Uğur Ayvaz
MSc in Information Systems Engineering (ongoing), Department of Information Systems Engineering, Research Assistant
Mugla Sitki Kocman University, Mugla, Turkey
ugur.ayvaz@mu.edu.tr

Hüseyin Gürüler
PhD in Electronic and Computer Education, Department of Information Systems Engineering, Assistant Professor
Mugla Sitki Kocman University, Mugla, Turkey
ORCID ID 0000-0003-1855-1882
hguruler@mu.edu.tr

Mehmet Osman Devrim
MSc in Information Systems Engineering (ongoing), Department of Information Systems Engineering, Student
Mugla Sitki Kocman University, Mugla, Turkey
mehmetosmandevrim@posta.mu.edu.tr

USE OF FACIAL EMOTION RECOGNITION IN E-LEARNING SYSTEMS

Abstract. Since the personal computer usage and internet bandwidth are increasing, e-learning systems are also widely spreading. Although e-learning has some advantages in terms of information accessibility, time and place flexibility compared to the formal learning, it does not provide enough face-to-face interactivity between an educator and learners. In this study, we are proposing a hybrid information system, which is combining computer vision and machine learning technologies for visual and interactive e-learning systems. The proposed information system detects emotional states of the learners and gives feedback to an educator about their instant and weighted emotional states based on facial expressions. In this way, the educator will be aware of the general emotional state of the virtual classroom and the system will create a formal learning-like interactive environment. Herein, several classification algorithms were applied to learn instant emotional state and the best accuracy rates were obtained using kNN and SVM algorithms.

Keywords: e-learning, facial emotion recognition, computer vision, machine learning, virtual classroom.

1. INTRODUCTION

The word e-learning was pronounced at the CBT Systems seminar in 1999 firstly. E-learning is called "blended learning" that picks all educational and technical items and brings them together in a rich and interactive environment [1]. As soon as the idea of e-learning was explained, expectations of converting the formal education to e-learning are increased. However, in the early 2000s, web technologies could not meet these high expectations in the sense of the e-learning system and course materials (Figure 1) [2]. Here, the preparation phase of the educational documents was an important problem, because it was a time consuming and expensive process. The contents of the e-learning course were not impressive and beneficial enough for keeping the students in the virtual environment for a long time [3]. In the last fifteen years, e-learning activities and web-based training documents have been getting common around the world with the increase of personal computers and the Internet usage.
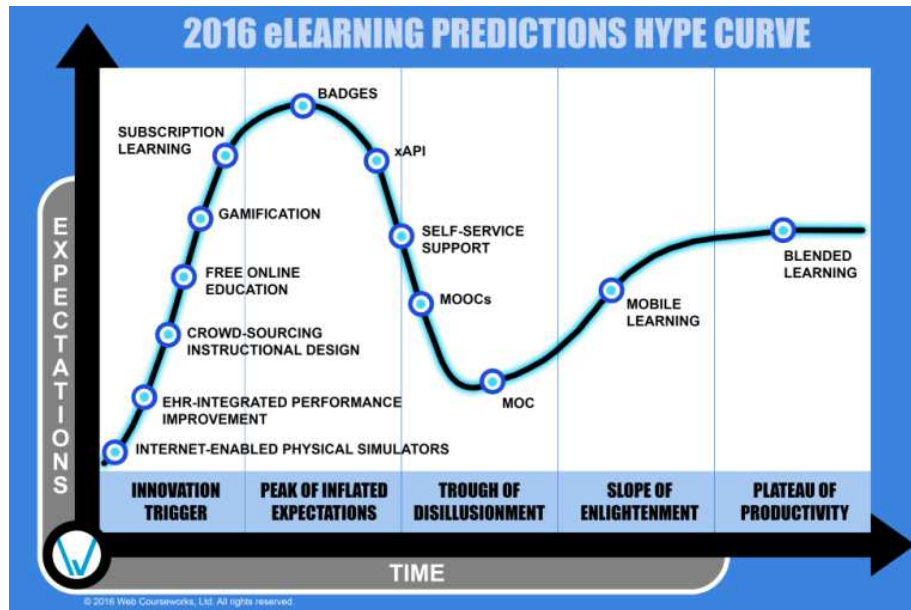
Figure 1. E-learning hype curve [4]

