Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 13, Issue 1, January 2022: 449-459

Facial and posture features assisted personality traits recognition from videos

[1]Sumaiya Samreen, [2]Prasadu Peddi

[1] PhD Research Scholar, JJT University Jhunjhunu Rajasthan 333001

[2]Professor, CSE Department, JJT University Jhunjhunu Rajasthan 333001

Abstract—There is a growing interest in detecting personality of an individual in a non-intrusive manner in areas like career development, counseling, interviews etc. Many solutions have been proposed for detection of personality from facial features alone. Different from it, this work proposes a model for detection of personality integrating the features learnt from face and human postures from videos. An integration of traditional and deep learning features in combination with machine learning algorithm is used to classify the personality of the individual. This work also proposes a frame of interest selection algorithm for selection of suitable frames in the video for personality assessment.

I. INTRODUCTION

Physiognomy is the practice of assessment of personality of an individual from the facial features or body structure of the person. It has its roots in ancient civilizations. Chinese named it as Mian Chiang meaning knowledge of facial features. It correlated facial features for various applications like disease diagnosis, assessment of personality type of individual, forecasting the future of the individual etc.

Modern day psychologists establish five core personality traits. The five personality traits are: openness, conscientiousness, extraversion, agreeableness and neuroticism. These personality factors represent a range between two extremes of low and high. The characteristics of each of the personality are listed below.

Personality	Characteristics
Openness	Very creative, open to trying new things, focused on tackling new challenges, happy to think about abstract concepts
Conscientiousness	Plan ahead, spends time preparing, finishes important tasks right away, pays attention to detail, enjoys having a set of schedule
Extraversion	Enjoys being a center of attention, likes to start conversations, enjoys meeting new people, likes to have wide social circle of friends and acquaintances, easy to make new friends, feels energized when around people, says things before thinking about them
Agreeableness	Have great interest in other people, cares about others, feel empathy, enjoys helping and contributing to happiness of other people, assist others who are in need of help
Neuroticism	Experiences a lot of stress, worries about many things, gets upset easily, experiences dramatic mood shifts, feels anxious

Many studies correlating the information in facial images to big five personality traits are available.

Static facial images are analyzed using image processing and machine learning algorithms to detect the personality traits of the person. Following are some of the challenges in personality trait classification using static facial clues.

1. Large number of features with some lacking the ability to be quantified.
2. Lack of information correlating many facial features to the personality trait
3. Non linear relation between the facial features and the personality traits

This work addresses this problem and proposes a personality prediction model integrating dynamic facial and posture features acquired from video. Following are some of the salient features of this work.

1. A model for salient frame selection from the videos which has feature rich information for personality trait classification.
2. An integrated model combining both facial and posture features for personality trait classification.
3. Identification of best set of features and machine learning classifier for predicting the personality

II. RELATED WORK

Awaja (2017) built a ontology based physiognomy system to provide personality traits from facial appearance. The ontology is built in OWL format mapping between the facial features to the big five personality traits. Semantic web rules are created between human face features to big five personality and this can be used as database for big five personality assessment systems. Al Moubayed et al (2014) proposed an approach for predicting big five personality from static face image. The face image is histogram equalized to increase brightness. Eigen face method is employed to do face representation and features around salient points of eyes, nostrils, chin tip, lip corners are extracted. Support Vector Machine (SVM) classifier with Gaussian radial basis function is trained to predict the personality from the salient features. Eddine et al (2017) proposed an approach for personality detection from facial features. Facial features from eye, lip , nose and ear regions are extracted in terms of pixel distance. Rule based mapping is done from facial features to the personality trait. Gavrilescu et al (2017) proposed a personality trait prediction system based on facial expression activities captured over a period of time. Hidden markov model (HMM) was used to guide the facial expression activity collection. Regression model is created mapping the activities to the big five personality traits. Gorbova et al (2017) proposed an automatic personality screening system based on visual, audio and lexical clues. Speech features like pitch, intensity, duration, Mel Frequency Cepstral Coefficients (MFCCs), formants, zero-crossing rate, and filter-bank energy parameters etc are extracted from short video clips. Facial features are extracted using OpenFace tool to extract around 416 features. Regression model is created associating speech and face features to the personality of the individual. G¨urpınar et al (2016) used face, video and scene features to predict the personality of the individual. Face features and scene features are extracted using deep learning VGG-network. Kernel ELM is then used to model the personality traits from the three features of face, video and scene. Ilmini et al (2016) using features extracted from facial landmarks to predict the personality trait of the individual. The traits values around the landmarks are used as input to artificial neural network (ANN) and SVM classifier to predict the personality traits. Kampman et al(2018) used three channel information of audio, video and text to predict the personality of the individual from video segments. Deep learning convolutional neural network is used to predict personality from each channel separately and then result is fused using weighted ensemble to predict the final personality trait of an individual. Kindiroglu et al (2017) predicted the extraversion personality trait necessary for leadership from the group interaction videos. Principal component analysis (PCA) is used for feature extraction; Maximum relevance maximum redundancy algorithm is used for feature selection and SVM is used for personality prediction. Levitan et al (2016) used speech features to predict the personality of the individual. Acoustic-prosodic and lexical features are extracted from speech and Adaboost algorithm is trained to classify the features to big five personality trait. Liu et al (2016) predicted the personality traits of an individual from the emotions

