

# Salient Bag of Feature for Skin Lesion Recognition

Pawan Kumar Upadhyay\* and Satish Chandra

*Department of Computer Science and Engineering, IIIT, Noida, 201301, India  
Jaypee Institute of Information Technology, Noida, 201309, India*

---

## Abstract

With the rapidly increasing incidence of various types of skin cancer, there is a need for decision support systems to detect abnormalities in the early stages and help reduce the mortality rate. Several computer-aided diagnosis (CAD) systems have been proposed in the last two decades for skin melanoma recognition. Continuous improvements have been made in the accuracy of melanoma diagnosis, but other classes of cancer, such as basal cell carcinoma and squamous cell carcinoma, are not very intact with the non-invasive diagnosis system. In this paper, a generic method of diagnostic system is proposed and is viable to classify the ten classes of a skin lesion. These lesion classes belong to cancer, pre-cancerous, and tumor categories of samples, as shown in a gold standard image dataset. The key idea of the proposed approach is to optimize the bag-of-SURF features by the non-linear Hessian matrix of HSV color descriptors. These features are combined to form a salient bag-of-features, which helps recognize the skin lesion classes more accurately. Experimental results show that the proposed method of skin lesion diagnosis significantly improves the accuracy of recognition up to 89% as compared to the current state-of-the-art accuracy of 81.8%. It does not require any complex pre-processing of images, which affects the performance of the recognition system.

**Keywords:** computer-aided diagnosis (CAD); bag of visual words (BoVW); speeded-up robust features (SURF); hue saturation value (HSV); bag of color (BOC); bag of robust features (BORF); support vector machine (SVM)

(Submitted on July 10, 2018; Revised on November 10, 2018; Accepted on March 15, 2019)

© 2019 Totem Publisher, Inc. All rights reserved.

---

## 1. Introduction

Dermoscopic skin lesion screening is a non-invasive imaging method that facilitates the detection of severe abnormalities such as cancer in the categories of cancer, pre-cancerous, and tumor. The most severe pigmented lesion is melanoma because of its propensity to metastasis, and it requires early diagnosis. The diagnosis of a pigmented skin cancer lesion such as melanoma follows certain standard rules of diagnostic in dermoscopy. The fundamental rule is ABCD (asymmetry, border, shape, color aspects, and structural dissimilarities on the lesion) for pigmented skin lesions [1]. In addition, there are standard methods used for melanoma detection such as CASH [2], Menzies [1, 3], 3-points [4], and 7-points checklist [2, 4]. They help classify the skin melanoma as malignant or benign. However, these methods are available only for melanoma diagnosis, and non-melanoma skin cancer recognition is still an untouched area for computer-aided diagnosis systems (CAD).

Non-invasive methods of radiology provide great support to medical experts. Moreover, radiology involves the visualization of abnormal regions in various layers of skin using image-guided methods. To improve the accuracy of computer-aided diagnosis in the domain of dermoscopy, the proposed algorithm provides more detailed information of skin lesions by the means of bag-of-features, and it is able to discriminate the skin lesion classes more accurately. The input image samples are analyzed in the image space with their local intensities. These intensities are characteristic features at standard points such as edges, corners, and regions, which are considered to be key points. These key points as low-level features are detected due to directional changes in intensity at each image pixel. For a certain region, changes in intensities are constant, and they are captured by the box or boundary box method of SURF descriptors. These regions are called blobs and represent constant features such as texture and color. The size of the blob increases with the scale space.

---

\* Corresponding author.

E-mail address: [pawan.upadhyay@iiit.ac.in](mailto:pawan.upadhyay@iiit.ac.in)

Minimizing this gap between high-level representation (as interpreted by medical experts) and low-level depiction (detected by CAD) is the key motivation of research in the Vivo clinical examination and has proven to be useful for skin lesion detection. The major challenge is to discriminate between the samples of skin cancer lesions with their associated precancerous and tumor class samples. The example of connected class labels is actinic keratosis caused by UV radiation, and it progresses to squamous cell carcinoma [5]. The disease evolution depends on a number of clinical features, such as bleeding and having a diameter that is greater than 1cm and rapidly enlarging in all direction [3, 5]. The irregular growth of actinic keratosis is a major cause of non-melanoma skin cancers, and these are the predecessors for squamous cell carcinoma [3, 6]. The various features vector and its directional growth are symbolized by the descriptor BoVW [7-8]. The proposed method is inspired from bag-of-visual-words. Most of the existing work on BoVW is directed towards melanoma classification. The usefulness of this method for skin lesion recognition is the description of an image locally by a unique feature vector that corresponds to each class label of it [7, 9-11]. It also describes image representation as a histogram, based on the occurrences of visual words, and generates a codebook. In this approach, the generation of vocabulary in quantized scale space is represented by the fusion of SURF and the color descriptor. In addition to this, both the descriptors use similar methods for feature detection, which are based on the image integral property and box filtering.

## 2. Related Works

Medical images usually show complex feature space distributions due to intra-class similarity and inter-class ambiguity. These aspects can be handled nicely by the feature descriptor, which provides the complete (per pixel) information of an image. In addition to this, feature extraction has been accomplished through texture, color, and shape descriptors. The most popular color descriptors are color moments and color histograms [7, 10], which are used to discriminate between the lesion patterns. The next most important descriptors are textural, having various categories such as Haar, Gabor filter, wavelet, and gray level co-occurrence matrix (GLCM). The vital descriptors are shaped and used in the active shape model (ASM) for an active instance. This low-level feature representation is ground information for predicting the class label, but processing it is tedious.