E-learning enables to access educational materials easily at any time and from anywhere. In addition, with the personalized learning technologies, the productivity in education towards the student groups in the heterogeneous structure is also increasing. Today, e-learning is not limited to computer-aided systems. It is also accessible from mobile devices. [5]. Thus, students can take virtual courses whether from their personal computer or mobile devices, at home or in a cafe. These elasticities allow the student to feel comfortable. This situation has led to increasing efficiency and facilitating the learning process.

Recent researches show that emotional states and motivation of students directly or indirectly influence the learning process [6]. Isen [7] argues that a positive mood not only makes us feel good, but it also leads to more accurate decision-making, greater creativity, and flexibility toward problem-solving. According to Isen, the human brain is not completely cognitive information. The human brain is evaluated as a system that integrates emotional and cognitive functions inseparably. Individuals need to be physically, mentally, emotionally and socially ready to be able to comprehend the information easily [8]. Feidakis [9] has presented a review of emotional awareness in the context of e-learning.

The problem statement. One of the biggest problems of e-learning systems is to maintain the motivation of the students in the virtual classrooms. This situation is easier in formal classrooms because the educator is in face-to-face contact with the students in the same environment [10]. In this aspect, it is much easier to observe their emotional states and motivations.

Analysis of recent studies and publications. Several face recognition studies have been utilized to evaluate the emotional state of participants in e-learning environments. Ashwin et al. [11] proposed multi-user face detection based e-learning system using SVM with different data sets (LFW, FDDB, and YFD). Al-Alwani et al. [12] performed mood extraction using a neural network, which classifies facial expressions according to facial features in order to improve students' learning tendencies in e-learning systems. Krithika et al. [13] developed a SERS using the local binary pattern (LBP) for facial recognition and detection. It finds exciting, disturbing and moving patterns and reveals meaningful information to understand the mood of the student studying in an e-learning environment. Recently, real-time emotion recognition studies have been encountered in e-learning. For instance, Magdin et al. [14] have designed software, which evaluates the emotional state of

the user using a webcam. The system works in real time using neural networks. An overall accuracy of their software based on the requested emotions and the recognized emotions is 78%. As an adaptive e-learning study that appeals to a different student population, an emotion recognition system was proposed for students with high-functioning autism [15].

The purpose of the study is to develop a new Facial Emotion Recognition System (FERS), which recognize the emotional states and motivation levels of students in video-conference type e-learning systems. In order to create a more interactive educational environment, this system transfers the emotional states of the students to the educator instantaneously. Our study is supportive of the studies that make possible to observe the motivation level of both the individual and the virtual classroom in the e-learning systems.

2. PROPOSED SYSTEM ARCHITECTURE

In this section, the hardware infrastructure and the software development stages of the proposed system are described.

2.1. Hardware infrastructure

This work is suitable for e-learning applications through video-conference. For this reason, experimental work has been carried out through the Skype program. Once the virtual classroom in the Skype program is launched, the screenshots of the educator's computer involving frontal face images of participants are recorded in ".png" or ".jpg" format instantaneously. Each recorded image is read instantly and the image processing process begins. In this process, GPU is important in terms of the proposed system performance [16]. We used a computer with Nvidia GeForce GTX 860M graphics card. The proposed FERS was developed on Python platform. Python has various powerful software packages for both image processing (OpenCV, dlib, etc.) and machine learning (scikit-learn, numpy, pandas, etc.) processes.

