

AUTOMATIC AGE ESTIMATION AND GENDER CLASSIFICATION IN THE WILD

*SE. Bekhouche** *A. Ouafi** *A. Benlamoudi†* *A. Taleb-Ahmed‡* *A. Hadid§*

*LESIA Laboratory, University of Biskra, Algeria

†Laboratory of LAGE, University of Ouargla, Algeria

‡LAMIH Laboratory, UMR CNRS 8201 UVHC, University of Valenciennes, France

§ Center for Machine Vision Research, University of Oulu, Finland

ABSTRACT

Automatic age estimation and gender classification through facial images are attractive topics in computer vision. They can be used in many real-life applications such as face recognition and internet safety for minors. In this paper, we present a novel approach for age estimation and gender classification under uncontrolled conditions following the standard protocols for fair comparison. Our proposed approach is based on Multi Level Local Binary Pattern (ML-LBP) features which are extracted from normalized face images. Two different Support Vector Machines (SVM) models are used to predict the age group and the gender of a person. The experimental results on benchmark Image of Groups dataset showed the superiority of our approach compared to that of the state-of-the-art methods.

Index Terms— Age estimation, Gender classification, Local Binary Pattern, Support Vector Machines

1. INTRODUCTION

Age estimation and gender classification belong to soft biometrics that provides ancillary information of an individual's identity information. Age estimation is defined as the task of determining the age of a person based on biometric features. It can be considered as either a multi-class problem (age classification) or a regression problem (age estimation) [1]. Gender classification is a binary classification problem (male vs. female).

There are several studies on automatic demographic classification (age, gender and ethnicity) from facial images. Recently, several applications have emerged to make the demographic classification useful. These applications include access control, re-identification in surveillance videos, law enforcement, integrity of face images in social media, intelligent advertising, and human-computer interaction, to cite few [2].

Most of age estimation and gender classification studies have utilized face images captured in controlled conditions. Many variations may appear in the image of a person's face. These variations may affect the ability of the computer vision system to estimate the age or recognize the gender. They

are due to many factors which can be divided into: human factors (race, facial expression etc.) and capture process factors (pose, illumination, quality etc.).

Gallagher et al. [3] studied contextual features for capturing the structure of people images. Instead of treating each face independently, they extracted features which cover the structure of the group from persons' faces. Firstly, they used the social context features, then they tried to use appearance features. Finally, they combined context and appearance features together achieving the accuracy of 42.9% and 74.1% for age and gender respectively.

Shan [4] investigated age estimation and gender classification by treating each face independently. He focused on appearance features exactly on Local Binary Patterns (LBP) and Gabor features as face representation, then he adopted Adaboost to learn the discriminative local features. The best performance in his experiments is with the boosted LBP based on SVM classifier with an accuracy of 50.3% for age and 74.9% for gender.

Li et al. [5] focused on facial age estimation based on ordinal discriminative feature learning. They tried to remove redundant information from both the locality information and ordinal information by minimizing non linear correlation and rank correlation, using different feature selection algorithms (Laplacian Score, LAR, Fisher Score, Rank Boost and PLO). The Piecewise Linear Orthonormal (PLO) gave the best results for the age estimation with an accuracy of 48.5%.

Ylioinas et al. [6] studied automatic age classification using LBP variants. They proposed a method based on a combination of LBP variants encoding the structure of elongated facial micro-patterns and their strength. Their experimental results gave an accuracy of 51.7% for age classification.

Fu et al. [7] wrote a survey about age synthesis and estimation. They divided the age estimation systems into two concatenated modules: age image representation and age estimation techniques. There are five main models in age image representation, they are: Anthropometric Models, Active Appearance Models, Aging Pattern Subspace, Age Manifold and Appearance Models. And the age estimation techniques are: Classification, Regression and Hybrid techniques. They

mentioned the majority of the aging databases including the Images of Groups Dataset. They summarize and compare different age estimation methods on different databases.

