

# Evaluation of LBP and HOG Descriptors for Clothing Attribute Description

J. Lorenzo\*, M. Castrillón, E. Ramón, and D. Freire

Instituto Universitario SIANI  
Campus Universitario de Tafira  
35017 Las Palmas - SPAIN  
Universidad de Las Palmas de Gran Canaria

**Abstract.** In this work an experimental study about the capability of the LBP, HOG descriptors and color for clothing attribute classification is presented. Two different variants of the LBP descriptor are considered, the original LBP and the uniform LBP. Two classifiers, Linear SVM and Random Forest, have been included in the comparison because they have been frequently used in clothing attributes classification. The experiments are carried out with a public available dataset, the *clothing attribute dataset*, that has 26 attributes in total. The obtained accuracies are over 75% in most cases, reaching 80% for the necktie or sleeve length attributes.

## 1 Introduction

In most cultures, clothing is related to gender, age or social status. In video analytics demographics is a key element, so the identification of clothes can be used as an additional cue providing information about the consumer style or interest. The clothing identification problem can be divided into three categories. On the one hand, there exists the problem of clothes segmentation whose aim is to segment the image into regions that corresponds to the same garment. Another problem is the clothing attribute classification where the garment is described by its attributes as color, pattern, neck type, sleeve and others. The last problem is the clothing recognition where the objective is to detect garment categories in the images as t-shirt, dress, trousers and so. This paper focuses on the problem of clothing attribute classification.

The use of clothing to improve other identification task is not new. Satta et al. [12] propose a variation of their Multiple Component Matching framework named Multiple Component Dissimilarity (MCD) in order to compute the dissimilarity between a prototype person built using textual description of clothing attributes and those that made up the gallery. Gallagher and Chen [6] compute color and texture features for each superpixel obtained from a previous

---

\* Work partially funded by the Institute SIANI and the Departamento de Informática y Sistemas at ULPGC. Contact author: [javier.lorenzo@ulpgc.es](mailto:javier.lorenzo@ulpgc.es)

segmentation process and finally the clothing mask is calculated. In an application of image retrieval, Borrás et al. [1] propose a method for describing upper body clothing making use of texture and color features. The different parts of the clothes are obtained with a split-and-merge approach using homogeneity measures as criteria. Also in content-based image retrieval, Weber et al. [14] introduce a novel approach by getting the mask of the clothing starting from a set of trained pose detectors in order to deal with occlusions and different poses inherent to humans. Manfredi et al. [8] present an approach for segmenting garments in fashion stores databases. As features, color and HOG are used and combined with a Gaussian Mixture model. On the other hand, a method based in computing the color and texture of the clothing regions obtained from a background subtraction phase is described by Yang and Yu [16]. Clothing has even been used in estimating the human occupation in images [13].

Recently some works have been proposed to describe clothing based on attributes. In this regard, Yamaguchi et al. [15] describe clothes based on a method which labels superpixels obtained from a previous segmentation process making use of a Conditional Random Field model. The method starts from a pose estimation that is re-evaluated using the clothing predictions. In a recommendation scenario, Kalantidis et al. [7] describe an approach for segmenting and recognizing clothing. They start from a segmentation of the person and then each segment is classified by computing the LSH index. A similar approach is introduced by Chen et al. [4] but the user assists in the segmentation process. Liu et al. [11] have a proposal for describing clothes which is based on pose estimation as [15] and using as features color, SIFT and HOG are able to classify clothes into 23 categories. A method for classifying upper body clothing using Random Forests where decisions in each node include a SVM to estimate the density around each decision boundary is described in [2]. Chen et al. [3] present a method for describing clothing by attributes. They define 11 attributes which are classified using 40 features obtained from the arms and torso. For each attribute a SVM is learned and the CRF is used in order to take into account the relation that exists among some attributes as collar and necktie.

The aim of this paper is to study the ability of two well known descriptors along with color to classify clothing attributes as sleeve length, the existence of collar, fabric pattern and color. In this work we only focus on upper body clothing. The paper is organized as follows. In section 2, a brief description of the LBP and HOG descriptor will be given. Section 3 describes the attributes that are considered in this work and the features used for classifying each one. In section 4, the results obtained for the different attributes under study are shown and compared in two different experimental setups. Finally in section 5, the conclusions are presented.