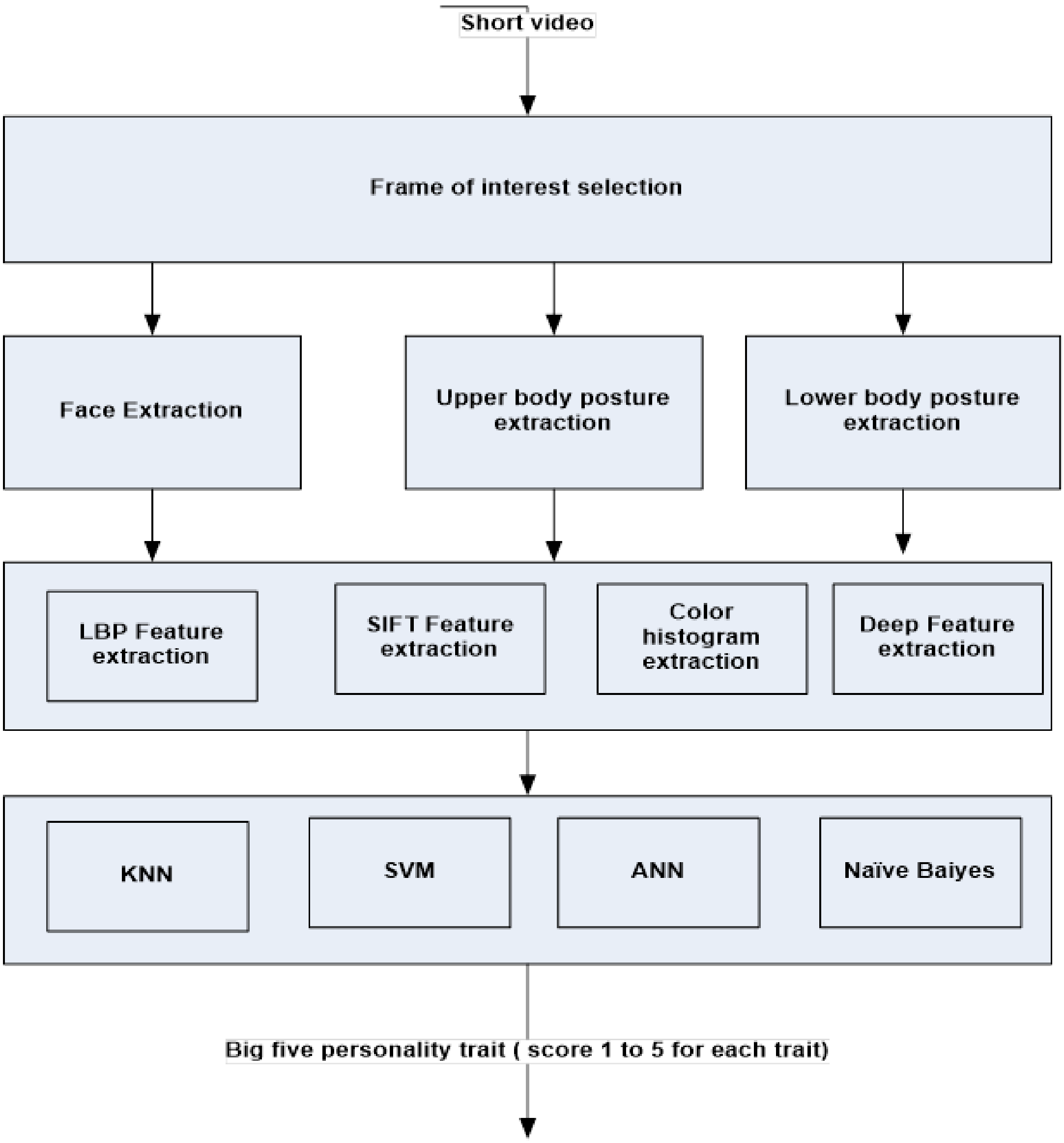


Figure 1 Architecture of proposed solution

expressed in the face. Emovu API is used to extract the facial expression features. Ekman's model of six discrete basic emotions and their mapping to big five personality trait is used for personality classification from the face images. Lingyun Wen et al (2015) proposed facial region visual-based nonverbal behavior analysis for depression diagnosis. Features are extracted from facial sub regions and these features are represented in sparse coding. The spare coded features SVM regression to diagnose depression. Madzlan et al (2014) applied machine learning for automatic prediction of attitudes from video blogs. Prosodic and visual features are extracted from the facial regions and it is used to classify the attitude. SVM is used to classify the facial appearance features to attitudes. Van Gerven et al (2016) used audio visual deep residual network for personality trait prediction. Audio and visual channels are directly passed as input to the deep residual network without any necessity for feature extraction and output is a value from 0 to 1 corresponding to a personality trait. Rai et al (2016) combined multiple modality specific model to provide the personality trait. Deep network extracts visual information from faces, supplementary information from background and acoustic features and fusion of all model results is done to predict the personality. Zhang et al (2017) built a end to end convolutional neural network to predict personality trait from the face image. In addition to personality trait classification, the approach is also able to measure the intelligence score by integrating regression to the VGG network.

III. PROPOSED SOLUTION

The architecture of the proposed personality trait prediction system is given in Figure 1.

Following are the important stages in the proposed solution

1. Extraction of relevant frames
