


# Evaluating color and texture features for forgery localization from illuminant maps

Divya S. Vidyadharan<sup>1,2</sup>  · Sabu M. Thampi<sup>3</sup>

Received: 31 March 2017 / Revised: 29 November 2017 / Accepted: 21 December 2017 /  
Published online: 4 January 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

**Abstract** Images are widely accepted as a record of events even when images are prone to easy manipulations. It is difficult to identify image alterations by the human visual system. Once an image is identified as forged, the next step is to locate forged regions. Recently, distribution of scene illumination across an image has been analyzed to detect forged images and to locate forged image regions. In this paper, we investigate the problem of locating spliced image region based on illumination inconsistency. We investigated the discriminative power of a number of color and texture descriptors in locating spliced image regions. During digital crime investigations, often it is required to detect the spliced face in a group photo. Here, we have selected forged images containing human facial regions where the regions to be compared are of similar object material, human skin regions. We evaluated various color, texture, and combined color-texture descriptors in an unsupervised manner by comparing the distance between the feature vectors to identify the inconsistent image region. We also investigated the performance of different histogram similarity measures including heuristic histogram distance measures, non-parametric test statistics, information theoretic divergences, and cross-bin measures. Experiments show that the Local Phase Quantization (LPQ) descriptor performs best in identifying the spliced image region from the illuminant map.

**Keywords** Image forgery localization · Illumination inconsistency · Color descriptor · Texture descriptor · Combined color-texture descriptor · Local phase quantization

---

✉ Divya S. Vidyadharan  
divya.s.vidyadharan@ieee.org  
Sabu M. Thampi  
smthampi@ieee.org

<sup>1</sup> College of Engineering-Trivandrum, Thiruvananthapuram, India

<sup>2</sup> LBS Centre for Science and Technology, University of Kerala, Thiruvananthapuram, Kerala, India

<sup>3</sup> Indian Institute of Information Technology and Management-Kerala, Thiruvananthapuram, India

## 1 Introduction

Image contents are considered as an authentic representation of events. Plenty of images are encountered in our day-to-day life as more and more hand-held devices are equipped with image capturing and image editing tools in addition to traditional digital still cameras. This substantiated with the handiness of a lot of image processing software makes it easier to manipulate images. This has attracted the attention of researchers and has led to the development of a lot of image forensic techniques that reveal image alterations.

Digital crimes involving forged images containing human facial regions are increasing recently. Once an image is proved as a forged image, the next step is to locate the forged facial region. The techniques that locate forged regions are termed as forgery localization techniques. For example, in a forged group photo, the task is to locate the spliced facial region among various facial regions in the image. Usually, forgery localization techniques detect the presence of irregularities in forged regions. Whenever an image content is altered by adding/modifying an image region, it changes the original patterns of scene information such as the color of scene illumination, thus creating an inconsistent region.

In this work, we analyze the inconsistency in the scene illumination across different image regions for detecting forged region. For this, we rely upon the scene illumination representation proposed by Riess and Angelopoulou [33]. The color of scene illumination is recorded in the pixels and if the scene were illuminated by multiple light sources, then different image regions will exhibit different illuminant color, introducing a pattern. The properties of this pattern will be different at a spliced (copy-pasted) image region when compared to the untouched regions within the image. Since the surface reflectance properties depend on the material of the object, only similar object materials can be compared for checking the inconsistency. Therefore, the illumination pattern and color, in the facial skin regions are analyzed to reveal the spliced facial region.

Forgery detection and localization in spliced images by exploiting inconsistencies in illumination is addressed earlier in [6, 8, 11, 13, 15, 16, 27, 33, 46, 47, 51, 52]. Gholap and Bora developed a technique based on the difference in the illuminant color observed from different image regions [16]. Here, an image is declared as authentic if different image regions report the same illuminant color, otherwise declared as forged. Cao et al. developed a forgery detection method that considers color histograms and illuminant color differences estimated from foreground and background image regions [6]. Wu and Fang proposed another forgery localization method [51], where the image is divided into different overlapping blocks and one of the blocks is selected as the reference block. The image block is declared as spliced if the illuminant color difference between the corresponding block and the reference block is greater than a threshold.

Fan et al. devised a forgery localization technique that automatically selects reference illuminant color [13]. Here, the image is divided into different vertical and horizontal regions, a region is declared as spliced if the difference between the estimated illuminant color and the reference illuminant color is greater than a threshold. Illuminant estimation is carried out using 5 different algorithms, and for each algorithm the inconsistent regions are identified. Finally, the intersection of all inconsistent regions is declared as the spliced region. Vidyadharan and Thampi proposed a forgery localization technique by applying histogram distance measures on brightness distribution obtained from facial skin regions [46].

Among the forensic techniques that consider illumination inconsistency, certain works consider forgery localization in spliced images by analyzing human facial regions [8, 11,

15, 27, 33, 47]. Riess and Angelopoulou, in their pioneering work that forensically analyze illuminant distribution across an image, found that a manual examination of illuminant representation could reveal forged image regions [33]. This finding is further explored by Carvalho et al., by utilizing edge and texture features generated from the facial regions extracted from the illuminant maps [11]. They automated forgery detection by developing a machine learning technique that classifies forged and authentic images based on the discrepancies in the texture and edge features from the illuminant representation. Later, Carvalho et al. improved forgery detection and considered forgery localization as well, by considering color, and shape features, in addition to texture features with the help of a classifier ensemble [8].

Meanwhile, there are certain techniques that followed non-machine learning approaches for forgery detection and localization [15, 27, 47] from images containing spliced human facial regions. Francis et al. developed a forgery localization technique based on differences in illuminant color estimated from the nose tips of different persons in a group photo [15]. Vidyadharan and Thampi proposed a technique where a Principal Component Analysis (PCA) is carried out on facial regions extracted from the illuminant maps for locating spliced facial region [47]. Mazumdar and Bora devised a Dichromatic Plane Histogram (DPH) based technique to detect forged images [27]. The DPH is considered as an illumination signature for a face, and DPHs obtained from facial regions captured at similar illumination will be similar. The similarity between DPHs are examined using correlation measure. If the correlation measure of any of the face pairs in an image is lower than a threshold the image is considered forged.

In the proposed work, we investigate the discriminative power of different color, texture and combined color-texture features in locating spliced facial region from the illumination representation of an image. This evaluation is inspired by Van de Sande's evaluation of color descriptors for scene recognition [36]. The attempt to consider combined color and texture features is motivated by the work of Khan et al. where a compact texture and color descriptor is used for texture classification [21].

The main contributions of the work are,

- An evaluation of 5 texture descriptors, 5 color descriptors, 3 descriptors that combined color and texture features, 5 color moments and histogram descriptors, and 5 color-shape descriptors for forgery localization from illuminant maps.
- Evaluation of the performance of various categories of histogram distance measures.
- A comparison exhibiting a better detection accuracy for forgery localization based on texture descriptors than existing non-machine learning approaches.

Even though we evaluated the discriminative power of various descriptors on forgery localization, the results of the study can be applied to other image processing domains that consider the similarity of segmented images. Also, the comparison of features to detect similarity in digital data can be used in different applications [34, 41, 53, 54].

The rest of the paper is organized as follows. In Section 2, we discuss the representation of scene illumination and illuminant maps briefly. The texture descriptors considered in the work are briefed in Section 3. The color descriptors used for locating spliced face is discussed in Section 4. The combination features, such as color-texture descriptors are discussed in Section 5. Additional features considered are mentioned in Section 6. The evaluation framework illustrating how various descriptors are generated from the illumination representation and how the descriptor representing the spliced face is located is mentioned in Section 7. The details of experiments including the experimental setup, distance metrics used for comparing the feature descriptors, performance evaluation criteria, and the

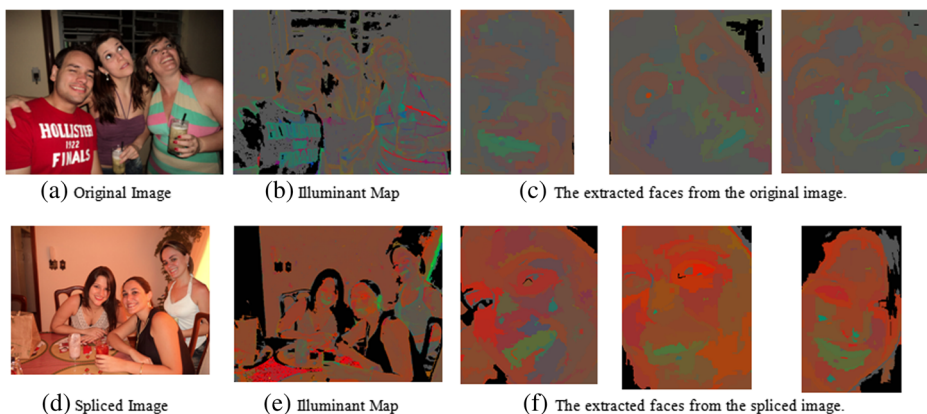
experimental analysis conducted are detailed in Section 8. Finally, the conclusion of the work is given in Section 9.

## 2 Representing scene illumination

When an image is captured by the camera sensor, the illumination present in the environment is recorded in the pixels. The scene illumination information present in an authentic unaltered image will be consistent whereas, in a forged image, the scene illumination will be inconsistent at the forged region. For studying the change in the pattern and color of illumination, we used illuminant maps - the scene illumination representation proposed by Riess and Angelopoulou [33].

For generating the illuminant map, an image is first segmented into regions of similar color. Each region is further divided into small patches. From each patch, the color of illumination, termed as illuminant color is estimated. Finally, a majority voted illuminant color is selected as the illuminant color of the region. Since illuminant color is estimated locally at a region this representation is capable of representing the multi-illuminant environment [33]. In [11], Carvalho et al. have used two variants of illuminant maps, such as Inverse Intensity Chromaticity (IIC) Map and Generalized Grey World (GGW) map. In GGW, the illuminant color in an image patch within a region is computed using the statistical approach followed in Generalized Grey Edge Framework proposed by Van de Weijer in [44]. In IIC map, the illuminant color in the image patch is computed using the Inverse Intensity Chromaticity space proposed by Tan et al. in [42] based on the physics-based Dichromatic Reflection Model [14]. Figure 1 shows an example of original unmodified image, a spliced image, corresponding illuminant maps, and facial regions extracted from the respective illuminant maps.

In our work, we evaluated the discriminative power of color, texture, and combined color-texture features in IIC and GGW illuminant maps. The illuminant maps show a texture pattern that varies in color, depending on the scene illumination intensity and the direction of light. Thus, the texture properties and color properties at a spliced region may differ from the rest of the untouched regions in the image. Here, we explore how the difference in



**Fig. 1** An example of the original image and spliced image, corresponding illuminant maps and extracted facial regions. Both the original and spliced image are taken from tifs-database [11]

the texture and color properties among different facial regions within an image can locate forged regions.

### 3 Texture descriptors

In this work, we evaluate the discriminative power of 5 popular texture descriptors.

*Local Binary Pattern (LBP)*. Ojala et al. proposed a local texture descriptor known as Local Binary Pattern capable of capturing texture patterns [30, 40]. Here, the pixels in a local neighborhood are compared with the central pixel. If the neighboring pixel is greater than the central pixel, that pixel is coded with a one and otherwise a zero. Finally, these binary values are concatenated to get the LBP code. All these LBP codes are counted and the distribution is represented as the LBP histogram (256-bin).

*Completed Local Binary Pattern (CLBP)*. Guo et al. devised a completed LBP descriptor that makes use of the Local Difference Sign-Magnitude Transform (LDSMT) [17]. In traditional LBP, texture pattern around the neighborhood of a pixel is represented by the sign of the difference between current and the central pixel. In complete LBP, both the magnitude of the difference and the central pixel value are also considered. The sign and magnitude CLBP representations, when joined, is represented as CLBP\_S/M (59049-bin).

*Local Phase Quantization (LPQ)*. T Ojansivu and Heikkilä [31] proposed blur invariant, local texture descriptor that performed better on non-blurred images as well. The phase of low-frequency components in the Fourier domain are used to represent texture. The phase values of four low-frequency components are decorrelated and quantized to get the LPQ codeword. Finally, the codewords are represented as LPQ histogram (256-bin).