Ng et al. [8] wrote a survey about human gender recognition from face, gait and body. They mentioned some face preprocessing methods that may be applied in gender classification systems like normalization for contrast and brightness. They also categorized feature extraction methods for face gender classification into geometric-based and appearance-based methods. As in any survey, a comparison of different methods and databases was summarized.

In our work, we propose a new approach to estimate the age and recognize the gender in uncontrolled conditions based on facial images. Our approach uses faces detected directly from multi-faces images without alignment, also it extracts the facial features using the multi level local binary patterns (ML-LBP) which are powerful means of texture description. These texture features are then classified by multi-class Support Vector Machine (SVM) via one-against-one strategy. We conduct experiments on the Images of Groups Dataset. We also compared multi block and multi level LBP results.

The rest of the paper is organized as follows: in Section 2 we explain our approach. The experiments and the results are given in Section 3. In section 4 we give the conclusion and the future plans.

2. PROPOSED APPROACH

Since the local binary patterns (LBP) was proposed, it became useful in texture classification like face recognition, authentication, face detection, facial expression recognition, gender classification and age estimation [9]. Due to the above we used LBP in our approach. The LBP operator detects microstructures such as edges, spots and flat areas. It is one of the best performing texture descriptors [10, 11].

The original LBP operator works in a 3×3 neighborhood pixels, each pixel can be labelled by using the center value as a threshold and considering the result as a binary number.

$LBP_{P,R}$ is almost used for pixel neighborhoods and it refers to P sampling points on a circle of radius R . The value of the LBP code of a pixel (x_c, y_c) is given by:

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

where g_c corresponds to the gray value of the center pixel (x_c, y_c) , g_p refers to gray values of P equally spaced pixels on a circle of radius R , and s defines a thresholding function as follows:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

In [12], the authors revealed that it is possible to use only a subset of the 2^P local binary patterns to describe textured

images. This subset is called uniform patterns. LBP is called uniform if the circular binary pattern contains in maximum 2 transitions from 0 to 1 and 1 to 0. In our work we used $LBP_{(8,2)}^{u2}$ which have 58 uniform bins plus a non-uniform one (see Fig 1).

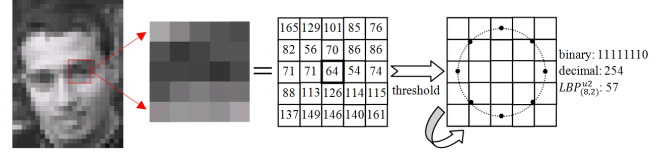


Fig. 1. Example of $LBP_{(8,2)}^{u2}$.

In [13], we divided the face ROI (region of interest) into 3×4 sub-blocks and applied LBP features on each sub-block. This method is called multi block local binary patterns (MB-LBP). The MB-LBP method provides the most important information, i.e it gives better results and these results differentiates according to the division of the face ROI and experiments proved this.

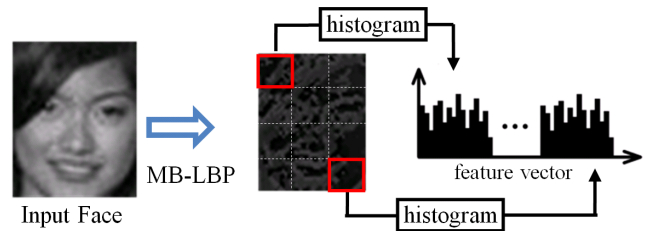


Fig. 2. Example of MB-LBP features extraction with 4×3 sub-blocks.

Based on the $LBP_{(8,2)}^{u2}$ method, instead of using single MB-LBP division we use ML-LBP which is firstly introduced in [14]. They used ML-LBP in age estimation to solve a regression problem. The main idea of ML-LBP is to extract features from different MB-LBP divisions and then combine them. In other words, extracting features from the whole image, then dividing the image into 2^2 sub-blocks and extracting the features from each sub-blocks and so on until we reach the intended level. The final result of ML-LBP is $1^2 + 2^2 + \dots + n^2$ histograms. We combine these histograms to get the feature vector. Figure 3 explains our approach.

