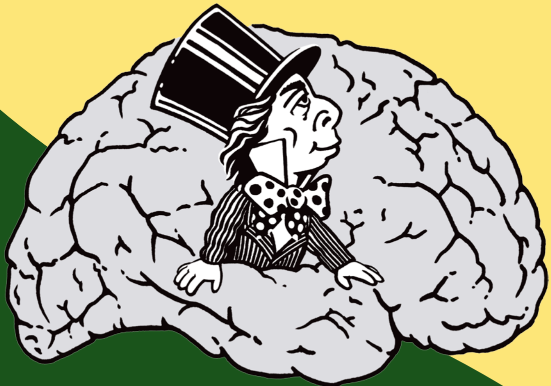
КОГНИТИВНАЯ НАУКА

В МОСКВЕ



НОВЫЕ ИССЛЕДОВАНИЯ

МАТЕРИАЛЫ КОНФЕРЕНЦИИ 2023

Под ред. Е.В. Печенковой, М.В. Фаликман, А.Я. Койфман

УДК 159.9
ББК 88.25
К57

Когнитивная наука в Москве: новые исследования. Материалы конференции 21 – 22 июня 2023 г. Под ред. Е.В. Печенковой, М.В. Фаликман, А.Я. Койфман. – М.: ООО «Буки Веди», Московский институт психоанализа. 2023 г. — 604 стр.

ISBN 978-5-4465-3880-5

УДК 159.9
ББК 88.25

ISBN 978-5-4465-3880-5 © Авторы статей, 2023

CULTURALLY SENSITIVE EMOTION DETECTION USING A LARGE PSYCHOPHYSIOLOGICAL DATASET WITH A RUSSIAN SAMPLE

E. O. Ivanina* (1), A. D. Tokmovtseva (1, 2), E. V. Akeleva (1, 3)
info@leo-neurolab.ru
1 — Neurollab, Moscow;
2 — Russian Orthodox University of Saint John the Divine, Moscow;
3 — Moscow State University, Moscow

Abstract. Machine learning algorithms for emotion detection require large amounts of data to make accurate predictions. However, common psychophysiological databases often have small samples and represent only western communities. Thus, we decided to collect a large psychophysiological dataset for emotion recognition in a Russian sample, using culturally relevant stimuli. We recorded 63,586 cases (174 participants; age: 27.8 ± 11.6 years) of eye tracking, galvanic skin response (GSR) and heart rate variability (HRV). Respondents viewed a random selection of 200 out of 800 pictures representing Russian culture and nature and assessed their emotional state after viewing using the Self-Assessment Manikin (SAM). Recordings were cleared of artifacts and noise and fed to the emotion classification model. We created a simple multimodal neural network to investigate the possibility of using the dataset to predict emotional states. Combining the visual (images, scanpath) and time series (eye tracking, GSR, HRV) data, we achieved prediction accuracies of 65.8% (valence) and 59.1% (arousal) for three classes of emotion. As a result of this study, a multimodal and culture-specific psychophysiological database was developed that achieved moderate but stable results in emotion recognition tasks.

Keywords: emotion detection, deep learning, cross-cultural studies, eye tracking, biometrics, galvanic skin response, heart rate

Introduction

Emotions are a significant part of human existence, crucial to peoples' understanding of interpersonal relationships and the world around them. As emotions and their accurate and nuanced understanding is essential in advertisement (Consoli, 2010), the demand in automatic, data-based approaches for emotion detection has been on the rise.

