Lips recognition for biometric identification systems

Boucetta Aldjia[1], Boussaad Leila[2]

[1]*Computer Science department, University of Batna2, Batna, Algeria*
[2]*Computer Science department, University of Batna2, Batna, Algeria*

Abstract
In recent years, researches in biometric methods have gained much attention and they have advanced to a wide scope in security concepts. Therefore, many biometric technologies have been developed and enhanced with many of the most successful security applications. Lately, lip-based biometric identification becomes one of the most relevant emerging tools, which comes from criminal and forensic real-life applications. The main purpose of this paper is to prove the benefit of lips as a biometric modality, by using both handcraft and deep-learning based feature extraction methods. So, we consider three different techniques, Histogram of Oriented Gradients(HOG), Local Binary Pattern(LBP) and pretrained Deep-CNN. All results are confirmed by a ten-fold cross-validation method using two datasets, NITR-LipV1 and database1. The mean accuracy is found to be very high in all the experiments carried out. Also the feature extraction using the Inceptionv3 model always achieve highest mean accuracy.

Keywords
human identification, lips recognition, Histogram of Oriented Gradients (HOG), Local Binary Pattern (LBP), Convolutional Neural Network (CNN)

1. Introduction

Biometric human identification methods have recently gained a lot of attention, since they easily address most traditional identification issues.

In biometric human identification systems, users are recognized by who they are and not by anything to keep in mind or take with them [1]. Several known methods of human identification, like face, iris, retina, etc. are developed and optimized, but there are still need to emerging and innovative solutions[2]. Some of the new biometrics modalities are: heartbeat (ECG)[3], EEG biometrics[4], dental radiograph[5] and finger-nails [6].

Recently, lips recognition [7] has been proposed as a new relevant emerging kind of biometrics, which derived from criminal and forensic real-life applications.

Studies on lip-prints date back to 1950's, where extensive studies on lip-traces have been performed by Japanese researchers, without indicating or proposing a useful application of them. At the beginning of the 1970's, based on the lip prints of 1364 people between the ages of 3 and 60 of both sexes, Yasuo Tsuchihachi and Kazuo Suzuki[8] demonstrated that lip prints are unique and stable for an individual. Further, they suggested the ability to use lip prints in human identification. Also, lip prints are used to gender determination of the examined subject [9].

Lip prints properties have been successfuly applied as a subdiscipline of dactyloscopy, to human identity confirmation by forensic specialists and police. Precisely, when examining the features of the human lips, the anatomical patterns on the lips are often considered. Different classifications have been developed by authors. However, no classification has yet been recognized internationally and each author creates more or less his own by modifying some already existing. For example, Yasuo Tsuchihachi and Kazuo Suzuki identified 6 lip models based on the patterns found, figure 1 and table 1 present this classification and their description.
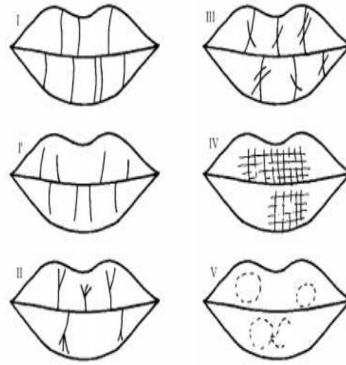

Figure 1: Yasuo Tsuchihachi and Kazuo Suzuki classification of lip patterns [8].

Table 1
Yasuo Tsuchihachi and Kazuo Suzuki description of lip patterns [8].

Type of lip pattern	Description
Type I	Complete straight grooves
Type I'	Partial straight grooves
Type II	Branched grooves
Type III	Intersected grooves
Type IV	Reticular grooves
Type V	Other patterns

Unfortunately, in image analysis based recognition system, these features cannot be used because they are difficult to extract from the acquired images. Therefore, in our approach, we do not use the features of lip prints, but we focus on the features extracted from the lips in a static face image.

Authors in [2] consider that the use of lips as modality for human identification is very interesting, since lips are passive biometrics, in which images can be obtained without the knowledge of the person being examined. Also, lips are usually visible and not hidden or covered with anything. Further lips can be implemented in a hybrid lips-face or lips-voice biometric systems.

