Yevheniia Zubrych, Sergii Stirenko.

EMOTION RECOGNITION METHOD FOR DISTANCE LEARNING SYSTEMS

This paper analyzes the existing emotion recognition solutions and the impact of emotions on the learning process. Based on the obtained results, the method of emotion recognition is proposed. It includes such steps as images collecting, image normalization, face detection, face identification, and emotion recognition. The collected information about audience emotional state could be used by a lecturer to indicate the level of understanding and adjust the teaching method to a specific audience or students.

Keywords: Emotion recognition, distance learning, convolutional neural network.

Fig.: 1. Tabl.: 0. Bibl.: 19.

Relevance of research topic. Emotions play a key role in learning. Thus, emotional prosody is one of the key conditions that determine a student` ability to memorize and focus on the learning process. Emotionally colored knowledge is remembered faster and stronger than the knowledge that is devoid of individuality and leaves a person indifferent.

On the other hand, the students` emotions is a good indicator of understanding of the material and of the involvement in the educational process. This information could help teachers and parents to learn more about the effectiveness of the learning process and help teachers to improve teaching.

Formulation of the problem. In conducting classes, teachers find it difficult to quickly read and analyze the emotional state of students, especially when the group of students is large, and they need to pay attention to many students at the same time.

Analysis of recent research and publications. Recent research shows that emotions significantly affect learning strategy and student motivation, there is a correlation between anxiety and learning effectiveness. Thus, students who experience high test anxiety have lower grades [1]. Emotions have a significant impact on human cognitive processes, including perception, attention, learning, memory, reasoning, and problem solving [2]. Therefore, emotions can guide and motivate students. Analyzing and understanding students' emotions can help teachers better organize the teaching process and increase their effectiveness [3].

There were proposed various approaches for recording and measuring students' emotions in learning scenarios. In general, they are divided into two categories: methods of self-reporting, which allow collecting data on the emotional state directly by students, and methods of assessing emotions by external participants or systems that do not require self-reporting. The advantage of the second type of methods is that they do not distract students from the learning process. At the same time, it will not be easy for teachers to analyze the emotional state of each group of students during the lesson [6]. Automated methods of emotion recognition solve this problem. Examples of this type include the method of recognizing emotions by speech analysis [4], as well as computer vision methods [5].

Methods of recognizing the emotions of one or more faces simultaneously, based on information from cameras, are becoming widespread [7]. Currently, some studies have adapted emotion recognition technology to rapidly changing situations, such as during a group discussion [8].

Selection of unexplored parts of the general problem. There are solutions for the application of emotion recognition methods during learning processes in the classroom [9] [10], but they are designed to recognize faces and emotions from a camera that is installed in the room and captures a group of students. In this paper, it is proposed to focus on distance learning systems such as online conferencing, where each student participates in the conference using their own webcam. In such systems, control over the shooting and the right to stop it is left to each of the users, which gives more personal freedom.

Setting objectives. This work aims to develop a method of students' emotional state recognition for distance learning systems that could be used for audience emotional state monitoring.

Presentation of the main material

Each video V is a series of images:

$$V = \{I_1, \dots, I_i, \dots, I_N\}, \tag{1}$$

where I_i- i-th frame of the V, and N - the total number of frames in V. Due to the fact that each video contains a large number of frames, and neighboring frames are generally very similar to each other, it makes sense to remove unnecessary frames from the video, thereby reducing the amount of data for processing. Today, the most common standard for video conferencing is up to 30 frames per second [11]. By reducing the frequency to 2 frames per second, we reduce computing costs by 15 times without significant information loss.

The process of recognizing emotions can be divided into several stages. The first step is to determine the area occupied by the person's face in the frame. In most cases, during online conferences, there is one person in each frame, but there may be cases where there is more than one person or none. It is also necessary to consider that a person moves almost constantly in the frame, and his camera can be at different angles.

The Multi-task Cascaded Convolutional Network, which can handle multiple faces at once and, unlike openCV and dlib, can find small faces in group photos and videos of several dozen, can best handle the task of finding a face in the frame. As a result, we obtain the position of each face p_j in frame I_i, ie the face F_i can be defined as follows:

$$Fi = \{ p_1, \dots, p_j, \dots, p_{N_i} \}, \tag{2}$$

where N_i - the total number of faces in the frame I_i, , and p_j - rectangular area, which is occupied by j-th face in the frame.

The second stage is face recognition, as it is necessary to identify each face according to a person. To do this, certain facial features from a given image should be compared with the faces in the database.

The last stage is the recognition of emotions on the face. Formal models of emotions aim to define emotions in a form that could be used in the design of robots. In general, these models are divided into two main categories: discrete emotional states and dimensional models of emotions.

In the theory of discrete emotions, all people have an innate set of basic emotions. These basic emotions are described as discrete because they differ in facial expressions and biological processes [14]. There have been done the research to determine which emotions are the main ones. According to a study by Paul Ekman and colleagues, there are six main emotions - anger, disgust, fear, happiness, sadness, and surprise [15].

In dimensional models, emotions are modeled as points in a continuous n-dimensional space, such as Pleasure, Arousal, and Dominance (PAD) [16] or Valence and Arousal (VA) [17].

Dimensional models are more flexible and cover a wider range of emotions, but they are less transparent and intuitive than discrete ones, so discrete models are better suited for this scenario.