To enhance the processing, we adopt the mid-level features for histogram-based representation of a feature vector as visual words. The detailed features space is formed by transforming the low-level feature vector into a global one, commonly known as mid-level features [12]. In addition, the most common mid-level representation is the bag-of-visual-words approach (BoVW). It is used for binary classification of skin melanoma with non-melanoma image samples and has some gaps related to the number of clusters and the size of the visual vocabulary [13].

Research is primarily on melanocytic lesions (i.e., malignant melanoma), and non-melanoma classes of cancer recognition need to be further explored. SURF features are used to discriminate between the various patterns of melanoma [12]. The present findings seem to be consistent with current research in the domain of dermoscopy. The dermoscopic image analysis follows three major steps: (1) border identification and detection [11, 14-15], (2) extraction of image features from the lesion [16-18], and (3) comparison of these feature with pre-calculated features of each class of skin lesions and performance of skin lesion classification [19]. In order to obtain better classification accuracy, we target the methods that reduce the first and second steps to perform the third step.

Image classification extensively depends upon lesion localization and detection and can be performed by bag-of-features. Similar to this approach, there are various other methods used for object categorization such as pLSA [20] and latent Dirichlet allocation [21-23]. In these approaches, image content features that are semantically pertinent to the user's perception are considered. In order to do so, semantic-based image retrieval methods are performed in two phases. In the first phase, the subsequent image is divided into regions by a fixed grid size similar to image segmentation, which helps identify the key points from the image space. In the second phase, there is an association of semantic labels with image regions [20, 23-24]. In succession to this, semantic-based image retrieval is used to extract two groups of features: local and global. Global features are commonly used for skin lesion classification including shape, color, and texture descriptors of a lesion, and these features have been used with great success in previous works [11, 18].

There are certain challenges left as a local feature, which can be handled nicely by the proposed method for skin lesion recognition. In addition, we are considering SURF and the color descriptor for describing the region of interest as lesions. The ROI for each class label is represented with gray level blobs, which are obtained from the concept of the determinant of the Hessian matrix. The scale-space blob tree represents the lesions of various sizes with similar visual cues for each class label, as described in a further section. As we have mentioned, most research focuses on melanoma classification rather than "all class" dataset, which includes cancer, precancerous, tumor, and nevus (mole) skin lesions. Although there are several works on skin lesion classification, they are not compared directly due to the adoption of different datasets and dissimilar validation measures.

### 3. Methodology

In this section, we describe the proposed method of the skin lesion recognition system. It relies on two feature descriptors, where one describes the shape of the skin lesion with scale and rotational invariance properties and the other shows the morphological information of melanin in dermoscopy images. The subsequent process is divided into the following steps: key patch identification, dermoscopic features extraction, bag-of-features fusion, and codebook generation. The classification is shown in Figure 1.

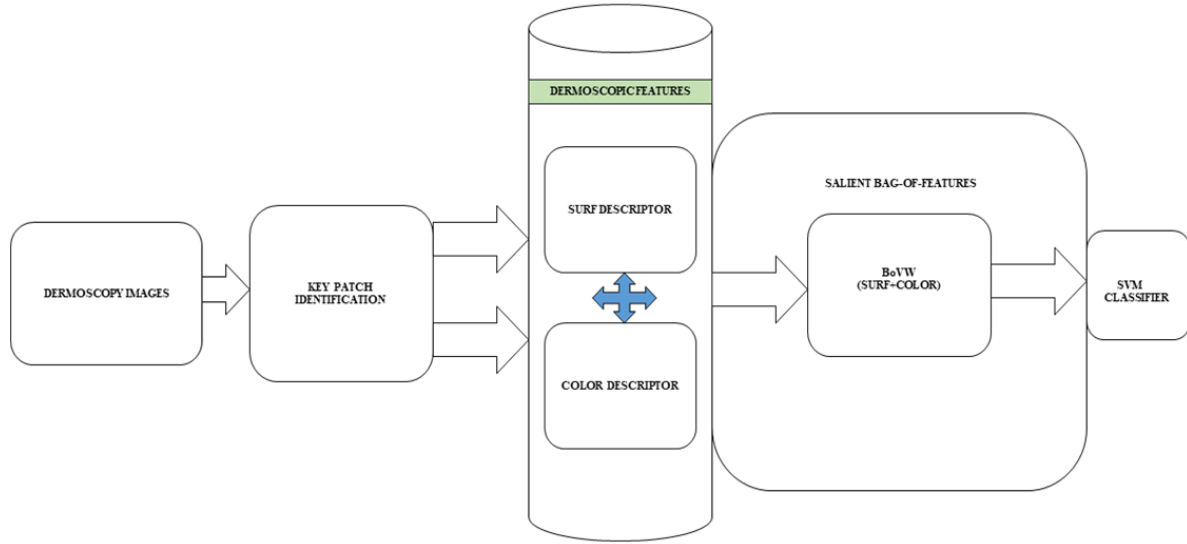


Figure 1. Skin lesion recognition system

#### 3.1. Key Patch Identification

For a given integral image, we first identify the number of points ( $N$ ) in the dermoscopic image and then succinct to appropriate key points having similar attributes. These are known as a patch, and the number of patches in the image is

$$P = [P_1, P_2, P_3, P_4, \dots, P_N] \quad (1)$$

Each patch represents

$$P = [0, 0, 0, 1, \dots, 1] \quad (2)$$

The key patches in the dermoscopy images represent different lesion structures due to the uniqueness of the features vectors in the skin lesion area, and these structures are called blobs.

