SERBIAN JOURNAL OF ELECTRICAL ENGINEERING
Vol. 20, No. 2, June 2023, 229-242

UDC: 004.032.26::159.9.019.4 **DOI: https://doi.org/10.2298/SJEE2302229V**

Trait Analysis Based on Multimodal Prediction and Optimization of the Output Parameters: A Survey

Milić Vukojičić[1], Mladen Veinović[1]

Abstract: This paper provides a survey of apparent personality trait analysis based on multimodal trait prediction and optimization of multimodal output parameters. This paper analyzes, synthesizes, and compares the development of different methods of Personality computing based on the personality trait analysis within the Big Five model (OCEAN model) and based on audio, visual and textual inputs. The results of the trait analysis and detection are different for different methods, from trait extracted from handwriting with an accuracy of 62-83%, trait extracted from with 56-62%, trait extracted from audio with 70%, trait extracted from visual element images and videos 87-91%, from aggregation function min and max 52 - 81%, mean and median 85%, and methods based on robustness and Huber function 72 - 77 %. Trait analysis is a problem with numerous real-life applications. Accurate trait recognition and analysis of various human traits are tasks that have been studied for a very long time in the field of psychology and lately represent an important area of study in the field of computer science and computing. The best results in the field of apparent personality trait analysis were achieved with convolutional neural networks. Multimodal trait prediction will give a better prediction than prediction based on a single modality, and optimization with aggregation functions and robust methods can achieve a better prediction of the models.

Keywords: Personality computing, First impressions, Big-five, Aggregation functions, Personality classification, Feature classification, Robust loss function.

1 Introduction

Human personalities are studied by psychologists and many theories are formed about how they can be better and closer represented and categorized. We can define a human personality as a form of psychological construct that can explain the way that humans are behaving. In order to represent human behavior effectively, we need these characteristics to be stable over time and consistent. In the paper represented by Costa & Mccrae [1] they started the introduction with the words: "*For anyone who truly wishes to understand human personality, trait psychology is not an option*". In their work, they showed that human traits can be

[1]Department of Advanced Systems Protection, Singidunum University, Belgrade, Serbia;
E-mails: vukojicic.milic@gmail.com; mveinovic@singidunum.ac.rs

measured from multiple aspects like patterns of human behavior, human thoughts, and human emotions. This can help us make better predictions over time since they are stable. Many of the models that are used for understanding human behavior and human traits are based on human perception. They are represented with the words like open-minded, shy, and enthusiastic. These words are often used when a person wants to describe himself, where semantics and the relation between adjectives are very similar to all of these descriptions.

Research based on the stability of traits is represented in the work proposed by Funder [2]. As he showed in his work, some human thoughts and behaviors depend on the situation and change over time. The problem that occurs is how and when we need to detect the human traits if they can vary over time. In this regard, the RAM (Realistic Accuracy Model) can provide a more accurate understanding of personality - which traits should be considered, when they should be considered, how they should be considered according to the situation, and who is giving the judgment.We can see that the personality judgments are very close to the accurate state that the person is in. In the next sections we will show how we can overcome this by adding the algorithm as an observer.

There are several trait models which are represented in the literature and which are mostly used when we want to describe personality. Models like Big-Two, Big-Five, and 16PF are mostly used and researched in the literature. The Big-Two model [3] which is derived from the Big-Five model shows how we can represent human personality from two factors: Plasticity and Stability. Other models like the 16PF [4] model that was represented in 1970. by Cattel, Eber and Tatsuoka show that human personality can be presented from the next 16 dimensions: Abstractedness, Apprehension, Dominance, Emotional stability, Liveliness, Openness to change, Perfectionism, Privateness, Reasoning, Rule-consciousness, Self-reliance, Sensitivity, Social boldness, Tension, Vigilance, and Warmth. Models like 16PF and Big-Two can be used as a valuable models for detecting and measuring people's traits, sometimes only two aspects for the detection can be misleading, and sometimes we can have a potential problem of detecting too many aspects. In the middle ground, we can see that we have a model called Big-Five or OCEAN model because this model uses five basic dimensions of personality. The model is proposed in the paper "*An introduction to the five-factor model and its applications*" [5] in 1986. by McCrae & John. O-CEAN is the second name used for a Big-Five model because of the dimensions that we are going to detect: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism, Fig. 1. The model can be used as a valuable source of human trait detection and remains stable for different observers. In our case the observer will be the neural network, and it also can be used across different cultures.