Today's commercial and non-profit solutions mostly utilize data or stimuli sets collected in Western Europe or USA samples (Wang et al., 2022) without considering the effect of culture on emotional expression. Based on the previous studies, it could be noticed that Russians have a tendency to show high levels of negative emotions such as shame, fear and distress and lower levels of joy (Wierzbicka, 1998) and might have a better self-control of emotion expression (Pankratova et al., 2013), possibly leading to less extreme rating in affect measurements. Thus, to make accurate predictions we should develop the methodological tools which accentuate the cultural impact on perception. Due to the specificity of our stimuli which could contain important symbolism to Russian participants, we needed to provide an additional measure free of possible socially desirable responses. Eye-tracking and biometrics is often applied in commercial research as an alternative to traditional methods such as interviews or questionnaires which could be influenced by social biases (Jordao et al., 2017). Eye movement deviations, for instance, fixation instability (Alshehri, Alghowinem, 2013) or the pattern of fixations (Liang et al., 2017) could be predictors of pleasant or unpleasant emotions or anxiety. Galvanic skin response (GSR), heart rate (HR) and pupil size all could capture the state of vegetative nervous system activation (Wang, Minor, 2008), often termed arousal. Dimensional theories of emotion often pair dimension of Arousal with Valence (an estimation of an emotional state from positive to negative) and are widely used in automatic emotion detection (Osuna et al., 2020). We also employed a dimensional model, namely Russel's circumplex model of affect (Posner et al., 2005), in our research, thus associating self-reported emotions with psychophysiological measures.

Thus, the main objective of this study was to create a dataset of psychophysiological reactions to affective culturally relevant stimuli for emotional classification machine learning tasks.

Method

One hundred seventy-four respondents (121 women; 27.8 ± 11.6 years) took part in the study. We collected demographics to assess the socio-economic status of our sample. All the participants had normal or corrected vision, given their written consent and were paid for participating in the experiment.

In the process of data collection and stimuli set development, we followed a common procedure for emotional recognition studies (Kurdi et al., 2017; Skaramagkas et al., 2023). We used Gazepoint GP3 60Hz eye tracker and 60Hz biometrics kit, including biometric signal tracker for capturing heart rate and galvanic skin response (Gazepoint, Canada, https://www.gazept.com). Pictures were presented on ASUS TUF Gaming VG249Q 24" monitor with 1920 × 1080 resolution. The 800 images were selected from open-source image libraries. Pictures were resized to 500 × 400 pixels which was the only modification. The influence of characteristics of the stimuli was controlled further using neural network output. The images represented Russian culture and nature and were divided into 4 categories, similarly to the OASIS study (Kurdi et al., 2017): animals, people, scenes, and objects, with an equal number of pictures in each category ($N = 200$).

The procedure was performed as follows: respondents were asked to take a comfortable position on the armchair in front of the computer during the experiment. After a nine-point calibration, respondents viewed 200 random items from 800 pictures of the set and were instructed to evaluate their emotional state after viewing the picture using the Self-Assessment Manikin (SAM). SAM is a seven-point scale, a non-verbal questionnaire that is commonly applied in emotion classification field (Sainz-de-Baranda Andujar et al., 2022). The pictures and the SAM were presented in OpenSesame software (Mathôt et al., 2012). Each picture was in the center of the screen for 10 seconds. Participants proceeded to the next picture after completing SAM presentation of which did not have timing constraints. The whole experimental procedure took on average from 25 min to an hour.

Figure 1. Stimuli example for "People" category (left) and "Scenes" category (right)

Final recordings were restored using linear interpolation (Sanchez-Comas et al., 2021), smoothed with a moving window, and normalized and cut into five-second epochs as the data beyond that mark reached a plateau. Final features included mean pupil diameter (mm), fixation coordinates X and Y (fraction of the screen size), GSR (kΩ), and HR (beats per minute). We also calculated mean for the SAM dimensions and standard deviation scores for each picture to analyze the resulting distribution.

Results

To probe the dataset applicability to emotion detection tasks we used a multimodal neural network consisting of small convolutional neural network (NN) for image-like data of scan paths, long short-term memory (LSTM) NN for table-like time-series data of fixation coordinates, GSR, HR (Sims, Conati, 2020), and pre-trained NN (VGG16) for image classification (Simonyan, Zisserman, 2015). The outputs of all three parts were concatenated and put through several fully connected layers. We achieved accuracy of prediction at 29.5% for valence and 27.5% for each of seven classes as in the SAM questionnaire. Allowing the error of one point in prediction of valence and arousal limited the amount of classes to three and increased the accuracies to 65.8% (valence) and 59.1% (arousal).

