

Video-Based Automated Emotional Monitoring In Mental Health Care supported by a Generic Patient Data Management System¹

Hayette Hadjar¹, Julian Lange¹, Binh Vu¹, Engel Felix², Mayer Gwendolyn³,

Paul Mc Kevitt⁴, and Matthias Hemmje²

¹ University of Hagen, Faculty of Mathematics and Computer Science, Hagen, Germany
{hayette.hadjar, julian.lange, binh.vu}@fernuni-hagen.de

² Research Institute for Telecommunication and Cooperation, Dortmund, Germany
{fengel, mhemmje}@ftk.de

³ Heidelberg University, Department of Internal Medicine II, General Internal Medicine and
Psychosomatics, Heidelberg, Germany
gwendolyn.mayer@med.uni-heidelberg.de

⁴ Ulster University, Derry/Londonderry, Northern Ireland
p.mckevitt@ulster.ac.uk

Abstract. The detection of emotion and expression from video streaming plays a very important role in the mental health care of a patient. The data obtained from it can be used to support the diagnosis of emotional needs related to depression or other kinds of mental illnesses. These data can provide useful emotion monitoring information for health monitoring systems using automatic calculation of this Affective Computing (AC) information and storing them in patient data management systems. This research has been developed in the context of the SenseCare project, in order to support the treatment of patients with primary or comorbid mental disorders. There are two processes for tracking emotion in video, real-time, and offline facial expression video analysis. Real-time video analysis uses streamed webcam videos as data input. Offline video analysis uses pre-recorded video files as input. We focus in this paper on the real-time video analysis process, and we employ deep learning in web browsers for face detection and recognition using JavaScript.

Keywords: Video Content Analysis, Affective Computing (AC), Emotion Recognition, Facial Expression Analysis, Emotions representation, Emotional Monitoring.

¹ Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0)

1 Introduction and Motivation

Understanding and utilizing psychological knowledge in order to e.g. automatically detect psychological events is one of the key research challenges in Affective Computing (AC), especially in Research and Development (R&D) of software related to automatic emotion detection. Furthermore, ambient assisted living and tele-monitoring health care technologies can facilitate the collection of vital signal data remotely (e.g., ECG, heart, and breath sounds) as well as the collection of software-based automatic assessment and monitoring signals of mental or emotional status [1].

The SenseCare (Sensor Enabled Affective Computing for Enhancing Medical Care) Platform [2] has been developed as a prototypical AC R&D platform providing software services applied to the care of patients with different support needs in the field of mental health care. This technology provides various opportunities for physicians, psychotherapists, clinicians, or other healthcare professionals. Such target user groups can e.g. be enabled to intervene early in the case of a critical mental state that could result in a crisis and thus a worsening of the patients' state of health. Hence, primary care professionals can achieve an improved overview of the emotional well-being of patients through the SenseCare [3] AC R&D platform's software services. SenseCare integrates data streams from multiple sensors and fuses data from these sensor signal streams to provide a global assessment that includes objective levels of understanding emotional expressions, as well as the corresponding well-being, and cognitive state of the patients. Several potential use cases for a system like SenseCare underline the topicality, of which the recent crisis due to COVID-19 is only one: Patients with mental disorders on isolation wards have to stay outside the support system, as e.g. psychiatrists, psychologists, and other clinical staff fear infection [4]. Remote, emotion-sensitive support could have supported these patients better and first solutions in tele-medical intervention and corresponding pathways have been developed in the meantime [5]. Additionally, patients in an online group therapy setting (e.g. due to rural provenience with low density of psychotherapists) can be supported by emotion-sensitive videoconference tools. Recent changes in the accounting system of e-health applications by health insurances will promote this development [6]. Furthermore, patients with complex psychosomatic diseases often suffer from a comorbid depression or anxiety, which leads to a vicious circle of deleterious effects. For example, every fifth patient with heart failure suffers from depression, which may lead to a lack of treatment adherence [7]. Continuous monitoring of these patients by software services of platforms like SenseCare may reduce high health related costs. Finally, elderly patients in ambient assisted living are in need for a continuous monitoring of their emotional state, as sudden changes in the mood can be a risk-marker for dementia [8]. Processing of voluminous data streams from video recordings, on the basis of the recently introduced Information Visualization for Big Data (IVIS4BigData) model [9] elaborates data stream types addressed by our visualization approach.

