An AI Application in E-Learning

Bozhan I. Orozov[1], Daniela A. Orozova[2,3]

[1] *Abertay University, Bell Street, Dundee, Scotland*
[2] *Burgas Free University, Department of Computer Science, Burgas, Bulgaria*
[3] *Institute of Information and Communication Technologies at BAS, Sofia, Bulgaria*

Abstract

The report presents a study in the field of machine learning with visual data. The main goal is to design and develop an application for recognizing human emotions by facial image, through the facial coding system (FACS) of Ekman and Friesen. The ML.Net framework implements a multi-class classification algorithm based on the ML agent of the Supervised Learning type. After training the model, he can recognize the emotional state of a submitted image (photo) of a person. The developed application will be integrated into an e-Learning environment. Changes in the emotional states of learners are sought as they cope with the cognitive tasks in the learning process.

Keywords

Artificial Intelligence, Machine Learning Algorithms, E-learning systems.

1 Introduction

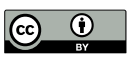
For more than a century, scientists have wondered whether all people experience the same range of basic emotions and, if so, whether they express them in the same way. In the 1970s, this was the main question that Charles Darwin explored in "On the Expression of Emotions in Man and Animals." In the 1960s, Paul Eckman, a respected psychologist at the University of California, San Francisco, came up with a methodology for researching this issue. Psychologist Carlos Crivelli (Madrid, 2011) explains that emotions trigger natural physiological reactions, and the limited number of ways in which facial muscles can move creates the basic set of expressions. The facial expression of emotions is determined by the muscles involved and the gestures that characterize it. There are certain distinctive patterns of emotional reactions that are summarized and shared by most people. They are considered basic emotions: happiness (joy), sadness, anger, surprise, fear and disgust.

Based on this basic set of expressions, studies by Ekman and Friesen have been published [5]. These authors developed the Facial Coding System (FACS) to categorize facial expressions. Through this system, facial movements are described by a set of muscle-based units of action (AU). These researchers analyze facial expressions by detecting gestures and measuring the number of facial movements. Other similar coding systems are EMFACS, MAX, AFFEX and CANDIDE-3. Ekman et al. also present a database called the FACS Affect Interpretation Dictionary (FACSAID), which allows the translation of FACS assessments related to emotions [10].

A number of software products have been created that can read and distinguish emotions, as well as nuances of smiles or grimaces. Here are some examples.

Amazon Web Services' cloud service (AWS) for analyzing faces and emotions captures different expressions and predicts people's emotions from photos. The service uses artificial intelligence to "learn" from the many data it processes. This tool includes gender recognition, emotions and approximate age (*https://aws.amazon.com*).

Education and Research in the Information Society, September 27–28, 2021, Plovdiv, Bulgaria
EMAIL: bozhanwork@gmail.com; orozova@bfu.bg



Interesting applications of facial expression recognition technologies are in the field of healthcare [4]. *CERT* is a software developed by researchers at the University of San Diego related to the recognition of pain expressions. In addition, researchers have been able to develop software capable of detecting whether the pain expressed by humans is genuine or false. A study [3] compared the effectiveness of humans and computers in expressing real or simulated pain.

Researchers at Carnegie Mellon University have developed a multimodal algorithm for analyzing facial expressions on 68 individual points, including eyebrows, corners of eyes, mouth and nose. The system, called MultiSense, can recognize the position of a person's head, the direction of his gaze, the orientation of the body, and even distinguish a smile caused by real emotions from one that appears due to politeness. Louis-Philippe Morensi and his colleagues developed the algorithm to automatically track the relationships between facial expressions and the emotional state of depressed people. The results turned out to be unexpected. For example, people who are depressed and those who are not depressed smile with the same frequency, but their smiles are different.

Computer vision is one of the main areas of research in the field of AI, due to its widespread use in the modern world. The developed project provokes the authors' desire for more experiments and deepening the understanding of the architectures of these systems because they are widely applicable in the field of E-learning.

The report presents a software application for detecting an emotional state by extracting information from facial expressions. An approach is applied to detect a person's emotion, based on facial marking and applying an algorithm for analyzing the expression on 68 separate points. The report is structured as: an introduction to the problem related to facial expression recognition is made and the main methods and results of conducted research in the field are briefly presented. The following is a description of the methodology of the experiment and the development of a software application for detecting an emotional state based on facial expression. The main challenges and applications in the field of electronic systems and learning tools are then discussed.

2 Methodology

The research focuses on building a methodology for detecting meaningful pre-defined changes in participants' emotional states while dealing with cognitive tasks in the context of learning.