Discussion and Conclusions

Emotion detection datasets could suffer from sample representativity (Paullada et al, 2021) and size issues which limits their applicability to machine learning tasks. As an attempt to overcome these limitations, we collected a large psychophysiological data corpus of affective reactions in an uncommon sample of Russian participants with detection accuracy on par with similar studies (Sims et al., 2019; Liu et al., 2016). Current research met two main issues: imbalanced classes, which was solved by training the same architecture separately for arousal and valence prediction, and an overcomplication of the results imposed by seven-class analysis, which we tried to overcome by limiting classes to three with a prediction error of one point.

In conclusion, we developed the combined stimuli and the psychophysiological dataset which could be used in affect detection tasks. We are planning to continue our research in order to increase the ability of the stimuli to induce emotions and reduce the visual biases of the selected stimuli, thus, hopefully, obtaining higher prediction accuracies and quality of detection. The final version of the dataset will be available for the research purposes on our website https://leo-neurolab.ru/.

References

Alshehri M., Alghowinem S. An exploratory study of detecting emotion states using eye-tracking technology // Science and Information Conference. London, UK: 2013. P. 428 – 433.

Consoli D. A new concept of marketing: The emotional marketing // BRAND. Broad Research in Accounting, Negotiation, and Distribution. 2010. Vol. 1. No. 1. P. 52 – 59.

Jordao I.L.D.S., Souza M.T.D., Oliveira J.H.C.D., Giraldi J.D.M.E. Neuromarketing applied to consumer behaviour: an integrative literature review between 2010 and 2015 // International Journal of Business Forecasting and Marketing Intelligence. 2017. Vol. 3. No. 3. P. 270. https://doi.org/10.1504/ijbfmi.2017.10006047

Kurdi B., Lozano S., Banaji M.R. Introducing the Open Affective Standardized Image Set (OASIS) // Behavior Research Methods. 2016. Vol. 49. No. 2. P. 457 – 470. https://doi.org/10.3758/s13428-016-0715-3

Liang C.-W., Tsai J.-L., Hsu W.-Y. Sustained visual attention for competing emotional stimuli in social anxiety: An eye tracking study // Journal of Behavior Therapy and Experimental Psychiatry. 2017. Vol. 54. P. 178 – 185. https://doi.org/10.1016/j.jbtep.2016.08.009

Liu M., Fan D., Zhang X., Gong X. Retracted: Human emotion recognition based on galvanic skin response signal feature selection and SVM // 2016 International Conference on Smart City and Systems Engineering (ICSCSE). IEEE, 2016. https://doi.org/10.1109/icscse.2016.0051

Mathôt S., Schreij D., Theeuwes J. OpenSesame: An open-source, graphical experiment builder for the social sciences // Behavior Research Methods. 2012. Vol. 44. No. 2. P. 314 – 324. https://doi.org/10.3758/s13428-011-0168-7

Osuna E., Rodríguez L.-F., Gutierrez-Garcia J.O., Castro L.A. Development of computational models of emotions: A software engineering perspective // Cognitive Systems Research. 2020. Vol. 60. P. 1 – 19. https://doi.org/10.1016/j.cogsys.2019.11.001

Pankratova A., Osin E., Lyusin D. Differences in emotional intelligence in Russian and Azerbaijani culture // Psychological Research. 2013. Vol. 6. No. 31. P. 11. https://doi.org/10.54359/ps.v6i31.673

Paullada A., Raji I.D., Bender E.M., Denton E., Hanna A. Data and its (dis)contents: A survey of dataset development and use in machine learning research // Patterns. 2021. Vol. 2. No. 11. P. 100336. https://doi.org/10.1016/j.patter.2021.100336

Posner J., Russell J.A., Peterson B.S. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology // Development and Psychopathology. 2005. Vol. 17. No. 03. P. 715 – 734. https://doi.org/10.1017/s0954579405050340

Sainz-De-Baranda Andujar C., Gutiérrez-Martín L., Miranda-Calero J., Blanco-Ruiz M., López-Ongil C. Gender biases in the training methods of affective computing: Redesign and validation of the Self-Assessment Manikin in measuring emotions via audiovisual clips // Frontiers in Psychology. 2022. Vol. 13. P. 955530. https://doi.org/10.3389/fpsyg.2022.955530

