Facial Makeup Detection Technique Based on Texture and Shape Analysis

Neslihan Kose[1], Ludovic Apvrille[2], Jean-Luc Dugelay[3]

[1,3] EURECOM, Sophia Antipolis, France

[2] Institut Mines-Telecom, Telecom ParisTech, CNRS LTCI, Sophia Antipolis, France

[1,3] {neslihan.kose, jean-luc.dugelay}@eurecom.fr,
[2] ludovic.apvrille@telecom-paristech.fr

Abstract— **Recent studies show that the performances of face recognition systems degrade in presence of makeup on face. In this paper, a facial makeup detector is proposed to further reduce the impact of makeup in face recognition. The performance of the proposed technique is tested using three publicly available facial makeup databases. The proposed technique extracts a feature vector that captures the shape and texture characteristics of the input face. After feature extraction, two types of classifiers (i.e. SVM and Alligator) are applied for comparison purposes. In this study, we observed that both classifiers provide significant makeup detection accuracy. There are only few studies regarding facial makeup detection in the state-of-the art. The proposed technique is novel and outperforms the state-of-the art significantly.**

I. INTRODUCTION

There are several challenges in face recognition such as illumination, pose, expression, occlusions, aging, plastic surgery and spoofing. Challenges related to variations in pose, illumination and expression have been identified and addressed by diverse advanced algorithms [1]. Although many challenges in face recognition have been explored extensively, some challenges such as aging [2], plastic surgery [3], [4], spoofing [5], [6] and makeup [7], [8] still continue to challenge the performance of face recognition systems.

The use of makeup to cover facial flaws and look more attractive is very common in many parts of the world. Makeup can also be used to impersonate or change appearance, which poses a major challenge to face recognition systems. Therefore it is important to mitigate the impact of makeup in face recognition in terms of both security and recognition accuracy. Also, age estimation and plastic surgery detection methods can utilize knowledge about the presence of makeup to optimize their outputs.

Appropriate databases are necessary for the evaluation of the impact of makeup in face recognition. Some databases (e.g. [13], [14]) provide images with applied facial cosmetics but there are no corresponding companion images without makeup which makes them inappropriate for face recognition tests. There are only a few studies which analyze the impact of facial makeup in face recognition. The studies [19], [20] demonstrate that makeup can affect a person's ability to recognize faces. In the recent study of Dantcheva et al. [7], the negative impact of facial makeup on the matching performance of four face recognition algorithms was demonstrated using two publicly available databases (i.e.

YMU and VMU databases [7]). Their experiments showed a significant decrease in matching accuracy in presence of facial makeup. In the study of Eckert et al. [8], a facial makeup database (i.e. FCD database [8]) was prepared and the preliminary tests concerning the impact of facial cosmetics on face recognition were conducted. In [8], the amount of applied facial cosmetics and the cosmetics' application area are evaluated in details. Again, significant decrease was observed in recognition performance in presence of facial makeup. In both [7], [8], the authors mention the need for the development of robust algorithms to facial makeup.

Since it has been shown that performances of face recognition systems degrade in presence of facial makeup, researchers started to work on developing techniques for the detection of makeup to further improve face recognition performances. To the best of our knowledge, the only works related to automatic makeup detection are the recent studies by Chen et al. [9], Varshovi [10], Guo et al. [21], and Hu et al. [22]. In [9], the authors designed a method to automatically detect the presence of makeup in face images. The proposed algorithm in [9] extracts a feature vector that captures the shape, texture and color characteristics of the input face, and employs a classifier to determine the presence or absence of makeup. They did experiments on two datasets consisting of 151 subjects (600 images) and 125 subjects (154 images), respectively, and obtained makeup detection rate of up to 93.5% (at a false positive rate of 1%). In [10], the authors tested their approach on 120 images of 21 frontal, neutral expression female subjects and obtained classification accuracies of 90.62% for eye-shadow detection, 93.33% for lip-stick detection and 52.5% for liquid foundation detection. The study [10] also explores texture and color features for makeup cues. In [21], Guo et al. use Partial Least Square (PLS) to learn the correlation between different facial parts separately and show that correlation learning is an essential step to improve verification performance in presence of makeup. In [22], the authors use Canonical Correlation Analysis (CCA) to learn the meta subspace, which can maximize the correlation of feature vectors belonging to the same individual.