The first step is to categorize the given photos with images of faces in separate subdirectories. There are seven of them, each of which is responsible for a specific emotion and contains about fifty sample images related to the respective emotion. The next step is creating software using a third-party library (DLib), through which the processing of the respective photos. Dlib is open source and allows free use in various applications. It is a library of machine learning algorithms and general-purpose software development tools and is used in many fields – industry, training, robotics, mobile phones, large high-performance computing environments and many more. Some of the main features for the library are: *High-Quality Portable Code*, *Machine Learning Algorithms*, *Numerical Algorithms*, *Graphical Model Inference Algorithms*, *Image Processing*, *Threading*, *Networking*, *Graphical User Interfaces*, *Data Compression and Integrity Algorithms*, *Testing*, *General Utilities* [6].

By using the *spape_predictor_68_face_landmarks.dat* file, the faces in the images are marked with corresponding dots. Based on the set points, in the next step, a number of mathematical calculations are performed, with the help of which 6 normalized vectors of each person are determined. Each facial feature is calculated as the data is normalized and the distances between key landmarks are collected. A key point is the normalization of the vectors relative to the smallest vector. The received information is saved in a document in *.csv* format. There are many approaches to constructing a vector characteristic. The applied approach is appropriate for the model training functions used in the next step.

Two classes are created for the purposes of the project. The first is called "*Face*" and is used to load the information that is stored in the facial *features.csv* file. The other class "*FacePredicted*" is used to predict a new image based on the created model.

Using the Interface *IDataView*, the data from facial features.csv is loaded and saved. The received data is saved as objects of class "*Face*". A pipeline is created, in which the main elements of the built model are set. This pipeline together with the *IDataView* Interface are used to train the ML agent.

A Microsoft framework called "ML.Net" is used to create the ML agent. It is used for the integration of machine learning models in built user applications, such as: mood analysis, price forecasting, product recommendations, sales forecasting, image classification, object detection and more. The advantage of the created tools is their ability to be used by developers without extensive experience in the field of Data Science.

For the purposes of the project, the *MulticlassClassification* method is used, which implements a multi-class classification algorithm. The method uses the submitted data to train the agent. After completing the agent training, an object of class *Predictor* is created using the *CreatePredictionEngine* function. This is a class that is used to predict the emotion of submitted new facial features, on an already trained model.

This is followed by loading the test data file, which is different from the data used for training. This file is used to calculate test metrics for the accuracy of predictions made.

3 Results

The project implements a classification algorithm based on ML agent of type *Supervised Learning*. The model is trained with data for which the type is indicated (with a label) and allows predictions to be made for future data.

After training the model, he can recognize the emotional state of a submitted image of a person. As a result of the work, the program returns a prediction about the emotional state. Based on the test data, the program derives a series of estimates for the accuracy of the prediction. The analysis of the results of the metrics for assessing accuracy shows that the highest accuracy of prediction is achieved in the emotional states *Joy*, *Surprise* and *Sadness*. The *Anger* state is often confused by the algorithm with *Sadness*, *Joy* or *Neutral*. The algorithm can be improved by learning with a much larger amount of training data than the different types of emotional states. But a balanced set of data must be used because classification algorithms can be affected by a shortage of data of a certain type, which can lead to erroneous predictions. But access to such data is associated with a number of problems because it is personal information. On the other hand, the incoming data can be in a different format, which creates the need to process different types of images.

An alternative approach is to apply an algorithm of the *Unsupervised Learning* type, which examines the structure of the data without instructions – the training data are not labelled. ML agents based on *Clustering algorithms* or *Deep Learning* can be used. Different classifiers are applied in the experiments and the results obtained are experimented with to increase the accuracy of different groups and categories of data [9]. In addition, different models can be applied sequentially to improve the quality of the forecast.

4 Application in E-Learning

In the study, the authors seek to integrate the created application in an e-learning environment by tracking the learning process and the learner's emotions. The aim is to use a multimodal approach based on data extracted from several sources (physiological data on keyboard and mouse interactions; subjective information provided by learners about their emotions, photos of facial expressions, etc.) to detect emotions and modelling of adapted system behaviour to provide emotional personalized support to learners. Of particular importance is the monitoring of the emotional state of learners in game-based learning [11, 12].

To provide personalized support in an educational context, a methodology for marking facial expressions and tracking changes in learners' emotional states while coping with cognitive tasks in the learning process is described. The ultimate goal is to combine the obtained data on the emotional state with additional information collected during experimental training sessions from various sources such as qualitative, physiological and behavioural information. This data is intended for training algorithms that serve to automatically identify changes in learners' emotions when dealing with cognitive tasks. On this basis, emotional personalized support should be provided.