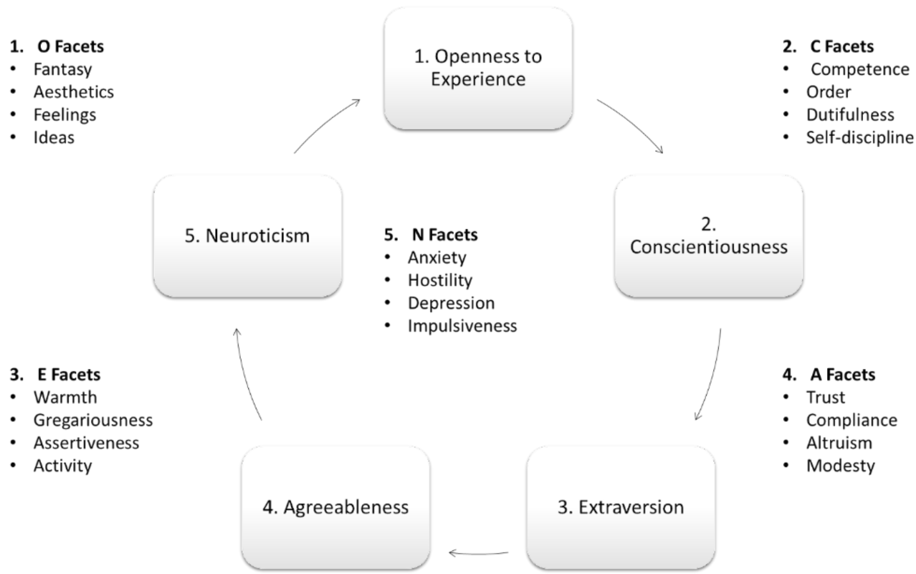

Fig. 1 – *Big 5 (OCEAN) personality trait model and trait facets.*

Moreover, the Big-Five model represents the most studied model in the field of psychology, and it is one of the most used models in the human resources department. It is mostly used in automatic personality detection in the field of computer science and computing. Psychologists are also improving the Big-Five model and they introduce facets for each of the OCEAN traits [5]. We can see many improvements in the facets in the paper "*Between facets and domains: 10 aspects of the Big Five*" published by DeYoung, Quilty and Peterson in 2007. The most used facets are:

– aesthetics, fantasy, feelings, ideas for openness to experience;

– competence, dutifulness, order, self-discipline, deliberation for conscientiousness;

– gregariousness, warmth, excitement-seeking, positive emotions for extroversion;

– straightforwardness, modesty, trust, altruism for agreeableness;

– depression, impulsiveness, self-consciousness, vulnerability for neuroticism.

To test the Big-Five model, Costa and McCrae invented the Revised NEO Personality Inventory (NEO PI-R), a personality inventory test [7]. NEO Personality Inventory can be used for detecting Big-Five personality traits: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism and also their facets.

2 Apparent Personality Trait Analysis in Computing

Apparent personality trait analysis is a part of computing that is called Personality Computing (PC). Personality Computing can be defined as a field of computer science that is used for the extraction and detection of personality information. In the paper proposed by Phan & Rauthmann, we can see that Personality Computing and the detection of the personality relevant information can be different and affected by the type of sensor that we are going to use for the reliability and validity of data [8]. In their paper, they also represented legal, moral, and ethical views in the field of Personal Computing and also societal implications of the field.

The main idea of extracting human personality traits from various modalities is not new to psychology. However, today's digital data is more and more dominating over the analog representation of data so PC is more and more useful and easier to use and process data. The modalities that are generally used in PC are written text, digital text, audio, visual, social media, smartphones, and wearables. Written and digital text can be extracted from essays, blog posts, transcripts, and many more data sources. There is a big difference between the written, analog and representation of human writing and digital text. We will see that we can extract more information from text which is written by hand than the digital representation of that text. Audio and visual modalities can be extracted from the conversation, vlogs, videos, and visual modalities, especially from images and self-presentations. The usage of these technologies is very broad. We can use them in real-time while someone is talking or while we are listening to someone. Modality like social media is very interesting from the research perspective because we can have a huge variety of data from users' profiles, posts, pictures, likes, etc. Modalities like smartphones and wearables are very interesting because of the number of sensors that they have, things like location, Bluetooth scanning and WiFi sharing, ambient sound, speech or physical activity, or any kind of interaction that can give us a lot of specific data.