In this paper, we designed a new method to detect makeup in unconstrained face images. Given a face image, the proposed method first extracts a set of features based on shape and texture characteristics of the input face. The features

Fig. 1. Reference Image and Makeup Series Images in their Original Form. This figure is taken from [8].

extracted using texture descriptor Local Gabor Binary Pattern (LGBP) [15] and shape descriptor Histogram of Gradients (HOG) [18] are concatenated and this feature set is used by a classifier to detect the presence or absence of makeup in the image. The contributions of this study can be listed as:

- It is the first time that LGBP and HOG decsriptors are combined to be used for makeup detection.
- The experiments are conducted on three challenging datasets, which include variations in facial pose, illumination, expression, and image resolution. Two of these databases (YMU and MIW databases) were proposed in [7] and the remaining database (FCD database) was proposed in [8]. In [8], FCD database was used to analyze the impact of makeup on face recognition. In this study, it is the first time that FCD database is used for facial makeup detection purposes. The results of this study show that FCD database is more challenging compared to YMU and MIW databases.
- Since FCD database [8] provides annotation for non-makeup, slight, intermediate and heavy makeup, it is the first time that the classification accuracy of a makeup detector is evaluated for multi-classes.
- The proposed facial makeup detection approach outperforms the state-of-the art significantly.
- In this study, we applied the same experiments as in [9] using the same databases. Hence an almost exact comparison with the study [9] was possible. The results of this comparison show that the selected descriptors in our study (LGBP and HOG) are more appropriate for makeup detection compared to the descriptors selected in [9].
- In this study, performances are evaluated using two types of classifiers. The first one is Support Vector Machine (SVM) [11], which is commonly used in biometrics domain. The second classifier is Alligator [12], which was previously developed for malware detection. Aligator is a tool that includes several classifiers. In this study, we observed that Alligator provides almost perfect results, which proves that fusion of classifiers has a positive impact in classification accuracy and also the features extracted with the proposed facial makeup detector provide very discriminative results in facial makeup detection.
- Different from the previous techniques [9], [10], [21], in this study, color information is not used for facial makeup detection.

The paper is organized as follows: Sections II introduces

the databases that were used for this study. Section III explains the proposed makeup detection method. Section IV shows the experiments and results. Finally, conclusions are provided in Section V.

II. FACIAL COSMETIC DATABASES

In this study, the performances of the proposed technique in facial makeup detection is tested using three publicly available databases.

A. Facial Cosmetics Database (FCD) [8]

Facial Cosmetics database is proposed in the study of Eckert et al. [8]. This database contains multiple images with and without facial cosmetics for each person and provides detailed information about the application area and amount of applied cosmetics for each image.

In this database, makeup tutorial videos are used as source hence images in this database are similar to images in real world scenarios (e.g. surveillance camera, snapshots). Most of the videos used in FCD database are found on YouTube, some on websites of facial cosmetic companies.

The persons in the database belong mainly to the Caucasian race and are between 20 and 30 years old. Hispanics, Asian and African Americans are also sparsely represented.

The database contains frontal facial images with limited expression. It provides one reference image and one or more progressive makeup series per person. The series always start with a non-makeup image. The reference non-makeup image does not belong to any makeup series and is taken in another setting, resulting in variations of illumination, skin and hair condition (Fig. 1).

In this database, each image is available in its original and pre-processed (i.e. cropped and aligned) version. The database currently contains 385 cropped images, whereas 50 are reference images and 335 belong to makeup series. 109 of the makeup serie images are classified into the makeup category 'no', 145 into 'slight', 52 into 'intermediate' and 79 into 'heavy' makeup. Each image is available in its original (Fig. 1) and pre-processed form. The size of the pre-processed faces available in the facial cosmetics database is 150×140 pixel. In this study, we used the pre-processed faces in FCD database as input. More information is available at $http://fcd.eurecom.fr/$.

B. YouTube Makeup Database (YMU)[7]

In this study, we utilized the database introduced by Dantcheva et al. [7], which contains the before and after makeup images of 151 Caucasian female subjects taken from

Fig. 2. Facial images showing variations in pose, illumination, expression and resolution from the YMU database [7].

Fig. 3. Sample images from the MIW database [9]. Top row shows images without makeup and the bottom row shows images with makeup.