*Binarized Statistical Image Features (BSIF)*. BSIF is a 256-bin dense descriptor where a binary code for a pixel is generated by convolving a neighborhood region of pixels with filters that are learned by prior training by independent component analysis [20]. Training is performed using image patches randomly sampled from a small set of natural images. Thus, the filters capture the statistical properties of natural images.

*Binary Gabor Pattern (BGP)*. Zhang et al. proposed a texture feature using Gabor filters [55]. Here, an image is convolved with even symmetric and odd symmetric Gabor filters at three different resolutions to obtain a 216 bin histogram termed as the Binary Gabor Pattern.

### 4 Color descriptors

Inconsistencies in illumination are visible as noticeable color changes in the illuminant maps of forged images. Thus, to study the discriminative power of color for forgery localization, we considered various color descriptors, including color name descriptors, color histograms and color moments.

*Color Names*. Color name descriptors represent colors based on human perception of color as linguistic terms such as ‘Red’ and ‘Blue’. Benavente et al. devised an 11-bin color name descriptor, Automatic Color Names (ACN), based on the parametric model with fuzzy set membership for different colors [2]. Van de Weijer et al. devised an improved version of linguistic color names, termed as Color Names (CN), by learning from real-world images [50].

*Discriminative Color Descriptors (DCD)*. Khan et al. proposed a discriminative color descriptors that represent the color features with clusters grouped based on their discriminative power in classifying images [22]. The color descriptors with 11, 25, or 50 clusters are available.

In addition to the above color descriptors, we have studied the performance of a few more color histograms and color moment descriptors mentioned in Table 1.

## 5 Combined color-texture descriptors

Although color and textures can be represented well as separate descriptors, certain attempts have been made to combine the features for computer vision tasks such as texture classification [21]. Currently, color and texture features are combined in either of two ways described as follows. In the first approach, color and texture features are computed separately and the final descriptors are combined together by concatenating the feature vectors. In the second approach, known as the joint approach, the texture descriptor is computed in different color channels separately and all these texture descriptors are concatenated later [1].

*Color LPQ.* Pedone and Heikkilä proposed an extension to LPQ that considers color features [32]. Here, the 1280-bin descriptor is computed from a multi-vector representation of color.

*Color Texton Descriptors.* Alvarez and Vanrell extended the traditional texton theoretic approach considering the color and shape of image blobs [1]. The basic image blob representation is as follows. A perceptual blob is defined as a region with similar color identified from opponent color space. Later, shape attributes and color attributes of blobs are extracted. Shape attributes include width, length, and orientation. Color attributes include the intensity, *rg* and *by* components extracted using the median of color information of all pixels belonging to the blob. Thus, there is a total of six attributes. Each attribute value is then quantized in *m* intervals resulting in  $6 \times m$  terms. Finally, the attributes are described by concatenating the probability distribution of all six attributes. This representation is further improved by adding the perceptual relationships between attributes and the co-occurrence of shape and color attributes of blobs. Perceptual relationships between shape attributes are added by transforming the attributes into shape space. The two axes in the shape space represent the width and length and the third axis represents the angle. Three quantization models such

**Table 1** Details of color histograms and color moments considered in the evaluation

Descriptor	Dimension	Based on
Opponent histogram [21, 36]	36	combined histograms of $O_1, O_2, O_3$ channels in opponent color space
Hue histogram [36, 49]	36	by weighing the hue samples by saturation
rgb histogram [21, 36]	45	combined histograms of r,g,b channels.
rg histogram [21, 36]	45	normalized RGB
Transformed [21, 36]	36	normalizing each channel of RGB histogram
Color moments [21, 36]	36	generalized color moments
Color moment invariants [29]	24	generalized color moments

as Cartesian, Cylindrical and Circular models are used for shape space representation. The color attributes are represented as HSI-Carron [7] and HSV-Smith spaces [38]. These color and shape attributes can be represented either as combined at the blob level or separate at the image level resulting in the following texton descriptors.

*Co-joint Texton Descriptor (JTD)*: This descriptor represents the color and texture attributes as a joint probability distribution. Thus the co-occurrence of color and shape attributes are taken care of.

*Semi-joint Texton Descriptor (STD)*: This descriptor represents color and texture by concatenating the probability distribution of color and texture.

## 6 Additional features

For analyzing the differences in underlying pixel statistics, we have used two more features such as edge features and Grey Level Run Length features (GLRLM) [43]. Edge features are represented by Histogram of Oriented Gradients (HOG) proposed by Dalal and Triggs [10]. HOG and GLRLM features are 81-bin dimensional and 44-bin dimensional respectively.

We also considered features based on Scale Invariant Feature Transform (SIFT). The SIFT, proposed by Lowe [25] represents the local features representing regions around identified keypoints within an image. Since keypoint based feature extraction is performed, SIFT represents the spatial information of image regions. Here, different variants of SIFT such as HSV-SIFT [4], Opponent-SIFT [36], C-SIFT [5], rg-SIFT [36] and rgb-SIFT [36] as discussed by van de Sande et al. in [36], are considered. Details are given in Table 2.

## 7 Evaluation framework

We used the evaluation framework shown in Fig. 2 for evaluating the forgery localization capability of different feature descriptors. At first, the given image is represented as an illuminant map showing the variation in the illumination pattern across the image. The facial regions are extracted from this illuminant map by specifying the bounding box around each face. Then, the feature descriptor to be evaluated is generated from each face. Finally, the distance between the descriptors are compared among themselves, to identify the descriptor

**Table 2** Details of color-shape descriptors considered for the evaluation

Descriptor	Dimension	Based on
SIFT [25]	128	local features around keypoints
Hue-SIFT [36]	165	concatenation of Hue histogram with SIFT feature vector
HSV-SIFT [4]	384	SIFT descriptors from 3 channels of HSV color space
Opponent SIFT [36]	384	SIFT descriptors from 3 channels of Opponent color space
rg-SIFT [36]	384	SIFT descriptors from the r,g components of normalized RGB color space
C-SIFT [5]	384	SIFT descriptors based on C-invariant
rgb-SIFT [36]	384	SIFT descriptors computed on R,G,B channels

that represents the spliced face. The descriptors representing faces captured at same illuminant environment will be similar compared to the descriptor representing spliced face. Therefore, the distance between descriptors representing authentic faces will be less compared to the descriptor representing spliced face. Accordingly, the spliced face is located. We evaluated different histogram distance measures to identify the most dissimilar feature descriptor. The steps involved in evaluating a feature descriptor with  $M$  number of distance measures is given in Algorithm 1.

---

**Algorithm 1** Locating spliced face from an image.

---

**Input** : Image containing  $N$  facial regions.

$M$  number of different distance measures.

**Output** : Identified spliced face with each distance measure.

```

1: Generate illuminant representation of the input image.
2: Extract the facial regions.
3: for  $i=1$  to  $N$  do
4:   Generate the feature descriptor,  $FD_i$ .
5: end for
6: for  $distMeasure=1$  to  $M$  do
7:   for  $i=1$  to  $N$  do
8:     for  $j=i+1$  to  $N$  do
9:       if  $i >= j$  then
10:         $distMatrix(i,j) = ComputeDistance(FD_i, FD_j, distMeasure)$ 
11:       end if
12:     end for
13:   end for
14:   for  $i=1$  to  $N$  do
15:     $dist(i) = max(distMatrix(i,:))$ 
16:   end for
17:   Detect the majority voted most distant face as the spliced face.
18: end for

```

---

## 8 Experiments and results

### 8.1 Datasets

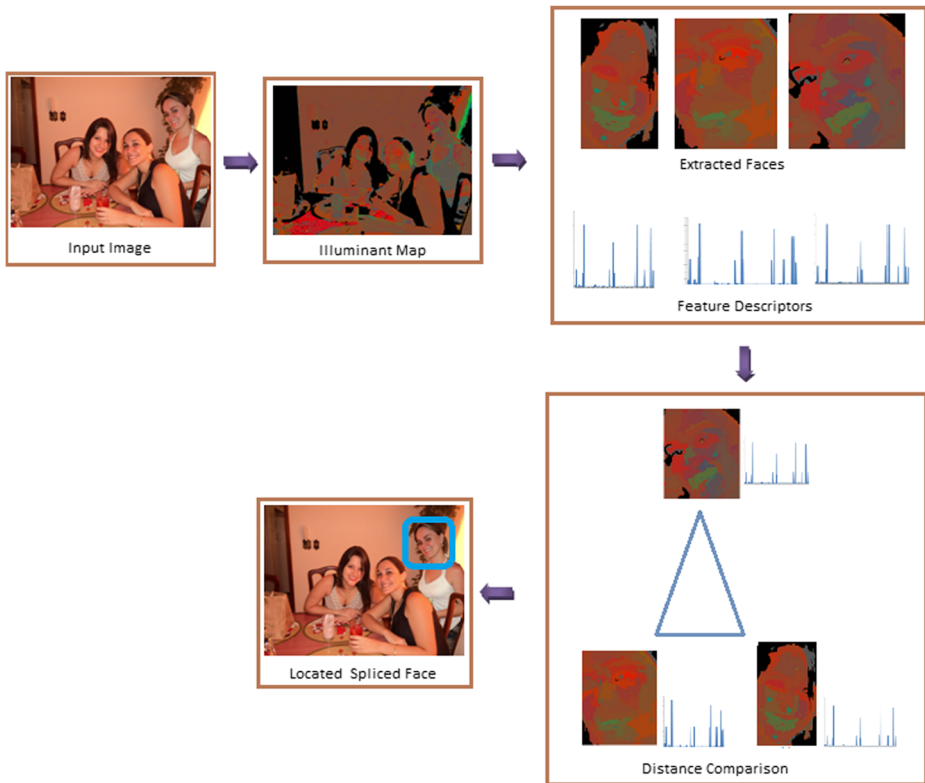
For experimental evaluation we used spliced images from three datasets, such as i) DSO-I ii) SwapMe and iii) FaceSwap. The DSO-I dataset is taken from tifs-database<sup>1</sup> [11]. The tifs-database contains 100 spliced images containing human facial regions, saved in Portable Network Graphics (PNG) with a resolution of 2,048 x 1536 pixels. We used a subset of 55 spliced images that contain more than two facial regions for evaluating the forgery localization capability of various color and texture descriptors.

The SwapMe and FaceSwap datasets contains spliced images created by exchanging a source facial region with a destination facial regions [56]. For our experiments, we selected

---

<sup>1</sup><http://ic.unicamp.br/~rocha/pub/downloads/2014-tiago-carvalho-thesis/tifs-database.zip>





**Fig. 2** The framework illustrating the steps involved in the evaluation process. For clarity, scaled version of feature descriptors are shown. The input image shown is taken from tifs-database [11]

a subset of 55 spliced images from the Swapme dataset. From the FaceSwap dataset, we selected 33 images that were not present in Swapme. Along with FaceSwap images, we combined our own set of 7 spliced images to create a Combined FaceSwap dataset. All the selected spliced images in both Swapme and Combined FaceSwap contained three or more facial regions.

We used two variants of illuminant maps such as GGW and IIC. In Sections 8.2–8.13, we present the feature extraction process, performance evaluation criteria, and the evaluation of various descriptors for forgery localization from illuminant maps.

## 8.2 Feature extraction

For illumination representation, the two variants of illuminant maps such as IIC and GGW maps were generated using the software<sup>2</sup> [33]. From the illuminant map, all the facial regions are extracted. Then, feature descriptors are computed for each facial region. The similarity among the feature vectors is measured using different histogram distance measures.

<sup>2</sup><http://www5.cs.fau.de/research/areas/computer-vision/image-forensics/scene-illumination-as-an-indicator-of-image-manipulation/>

### 8.3 Distance measures considered

Feature descriptors extracted from the facial regions in the illuminant maps can be considered as a distribution. As studied in the work of Meshgi and Ishii [28], we compared feature descriptors using various categories of histogram distance measures such as, heuristic distance measures, non-parametric test statistics, information theoretic divergences, and cross-bin distance measures.

From heuristic distance measures, we considered the L2 distance and the Pearson Correlation Coefficient [3]. From the non-parametric test statistics, we used the Kolmogorov-Smirnov distance (KS) [26], Cramer-von Mises Statistics (CM) [12], Chi-square (CS) statistics [23, 39], and Bhattacharya Distance (BD) [19]. The Kullback-Leibler divergence (KL) measure, is based on information theoretic divergences [23]. The Diffusion Distance (DF) [24] and the Earth Mover's Distance (EMD) consider cross-bin information capable of capturing the perceptual similarity of images [23, 35].