Two papers, released by authors Vinciarelli and Mohammadi: "*A survey of personality computing*" and "*More personality in personality computing*" are published in 2014. They showed that Personality Computing is growing fast and approaches to how we can extract human traits from different modalities are vast. Also, they showed the main problem of Personality Computing like Automatic Personality Recognition (ARP), Automatic Personality Perception (APP) and Automatic Personality Synthesis (APS). ARP will represent the true personality of the individual based on their behavior. APP tries to represent the personality of an individual based on their observable behavior and monitored by others. APS will represent the creation of the fabricated personality of the individual judged by the agents [9, 10].

The Fig. 2 shows the relationship between the Brunswik Lens and the three main problems addressed in Personality Computing. APR is the inference of self-assessments (µS in the figure) from distal cues, APP is the inference of assessments (µP in the figure) from proximal cues, APS is the generation of artificial cues aimed at eliciting the attribution of predefined traits [9, 11]. In the next section, we will explain the three main problems in Personality Computing: ARP, APP and APS and their relation with the Brunswik Lens model [11].

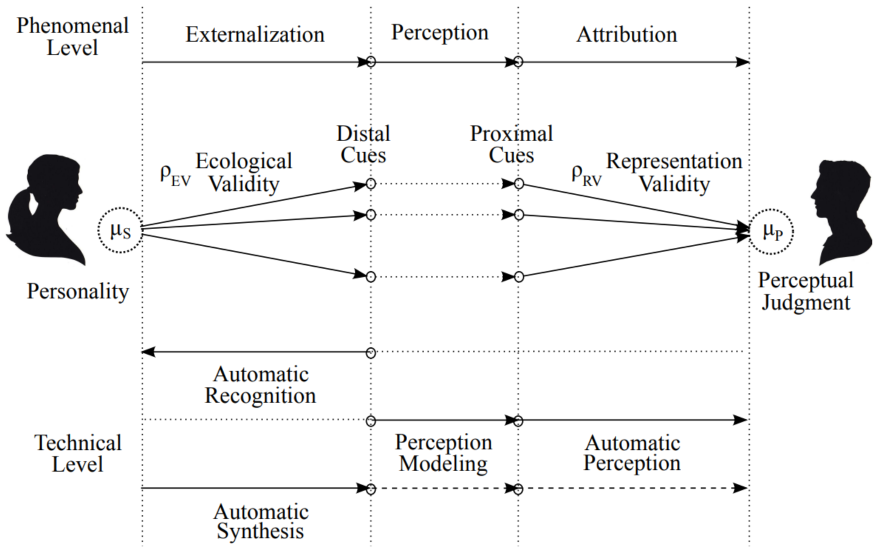

Fig. 2 – *Brunswik Lens and the three main problems addressed in Personality Computing* [9, 11].

2.1 APR, APP and APS

APR is related to Personality Externalization - a person's personality can be shown by using distal cues, and this personality can be observed by others. The individual will externalize their personality, in a way that is not observable by them, but their personality will be observable by others from the physical markers, these markers will be shown through everything that person does [12]. APR will represent a way of detecting the distal cues, and the traits that are detected from the self-assessment tools. APR will use methodologies like Social Signal Processing (SSP), Affective Computing (AC), and many others that will be related to the social and emotional fields that machines can detect. Ecological Validity is shown on Fig. 2 will represent the covariation between distal cues and individual personality measurement.

APP will be connected to personality attribution. Here we will make a distinction between distal cues and proximal cues, as the cues that person, observer, actually perceives. The attribution process will be connected to the proximal cues and later on in the creation of the perceptual judgment. APP will be a process in which the observer will create perception of the observed person based on the proximal cues. The result of this will not be the real, or true personality of the observed person, but only the perception of the observer. In the Personality Perception process, the APP can also include multiple observers, and the traits that are extracted from the multiple observers can be used as an average.

APS will represent the part that automatically generates distal cues from personal traits. If we take a look at Fig. 2 and Brunswik Lens we can say that APS will include the phase of externalization and attribution. The main interest for the computation files and computer science field is that these cues can be generated not only by humans but by any entity that has human-like behavior. This process aims at making sure that the characteristics of the person that observes are those that machine designers envisioned [9].