#### 3.2. Dermoscopic Features Extraction

Bag-of-features consists of the unique and compact summarization of local features. In this section, a numeric detector and descriptor called SURF for local image features (commonly known as interest points, key points, and salient features) can be used to characterize a definite pattern of skin lesions having a distinctness from its immediately close pixels. It is generally correlated with an image property (edge, corner, and region) as a region. The unique features obtained from this region of the grayscale image are concatenated with a bag of color and form a salient bag-of-features. The salient bag-of-features gives the optimal result for skin lesion classification using an SVM classifier having 10 class samples of skin lesions, which majorly belong to cancerous, precancerous, and tumor.

##### 3.2.1. Speedup Robust Feature Descriptor

A SURF descriptor is used to describe the detected patches as blobs in section 3.1. The identified region in an image space describes unique features, which are the scale and rotational invariant. They are key dermoscopic features that describe the lesion as a region of circular shape having different attributes such as shape and texture. In addition to this, most of the

lesion patterns are an irregular shape, so the direction of growth of the lesion must be identified by the gradient method. This can be performed in three steps: (1) the key patch in a blob of circular region, (2) the size of the blob describes the density of the lesion, and (3) the size of the blob increases at different scales of a scale space called the multi-scale space.

**Radial blob detection method:** The blobs in dermoscopy images are computed by selecting the centre at some location  $(n, n)$  and sampling these points at certain pixel intensities in each class of skin lesions. Points in different boxes are connected to the centre pixel and marked in a radial shape. The radius of the radial space is described by Equation (3).

$$R = L/\sigma \quad (3)$$

Where  $R$  is radius,  $\sigma$  is scale space, and  $L$  is number of elements in each direction from the centre.

**Lesion density vs. blob size:** The size of the blob is a crucial parameter, and it depends on the actual size of the Gaussian kernel that is convolving with an input image of a skin lesion. If the minimum scale space is  $4 \times 4$  and is convolving with a  $5 \times 5$  filter, 20 grids are required to generate the fastest Hessian matrix for blob detection. This circular shape of the descriptor is SURF, and the elements of matrix  $(L_{xx}, L_{xy}, L_{xy}, L_{yy})$  are second order derivatives. It is commonly known as a Hessian matrix. This element of the Hessian matrix represents the growth of the skin lesion in each direction. The responses of each blob are four Haar features in radial space and are represented as a Hessian matrix, as shown in Figure 2.

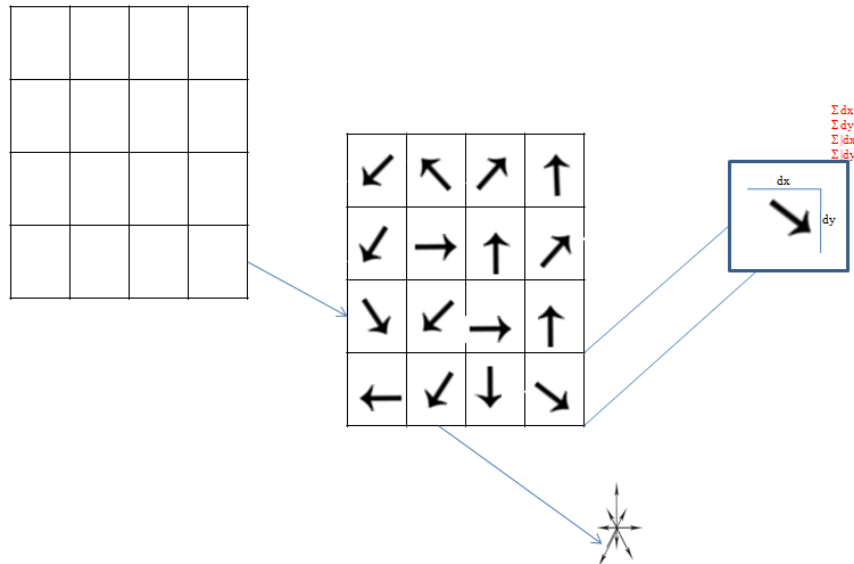


Figure 2. Key points detector using SURF [Adopted from Bay et al.]

The Hessian matrix  $H(X, \sigma)$  at point  $X = (x, y)$  in an image is expressed as

$$H(X, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix} \quad (4)$$

Where  $\sigma$  is the scale factor.

In the SURF descriptor, local features are computed from the image using box filter and its integral property.

**Bag of multi-scale speedup robust features:** The salient features for each class of skin lesions needs to be found at various scales. They are generally implemented as an image pyramid. It consists of a repetitive convolution of the same image with a filter of increasing size and can be represented as

$$\text{Binomial filter approximation} = \sqrt{N-1}/2 \quad (5)$$

Where  $N$  = block width + 1.

Due to its image integral property, it helps to up-scale the size of the filter in an order of 2 and generate pair-wise elements, which represent the spatial distribution of the gradient and provide information about each pixel in the dermoscopy image. In order to see the reproducible orientation of the lesion feature, calculate the Haar wavelet responses in the direction  $(x, y)$  within the circular neighborhood of radius of  $R$  around the key points, as shown in Equation (2). The calculated blob responses of a skin lesion in the horizontal and vertical directions are strengthened by its vector length of paired elements by some scale in image space.