### 8.4 Performance evaluation criteria

In forgery localization, the objective is to locate forged image region within spliced images. Hence, the performance of a method can be measured by the detection rate of detecting forged facial regions within spliced images. Here, the performance is evaluated using the following performance metrics,

$$\text{Sensitivity or Recall or TPR} = \frac{\text{Faces}_{\text{Located}}}{\text{Images}_{\text{Spliced}}} \quad (1)$$

where  $\text{Faces}_{\text{Located}}$  = No. of spliced faces located correctly,

$\text{Images}_{\text{Spliced}}$  = Total no. of spliced faces, and TPR is the True Positive Rate.

$$\text{Specificity or TNR} = \frac{\text{AuthenticFaces}_{\text{Detected}}}{\text{Faces}_{\text{Authentic}}} \quad (2)$$

where  $\text{AuthenticFaces}_{\text{Detected}}$  = No. of authentic faces detected correctly, and

$\text{Faces}_{\text{Authentic}}$  = Total no. of authentic faces.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$F - \text{Score} = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

### 8.5 Experiment 1: evaluating texture descriptors

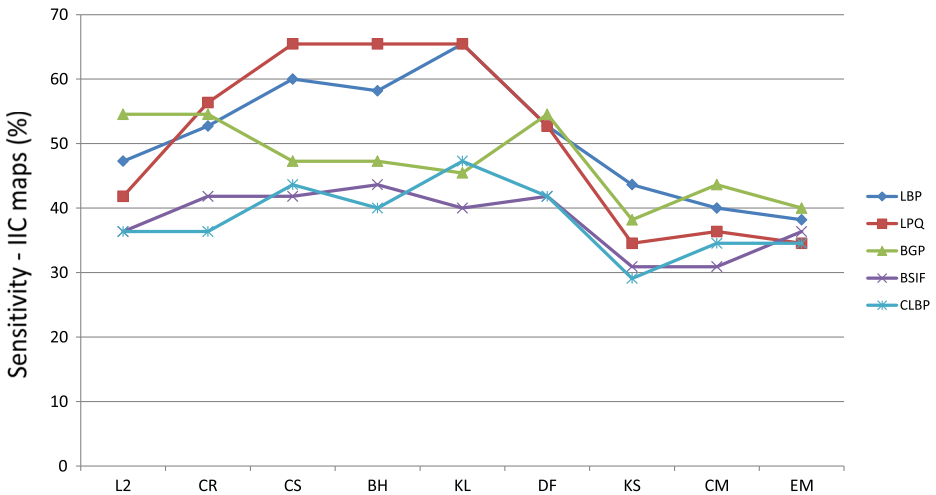
LBP and LPQ descriptors are obtained using the software provided by the respective authors<sup>3</sup>. The BSIF<sup>4</sup> and BGP<sup>5</sup> features are estimated using the source code provided by the corresponding authors. The CLBP features are generated by using the source code provided by Guo et al.<sup>6</sup> [17].

<sup>3</sup><http://www.cse.oulu.fi/CMV/Downloads/LBP Matlab>

<sup>4</sup><http://www.ee.oulu.fi/~jkannala/bsif/bsif.html>

<sup>5</sup><http://sse.tongji.edu.cn/linzhang/BGP/BGP.htm>

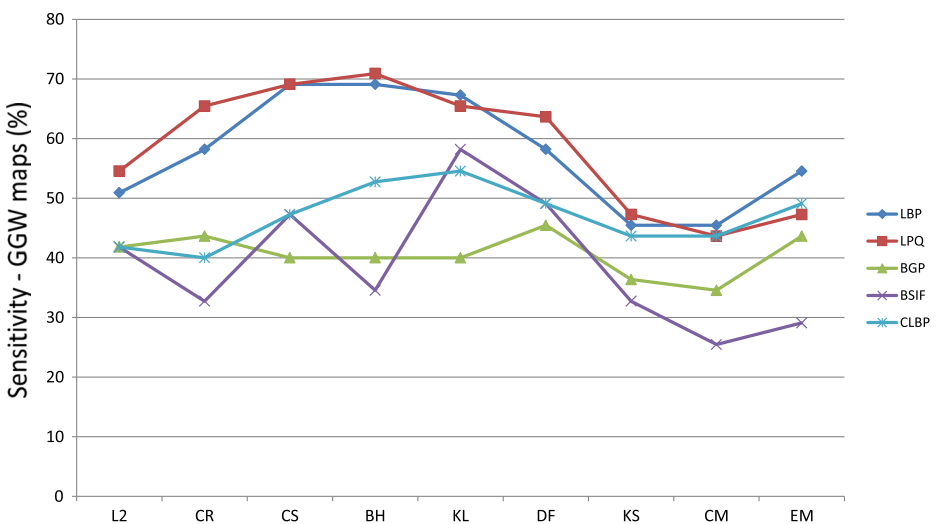
<sup>6</sup><http://www.comp.polyu.edu.hk/~cslzhang/code/CLBP.rar>.



**Fig. 3** Sensitivity obtained for texture descriptors from IIC maps using different distance measures

First, we considered texture descriptors and the sensitivity obtained from IIC and GGW illuminant maps are shown in Figs. 3 and 4 respectively. It is interesting to note that, the distance measure providing the best sensitivity is different for different texture features. The reason is that the nature of the feature vectors and the discriminative capability of different texture descriptors varies. Among the various distance measures, CS, BH, and KL yielded the best results for LBP and LPQ descriptor in IIC map. For BGP, the distance measures L2, and DF provided the best results. BSIF, and CLBP showed highest sensitivity with the distance measures BH and, KL respectively.

In GGW map, the distance measures-CS and BH yielded best result for LBP and LPQ (see Fig. 4). BGP showed good performance with CR, DF, EM distance measures and BSIF



**Fig. 4** Sensitivity obtained for texture descriptors from GGW maps using different distance measures

**Table 3** Evaluation of texture descriptors on different datasets using IIC maps

Texture Descriptor	DSO-I dataset				
	Precision	Recall	TNR	Accuracy	F-Score
LBP	54.24	58.18	80.15	73.82	56.14
<b>LPQ</b>	<b>61.02</b>	<b>65.46</b>	<b>83.09</b>	<b>78.01</b>	<b>63.16</b>
BGP	44.07	47.27	75.74	67.54	45.61
BSIF	41.38	43.64	75.00	65.97	42.48
CLBP	36.67	40.00	72.06	62.83	38.27
	SwapMe Dataset				
LBP	33.87	38.18	76.16	66.96	35.90
LPQ	33.90	36.36	77.33	67.40	35.09
<b>BGP</b>	<b>36.07</b>	<b>40.74</b>	<b>76.92</b>	<b>68.16</b>	<b>38.26</b>
BSIF	16.67	18.18	70.76	57.96	17.39
CLBP	30.16	34.55	74.27	64.60	32.20
	Combined FaceSwap Dataset				
LBP	31.92	37.50	76.30	67.43	34.48
<b>LPQ</b>	<b>36.36</b>	<b>40.00</b>	<b>79.26</b>	<b>70.29</b>	<b>38.10</b>
BGP	25.53	30.77	73.28	63.5294	27.91
BSIF	14.00	17.50	68.15	56.5714	15.56
CLBP	25.00	30.00	73.33	63.43	27.27

The texture descriptor LPQ gave best performance in DSO-I dataset and Combined FaceSwap dataset. In SwapMe dataset, BGP gave best performance

The bold values show the highest values obtained for a particular dataset

exhibited highest sensitivity with KL distance measure. For comparing the performance of the texture descriptors in different datasets such as DSO-I, SwapMe, and Combined FaceSwap, we selected the distance measure-BH that provided better sensitivity in both IIC and GGW maps.

The precision, sensitivity, specificity, accuracy and F-Score obtained for texture descriptors using BH distance on IIC and GGW maps are shown in Tables 3 and 4, respectively. Compared to the performance of texture descriptors on IIC maps, the descriptors exhibited better performance on GGW maps on all the three datasets. This means that the GGW maps carried more texture variations capable of discriminating spliced regions from authentic regions.

## 8.6 Experiment 2: evaluating color descriptors

For evaluating various color descriptors for forgery localization, we used the software provided by Joost van de Weijer<sup>7</sup>. The sensitivity obtained for various color descriptors on IIC and GGW illuminant maps of spliced images in DSO-I dataset are shown in Figs. 5 and 6 respectively.

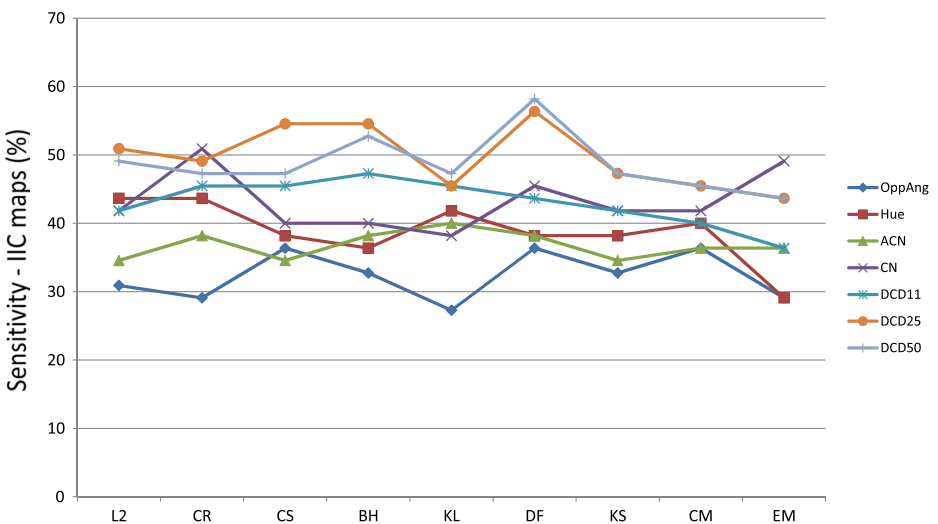
<sup>7</sup><http://cat.cvc.uab.es/~joost/software.html>

**Table 4** Evaluation of texture descriptors on different datasets using GGW maps

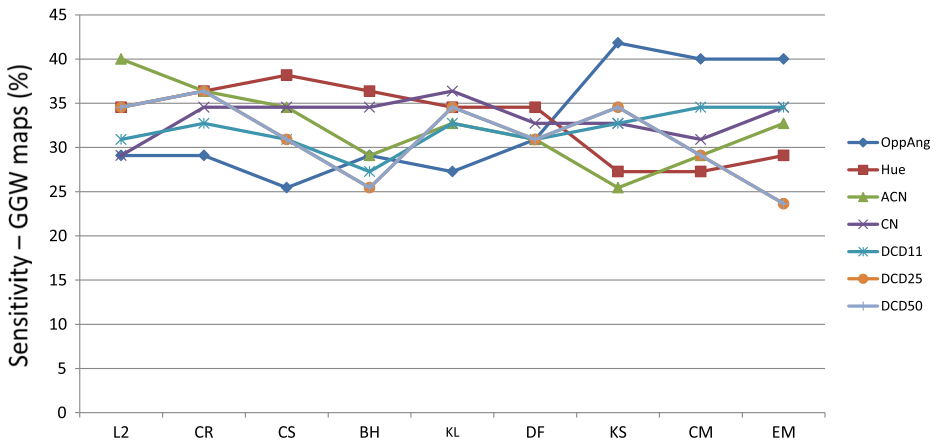
Texture Descriptor	DSO-I dataset				
	Precision	Recall	TNR	Accuracy	F-Score
LBP	61.29	69.09	82.35	78.53	64.96
<b>LPQ</b>	<b>68.42</b>	<b>70.91</b>	<b>86.76</b>	<b>82.20</b>	<b>69.64</b>
BGP	37.29	40.00	72.79	63.35	38.60
BSIF	33.93	34.55	72.79	61.78	34.23
CLBP	47.54	52.73	76.47	69.63	50.00
	SwapMe Dataset				
LBP	33.90	36.36	76.92	66.96	35.09
LPQ	27.12	29.09	74.56	63.39	28.07
<b>BGP</b>	<b>40.68</b>	<b>44.44</b>	<b>78.92</b>	<b>70.45</b>	<b>42.48</b>
BSIF	36.21	38.18	77.98	68.16	37.17
CLBP	21.31	23.64	71.43	59.64	22.41
	Combined FaceSwap Dataset				
<b>LBP</b>	<b>28.89</b>	<b>32.50</b>	<b>76.30</b>	<b>66.29</b>	<b>30.59</b>
LPQ	23.91	27.50	74.07	63.43	25.58
BGP	28.57	30.77	77.10	66.47	29.63
BSIF	25.53	30.00	74.07	64.00	27.59
CLBP	20.45	22.50	74.07	62.29	21.43

Texture descriptors - LPQ, BGP, and LBP gave best performance in DSO-I, SwapMe and Combined FaceSwap datasets respectively

The bold values show the highest values obtained for a particular dataset



**Fig. 5** Sensitivity obtained for color descriptors from IIC maps using different distance measures



**Fig. 6** Sensitivity obtained for color descriptors from GGW maps using different distance measures

As in the case of texture descriptors, here also, for both IIC and GGW maps, the highest sensitivity obtained for each color descriptor is different for different distance measures. This indicates that the nature of the feature vector obtained with different descriptors are different. In IIC map, distant measure DF resulted in the best sensitivity with DCD25 and DCD50 descriptors. But, on GGW, the descriptor - Opponent histogram yielded highest sensitivity using KS distance measure. This shows that both IIC and GGW may contain varied color features, that can be captured using different color descriptors. Hence, it would be better if both illuminant maps are considered in forgery localization techniques.