We can conclude that Personality Computing can help in the trait detection of the observed individual by minimizing the potential problems that an observer or multiple observers can create. Since all of the computation methods are still relying on their creators, we can also transfer some of the same problems that we have in the field of Personality Science to the field of Personality Computing. The part of personality detection where Personality Computing is best implemented is APR, APP, and APS, especially in the fields of signal processing and machine learning [10].

3 Multimodal Trait Prediction

Using multimodal data for automatic trait prediction is one of the best approaches in the field of Personality Computing. Multimodal trait prediction is often related to several input parameters like images, sound, audio, and video parameters which are combined into a single discrete value that represents every single one of OCEAN/Big-5 characteristics.

The oldest method of personal trait extraction will be the extraction based on an individual's handwriting. Personality based on the individual's handwriting will be represented as one of the hardest methods of trait prediction and human personality trait analysis. This method is mostly based on the patterns that can occur in the individual's handwriting. The paper with the name "*Automatic personality identification using writing behaviors: an exploratory study*" published by Zhi Chen & Tao Lin showed that if we combine the writing methods with the machine learning methods the result of the prediction from the handwriting can lead to 62.5% to 83.9% accuracy [13]. The result also showed that writing parameters is a good and stable indicator of traits over time. In the paper "*Identify* *human personality parameters based on handwriting using neural network*” authors Fallah & Khotanlou showed that handwriting can also include indicators for happiness, excitement, different emotional states, anger, and calm. The work in their paper showed neural networks based on the Hidden Markov Model [14] where the Hidden Markov Model is used for classification of the individual traits. The results from the Hidden Markov Model and neural networks are showing the accuracy of 72%.

The next modality of multimodal trait prediction is the method based on personality trait extraction from the text. Here we will define a text as digital data and it can be in the form of a typed short sentence, tweet, post, essay, or longer written content. In the domain of personality trait detection and extraction, we have popular models based on word vectorization [15] shown in the paper “*Linguistic regularities in continuous space word representations*”, convolutional neural networks in combination with Mairesse shown in the paper “*Deep learning-based document modeling for personality detection from the text*” [16]. Here the result that the authors are representing is showing the accuracy of 56% to 62%.

Personality trait extraction based on audio is mostly based on the time domain features and frequency domain features. Personality Computing and trait extraction will be based on the Mel Frequency Cepstral Coefficients (MFCC) and Linear Prediction Cepstral Coefficient (LPCC) [17, 18]. These models are mostly giving an accuracy of 70%. The main problem with trait detection and extraction from the audio signal will be the quality of the signal and the noise presented in the signal itself.

The modalities that are mostly used in Personality Computing for automatic train detection and extraction is images. Here we need to take into consideration sequences of images (video material) without audio, or individual images. This modality in the literature is also visual modality. The main research is mostly based on YouTube and Facebook and other video sharing platforms, on which authors are undertaking multimodal sentiment analysis, with the usage of deep learning, Gated Multimodal Embedding LSTM with Temporal Attention (GME-LSTM(A)) [19]. Other approaches are based on the convolutional neural networks and done on images [20] where authors achieved an error rate of 37% to 17%, with the neural network with 650.000 neurons. The design of a network was this: five convolutional layers, a max-pooling layer, three fully-connected layers, a softmax layer, and output. Other approaches were using Deep Bimodal Regression (DBR) framework, with the convolutional neural network for visual cues [21] with the combination of the linear regression for audio cues. Some authors are basing their models not only on the audio, and visual cues but also on the surrounding of the subject based on the deep convolutional neural networks, with Local Gabor Binary Patterns from Three Orthogonal Planes video descriptor an openSMILE for the acoustic features [22]. The accuracy of the models shown in these papers is from 0.87 to 0.91.

Table 1
OCEAN results extracted from audio-visual and surrounding elements [22] *extracted from the First impressions (ECCV'16, ICPR'16) challenge.*

Team name	O	C	E	A	N
BU-NKU	0.914	0.915	0.918	0.907	0.911
evolgen	0.911	0.914	0.916	0.911	0.910
pandora	0.904	0.901	0.904	0.905	0.900

The most popular results based on the multimodal trait prediction with the state-of-the-art models are based on two or three different modalities. The main problem is that some modalities represent some traits better and some traits worse.

4 Optimization of The Trait Prediction with Aggregation Model

The framework proposed in the paper "*Apparent Personality Analysis based on Aggregation Model*" [23] for the trait extraction model based on the aggregation function is shown in Fig. 3.