The creation of the scale space begins at 32, and the size of the filters increases to 64 ( $32 \times 2$ ), 96 ( $32 \times 3$ ), and 128 ( $32 \times 4$ ). The use of integral images permits the up-scaling of the filter at a constant interval from 32 to 128. The selection of accurate scale space for the blobs is computed from the local maxima of the determinant of Hessian (DOH), which signifies the dominance change in skin lesion mass at the different scales. Investigate the maximum change of pixel intensity in all four directions, and segment the lesion from the input image. In addition, this method helps achieve maximum interclass variability at the sub-region space of  $(4 \times 4)$  with a grid step size of 64. SURF descriptors are computationally faster than any other spectra descriptors.

They are used to describe the distribution of each class of skin lesions at a certain scale space and compute the surface curvature of each lesion pattern. The distribution of each class label is dependent on certain parameters such as the Hessian threshold and location of strongest features at different scales, as described by Equation (3). It also reduces the significance of individual pixels, as discussed in a further section. Color features are equally significant with SURF features for skin lesion classification. In order to detection skin lesion classes, these two features are required for skin lesion classification.

### 3.2.2. Color Feature Extraction

As an extension to the previous section, SURF features are extracted from the grayscale version of color dermoscopy images and are unable to preserve the chromatic saliency (distinctiveness in chromatic regions). To converge the chrominance attribute of image pixels in isoluminance space, it is required to map the RGB color space model to HSV and explicitly represent the color saliency in multi-scale space for this system.

## 4. Color Hessian Matrix and Salient Bag-of-Features

A method is proposed in which SURF bag-of-features are boosted with a color feature vector based on the HSV color model. The saliency in an improved bag-of-features for skin lesion recognition is performed in three steps: (1) **feature detector** using a non-linear tensor matrix for the HSV color model, (2) **feature descriptor**: representation and fusion of a multi-scale bag of color and SURF features, and (3) **skin lesion recognition using SVM**: classification of all ten classes of skin lesions using an SVM classifier with radial basis function.

### 4.1. Feature Detector: Color Hessian Matrix

Let  $C$  be the color image having primary colors  $(R_x, G_y, B_z)$  as information, and its spatial information derivatives chrominance and luminance compute from the Hessian matrix similar to the extraction method used for the SURF feature vector. Extending the luminance to color signals in the direction of Haar features generates optimized feature vectors with all the characteristic properties to categorize the skin classes. In this paper, a linear tensor for the RGB color model transforms to a non-linear tensor based on the HSV color model, and a matrix called the Hessian matrix is developed. The reason for conversion is to extract chrominance features at the curvature point from each class of skin lesions.

In the non-linear tensor, the linear values of three bands of color are mapped to the vector space  $(V_x, V_y, V_z)$  with polar coordinate  $V(\Theta, \emptyset)$  and  $V \in [0, 1]$ . The obtained values of color from each of the samples are the new values as  $R', G'$ , and  $B'$ .

The procedure for non-linear tensors of the HSV color model is described below:

**Step 1** Normalize the  $R' = R/255$ ,  $G' = G/255$ ,  $B' = B/255$ , and  $(R', G', B') \in [0, 1]$  scale.

**Step 2** Compute the hue that represents the color component at each of the pixel  $(x, y)$  in the dermoscopic image. It is in the range of minimum  $(\emptyset_1 \geq 0^\circ)$  and maximum  $(\emptyset_2 \leq 360^\circ)$ .

**Step 3** The saturation value is  $S = |\emptyset_2 - \emptyset_1|$ , the level of chrominance describes the color components, and the change of second order gradient at each of the pixels is  $(C_{xx}, C_{xy}, C_{yx}, C_{yy})$ . The Hessian matrix  $D(X, \sigma)$  at point  $X = (x, y)$  in an image is expressed as

$$D(X, \sigma) = \begin{bmatrix} C_{xx}(X, \sigma) & C_{xy}(X, \sigma) \\ C_{xy}(X, \sigma) & C_{yy}(X, \sigma) \end{bmatrix} \quad (6)$$

Where  $\sigma$  is the scale factor.

This definition is simply an extension of the second order moment matrix. To color, we compute the color blob detector similar to the Hessian matrix. Two parameters,  $\alpha$  and threshold, are described below in Equation (7).

$$\det(D) - \alpha \text{trace}^2(D) > \text{Threshold} \quad (7)$$

To extend this for multi-scale,

$$\text{color LOG}(\sigma) = \sigma^2 (C_{xx} + C_{yy}) \quad (8)$$

The color hessian is

$$\text{Color hessian}(\sigma) = \sigma^2 \|C_{xx} C_{yy} - C_{xy}^2\| \quad (9)$$

Where  $C$  is the color image,  $C_{xx}$ ,  $C_{yy}$ , and  $C_{xy}$  are the second order derivatives, and  $\|\cdot\|$  is a vector norm. This extension leads to a scale-space representation of color feature vectors.

In the next section, we combine the color bag-of-features with the SURF feature bag. The robust bag of features is boosted by the contribution of a color descriptor. The level of the color component is the chrominance weight vector, which helps improve the rate of information captured from each single unit of the skin lesion in image space. The detail information helps increase the accuracy of the lesion classification system.