To classify each subject into his/her age group and gender, we use multi-class Support Vector Machine (SVM) via one-against-one strategy with a non-linear RBF kernel. Multiclass SVMs are generally considered as unreliable for arbitrary data, since in many cases no single mathematical function exists to separate all classes of data from one another. The one-against-one strategy, also known as one-versus-one method, an SVM is constructed for every pair of classes by training it to discriminate the two classes. Thus, for a problem with c classes, $c(c-1)/2$ SVMs are trained to distinguish the

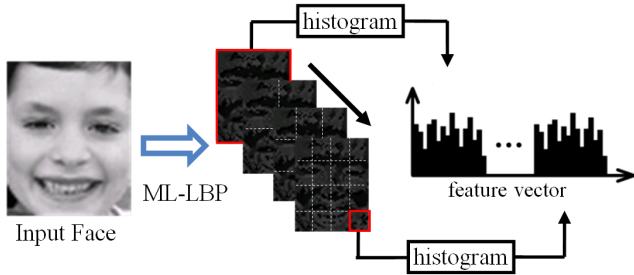


Fig. 3. Example of ML-LBP features extraction with 4 level.

samples of one class from the samples of another class. The maxwins strategy is commonly used to determine the class [15].

3. EXPERIMENTS

To evaluate the performance of the proposed approach, we use the Images of Groups database¹ (see Fig. 4) which contains 5080 images with 28231 faces labeled with age and gender. There are seven age categories as follows: 0-2, 3-7, 8-12, 13-19, 20-36, 37-65, and 66+ years, roughly corresponding to different life stages. In some images people are sitting, laying, or standing on elevated surfaces. People often have dark glasses, face occlusions, or unusual facial expressions [3].



Fig. 4. Example of detected faces from Image of Groups database.

3.1. Experimental Settings

As we are not studying face detection, in [3] they manually add missed faces, but 86% of the faces are automatically

¹<http://chenlab.ece.cornell.edu/people/Andy/ImagesOfGroups.html>

found. Following the BeFIT² recommendations, we trained the classifiers using a random selection of 3500 faces, having equal number of samples per each age category (500 per class). Tests was performed on an independent uniformly distributed set of 1050 face samples (150 per class). For the gender classification we used 23218 face samples as a train and 1881 as a test as in [3]. The table below shows the distribution of the 25099 faces used in the experiments.

	0-2	3-7	8-12	13-19	20-36	37-65	+66
Male	382	664	386	704	6452	2887	501
	44	56	45	61	440	241	56
Female	337	625	365	641	5988	2799	487
	44	55	44	60	439	241	55
Total	719	1289	751	1345	12440	5686	988
	88	111	89	121	879	482	111

Table 1. The age/gender distribution of the faces used for the gender classification

3.2. Experimental Results

The measures of the age estimation are evaluated by the accuracy of an exact match (AEM) and the accuracy of allowing an error of one age category (AEO) (e.g. an 0-2 year old predicted as 3-7 year old). They are defined as follows [5]:

$$AEM = \frac{\hat{N}_m}{N'} \times 100\% \quad (3)$$

$$AEO = \frac{\hat{N}_o}{N'} \times 100\% \quad (4)$$

where \hat{N}_m is the number of an exact match with N' test images, and \hat{N}_o is the number of correct prediction when allows an error of one age category. The measure of the gender classification is evaluated by:

$$Acc = \frac{\hat{N}}{N'} \times 100\% \quad (5)$$

where \hat{N} is the number of true classification.

To gain the best results, we compare different LBP variants, we find that $LBP_{(8,2)}^{u2}$ gives the best results. We also compare different MB-LBP uniform divisions, the results are shown in Table 2.

