. He is currently acting as Lecturer in Computer Science and Technology Department at Dr B.C. Roy Polytechnic, Durgapur.

His research interest include are deep learning, medical image processing, computer vision and object identification.



140 peer-reviewed research articles. He serves as an associate editor of the *Sadhana Academy Proceedings in Engineering Sciences*.

Nibaran Das received his B.Tech. degree in Computer Science and Technology from Kalyani Govt. Engineering College under Kalyani University in 2003. He received his M.C.S.E. and Ph.D. (Engg.) degree from Jadavpur University in 2005 and 2012, respectively. He joined Jadavpur University as a faculty member in 2006. His areas of current research interest are OCR of handwritten text, optimization techniques, deep learning and medical image processing. He published more than



Dr Santosh worked as a research fellow at the U.S. National Library of Medicine (NLM), National Institutes of Health (NIH). He worked as a postdoctoral research scientist at the LORIA research centre, Université de Lorraine in direct collaboration with industrial partner ITE-SOFT, France. Dr Santosh has demonstrated expertise in artificial intelligence, machine learning, pattern recognition, computer vision, image processing, data mining and big data with various application domains, such as healthcare informatics and medical imaging, document imaging, biometrics, forensics, speech analysis and Internet of Things. As of now (August 2020), he published more than 180 research works that include journal articles (70), conference proceedings (100) and book chapters (11). He authored two books and edited five books, 14 journal issues and six conference proceedings. His research projects are funded by multiple agencies, such as SDCRGP, Department of Education (DOE) and National Science Foundation (NSF). Dr Santosh is the proud recipient of the President's Research Excellence Award (USD, 2019) and an award from the US Department of Health & Human Services (2014).

K.C. Santosh is the Chair of the Department of Computer Science (CS) at the University of South Dakota (USD). Dr Santosh served School of Computing and IT, Taylor's University, as a Visiting Associate Professor, for one year (2019/2020). Before joining USD, Dr Santosh worked as a