Automated Recognition of Facial Expressions Authenticity

Krystian Radlak
Institute of Automatic Control
Silesian University of Technology
Akademicka 2A, 44-100 Gliwice, Poland
krystian.radlak@polsl.pl

Bogdan Smolka
Institute of Automatic Control
Silesian University of Technology
Akademicka 2A, 44-100 Gliwice, Poland
bogdan.smolka@polsl.pl

ABSTRACT

Recognition of facial expressions authenticity is quite troublesome for humans. Therefore, it is an interesting topic for the computer vision community, as the developed algorithms for facial expressions authenticity estimation may be used as indicators of deception. This paper discusses the state-of-the art methods developed for smile veracity estimation and proposes a plan of development and validation of a novel approach to automated discrimination between genuine and posed facial expressions. The proposed fully automated technique is based on the extension of the high-dimensional Local Binary Patterns (LBP) to the spatio-temporal domain and combines them with the dynamics of facial landmarks movements. The proposed technique will be validated on several existing smile databases and a novel database created with the use of a high speed camera. Finally, the developed framework will be applied for the detection of deception in real life scenarios.

CCS Concepts

•**Computing methodologies** → **Activity recognition and understanding;** *Computer vision;* •**Applied computing** → Psychology;

Keywords

facial expressions recognition; facial expressions spontaneity; smile genuineness; deception detection; human-computer interaction

1. INTRODUCTION

Facial expressions play a major role in interpersonal interaction and reflect the human emotional state. Emotions play also a vital role in addition to verbal communication, and how well these emotions are expressed and understood is important to interpersonal relationships and individual well-being [14]. However, people are able to alter their facial expressions in order to fabricate or conceal their true emotional state. Therefore, the identification of falsified and unnatural expressions of emotions is important in everyday life, the courts, parole hearings and crime prevention [28].

The huge potential of non-invasive methods intended for the recognition of human emotions, based on the analysis of visual information, was quickly spotted by commercial companies providing services in the field of screening the employees and customers of large industrial corporations, banks and insurance companies, as well as by government agencies, which perceive intelligent monitoring systems as an effective tool to fight against crimes [35].

Despite the ubiquity of deception in our daily lives, the researcher shows that a typical observer is not better in deception detection than flipping a coin [34]. According to Bond and DePaulo [3], the average accuracy of trustworthiness assessment is equal to 54%. Several studies suggest that some people with special talent or law enforcement practitioners called "wizards", due to their extraordinary skills, can reach even 90% accuracy in deception detection [13, 12, 25]. However, despite the fact that the results presented in [4] show that this ability is very rare, even in the group of the well-trained officers, (only two female officers from the group of 112 participants of the law enforcement and 122 undergraduate students were over 80% accurate in deception assessment of truthful and deceptive videos), in general it is possible to estimate the trustfulness using only non-verbal cues.

Although a plethora of literature has been published in the recent years on deception detection, only a few papers are connected with automated deception detection using computer vision methods [21, 22, 2, 32, 38]. Our personal experience in this topic suggests that the main problem can be the availability and difficulty in developing a well-designed deception database [29]. For example, the publicly available deception database, developed by our research group and described in [29], has several limitations, such a as low emotional engagement of the students and many mistakes made in their answers during the experiments. This may not guarantee that the trained algorithms will really reflect the discrimination power in recognition of truth-tellers and deceivers. Therefore, in the first step of our research, we decided to focus on the recognition of facial expressions authenticity on well-established, validated and publicly available databases of spontaneous and deliberate facial expressions and then to apply the same approaches directly to our deception database.

The main aim of the proposed PhD dissertation is to develop novel computer vision algorithms that are able to de-

ICMI'16, November 12–16, 2016, Tokyo, Japan

http://dx.doi.org/10.1145/2993148.2997624

tect some hardly noticeable cues that are presented on human face in order to discriminate posed from genuine facial expressions. Our fully automated approach is based on the novel extension of the high-dimensional Local Binary Patters (LBP) [6] to the spatio-temporal domain and combines them with the features that are determined using the automatically estimated facial landmarks dynamics. Finally, the developed framework will be applied to the available deception databases.

2. RELATED WORK

The problem of the automated facial expressions recognition has been intensively studied in the literature [5, 19]. The first approach was published in 1978 by Suwa et al. in [33], but a significant interest in this field have been observed since the last decade of XX century.