2. Extraction of Features
3. Personality trait prediction

A. Relevant frame extraction

The frames of interest which has capability to predict the personality of an individual must be selected from the videos. The proposed relevant frame extraction algorithm has following steps

1. Extract Frames from videos
2. Skip the frame if human face cannot be found in frame using Voila Jones method.
3. Calculate the difference of frame with previous frame found.
4. If the difference is less than 10% skip the frame
5. Frames not skipped in step 4 are passed to the next stage as relevant frames.
6. Select key frames from the relevant frames using clustering approach given below

From each of the relevant frames, OpenFace [20] AU features and presence of upper and lower body parts in the frame is extracted. Grouping of frames based on feature similarity is done using K-Means clustering. From the cluster group, a frame close to the centroid of the cluster is selected as key-frames.

Frame allocation to cluster is done with a aim of minimization of sum of squared errors between cluster centroid and feature vector.

$$J(X_i, C) = \sum_{j \in N_i} ||x(j) - C(a(j))||^2$$

Where C is the set of centroids, a is the assigned center of each relevant frame.

B. Feature Extraction

We propose a Deep Feature extraction algorithm and experiment with following features to decide the best feature for personality classification.

Local binary patterns (LBP): A 8 point LBP with radius 1 is applied on the image and a 10 bin histogram is got as result. L2- normalized histogram is applied on 10 bin histogram providing the 10 dimension feature as output.

Scale invariant feature transform: A neighborhood around the center of the image is taken. This neighborhood is split into 16 sub-blocks. An 8 bin histogram is got from each of the sub-block and l2 normalization is applied to result in 128 dimensional feature vector.