YouTube makeup tutorials. 99 subjects were used in their work in [7] but then in [9], all the 151 subjects (600 images) in the YMU database are used.

After their study in [9], they did slight changes in YMU database. The total number of images in the current YMU dataset is 604, with 302 makeup images and 302 non-makeup images for 151 subjects. There are four shots per subject, two shots before the application of makeup and two shots after the application of makeup. However, for comparison purposes with the study in [9], we used the images of the first 150 subjects in the database which corresponds to 300 makeup images and 300 non-makeup images (600 images in total). We note that the degree of makeup in this database varies from slight to heavy. The database is relatively unconstrained, exhibiting variations in facial expression, pose and resolution. Examples are shown in Fig. 2 (after face cropping and alignment).

C. Makeup in the wild database (MIW)[9]

In [9], the authors assembled another database of 154 images (77 with makeup, and 77 without makeup) corresponding to 125 subjects. The images are obtained from Internet. A few examples are shown in Fig. 3. In [9], the purpose of using this database was to evaluate the generalization capability of the proposed makeup face detector where the training is performed using the YMU database and testing is done on the MIW database. In our study, the purpose of using YMU and MIW databases is to apply the proposed makeup detection approach similar to the protocol explained in [9] for an exact comparison between the proposed approach and the most recent facial makeup detector explained in [9].

For both the YMU and MIW databases, the pre-processed images are provided online in the authors' webpage: $http://www.antitza.com/makeup-datasets.html$. In this study, we used these pre-processed images of size 150×130 pixel as input. Also, in both databases, an image is labeled as 'makeup' even if cosmetic details are present in only a portion of the face therefore there are only two classes which are 'makeup' and 'non-makeup'.

The advantage of FCD database compared to YMU and MIW databases is that it is possible to classify images as 'no', 'slight', 'intermediate' and 'heavy' makeup since the annotations for all four classes are provided in FCD database.

III. THE PROPOSED FACIAL MAKEUP DETECTION TECHNIQUE

In this section, the proposed facial makeup detector is explained in two parts as feature extraction and classification.

Fig. 4 shows the flowchart of the proposed makeup detector. In this approach, the following steps are applied to discriminate makeup and non-makeup images (1) An input face image is normalized (for this step we used the cropped face images provided by the publicly available facial makeup databases) (2) Local Gabor Binary Pattern (LGBP) technique is applied on 5×5 non-overlapping blocks and a feature histogram of 59000 features ($59 \times 40 \times 25$) is obtained. (3) Histogram of Gradients (HOG) technique is applied on 5×5 non-overlapping blocks and a feature histogram of 2025 features (81×25) is obtained. (4) The two types of histograms are concatenated to form the final histogram sequence. (5) Finally the classifier detects whether the input image is a face with or without makeup.

A. Feature Extraction

The proposed features for makeup detection are based on shape and texture descriptors. In this study, we used HOG and LGBP techniques to extract shape and texture features, respectively. The reason of selecting LGBP and HOG descriptors for facial makeup detection in this study can be summarized as follows. In the literature, we observed that the fusion of HOG and LBP descriptors is common in many domains such as people detection [23], object localization [24], pedestrian detection [25], [26], palm tracking [27], object detection [28] and signature verification [29]. However, there is a lack of studies for the fusion of LGBP and HOG features. In LGBP technique, LBP technique is applied on Gabor pictures. In this study, we observed that the fusion of LGBP and HOG descriptors provides better results compared to the fusion of LBP and HOG descriptors. Since HOG is a shape and LGBP is a texture descriptor, they provide complementary information when they are used together, and much better detection accuracies are obtained in facial makeup detection.

1) LGBP Features [15]: In the study [16], LGBP technique is selected as baseline for face recognition and the impact of facial makeup on recognition performance of LGBP technique is evaluated. In [16], a significant increase is observed in Equal Error Rate (EER) after application of makeup. This proves that application of makeup reduces the recognition performance of LGBP technique.

In the present study, different from its utilization as baseline technique in [16], we applied this technique for feature extraction as a part of the proposed makeup detector (Fig. 4).