2.2. Software development stages

2.2.1. Face Detection
Face detection is the first step in the FERS. Face detection can be realized on images and on real time video. The Viola-Jones algorithm [17] is preferred in facial detection in terms of real-time operation and fast detection. There are six categorized emotions (Happiness, Fear, Sadness, Anger, Surprise, and Disgust) that are universally accepted in FER [18].

2.2.2. Obtaining the Facial Landmarks and Attributes
Humans' faces give an idea about their emotional state. As the emotional moods change, there will also be changes in the sections of the face. The edges of these sections are determined and followed by facial landmarks. In this study, a facial landmarks detection algorithm proposed by Sagonas et al. [19] was used. Sagonas performed this algorithm on different face data sets [20] – [23] and aimed to obtain a standard in determining facial landmarks.
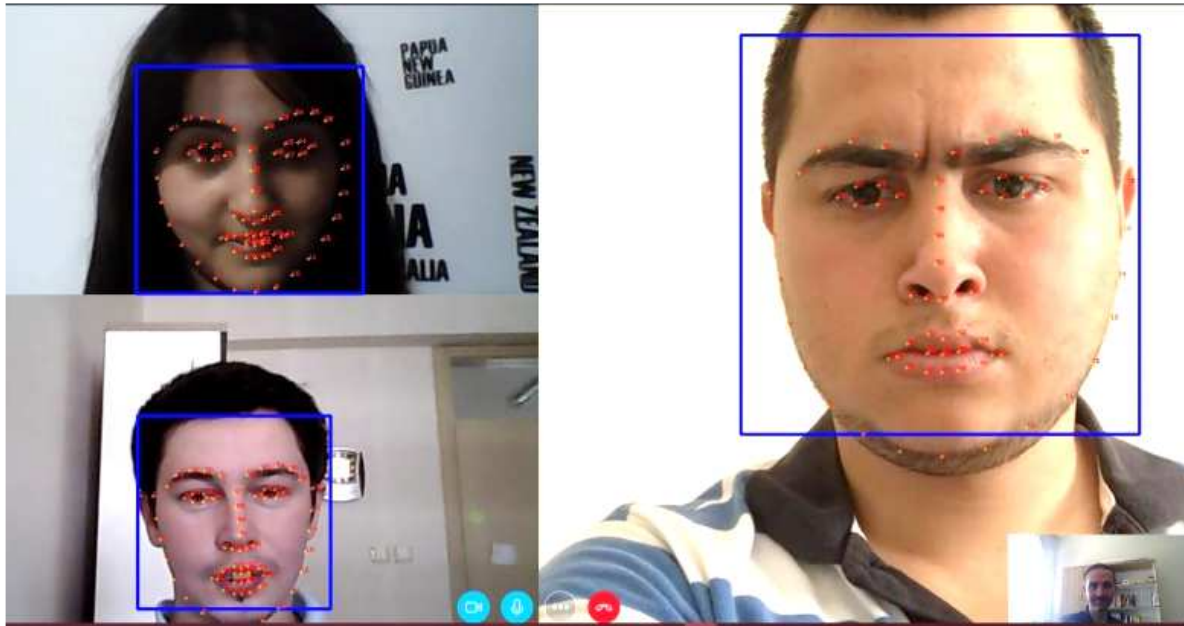
Figure 2. Displaying facial landmarks on multi-participants' faces in Skype video conference

This algorithm yields 68 facial landmarks on the detected face indicating the edges of sections of the face (Figure 2). In this study, the new distinctive attributes obtained using these facial landmarks (Figure 3).