The so-called Knowledge Management Ecosystem Portal (KM-EP) is the backbone system of the SenseCare platform [10] and is comprised of five subsystems where each of them has several components of its own. The Information Retrieval Subsystem (IRS) of the SenseCare KM-EP indexes AC content and enables users to

search for AC content using keywords, faceted search, and taxonomies. The Learning Management Subsystem (LMS) of the SenseCare KM-EP provides tools for authors and trainers to create AC-related e-learning courses using content in the SenseCare KM-EP. SenseCare KM-EP users can later register in these SenseCare courses to obtain new AC knowledge. The Content and Knowledge Management Subsystem (CKMS) of the SenseCare KM-EP acts as a central repository for AC publications, AC multimedia, AC software, AC R&D dialogs, or AC-related medical records in the SenseCare KM-EP. Producers of AC content can use components in this KM-EP subsystem to import, create, manage, and classify their AC contents. Furthermore, SenseCare KM-EP users can access these AC contents and rate their quality. The User Management Subsystem (UMS) of the SenseCare KM-EP manages all users and groups of the SenseCare KM-EP. Other systems can ask to authenticate SenseCare KM-EP user's identity using OpenID Connect [11] integrated into this SenseCare KM-EP subsystem. The Storage Management Subsystem (SMS) of the SenseCare KM-EP provides storage for files and documents. They can either be stored in a local server or on the cloud for better access speed and availability.

Solutions already exist for the administration of medical data and processes. Incorporated within a specialist internship at the FernUniversität in Hagen were exemplary projects: IndivoHealth [12] and Tolven [13] considered as a solution for electronic patient records and validated with regard to the requirements of SenseCare.

On the practical side, our objective is to develop and implement further new software modules as R&D prototypes and corresponding AC software services that can be integrated with the SenseCare KM-EP. Such R&D results can then be re-used to achieve some directions for future R&D work in this domain. The main contributions of this paper are:

- Implementation of a prototype module that collects patients' facial expressions and corresponding emotion data in real-time, during treatment sessions, or at home for cases of patients with or at risk of a mental disorder. The software categorizes emotional states according to the seven basic emotions described by Paul Ekman (anger, contempt, disgust, enjoyment, fear, sadness, and surprise) [14].
- The prototype employs deep learning in browsers by using JavaScript, and stores results (Date - Time- Detected emotion) in a MongoDB.

The remainder of this paper is organized as follows. Section 2 discusses the state of the art of using sensors in healthcare, existing tools, and Convolutional Neural Networks (CNNs). In section 3 we detail the conceptual design of modeling API, information model and implementation of the solution, section 4 discusses our findings, and finally we conclude in section 5.

2 Selected State of the Art and Related Work

Research into wireless sensor networks and smart environments for remote monitoring for healthcare applications [15] employs wearable micro-machined sensors for providing accurate biomechanical analysis under ambulatory conditions.

In the continuous monitoring of human activities, wearable sensors can e.g. detect abnormal and/or unforeseen situations by monitoring physiological parameters along with other symptoms [16]. There are many software tools that employ methods of machine learning to assist people in the areas of health.

Eq-Radio [17]: Researchers from MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) have developed EQ-Radio, a device that can detect a person's emotions using wireless signals. It transmits an RF signal and analyzes its reflections off a person's body to recognize his emotional state (e.g. happy, sad). The key enabler underlying EQ-Radio is a new algorithm for extracting the individual heartbeats from the wireless signal at an accuracy comparable to on-body ECG monitors. EQ-Radio has three components: a radio for capturing RF reflections, a heartbeat extraction algorithm, and a classification subsystem that maps the learned physiological signals to emotional states [17].