It should also be taken into account that not everything that appears on the face necessarily shows an inner experience. A facial expression depends on the individual's ability to control his emotions [2]. Thus, the same emotion can cause different facial expressions. In addition, a number of factors such as: the distance between the camera and the face, the parameters of the camera, slight tilts of the head, barely hinted expressions and many other seemingly insignificant differences are reflected in the reading of the characteristics and can affect the result.

In the experiment, the learners in the electronic environment used webcams that capture their faces, and the camera is located at a distance of about 75 cm from it. In addition, participants provide answers to several questions related to the emotions experienced, difficulties encountered, and strategies used to address them. These data become part of the user model, which takes into account the emotions of the learner.

In the next stages of the research, data that can be collected in files will be included: a file with all the keystrokes performed by the participant (keystroke and release) and another file with all the actions performed with the mouse (keystrokes and mouse movements), as well as the time of each interaction. Of course, a number of other sensors and sensors can be included, and the data can be stored in a file with all recorded physiological signals and timestamps added during the experiment.

5 Conclusion

The successful implementation of the project poses new challenges for the implementation of various applications related to recognizing the emotional state of the user and adapting the system to him. Thus, in the case of a training system, the difficulty of the proposed task can be corrected if the user's expression shows that it is too stressful for him. It can be used to maintain a lie detector, as a security check, or as a diagnostic tool.

Facial expression is a powerful engine capable of regulating behaviour and opening up new scenarios in human-machine relationships. For this reason, the possibility of automatic detection and correct understanding of facial expressions has been of great interest for years and has become the focus of applied research in various fields. An interesting trend in more complex applications is the use of different emotion recognition techniques, which are combined to achieve a more accurate result.

Acknowledgement

The work is funded by the SRF of the Burgas Free University as a part of the project “Data Science in the Learning Space for a Blue Career”.

References

[1] B T. Dalgleish & T. Power (Eds.), The Handbook of Cognition and Emotion, Sussex, У. K. John Wiley & Sons, Ltd, 1999, ISBN 0-471-97836-1.

[2] O’Keefe, B. J., & Lambert, B. L. (1995). Managing the Flow of Ideas: A Local Management Approach to Message Design. B B. R. Burleson (Ed.), pp. 54-82. https://doi.org/10.1080/23808985.1995.11678907.

[3] M. S. Bartlett, G. C. Littlewort, M. G. Frank, K. Lee, Automatic Decoding of Facial Movements Reveals Deceptive Pain Expressions, Current Biology, Volume 24, Issue 7, 2014, pp. 738-74324, https://doi.org/10.1016/j.cub.2014.02.009.

[4] F. Canento, H Silva, A Fred – Computing Paradigms for Mental Health, 2012Sedentary Work Style and Heart Rate, Variability: A Short Term Analysis, B P. Cipresso, M. Hoogendoorn, M. Klein & A. Matic (Eds.), Computing Paradigms for Mental Health. Proceedings of the 2nd International Workshop on Computing Paradigms for Mental Health, ISBN: 978-989-8425-92-8.

[5] P. Ekman, W.V. Friesen, Constants across cultures in the face and emotion. Journal of Personality and Social Psychology, 1971, 17(2), 124–129, https://doi.org/10.1037/h0030377.

[6] DLib library: http://dlib.net/

[7] S.J. Russell, P. Norvig,. Artificial Intelligence: A Modern Approach (3rd ed.), Prentice Hall, 2010, p.649, ISBN 978-0136042594.

[8] E. Charniak, D. McDermott, Introduction to Artificial Intelligence, Addison-Wesley, Reading, Massachusetts, 1985.

[9] E. Charniak, C.K. Riesbeck, D.V. McDermott, J.R. Meehan, Artificial Intelligence Programming (2nd ed.), Lawrence Erlbaum, 1987, ISBN 0-89859-609-2.

[10] M. Hornick, E. Marcade, S. Venkayala, Java data mining: strategy, standard and practice, A practical Guide for Architecture, Design, and Implementation, 2006, ISBN-10:0-12-370452-9.

[11] Gocheva M., N. Kasakliev, E. Somova, An Attempt of Adaptability of the Learning Process and Content in Mobile Math Educational Game, published: AIP Conference Proceedings, e-ISSN: 1551-7616, 10-th International scientific conference Engineering, technologies and systems – TechSys'2021.

[12] Gocheva M., E. Somova, N. Kasakliev, Game-based Approach in Mobile Learning for Primary School, published: AIP Conference Proceedings, e-ISSN: 1551-7616, 10-th International scientific conference Engineering, technologies and systems – TechSys'2021.

[13] Pavlov N., A. Iliev, A. Rahnev, N. Kyurkchiev, Some Software Reliability Models: Approximation and Modeling Aspects; LAP LAMBERT Academic Publishing: Saarbrucken, Germany, 2018.