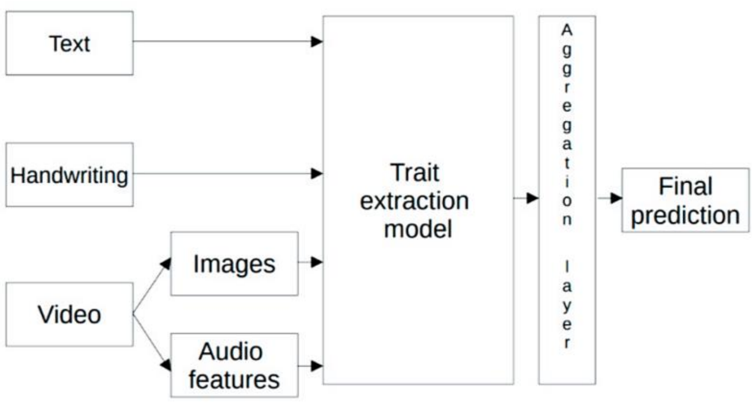

Fig. 3 – *Framework based on Aggregation layer* [24].

Fig. 3 shows the proposed model based on an aggregation layer that has four different inputs: video, is treated as having two natural modalities, text and handwriting one natural modality. The final predicted personality traits are created from the aggregation layer by the fusion of all outputs of text, handwriting, images and audio features. [24]. The main idea of introducing the aggregation function to the trait prediction model is to use the power of multiple inputs such as image, handwriting, video, and audio, and produce the output that can represent the final trait based on these inputs. The final output needs to be a single discrete value for the aggregation model. Authors are often using aggregation functions min, max, mean and median. These aggregation functions are very easily implemented in the already existing models. Functions can define as min as the smallest value that can be found in the input and max as the biggest value that can be found in the input of these functions and together mean and median [24].

Authors Vukojičić and Veinović are using aggregation functions to get better predictions of the several models that have different input parameters: handwriting, text, image, audio, and video. The authors are giving a new perspective to the aggregation function and in the paper, we can see that some aggregation functions like mean and median can be very effective in the domain of multimodal trait prediction. The problem that can occur is that min and max aggregation functions can create outliers that can be wrongly interpreted. Moreover, the models based on the aggregation functions can be very rigid. As we mentioned in the introduction section of this paper human traits are very constant over time but there can be variations of certain traits. The results of the min and max functions are giving relative errors between 52% and 81% for all traits, and from 1% to 15% for mean and median aggregation functions.

5 Optimization of the Trait Prediction Based on Robust Estimation

Optimization of the trait prediction based on the robust estimation is represented in the paper "*Apparent personality analysis based on robust estimation*" [25]. The framework shown in the paper is presented in Fig. 4.

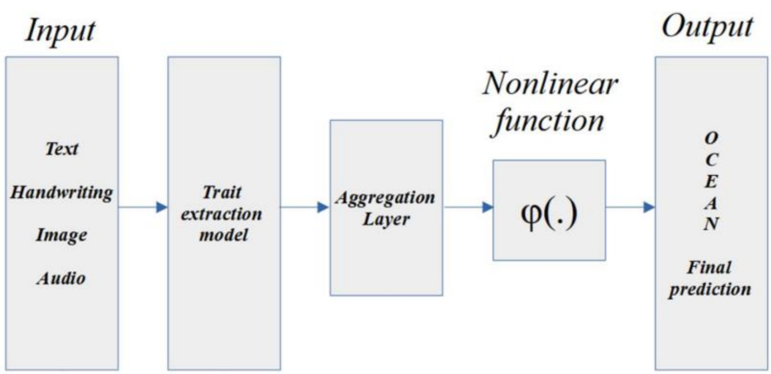

Fig. 4 – *Framework based on nonlinear Huber function* [25].

Fig. 4 shows the framework with inputs: text, handwriting, image, and audio. These traits are then extracted and on them will be used different aggregation functions (min, max, mean and median) will be applied to them. An aggregation function framework is showing optimization with nonlinear function and then extraction of final OCEAN characteristics parameters between 0 and 1 [25]. The main concept of robust estimation is that we need to move from the linear functions represented in Section 4 to a more natural way of optimizing the parameters. The idea here is to use the linear parameters as a starting point and then introduce the Huber nonlinear function [26] to these parameters. By introducing Huber nonlinearity authors represented the idea of a more flexible and natural way of optimization of the multimodal trait prediction. The Huber loss function can be defined as:

$$L_\delta(a) = \begin{cases} \frac{1}{2}a^2, & \text{for } |a| \le \delta; \\ \delta(|a| - \frac{1}{2}\delta), & \text{otherwise.} \end{cases} \tag{1}$$

The authors are showing that we can use Huber nonlinear function to get a better optimization of the trait prediction algorithms. Using a nonlinear system is showing that we can have a more natural way of predicting traits, and they showed that we can overcome the problem created by min and max function, the problem of outliers. The results shown in the paper are showing that by using Huber nonlinear function we can improve the results from the min and max function giving relative error between 52% and 81% to the error between 23% and 28%.