## 2 LBP and HOG descriptors

In this section a brief description of the Local Binary Patterns (LBP) and the Histogram of Oriented Gradients (HOG) is given. LBPs have been used su-

ccessfully in different Computer Vision problems since its original application to texture classification [10]. In their definition, each pixel is encoded taking into consideration its neighborhood by means of a threshold. LBPs are therefore easily computed, and have shown their capacity of discrimination in different real world problems, while exhibiting a notorious robustness to monotonic gray-scale changes. In the original definition by Ojala et al. [9], the operator labels each image pixel comparing it with its  $3 \times 3$  neighborhood, and encodes the pixel as a binary number. The resulting codes are used as a histogram to represent the texture. This definition has recently been extended to a set of arbitrary circular neighborhoods. The expression to compute the generalized LBP is the following:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad (1)$$

where,  $g_c$  is the gray level of  $P$  neighbors,  $g_p (p = 0, 1, \dots, P - 1)$ . The function  $s(x)$  is defined as:

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

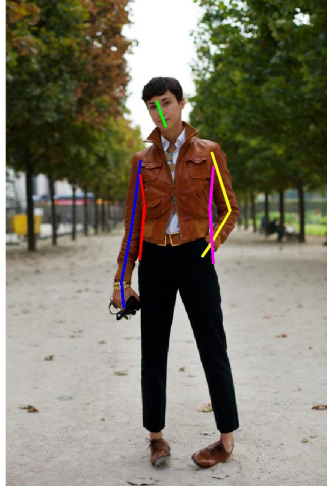
An extension of the LBP is the uniform LBP,  $LBP^u$ . A LBP code is considered uniform if the binary pattern contains at most two bitwise transitions, from 0 to 1 or vice versa when the bit pattern is considered circular. For example, patterns of 00000000 (0 transitions), number 00110000 to (2 transitions) and 11100111 (2 transitions) are uniform, while the patterns of 11001001 (4 transitions) and 01001010 (6 transitions) are not uniform. Uniform patterns indicate that some patterns are more frequent than others. Among the 256 original LBPs using the 8 neighbors with radius 1, just 58 of them are uniform, and the rest are non uniform, there are a total number of 59 different labels considering uniform LBPs.

HOG encloses a histogram in its definition [5], as the gradient orientations of a regular area are represented by means of a histogram. A whole image is represented by the concatenation of the histograms of the collection of regular areas, i.e. cells. The illumination influence is reduced normalizing each cell considering its neighborhood, called block. In their work, Dalal et al. made use of a cell size of  $8 \times 8$  pixels, and the block is  $2 \times 2$  cells. We have assumed that implementation, restricting the number of bins to 9. For the normalization stage we have used L1.

### 3 Attribute Clothing Classification

Pose estimation is an important element in clothing attribute segmentation and classification. Some authors have used the state of the art pose estimator [17] but in several situations the correct pose is not retrieved. So in other to avoid the influence of a bad pose estimation, in this work we have annotated manually the dataset. As we focus on upper body clothing, 16 annotation points are defined for each person: 2 for the face, 4 for each arm and 6 for the torso (Figure 1).

**Fig. 1.** Manual annotation of the upper body: face (green), right arm (blue), left arm (yellow), right torso (red) and left torso (magenta).



