2016 23rd International Conference on Pattern Recognition (ICPR)
Cancún Center, Cancún, México, December 4-8, 2016

Multimodal Fusion of Audio, Scene, and Face Features for First Impression Estimation

Furkan Gürpınar
Program of Computational Science and Engineering
Boğaziçi University
Bebek, Istanbul, Turkey
Email: furkan.gurpinar@boun.edu.tr

Heysem Kaya
Department of Computer Engineering
Namık Kemal University
Çorlu, Tekirdağ, Turkey
Email: hkaya@nku.edu.tr

Albert Ali Salah
Department of Computer Engineering
Boğaziçi University
Bebek, Istanbul, Turkey
Email: salah@boun.edu.tr

***Abstract*—Affective computing, particularly emotion and personality trait recognition, is of increasing interest in many research disciplines. The interplay of emotion and personality shows itself in the first impression left on other people. Moreover, the ambient information, e.g. the environment and objects surrounding the subject, also affect these impressions. In this work, we employ pre-trained Deep Convolutional Neural Networks to extract facial emotion and ambient information from images for predicting apparent personality. We also investigate Local Gabor Binary Patterns from Three Orthogonal Planes video descriptor and acoustic features extracted via the popularly used openSMILE tool. We subsequently propose classifying features using a Kernel Extreme Learning Machine and fusing their predictions. The proposed system is applied to the ChaLearn Challenge on First Impression Recognition, achieving the winning test set accuracy of 0.913, averaged over the "Big Five" personality traits.**

I. Introduction and Related Work

Automatic prediction of apparent personality is an interesting and challenging topic for researchers from a range of backgrounds. Machines that are able to recognize apparent personality traits can be useful in many applications such as computer assisted tutoring systems, forensics and user recommendation systems. The complexity of the personality formation makes it also hard to automatically recognize [1], [2]. A way to handle this issue is working on the *impressions* (apparent personality) instead of the personality itself [3].

In this work, we tackle the problem of predicting the apparent personality using the data and protocol from the ChaLearn Looking at People 2016 First Impression Challenge [3]. Our aim is to benefit from influences of emotional facial expressions, as well as ambient cues on first impressions.

The apparent personality is assessed along the "Big Five" personality traits that are Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (OCEAN). The formation of a personality impression affects decisions in humans, and it is an interesting question whether computers can automatically estimate how a certain person is perceived by others. There are several recent approaches for recognizing apparent personality traits from different modalities such as audio [4], [5], text [6], [7], [8] and visual information [9], [10]. To increase the robustness of predictions, multimodal systems are also investigated [11], [12], [13], [14], [15].

In our previous work, we have shown that for first impression prediction, deep learning approaches for face processing can be fused profitably with features that describe the scene, i.e. the context of the perceived image [15]. For classification, Support Vector Machines (SVM) approaches are widely used [5], [12], [14], but we have used Extreme Learning Machines (ELM) [16], which achieved good results with rapid classification of new samples [15]. Similar approaches have been used on related tasks like facial age estimation [17] and emotion recognition [18], with good results.

All three winners of the first round of this challenge extensively used deep learning in their bimodal systems, while the overall approach and the type of the network was different [19], [20], [21]. In [19] single hidden layer neural network (NN) was used for audio modality regression, while deep convolutional neural networks (DCNN) were used for video representation and regression. The choice of the first runner up [20] was a Recurrent DCNN for both modalities. The learning system of the second runner up [21] was based on Residual Networks. The winner [19] and the second runner up [21] did not employ face alignment in the preprocessing step, but both of these works applied late fusion of modality based scores. For facial feature extraction, the winning system of Zhang et al. [19] used the VGG-Face pre-trained DCNN model [22], which we also employ here, and in our submission to the first round of the challenge [15].