In this paper, we study the efficiency of lips based biometrics systems using three feature extraction techniques, i.e., Histogram of Oriented Gradients (HOG), Local Binary Pattern (LBP) and inceptionv3 pre-trained Deep-CNN as automatic feature extractor.

In the recognition step, two classifiers are used for lip recognition, which are K-nearest neighbor (K-nn) and support vector machine (SVM).

The rest of this paper is organized like this: Literature survey is presented in section 2. The proposed lip based biometrics system is presented in section 3. The experimental results and discussion in section 4. Finally, conclusions and future works are drawn in section 5.

2. Literature survey

Lip biometrics have not been very much studied so far. In this section, we will cite some works that have been realized in this field:

In [10], authors consider both physiological and behavioral information of the lip as biometric. Their results prove that both the static texture feature of lips and the dynamic shape deformation feature can give satisfactory accuracy.

Also, in [2], there have been several investigations to recognize a person directly from the shape and contour of lips, where lip region of interest is determined based on the color distribution around the lip area.

In addition, Choras in his various researches [11] [1] [2], proved that the lip can be used as a strong biometric trait.

Further, Hsu et al. [12] proposed lip recognition method based on active basis model, Particle Swarm Optimization (PSO) algorithm to define the best combination of parameters, and SVM to obtain classified results.

In th same context, Bakshi et al. [13] applied tow techniques to extract the local features from grayscale lip images, viz. SIFT (Scale Invariant Feature Transform) and SURF (Speeded Up Robust Features). The precision turns out to be very high ($> 90\%$) in the two experiments carried out.

In recent work, Bakshi et al. [14] studied the fusion of the shape and texture characteristics of the lip image to verify a person's identity in mobile devices.

Moreover, Wrobel et al. [15] proposed an efficient lip-based biometric recognition approach using a Probabilistic Neural Network (PNN) for verification purpose. The results obtained by PNN are improved by a Particle Swarm Optimization (PSO) technique.

Furthermore, in the context of age progression, Clare and Jain [16] proved that lips have more discriminating information than the nose, and it seem to be more stable and identifiable facial part. Also, Boussaad et al. [17] proposed a component-based approach for age invariant face recognition in which Deep-based features are computed from separated facial parts including lips.

3. Proposed lip biometric approach

In this section, we present the proposed lip biometric system, its overview is illustrated in Fig. 2 and the details of each step are described in the following subsections:
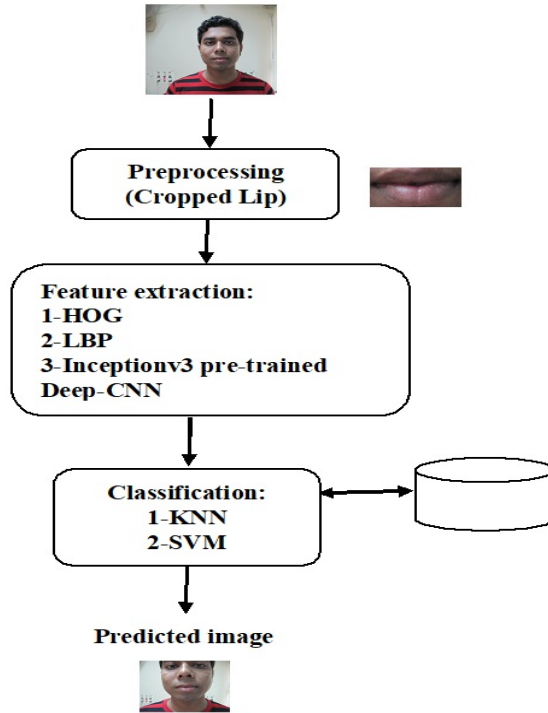
Figure 2: Basic Design of the Proposed Lip Biometric System.

3.1. Feature extraction phase

In this study, three different feature extraction algorithms are used. The details of these algorithms are given below.

1. Histogram of Oriented Gradients(HOG): HOG is a feature descriptor that focuses on the structure or shape of an object. It is considered as one of the well recognized features due to its superior performance and relatively simple computation [18].
   It is initially proposed for the detection of pedestrians [19]. It counts the occurrences of gradient orientation in localized parts of an image. The main steps to calculate HOG features can be summarized as follows [18].