### 3.1 Sleeve Length

The hypothesis in this work is that computing the descriptors in those segments can give more information for the attribute in study. So, once the different parts of the body are obtained the sleeve length attribute is related to the portion of arm that is covered by fabric. In order to detect the amount of exposed skin in the arms there is the possibility of training a skin detector but it has to face up with different illumination and color skin. Here, a different approach is proposed. Normally sleeves are made of the same fabric than the rest of the upper body clothing, so we compare the color histogram of each arm part (hand, arm and forearm) with the color histogram of the torso. According to this, a six element feature vector is defined as

$$x_{arm\_segment} = diff(color\ histogram_{torso}, color\ histogram_{arm\_segment}) \quad (3)$$

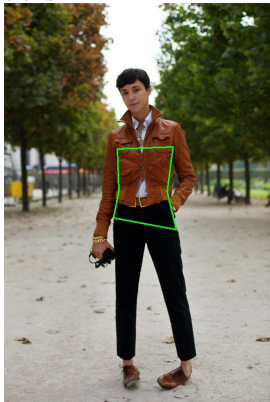
where  $arm\_segment \in \{right\ hand, right\ arm, right\ forearm, left\ hand, left\ arm, left\ forearm\}$  and  $diff$  is the Bhattacharyya distance between the two histograms.

### 3.2 Collar and Necktie

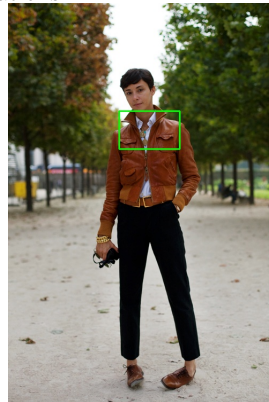
In order to classify the existence of collar and necktie the region of interest to consider is the neck and the upper chest. Considering the annotations of the images, this region can be obtained as the rectangle with base located in the armpits and height in the chin (Figure 3). To compute the HOG descriptor in this region, it is previously normalized to 32x24 pixels and then HOG is computed using a cell size of 8x8 pixels, block size of 16x16 pixels, 9 bins for

orientation and 8x8 pixels overlapping between blocks. The two variants of the LBP descriptor, original and uniform, are computed on the original rectangle obtaining 256 and 59 bin histograms respectively. The descriptors are obtained from the histograms after a  $\ell_1$  normalization stage.

**Fig. 2.** Torso region obtained from the annotations.



**Fig. 3.** Neck region obtained from the annotations.



### 3.3 Fabric Pattern

The pattern of the fabric is mainly defined by the torso region so this region will be used to compute the features used to classify the different patterns. The torso region is defined by six points: 2 for armpits, 2 for waist and 2 for hips (Figure 2). The histogram of the LBP in this region is computed and it is used as feature vector for pattern classification.

### 3.4 Color

The last attribute we have considered in this work is the color of the clothes. The histogram of the CIE  $L^*a^*b^*$  is computed for the torso (Figure 2) and it is used as feature for color classification.

## 4 Experiments

The experiments has been carried out using the *clothing attribute dataset* [3] which has 1856 images. It has 26 attributes in total, including 23 binary-class attributes: 6 for pattern, 11 for color and 6 miscellaneous; and 3 multi-class attribute: sleeve length, neckline shape and clothing category.

Some attributes exhibit a very unbalanced class distribution where a majority classifier can yield a high accuracy. For example the class distribution

**Table 1.** Results for necktie, collar and sleeve length attributes with the balanced dataset with 10-CV.

	Linear SVM				Random Forest			
	Acc.	Prec.	Rec.	F-measure	Acc.	Prec.	Rec.	F-measure
Necktie (HOG)	<b>82.47%</b>	0.83	0.83	0.83	80.48%	0.79	0.85	0.82
Necktie (LBP)	72.60%	0.74	0.74	0.74	65.41%	0.64	0.74	0.69
Necktie ( $LBP^{u2}$ )	67.12%	0.69	0.67	0.68	68.49%	0.68	0.75	0.71
Collar (HOG)	<b>78.40%</b>	0.78	0.78	0.78	74.35%	0.72	0.79	0.75
Collar (LBP)	66.91%	0.68	0.64	0.66	60.04%	0.58	0.67	0.62
Collar ( $LBP^{u2}$ )	65.61%	0.69	0.55	0.61	59.67%	0.58	0.69	0.63
Sleeve Length (RGB)	70.28%	0.70	0.70	0.70	70.60%	0.69	0.74	0.71
Sleeve Length (HSV)	<b>77.83%</b>	0.78	0.78	0.78	77.64%	0.79	0.79	0.79

**Table 2.** Results for necktie, collar and sleeve length attributes training with the balanced dataset and testing with the unbalanced dataset.