#### 4.2. Feature Descriptor: Salient Bag-of-features

Dermoscopic key features are represented and fused as a bag of SURF and color. The color adds saliency in robust feature bags and helps develop an improved bag, which is appropriate for dermoscopy skin lesion recognition. The salient bag-of-features consists of the intersection of union between two features vector obtained from the samples of skin lesions. A bag of salient features form a vocabulary  $g = \{g_1, \dots, g_n\}$  of SURF and color feature vector. The fused vector helps select the input image from the subset of the standard image database and generate a matching score. SURF is the most robust feature descriptor with respect to geometrical changes, and its vocabulary is improvised with the added weight of color features.

The following steps describe the generation of proposed bag-of-features:

**Step 1** Divide the dermoscopy image into a set of independent local descriptors extracted from the patches (small image region).

**Step 2** The extracted descriptors from each training sample are SURF and color feature vectors, which are transformed as a histogram.

**Step 3** The color features demonstrate a great impact on textural SURF features because of the similar procedure of feature detection.

**Step 4** The descriptors are computed in two steps. The first is feature detection using the Hessian matrix, and its determinant generates blobs. The second is the added descriptor of color, which gives various shades of skin lesions along with the shape and describes each class of skin lesions more clearly.

**Step 5** Concatenation of the color and SURF features at each pixel location are performed by  $k$ -mean clustering, and visual vocabulary is constructed.

**Step 6** The number of strong features vectors, which are extracted from each class of skin lesions, and the number of sample selection are based on the class that has the minimum number of samples.

**Step 7** The visual codebook is a decision box for the input image, and it is assigns the image to the closest class of skin

lesions in a visual vocabulary based on the obtained classifier score.

The skin lesion pigmentation, along with shape, is an important factor for its recognition. These steps signify that the color adds saliency in the standard bag-of-SURF features. In the next section, the SVM classifier is employed for the classification steps and the size of the vocabulary is set to 700. In order to see the results of this approach, we use dermoscopy data that has ten classes of lesions.

#### 4.3. Skin Lesion Recognition using Cascaded SVM Classifier

In the previous section, Blob identification is based on their pixels frequencies as region of interest (ROI). The size of the blob depends upon the size of the Gaussian kernel which is used to compute the Hessian matrix. We are capturing the centered Haar feature and color features at multi-scale space with a grid width of [32 64 128]. The cascaded features at different scale space are stored in a form grid and size of the grid step is 64.

Pixels frequencies in the skin lesion area are represented as a non-linear basis set functions having Haar and chrominance feature vectors as a variable, obtained from the SURF and color descriptor. Each region of the image is searched for well-defined Haar and color features using a sliding sector window method, and the grid step is chosen at an interval of 64 (such as for every 64 pixels). The detected features are fed into a channel of classification funnel after dividing the data samples into parts known as a Haar cascaded classifier, as shown in Figure 3. The strongest Haar features are on the top of the funnel and have a low rate of false positives and false negatives, so the first order derivatives of the cascaded classifier contain max-probability regions of the image sample for further analysis. Haar is a wavelet-based feature and becomes more complex as we progress deeper into the funnel of the cascaded classifier, so the features at the bottom of the funnel will spill out. For classification, the Haar features and the dominance color features of pixels are important. In addition to this, another important factor in relation to the lesion classification is pixel linear, and angular distance measure is also considered by these two feature bags. For the proposed method, we select 70% of strong features from the top of the funnel, which helps to reveal the optimal results for skin lesion classification. The feature vectors obtained from two the image samples are passed to the SVM classifier with RBF kernel for the classification system and are used to classify the skin lesions more accurately.

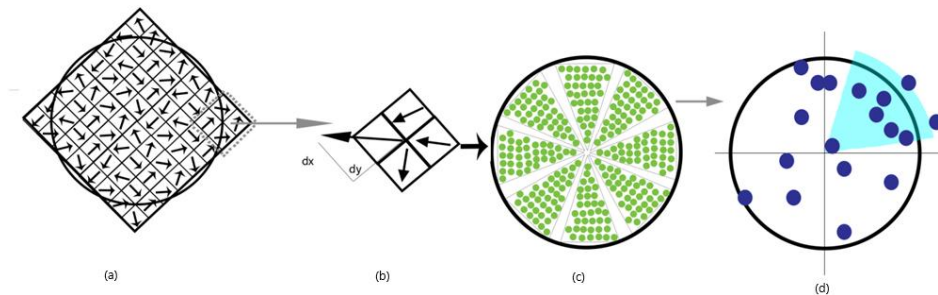


Figure 3. Centre Haar features grows at each level of scale space: (a) from  $2 \times 2$  to  $4 \times 4$  grid space (i.e., 4 to 16 grids) in circular region filter; (b) directional change in  $(x, y)$ ; (c) funnel according to the eight directions as shown in Figure 1 and connect the first element of each funnel to give the circular space of the object, which grows subsequently; (d) all the features in the funnels follow binomial filter approximation (Equation (5)), and cascaded features feed to the classifier [Adopted from Bay et al.]

When we match the patterns among two images, they have similar attributes (such as texture and shape) due to the constant pixel intensities inside the radial region and the obtained score of classification deciding the actual class of sample images.

#### Dataset

The image dataset represents "all class data" and has samples of nevi (mole), tumor, pre-cancerous, and cancerous. The adopted database of dermoscopic images is gold standard (verified by dermatopathologists and dermatologists) and is under a license agreement (reference number: EC2813) with Edinburgh Research and Innovation [13].