   - Gradient calculation: In this step, the spatial gradients in the vertical and horizontal directions are calculated, then used to calculate the gradient magnitudes and angles.
   - Orientation binning: In this step, the image is divided into small connected regions called cells and according to the gradient angle, the gradient magnitude of each pixel in a cell is voted into different orientation bins.
   - Feature description: In this step, the adjacent cells are grouped into blocks . Each block is normalized by its L2 norm, then to form a descriptor, the normalized block histograms in a detection window are concatenated.

An example of HOG features over the original image is illustrated in Fig.3.
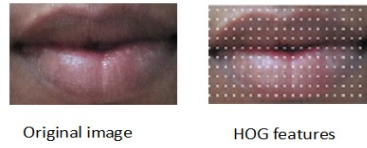
Figure 3: An Example of HOG Features over the Original Image.

2. Local Binary Pattern(LBP): LBP features are originally proposed for texture analysis, which labels the image pixels by thresholding the neighborhood of each pixel and considers the result as a binary number [20]. The most important properties of LBP features are their tolerance to illumination changes and their ease of calculation. LBP proceeds as illustrated in Fig. 4; each pixel is compared to its eight neighbors in a 3×3 neighborhood by subtracting the value of the central pixel; Strictly negative values are coded with 0 and others with 1; A binary number is obtained by concatenating all these binary codes clockwise from the one at the top left. Derived binary numbers are called Local Binary Patterns or LBP codes [21].
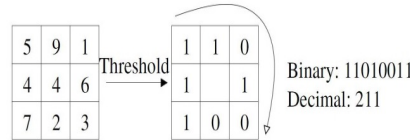

Figure 4: An Example of the Basic LBP [21].

Initially, the size of the LBP operator was limited to only 3×3 neighborhood, it cannot capture dominant features with large scale structures. To overcome this limitation, the descriptor was generalized to use neighborhoods of different sizes [22].

A local neighborhood is defined as a set of evenly spaced sample points on a circle centered on the pixel to be labeled and the sample points that do not fall in the pixels are interpolated by bilinear interpolation, thus allowing any radius and any number of sampling points in the neighborhood.

Fig.5 illustrates some examples of the extended LBP operators, where (P, R) denotes a neighborhood of P sampling points on a circle with radius R.

The LBP operator used in this paper is the circular (8,1) neighborhood.

3. Pre-trained Inception-v3 Deep CNN model:
Inceptionv3 [23] is a CNN architecture from the Inception family, including three types of Inception modules (Inception A, Inception B and Inception C) as shown in Fig. 6. Each Inception module is composed of several convolutional layers and pooling layers in parallel, which can generate discriminatory features and reduce the number of parameters. In Inception-v3, three Inception A modules, five Inception B modules and two Inception C modules are stacked in series.

In our experiment, feature extraction is computed from the pooling layer 'avg_pool' of the pre-trained Deep-CNN InceptionV3 model.

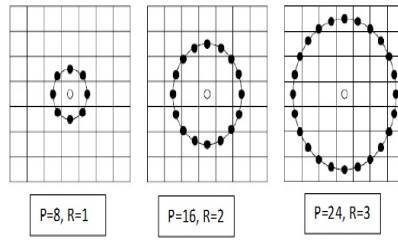A general diagram of the Inception-v3 model is shown in Fig.7.

Figure 5: Three Examples of the extended LBP operator with circular neighborhood [22].
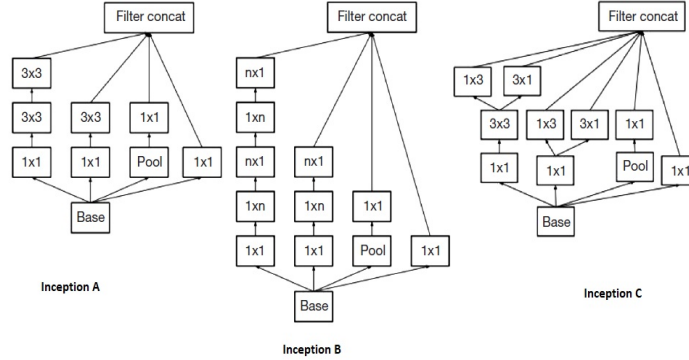

Figure 6: The Inception Modules of Inception-v3 [23].