Color histogram: A 64 bin is constructed in RGB color space for the input face image. L2 normalization is color bin to get the 64 dimension color feature.

Deep Features: A modified VGG 16 Convolutional Neural Network architecture is proposed in this work to extract the features.

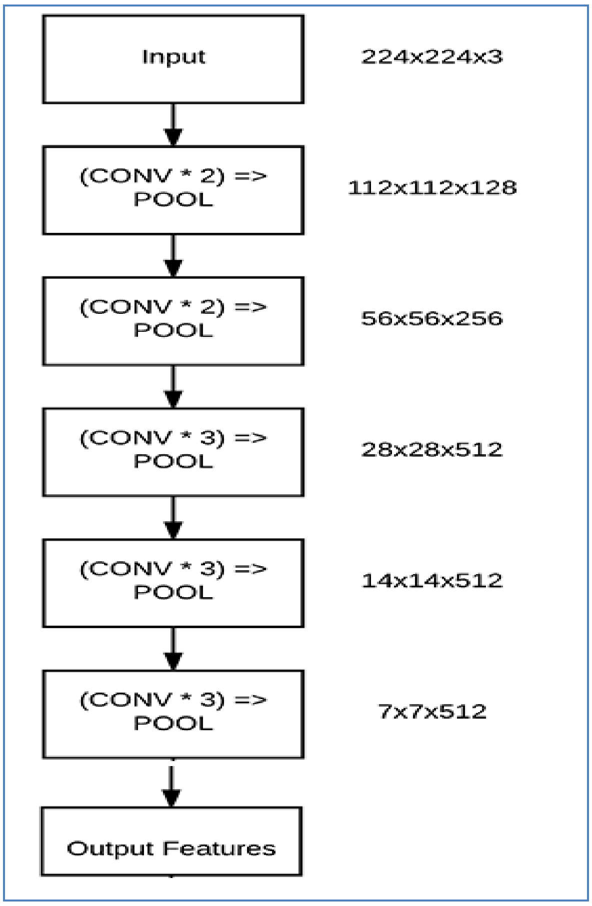


Figure 2 VGG Net for Feature extraction

A modified VGG16 deep neural network is applied to extract the feature. Feature is extracted at the final max pooling layer. Max pooling layer is a matrix of size 7 x 7 x 512. It is flatted in row major order to a feature vector of dimension 21,055.

This feature is extracted for all N images in the dataset yielding a feature matrix of size of 21,055*N.

Face image is segmented using Voila jones detector on the noise removed frames and features (all the above four) are extracted from the face image.

The human body upper and lower part voila jhones detector is run on the key frames and the posture is segmented. On the segmented posture, all the four features are extracted.

C. Personality trait prediction

This work experiments with following machine learning classifiers in different combination with features as shown below

	LBP	SIFT	Color histogram	Deep feature
KNN	×	×	×	×
SVM	×	×	×	×
ANN	×	×	×	×
Naïve Baiyes	×	×	×	×

The machine learning classifiers are trained to classify the features to big 5 personality traits with values for 1 to 5 for each trait. The maximum occurring personality found from processing of each of significant frame is provided as the result.

IV. NOVELTY IN PROPOSED SOLUTION

The proposed solution is different from existing solutions is following aspects

1. An integrated model combining facial and posture features is proposed for personality trait prediction
2. Frame of interest algorithm is proposed to select the frames with high correlation to personality assessment from the videos.
3. Best set of features with higher correlation to personality is selected to reduce the prediction time and improve the accuracy of prediction.