The precision, recall, TNR, accuracy and F-Score obtained for color descriptors on the three datasets using IIC and GGW maps (with DF distance measure) are shown in Tables 5 and 6, respectively. For DSO-I dataset, the IIC maps showed better performance for color descriptors compared to GGW maps. But, for SwapMe and Combined FaceSwap, the variation in the performance of descriptors are less noticeable. This is due to the fact that, DSO-I dataset contains images in uncompressed PNG format, and the IIC and GGW maps showed visible color variations. In addition to this, the clarity of facial regions in DSO-I dataset were better compared to the clarity of facial regions in both SwapMe and Combined FaceSwap.

### 8.7 Experiment 3: evaluating combined color and texture descriptors

In this section, the discriminative power of descriptors that consider both color and texture features, such as Color-LPQ, two variants of color textons - JTD and STD are evaluated. The source code for computing Color LPQ descriptor was provided by Pedone [32]. The JTD and STD descriptors are generated using the source code provided by Alvarez and Vanrell<sup>8</sup> [1]. The sensitivity obtained for various combined color and texture descriptors on IIC and GGW illuminant maps of spliced images in DSO-I dataset are shown in Figs. 7 and 8 respectively.

In IIC map, Color LPQ and STD provided highest sensitivity with the distance measure CR as shown in Fig. 7. Similarly, Color LPQ showed highest sensitivity with CR distance measure in GGW maps too. The precision, recall, TNR, and F-Score obtained for combined

<sup>8</sup><http://www.cat.uab.cat/Research/ColorTextureDescriptors/>

**Table 5** Evaluation of color descriptors on different datasets using IIC maps

Texture Descriptor	DSO-I dataset				
	Precision	Recall	TNR	Accuracy	F-Score
Opponent	35.71	36.36	73.53	62.83	36.04
Hue	35.59	38.18	72.06	62.30	36.84
ACN	33.87	38.18	69.85	60.73	35.90
CN	41.67	45.45	74.26	65.97	43.48
DCD11	46.26	56.36	73.53	68.59	50.82
<b>DCD25</b>	<b>54.24</b>	<b>58.18</b>	<b>80.15</b>	<b>73.82</b>	<b>56.14</b>
<b>DCD50</b>	<b>54.24</b>	<b>58.18</b>	<b>80.15</b>	<b>73.82</b>	<b>56.14</b>
	SwapMe Dataset				
Opponent	18.97	20.00	72.51	59.73	19.47
Hue	22.73	27.27	70.18	59.73	24.79
<b>ACN</b>	<b>28.81</b>	<b>30.91</b>	<b>75.44</b>	<b>64.60</b>	<b>29.82</b>
CN	17.59	18.18	72.51	59.29	17.86
DCD11	22.41	23.64	73.68	61.50	23.01
DCD25	20.00	23.64	69.59	58.41	21.67
DCD50	20.00	23.64	69.59	58.41	21.67
	Combined FaceSwap Dataset				
Opponent	19.15	22.50	71.85	60.57	20.69
Hue	16.00	20.00	68.89	57.71	17.78
<b>ACN</b>	<b>27.27</b>	<b>30.00</b>	<b>76.30</b>	<b>65.71</b>	<b>28.57</b>
CN	19.15	22.50	71.85	60.57	20.69
DCD11	15.22	17.50	71.11	58.86	16.28
DCD25	12.77	15.00	69.63	57.15	13.79
DCD50	12.77	15.00	69.63	57.15	13.79

The descriptors -DCD25 and DCD50 showed better performance in DSO-I dataset, and ACN descriptor showed best performance SwapMe dataset and Combined FaceSwap dataset

The bold values show the highest values obtained for a particular dataset

color-texture descriptors on various datasets using CR measure on both IIC and GGW maps are shown in Tables 7 and 8, respectively.

## 8.8 Experiment 4: evaluating color histograms, moments and color shape descriptors

Color histograms, moments and shape descriptors are generated using the source code provided by Koen van de Sande<sup>9</sup> [37].

Figure 9 shows the sensitivity obtained for color histogram moments and moment descriptors in IIC map. The color moment invariants and rgb histogram yielded the highest sensitivity of 64.58% with CS and KL distance measures. Table 9 shows the evaluation using various performance metrics obtained on different datasets using CS distance measure. Fig. 10 shows the sensitivity evaluation of color-shape descriptors. Among the different

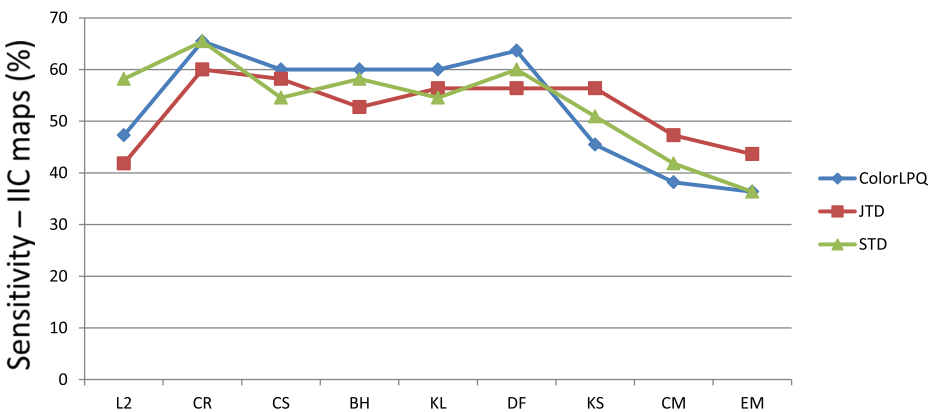
<sup>9</sup><http://koen.me/research/colordescriptors/readme>

**Table 6** Evaluation of color descriptors on different datasets using GGW maps

Texture Descriptor	DSO-I dataset				
	Precision	Recall	TNR	Accuracy	F-Score
Opponent	28.81	30.91	69.12	58.12	29.83
<b>Hue</b>	<b>31.15</b>	<b>34.55</b>	<b>49.21</b>	<b>59.16</b>	<b>32.70</b>
ACN	29.82	30.91	70.59	59.16	30.36
CN	30.00	32.73	69.12	58.64	31.30
DCD11	28.81	30.91	69.12	58.12	29.85
DCD25	29.82	30.91	70.59	59.16	30.36
DCD50	29.82	30.91	70.59	59.16	30.36
	SwapMe Dataset				
Opponent	24.14	25.45	73.96	62.05	24.78
Hue	28.81	30.91	75.15	64.29	29.82
ACN	25.81	29.09	72.78	62.06	27.35
<b>CN</b>	<b>31.67</b>	<b>34.55</b>	<b>75.74</b>	<b>65.63</b>	<b>33.04</b>
DCD11	22.41	23.64	73.37	61.16	23.09
DCD25	22.73	27.27	69.83	59.38	24.79
DCD50	22.73	27.27	69.83	59.38	24.79
	Combined FaceSwap Dataset				
Opponent	20.00	25.00	70.37	60.00	22.22
Hue	13.46	17.50	66.67	55.43	15.22
<b>ACN</b>	<b>21.74</b>	<b>25.00</b>	<b>73.33</b>	<b>62.29</b>	<b>23.26</b>
CN	5.77	7.50	63.70	50.86	06.52
DCD11	11.36	12.50	71.11	57.71	11.90
DCD25	11.11	12.50	70.37	57.14	11.76
DCD50	11.11	12.50	70.37	57.14	11.76

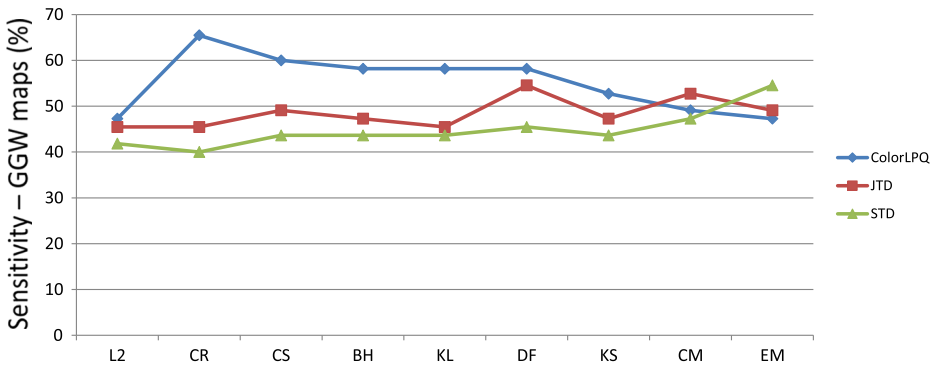
Among the various descriptors Hue, CN and ACN gave comparatively better results in DSO-I, SwapMe, and FaceSwap datasets respectively

The bold values show the highest values obtained for a particular dataset



**Fig. 7** Sensitivity obtained for combined color texture descriptors from IIC maps using different distance measures





**Fig. 8** Sensitivity obtained for combined color texture descriptors from GGW maps using different distance measures

variants of color shape descriptors, Hue-SIFT exhibited the highest sensitivity of 54.17% using L2 distance measure. Table 10 shows the evaluation of various performance metrics obtained on DSO-I dataset using L2 distance measure.