3.2. Classification phase

1. K-Nearest Neighbor (K-nn) classifier:
   The KNN classifier is a very simple non-parametric classification method proposed by Cover and Hart in 1968 [25]. Despite the simplicity of the algorithm, it works very well and it is an important reference method. Due to its clear principles and excellent classification performance, it is used in several applications.

   The KNN method is based on K which means the number of nearest neighbors. Decision rules can be described as follows [26]:

   - If K = 1, the KNN method is called NN (nearest neighbor) method. Firstly, calculate the distances between the test sample x and all training samples by a distance function (Euclidean, Manhattan,...). Secondly, find the nearest neighbor, that is, the nearest training sample to x. Finally, give x the class label identical to nearest neighbors.
   - When $K \neq 1$, KNN tries to find the K nearest neighbors of x. Among these K nearest neighbors, if the samples belonging to class i has the largest quantity, the class label of x can be marked with i.
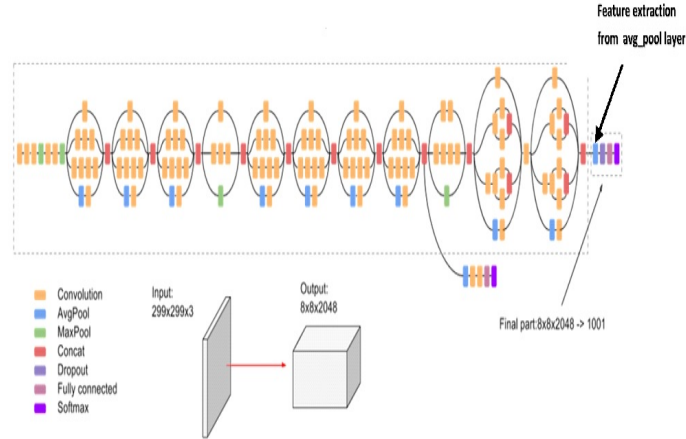
Figure 7: Inception-v3 model [24].

2. Support Vector Machine(SVM) classifier:
   SVM is a related supervised learning method used for classification and regression, it was originally proposed by Vapnik [27]. It is the most preferred by many because it can provide significant accuracy with computational efficiency.
   The objective of the SVM algorithm is to find the best decision line or boundary (hyperplane) that can separate the n-dimensional space into classes so that we can easily place the new data point in the correct category in the future. Fig. 8 shows an example of the classification process of SVM.
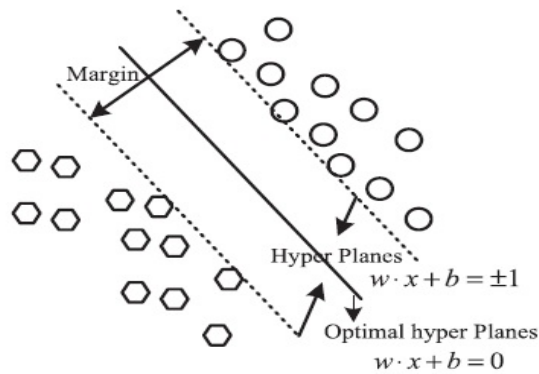

Figure 8: The SVM Classification Process [28].

In addition, the SVM can efficiently performs nonlinear classification using the kernel function, by mapping its inputs into large feature spaces. The kernel function plays a crucial role in SVM, because it is a kind of measure of similarity between the input object. The proper selection of the kernel function will affect the accuracy of the model. There are four types of kernel function available for SVM which include linear, Radial Basic

Function (RBF), polynomial and sigmoid [29]. Among these popular kernel functions, RBF is the most popular choice due to its less numerical difficulties and less hyperparameters than the polynomial kernel.

4. Experiments and results

In this section we will discuss the used datasets , and the experimental results that were generated.

4.1. Description of databases

The proposed identification system is evaluated on two publicly-available lip databases, namely NITRLipV1 [30] and Database1 [31] databases.

1. NITRLipV1 database captured by Canon PowerShot A1100IS with F2.7 aperture and shutter speed varying from 1/60s to 1/25s. The database images were collected from 15 Indian volunteers, including men and women with age ranged from 20 to 40 years. This database is composed of 109 color images characterized by a variety of illumination conditions saved in JPEG format.
2. The Database1 database contains 23 color and grayscale images of objects, 5 images per object. Images have different sizes from the range 3096×3456 to 4128×4608 pixels and have various illumination and position conditions.