This dataset has 1300 image sample with a pixel resolution of 0.03mm. These samples are divided into ten classes, and the number of samples in each class is as follows: (a) Actinic keratosis (AK): 45, (b) Basal cell carcinoma (BCC): 239, (c) Melanocytic nevus (ME): 331, (d) Seborrhoea keratosis (SK): 257, (e) Squamous cell carcinoma (SCC): 88, (f) Intraepithelial carcinoma (IEC): 78, (g) Pyogenic granuloma (PYO): 24, (h) Haemangioma (HAEM): 97, (i)

Dermatofibroma (DF): **65**, and (j) Malignant melanoma (ML): **76**. Pre-process the images and keep the size of each image sample at  $227 \times 227$  for validation of the proposed method.

## 5. Results and Discussions

The novel method of skin lesion classification does not require any kind of pre-processing such as lesion segmentation and lesion border detection, and it is directly applied to homogeneous image bundles for feature extraction and classification. The related experiments are performed on the platform MATLAB-2014b with a system configuration of Dual Core I-5 Processor, 8GB RAM, GPU of 920M series, and size of 2GB.

**Parametric evaluation for salient bag-of-features:** It primarily performs the extraction of SURF (Speeded Up Robust Features). It involves invariant to image rotation and faster generation of a bag of robust features (BoRF) and optimizes by adding chrominance feature weights obtained from bag-of-color features (BOC). To reduce the computation time and increase the classification accuracy, the following measures are considered: (a) number of clusters, (b) number of features, and (c) number of iterations or time to converge. By considering the measure as described in Table 1, an efficient visual codebook can be generated.

Table 1. Parameters for visual codebook that are obtained from 1300 image samples of 10 class with grid step sizes of 32, 64, and 128: (1) No. of clusters, (2) No. of features, (3) No. of iteration and convergence time of each iteration

S.No	Parameters	Grid32	Grid64	Grid128
1.	No of features	55989	17280	17280
2.	No of iterations	34	31	31
3.	Converge time (seconds/iteration)	0.23	0.11	0.11

In addition to this, obtained feature vectors are scaled down to strong feature vector space using *k*-mean clustering applied with the *k*-d tree method, and the results are described in Table 1. Moreover, various other measures such as the block width of the grid help decide the number of course features for each class of skin lesion available in the image database, as shown in Table 2. Extracted feature vectors are proportionate with a number of samples of each class at a grid step size of 64. Feature vectors are extracted using the box filtering method of integral images, in which the size of the singleton sub-region is  $4 \times 4$  and the directions are described as (*dx*, *dy*, *|dx|*, and *|dy|*) direction points. Thus,  $4 \times 4 \times 4$ , or 64 (size of the grid step), is a more stable representation to generate proportionate feature values (i.e., 45 samples of AK have 4500), as shown in Table 2. The bag of visual words consists of 80% of the strong features (SURF) to classify skin lesions individually or as connected pairs of classes.

Classification is performed by the SVM classifier. By keeping the training dataset of 70% and a testing dataset of 30%, skin lesion classification is performed. If pyogenic granuloma (PYO) is a class of minimum sample, then the considered dataset split performed as a 190 sample for training and 50 for testing. It generates the most accurate prediction of label classes, which are described in the confusion matrix with an average accuracy of 89%, as shown in Table 3. The confusion matrix adopts a one versus all classification approach for 10 class data of skin lesions.

Table 2. SURF extracted features from each class of skin lesion

S.No	Dermoscopic sample	#Samples	#SURF32	#SURF64	#SURF128
1.	Actinic Keratosis	45	14580	4500	720
2.	Basal Cell Carcinoma	239	71826	23900	3824
3.	Melanocytic Nevus (mole)	331	107244	33100	5296
4.	Seborrheic Keratosis	257	74520	25700	3680
5.	Squamous Cell Carcinoma	88	125628	8800	1392
6.	Intraepithelial Carcinoma	78	24948	7800	1232
7.	Pyogenic Granuloma	24	7776	2400	384
8.	Haemangioma	97	31104	9700	1536
9.	Dermatofibroma	65	21060	6500	1040
10.	Malignant Melanoma	76	24624	7600	1216

The results imitate the importance of low-level features, which are used to construct a visual vocabulary for skin lesion classification. Individual class level as well as connected pair (pre-cancerous and cancerous) classification is effectively

computed by the proposed method of salient bag-of-features.

Table 3. Confusion matrix (average accuracy: 89%)

TRUE	AK	DF	HAEM	IEC	ME	ML	PYO	SCC	SK	BCC
AK	0.98	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00
DF	0.00	0.92	0.03	0.03	0.00	0.02	0.00	0.00	0.00	0.00
HAEM	0.04	0.00	0.83	0.04	0.01	0.05	0.02	0.00	0.01	0.00
IEC	0.03	0.01	0.01	0.91	0.00	0.01	0.01	0.01	0.01	0.00
ME	0.05	0.03	0.02	0.02	0.72	0.01	0.08	0.05	0.01	0.01
ML	0.00	0.04	0.03	0.05	0.00	0.86	0.01	0.00	0.01	0.00
PYO	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
SCC	0.06	0.00	0.02	0.01	0.00	0.00	0.00	0.91	0.00	0.01
SK	0.10	0.02	0.04	0.04	0.17	0.02	0.04	0.07	0.49	0.01
BCC	0.05	0.06	0.03	0.05	0.02	0.10	0.03	0.02	0.01	0.63