Since local shape and texture information are impacted by makeup, initially, an input face image is transformed to ob-
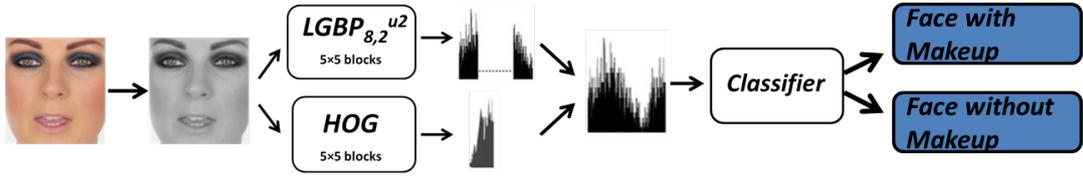
Fig. 4. Flowchart of the proposed facial makeup detection technique.

tain multiple Gabor Magnitude Pictures (GMPs) in frequency domain by applying multi-scale and multi orientation Gabor filters; (2) Each GMP is converted to Local Gabor Binary Pattern (LGBP) map; (3) Each LGBP Map is further divided into 5×5 non-overlapping rectangle regions, and histogram is computed for each region; (4) The LGBP histograms of all the LGBP Maps are concatenated to form the final histogram sequence as the model of the face.

The Gabor filters we used are defined as follows [15]:

$$\psi_{\mu,\nu}(z) = \frac{\left\|k_{\mu,\nu}\right\|^2}{\sigma^2} e^{-\frac{\|k_{\mu,\nu}\|^2 \|z\|^2}{2\sigma^2}} \left[e^{ik_{\mu,\nu}z} - e^{-\frac{\sigma^2}{2}} \right] \quad (1)$$

μ and ν define the orientation and scale of the Gabor filters, $z = (x, y)$, $\|.\|$ denotes the norm operator, and the wave vector $k_{\mu,\nu} = k_\nu e^{i\phi_\mu}$ where $k_\nu = k_{max}/\lambda^\nu$ and $\phi_\mu = \pi\mu/8\lambda$ is the spacing factor between filters in the frequency domain.

The Gabor representation of a face image is derived by convolving the face image with the Gabor filters. Let $f(x,y)$ be the face image, its convolution with a Gabor filter $\psi_{\mu,\nu}(z)$, is defined as follows:

$$G_{\psi f}(x, y, \mu, \nu) = f(x, y) * \psi_{\mu,\nu}(z) \quad (2)$$

$*$ denotes the convolution operator. Five scales $\nu \in \{0, ..., 4\}$ and eight orientations $\mu \in \{0, ..., 7\}$ Gabor filters are used. Convolving the image with each of the 40 Gabor filters can then generate the Gabor features. Note that, because the phase information of the transform is time-varying, generally, only its magnitude is explored. Thus, for each Gabor filter, one magnitude value will be computed at each pixel position, which will totally result in 40 Gabor Magnitude Pictures.

In LGBP technique, the magnitude values are encoded with LBP operator. The original LBP [17] forms labels for the image pixels by thresholding the 3×3 neighborhood of each pixel with the center value and considering the result as a binary number.

The LBP operator has been extended to use neighborhoods of different sizes. $LBP_{P,R}$ is computed such that for a given central pixel in an image, a pattern number is computed by comparing its value with those of its neighbors. In Eq. (3), g_c is the gray value of the central pixel, g_p is the value of its neighbors, P is the number of neighbors around a circle of radius R. $LBP_{P,R}$ calculation is shown in Eq. (3) and (4):

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad (3)$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (4)$$

Another extension to the original operator is the use of uniform patterns, which are verified to be the fundamental patterns of local image texture. A local binary pattern is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa when the bit pattern is traversed circularly. The notation is $LBP_{P,R}^{u2}$. u2 stands for using only uniform patterns and labeling all remaining patterns with a single label.

The operator LGBP denotes the LBP operates on GMP. In this study, local feature histogram is exploited to summarize the region property of the LGBP patterns. Each LGBP Map is spatially divided into 5×5 non-overlapping regions. Then, histogram is extracted from each region by utilization of $LBP_{8,2}^{u2}$. Finally, all the histograms estimated from the regions of all the LGBP Maps are concatenated into a single histogram sequence to represent the given face image. The total number of features extracted from one image is equal to $59 \times 5 \times 5 \times 40 = 59000$ features.