Table 2

OCEAN results extracted from audio-visual and surrounding elements [24, 25] *based on aggregation and nonlinear models.*

Method	**O**	**C**	**E**	**A**	**N**
Min	81%	79%	75%	71%	77%
Max	52%	60%	72%	63%	68%
Mean	10%	6%	7%	1%	1%
Median	5%	3%	15%	5%	5%
Huber	24%	28%	25%	24%	23%

In the papers "*Apparent Personality Analysis based on Aggregation Model*" and "*Apparent personality analysis based on robust estimation*" the authors deducted an experiment involving the comparison of the results taken from the subject's NEO-PI-R tests and frameworks shown in Figs. 3 and 4. The database was constructed with 64 subjects, 33 male and 31 female subjects, between 18 and 70 years old. The database is constructed with people of different job titles and education levels.

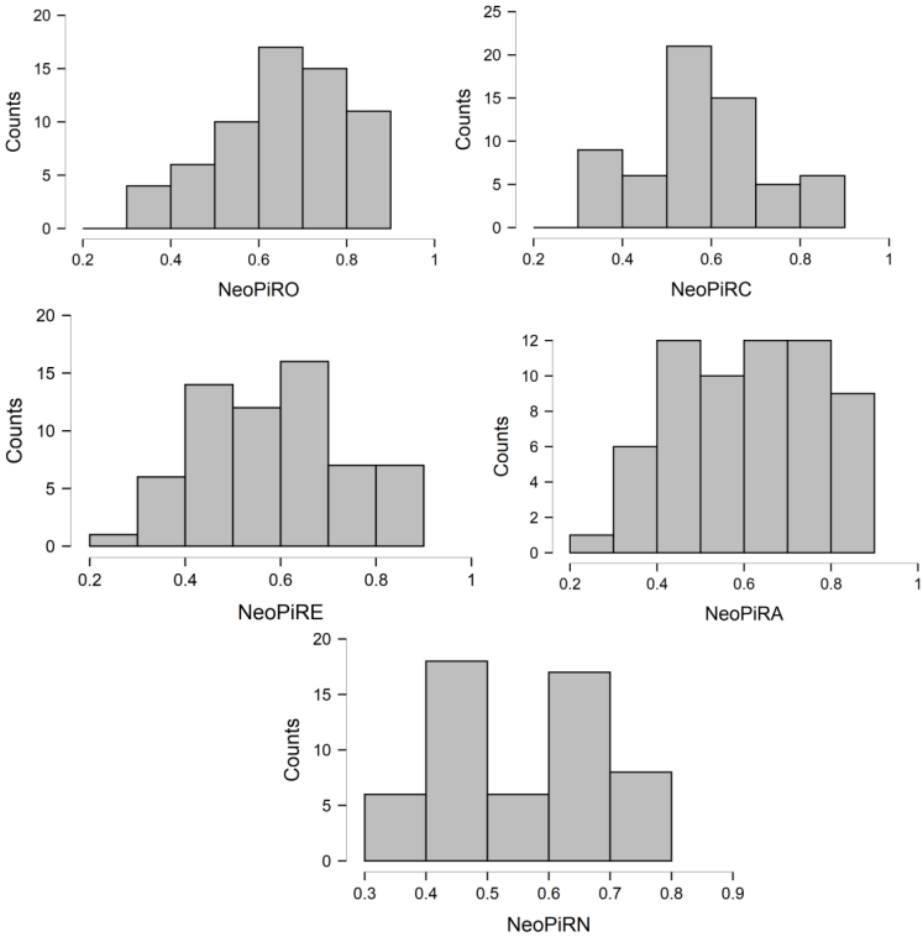

Fig. 5 – *Distribution of NEO-PI-R test results* [25].