Attribute Description	Ratio	Name
The ratio of face width to mouth width	d1/d2	a1
The ratio of mouth width to mouth openness	d2/d3	a2
The ratio of eye width to eye openness	d4/d5	a3
The ratio of eyebrow length to distance between two eyebrows	d6/d7	a4
The ratio of eyebrow length to distance between eye and eyebrow	d6/d8	a5
The ratio of the length of the nose to the distance between the nose and the upper lip	d9/d10	a6
The ratio of the distance between the extreme points of the eyebrows to the length of the nose	d11/d9	a7
The ratio of distance between two irises to mouth width	d12/d2	a8

Figure 3. Distinctive attributes extracted from facial expressions

2.2.3. Training Data Set

The training data set has been described in detail in our previous study [24]. The data set has the size of 11680x10 and consists of attributes obtained from the students' facial expressions. A section of this data set is given in Figure 4.

a1	a2	a3	a4	a5	a6	a7	a8	gender	emotional_state
2,2565	3,1814	3,4315	1,9529	2,6196	2,9567	1,1655	3,1065	1	Happiness
2,2219	3,3773	3,4315	2,0676	2,6196	2,9567	1,1665	3,0018	0	Happiness
2,3032	2,4596	2,9872	1,9702	2,6798	2,208	1,2244	3,0376	0	Surprise
2,2494	2,6616	2,7613	2,5886	1,5455	1,7676	1,2773	3,2875	1	Fear
2,219	2,5633	2,754	2,5903	1,6172	1,7676	1,2489	3,2867	1	Fear
1,9155	2,1664	3,2888	3,7272	1,4998	2,4759	1,0962	3,0643	1	Disgust
1,9576	2,2183	3,2888	3,2062	1,6106	2,5466	1,1184	3,0643	1	Disgust
2,3079	4,9497	5,6874	3,1226	2,8932	3,6409	1,312	2,698	0	Anger
2,457	2,1148	5,6874	2,6519	2,9019	3,4933	1,4337	2,8067	0	Anger
2,4334	2,9872	5,3413	1,7897	2,8685	2,4703	1,2279	3,3539	0	Sadness
2,5714	2,1523	3,4	2,3637	2,4434	2,6138	1,4288	3,2292	0	Neutral
2,526	2,1926	3,0922	2,3403	2,4086	2,4641	1,4218	3,2551	0	Neutral
2,2589	3,1786	3,4315	2,0155	2,6542	2,9995	1,1665	3,1055	0	Happiness

Figure 4. A section of training data set

2.2.4. Emotion Recognition

Each instance in training data set contains nine attributes (Figure 4). Eight of them are facial attributes and the other attribute is gender information represented numerically as "0" and "1" for "male" and "female. These instances are used in both training and testing process of emotional class prediction. The goal of this phase is to predict the emotional class of the each instance obtained after face detection and feature extraction phases. The prediction performances have been compared by using four different types of classification methods in order to determine emotional classes.

Support Vector Machine (SVM) has proven that core-based non-linear classifiers give good results in practical problems in machine learning and data mining [25]. SVM determines the boundary surface between the two classes, maximizes the distance between the training set and the nearest points to the discrimination surface.

The k-nearest neighbor (kNN) classification method is based on the similarity between the sample to be classified and the samples in the class. Training samples are defined by n-dimensional numerical qualities. Each sample points to n-dimensional space [26]. In this way, all training samples are stored in n-dimensional sample space. These k-education examples are the closest neighbors to the unknown sample A. Proximity is found by calculating Euclidean Distance as given in equation (1).

$$D(X,Y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$
(1)

Random Forest (RF) is a classifier made by Breiman and Cutler, consisting of multiple trees produced from random samples of states or variables. In this method, the tree starts with a single node that all the data has formed, and if all of the samples belong to the same class, the node ends up as leaf and the class label is given. If the samples are not included in the same class, the samples are chosen to best fit the classes [27].

Classification & Regression Trees (CART) algorithm has a structure that forms binary decision trees by dividing the corresponding cluster into two sub-clusters that are more homogeneous at each step. The best independent variable is selected using the impurity and

the variance in the variation measure. Here, the goal is to produce the most homogeneous data subgroups possible for the target variable [28].