### 8.9 Experiment 5: evaluating HOG and GLRLM descriptors

GLRLM and HOG features are computed using the Matlab source codes provided by Wei [48] and Ludwig et al. [18] respectively. The highest sensitivity obtained for HOG is 43.64% in IIC map using CR, DF distance measures, and 52.73% in GGW map with distance measure DF. In both, IIC and GGW, the HOG descriptors performed better than GLRLM descriptors. Figures 11 and 12 shows the sensitivity obtained for various distance measures

**Table 7** Evaluation of combined color-texture on different datasets using IIC maps

Texture Descriptor	DSO-I dataset				
	Precision	Recall	TNR	Accuracy	F-Score
<b>Color LPQ</b>	<b>61.02</b>	<b>65.46</b>	<b>83.09</b>	<b>78.01</b>	<b>63.16</b>
JTD	56.90	60.00	81.62	75.39	58.41
STD	59.02	65.46	81.62	76.96	62.07
SwapMe Dataset					
ColorLPQ	32.00	32.65	74.81	63.59	32.32
<b>JTD</b>	<b>37.70</b>	<b>46.94</b>	<b>71.21</b>	<b>64.64</b>	<b>41.82</b>
STD	30.91	34.69	71.21	61.33	32.69
Combined FaceSwap Dataset					
ColorLPQ	34.15	35.90	79.23	69.23	35.00
<b>JTD</b>	<b>34.88</b>	<b>38.46</b>	<b>78.29</b>	<b>69.05</b>	<b>36.59</b>
STD	34.09	38.46	77.52	68.45	36.14

The descriptor Color LPQ exhibited highest performance in DSO-I dataset, and JTD exhibited highest performance in SwapMe and Combined FaceSwap

The bold values show the highest values obtained for a particular dataset

**Table 8** Evaluation of combined color-texture on different datasets using GGW maps

Texture Descriptor	DSO-I dataset				
	Precision	Recall	TNR	Accuracy	F-Score
<b>Color LPQ</b>	<b>63.16</b>	<b>65.45</b>	<b>84.56</b>	<b>79.06</b>	<b>64.29</b>
JTD	43.86	45.46	76.47	67.54	44.64
STD	36.67	40.00	72.06	62.83	38.26
	SwapMe Dataset				
ColorLPQ	36.36	37.04	76.03	65.50	36.70
<b>JTD</b>	<b>43.33</b>	<b>48.15</b>	<b>76.22</b>	<b>68.53</b>	<b>45.61</b>
STD	35.71	37.04	74.83	64.47	36.36
	Combined FaceSwap Dataset				
<b>ColorLPQ</b>	<b>30.23</b>	<b>34.21</b>	<b>76.56</b>	<b>66.87</b>	<b>32.10</b>
JTD	22.92	28.95	71.09	61.45	25.58
STD	25.58	28.95	75.00	64.46	27.16

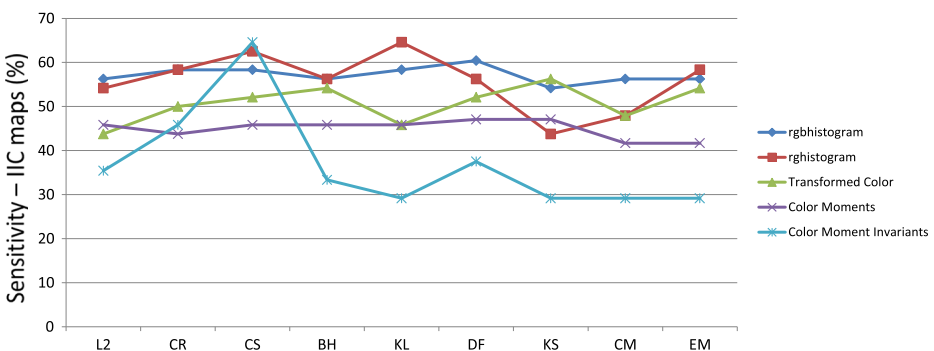
The descriptor ColorLPQ showed better performance in DSO-I and Combined FaceSwap datasets. On SwapMe dataset, JTD showed better performance

The bold values show the highest values obtained for a particular dataset

in IIC and GGW maps respectively. The evaluation of HOG and GLRLM descriptors on different datasets using IIC maps and GGW maps is shown in Tables 11 and 12 respectively.

### 8.10 Evaluation of deep features

For evaluating the performance of deep features, we used the pretrained Convolutional Neural Network (CNN) model, vgg-f [9] available in MatConvNet library [45]. The CNN model, vgg-f consists of 8 layers - 5 convolutional, and 3 fully-connected layers. The input image is resized to 224 x 224 pixel regions. The 4096 dimensional deep features are extracted from the 7th layer. We evaluated the deep features obtained from both IIC and GGW maps with various distance measures. Tables 13 and 14 shows the performance obtained using IIC and GGW maps with deep features extracted from the vgg-f model.



**Fig. 9** Sensitivity obtained for color histogram and moment descriptors from IIC maps using different distance measures.

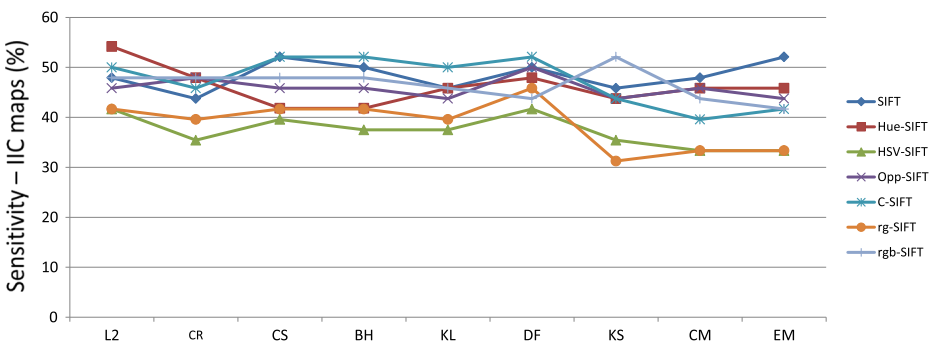
**Table 9** Evaluation of color histogram and moment descriptors on different datasets using IIC maps

Texture Descriptor	DSO-I dataset				
	Precision	Recall	TNR	Accuracy	F-Score
<b>rgb histogram</b>	<b>56.00</b>	<b>58.33</b>	<b>80.18</b>	<b>73.59</b>	<b>57.14</b>
rg histogram	54.90	58.33	79.28	72.96	56.57
Transformed color	47.06	50.00	75.68	67.92	48.48
Color moments	42.00	43.75	73.87	64.78	42.86
Color moment Invariants	44.00	45.83	74.77	66.04	44.90
	SwapMe Dataset				
rgb histogram	28.57	28.57	69.70	57.45	28.57
rg histogram	42.86	42.86	75.76	65.96	42.86
Transformed color	35.71	35.71	72.73	61.70	35.71
<b>Color moments</b>	<b>50.00</b>	<b>57.14</b>	<b>75.76</b>	<b>70.21</b>	<b>53.33</b>
Color moment Invariants	28.57	28.57	69.70	57.45	28.57
	Combined FaceSwap Dataset				
rgbHistogram	27.27	33.33	63.64	54.84	30.00
rgHistogram	22.22	22.22	68.18	54.84	22.22
Transformed color	11.11	11.11	63.64	48.39	11.11
Color moments	33.33	33.33	72.73	61.29	33.33
<b>Color moment Invariants</b>	<b>44.44</b>	<b>44.44</b>	<b>77.27</b>	<b>67.74</b>	<b>44.44</b>

The descriptor, rgb histogram, color moments, and color moment invariants exhibited better performance in DSO-I, SwapMe and Combined FaceSwap datasets respectively

The bold values show the highest values obtained for a particular dataset

In IIC maps, the BH distance measure yielded best performance on DSO-I dataset, whereas the distance measure KL resulted in best performance on SwapMe dataset. The distance measure L2 provided best results on Combined FaceSwap dataset. Alternatively, on GGW maps, CR and KL distance measures showed good results on DSO-I dataset, and for SwapMe, the best performance is obtained with CR and DF distance measures. For Combined FaceSwap dataset, the best results are obtained with L2 distance. However, in both IIC and GGW maps, the performance of deep features was less compared to the texture



**Fig. 10** Sensitivity obtained for color shape descriptors from IIC maps using different distance measures

**Table 10** Evaluation of color shape descriptors on DSO-I dataset using IIC maps

Color-Shape Descriptor	Precision/Specificity	Recall/Sensitivity	TNR	Accuracy	F-Score
SIFT	42.59	47.92	72.07	64.78	45.10
<b>Hue-SIFT</b>	<b>52.00</b>	<b>54.17</b>	<b>78.38</b>	<b>71.07</b>	<b>53.06</b>
HSV-SIFT	40.00	41.67	72.97	63.52	40.82
Opp-SIFT	44.00	45.83	74.77	66.04	44.90
C-SIFT	47.06	50.00	75.68	67.93	48.49
rg-SIFT	40.00	41.67	72.97	63.52	40.82
rgb-SIFT	46.94	47.92	76.58	67.93	47.42

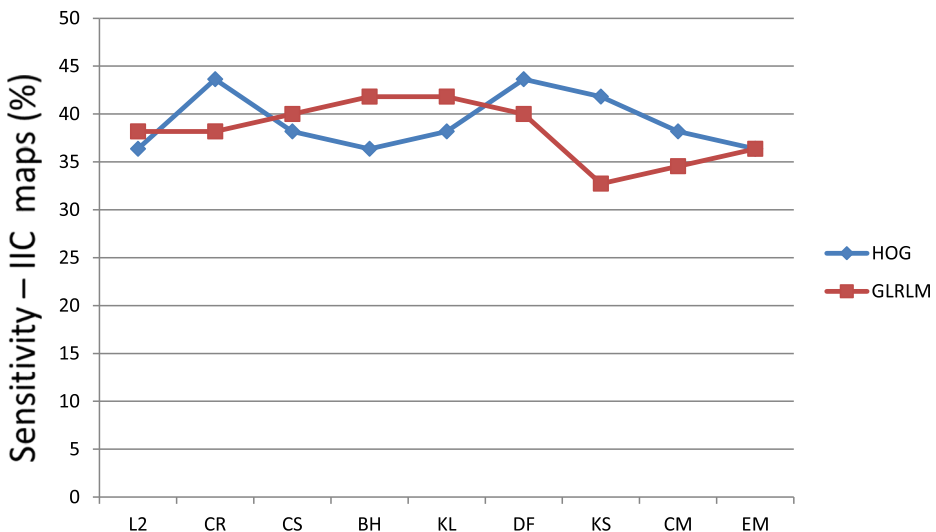
Among the various descriptors, Hue-SIFT showed best performance

The bold values show the highest values obtained for a particular dataset

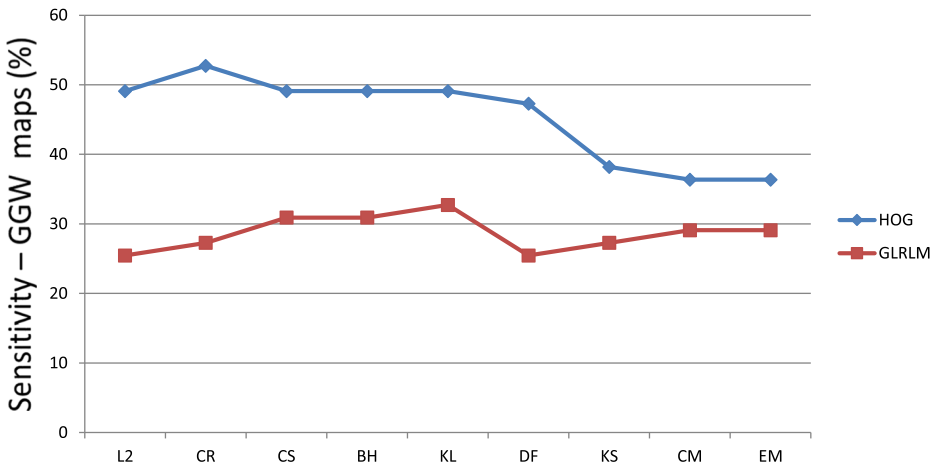
descriptor-LPQ in DSO-I. But, for Swapme, and Combined FaceSwap datasets deep features resulted in comparable performance. Therefore, as in various computer vision domains were deep features yielded better performance, an application specific model could provide better performance for forgery localization from illumination maps too.

### 8.11 Effect of JPEG compression

The robustness of texture, color, and combined color-texture descriptors are evaluated against JPEG compression. Experiments are conducted on DSO-I dataset as the dataset contains images in uncompressed PNG format. For evaluation, the images in the DSO-I dataset are compressed with JPEG quality factors 60, 70, 80, and 90. Figure 13 shows the performance of various texture features at different compression levels (60-90). Experiments are conducted on DSO-I dataset using IIC maps for exploring the effect of JPEG compression, since images in DSO-I dataset are in uncompressed PNG format. In Figs. 13, 14



**Fig. 11** Sensitivity obtained for HOG and GLRLM descriptors from IIC maps using different distance measures



**Fig. 12** Sensitivity obtained for HOG and GLRLM descriptors from IIC maps using different distance measures

and 15, the performance on the uncompressed version is marked as '100'. As depicted in Fig. 13, the texture descriptors such as LBP, LPQ, and BSIF showed a performance variation on JPEG compressed images (marked as 60-90) compared to the uncompressed images (marked as '100'). When the images are compressed, the JPEG boundary artifacts and the encoding scheme alter the underlying pixel boundaries, thereby affecting the segmentation of image during the generation of illumination map. An interesting observation is that the BGP and CLBP features perform relatively consistently in all JPEG compression levels and uncompressed version.