Given the success of deep learning and the speed of ELM, we propose to fuse ELM models trained on audio, deep face, and scene features. Our contribution to the first round of this challenge proposed combining emotion related and ambient features that are efficiently extracted from pre-trained/fine-tuned DCNN models [15]. Here, we further improve this system by investigating i) other visual descriptors; ii) the audio modality; and iii) weighted score level fusion strategy. Our method is illustrated in Figure 1.

The remainder of this paper is organized as follows. In the next section we provide background and details on the methodology. Then in Section III, we present the experimental results. Finally, Section IV concludes the paper with remarks on the proposed approach in context.

978-1-5090-4847-2/16/$31.00 ©2016 IEEE

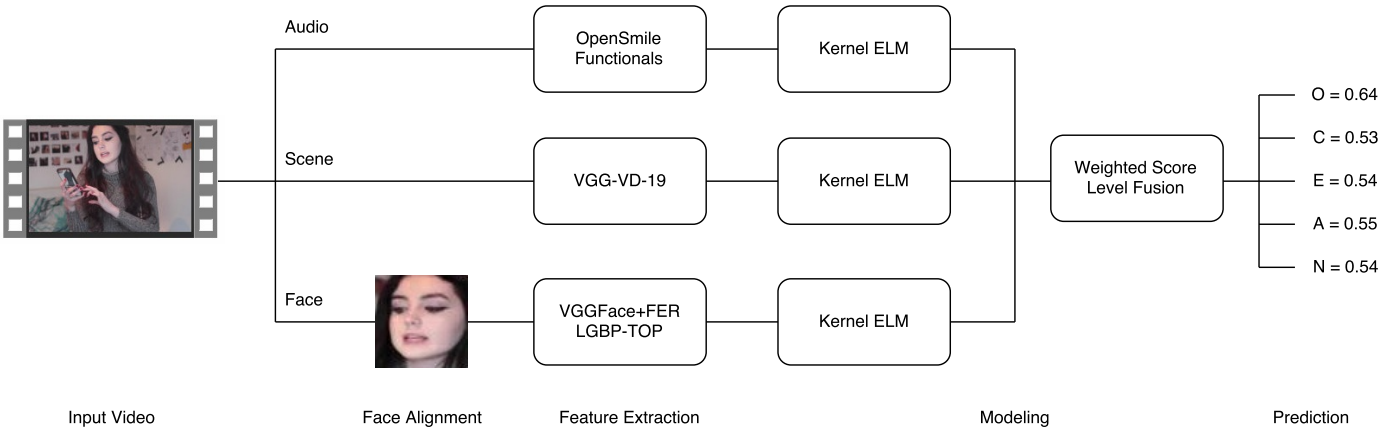


Fig. 1. Flowchart of the proposed method.

II. Methodology

Our proposed approach evaluates a short video clip that contains a single person, and outputs an estimate of apparent personality traits in the five dimensions mentioned earlier. Both visual and audio features are used in the proposed approach. In this section, we describe the main steps of our pipeline, namely face alignment, feature extraction, and modeling. The described approach is very similar to [15], but includes acoustic features in addition to face and scene features. We provide a comparison of the proposed approach with our previous work in Section III.

A. Face Alignment

For face alignment, Xiong and de la Torre's Supervised Descent Method (SDM) is used [23]. 49 landmarks are located on the face. The roll angle is estimated from the eye corners to rotate the image accordingly. Then a margin of 20% interocular distance around the outer landmarks is added to crop the facial images and each image is resized to 64×64 pixels.

B. Visual Feature Extraction

Facial features are extracted over an entire video segment and summarized by functionals. Scene features, however, are exracted from the first image of each video only. The assumption is that videos do not stretch over multiple shots.

1) Face Features: After aligning the faces, image-level deep features are extracted from a network trained for facial emotion recognition, as explained in Section II-C. This network has a 37-layer architecture. The response of the 33^{rd} layer is used in this work, which is the lowest-level 4096-dimensional descriptor.

We compare deep features with traditional appearance descriptors and geometric information that is shown to be effective in emotion recognition [18]. A detailed evaluation of each approach was given in [15].