MB-LBP divisions	AEM	AEO	ACC
2×2	38.33%	68.19%	62.95%
3×3	43.14%	79.14%	70.09%
4×4	49.19%	82.38%	73.23%
5×5	46.57%	81.61%	71.61%

Table 2. Comparison between the different MB-LBP uniform divisions

²<http://fipa.cs.kit.edu/425.php>

Finally, we apply ML-LBP and change the level until the results began to fall. Table 3 shows the changing in the results between different levels. All these experiments are evaluated by multi-class SVM based on one-class SVM, the implementation of matlab was used [16].

ML-LBP level	AEM	AEO	ACC
1	29.14%	54.76%	63.10%
2	39.71%	69.33%	69.91%
3	51.14%	83.52%	75.06%
4	55.05%	88.19%	78.25%
5	53.52%	85.81%	76.50%

Table 3. Comparison between the different levels of ML-LBP

We have obtained $AEM = 55.05\%$ and $AEO = 88.19\%$. The confusion matrix of the age estimation is shown in Table 4. From the table, the classification of infants (0-2 years old) and over-ages (66+ years old) have the highest accuracy compared to the other age groups. However for other groups we remark that it is difficult to determine the exact group due to several reasons. The most important is the converged ages like 3-7 and 8-12 groups which represent the childhood.

	0-2	3-7	8-12	13-19	20-36	35-65	66+
0-2	113	27	4	3	1	1	1
3-7	22	82	34	7	1	4	0
8-12	3	43	64	25	5	4	6
13-19	1	11	26	74	27	7	4
20-36	2	6	14	29	68	24	7
35-65	2	3	7	7	38	66	27
66+	1	2	4	2	4	26	111

Table 4. Confusion matrix of age estimation

The confusion matrix of the gender classification is shown in Table 5. From this table, we see that our approach recognize males easily compared to females. For this classification, We have obtained an accuracy of **78.25%** (cf. Table 6).

	male	female
male	789	154
female	255	683

Table 5. Confusion matrix of gender classification

We conducted experiments on the Images of Groups dataset [3] for age estimation and gender classification. Table 6 shows the comparison against the state-of-the-art. Its confirm the superiority of our approach in the case of gender classification and age classification with exact match.

Approach	Age		Gender
	<i>AEM</i>	<i>AEO</i>	<i>ACC</i>
Context [3]	32.9%	64.4%	66.9%
LaplacianScore [5]	35.5%	74.5%	N/A
Appearance [3]	38.4%	71.3%	69.6%
FisherScore [5]	42.8%	83.7%	N/A
Appearance+Context [3]	42.9%	78.1%	74.1%
Gabor + Adaboost [4]	43.7%	80.7%	70.2%
LAR [5]	44.8%	84.9%	N/A
RankBoost [5]	44.8%	84.5%	N/A
LBP + Adaboost [4]	44.9%	83.0%	71.0%
boosted Gabor + SVM [4]	48.4%	84.4%	73.3%
PLO [5]	48.5%	88.0%	N/A
boosted LBP + SVM [4]	50.3%	87.1%	74.9%
CLBP_M + SVM [6]	51.7%	88.7%	N/A
Ours	55.05%	88.19%	78.25%

Table 6. Age estimation and gender classification results and comparison to state-of-the-art

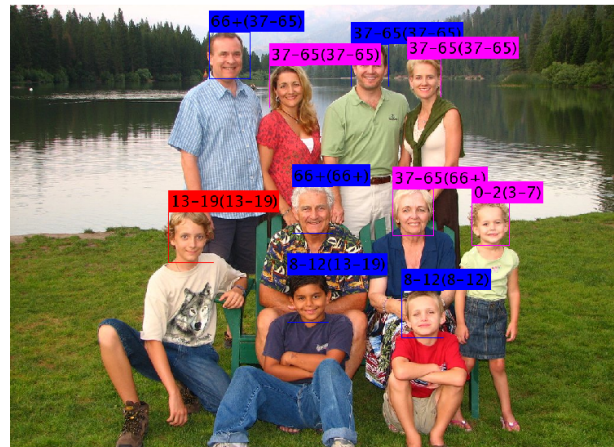


Fig. 5. Examples of age estimation (ground truth in parentheses) and gender classification (male: blue, female: magenta and wrong gender: red).