The current research in the facial expressions analysis domain is focused on several problems: recognition of different basic types of facial expressions, automated facial action units detection, micro-expressions recognition, among many others [19]. Another interesting field connected with facial expressions analysis is the recognition of their authenticity, due to the fact that the same facial features that allow to distinguish spontaneous and posed facial behaviors may be used as an indicator of deception.

Most of the state-of-the-art algorithms that were designed for recognition of genuine and deliberate facial expressions are focused on the analysis of the temporal changes of facial display [?]. In [7], the authors conclude that genuine smile has smaller amplitude and the ratio of duration to amplitude is larger. Another approach is presented in [37], in which the discriminative Completed Local Binary Patterns (CLBP) are applied around five facial landmarks from three orthogonal planes and separately in three smile phases (rise, sustain and decay), which are detected using the smile detector from OpenCV. In another study [31], the authors show that the maximum speed of the smile onset and offset phases are higher in posed samples. In the recent work [9], the authors utilized the facial tracking of the eyelid, cheek, and lip corner landmarks and extracted 25 features, such as duration of the smile phases, the amplitude or the mean speed that incorporates the dynamics of the facial landmarks movements.

3. PROPOSED METHOD AND AUTHOR'S CONTRIBUTION

Previously discussed works suggest that studying dynamic properties of a smile allows to discriminate between falsified and sincere one. In 1862 Duchenne de Boulogne, a French neurologist, concluded that the deliberate smile did not involve the contraction of the muscles surrounding the eyes. This finding is also confirmed in [11], what may suggest that spatial information around the eyes is also a discriminative feature and fusion of the spatio-temporal features and facial landmarks dynamics should improve the accuracy of the final algorithm.

In general, the automated analysis of facial expressions can be divided into several phases. The first step is facial region detection and the facial landmarks localization. Then, the facial display is normalized and typically frontal view of the face is reconstructed, as the recent reports suggested that this process may improve the performance of face recognition systems [15]. Afterwards, the facial features are extracted and finally the facial expression is recognized.

In order to obtain the competing results, we evaluated experimentally on several databases the state-of-the-art algorithms at each stage in the face analysis pipeline. As the most promising, we chose the Histograms of Oriented Gradients (HOG), trained with structural Support Vector Machines (SVM) [17] and the HeadHunter technique [20] for face detection. Then, the facial landmarks are localized using an ensemble of regression trees [16] and the coarse-to-fine shape searching method proposed in [41]. Next, the frontal face is reconstructed using the algorithm proposed in [15] or the face image is normalized applying an affine transformation. The facial features are extracted using the LBP [23] and its modifications, such as the uniform LBP [24] and the high-dimensional LBP [6]. In order to reduce the dimensionality of the feature vector, we used the Random Frog feature selection algorithm [18] and PCA feature extraction. Finally, the facial expressions are recognized using SVM classifier with optimized parameters. The proposed pipeline was applied in our recent work connected with the automated facial expressions recognition in the wild [30].

We also developed a very robust smile detector that was trained on the very challenging GENKI-4K database with a large variability in pose, illumination and imaging conditions. The application of the SVM with the combination of the state-of-the-art algorithms of face detection, facial landmarks localization and feature selection yields highly competitive results and outperforms the latest results that were obtained on the GENKI-4K dataset, including a well-designed deep convolutional networks (CNN) [39], (these results have been unpublished yet). We also confirmed the results presented in [36], that the distance from the hyperplane for the SVM classifier trained on binary task of the smiling and neutral facial display corresponds to the smile intensity. It allows us to find the intervals in video recordings, in which the smile is presented.

The proposed pipeline will be extended to the spatio-temporal domain similarly to the LBP-TOP technique [40] and combined with the features calculated using facial landmarks dynamics in different facial expressions phases (onset, apex, offset).

Our approach towards facial expressions authenticity will be evaluated on several smile databases as UvA-NEMO smile database [10], MAHNOB-Laughter database [26] and BBC smile database [1] and spontaneous vs. posed Facial Expression Database (SPOS) database [27]. The proposed techniques will be also applied to the recently prepared smile database gathered using a monochrome high speed camera working at 350 fps at resolution 2320×1726. Finally, we plan to verify the usefulness of the proposed algorithms on the *Silesian Deception Detection Database* elaborated by our research group [29] and on the recently published MAFIA deception detection database presented in [8].