2) Video Features: After extracting frame-level features from each registered face, we summarize the videos by computing functional statistics of each dimension over time. The functionals include mean, standard deviation, offset, slope, and curvature. Offset and slope are calculated from the first order polynomial fit to each feature contour, while curvature is the leading coefficient of the second order polynomial. An empirical comparison of the individual functionals is given in [15].

3) Scene Features: In order to use ambient information in the images to our advantage, we extract features using the VGG-VD-19 network [24], which is trained for an object recognition task on the ILSVRC 2012 dataset. Similar to face features, we use the 4096-dimensional feature from the 39^{th} layer of the 43-layer architecture, hence we obtain a description of the overall image that contains both face and scene.

C. CNN Finetuning

We start with the VGG-Face network [22], changing the final layer (originally a 2622-dimensional recognition layer), to a 7-dimensional emotion recognition layer, where the weights are initialized randomly. We finetune this network with the softmax loss function using more than 30K training images in the FER-2013 dataset [25]. We choose an initial learning rate of 0.0001, a momentum of 0.9 and a batch size of 64. We train the model only for 5 epochs.

D. Acoustic Feature Extraction

The open-source openSMILE tool [26] is popularly used to extract acoustic features in a number of international paralinguistic and multi-modal challenges. The idea is to obtain a large pool of potentially relevant features by passing an extensive set of summarizing functionals on the low level descriptor contours (e. g. Mel Frequency Cepstral Coefficients, pitch, energy and formants). We use the toolbox with standard feature configurations that served as the challenge baseline sets in INTERSPEECH 2009, 2010, 2012 and 2013 Computational Paralinguistics Challenges [27], [28], [29], [30], respectively.

E. Regression with Kernel ELM

In order to model personality traits from visual features, we used kernel extreme learning machines (ELM), due to the learning speed and accuracy of the algorithm. In the following paragraphs, we briefly explain the learning strategy of ELM.

Initially, ELM is proposed as a fast learning method for Single Hidden Layer Feedforward Networks (SLFN): an alternative to back-propagation [31]. To increase the robustness and the generalization capability of ELM, a regularization coefficient $\mathbf{C}$ is included in the optimization procedure. Therefore, given a kernel $\mathbf{K}$ and the label vector $\mathbf{T} \in \mathbb{R}^{N\times 1}$ where N denotes the number of instances, the projection vector β is learned as follows:

$$\beta = (\frac{\mathbf{I}}{C} + \mathbf{K})^{-1}\mathbf{T}. \quad (1)$$

In order to prevent parameter over-fitting, we use the linear kernel $\mathbf{K}(x, y) = x^T y$, where x and y are the original feature vectors after min-max normalization of each dimension among the training samples. With this approach, the only parameter of our model is the regularization coefficient C, which we optimize with a 5-fold subject independent cross-validation on the training set.

III. Experiments

The "ChaLearn LAP Apparent Personality Analysis: First Impressions" challenge consists of 10,000 clips collected from 5,563 YouTube videos, where the poses are more or less frontal, but the resolution, lighting and background conditions are not controlled, hence providing a dataset with in-the-wild conditions. Each clip in the training set is labeled for the Big Five personality traits.

For brevity, we skip corpus related information here and refer the reader to in [3] for details on the challenge. The performance score in this challenge is the Mean Absolute Error subtracted from 1, which is formulated as follows:

$$1 - \sum_{i}^{N} \frac{|\hat{y}_i - y_i|}{N}, \quad (2)$$

where N is the number of samples, $\hat{y}$ is the predicted label and y is the true label ($0 \leq y \leq 1$). This score is then averaged over five tasks. This means the final score varies between 0 (worst case) and 1 (best case).