	Linear SVM				Random Forest			
	Acc.	Prec.	Rec.	F-measure	Acc.	Prec.	Rec.	F-measure
Necktie (HOG)	84.02%	0.96	0.84	0.88	<b>86.39%</b>	0.99	0.86	0.92
Necktie (LBP)	76.02%	0.99	0.76	0.86	73.32%	0.97	0.74	0.84
Necktie ( $LBP^{u2}$ )	71.60%	0.99	0.71	0.83	74.71%	0.98	0.75	0.85
Collar (HOG)	<b>78.76%</b>	0.80	0.79	0.79	71.97%	0.55	0.80	0.65
Collar (LBP)	69.70%	0.53	0.66	0.59	63.85%	0.47	0.74	0.57
Collar ( $LBP^{u2}$ )	69.37%	0.53	0.56	0.54	60.50%	0.44	0.72	0.54
Sleeve Length (RGB)	75.62%	0.87	0.76	0.78	71.08%	0.17	0.71	0.27
Sleeve Length (HSV)	<b>81.62%</b>	0.90	0.82	0.85	78.74%	0.23	0.79	0.35

for yellow color is 67 positive samples and 1677 negative samples so a majority classifier achieves a 96% of accuracy. To avoid this, for each attribute the samples were divided into two datasets. The balanced dataset where the classes has a balanced distribution, and the rest of the samples in another dataset, the unbalanced dataset. For example, for the yellow attribute the balanced dataset has 58 positive samples and 57 negative samples, and the unbalanced dataset has 19 positive samples and 1620 negative samples.

The results were obtained in two different settings for the training of the classifier. One experiment was to test the accuracy of the classifier with a 10-fold cross validation using the balanced dataset. The other experiment was to train the classifier with the balanced dataset and test it with the unbalanced dataset. For the experiments a linear SVM classifier with  $C=5$  and a Random Forest were used.

Table 1 and Table 2 show the results for the necktie, collar and sleeve length for the two experimental settings: 10-fold CV with the balanced dataset and train with the balanced dataset and test with the unbalanced dataset. It can be observed that for the necktie attribute the highest accuracy, 82.47%, is obtained with the SVM classifier with the balanced dataset and with the Random Forest,

**Table 3.** Accuracy for necktie and collar attributes combining HOG and LBPs descriptors.

	10-CV setup		Train/Test setup	
	SVM Linear	Random Forest	SVM Linear	Random Forest
Necktie (HOG+LBP)	<b>82.88%</b>	<b>82.88%</b>	<b>85.21%</b>	<b>85.21%</b>
Necktie (HOG+ $LBP^{u2}$ )	80.48%	79.11%	85.28%	84.80%
Collar (HOG+LBP)	<b>78.07%</b>	74.53%	78.03%	74.46%
Collar (HOG+ $LBP^{u2}$ )	77.30%	73.99%	<b>78.90%</b>	73.05%

**Table 4.** Results for pattern attribute using HOG with the balanced dataset with 10-CV.

	Linear SVM				Random Forest			
	Acc.	Prec.	Rec.	F-measure	Acc.	Prec.	Rec.	F-measure
Floral	<b>86.49%</b>	0.91	0.83	0.87	81.08%	0.77	0.92	0.84
Graphics	<b>81.76%</b>	0.81	0.86	0.83	78.24%	0.76	0.86	0.81
Plaid	<b>69.05%</b>	0.70	0.72	0.71	68.45%	0.68	0.76	0.72
Solid	70.75%	0.70	0.74	0.72	<b>75.82%</b>	0.74	0.80	0.77
Spot	71.26%	0.75	0.69	0.72	<b>74.85%</b>	0.73	0.84	0.78
Stripe	62.72%	0.64	0.62	0.63	<b>68.86%</b>	0.65	0.86	0.74
Average	73.67%	0.75	0.74	0.75	74.55%	0.72	0.84	0.78

86.39%, in the second setting. Furthermore, the F-measure is above 0.80 for both classifiers. With both classifiers, HOG descriptors exhibits a better performance than LBP and  $LBP^{u2}$  than can be explained for the typology of the necktie attribute that is defined by straight lines.