Sanchez-Comas A., Synnes K., Molina-Estren D., Troncoso-Palacio A., Comas-González Z. Correlation analysis of different measurement places of galvanic skin response in test groups facing pleasant and unpleasant stimuli // Sensors. 2021. Vol. 21. No. 12. P. 4210. https://doi.org/10.3390/s21124210

Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition // International conference on learning representations. arXiv, 2015. https://doi.org/10.48550/arXiv.1409.1556

Sims S., Conati C. A neural architecture for detecting user confusion in eye-tracking data // Proceedings of the 2020 International Conference on Multimodal Interaction. ACM, 2020. https://doi.org/10.1145/3382507.3418828

Sims S., Putnam V., Conati C. Predicting confusion from eye-tracking data with recurrent neural networks // 2nd Workshop on Humanizing AI (HAI) at IJCAI'19. 2019. https://doi.org/10.48550/arXiv.1906.11211

Skaramagkas V., Ktistakis E., Manousos D., Kazantzaki E., Tachos N.S., Tripoliti E., Fotiadis D.I., Tsiknakis M. eSEE-d: Emotional State Estimation based on Eye-Tracking Dataset // Brain Sciences. 2023. Vol. 13. No. 4. P. 589. https://doi.org/10.3390/brainsci13040589

Wang Y., Song W., Tao W., Liotta A., Yang D., Li X., Gao S., Sun Y., Ge W., Zhang W., Zhang W. A systematic review on affective computing: Emotion models, databases, and recent advances // Information Fusion. 2022. Vol. 83-84. P. 19 – 52. https://doi.org/10.1016/j.inffus.2022.03.009

Wang Y.J., Minor M.S. Validity, reliability, and applicability of psychophysiological techniques in marketing research // Psychology and Marketing. 2008. Vol. 25. No. 2. P. 197 – 232. https://doi.org/10.1002/mar.20206

Wierzbicka A. Russian emotional expression // Ethos. 1998. Vol. 26. No. 4. P. 456 – 483. https://doi.org/10.1525/eth.1998.26.4.456

ДЕТЕКЦИЯ ЭМОЦИЙ ДЛЯ КУЛЬТУРНО ЗНАЧИМЫХ СТИМУЛОВ С ИСПОЛЬЗОВАНИЕМ БОЛЬШОГО КОРПУСА ПСИХОФИЗИОЛОГИЧЕСКИХ ДАННЫХ РОССИЙСКОЙ ВЫБОРКИ

Е. О. Иванина (1), А. Д. Токмовцева (1, 2), Е. В. Акельева (1, 3)
info@leo-neurolab.ru
1 — Нейроллаб, Москва; 2 — Российский православный университет, Москва;
3 — Московский государственный университет, Москва

Аннотация. Точность алгоритмов машинного обучения во многом зависит от величины датасета, и область распознавания эмоций не исключение. Однако, часто корпусы психофизиологических данных имеют небольшой объем выборки и собраны на респондентах из США или Европы. Исходя из этого, мы решили разработать психофизиологический датасет аффективных реакций на российской выборке, используя культурно значимые стимулы. Было собрано 63586 записей (174 участника; возраст: 27.8 ± 11.6 лет) движений глаз, кожно-гальванической реакции (GSR) и вариабельности сердечного ритма (HRV). Респонденты просматривали случайные 200 из 800 картинок, демонстрирующих отечественную культуру и природу и оценивали свое эмоциональное состояние с помощью методики моделей для самооценки эмоций (Self-Assessment Manikin, SAM). Записи были очищены от артефактов и шума. Была создана мультимодальная нейронная сеть, совмещающая данные об изображениях, путях взгляда (scan path) и временных рядах (движения глаз, GSR, HRV), для оценки возможности использования датасета для предсказания эмоциональных состояний. Была достигнута точность в 65.8 % для валентности и 59.1 % для возбудимости при предсказании трех классов эмоциональных состояний. Таким образом, в результате исследования был разработан мультимодальный датасет, использующий культурно-сенситивную базу стимулов, позволяющую получать стабильные предсказания эмоций средней точности.

Ключевые слова: детекция эмоций, машинное обучение, кросс-культурные исследования, айтрекинг, биометрия, кожно-гальваническая реакция, фотоплетизмограмма