2) HOG Features [18]: As explained in [18], to compose HOG, the cell histograms of each pixel within the cell casts a weighted vote, according to the gradient L2-norm, for an orientation-based histogram channel. Similar to [18], in this work the histogram channels are calculated over rectangular cells (i.e. R-HOG) by the computation of unsigned gradient. The cells overlap half of their area, which means that each cell contributes more than once to the final feature vector. In order to account for changes in illumination and contrast, the gradient strengths were locally normalized over each cell. As suggested in [18], we used 9 rectangular cells, and 9 bin histogram per cell is computed. The nine histograms with nine bins were then concatenated to make a 81-dimensional feature vector. We applied HOG technique on 5×5 non-overlapping regions. Therefore, the total number of features extracted from one image is equal to $81 \times 5 \times 5 = 2025$ features.

In this paper, we used the publicly available code developed for the study [18] for HOG implementation.

Feature Level Fusion: The HOG feature values are between 0 and 1 range due to the normalization in its implementation. This is why, first we also apply normalization for LGBP features. We choose the maximum feature value appeared in the training set of makeup and non-makeup images consecutively and divide the features of all makeup and non-makeup images with these two corresponding maximum values, respectively. After this normalization step, HOG and

LGBP features are concatenated and the classifier is applied.

B. Classification

In [9], the authors tested the performance of their makeup detector using two types of classifiers which are SVM and Adaboost. They claim that they obtain better results using SVM classifier.

In this study, we evaluated the performance of the proposed approach using the powerful SVM [11] classifier for comparison purposes with the previous studies. However, in Experiment 3, we also tested the performance of the proposed approach with Alligator [12] classifier, which was previously developed for malware detection to solve security problems. In this study, it is the first time the performance of this classifier is evaluated in another domain.

1) Support Vector Machine (SVM) [11]: In this study, we tested the performances using both linear and Gaussian RBF kernel SVM classifiers. We obtained the same results with the linear SVM and RBF kernel using the optimum parameters. Hence, we reported the results using linear SVM only.

2) Alligator [12]: Alligator is a tool that includes several classifiers (SVM but also k-NN, epsilon-cluster and so on) and that provides an optimal fusion based on learning from the training set. It has been first introduced for the classification of malware/non-malware applications [12], [30]. At training, alligator takes as input a list of classification algorithms and their respective parameters (example : 20-NN). Then, based on heuristics, it determines a best way to combine those algorithms, and thus outputs a weight to apply to each algorithm and for each set. Then, Alligator can classify unknown sets of elements based on combination of weighted algorithms. The selection of weights to apply to classification algorithms can be automatically performed by Alligator. Nonetheless the list of classification algorithms has to be selected by your own.

IV. EXPERIMENTS

In order to evaluate the performance of the proposed makeup detection technique, three experiments were done.

A. Experiment 1

For this experiment, we applied a 5-fold cross-validation scheme similar to the study in [9] using the YMU database with the same evaluation protocol for comparison purposes. In [9], the YMU dataset is divided into 5 folds with approximately 30 subjects in each fold. 4 folds are used for training the makeup detector, and the remaining fold is used for testing it. This is repeated 5 times. The subjects in the training and the test set are not overlapping. We do not know which subjects were selected for training and test sets in [9]. However since the total number of subjects involved in each set for each evaluation is similar, we claim that the comparison of our results with the reported results in [9] is almost an exact comparison.

In our training and test set partioning, we divide the YMU dataset into 5 folds with 30 subjects in each fold. Then at each trial, one fold is used for testing and the remaining folds are used for training.

Trial	Train	Test	SVM(%)
1	480(240/240)	120(60/60)	100
2	480(240/240)	120(60/60)	93.33
3	480(240/240)	120(60/60)	100
4	480(240/240)	120(60/60)	100
5	480(240/240)	120(60/60)	99.17
Average			**98.5**

The performance of the makeup detector is reported using the metrics Classification Rate (CR), which is the percentage of makeup and non-makeup images that are correctly classified by the detector.

Table I shows the results of the proposed facial makeup detector using two types of classifiers. In [9], the authors' proposed approach achieves the average classification rate of 91.20% using SVM classifier on YMU database. Since in this study we use the same evaluation protocol in [9], we can say that the proposed approach outperforms the results of [9] with the classification rate of 98.5% using the SVM classifier.