Currently, we are working on the preparation of the annotations for the prepared smile database and implement the proposed extension of the high-dimensional LBP and the facial landmarks movements dynamics. Examples of the change of smile intensity in time for genuine and posed smile that was obtained on the UvA-NEMO database are presented in Fig. 1. In our approach, the smile intensity is defined as the mean amplitude of right and left lip corners, normalized by the length of the lip. These results were obtained using the face frontalization algorithm [15] and the facial landmarks were localized using an ensemble regression trees algorithm presented in [16] .

4. CONCLUSION

The main aim of the proposed research is to increase the understanding of the fundamental principles behind the link between the emotional arousal and the occurrence of subtle facial movements, in order to verify the facial expressions authenticity displayed by the observed subject. Such studies have been conducted only in limited range and we hope that the research increases an interest of the computer vision community in this topic.

The main contribution of our work consists in: (i) extension of the concept of high-dimensional Local Binary Patterns around facial landmarks [6] to the three orthogonal planes, so as to use the spatio-temporal properties of the analyzed video sequence similarly to the well-established LBP-TOP technique [40], (ii) incorporating facial landmarks dynamics of eyes, eyebrows and mouth localized using the state-of-the-art face alignment algorithms, (iii) fusion of these features with those that are based on the facial landmarks dynamics, which are estimated using the state-of-the-arts face analysis algorithms, (iv) elaboration of a novel publicly available deception database and (v) preparation of a database that contains the spontaneous and posed video recordings of smile gathered using a high speed camera working at 350 fps.

We believe that the performed research will shows that it is possible to develop a computer system that is able to perceive some hardly noticeable cues that are presented on face, in order to discriminate posed from genuine facial expressions. We do hope that the created system could be used to test the truthfulness of persons and would enable the detection of suspicious or anomalous human behavior.

5. ACKNOWLEDGMENTS

This material is based on work supported by the Polish National Science Center (NCN) under the Grant: DEC-2012/07/B/ST6/ 0122. Additionally, Krystian Radlak was supported by statutory funds for young researchers (BKM-506/RAU1/2016) of the Institute of Automatic Control, Silesian University of Technology, Poland.

6. REFERENCES

[1] BBC. Spot The Fake Smile. http://www.bbc.co.uk/science/humanbody//mind/surveys/smiles/. Online; accessed 20 September 2016.

[2] N. Bhaskaran, I. Nwogu, M. Frank, and V. Govindaraju. Lie to me: Deceit detection via online behavioral learning. In *IEEE International Conference on Automatic Face Gesture Recognition and Workshops*, pages 24–29, 2011.

[3] C. F. Bond and B. M. DePaulo. Accuracy of deception judgments. *Personality and Social Psychology Review*, 10(3):214–234, 2006.

[4] G. D. Bond. Deception detection expertise. *Law and Human Behavior*, 32(4):339–351, 2008.

[5] A. J. Calder and A. W. Young. Understanding the recognition of facial identity and facial expression? *Nature Reviews Neuroscience*, 6:641–651, 2005.

[6] D. Chen, X. Cao, F. Wen, and J. Sun. Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3025–3032, 2013.

[7] J. F. Cohn and K. L. Schmidt. The timing of facial motion in posed and spontaneous smiles. *International Journal of Wavelets, Multiresolution and Information Processing*, 02(02):121–132, 2004.

[8] S. Demyanov, J. Bailey, K. Ramamohanarao, and C. Leckie. Detection of deception in the mafia party game. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, pages 335–342, 2015.

[9] H. Dibeklioglu, A. Salah, and T. Gevers. Recognition of genuine smiles. *IEEE Transactions on Multimedia*, 17(3):279–294, 2015.

[10] H. Dibeklioğlu, A. A. Salah, and T. Gevers. *Are You Really Smiling at Me? Spontaneous versus Posed Enjoyment Smiles*, pages 525–538. 2012.

[11] P. Ekman, R. J. Davidson, and W. V. Friesen. The duchenne smile: Emotional expression and brain physiology: Ii. *Journal of Personality and Social Psychology*, 58(2):342–353, 1990.

[12] P. Ekman, M. O'Sullivan, and M. G. Frank. A few can catch a liar. *Psychological Science*, 10(3):263–266, 1999.

[13] M. G. Frank and P. Ekman. The ability to detect deceit generalizes across different types of high-stake lies. *Journal of Personality and Social Psychology*, 72(6):1429–1439, 1997.

[14] S. Fussell. *The Verbal Communication of Emotions: Interdisciplinary Perspectives*. Taylor & Francis, 2002.

[15] T. Hassner, S. Harel, E. Paz, and R. Enbar. Effective face frontalization in unconstrained images. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4295–4304, 2015.