A. Experimental Results

Since the video modality features were primarily investigated in our former work [15], in this study we first analyze the performance of acoustic features obtained from the open source openSMILE tool [26]. We extracted supra-segmantal acoustic features that represent the whole utterance using four standard feature sets, whose dimensionalities ranged from 384 (IS09) to 6 373 (IS13). The results from 5-fold cross-validation on the training set are shown in Table I.

TABLE I
Performance of standard acoustic feature sets using 5-fold cross-validation on the training set

Features	Dim	Mean	Extr.	Agre.	Cons.	Neur.	Open.
IS09 [27]	384	0.894	0.893	0.898	0.886	0.892	0.898
IS10 [28]	1582	0.895	0.895	0.899	0.889	0.895	0.899
IS12 [29]	6125	0.895	0.895	0.900	0.889	0.895	0.899
IS13 [30]	6373	0.896	0.894	0.900	0.890	0.895	0.899

Acoustic feature sets provide a similar overall performance, with around 0.105 MAE for each trait. As stated earlier, the performance is reported as 1-MAE. In the subsequent experiments and on our validation/test set submissions, we use the IS13 baseline set that gives the highest accuracy on the training set.

In Table II, we report the validation set performances of individual features, as well as their feature-, score- and multi-level fusion alternatives. Here, System 5 corresponds to our best submission in the first round of this challenge [15]. Score fusion weights are searched in a pool of randomly generated fusion vectors, each of which sum up to 1. The selective fusion simply selects the best performing system for each of the five categories. We show the cross-validation estimations of all the modalities as well as the multimodal fusion in Figure 2. In general, fusion scores are observed to benefit from complementary information of individual sub-systems. On the other hand, although images are not always frontal, the predictions of the model learned from the deep facial features are individually closest to the ground truth.

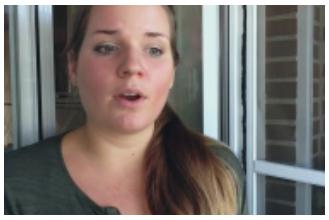
The best fusion system (System 8 in Table II) is obtained by augmenting the Agreeableness scores of System 6 with predictions of System 7 for the remaining four personality traits. This gives a test set mean accuracy of 0.913, which ranks the first in the ICPR 2016 ChaLearn LAP First Impression contest. Considering the obtained test set performance in comparison to other top competitors' accuracies (see Table III), we observe that the performances are around 0.91.

The test set results of top ranking teams are both high and competitive. When individual personality dimensions are analyzed, we see that our system ranks the first in all but one dimension (Agreeableness), exhibiting the highest performance in Extraversion. We also observe that the proposed system's validation and test accuracies are very similar: in four dimensions the absolute difference is smaller than or equal to 0.1% and in Agreeableness it is 0.7%. Therefore, we can conclude that the generalization ability of the proposed system is high.

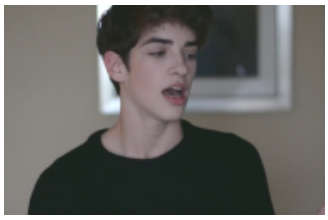
IV. Conclusions

In this paper, we improve on our recent work [15], which proposed to fuse deep convolutional neural networks (DCNN) that are originally trained for other tasks such as face, object, and emotion recognition. In addition to the features extracted from pre-trained/fine-tuned DCNNs, in this work we contrast and combine other visual descriptors and the audio modality, using multi-level, feature and score fusion. The best results are achieved with multi-level (particularly weighted score level) fusion of one acoustic and three visual sub-systems.

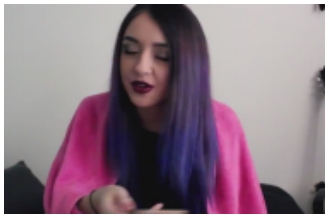
We observe that the addition of the audio modality does not bring about a significant improvement, and there are diminishing returns. The first round of the challenge reported ten systems with accuracies varying from 0.876 to 0.913. We observe that the top systems, including the system described in this paper, nonetheless make use of the audio information.