**Performance analysis of connected class samples as precancerous and cancerous:** The proposed method is able to classify the most connected lesion classes such as actinic keratosis (**AK**), which is pre-cancerous, grows abruptly, and develops into the most severe class of cancer, squamous cell carcinoma (**SCC**). In addition to this, the prediction average accuracy after ten trials for AK is 98% and SCC is 91%, as described above in **Table 3**. Moreover, normal mole (ME) and malignant melanoma (ML) are also predicted nicely with an average accuracy of 72% and 86%, respectively. It is difficult to predict the basal cell carcinoma (BCC) cancer class, which is connected with an available class of tumor called SK or pre-cancerous class AK. The obtained accuracy of third connected classes is less than 80%, although remaining classes have an accuracy of greater than 80%. The SURF feature-based vocabulary method gives promising results for all class of abnormalities occurring in melanocytic, keratosis, and carcinoma cells of the skin.

The results obtained from the proposed method are optimal compared to previous state-of-the-art methods for the 10 class dermoFit dataset. The paired class label as precancerous-cancerous can be easily classified with proposed method and achieves consistent accuracy. It also describes the structure similarity with high covariant feature vectors among each class sample. This approach is quite useful if the samples of connected classes belong to a single entity (patient). Our proposed approach is able to classify the samples of melanocytic cells such as ME (nevus) and ML (malignant melanoma) in increasing order, i.e., (72% to 86%) shows that recognition is improving related to disease prediction. Meanwhile, for non-melanoma-connected classes as actinic keratosis (**AK**) and squamous cell carcinoma (**SCC**), the average accuracy is slightly reduced from 98% to 91% but can be acceptable because the change is more than 5%. In addition to this, it is found that the proposed technique outperforms the others, not only for individual class labels but also for connected classes.

**Comparison of improved bag-of-features with BORF and BOC:** The result reveals that the proposed bag-of-features is more densely bounded. In addition to this, it also helps mark the dermoscopic key points more accurately as compared to the standard bag features. The comparison results of skin lesion recognition are validated with the following performance measure: sensitivity, specificity, and accuracy, as shown in the Table 4.

$$\text{Sensitivity} = \text{TP}/(\text{TP}+\text{FN}) \quad (10)$$

$$\text{Specificity} = \text{TN}/(\text{FP}+\text{TN}) \quad (11)$$

$$\text{Accuracy} = (\text{TP}+\text{TN})/(\text{TP}+\text{TN}+\text{FP}+\text{FN}) \quad (12)$$

Table 4. Performance of bag-of-features using SVM classifier

S.No	Bag-of-features/Measures	Sensitivity	Specificity	Accuracy (%)
1.	BORF	86	85	84.5
2.	BOC	82	79	80.1
3.	Proposed BAG	86	92	89.0

The accurate detection of skin lesion class is crucial for successful treatment. There are various methods for skin lesions recognition, but the proposed method has the potential of enhancing the current clinical paradigm in the domain of dermoscopy.

In our experiments, we used the pigmented image dataset of 1300 samples without the removal of poor quality images that have various artifacts such as hair or nails. We do not perform any ad-hoc preprocessing (except keeping the size of

image 227×227) to promote the classifier, which was proven to be the robust for the proposed bag-of-features. Our method is comparatively better than [7-8, 13, 19, 25], as shown in Table 5.

Table 5. Result reported in the literature: Direct comparison is not possible because of different dataset and validation measures

Ref.	Authors	Method	Dataset	Accuracy (%) / [Sensitivity (%), Specificity(%)]
11.	Ballerini et al. 2013	Color and texture, K-NN	5 class (705 images)	74
25.	Kawahar et al. 2016	Deep features, CNN	10 class (1300 images)	81.8
20.	Celebi et al. 2007	Color and texture, SVM	2 class (546 images)	[93.33, 92.34]
7.	Situ et al. 2008	Color histogram, Gabor filter, BoVW, K-NN	2 class (100 images)	82.21
8.	Barata et al. 2013	Color histogram, Gabor filter, BoVW, K-NN	2 class (176 images)	[93, 85]
-	<b>Proposed method</b>	<b>SURF and color descriptor, BoVW, SVM</b>	<b>10 class (1300 images)</b>	<b>89/[86, 92]</b>

## 6. Conclusions

In this paper, a new method of fused descriptor based on the color Hessian matrix is proposed for capturing more informative regions in a dermoscopy image space as compared to the luminance only detector using the SURF descriptor. The color boosted detector and luminosity detector have identical procedures using the Hessian matrix for feature detection. The structure similarity of the two descriptors is able to unite them into a bag of visual word (BoVW). The improved bag-of-features leads to a large number of histogram bins for each class of lesions. Due to this reason, intra-class similarity and inter-class variability can be easily achieved by the proposed bag of features. These feature bags are denser and significantly increase the discrimination of skin lesion classes. They assist the SVM classifier and help improve the average accuracy of skin lesion recognition.