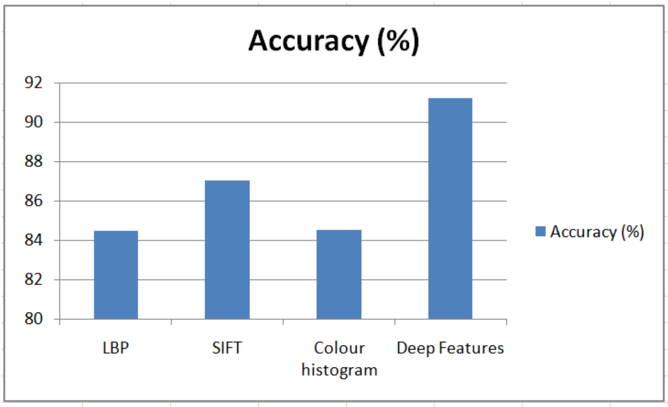
V. **RESULTS**

In 2007, researchers at Cambridge and Stanford [19] embarked on collecting an online dataset to study social/psychological trends. Their work resulted in the beginnings of the myPersonality project. Over the years, the project has collected several samples from Facebook user attributes such as likes, status updates, private messages, profile pictures and demographic information such as age, sex, marital status, geographical check ins, as well as personality test scores. Facebook users participated in the study out of their voluntary contribution. In our exploration, we use this dataset and our primary attributes of interest are Big5 (Five-Factor Model) personality scores and Facebook profile pictures. The score for each trait is a number between 1 and 5 (inclusive). The dataset distribution used for testing the performance of the proposed solution is as below

Parameters	Training	Validation
Number of images	1000	200
Age distribution of dataset	40:30:30	40:30:30
Gender distribution	60:40	60:40

The performance of the proposed features for classification of personality is measured and the result is given below.

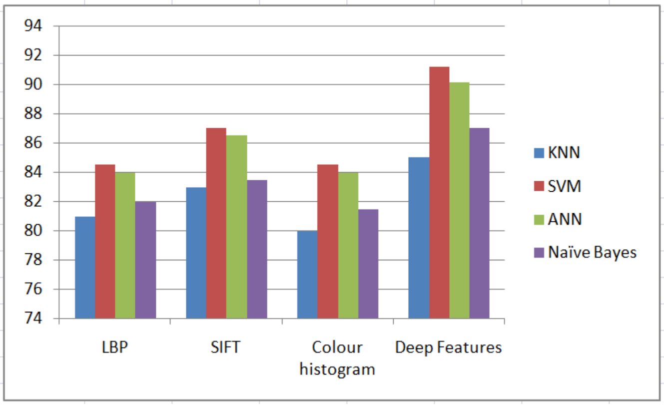
Features	Accuracy (%)	Positive rate (%)	Negative rate (%)	Unbiased metric (%)
LBP	84.50	46.89	84.06	65.47
SIFT	87.02	65.24	87.13	76.18
Colour histogram	84.52	55.48	84.33	69.90
Deep Features	91.22	78.35	91.89	85.12



Among all the features, deep learning provided a highest accuracy of 91.22% followed by SIFT features with a accuracy of 87.02%.

The accuracy results for each of classifier in combination with different features is given below

	LBP	SIFT	Colour histogram	Deep Features
KNN	81	83	80	85
SVM	84.50	87.02	84.52	91.22
ANN	84	86.50	84	90.12
Naïve Bayes	82	83.50	81.50	87



As seen from the results, SVM classifier has higher accuracy compared with rest of the classifiers. The highest accuracy is achieved in combination with Deep features. SVM classifier with Deep feature achieves an peak accuracy of 91.22%.

The performance of the proposed Deep features + SVM classifier is compared with

1. Deep Inference of Personality Traits by Integrating Image and Word Use in Social Networks [3]

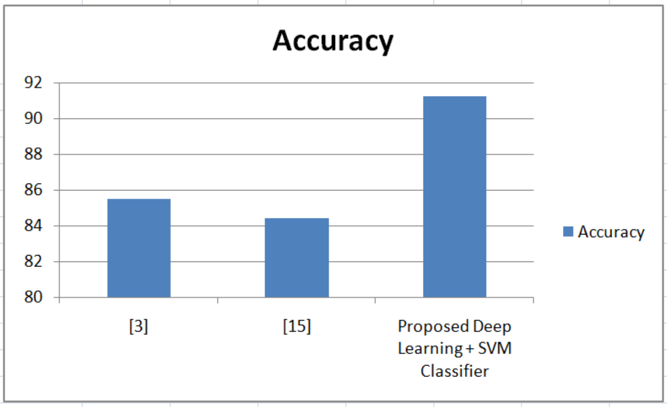
2. Predicting Social Perception from Faces: A Deep Learning Approach [15]