6 Optimization of the Trait Prediction Based on Particle Swarm Optimization

Paper "*Optimization of Multimodal Trait Prediction Using Particle Swarm Optimization*" introduce the Soft Aggregation Model (SAM) [26]. SAM is based on the Min, Max, Mean and Median aggregation functions, a Huber function to minimize outliers, and a Particle Swarm Optimization algorithm used on nonlinear functions. Previous work used linear aggregation functions and nonlinear functions, but they resulted in rigid values and a less natural performance. The new model uses swarm intelligence to provide a more natural approach and a less rigid result. The Huber function is employed to make the system less sensitive to outliers, and therefore improve the results of the PSO algorithm.

This paper presents the usage of the Particle Swarm Optimization (PSO) algorithm proposed by Kennedy & Eberhart (1995) [27] which simulates the behavior of a flock of birds or a shoal of fish and is based on multiple agents called particles.

The paper showed that SAM leads to better prediction of personal traits. The Particle Swarm Optimization (PSO) algorithm is used to get better results with different iteration and swarm size parameters. It is revealed that the optimal values of the swarm size are between 4 and 25 and of the several iteration parameters is between 10 and 25. Results of the absolute and relative error of the SAM model are shown in **Table 3**.

Table 3
OCEAN results extracted from audio-visual and surrounding elements [26] *based on SAM model.*

SAM model	**O**	**C**	**E**	**A**	**N**
Absolute error	0.11	0.14	0.10	0.14	0.13
Relative error	18%	20%	22%	18%	19%

7 Conclusion

Apparent personality analysis and detection are very important part of the Computer Science and Personality Computing domain. In the last decade trait analysis based on different inputs (text, audio, video, and handwriting) are increasingly popular among researchers. Trait analysis from handwriting based on Hidden Markov Models, text-based on convolutional neural networks, audio based on Mel Frequency Cepstral Coefficients (MFCC) and Linear Prediction Cepstral Coefficient (LPCC), and images based on convolutional networks are rising in popularity in the research field. Some models have an accuracy of 91% based on the multimodal trait analysis. Novel research in the field of apparent personality analysis is conducted on the optimization of output parameters in the multimodal trait prediction based on aggregation models and robust methods with a relative error rate from 5% for aggregation function to 25% for models based on robust methods.

The very big interest in the domain of Personal Computing and trait analysis is the real-life application of the field. There are various applications in the field of psychology, trait detection and extraction can help psychologists in this area, in computer science where researchers are improving models for detection and analysis. The application of these kinds of models can be in the field of human resources for candidate selection, psychological assessments, selection of candidates for universities and schools, etc. Models can also be used in the management departments of the companies for monitoring the employees to achieve better performance in the workplace.

Future work, that is proposed, in the field of apparent trait analysis is based on soft computing methods like swarm intelligence and a mixture of experts networks.

7 References

[1] P. T. Costa, R. R. McCrae: Trait Theories of Personality, Ch. 5, Advanced Personality, Springer, Boston, 1998.

[2] D. C. Funder: Accurate Personality Judgment, Current Directions in Psychological Science, Vol. 21, No. 3, June 2012, pp. 177 – 182.

[3] A. E. Abele, B. Wojciszke: Agency and Communion from the Perspective of Self Versus Others, Journal of Personality and Social Psychology, Vol. 93, No. 5, November 2007, pp. 751 – 763.

[4] R. B. Cattell, S. E. Krug: The Number of Factors in the 16PF: A Review of the Evidence with Special Emphasis on Methodological Problems, Educational and Psychological Measurement, Vol. 46, No. 3, Autumn 1986, pp. 509 – 522.

[5] R. R. McCrae, O. P. John: An Introduction to the Five-Factor Model and Its Applications, Journal of Personality, Vol. 60, No. 2, June 1992, pp. 175 – 215.

[6] C. G. DeYoung, L. C. Quilty, J. B. Peterson: Between Facets and Domains: 10 Aspects of the Big Five, Journal of Personality and Social Psychology, Vol. 93, No. 5, November 2007, pp. 880 – 896.

[7] P. T. Costa Jr., R. R. McCrae: The Revised NEO Personality Inventory (NEO-PI-R), Ch. 9, The SAGE Handbook of Personality Theory and Assessment, Vol. 2, pp. 179 – 198, SAGE Publications Ltd, London, 2008.

[8] L. V. Phan, J. F. Rauthmann: Personality Computing: New Frontiers in Personality Assessment, Social and Personality Psychology Compass, Vol. 15, No. 7, July 2021, p. e12624.