Some examples from the NITRLipV1 and Database1 databases are shown in Fig. 9.

Figure 9: Examples from the NITRLipV1 and Database1 Databases.

4.2. Experimental evaluations

The entire algorithm was evaluated using the Matlab (R2018b) environment. All images in the NITRLipV1 and Database1 databases are used for training and test and the experiments results

are reported in terms of average recognition accuracy rates following a 10-fold cross-validation scheme.

The accuracy rate is defined by Eq. 1.

$$\text{Accuracy} = \frac{TPR + TNR}{TPR + TNR + FPR + FNR} \times 100 \tag{1}$$

where TPR (True Positive Rate) is the probability that authorized users are correctly recognized on the total number tested, TNR (True Negative Rate) is the probability of authorized users who are not recognized on the total number tested.

FPR (False Positive Rate) describes the probability of unauthorized users who are recognized over the total number tested.

FNR (False Negative Rate) describes the probability of unauthorized users who are not falsely recognized over the total number tested.

The obtained results from the lip biometric system using the three feature extraction techniques, namely, HOG, LBP and inceptionv3 and two classifiers KNN and SVM are shown in Table 2, Fig. 10 for the NITRLipV1 database and Table 3, Fig. 11 for the Database1 database.

Table 2
The Obtained Results for the NITRLipV1 Database.

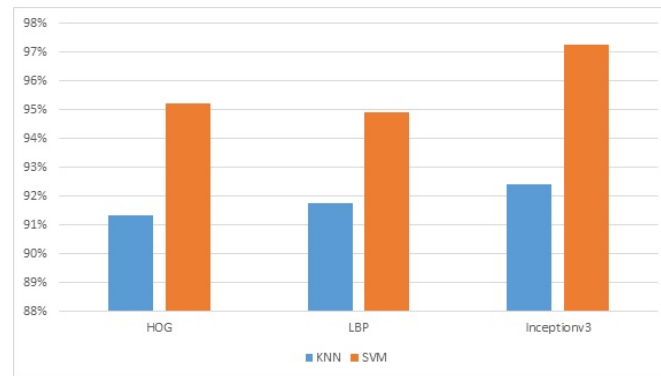| Feature Extraction Technique	KNN	SVM
HOG	91.33%	95.20%
LBP	91.76%	94.91%
Inceptionv3	92.42%	97.26 %

Figure 10: Recognition Accuracy Rates for the NITRLipV1 Database.

From the results reported in Table2, Table3, Fig. 10 and Fig. 11, we can make the following observations:

1. The accuracy rates provided with the SVM classifier always exceed the results given by the KNN classifier, which clearly proves the powerful of SVM.

Table 3
The Obtained Results for the Database1 Database.

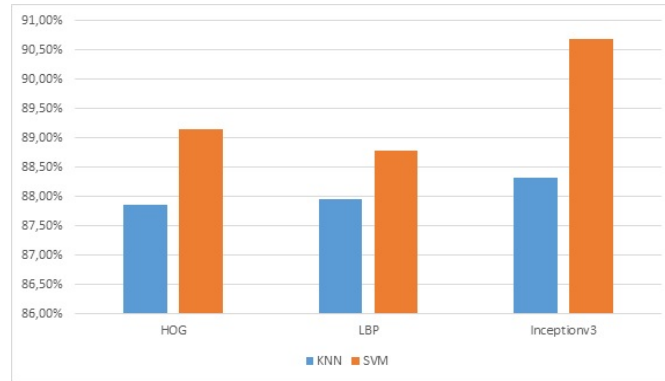| Feature Extraction Technique	KNN	SVM
HOG	87.85%	89.14%
LBP	87.96%	88.78%
Inceptionv3	88.31%	90.68 %

Figure 11: Recognition Accuracy Rates for the Database1 Database.

2. The pretrained Inceptionv3 CNN model appears to be a great tool for feature extraction, where the highest accuracy rate that is 97.26 % is obtained by SVM classifier for the NITRLipV1 Database.
3. These results show that lips may be effective biometric modality for identification system.