	Face	Scene	Audio	Fusion	True
O	0.53	0.63	0.54	0.55	0.54
C	0.45	0.69	0.45	0.49	0.49
E	0.41	0.53	0.47	0.43	0.46
A	0.56	0.68	0.51	0.58	0.59
N	0.44	0.61	0.49	0.47	0.48

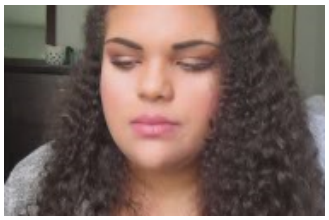


	Face	Scene	Audio	Fusion	True
O	0.53	0.45	0.56	0.51	0.49
C	0.48	0.44	0.54	0.48	0.46
E	0.45	0.37	0.49	0.44	0.47
A	0.54	0.49	0.56	0.53	0.53
N	0.48	0.44	0.52	0.48	0.48



	Face	Scene	Audio	Fusion	True
O	0.70	0.58	0.68	0.68	0.67
C	0.64	0.49	0.65	0.62	0.62
E	0.62	0.47	0.59	0.60	0.62
A	0.59	0.49	0.62	0.57	0.57
N	0.62	0.51	0.62	0.60	0.64

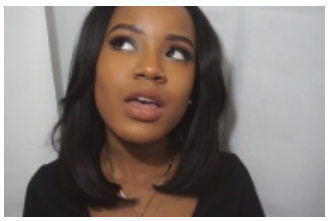
	Face	Scene	Audio	Fusion	True
O	0.62	0.62	0.64	0.62	0.63
C	0.59	0.52	0.55	0.58	0.58
E	0.56	0.54	0.50	0.56	0.51
A	0.61	0.61	0.58	0.61	0.60
N	0.57	0.57	0.60	0.57	0.56



	Face	Scene	Audio	Fusion	True
O	0.55	0.57	0.61	0.56	0.56
C	0.53	0.48	0.43	0.52	0.51
E	0.46	0.50	0.51	0.47	0.50
A	0.55	0.57	0.53	0.55	0.58
N	0.49	0.57	0.53	0.50	0.51

	Face	Scene	Audio	Fusion	True
O	0.56	0.49	0.58	0.55	0.56
C	0.51	0.45	0.50	0.50	0.46
E	0.49	0.41	0.49	0.47	0.48
A	0.57	0.57	0.54	0.57	0.56
N	0.54	0.51	0.52	0.54	0.51

	Face	Scene	Audio	Fusion	True
O	0.42	0.56	0.46	0.45	0.46
C	0.49	0.55	0.37	0.50	0.51
E	0.33	0.44	0.36	0.34	0.36
A	0.43	0.54	0.45	0.45	0.40
N	0.36	0.49	0.39	0.38	0.36



	Face	Scene	Audio	Fusion	True
O	0.59	0.60	0.62	0.59	0.60
C	0.49	0.56	0.62	0.50	0.51
E	0.39	0.52	0.53	0.41	0.44
A	0.51	0.56	0.58	0.51	0.55
N	0.48	0.52	0.58	0.49	0.51

Fig. 2. Eight examples from the training set where our approach produced good estimations for the traits.

TABLE II
REGRESSION PERFORMANCE OF VARIOUS SYSTEMS ON THE VALIDATION SET. FF: FEATURE-LEVEL FUSION, WF: WEIGHTED SCORE-LEVEL FUSION, SF: SELECTIVE FUSION

System	Feature	Dim	Mean	Extr.	Agre.	Cons.	Neur.	Open.
1	LGBP-TOP	100224	0.912	0.915	0.913	0.910	0.910	0.911
2	VGGFace+FER	20480	0.910	0.915	0.911	0.906	0.907	0.910
3	VGG-VD-19	4096	0.899	0.895	0.906	0.899	0.893	0.900
4	Audio-IS13	6373	0.899	0.898	0.907	0.892	0.898	0.902
5	FF(2,3)	24576	0.911	0.914	0.913	0.913	0.906	0.910
6	WF(1:4)	-	**0.914**	0.918	**0.914**	0.912	**0.912**	0.913
7	WF(FF(1,2),3)	-	**0.914**	**0.919**	0.913	**0.914**	**0.912**	**0.914**
8	SF(6,7)	-	**0.915**	**0.919**	**0.914**	**0.914**	**0.912**	**0.914**