The results for the collar attribute show lower accuracy than for the necktie attribute with an accuracy of 78% in both settings for the SVM classifier and in both cases the accuracy is higher than the one obtained with the Random Forest. Moreover, the presence of collar is evidenced with clear contours that are better captured by the HOG descriptors than by the LBP one due to their different nature.

In order to test if LBP and HOG descriptors encode complementary information, an experiment using as descriptor the combination of the LBP and HOG in only one descriptor is carried out. Table 3 shows the results obtained in this experiments and except in one case the accuracy is very similar or even lower than using only HOG as features.

The features for the sleeve length attribute were computed using two different color spaces in equation (3): RGB and HSV. In Tables 1 and 2 can be observed the improvement of the performance when the HSV color space is used to compute the difference between the arm segments and the torso region with both classifiers. On the other hand, for HSV, accuracies of 70.28% and 81.62% are obtained for the two experimental settings with SVM classifier. Note that the performance of the classifier does not decrease in both experimental setups. For

**Table 5.** Results for pattern attribute using HOG training with the balanced dataset and testing with the unbalanced dataset.

	Linear SVM				Random Forest			
	Acc.	Prec.	Rec.	F-measure	Acc.	Prec.	Rec.	F-measure
Floral	76.29%	1.00	0.76	0.86	<b>88.43%</b>	1.00	0.89	0.94
Graphics	<b>85.76%</b>	0.99	0.86	0.92	85.20%	0.99	0.86	0.92
Plaid	67.43%	0.99	0.68	0.80	<b>80.28%</b>	0.99	0.81	0.89
Solid	<b>75.71%</b>	0.35	0.78	0.49	67.08%	0.29	0.83	0.43
Spot	68.19%	0.99	0.68	0.81	<b>83.45%</b>	0.99	0.84	0.91
Stripe	60.24%	0.98	0.60	0.75	<b>79.18%</b>	0.99	0.80	0.88
Average	72.27%	0.88	0.73	0.77	80.60%	0.88	0.84	0.83

**Table 6.** Results for pattern attribute using LBP with the balanced dataset with 10-CV.

	Linear SVM				Random Forest			
	Acc.	Prec.	Rec.	F-measure	Acc.	Prec.	Rec.	F-measure
Floral	77.98%	0.79	0.78	0.77	<b>78.38%</b>	0.78	0.83	0.81
Graphics	78.82%	0.79	0.79	0.79	<b>84.71%</b>	0.83	0.89	0.86
Plaid	<b>77.38%</b>	0.78	0.77	0.77	75.00%	0.74	0.82	0.78
Solid	73.53%	0.74	0.74	0.74	<b>77.45%</b>	0.75	0.84	0.79
Spot	<b>78.44%</b>	0.78	0.78	0.78	71.86%	0.71	0.80	0.75
Stripe	<b>74.56%</b>	0.76	0.74	0.74	68.42%	0.66	0.79	0.72
Average	76.79%	0.77	0.77	0.77	75.97%	0.75	0.83	0.79

RGB the performance is lower but the behavior is similar to the HSV, increasing the accuracy in the second experimental setting.

Tables 4 and 5 present the results obtained in the pattern attribute classification using HOG descriptors. It can be observed that Random Forest performs better than SVM because in both experimental setups the average accuracy is higher. In the second experimental setup the difference in performance is more emphasized with 80.60% for Random Forest versus 72.27% for SVM.