2.2.5. Determining the Motivation Level

In this section, instant and weighted emotional states of students in a virtual class have been determined. From the beginning of the video conference, the emotional state of the students transferred to the educator. The emotional state of the majority of the learners indicates the emotional mood of the virtual class. Herein, some of the six universal emotions are considered as positive and some of them negative. For instance; happiness and surprise are considered as positive emotions; sadness, disgust, fear, anger are negative emotions. The instantaneous emotional states of the students are displayed with different emoji (Figure 5). The educator is also able to see the general motivation level of the virtual classroom with the chart of motivation level.
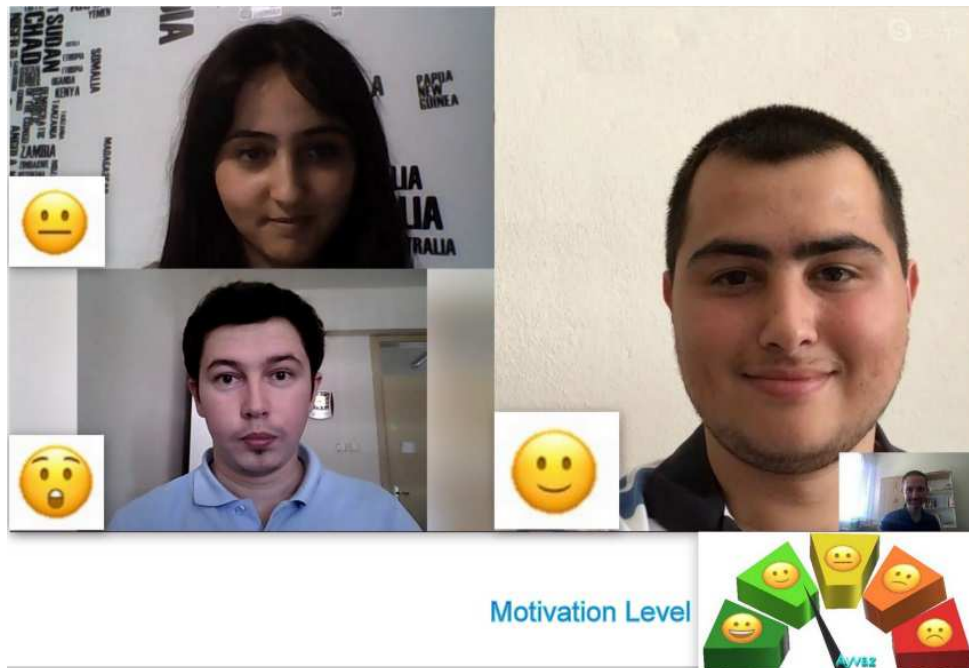

Figure 5. General motivation level of the virtual-classroom

3. THE RESULTS AND DISCUSSION

A two-stage test process was performed to determine the best of the prediction algorithms given in Section 2.2.4. In the first stage, the entire data set was statically classified. The results obtained in the first stage are given in Table 1.

Table 1

The classification results of machine learning algorithms

Tested Algorithms	Accuracy (%)
kNN	96.38
Random Forest	93.36
CART	89.11
SVM	97.15

The accuracy rates were sorted as CART, RF, kNN and SVM algorithms in ascending order. The reason for the higher performance of SVM and kNN algorithms is that the features are extracted from the Euclidean distance between two facial points. It is known that KNN and SVM algorithms are very successful in classifying data sets with such spatial values.

In the second stage, the emotional class of the instances is predicted with algorithms mentioned at Table 1. The emotional classes are happiness, surprise, sadness, anger, disgust, fear and neutral. They are enumerated 0 to 6 respectively. The instances whose actual emotional classes are already known were randomly selected from entire data set. A total of 1650 instances were predicted separately and the prediction results of the algorithms were recorded. Figure 6 presents the actual emotional class and prediction of the algorithms.

Figure 6. Predicted emotional classes of instances

As shown in Figure 6, different prediction results can be obtained when distinct instances are tested with different algorithms. Prediction accuracy was obtained by calculating how many pieces of accurate predictions were made for each algorithm. Table 2 gives the prediction accuracies of these algorithms.