[16] V. Kazemi and J. Sullivan. One millisecond face alignment with an ensemble of regression trees. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1867–1874, 2014.

[17] D. E. King. Max-margin object detection. *CoRR*, abs/1502.00046, 2015.

[18] H.-D. Li, Q.-S. Xu, and Y.-Z. Liang. Random frog: An efficient reversible jump Markov chain Monte Carlo-like approach for variable selection with applications to gene selection and disease classification. *Analytica Chimica Acta*, 740:20 – 26, 2012.

[19] B. Martinez and M. F. Valstar. *Advances, Challenges, and Opportunities in Automatic Facial Expression Recognition*, pages 63–100. 2016.

[20] M. Mathias, R. Benenson, M. Pedersoli, and L. Gool. *13th European Conference on Computer Vision (ECCV)*, chapter Face Detection without Bells and Whistles, pages 720–735. 2014.

[21] T. Meservy, M. Jensen, J. Kruse, J. Burgoon, and J. Nunamaker. Automatic extraction of deceptive behavioral cues from video. In P. Kantor, G. Muresan, F. Roberts, D. Zeng, F.-Y. Wang, H. Chen, and R. Merkle, editors, *Intelligence and Security Informatics*, volume 3495 of *Lecture Notes in*

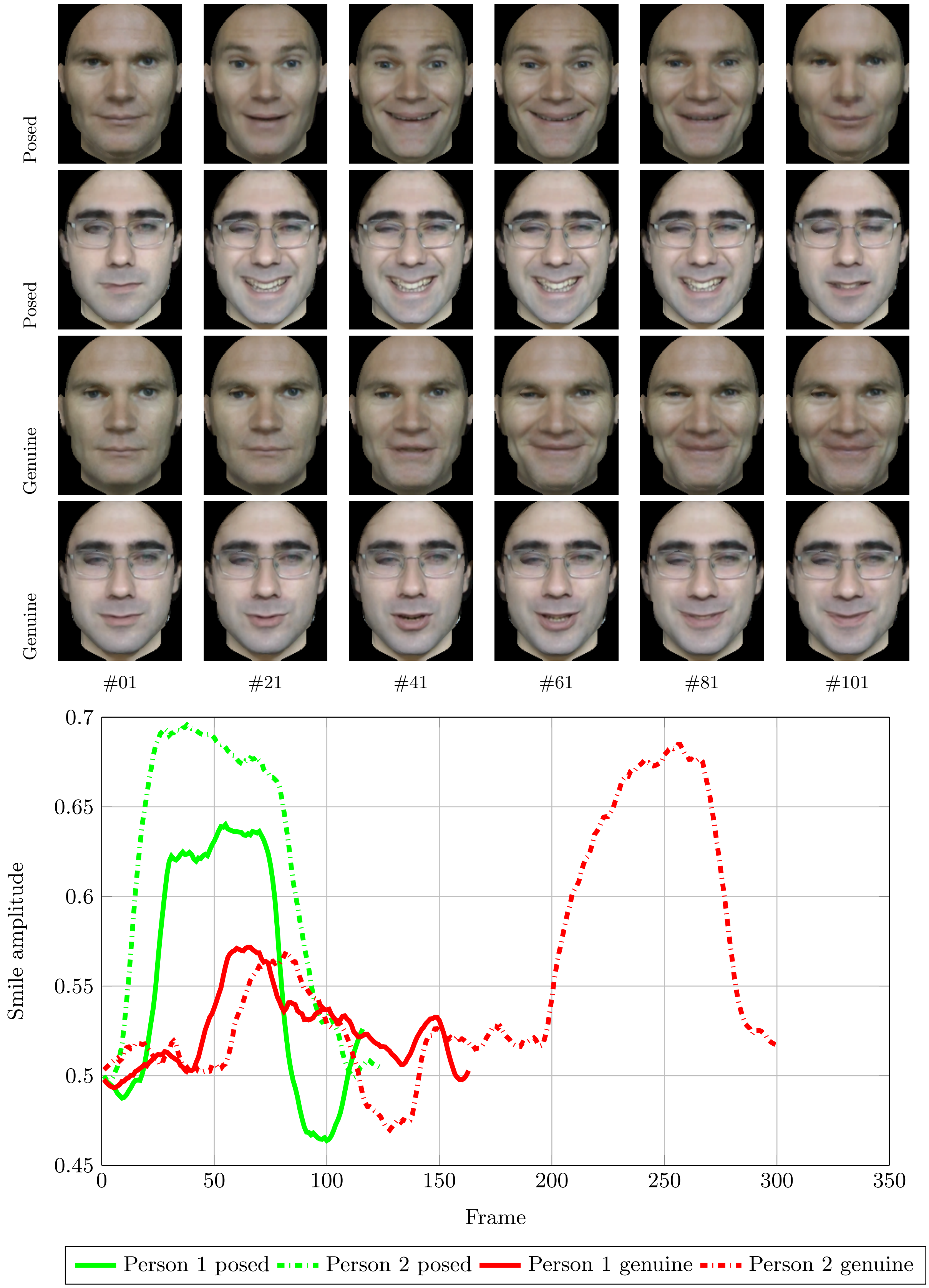

Figure 1: Smile intensity of genuine and posed smiles for the first two persons in UVA-NEMO database. Selected frames after face frontalization are shown above the graphs.

Computer Science, pages 198–208. Springer Berlin Heidelberg, 2005.

[22] T. Meservy, M. Jensen, W. Kruse, J. Burgoon, and J. Nunamaker, JayF. Automatic extraction of deceptive behavioral cues from video. In *Terrorism Informatics*, volume 18 of *Integrated Series In Information Systems*, pages 495–516. Springer US, 2008.

[23] T. Ojala, M. Pietikainen, and D. Harwood. Performance evaluation of texture measures with classification based on kullback discrimination of distributions. In *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, volume 1, pages 582–585, 1994.

[24] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

[25] M. O'Sullivan and P. Ekman. The wizards of deception detection. In *The Detection of Deception in Forensic Contexts*, pages 269–286. 2004.

[26] S. Petridis, B. Martinez, and M. Pantic. The {MAHNOB} laughter database. *Image and Vision Computing*, 31(2):186–202, 2013.

[27] T. Pfister, X. Li, G. Zhao, and M. PietikÃd'inen. Differentiating spontaneous from posed facial expressions within a generic facial expression recognition framework. In *IEEE International Conference on Computer Vision Workshops*, pages 868–875, 2011.