The results for the different LBP descriptors that were considered in the classification of the fabric pattern attribute, original LBP and the  $LBP^{u2}$ , are shown in Tables 6, 7, 8 and 9. Tables 6 and 7 show the results obtained with the original LBP descriptor for each pattern. The accuracy using SVM classifier for all patterns is around the 77% except for the *Solid* and *Stripe* patterns that drop to 73.53% and 74.56%. In the second experimental setup with the same classifier, the performance is worse except for the *Spot* and *Stripe* patterns that yield higher accuracy. When Random Forest is used as classifier, results have shown a higher variance because there are better accuracies as the one for *Graphics* with 84.71% but also there are worse accuracies like 68.42% for *Stripe* and 71.86% for *Spot*. The same behavior is observed when the Random Forest is trained with the balanced dataset and tested with the unbalanced one. On average the SVM classifier is better than the Random Forest classifier. However, in a multiple bi-class problem like this, a solution can be the combination of



**Table 7.** Results for pattern attribute using LBP training with the balanced dataset and testing with the unbalanced dataset.

	Linear SVM				Random Forest			
	Acc.	Prec.	Rec.	F-measure	Acc.	Prec.	Rec.	F-measure
Floral	71.56%	0.99	0.71	0.82	<b>78.16%</b>	1.00	0.78	0.88
Graphics	73.45%	0.98	0.73	0.83	<b>88.00%</b>	1.00	0.88	0.94
Plaid	76.65%	0.98	0.76	0.85	<b>85.50%</b>	0.99	0.86	0.92
Solid	<b>75.26%</b>	0.85	0.75	0.78	71.17%	0.32	0.82	0.45
Spot	<b>86.73%</b>	0.98	0.87	0.92	83.26%	0.99	0.83	0.91
Stripe	<b>87.00%</b>	0.97	0.87	0.91	84.44%	0.99	0.85	0.91
Average	78.44%	0.96	0.78	0.85	81.76%	0.88	0.84	0.84

**Table 8.** Results for pattern attribute using  $LBP^{u2}$  with the balanced dataset with 10-CV.

	Linear SVM				Random Forest			
	Acc.	Prec.	Rec.	F-measure	Acc.	Prec.	Rec.	F-measure
Floral	74.77%	0.77	0.75	0.75	<b>77.48%</b>	0.81	0.77	0.79
Graphics	70.00%	0.70	0.70	0.70	<b>77.06%</b>	0.76	0.82	0.79
Plaid	<b>75.00%</b>	0.76	0.75	0.75	73.21%	0.73	0.78	0.75
Solid	73.37%	0.74	0.74	0.73	<b>73.69%</b>	0.71	0.83	0.76
Spot	<b>76.05%</b>	0.76	0.76	0.76	71.26%	0.71	0.79	0.74
Stripe	<b>73.68%</b>	0.76	0.74	0.75	71.26%	0.71	0.79	0.74
Average	73.81%	0.75	0.74	0.74	73.99%	0.74	0.80	0.76

classifiers depending on the pattern, so if the best classifier for each pattern is considered we obtain on average an accuracy of 78.49%.

The results for the pattern attribute using the  $LBP^{u2}$  descriptor are shown in Tables 8 and 9 respectively. As can be seen, the performance for the SVM classifier is worse than with the LBP descriptor, being the accuracy lower for all patterns except for the *Graphics* pattern in the second experimental setup. It is shown that the use of the Random Forest classifier increases the performance in both experimental settings for most of the patterns. Considering both classifiers it can be observed that the Random Forest classifier performs better than the SVM one and if the best accuracy for each pattern is considered as in the previous case, the average accuracy increases to 84.44%.

Table 10 and Table 11 show the results of the accuracy obtained with the color attribute. As it was said in section 3.4, the histograms of the CIE  $L^*a^*b^*$  color space have been used as features. With both classifiers in the first experimental setup, the accuracy for this attribute does not have a similar behavior. On the one hand, there is a color, *Yellow*, with an accuracy 90.48% for the Random Forest classifier. On the other hand, for the same setting we observe an accuracy of 67.43% for the *White* color with the SVM classifier. In the first experimental setup both classifiers perform similar on average and if only the best classifier for each color is considered the accuracy reaches until 79.67%. When the classifiers

**Table 9.** Results for pattern attribute using  $LBP^{u2}$  training with the balanced dataset and testing with the unbalanced dataset.