Figures 14 and 15 show the performance of various color and combined color-texture features respectively at the different compression levels. In Fig. 14, the color descriptors such as Opponent, Hue, ACN, CN, and DCD11 performs evenly at different JPEG compression levels and the uncompressed version. But, for color descriptors-DCD25 and DCD50, the

**Table 11** Evaluation of HOG and GLRLM descriptors on different datasets using IIC maps

Texture Descriptor	DSO-I dataset				
	Precision	Recall	TNR	Accuracy	F-Score
<b>HOG</b>	<b>42.86</b>	<b>43.64</b>	<b>76.47</b>	<b>67.02</b>	<b>43.24</b>
GLRLM	34.92	40.00	69.85	61.26	37.29
	SwapMe Dataset				
HOG	32.79	36.36	76.02	66.37	34.48
<b>GLRLM</b>	<b>36.07</b>	<b>40.00</b>	<b>77.19</b>	<b>68.14</b>	<b>37.93</b>
	Combined FaceSwap Dataset				
<b>HOG</b>	<b>40.00</b>	<b>45.00</b>	<b>80.00</b>	<b>72.00</b>	<b>42.35</b>
GLRLM	20.45	22.50	74.07	62.29	21.43

The HOG descriptor showed best performance in DSO-I and Combined FaceSwap datasets, and GLRLM descriptor showed best performance in SwapMe dataset

The bold values show the highest values obtained for a particular dataset

**Table 12** Evaluation of HOG and GLRLM descriptors on different datasets using GGW maps

Texture Descriptor	DSO-I dataset				
	Precision	Recall	TNR	Accuracy	F-Score
<b>HOG</b>	<b>51.79</b>	<b>52.73</b>	<b>80.15</b>	<b>72.25</b>	<b>52.25</b>
GLRLM	24.59	27.27	66.18	54.97	25.86
	SwapMe Dataset				
<b>HOG</b>	<b>30.00</b>	<b>32.73</b>	<b>75.15</b>	<b>64.73</b>	<b>31.30</b>
GLRLM	23.33	25.45	72.78	61.16	24.35
	Combined FaceSwap Dataset				
<b>HOG</b>	<b>38.10</b>	<b>40.00</b>	<b>80.74</b>	<b>71.43</b>	<b>39.02</b>
GLRLM	27.90	30.00	77.04	66.29	28.92

The HOG descriptor showed best performance all the three datasets

The bold values show the highest values obtained for a particular dataset

**Table 13** Evaluation of deep features extracted from pre-trained model vgg on IIC maps of different datasets

Dataset	Distance/ Measure	Precision/ Specificity	Recall/ Sensitivity	TNR	Accuracy	F-Score
DSO-I	L2	42.37	45.45	75.00	66.49	43.86
	CR	41.94	47.27	73.53	65.97	44.44
	CS	41.27	47.27	72.79	65.45	44.07
	<b>BH</b>	<b>43.55</b>	<b>49.10</b>	<b>74.27</b>	<b>67.02</b>	<b>46.15</b>
	KL	43.33	47.27	75.00	67.02	45.22
	DF	40.68	43.64	74.26	65.45	42.11
	KS	34.92	40.00	69.85	61.26	37.29
	CM	32.20	34.55	70.59	60.21	33.33
	EM	35.59	38.18	72.06	62.31	36.85
	SwapMe	L2	36.67	40.00	77.78	68.59
CR		36.36	43.64	75.44	67.70	39.67
CS		34.38	40.00	75.44	66.81	36.97
BH		35.39	41.82	75.44	67.26	38.36
<b>KL</b>		<b>53.45</b>	<b>56.36</b>	<b>84.21</b>	<b>77.43</b>	<b>54.87</b>
DF		34.92	40.00	76.02	67.26	37.29
KS		27.87	30.91	74.27	63.72	29.31
CM		25.81	29.09	73.10	62.39	27.36
EM		26.56	30.91	72.51	62.39	28.57
Combined FaceSwap		<b>L2</b>	<b>40.38</b>	<b>52.50</b>	<b>77.04</b>	<b>71.43</b>
	CR	40.43	47.50	79.26	72.00	43.68
	CS	39.22	50.00	77.04	70.86	43.96
	BH	40.82	50.00	78.52	72.00	44.94
	KL	31.48	42.50	72.59	65.71	36.17
	DF	36.54	47.50	75.56	69.14	41.30
	KS	36.96	42.50	78.52	70.29	39.53
	CM	38.78	47.50	77.78	70.86	42.70
	EM	37.21	40.00	80.00	70.86	38.55

The best performance for DSO-I, SwapMe and FaceSwap datasets were obtained with BH, KL, and L2 distance measures respectively

The bold values show the highest values obtained for a particular dataset

**Table 14** Evaluation of deep features extracted from pre-trained model vgg on GGW maps of different datasets

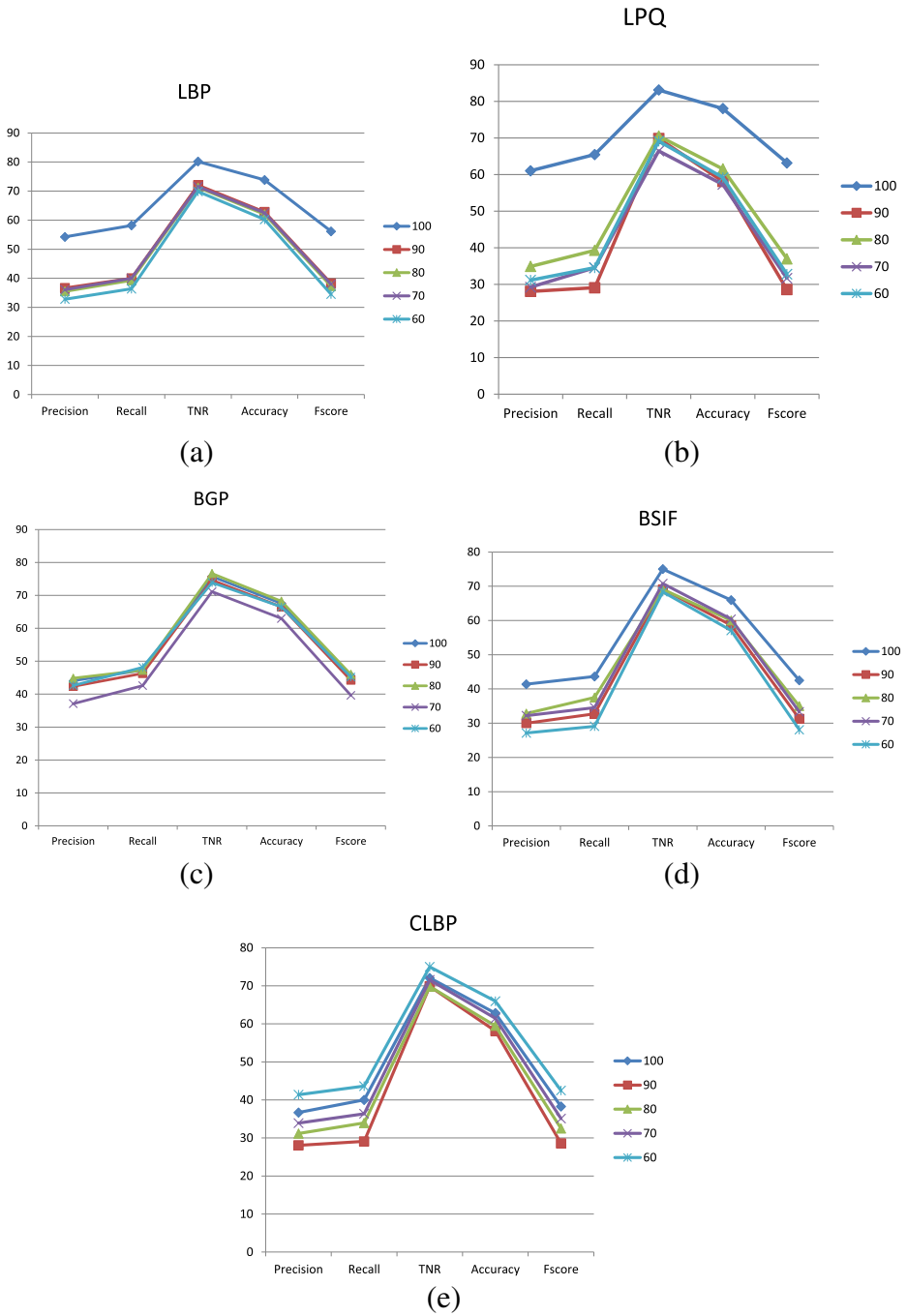
Dataset	Distance/ Measure	Precision/ Specificity	Recall/ Sensitivity	TNR	Accuracy	F-Score
DSO-I	L2	53.33	58.18	79.41	73.30	55.65
	<b>CR</b>	<b>59.32</b>	<b>63.64</b>	<b>82.35</b>	<b>76.96</b>	<b>61.40</b>
	CS	54.24	58.18	80.15	73.82	56.14
	BH	50.85	54.55	78.68	71.73	52.63
	<b>KL</b>	<b>59.32</b>	<b>63.64</b>	<b>82.35</b>	<b>76.96</b>	<b>61.40</b>
	DF	51.67	56.36	78.68	72.25	53.91
	KS	40.98	45.45	73.53	65.45	43.10
	CM	34.43	38.18	70.59	61.26	36.21
SwapMe	EM	31.67	34.55	69.85	59.69	33.04
	L2	42.37	45.45	79.88	71.43	43.86
	<b>CR</b>	<b>49.12</b>	<b>50.91</b>	<b>82.84</b>	<b>75.00</b>	<b>50.00</b>
	CS	47.37	49.09	82.25	74.11	48.21
	BH	46.55	49.09	81.66	73.66	47.79
	KL	38.24	47.27	75.15	68.30	42.28
	<b>DF</b>	<b>49.12</b>	<b>50.91</b>	<b>82.84</b>	<b>75.00</b>	<b>50.00</b>
	KS	32.76	34.55	76.92	66.52	33.63
Combined FaceSwap	CM	28.33	30.91	74.56	63.84	29.57
	EM	31.59	32.73	76.92	66.07	32.14
	<b>L2</b>	<b>40.38</b>	<b>52.50</b>	<b>77.04</b>	<b>71.43</b>	<b>45.65</b>
	CR	40.43	47.50	79.26	72.00	43.68
	CS	39.22	50.00	77.04	70.86	43.96
	BH	40.82	50.00	78.52	72.00	44.94
	KL	31.48	42.50	72.59	65.71	36.17
	DF	36.54	47.50	75.56	69.14	41.30
	KS	36.96	42.50	78.52	70.29	39.53
	CM	38.78	47.50	77.78	70.86	42.70
	EM	37.21	40.00	80.00	70.86	38.55

The bold values show the highest values obtained for a particular dataset

performance degraded on JPEG compression. In Fig. 16, the combined color-descriptors-Color LPQ, JTD, and STD showed a performance degradation with JPEG compression. Also, the performance lowers at lower JPEG compression levels. This indicates that the combined color-texture descriptors are affected by JPEG encoding scheme.