The performance is compared in terms of

1. Accuracy
2. Time for classification
3. Mean Square Error (MSE)

The results for accuracy is as below

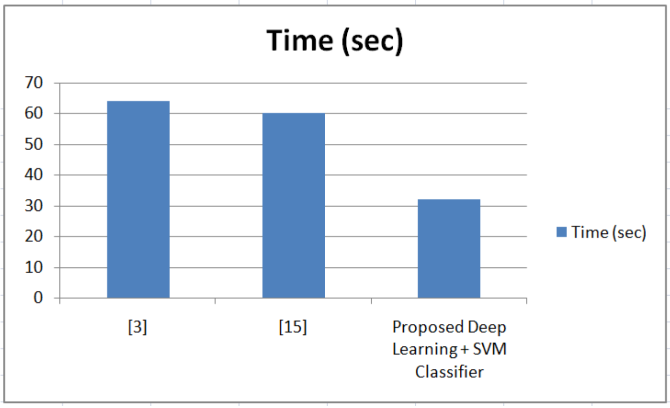
Solution	Accuracy
[3]	85.50
[15]	84.40
Proposed Deep Learning + SVM Classifier	91.22



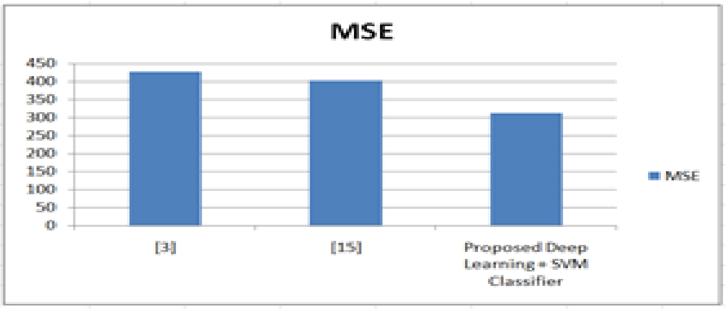
Deep learning in combination with SVM classifier achieves the highest accuracy. It is 6.82% more than [3] and 5.72% more than [15].

The average time for classification is measured and given below

Solution	Time (sec)
[3]	64
[15]	60
Proposed Deep Learning + SVM Classifier	32



The classification time is 87.5% lower compared to [3] and 100% lower compared to [15]. The MSE is calculated between actual and predicted personality. The result is given below



The proposed solution has 28.93% lower MSE compared to [24] and 37.29% lower MSE compared to [23].

VI. CONCLUSION

This work proposed a deep learning feature extraction along with SVM classifier for personality prediction. The salient feature of this approach is that it used integrated features extracted from both face and postures for personality classification. With use of Deep learning features and SVM classifier, the proposed solution was able to a achieve a peak accuracy of 91.22% which is 5.72% more than existing solutions at comparatively lower execution time of 87.5% lower than existing solution. In future, we plan to integrate demographic and age specific modalities to personalize the personality trait classification for different demographics and age.