[9] A. Vinciarelli, G. Mohammadi: A Survey of Personality Computing, IEEE Transactions on Affective Computing, Vol. 5, No. 3, July - September 2014, pp. 273 – 291.

[10] A. Vinciarelli, G. Mohammadi: More Personality in Personality Computing, IEEE Transactions on Affective Computing, Vol. 5, No. 3, July - September 2014, pp. 297 – 300.

[11] E. Brunswik: Perception and the Representative Design of Psychological Experiments, 2nd Edition, University of California Press, Berkeley, Los Angeles, 1956.

[12] Social Markers in Speech, Edited by K. R. Scherer, H. Giles, Cambridge University Press, Cambridge, London, New York, Melbourne, 1979.

[13] Z. Chen, T. Lin: Automatic Personality Identification Using Writing Behaviours: An Exploratory Study, Behaviour & Information Technology, Vol. 36, No. 8, March 2017, pp. 839 – 845.

[14] B. Fallah, H. Khotanlou: Identify Human Personality Parameters based on Handwriting Using Neural Network, Proceedings of the Artificial Intelligence and Robotics (IRANOPEN), Qazvin, Iran, April 2016, pp. 120 – 126.

[15] T. Mikolov, W.- T. Yih, G. Zweig: Linguistic Regularities in Continuous Space Word Representations, Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Atlanta, USA, June 2013, pp. 746 – 751.

[16] N. Majumder, S. Poria, A. Gelbukh, E. Cambria: Deep Learning-Based Document Modeling for Personality Detection from Text, IEEE Intelligent Systems, Vol. 32, No. 2, March – April 2017, pp. 74 – 79.

[17] L. H. Gilpin, D. M. Olson, T. Alrashed: Perception of Speaker Personality Traits Using Speech Signals, Proceedings of the CHI Conference on Human Factors in Computing Systems, Montreal, Canada, April 2018, pp. 1 – 6.

[18] S. Davis, P. Mermelstein: Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 28, No. 4, August 1980, pp. 357 – 366.

[19] M. Chen, S. Wang, P. P. Liang, T. Baltrušaitis, A. Zadeh, L.- P. Morency: Multimodal Sentiment Analysis with Word-Level Fusion and Reinforcement Learning, Proceedings of the 19th ACM International Conference on Multimodal Interaction, Glasgow, UK, November 2017, pp. 163 – 171.

[20] A. Krizhevsky, I. Sutskever, G. E. Hinton: ImageNet Classification with Deep Convolutional Neural Networks, Communications of the ACM, Vol. 60, No. 6, June, 2017, pp. 84 – 90.

[21] C.- L. Zhang, H. Zhang, X.- S. Wei, J. Wu: Deep Bimodal Regression for Apparent Personality Analysis, Proceedings of the European Conference on Computer Vision, Amsterdam, Netherlands, October 2016, pp. 311 – 324.

[22] F. Gürpinar, H. Kaya, A. A. Salah: Multimodal Fusion of Audio, Scene, and Face Features for First Impression Estimation, Proceedings of the 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, December 2016, pp. 43 – 48.

[23] M. Grabisch, J.- L. Marichal, R. Mesiar, E. Pap: Aggregation Functions, Cambridge University Press, Cambridge, New York, Melbourne, Madrid, 2009.

[24] M. Vukojičić, M. Veinović: Apparent Personality Analysis based on Aggregation Model, Proceedings of the International Scientific Conference on Information Technology and Data Related Research, Belgrade, Serbia, June 2021, pp. 220 – 225.

[25] M. Vukojičić, M. Veinović: Apparent Personality Analysis based on Robust Estimation, Proceedings of the 5th International Conference on Applied Informatics - Imagination, Creativity, Design, Development (ICDD), Sibiu, Romania, November 2021, pp. 1 – 10

[26] B. Kovačević, M. M. Milosavljević, M. Veinović, M. Marković: Robust Digital Processing of Speech Signals, Springer, Academic Mind, Cham, Belgrade, 2017.

[27] M. Vukojičić, M. Veinović: Optimization of Multimodal Trait Prediction Using Particle Swarm Optimization, Studies in Informatics and Control, Vol. 31, No. 4, December 2022, pp. 25 – 34.

[28] J. Kennedy, R. Eberhart: Particle Swarm Optimization, Proceedings of the International Conference on Neural Networks (ICNN), Perth, Australia, November 1995, pp. 1942 – 1948.