	Linear SVM				Random Forest			
	Acc.	Prec.	Rec.	F-measure	Acc.	Prec.	Rec.	F-measure
Floral	65.53%	0.98	0.65	0.78	<b>82.45%</b>	1.00	0.82	0.90
Graphics	76.55%	0.98	0.77	0.85	<b>82.96%</b>	0.99	0.83	0.91
Plaid	76.40%	0.98	0.76	0.85	<b>79.52%</b>	0.99	0.80	0.88
Solid	69.24%	0.84	0.96	0.73	<b>71.17%</b>	0.31	0.81	0.45
Spot	79.97%	0.98	0.80	0.87	<b>82.94%</b>	1.00	0.83	0.91
Stripe	<b>88.94%</b>	0.97	0.88	0.92	83.68%	0.99	0.84	0.91
Average	76.11%	0.96	0.80	0.83	80.45%	0.88	0.82	0.83

**Table 10.** Results for color attribute with the balanced dataset with 10-CV.

	Linear SVM				Random Forest			
	Acc.	Prec.	Rec.	F-measure	Acc.	Prec.	Rec.	F-measure
Brown	<b>80.00%</b>	0.80	0.80	0.80	76.89%	0.74	0.84	0.79
Red	75.18%	0.77	0.75	0.75	<b>85.11%</b>	0.84	0.89	0.86
Yellow	81.90%	0.82	0.82	0.82	<b>90.48%</b>	0.91	0.91	0.91
Green	76.03%	0.76	0.76	0.76	<b>79.34%</b>	0.78	0.82	0.80
Cyan	77.46%	0.77	0.77	0.77	<b>80.92%</b>	0.82	0.86	0.84
Blue	<b>79.55%</b>	0.79	0.79	0.79	<b>79.55%</b>	0.80	0.87	0.83
Purple	<b>76.52%</b>	0.77	0.77	0.77	72.17%	0.71	0.80	0.75
White	67.43%	0.68	0.67	0.67	<b>71.43%</b>	0.69	0.75	0.72
Gray	<b>73.97%</b>	0.74	0.74	0.74	71.58%	0.69	0.77	0.73
Black	79.00%	0.79	0.79	0.79	<b>79.34%</b>	0.78	0.82	0.80
Average	75.66%	0.76	0.76	0.76	75.83%	0.75	0.81	0.78

are trained with the balanced dataset and tested with the unbalanced one, the Random Forest behaves better than SVM although the same variations in accuracy are observed in this experimental setting, ranging from 70.23% for *Purple* with SVM to 89.93% for *Red* with Random Forest. Considering again the use of a combination of classifiers with the best one for each color, the average accuracy increases 84.05%.

## 5 Conclusions

In this work, LBP and HOG descriptors and color have been used for clothing attribute identification obtaining good results. The experiments were carried out with a public dataset, the *clothing attribute dataset*, to allow other researchers to compare their results with those published in this paper. Instead of using a pose estimator, the dataset images were manually annotated to avoid the influence of the bad pose estimations in the results that can hinder the actual discriminative power of the descriptors. For each individual, the head, torso and arms were annotated.

**Table 11.** Results for color attribute training with the balanced dataset and testing with the unbalanced dataset.

	Linear SVM				Random Forest			
	Acc.	Prec.	Rec.	F-measure	Acc.	Prec.	Rec.	F-measure
Brown	76.77%	0.96	0.77	0.84	<b>82.09%</b>	0.98	0.83	0.90
Red	80.91%	0.98	0.81	0.88	<b>89.39%</b>	1.00	0.89	0.94
Yellow	79.44%	0.99	0.79	0.87	<b>85.60%</b>	1.00	0.86	0.92
Green	83.55%	0.98	0.84	0.90	<b>87.74%</b>	1.00	0.88	0.93
Cyan	76.38%	0.99	0.76	0.86	<b>86.31%</b>	1.00	0.86	0.93
Blue	<b>88.65%</b>	0.96	0.89	0.92	86.94%	0.99	0.88	0.93
Purple	70.23%	0.98	0.70	0.81	<b>80.36%</b>	0.99	0.81	0.89
White	72.37%	0.84	0.72	0.77	<b>78.04%</b>	0.95	0.80	0.86
Gray	73.50%	0.90	0.75	0.79	<b>79.81%</b>	0.96	0.82	0.88
Black	79.15%	0.86	0.79	0.81	<b>82.26%</b>	0.94	0.83	0.88
Average	76.71%	0.92	0.77	0.83	82.29%	0.97	0.83	0.90