Table 2

The prediction results of the algorithms

	kNN	RF	CART	SVM
# of Actual Class/ # of Correctly Classified	1650 / 1574	1650 / 1534	1650 / 1585	1650 / 1621
Accuracy (%)	95.39	92.97	96.06	98.24

With the SVM algorithm, only 29 of 1650 different data samples were incorrectly estimated and a high accuracy value of 98.24% was obtained. The SVM algorithm with the RBF kernel function, which gives the best performance as a result of these two steps, has been preferred for instant emotion recognition process.

4. CONCLUSIONS AND PROSPECTS FOR FURTHER RESEARCH

In this study, an FERS was developed that sends emotional states of the students in the virtual classroom to the educator. This supportive system makes possible to observe the motivation level of both the individual and the whole virtual classroom in the e-learning systems. In the experimental phase, SVM was chosen for emotional state classification because it gave the best prediction accuracy. In conclusion, evaluating the emotion of a learner in the virtual classroom can help in enhancing the learning experience. Skype is preferred as a test platform because it is free of charge. Moreover, it allows multi-user video conferencing. Nevertheless, this work can also be adapted for other video conferencing software in the future. As the camera becomes standard equipment for personal computers, the usability of this system increases.

ACKNOWLEDGEMENT

REFERENCES

[1]   A. Nagy, *"The* Impact of E-Learning", in: P.A. Bruck, A. Buchholz, Z. Karssen, A. Zerfass, (Eds). *E-Content: Technologies and Perspectives for the European Market*. Berlin: Springer-Verlag, pp. 79–96, 2005. (in English)

[2]   M. D. Lytras and A. Pouloudi, "E-Learning: Just a Waste Of Time," in *Seventh Americas Conference on Information Systems*, America, 2001. (in English)

[3]   S. Carliner and P. Shank, The e-Learning Handbook: Past Promises, Present Challenges, John Wiley & Sons, 2008. (in English)

[4]   A. Hicken, "*2016 Elearning Hype Curve Predictions*", 23.12.2015. [Online]. Available: http://www.webcourseworks.com/2016-elearning-hype-curve-predictions/. [Accessed 06.07.2017]. (in English)

[5]   K. I. Benta, C. Marcel and F. M. Vaida, "A multimodal affective monitoring tool for mobile learning," *Institute of electrical and electronics engineers computer society*, pp. 34-38, 2015. (in English)

[6]   M. Lewis, H. Jones and F. Barrett*, Handbook* Of Emotions, New York, 2000. (in English)

[7]   A. M. Isen, "Positive Affect Influences Decision Making," in Handbook of emotions, M. Lewis and J. Haviland, Eds., Guilford, New York, The Guilford Press, p. 720, 2000. (in English)

[8]   N. Razon, "*Okul Başarisini Etkileyen Faktörler*", [Online]. Available: http://www.ekipnormarazon.com/makale-detay/okul-basarisini-etkileyen-faktorler [Accessed 06.07.2017]. (in Turkish)

[9]   M. Feidakis, "A Review of Emotion-Aware Systems for e-Learning in Virtual Environments," in Formative Assessment, Learning Data Analytics and Gamification, S. Caballé and R. Clarisó, Eds., *Academic Press*, Boston, p.217-242, 2016. (in English)

[10] W. K. Horton, Leading e-learning, American Society for Training and Development, 2001. (in English)

[11] T. S. Ashwin, J. Jose, G. Raghu and G. R. Reddy, "An E-learning System With Multifacial Emotion Recognition Using Supervised Machine Learning," in IEEE Seventh International Conference on Technology for Education, 2015. (in English)

[12] A. Al-Awni, "Mood Extraction Using Facial Features to ImproveLearning Curves *of Students in E-Learning Systems," International* Journal of Advanced Computer Science and Applications, vol. 7, no. 11, pp. 444-453, 2016. (in English)