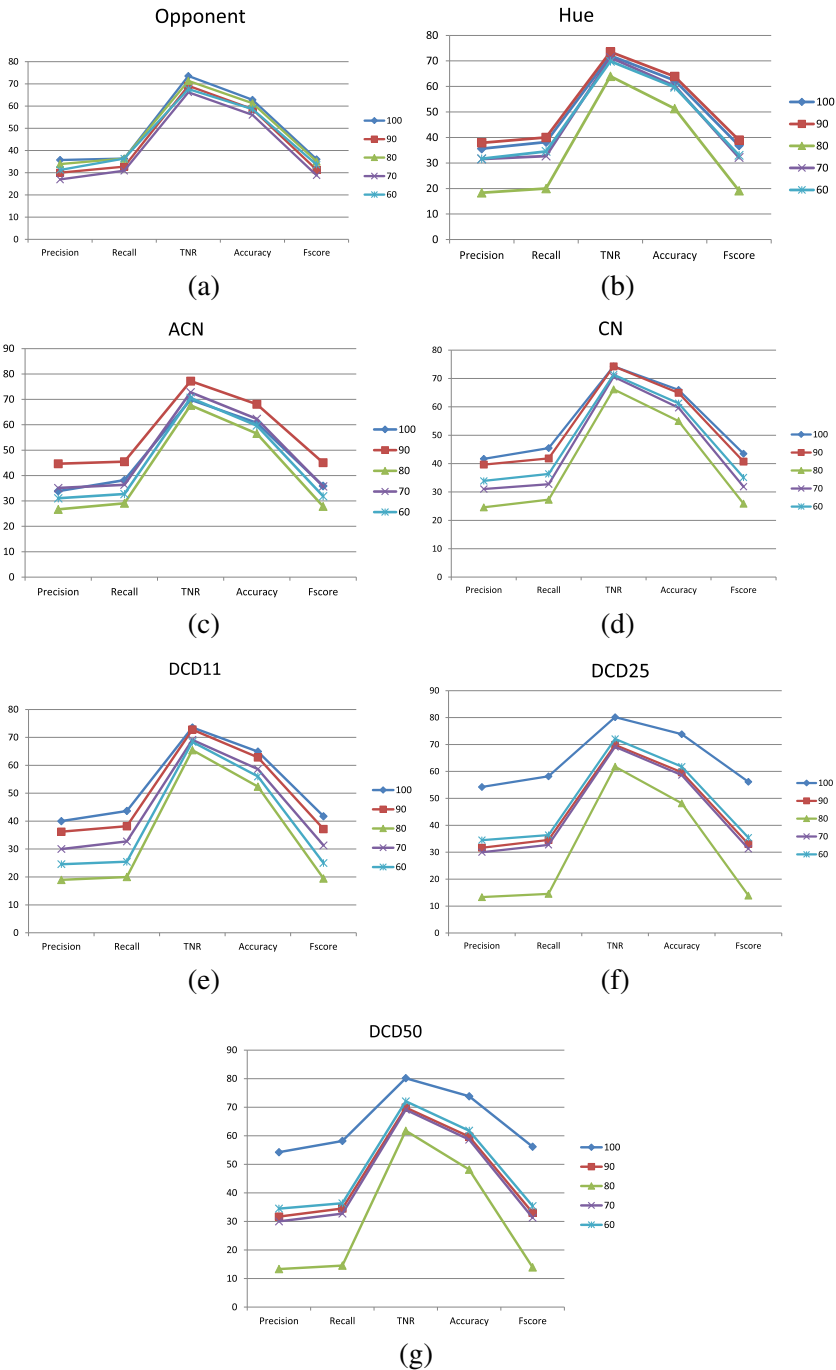
## 8.12 Comparison with other illumination inconsistency based methods

We found that the descriptors-LPQ, JTD, and HOG exhibited better performance on DSO-I, SwapMe and Combined FaceSwap datasets respectively. The performance of the descriptors are compared with the performance of two previous works proposed by Gholap and Bora [16], and Mazumdar and Bora [27]. Figures 16, 17 and 18 show the performance comparison on DSO-I dataset, SwapMe and Combined FaceSwap datasets respectively. It is clear that compared to both the previous works, the feature descriptors such as LPQ, JTD, and HOG showed better results on DSO-I, SwapMe and Combined FaceSwap datasets respectively.

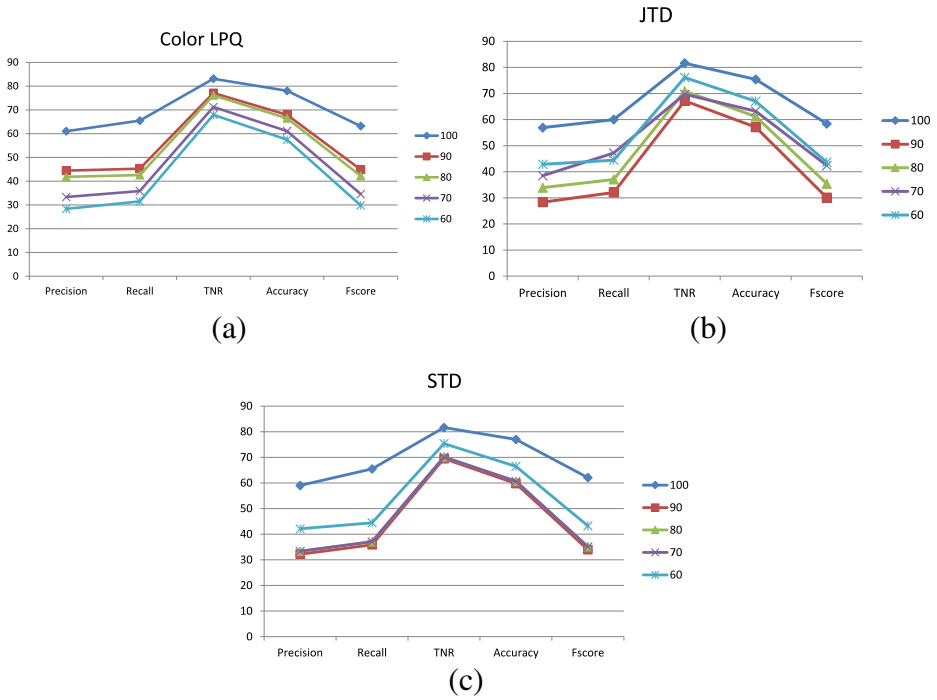


**Fig. 13** The performance of various texture descriptors at different JPEG compression levels. Quality factors considered are 60, 70, 80, 90. The uncompressed level 100 indicates the image without any compression. Experiments are carried on DSO-I dataset using IIC maps. **a** LBP **b** LPQ **c** BGP **d** BSIF **e** CLBP

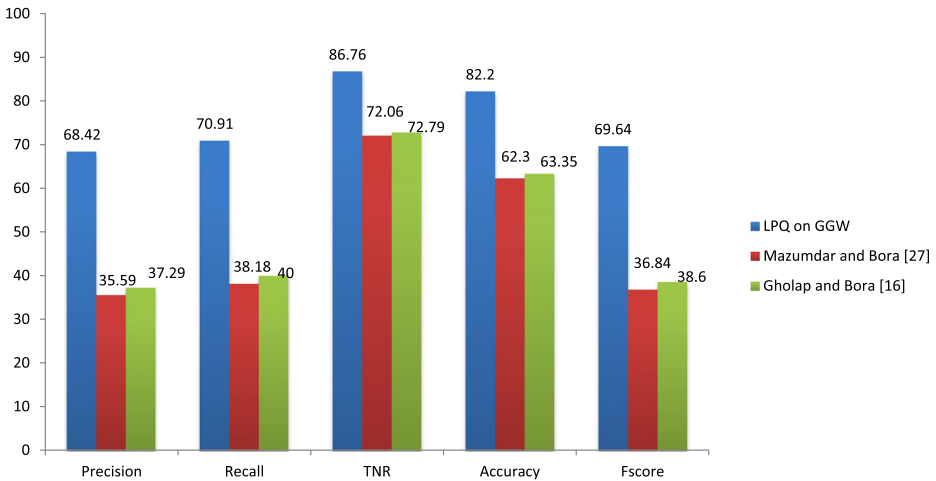




**Fig. 14** The performance of various color descriptors at different JPEG compression levels. Quality factors considered are 60, 70, 80, 90. The uncompressed level 100 indicates the image without any compression. Experiments are carried on DSO-I dataset using IIC maps. **a** Opponent **b** Hue **c** ACN **d** CN **e** DCD11 **f** DCD25 **g** DCD50



**Fig. 15** The performance of various combined color-texture descriptors at different JPEG compression levels. Quality factors considered are 60, 70, 80, 90. The uncompressed level 100 indicates the image without any compression. Experiments are carried on DSO-I dataset using IIC maps. **a** ColorLPQ **b** JTD **c** STD



**Fig. 16** Comparison of LPQ features with previous methods on DSO-I dataset

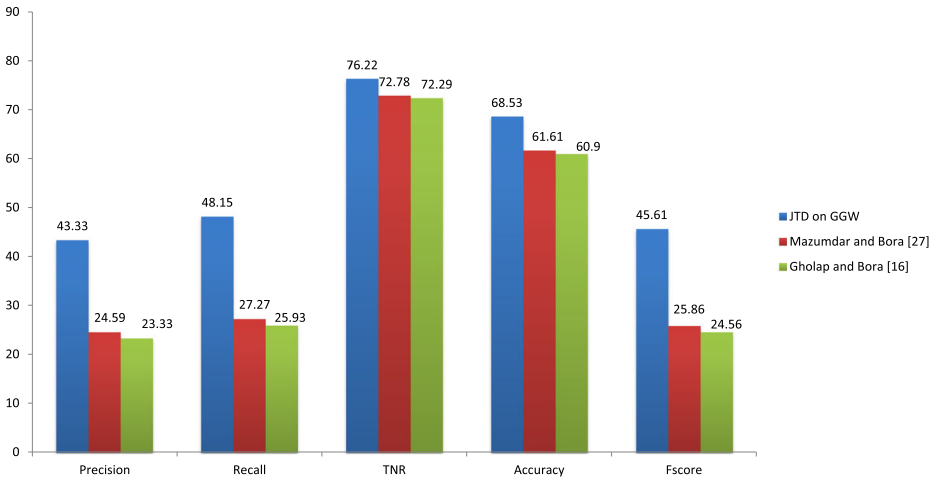


Fig. 17 Comparison of JTD features with previous methods on SwapMe dataset

### 8.13 Discussion

In general, the experimental results reveal that the texture features and combined color-texture features are better at locating forged image regions from illuminant maps than color features. The reason is that, the changes in texture patterns are more prominent than the variations in color. In many of the images in the dataset, the color variations are too subtle to be captured by the color descriptors. The color variation will be subtle to detect unless there is a drastic difference in the illumination environment where the two photographs are captured. For example, in a spliced image composed of image regions captured at indoor and outdoor environments, there will be prominent variations in the color distribution.

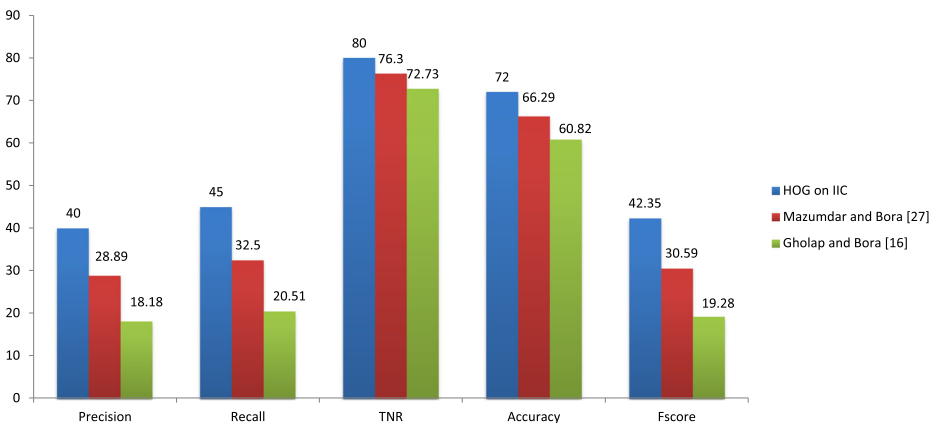


Fig. 18 Comparison of HOG features with previous methods on Combined FaceSwap dataset

## 9 Conclusion

Inconsistencies in illumination distribution can reveal forged image regions. In the proposed work, we carried out a comprehensive evaluation of the discriminative power of a number of texture, color and combined color-texture descriptors in forgery localization. For evaluation, forged images containing human facial regions were used and two variants of illumination distribution, such as IIC and GGW maps were considered. The discriminating capability of different features are assessed using 9 different histogram distance measures.

From the experiments, it is clear that texture descriptors are more capable of locating forged region compared to color features, and combined color-texture features. Also, among the various descriptors evaluated, we found that, LPQ descriptor showed the highest sensitivity of 70.91% in GGW map and 65.45% in IIC maps. We also evaluated deep features based on the pre-trained CNN model, vgg-f. But, evaluation showed that the performance of texture features is better compared to the deep features from the pre-trained model using the illuminant maps on DSO-I dataset.

We observed that the detection performance varied with different histogram distance measures for different descriptors indicating the differences in the nature of color and texture patterns captured by the feature descriptors. This suggests that a combination of features, such as a multi-feature representation may improve the detection accuracy. Hence, in future, we plan to consider future fusion for improving the accuracy in forgery localization.