We are also conducting additional experiments. We use different sizes of training sets to study how the amount of training data affects the accuracy of the test data sets. We randomly choose 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90% images for training. The results are shown in Fig. 12 and Fig.13.

As shown in Fig. 12 and Fig. 13, the accuracy of all test sets increases as the size of the training set increases and always the results obtained with inceptionv3 are better than HOG and LBP.

5. Conclusion and future works

In this paper, we have study the efficiency of lips as biometric modality for identification systems using three different techniques for feature extraction, namely, histogram of oriented (HOG), local binary pattern (LBP) and inceptionv3 pre-trained Deep-CNN.

In the recognition step, these feature vectors are used as input data for a K-nearest neighbor (K-nn) or Support Vector Machine (SVM) classifier. From the obtained results, we can conclude that lip based biometric system offers a promising accuracy rate.
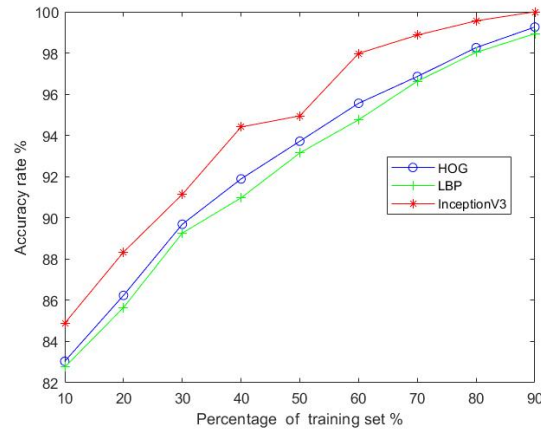
Figure 12: Accuracy Rate for Different Sizes of Training Set for the NITRLipV1 Database using SVM Classifier.
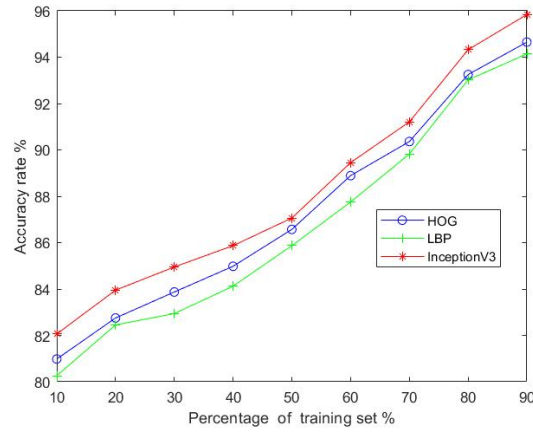

Figure 13: Accuracy Rate for Different Sizes of Training Set for the Database1 Database using SVM Classifier.

This motivates us to investigate in further researches in this field. In future studies, we aims to expand this research with other lip features and to evaluate the performance of the proposed method using other databases. Also, it is very interesting to implement it in a multimodal biometric system to improve performance of other biometrics.

References

[1] M. Choras, Human lips recognition, in: Computer recognition systems 2, Springer, 2007, pp. 838–843.

[2] M. Choraś, The lip as a biometric, Pattern Analysis and Applications 13 (2010) 105–112.

[3] M. Ingale, R. Cordeiro, S. Thentu, Y. Park, N. Karimian, Ecg biometric authentication: A comparative analysis, IEEE Access 8 (2020) 117853–117866.

[4] B. Goudiaby, A. Othmani, A. Nait-Ali, Eeg biometrics for person verification, in: Hidden Biometrics, Springer, 2020, pp. 45–69.

[5] M. Banday, A. H. Mir, Dental biometric identification system using ar model, in: TENCON 2019-2019 IEEE Region 10 Conference (TENCON), IEEE, 2019, pp. 2363–2369.

[6] S. N. G. Dessai, S. Borkar, Finger nail recognition system using neural network, in: 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2018 2nd International Conference on, IEEE, 2018, pp. 656–659.

[7] S. Das, K. Muhammad, S. Bakshi, I. Mukherjee, P. K. Sa, A. K. Sangaiah, A. Bruno, Lip biometric template security framework using spatial steganography, Pattern Recognition Letters 126 (2019) 102–110.