4. CONCLUSION

In this paper, we described a novel approach for demographic (age group and gender) classification based on Multi Level Local Binary Pattern features and One-class SVM with non-linear kernel (RBF) classifier. The experimental results showed that our approach provides a better performance than previous approaches. As a future work, we will study the effect of age on gender recognition and the other human factors (race, facial expression, ..etc.) on age estimation. We will also consider feature selection to boost the performance.

5. REFERENCES

- [1] Petra Grd, "Introduction to human age estimation using face images," *Research Papers Faculty of Materials Science and Technology Slovak University of Technology*, vol. 21, no. Special Issue, pp. 24–30, 2013.
- [2] Hu Han and Anil K. Jain, "Age, gender and race estimation from unconstrained face images," Tech. Rep. MSU-CSE-14-5, Department of Computer Science, Michigan State University, East Lansing, Michigan, July 2014.
- [3] A.C. Gallagher and Tsuhan Chen, "Understanding images of groups of people," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, June 2009, pp. 256–263.
- [4] Caifeng Shan, "Learning local features for age estimation on real-life faces," in *Proceedings of the 1st ACM International Workshop on Multimodal Pervasive Video Analysis*, New York, NY, USA, 2010, MPVA '10, pp. 23–28, ACM.
- [5] Changsheng Li, Qingshan Liu, Jing Liu, and Hanqing Lu, "Learning ordinal discriminative features for age estimation," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, June 2012, pp. 2570–2577.
- [6] J. Ylioinas, A. Hadid, and M. Pietikainen, "Age classification in unconstrained conditions using lbp variants," in *Pattern Recognition (ICPR), 2012 21st International Conference on*, Nov 2012, pp. 1257–1260.
- [7] Yun Fu, Guodong Guo, and Thomas S Huang, "Age synthesis and estimation via faces: A survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 11, pp. 1955–1976, 2010.
- [8] ChoonBoon Ng, YongHaur Tay, and Bok-Min Goi, "Recognizing human gender in computer vision: A survey," in *PRICAI 2012: Trends in Artificial Intelligence*, Patricia Anthony, Mitsuru Ishizuka, and Dickson Lukose, Eds., vol. 7458 of *Lecture Notes in Computer Science*, pp. 335–346. Springer Berlin Heidelberg, 2012.
- [9] Matti Pietikäinen, Abdenour Hadid, Guoying Zhao, and Timo Ahonen, *Computer vision using local binary patterns*, vol. 40, Springer, 2011.
- [10] Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen, "Face recognition with local binary patterns," in *Computer vision-eccv 2004*, pp. 469–481. Springer, 2004.
- [11] A. Gunay and V.V. Nابیev, "Automatic age classification with lbp," in *Computer and Information Sciences, 2008. ISCIS '08. 23rd International Symposium on*, Oct 2008, pp. 1–4.
- [12] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971–987, Jul 2002.
- [13] S. Bekhouche, A. Ouafi, A. Taleb-Ahmed, and A. Hadid, "Facial age estimation using bsif and lbp," in *Proceeding of the first International Conference on Electrical Engineering ICEEB14*, in press.
- [14] DatTien Nguyen, SoRa Cho, and KangRyoung Park, "Human age estimation based on multi-level local binary pattern and regression method," in *Future Information Technology*, James J. (Jong Hyuk) Park, Yi Pan, Cheon-Shik Kim, and Yun Yang, Eds., vol. 309 of *Lecture Notes in Electrical Engineering*, pp. 433–438. Springer Berlin Heidelberg, 2014.
- [15] Jonathan Milgram, Mohamed Cheriet, Robert Sabourin, et al., "one against one or one against all: Which one is better for handwriting recognition with svms?," in *Tenth International Workshop on Frontiers in Handwriting Recognition*, 2006.
- [16] MATLAB 8.4 and Statistics Toolbox 9.1, The MathWorks Inc., Natick, Massachusetts, 2014, United States.