TABLE III
FINAL RANKINGS ON THE SEQUESTERED TEST SET

Rank	Team	Mean	Extr.	Agre.	Cons.	Neur.	Open.
1	BU-NKU (ours)	**0.913**	**0.918**	0.907	**0.915**	**0.911**	**0.914**
2	evolgen	0.912	0.916	**0.911**	0.914	0.910	0.911
3	pandora	0.903	0.904	0.905	0.901	0.900	0.904
4	PILAB	0.898	0.895	0.904	0.896	0.894	0.901

REFERENCES

[1] H. Kaya and A. A. Salah, "Continuous mapping of personality traits: A novel challenge and failure conditions," in *Proceedings of the 2014 ICMI Workshop on Mapping Personality Traits Challenge*. ACM, 2014, pp. 17–24.

[2] O. Celiktutan and H. Gunes, "Automatic prediction of impressions in time and across varying context: Personality, attractiveness and likeability," *IEEE Transactions on Affective Computing*, 2016.

[3] V. P. Lopez, B. Chen, A. Clapes, M. Oliu, C. Corneanu, X. Baro, H. J. Escalante, I. Guyon, and S. Escalera, "Chalearn lap 2016: First round challenge on first impressions - dataset and results," in *ChaLearn Looking at People Workshop on Apparent Personality Analysis, ECCV Workshop Proceedings*, 2016.

[4] F. Valente, S. Kim, and P. Motlicek, "Annotation and recognition of personality traits in spoken conversations from the ami meetings corpus." in *INTERSPEECH*, 2012, pp. 1183–1186.

[5] N. Madzlan, J. Han, F. Bonin, and N. Campbell, "Towards automatic recognition of attitudes: Prosodic analysis of video blogs," *Speech Prosody, Dublin, Ireland*, pp. 91–94, 2014.

[6] F. Alam, E. A. Stepanov, and G. Riccardi, "Personality traits recognition on social network-facebook," *WCPR (ICWSM-13), Cambridge, MA, USA*, 2013.

[7] S. Nowson and A. J. Gill, "Look! who's talking?: Projection of extraversion across different social contexts," in *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition*. ACM, 2014, pp. 23–26.

[8] S. Gievska and K. Koroveshovski, "The impact of affective verbal content on predicting personality impressions in youtube videos," in *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition*. ACM, 2014, pp. 19–22.

[9] T. Fernando *et al.*, "Persons personality traits recognition using machine learning algorithms and image processing techniques," *Advances in Computer Science: an International Journal*, vol. 5, no. 1, pp. 40–44, 2016.

[10] R. Qin, W. Gao, H. Xu, and Z. Hu, "Modern physiognomy: An investigation on predicting personality traits and intelligence from the human face," *arXiv preprint arXiv:1604.07499*, 2016.

[11] C. Sarkar, S. Bhatia, A. Agarwal, and J. Li, "Feature analysis for computational personality recognition using youtube personality data set," in *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition*. ACM, 2014, pp. 11–14.

[12] F. Alam and G. Riccardi, "Predicting personality traits using multimodal information," in *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition*. ACM, 2014, pp. 15–18.

[13] G. Farnadi, S. Sushmita, G. Sitaraman, N. Ton, M. De Cock, and S. Davalos, "A multivariate regression approach to personality impression recognition of vloggers," in *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition*. ACM, 2014, pp. 1–6.