In order to show the impact of fusion in the proposed approach, we also evaluated the average classification rates using HOG and LGBP features only. When the SVM classifier is applied only using HOG features and LGBP features, the average classification rates are found to be 89.5% and 82.5%, respectively. The feature level fusion of HOG and LGBP features increases the result up to 98.5%, which proves that HOG (shape descriptor) and LGBP (texture descriptor) are complementary features.

B. Experiment 2

Next, the proposed makeup detector that is trained on the YMU database (300 makeup and 300 non-makeup images), is tested on the MIW database (77 makeup and 77 non-makeup images). In the study [9], the same test is done with the same databases. Since the training and test sets consist of almost the same images in our study and the study in [9], an exact comparison is then possible for this experiment.

Fig. 5. The only wrongly detected sample in MIW database.

A classification rate of 95.45% for SVM and 92.21% for Adaboost was obtained in [9]. Also, at 1% FPR, a TPR of 93.51% and 84.42% was obtained for SVM and Adaboost, respectively in [9]. We applied the same experiment with our proposed facial makeup detector using the same databases and the same protocol as in [9]. A classification rate of 99.35% for SVM was obtained. Also, at 1% FPR, a TPR of 97.7% was obtained for SVM and Area Under Curve

is computed as 0.99983. The exact comparison with the technique in [9] shows that our technique is very successful in detecting facial makeup by providing almost perfect detection accuracy. Hence we can say that the proposed technique outperforms the state-of-the art significantly. The proposed facial makeup detector failed only for 1 image of the 154 images. The wrongly detected sample is shown in Fig. 5. From this figure it is clear that the reason of this wrong detection is mainly due to the significant pose variation in the image.

This experiment confirms the generalization ability of the proposed approach since the training and test sets are constructed from different databases (i.e. YMU and MIW databases).

C. Experiment 3

The first two experiments show that the proposed facial makeup detector is very successful in detecting facial makeup compared to the very recent studies on this topic.

In this experiment, we tested the performance of the proposed facial makeup detector on another publicly available database proposed in [8]. The reason of this experiment is to prove that the proposed facial makeup detector is also successful on other databases.

For this test, we again applied 5-fold cross-validation scheme similar to Experiment 1. The results using the FCD database [8] are shown in Table II.

TABLE II

CLASSIFICATION RATES OF SVM AND ALLIGATOR CLASSIFIERS ON THE FCD DATABASE. THE NUMBERS IN PARENTHESES INDICATE THE NUMBER OF "MAKEUP" AND "NON-MAKEUP" IMAGES IN EACH TRIAL.

Trial	Train	Test	SVM(%)	Alligator(%)
1	270(189/81)	115(87/28)	90.43	100
2	325(236/89)	60(40/20)	83.33	100
3	324(235/89)	61(41/20)	96.72	100
4	319(230/89)	66(46/20)	87.88	96.05
5	302(214/88)	83(62/21)	87.95	97.60
Average			89.26	98.73

The results of this experiment prove the following:

- The proposed facial makeup detector provides very significant classification accuracy using challenging databases.
- The average classification rate computed in this experiment using SVM classifier is less compared to the average classification rate computed in Experiment 1. This result shows that the FCD database is more challenging compared to the YMU database.
- In this experiment, we obtained almost perfect classification accuracy (98.73%) with Alligator, which proves that fusion of different classifiers (as in the case of Alligator) improve the results significantly.
- This experiment also shows that the features extracted with the proposed approach are very appropriate in facial makeup detection since both classifiers provide significant classification accuracies using these features.

- After feature extraction step, the classification part for only the first trial in Experiment 3 takes 16 seconds with SVM classifier and inferior than 1 minute with Alligator.

Fig. 6. Non-makeup and slight makeup images for two examples in the FCD database. The images in the left are non-makeup images, the images in the right are makeup images.

In the YMU database, the number of makeup and non-makeup images are equal. This is why in training set, we used the same number of images for makeup and non-makeup image sets. However in FCD database the number of non-makeup images is less compared to the number of makeup images. This is why in the training sets, we have less samples for non-makeup images which may affect the performance of the proposed classifier. Furthermore, in the YMU database, the degree of makeup varies from subtle to heavy. The makeup in the images is quite visible for human eye. However in the FCD database, the slight makeup images are quite similar to the non-makeup images. Figure 6 shows some examples for this similarity. Since 145 out of the 335 makeup serie images are classified into the makeup category 'slight', we can claim that the FCD database is very challenging hence the performance of the proposed facial makeup detector is less in Experiment 3 compared to Experiment 1 using the same SVM classifier.