Some of the five attributes that have been considered in this work exhibits a very unbalanced class distribution in the dataset so we have defined two experimental setups. The first one with a dataset with equal proportion for each class and the second one with the rest of samples which results in an unbalanced distribution. Two of the most used classifier in clothing recognition field were compared, SVM and Random Forest. When an attribute, as pattern and color, has several classes neither of them performs better than the other for all the attributes so our conclusion is that the use of a combination of classifier will be the preferred option. With the combination of classifier higher accuracies are obtained, reaching 84.05% and 84.44% for color and pattern respectively. For necktie, collar and sleeve length attributes the best accuracies were 86.39%, 78.76% and 81.62% respectively.

## References

1. Borrás, A., Tous, F., Lladós, J., Vanrell, M.: High-level clothes description based on colour-texture and structural features. In: 1st Iberian Conference on Pattern Recognition and Image Analysis IbPRIA (2003)
2. Bossard, L., Dantone, M., Leistner, C., Wengert, C., Quack, T., Gool, L.V.: Apparel classification with style. In: 11th Asian Conference on Computer Vision. Daejeon, Korea (November 5-9 2012)
3. Chen, H., Gallagher, A., Girod, B.: Describing clothing by semantic attributes. In: European Conference on Computer Vision (ECCV) (2012)
4. Chen, J.L., Chen, W.Y., Chen, I.K., Chi, C.Y., Chen, J.L.: Interactive clothing retrieval system. In: IEEE International Conference on Consumer Electronics (ICCE) (2014)
5. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Schmid, C., Soatto, S., Tomasi, C. (eds.) International Conference on Computer Vision & Pattern Recognition. vol. 2, pp. 886–893. INRIA Rhône-Alpes, ZIRST-655, av. de l’Europe, Montbonnot-38334 (June 2005)

6. Gallagher, A., Chen, T.: Clothing cosegmentation for recognizing people. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR) (2008)
7. Kalantidis, Y., Kennedy, L., Li, L.J.: Getting the look: Clothing recognition and segmentation for automatic product suggestions in everyday photos. In: International Conference on Multimedia Retrieval (ICMR) (2013)
8. Manfredi, M., Granata, C., Calderara, S., Cucchiara, R.: A complete system for garment segmentation and color classification. *Machine Vision and Applications* November (2103)
9. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition* 29, 51–59 (1996)
10. Pietikäinen, M., Hadid, A., Zhao, G., Ahonen, T.: *Computer Vision Using Local Binary Patterns*. Springer (2011)
11. S Liu, J Feng, C.D.H.X.J.H.Z.H.S.Y.: Fashion parsing with weak color-category labels. *IEEE TRANSACTIONS ON MULTIMEDIA*, 16, 253–265 (2014)
12. Satta, R., Pala, F., Fumera, G., , Roli, F.: Person Re-Identification, chap. People search with textual queries about clothing appearance attributes, pp. 371–390. Springer (2014)
13. Song, Z., Wang, M., Hua, X.S., Yan, S.: Predicting occupation via human clothing and contexts. In: IEEE International Conference on Computer Vision (ICCV) (2011)
14. Weber, M., Buml, M., Stiefelhagen, R.: Part-based clothing segmentation for person retrieval. In: International Conference on Advanced Video and Signal-based Surveillance (AVSS) (2011)
15. Yamaguchi, K., Kiapour, M.H., Ortiz, L.E., Berg, T.L.: Parsing clothing in fashion photographs. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2012)
16. Yang, M., Yu, K.: Real-time clothing recognition in surveillance videos. In: 18th IEEE International Conference on Image Processing (2011)
17. Yang, Y., Ramanan, D.: Articulated pose estimation using flexible mixtures of parts. In: 24th IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2011)