REFERENCES

1. Awaja, K. (2017). Human Personality Derivation Using Ontology-Based Modern Physiognomy. Master thesis. Islamic University. Gaza. http://hdl.handle.net/20.500.12358/20163
2. Al Moubayed N, Vazquez-Alvarez Y, McKay A, Vinciarelli A (2014) Face-based automatic personality perception. In: Proceedings of the 22nd ACM international conference on multimedia. ACM, pp 1153–1156. https://doi.org/10.1145/2647868.2655014
3. Cucurull, Guillem & Rodriguez, Pau & Yazici, V. & Gonfaus, Josep & Roca, Xavier & Gonzàlez, Jordi. (2018). Deep Inference of Personality Traits by Integrating Image and Word Use in Social Networks.
4. Eddine Bekhouche S, Dornaika F, Ouafi A, Taleb-Ahmed A (2017) Personality traits and job candidate screening via analyzing facial videos. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 10–13. DOI: 10.1109/CVPRW.2017.211
5. Gavrilescu, M., & Vizireanu, N. (2017). Predicting the Sixteen Personality Factors (16PF) of an individual by analyzing facial features. Journal on Image and Video Processing, 59, 1-19. https://doi.org/10.1186/s13640-017-0211-4
6. Gorbova J, Lusi I, Litvin A, Anbarjafari G (2017) Automated screening of job candidate based on multimodal video processing. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 29–35. DOI: 10.1109/CVPRW.2017.214
7. Gürpınar F., Kaya H., Salah A.A. (2016) Combining Deep Facial and Ambient Features for First Impression Estimation. In: Hua G., Jégou H. (eds) Computer Vision – ECCV 2016 Workshops. ECCV 2016. Lecture Notes in Computer Science, vol 9915. Springer, Cham. https://doi.org/10.1007/978-3-319-49409-8_30
8. Ilmini, Kalani & Fernando, T.G.I.. (2016). Persons' Personality Traits Recognition using Machine Learning Algorithms and Image Processing Techniques. Advances in Computer Science. 5. 40-44. ISSN : 2322-5157
9. Kampman O, Barezi EJ, Bertero D, Fung P (2018) Investigating audio, video, and text fusion methods for endto-end automatic personality prediction. In: Proceedings of the 56th annual meeting of the association for computational linguistics (volume 2: short papers), vol 2, pp 606–611.DOI: 10.18653/v1/P18-2096
10. Kindiroglu AA, Akarun L (2017) Aran O (2017) Multi-domain and multi-task prediction of extraversion and leadership from meeting videos. EURASIP J Image Video Proc 1:77.

DOI: 10.1186/s13640-017-0224-z

11. Levitan SI, Levitan Y, An G, Levine M, Levitan R, Rosenberg A, Hirschberg J (2016) Identifying individual differences in gender, ethnicity, and personality from dialogue for deception detection. In: Proceedings of the second workshop on computational approaches to deception detection, pp 40–44.DOI: 10.18653/v1/W16-0806
12. Liu L, Preotiuc-Pietro D, Samani ZR, Moghaddam ME, Ungar LH (2016) Analyzing personality through social media profile picture choice. In: ICWSM, pp 211–220. DOI: 10.1109/ICICT.2017.8320184
13. Lingyun Wen, Xin Li, Guodong Guo, Yu Zhu, "Automated Depression Diagnosis Based on Facial Dynamic Analysis and Sparse Coding",IEEE Transactions on Information Forensics and Security, volume 10, issue 7, pp. 1432-441, March 2015. DOI: 10.1109/TIFS.2015.2414392
14. Madzlan NA, Han J, Bonin F, Campbell N (2014) Automatic recognition of attitudes in video blogs—prosodic and visual feature analysis. In: Fifteenth annual conference of the international speech communication association. 1826-1830, 37060642
15. Messer, Uwe & Faußer, Stefan. (2019). Predicting Social Perception from Faces: A Deep Learning Approach.
16. M.A.J. van Gerven and R. van Lier, Deep Impression: Audiovisual Deep Residual Networks for Multimodal Apparent Personality Trait Recognition, Cham:Springer International Publishing, pp. 349-358, 2016. DOI: 10.1007/978-3-319-49409-8_28
17. Rai N (2016) Bi-modal regression for apparent personality trait recognition. In: 23rd International conference on pattern recognition (ICPR). IEEE, pp 55–60. DOI: 10.1109/ICPR.2016.7899607
18. Zhang, T., Qin, R., Dong, Q., Gao, W., Xu, H., & Hu. Z. (2017). Physiognomy: Personality traits prediction by learning. International Journal of Automation and Computing, 14(4), 386–395. DOI: 10.1007/s11633-017-1085-8
19. M. Kosinski, S. C. Matz, S. D. Gosling, V. Popov, and D. Stillwell. Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. American Psychologist, 70(6):543, 2015
20. OpenFace 2.0: Facial Behavior Analysis Toolkit Tadas Baltrušaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency, IEEE International Conference on Automatic Face and Gesture Recognition, 2018