[13] L. B. Krithika and G. G. Lakshmi Priyya, "Student Emotion Recognition System (SERS) for e-learning," *Procedia Computer Science,* p. 767 – 776, 2016. (in English)

[14] M. Magdin, M. Turcani and L. Hudec, "Evaluating the Emotional State of a User Using a Webcam," *Special Issue on Artificial Intelligence Underpinning*, vol. 4, no. 1, pp. 61-68, 2016. (in English)

[15] C. H. Chu, W. J. Tsai, M. J. Liao and Y. M. Chen, "Facial emotion recognition with transition detection for students with high-functioning autism in adaptive e-learning," Soft Computing, pp. 1-27, 2017. (in English)

[16] A. W. Wai, S. M. Tahir and Y. C. Chang, "GPU Acceleration of Real Time Viola-Jones Face Detection," in IEEE International Conference on Control System, Computing and Engineering, 2015. (in English)

[17] P. A. Viola and M. J. Jones, "Rapid object detection using a boosted cascade," IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR, pp. 511-518, 2001. (in English)

[18] P. Ekman and D. Keltner, "Universal Facial Expression of Emotion:An Old Controversy and New Findings," *Nonverbal Behav*,, vol. 21, no. 1, pp. 3-21, 1997. (in English)

[19] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou and M. Pantic, "300 Faces in-the-Wild Challenge: The first facial landmark localization Challenge," in The IEEE International Conference on Computer Vision, 2013. (in English)

[20] P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman and N. Kumar, "Localizing parts of faces using a consensus of exemplars," in The IEEE International Conference on Computer Vision and Pattern Recognation (CVPR), 2011. (in English)

[21] V. Le, J. Brandt, L. Bourdev and T. S. Huang, "Interactive Facial Feature Localization," in European Conference on Computer Vision, 2012. (in English)

[22] X. Zhu and D. Ramanan, "Face detection, pose estimation, and landmark localization in the wild," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012. (in English)

[23] M. Kostinger, P. Wohlhart, P. M. Roth and H. Bischof, "Annotated facial landmarks in the wild: A large-scale, realworld database for facial landmark localization," in IEEE International Conference on ComputerVision Workshops (ICCV Workshops), 2011. (in English)

[24] U. Ayvaz and H. Gürüler, "The Detection of Emotional Expression towards Computer Users," *International Journal of Informatics Technologies*, vol. 10, no. 2, pp. 231-239, 2017. (in Turkish)

[25] V. N. Vapnik, "An overview of statistical learning theory," *IEEE transactions on neural networks* , vol. 10, no. 5, pp. 988-999, 1999. (in English)

[26] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang and H. Motoda, "Top 10 algorithms in data mining," *Knowledge and information systems*, vol. 14, no. 1, pp. 1-37, 2008. (in English)

[27] A. Liaw and M. Wiener, "Classification and regression by random Forest," *R news*, vol. 2, no. 3, pp. 18-22, 2002. (in English)

[28] D. Steinberg and P. Colla, "CART: classification and regression trees," in The Top Ten Algorithms in Data Mining, X. Wu and V. Kumar, Eds., Chapman & Hall/CRC, 2009, p. 179. (in English)

ВИКОРИСТАННЯ ВИРАЗІВ ЕМОЦІЙ НА ОБЛИЧЧІ ЛЮДИНИ В СИСТЕМАХ ЕЛЕКТРОННОГО НАВЧАННЯ

Угур Айваз
здобувач ступеня магістра (Інформаційні системи інженерії), молодший науковий співробітник,
кафедра інформаційних систем інженерії
Університет Muğla Sıtkı Koçman, Мугла, Туреччина
Ugur.ayvaz@mu.edu.tr