[28] S. Porter, N. Korva, and A. Baker. Secrets of the human face: New insights into the face and covert emotions. *Psychology Aotearoa*, 3(2):94–99, 2011.

[29] K. Radlak, M. Bozek, and B. Smolka. Silesian deception database: Presentation and analysis. In *Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection*, pages 29–35, 2015.

[30] K. Radlak and B. Smolka. High dimensional local binary patterns for facial expression recognition in the wild. In *2016 18th Mediterranean Electrotechnical Conference (MELECON)*, pages 1–5, 2016.

[31] K. L. Schmidt, Z. Ambadar, J. F. Cohn, and L. I. Reed. Movement differences between deliberate and spontaneous facial expressions: Zygomaticus major action in smiling. *Journal of Nonverbal Behavior*, 30(1):37–52, 2006.

[32] L. Su and M. Levine. High-stakes deception detection based on facial expressions. In *22nd International Conference on Pattern Recognition*, pages 2519–2524, 2014.

[33] M. Suwa, N. Sugie, and K. Fujimora. A preliminary note on pattern recognition of human emotional expression. In *International Joint Conference on Pattern Recognition*, pages 408–410, 1978.

[34] A. Vrij, P. A. Granhag, and S. Porter. Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest*, 11(3):89–121, 2010.

[35] S. Weinberger. Airport security: Intent to deceive? *Nature*, 465:412–415, 2010.

[36] J. Whitehill, G. Littlewort, I. Fasel, M. Bartlett, and J. Movellan. Toward practical smile detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(11):2106–2111, 2009.

[37] P. Wu, H. Liu, and X. Zhang. Spontaneous versus posed smile recognition using discriminative local spatial-temporal descriptors. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1240–1244, 2014.

[38] X. Yu, S. Zhang, Z. Yan, F. Yang, J. Huang, N. Dunbar, M. Jensen, J. Burgoon, and D. Metaxas. Is interactional dissynchrony a clue to deception? insights from automated analysis of nonverbal visual cues. *IEEE Transactions on Cybernetics*, 45(3):506–520, 2015.

[39] K. Zhang, Y. Huang, H. Wu, and L. Wang. Facial smile detection based on deep learning features. In *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, pages 534–538, 2015.

[40] G. Zhao and M. Pietikainen. Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):915–928, 2007.

[41] S. Zhu, C. Li, C. C. Loy, and X. Tang. Face alignment by coarse-to-fine shape searching. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4998–5006, 2015.