D. Experiment 4

In this experiment, we evaluated the classification rate for non-makeup, slight makeup and heavy makeup images using the FCD database. In the previous experiments, we analyzed binary classification whereas in this experiment, we analyze multi-class classification. There are only 52 images which are categorized in intermediate class in FCD database. Since the number of images in this class is quite low, we fused intermediate makeup images and heavy makeup images and named this set as heavy makeup set. Then, we compute the classification rates for the three classes which are non-makeup, slight makeup and heavy makeup.

In this experiment, we selected the first 40 subjects for training sets of 'no', 'slight', and 'heavy' makeup classes. The images of the remaining subjects are involved for testing sets of the three classes. The total classification rate for the classes 'no', 'slight' and 'heavy makeup' is evaluated as 72.29%. 60 sample out of 83 sample are correctly classified in their class. The number of wrongly classified samples are 5 out of 27 for 'slight', 2 out of 21 for 'no' and 16 out of 35 for 'heavy' makeup classes. This shows that the performance of the proposed approach is less for multi-class case compared to binary classification as expected. However we can still claim that the proposed approach provides significant multi-class classification accuracy.

Since FCD database provides annotations for different classes, the utilization of this database is advantageous for multi-class classification analysis compared to the other existing databases.

V. CONCLUSIONS

In this paper, an automated makeup detector is proposed for unconstrained facial images. The proposed detector utilizes shape and texture features extracted from the entire face, to determine the presence of makeup.

Different from the most recent studies [9], [10], [21], we do not benefit from color information and also analysis are done only on entire face. We do not extract information from facial subregions.

Experiments are conducted on three unconstrained face datasets. The experiments resulted in makeup detection rates of up to 97.7% (at 1% false positive rate) and overall classification rates of up to 99.35% when YMU database is used for training and MIW database is used for testing. The same experiment is conducted in [9], which resulted in makeup detection rates of up to 93.5% (at 1% false positive rate) and overall classification rates of up to 95.45%. This result proves that the proposed make detector outperforms the state-of-the art significantly.

Another novelty of this study is that it is the first time that the degree of makeup applied to the face is analyzed. Since FCD database provides annotations for 'no', 'slight', 'intermediate' and 'heavy' makeup, we were able to analyze multi-class classification in makeup analysis. Furthermore, it is the first time we analyzed the performance of Alligator, which was previously developed for malware detection, in image analysis domain and obtained almost perfect makeup detection accuracy using this classifier.

Since the proposed facial makeup detector is very successful in detecting facial makeup, the output of the makeup detector can be used to perform adaptive pre-processing in the context of face recognition as suggested in [9]. Our future work is to integrate the facial makeup detector to existing face recognition techniques in order to improve the recognition accuracy of face matchers when matching makeup images against non-makeup images. Finally, the makeup detector can be used to optimize the output of age estimation and beauty assessment algorithms that may also be impacted by the application of makeup.

REFERENCES

[1] S. Z. Li and A. K. Jain, editors. Handbook of Face Recognition, 2nd Edition. Springer, 2011.

[2] Y. Fu, G. Guo and T. S. Huang. "Age synthesis and estimation via faces: a survey", *IEEE Trans. on PAMI*, vol. 32, pp. 1955-1976, 2010.

[3] H. S. Bhatt, S. Bharadwaj and et al. "Recognizing surgically altered face images using multiobjective evolutionary algorithm", *IEEE Trans. on Information Forensics and Security*, vol. 8, pp. 89-100, 2013.

[4] N. Kose, N. Erdogmus and J.-L. Dugelay. "Block based face recognition approach robust to nose alterations", *IEEE Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS)*, pp. 121-126, 2012.

[5] J. Maatta, A. Hadid and M. Pietikainen. "Face spoofing detection from single images using texture and local shape analysis", *IET Biometrics*, vol. 1, pp. 3-10, 2012.

[6] N. Kose and J.-L. Dugelay. "Mask spoofing in face recognition and countermeasures", *Image and Vision Computing Journal - Elsevier, Special Issue of IMAVIS - Best of Automatic Face and Gesture Recognition 2013*, vol. 32, no. 10, pp. 779-789, July 2014, doi: 10.1016/j.imavis.2014.06.003.