**Acknowledgment** Authors acknowledge the Department of Higher Education, Government of Kerala for funding the research and the Department of Computer Science and Engineering, College of Engineering-Trivandrum for providing lab facilities to carry out the work. The authors would like to thank Dr. Tiago José De Carvalho for sharing the database. Also, authors would like to thank Mr. Aniruddha Mazumdar for sharing the source code of previous works for comparison.

## References

1. Alvarez S, Vanrell M (2012) Texton theory revisited: a bag-of-words approach to combine textons. *Pattern Recogn* 45(12):4312–4325
2. Benavente R, Vanrell M, Baldrich R (2008) Parametric fuzzy sets for automatic color naming. *JOSA A* 25(10):2582–2593
3. Benesty J, Chen J, Huang Y, Cohen I (2009) Pearson correlation coefficient. In: *Noise reduction in speech processing*. Springer, pp 1–4
4. Bosch A, Zisserman A, Muoz X (2008) Scene classification using a hybrid generative/discriminative approach. *IEEE Trans Pattern Anal Mach Intell* 30(4):712–727. <https://doi.org/10.1109/TPAMI.2007.70716>
5. Burghouts GJ, Geusebroek JM (2009) Performance evaluation of local colour invariants. *Comput Vis Image Underst* 113(1):48–62
6. Cao G, Zhao Y, Ni R (2008) Image composition detection using object-based color consistency. In: 2008 9Th international conference on signal processing, pp 1186–1189. <https://doi.org/10.1109/ICOSP.2008.4697342>
7. Carron T, Lambert P (1994) Color edge detector using jointly hue, saturation and intensity. In: 1994 Proceedings of IEEE international conference on Image processing, ICIP-94, vol 3. IEEE, pp 977–981
8. Carvalho T, Faria FA, Pedrini H, Torres RdS, Rocha A (2016) Illuminant-based transformed spaces for image forensics. *IEEE Trans Inf Forensics Secur* 11(4):720–733
9. Chatfield K, Simonyan K, Vedaldi A, Zisserman A (2014) Return of the devil in the details: delving deep into convolutional nets. In: *British machine vision conference*
10. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: 2005 IEEE Computer society conference on computer vision and pattern recognition (CVPR'05), vol 1, pp 886–893. <https://doi.org/10.1109/CVPR.2005.177>

11. De Carvalho TJ, Riess C, Angelopoulou E, Pedrini H, de Rezende Rocha A (2013) Exposing digital image forgeries by illumination color classification. *IEEE Trans Inf Forensics Secur* 8(7):1182–94
12. Durbin J, Knott M (1972) Components of cramer-von mises statistics. *i. J R Stat Soc Ser B Methodol* 34:290–307
13. Fan Y, Carr P, Fernandez-Maloigne C (2015) Image splicing detection with local illumination estimation. In: 2015 IEEE international conference on Image processing (ICIP), pp 2940–44. <https://doi.org/10.1109/ICIP.2015.7351341>
14. Finlayson GD, Schaefer G (2001) Solving for colour constancy using a constrained dichromatic reflection model. *Int J Comput Vis* 42(3):127–144
15. Francis K, Gholap S, Bora PK (2014) Illuminant colour based image forensics using mismatch in human skin highlights. In: 2014 twentieth national conference on Communications (NCC). IEEE, pp 1–6
16. Gholap S, Bora PK (2008) Illuminant colour based image forensics. In: TENCON 2008 - 2008 IEEE Region 10 conference, pp 1–5. <https://doi.org/10.1109/TENCON.2008.4766772>
17. Guo Z, Zhang L, Zhang D (2010) A completed modeling of local binary pattern operator for texture classification. *IEEE Trans Image Process* 19(6):1657–1663. <https://doi.org/10.1109/TIP.2010.2044957>
18. Junior OL, Delgado D, Goncalves V, Nunes U (2009) Trainable classifier-fusion schemes: an application to pedestrian detection. In: 2009 12Th international IEEE conference on intelligent transportation systems, pp 1–6. <https://doi.org/10.1109/ITSC.2009.5309700>
19. Kailath T (1967) The divergence and bhattacharyya distance measures in signal selection. *IEEE Trans Commun Technol* 15(1):52–60
20. Kannala J, Rahtu E (2012) Bsif: binarized statistical image features. In: 2012 21st International Conference on Pattern Recognition (ICPR). IEEE, pp 1363–1366
21. Khan FS, Anwer RM, van de Weijer J, Felsberg M, Laaksonen J (2015) Compact color–texture description for texture classification. *Pattern Recogn Lett* 51:16–22
22. Khan R, van de Weijer J, Khan FS, Muselet D, Ducottet C, Barat C (2013) Discriminative color descriptors. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 2866–2873. <https://doi.org/10.1109/CVPR.2013.369>
23. Levina E, Bickel P (2001) The earth mover’s distance is the mallows distance: Some insights from statistics. In: 2001 Proceedings of Eighth IEEE international conference on Computer Vision (ICCV), vol 2. IEEE, pp 251–256
24. Ling H, Okada K (2006) Diffusion distance for histogram comparison. In: 2006 IEEE computer society conference on Computer vision and pattern recognition, vol 1. IEEE, pp 246–253
25. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110
26. Massey FJ Jr (1951) The kolmogorov-smirnov test for goodness of fit. *J Am Stat Assoc* 46(253):68–78
27. Mazumdar A, Bora PK (2016) Exposing splicing forgeries in digital images through dichromatic plane histogram discrepancies. In: Proceedings of the tenth indian conference on computer vision, graphics and image processing. ACM, p 62
28. Meshgi K, Ishii S (2015) Expanding histogram of colors with gridding to improve tracking accuracy. In: 2015 14th IAPR International Conference on Machine Vision Applications (MVA). IEEE, pp 475–479
29. Mindru F, Tuytelaars T, Van Gool L, Moons T (2004) Moment invariants for recognition under changing viewpoint and illumination. *Comput Vis Image Underst* 94(1):3–27
30. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intell* 24(7):971–987. <https://doi.org/10.1109/TPAMI.2002.1017623>
31. Ojansivu V, Heikkilä J (2008) Blur insensitive texture classification using local phase quantization. In: Pedone national conference on image and signal processing. Springer, pp 236–243
32. Pedone M, Heikkilä J (2012) Local phase quantization descriptors for blur robust and illumination invariant recognition of color textures. In: 2012 21st International Conference on Pattern Recognition (ICPR), pp 2476–2479
33. Riess C, Angelopoulou E (2010) Scene illumination as an indicator of image manipulation. In: Information hiding, vol 6387, pp 66–80
34. Roemer J, Groman M, Yang Z, Wang Y, Tan CC, Mi N (2014) Improving virtual machine migration via deduplication. In: 2014 IEEE 11th International Conference on Mobile ad hoc and Sensor Systems (MASS). IEEE, pp 702–707
35. Rubner Y, Tomasi C, Guibas LJ (2000) The earth mover’s distance as a metric for image retrieval. *Int J Comput Vis* 40(2):99–121
36. van de Sande K, Gevers T, Snoek C (2010) Evaluating color descriptors for object and scene recognition. *IEEE Trans Pattern Anal Mach Intell* 32(9):1582–1596. <https://doi.org/10.1109/TPAMI.2009.154>

37. van de Sande KEA, Gevers T, Snoek CGM (2011) Empowering visual categorization with the gpu. *IEEE Trans Multimedia* 13(1):60–70. <http://www.science.uva.nl/research/publications/2011/vandeSandeITM2011>
38. Smith AR (1978) Color gamut transform pairs. *ACM Siggraph Comput Graph* 12(3):12–19
39. Steiger JH, Shapiro A, Browne MW (1985) On the multivariate asymptotic distribution of sequential chi-square statistics. *Psychometrika* 50(3):253–263
40. Ojala T, Pietikäinen M, Mäenpää T (2001) A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification. In: *Advances in Pattern Recognition, ICAPR 2001 Proceedings, Lecture Notes in Computer Science 2013*. Springer, pp 397–406
41. Tai J, Liu D, Yang Z, Zhu X, Lo J, Mi N (2017) Improving flash resource utilization at minimal management cost in virtualized flash-based storage systems. *IEEE Trans Cloud Comput* 5(3):537–549
42. Tan RT, Nishino K, Ikeuchi K (2004) Color constancy through inverse-intensity chromaticity space. *J Opt Soc Am A* 21(3):321–34
43. Tang X (1998) Texture information in run-length matrices. *IEEE Trans Image Process* 7(11):1602–1609. <https://doi.org/10.1109/83.725367>
44. Van De Weijer J, Gevers T, Gijzenij A (2007) Edge-based color constancy. *IEEE Trans Image Process* 16(9):2207–14
45. Vedaldi A, Lenc K (2015) Matconvnet: convolutional neural networks for matlab. In: *Proceedings of the 23rd ACM international conference on multimedia*. ACM, pp 689–692
46. Vidyadharan DS, Thampi SM (2015) Brightness distribution based image tampering detection. In: *2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)*. IEEE, pp 1–5
47. Vidyadharan DS, Thampi SM (2015) Detecting spliced face in a group photo using pca. In: *2015 7th International Conference of Soft Computing and Pattern Recognition (SOCPAR)*. IEEE, pp 175–180
48. Wei X (2007) Gray level run length matrix toolbox v1.0. Software Beijing Aeronautical Technology Research Center. <http://www.mathworks.com/matlabcentral/fileexchange/download.do?objectId=17482&fn=RunLengthMatrixToolboxver1.0&fe=.zip&cid=1101680>
49. Van de Weijer J, Gevers T, Bagdanov AD (2006) Boosting color saliency in image feature detection. *IEEE Trans Pattern Anal Mach Intell* 28(1):150–156
50. van de Weijer J, Schmid C, Verbeek J, Larlus D (2009) Learning color names for real-world applications. *IEEE Trans Image Process* 18(7):1512–1523. <https://doi.org/10.1109/TIP.2009.2019809>
51. Wu X, Fang Z (2011) Image splicing detection using illuminant color inconsistency. In: *2011 third International Conference on Multimedia Information Networking and Security (MINES)*. IEEE, pp 600–03
52. Yan-li H, Shao-Zhang N, Jian-Cheng Z, Lin-Na Z (2014) Forensics of image tampering based on the consistency of illuminant chromaticity. In: *2014 Annual Summit and Conference on Asia-Pacific Signal and Information Processing Association (APSIPA)*. IEEE, pp 1–4
53. Yang Z, Awasthi M, Ghosh M, Mi N (2016) A fresh perspective on total cost of ownership models for flash storage in datacenters. In: *2016 IEEE International Conference on Cloud Computing Technology and Science (CLOUDCOM)*. IEEE, pp 245–252
54. Yang Z, Tai J, Bhimani J, Wang J, Mi N, Sheng B (2016) Grem: dynamic ssd resource allocation in virtualized storage systems with heterogeneous io workloads. In: *2016 IEEE 35th International on Performance Computing and Communications Conference (IPCCC)*. IEEE, pp 1–8
55. Zhang L, Zhou Z, Li H (2012) Binary gabor pattern: an efficient and robust descriptor for texture classification. In: *2012 19th IEEE international conference on image processing*, pp 81–84. <https://doi.org/10.1109/ICIP.2012.6466800>
56. Zhou P, Han X, Morariu VI, Davis LS (2017) Two-stream neural networks for tampered face detection. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, pp 1831–1839



**Divya S. Vidyadharan** received her B.Tech. degree in Computer Science and Engineering from Sree Chitra Thirunal College of Engineering, Trivandrum in 2000 and M.Tech. degree in Computer Science, with specialization in Digital Image Computing, from the Department of Computer Science, University of Kerala in 2004. She is currently a doctoral research scholar at the Department of Computer Science and Engineering at College of Engineering, Kerala, India. Her research interests include digital forensics, doctored image detection, and forgery localization.



**Sabu M. Thampi** is an Associate Professor at the Indian Institute of Information Technology and Management, Kerala (IIITM-K), Trivandrum, India. He has completed his Ph.D in computer engineering from the National Institute of Technology, Karnataka. His research interests include network security, security informatics, bio-inspired computing, video surveillance, cloud security, secure information sharing, secure localization, and distributed computing. He has authored and edited a number of books published by reputed international publishers and published papers in academic journals and international and national proceedings. He has served as a guest editor for special issues in many international journals and as a program committee member for many international conferences and workshops.