[8] K. Suzuki, Y. Tsuchihashi, Personal identification by means of lip prints, J Forensic Med 17 (1970) 52–57.

[9] S. Vahanwala, C. Nayak, S. Pagare, Study of lip prints as aid for sex determination, Medico-legal update 5 (2005) 93–98.

[10] S.-L. Wang, A. W.-C. Liew, Physiological and behavioral lip biometrics: A comprehensive study of their discriminative power, Pattern Recognition 45 (2012) 3328–3335.

[11] M. Choraś, Emerging methods of biometrics human identification, in: Second International Conference on Innovative Computing, Informatio and Control (ICICIC 2007), IEEE, 2007, pp. 365–365.

[12] C.-Y. Hsu, C.-H. Yang, Y.-C. Chen, M.-c. Tsai, A pso-svm lips recognition method based on active basis model, in: 2010 Fourth International Conference on Genetic and Evolutionary Computing, IEEE, 2010, pp. 743–747.

[13] S. Bakshi, R. Raman, P. K. Sa, Lip pattern recognition based on local feature extraction, in: 2011 Annual IEEE India Conference, IEEE, 2011, pp. 1–4.

[14] R. Raman, P. K. Sa, B. Majhi, S. Bakshi, Fusion of shape and texture features for lip biometry in mobile devices, Mobile Biometrics 3 (2017) 155.

[15] K. Wrobel, R. Doroz, P. Porwik, J. Naruniec, M. Kowalski, Using a probabilistic neural network for lip-based biometric verification, Engineering Applications of Artificial Intelligence 64 (2017) 112–127.

[16] B. Klare, A. K. Jain, Face recognition across time lapse: On learning feature subspaces, in: 2011 International Joint Conference on Biometrics (IJCB), IEEE, 2011, pp. 1–8.

[17] L. Boussaad, A. Boucetta, An effective component-based age-invariant face recognition using discriminant correlation analysis, Journal of King Saud University-Computer and Information Sciences (2020).

[18] W. Zhou, S. Gao, L. Zhang, X. Lou, Histogram of oriented gradients feature extraction from raw bayer pattern images, IEEE Transactions on Circuits and Systems II: Express

Briefs 67 (2020) 946–950.

[19] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), volume 1, Ieee, 2005, pp. 886–893.

[20] H. Wang, J. Hu, W. Deng, Face feature extraction: a complete review, IEEE Access 6 (2017) 6001–6039.

[21] D. Huang, C. Shan, M. Ardabilian, Y. Wang, L. Chen, Local binary patterns and its application to facial image analysis: a survey, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 41 (2011) 765–781.

[22] S. Dey, D. N. Tibarewala, S. P. Maity, A. Barui, Automated detection of early oral cancer trends in habitual smokers, in: Soft Computing Based Medical Image Analysis, Elsevier, 2018, pp. 83–107.

[23] Q. Guan, X. Wan, H. Lu, B. Ping, D. Li, L. Wang, Y. Zhu, Y. Wang, J. Xiang, Deep convolutional neural network inception-v3 model for differential diagnosing of lymph node in cytological images: a pilot study, Annals of translational medicine 7 (2019).

[24] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2818–2826.

[25] T. Cover, P. Hart, Nearest neighbor pattern classification, IEEE transactions on information theory 13 (1967) 21–27.

[26] Y. Li, B. Cheng, An improved k-nearest neighbor algorithm and its application to high resolution remote sensing image classification, in: 2009 17th International Conference on Geoinformatics, Ieee, 2009, pp. 1–4.

[27] C. Cortes, V. Vapnik, Support-vector networks, Machine learning 20 (1995) 273–297.

[28] J. Wei, Z. Jian-Qi, Z. Xiang, Face recognition method based on support vector machine and particle swarm optimization, Expert Systems with Applications 38 (2011) 4390–4393.

[29] R. Soentpiet, Advances in kernel methods: support vector learning, MIT press, 1999.

[30] S. Bakshi, R. Raman, P. K. Sa, Nitrlipv1: a constrained lip database captured in visible spectrum, ACM SIGBioinformatics Record 6 (2016) 1–1.

[31] Database1, 2001. URL: http://www.biometrics.us.edu.pl/uploads/database1.zip.