[7] A. Dantcheva, C. Chen and A. Ross. "Can facial cosmetics affect the matching accuracy of face recognition systems?", *IEEE Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS)*, 2012.

[8] M.-L. Eckert, N. Kose and J.-L. Dugelay. "Facial cosmetics database and impact analysis on automatic face recognition", *IEEE Int. Workshop on Multimedia Signal Processing (MMSP)*, pp. 434-439, 2013.

[9] C. Chen, A. Dantcheva, and A. Ross. "Automatic Facial Makeup Detection with Application in Face Recognition", *Int. Conf. on Biometrics (ICB)*, pp. 1-8, 2013.

[10] S. Varshovi. "Facial makeup detection using HSV color space and texture analysis", *Master's thesis*, Concordia University, Canada, 2012.

[11] C.-C. Chang and C.-J. Lin. "LIBSVM: A library for support vector machines", *ACM Trans. on Intelligent Systems and Technology*, vol. 2, pp. 1-27, 2011.

[12] L. Apvrille and A. Apvrille. "Pre-filtering Mobile Malware with Heuristic Techniques", *Proceedings of GreHack' 2013*, Grenoble, Nov. 2013.

[13] "Face recognition homepage." [Online]. Available: $http://www.facerec.org/databases$

[14] "Face-place." [Online]. Available: $http://wiki.cnbc.cmu.edu/Face_{place}$

[15] W. Zhang, S. Shan, et al. "Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition", *Proc. ICCV*, pp. 786-791, 2005.

[16] A. Dantcheva, A. Ross, C. Chen, "Makeup challenges automated face recognition systems", *SPIE Newsroom*, 2013.

[17] T. Ojala, M. Pietikäinen, and T. Mäenpää. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", *PAMI*, vol 24, pp. 971-987, 2002.

[18] O. Ludwig, D. Delgado, et al. "Trainable Classifier-Fusion Schemes: An Application To Pedestrian Detection", *Int. IEEE Conf. On Intelligent Transportation Systems*, pp. 432-437, 2009.

[19] G. Rhodes, A. Sumich and G. Byatt. "Are average facial configurations attractive only because of their symmetry?", *Psychological Science*, vol. 10, pp. 52-58, 1999.

[20] S. Ueda and T. Koyama. "Influence of makeup on facial recognition", *Perception*, vol. 39, pp. 260, 2010.

[21] G. Guo, L. Wen and S. Yan. "Face Authentication with Makeup Changes", *IEEE Trans. on Circuits and Systems for Video Technology*, 2013.

[22] J. Hu, Y. Ge, et al. "Makeup-robust face verification", *ICASSP*, pp. 2342-2346, 2013.

[23] X. Wang, T. X. Han, S. Han. "An HOGLBP Human Detector with Partial Occlusion Handling", *IEEE Int. Conf. on Computer Vision*, pp. 32-39, 2009.

[24] J. Zhang, K. Huang, et al. "Boosted Local Structured HOG-LBP for Object Localization", *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 1393 - 1400, 2011.

[25] G. Gan, J. Cheng. "Pedestrian Detection Based on HOG-LBP Feature", *Int. Conf. on Computational Intelligence and Security*, pp. 1184 - 1187, 2011.

[26] R.-w. Sun, T. Rui, et al. "Pedestrian Detection by PCA-Based Mixed HOG-LBP Features", *ICIMCS*, pp. 92 - 95, 2012.

[27] J. Gao, Q. Cao. "Adaptive HOG-LBP Based Learning for Palm Tracking", *Int. Conf. on Computer and Information Application (ICCIA)*, pp. 543 - 546, 2012.

[28] Y. Yu, J. Zhang, et al, "Object Detection by Context and Boosted HOG-LBP", *ECCV workshop on PASCAL VOC*, 2010.

[29] M.B. Yilmaz, B. Yanikoglu, et al., "Offline Signature Verification Using Classifier Combination of HOG and LBP Features", *IEEE Int. Joint Conf. on Biometrics (IJCB)*, pp. 1-7, 2011.

[30] "Alligator" [Online]. Available: $http://perso.telecom-paristech.fr/apvrille/alligator.html$