CNN training and testing require a dataset of facial expressions. For this scenario, it is advisable to use the Radboud Faces Database (RaFD) [18], which contains 8040 images from 5 different camera angles, and FER-2013, which consists of 35685 examples of facial images of 48x48 pixels, all images of which are divided into categories according to emotions shown (happiness, neutral, sadness, anger, surprise, disgust, fear) [19].

The result of this stage will be a set of emotions with probability values:

$$P = \{P(e) \mid \Sigma\, P(e) = 1, e \in\ emotions\}, \tag{3}$$

where *emotions* is a discrete set of emotions, and *P (e)* is the value of the probability that the corresponding emotion is shown in the image. Emotions are based on dataset categories, such as *emotions = {happiness, neutral, sadness, anger, surprise, disgust, fear}*.

After the probabilities *P(e)* for the corresponding image have been calculated, the emotion with the highest value of *P(e)* is selected, and it will be defined as the emotion depicted on I_i. Once the face is identified and its emotion is determined for each frame, this data is collected and processed according to the requirements of the system in which this method is used. For example, for a distance learning system, the lecturer receives detailed values of the emotional state for each of the participants in the system and thus could adjust the teaching method. The detailed sequence of steps is shown in fig.1.

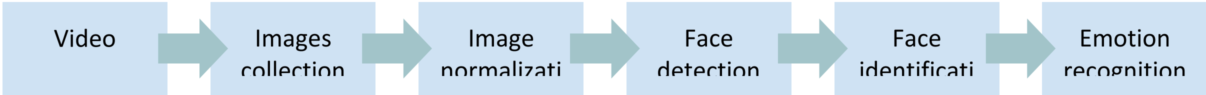

Fig. 1. The emotion recognition method steps.

Conclusions. This paper analyzes the existing emotion recognition solutions and the impact of emotions on the learning process. Based on the obtained results, the method of emotion recognition is proposed. It includes such steps as images collecting, normalization, face detection, face identification, and emotion recognition. The information about students' or audience emotional state could be used by a lecturer to indicate the level of understanding and adjust the teaching method to a specific audience or students.

References

1. Adesola S.A., Li Y., Liu X. Effect of Emotions on Students Learning Strategies // Proceedings of the 2019 8th International Conference on Educational and Information Technology - ICEIT 2019. 2019.

2. Tyng C.M. и др. The Influences of Emotion on Learning and Memory // Frontiers in Psychology. 2017. T. 8.

3. Pekrun R. Emotions: Functions and Effects on Learning // Encyclopedia of the Sciences of Learning. 2012. C. 1141–1146.

4. Ramakrishnan S. Recognition of Emotion from Speech: A Review // Speech Enhancement, Modeling and Recognition- Algorithms and Applications. 2012.

5. Kalaiselvi R., Kavitha P., Shunmuganathan K.L. Automatic emotion recognition in video // 2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE). 2014.

6. Meyer D.K. Situating Emotions in Classroom Practices // International Handbook of Emotions in Education.

7. Cohen I. et al. Facial expression recognition from video sequences: temporal and static modeling // Computer Vision and Image Understanding. 2003. Vol. 91, № 1-2. P. 160–187.

8. Weber, H., Cruz Rodriguez, A., Mateus, A. // Emotion and Mood in Design Thinking.

9. Zeng H. et al. Emotion Cues: Emotion-Oriented Visual Summarization of Classroom Videos // IEEE Transactions on Visualization and Computer Graphics. 2020. P. 1–1.

10. Putra W.B., Arifin F. Real-Time Emotion Recognition System to Monitor Student's Mood in a Classroom // Journal of Physics: Conference Series. 2019. Vol. 1413. P. 012021.

11. Meeting and phone statistics [Electronic resource] // Zoom Help Center. URL: https://support.zoom.us/hc/en-us/articles/202920719-Meeting-and-phone-statistics.

12. Zhang K. et al. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks // IEEE Signal Processing Letters. 2016. Vol. 23, № 10. P. 1499–1503.

13. Tarasov A.V., Savchenko A.V. Emotion Recognition of a Group of People in Video Analytics Using Deep Off-the-Shelf Image Embeddings // Lecture Notes in Computer Science Analysis of Images, Social Networks and Texts. 2018. P. 191–198.

14. Colombetti G. From affect programs to dynamical discrete emotions // Philosophical Psychology. 2009. Vol. 22, № 4. P. 407–425.

15. Ekman P. An argument for basic emotions // Cognition and Emotion. 1992. Vol. 6, № 3-4. P. 169–200.

16. Russell J.A., Mehrabian A. Evidence for a three-factor theory of emotions // Journal of Research in Personality. 1977. Vol. 11, № 3. P. 273–294.

17. Sun K. et al. An improved valence-arousal emotion space for video affective content representation and recognition // 2009 IEEE International Conference on Multimedia and Expo. 2009.

18. Radboud Faces Database [Electronic resource] // Radboud University Nijmegen. URL: http://www.socsci.ru.nl:8180/RaFD2/RaFD.

19. FER-2013: Wolfram Data Repository [Electronic resource] // FER-2013 | Wolfram Data Repository. URL: https://datarepository.wolframcloud.com/resources/FER-2013.

AUTHORS

Yevheniia Zubrych – student, Computer Engineering Department, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: evg.zubrich@gmail.com

Sergii Stirenko (supervisor) - head of Computer Engineering Department, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"