Хюсеїн Гюрюлер
PhD з Електронної і Комп'ютерної освіти, доцент кафедри інформаційних систем інженерії
Університет Muğla Sıtkı Koçman, Мугла, Туреччина
ORCID ID 0000-0003-1855-1882
Hguruler@mu.edu.tr

Мехмет Осман Деврім
здобувач ступеня магістра (Інформаційні системи інженерії), студент,
кафедра інформаційних систем інженерії
Університет Muğla Sıtkı Koçman, Мугла, Туреччина
Mehmetosmandevrim@posta.mu.edu.tr

Анотація. Оскільки використання персонального комп'ютера та пропускна здатність Інтернету зростають, широко поширюються системи електронного навчання. Незважаючи на те, що електронне навчання має певні переваги щодо доступності інформації, гнучкості

часу та місця порівняно з формальним навчанням, воно не забезпечує достатньої особистої взаємодії між викладачем та учнями. У цьому дослідженні ми пропонуємо гібридну інформаційну систему, яка об'єднує комп'ютерне бачення та технології машинного навчання для візуальних та інтерактивних систем електронного навчання. Запропонована інформаційна система виявляє емоційний стан учнів і дає зворотний зв'язок викладачу щодо їх емоційного стану у даний момент на основі виразів обличчя. Таким чином, викладач має можливість усвідомлювати загальний емоційний стан віртуального класу, і система створить формальне навчання, як інтерактивне середовище. У запропонованій системі застосовано кілька класифікаційних алгоритмів для вивчення миттєвого емоційного стану, а найбільший рівень точності було досягнуто за допомогою алгоритмів kNN та SVM.

Ключові слова: електронне навчання; розпізнавання емоцій особи; комп'ютерне бачення; машинне навчання; віртуальна класна кімната.

ИСПОЛЬЗОВАНИЕ ВЫРАЖЕНИЯ ЭМОЦИЙ НА ЛИЦЕ ЧЕЛОВЕКА В СИСТЕМЕ ЭЛЕКТРОННОГО ОБУЧЕНИЯ

Угур Айваз
соискатель степени магистра (Информационные системы инженерии), младший научный сотрудник, кафедра информационных систем инженерии
Университет Muğla Sıtkı Koçman, Мугла, Турция
Ugur.ayvaz@mu.edu.tr

Хюсеин Гюрюлер
PhD по Электронной и Компьютерной образования, доцент кафедры информационных систем инженерии
Университет Muğla Sıtkı Koçman, Мугла, Турция
ORCID ID 0000-0003-1855-1882
Hguruler@mu.edu.tr

Мехмет Осман Девра
соискатель степени магистра (Информационные системы инженерии), студент,
кафедра информационных систем инженерии
Университет Muğla Sıtkı Koçman, Мугла, Турция
Mehmetosmandevrim@posta.mu.edu.tr

Аннотация. Поскольку использование персонального компьютера и пропускная способность Интернета растут, широко распространяются системы электронного обучения. Несмотря на то, что электронное обучение имеет определенные преимущества по доступности информации, гибкости времени и места по сравнению с формальным обучением, оно не обеспечивает в достаточной мере личное взаимодействие между преподавателем и учениками. В этом исследовании мы предлагаем гибридную информационную систему, которая объединяет компьютерное видение и технологии машинного обучения для визуальных и интерактивных систем электронного обучения. Предложенная информационная система обнаруживает эмоциональное состояние учеников и дает обратную связь преподавателю их эмоционального состояния в данный момент на основе выражений лица. Таким образом, преподаватель имеет возможность осознавать общее эмоциональное состояние виртуального класса, и система создаст формальное обучение, как интерактивную среду. В предлагаемой системе применено несколько классификационных алгоритмов для изучения мгновенного эмоционального состояния, а наибольший уровень точности был достигнут с помощью алгоритмов kNN и SVM.

Ключевые слова: электронное обучение; распознавание эмоций лица; компьютерное видение; машинное обучение; виртуальная классная комната.