[14] M. Sidorov, S. Ultes, and A. Schmitt, "Automatic recognition of personality traits: A multimodal approach," in *Proceedings of the 2014 Workshop on Mapping Personality Traits Challenge and Workshop*. ACM, 2014, pp. 11–15.

[15] F. Gürpınar, H. Kaya, and A. A. Salah, "Combining deep facial and ambient features for first impression estimation," in *ChaLearn Looking at People Workshop on Apparent Personality Analysis, ECCV Workshop Proceedings*, 2016.

[16] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.

[17] F. Gürpınar, H. Kaya, H. Dibeklioğlu, and A. A. Salah, "Kernel ELM and CNN Based Facial Age Estimation," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, Las Vegas, Nevada, USA, June 2016, pp. 80–86.

[18] H. Kaya, F. Gürpınar, S. Afshar, and A. A. Salah, "Contrasting and combining least squares based learners for emotion recognition in the wild," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM, 2015, pp. 459–466.

[19] C.-L. Zhang, H. Zhang, X.-S. Wei, and J. Wu, "Deep bimodal regression for apparent personality analysis," in *ChaLearn Looking at People Workshop on Apparent Personality Analysis, ECCV Workshop Proceedings*, 2016.

[20] A. Subramaniam, V. Patel, A. Mishra, P. Balasubramanian, and A. Mittal, "Bi-modal first impressions recognition using temporally ordered deep audio and stochastic visual features," in *ChaLearn Looking at People Workshop on Apparent Personality Analysis, ECCV Workshop Proceedings*, 2016.

[21] Y. Güçlütürk, U. Güçlü, M. van Gerven, and R. van Lier, "Deep impression: Audiovisual deep residual networks for multimodal apparent personality trait recognition," in *ChaLearn Looking at People Workshop on Apparent Personality Analysis, ECCV Workshop Proceedings*, 2016.

[22] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *British Machine Vision Conference*, 2015.

[23] X. Xiong and F. De la Torre, "Supervised Descent Method and Its Application to Face Alignment," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 532–539.

[24] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[25] I. J. Goodfellow, D. Erhan, P. L. Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee *et al.*, "Challenges in representation learning: A report on three machine learning contests," in *International Conference on Neural Information Processing*. Springer, 2013, pp. 117–124.

[26] F. Eyben, M. Wöllmer, and B. Schuller, "OpenSMILE: the Munich versatile and fast open-source audio feature extractor," in *ACM International Conference on Multimedia*, 2010, pp. 1459–1462.

[27] B. Schuller, S. Steidl, and A. Batliner, "The INTERSPEECH 2009 Emotion Challenge," in *INTERSPEECH*, Brighton, UK, September 2009, pp. 312–315.

[28] B. Schuller, S. Steidl, A. Batliner, F. Burkhardt, L. Devillers, C. A. Müller, and S. S. Narayanan, "The interspeech 2010 paralinguistic challenge." in *Proc. INTERSPEECH*, Makuhari, Japan, 2010, pp. 2794–2797.

[29] B. Schuller, S. Steidl, A. Batliner, E. Nöth, A. Vinciarelli, F. Burkhardt, R. Van Son, F. Weninger, F. Eyben, T. Bocklet *et al.*, "The INTERSPEECH 2012 speaker trait challenge," in *INTERSPEECH*, 2012, pp. 254–257.

[30] B. Schuller, S. Steidl, A. Batliner, A. Vinciarelli, K. Scherer, F. Ringeval, M. Chetouani, F. Weninger, F. Eyben, E. Marchi, M. Mortillaro, H. Salamin, A. Polychroniou, F. Valente, and S. Kim, "The INTERSPEECH 2013 Computational Paralinguistics Challenge: Social Signals, Conflict, Emotion, Autism," in *INTERSPEECH*, Lyon, France, 2013, pp. 148–152.

[31] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme Learning Machine: a new learning scheme of feedforward neural networks," in *IEEE International Joint Conference on Neural Networks*, vol. 2, 2004, pp. 985–990.