## References

1. D. S. Rigel, J. Russak, and R. Friedman, "The Evolution of Melanoma Diagnosis: 25 Years Beyond the ABCDs," *CA: A Cancer Journal for Clinicians*, Vol. 60, No. 5, pp. 301-316, 2010
2. T. Wadhawan, N. Situ, H. Rui, K. Lancaster, X. Yuan, and G. Zouridakis, "Implementation of the 7-Point Checklist for Melanoma Detection on Smart Handheld Devices," in *Proceedings of 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3180-3183, 2011
3. R. Marks, "Epidemiology of Non-Melanoma Skin Cancer and Solar Keratosis in Australia: A Tale of Self-Immolation in Elysian Fields," *The Australasian Journal of Dermatology*, Vol. 38, pp. 26-29, June 1997
4. C. Barata, M. Ruela, M. Francisco, T. Mendonça, and J. S. Marques, "Two Systems for the Detection of Melanomas in Dermoscopy Images using Texture and Color Features," *IEEE Systems Journal*, Vol. 8, No. 3, pp. 965-979, 2014
5. P. J. Quaedvlieg, E. Tirisi, M. R. Thissen, and G. A. Krekels, "Actinic Keratosis: How to Differentiate the Good from the Bad Ones," *European Journal of Dermatology*, Vol. 16, No. 4, pp. 335-339, 2006
6. R. R. Neto and B. B. Yates, "Modern Information Retrieval," 1st Edition, Addison Wesley, 1999
7. N. Situ, X. Yuan, J. Chen, and G. Zouridakis, "Malignant Melanoma Detection by Bag-of-Features Classification," in *Proceedings of 2008 30<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3110-3113, 2008
8. C. Barata, J. Marques, and T. Mendonça, "Bag-of-Features Classification Model for the Diagnose of Melanoma in Dermoscopy Images using Color and Texture Descriptors," in *Proceedings of International Conference on Image Analysis and Recognition (ICIAR)*, pp. 547-555, 2013
9. J. Sivic and A. Zisserman, "Video Google: A Text Retrieval Approach to Object Matching in Videos," in *Proceedings of the Ninth IEEE International Conference on Computer Vision*, Vol. 2, 2013
10. Y. Boureau, F. Bach, Y. LeCun, and J. Ponce, "Learning Mid-Level Features for Recognition," in *Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2559-2566, 2010
11. H. Iyatomi, G. Schaefer, W. V. Stoecker, and M. E. Celebi, "Lesion Border Detection in Dermoscopy Images," *Computerized Medical Imaging and Graphics*, Vol. 33, No. 2, pp. 148-153, 2009
12. H. Bay, T. Tuytelaars, and L. V. Gool, "Speeded-Up Robust Features (SURF)," in *Proceedings of European Conference on Computer vision*, pp. 404-417, 2006
13. L. Ballerini, R. B. Fisher, B. Aldridge, and J. Rees, "A Color and Texture based Hierarchical K-NN Approach to the Classification of Non-Melanoma Skin Lesions," *Color Medical Image Analysis*, pp. 63-86, 2013
14. S. Gupta, S. K. Chakarvarti, and Zaheeruddin, "Medical Image Registration based on Fuzzy C-Means Clustering Segmentation Approach using SURF," *International Journal of Biomedical Engineering and Technology*, Vol. 20, No. 1, pp. 33-50, 2016
15. M. Rastgoo, R. Garcia, O. Morel, and F. Marzani, "Automatic Differentiation of Melanoma from Dysplastic Nevi," *Computerized Medical Imaging and Graphics*, Vol. 43, pp. 44-52, 2015
16. H. Bay, T. Tuytelaars, and L. V. Gool, "Speeded-up Robust Features (SURF)," *Computer Vision and Image Understanding*, Vol. 110, No. 3, pp. 346-359, June 2008
17. H. Iyatomi, H. Oka, M. E. Celebi, M. Hashimoto, M. Hagiwara, and M. T. K. Ogawa, "An Improved Internet-based Melanoma

- Screening System with Dermatologist-Like Tumor Area Extraction Algorithm,” *Computerized Medical Imaging and Graphics Journal*, Vol. 32, No. 7, pp. 566-579, 2008
18. M. Ruela, M. Francisco, and C. Barata, “Two Systems for the Detection of Melanomas in Dermoscopy Images using Texture and Color Features,” *IEEE Systems Journal*, Vol. 8, No. 3, pp. 965-979, 2014
  19. M. E. Celebi, H. A. Kingravi, B. Uddin, H. Iyatomi, Y. A. Aslandogan, W. V. Stoecker, et al., “A Methodological Approach to the Classification of Dermoscopy Images,” *Computerized Medical Imaging and Graphics*, Vol. 31, No. 6, pp. 362-373, 2007
  20. J. Luo, M. Boutell, and C. Brown, “Pictures are Not Taken in a Vacuum-an Overview of Exploiting Context for Semantic Scene Content Understanding,” *Signal Processing Magazine*, Vol. 23, pp. 101-114, 2006
  21. T. Hofmann, “Unsupervised Learning by Probabilistic Latent,” *Machine Learning Journal*, Vol. 42, pp. 177-196, 2001
  22. D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent Dirichlet Allocation,” *Machine Learning Research*, Vol. 3, pp. 993-1022, 2003
  23. Y. Sun and S. OZawa, “Semantic-Meaningful Content-based Image Retrieval in Wavelet Domain,” in *Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval*, pp. 122-129, 2003
  24. Y. Chen and J. Z. Wang, “A Region-based Fuzzy Feature Matching Approach to Content-based Image Retrieval,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, pp. 1252-1267, 2002
  25. J. Kawahara, A. BenTaieb, and G. Hamarneh, “Deep Features to Classify Skin Lesions,” in *Proceedings of IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 1397-1400, 2016