Valossa AI [18] is qualified to recognize sentiments and emotions from facial expressions and speech, either from recorded video content or live feed. Mika Rautiainen, founder and CEO of Valossa says that going through a video of a therapy session takes a whole day from a human being. But AI tells her in a real-time analysis what happens on the patient's face.

FaceReady [19]: from Noldus Information Technology is a facial expression analysis software. It can automatically analyze the expressions *happy, sad, angry, surprised, scared, disgusted, and neutral*. It can also calculate *Action Units, valence, arousal, gaze direction, head orientation, and personal characteristics* such as gender and age.

SHORE@ [20] Software of Fraunhofer IIS enables the quick detection of faces and objects as well as for the analysis of faces in image sequences, videos and single frames. It can estimate gender, age and facial expressions in real time. The software runs on standard Convolutional Neural Networks (CNNs) are a type of deep neural network designed to process multiple data types, but it was initially designed to analyze images [21] CNNs are the most popular neural network model employed in image classification [22], CNNs comprise several layers, such as the Convolutional Layer, Non-Linearity Layer, Rectification Layer, Rectified Linear Units (ReLU), Pooling Layer, Fully Connected Layer, and Dropout Layer.

Existing solutions stream frames from a video stream over a network with OpenCV [23] for the following advantages: (i) firstly, building a security application that requires all frames to be sent to a central hub for additional processing and logging, and (ii) secondly, the client machine may be highly resource-constrained (such as a Raspberry Pi) and lack the necessary computational horsepower required to run computationally expensive algorithms (such as CNNs).

3 Conceptual Architecture of Health Information Subsystems

The conceptual architecture of the Health Information System (HIS) within the SenseCare KM-EP can be characterized by several subsystems which organize and

process information by specifying the type of data processed in each subsystem independently of the others.

Within the SenseCare KM-EP's HIS, the Carna subsystem for data management and information systems can run workflows for processing different types of AC data (offline data and real-time data). Hence, it utilizes a workflow engine and enables implementation of customized workflow action steps by Java code. The system consists of different modules, the most important of which are Carna.dms (Data Management System), Carna.process[emotion detection] (support processes, using the example of Emotion Detection), and Carna.tenantmodules (general tenant-based modules). Each Carna module within the SenseCare KM-EP's HIS implements a REST-API to access its functionality. Fig. 1 shows the most important of the implemented REST interfaces.

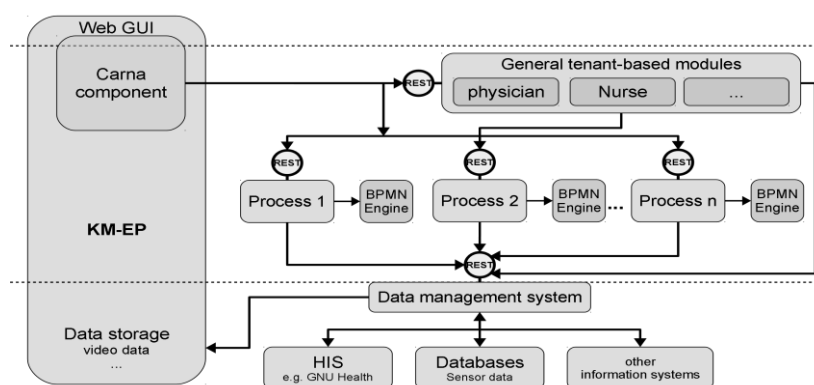


Fig. 1. SenseCare KM-EP HIS's conceptual architecture of the Carna modules supporting the integration of Health Care support processes [24].

In the `carna.dms` module, among other data, process-related data and registered processes are saved. When a workflow process is started for a patient, a new process instance is initialized by the process module and an associated data record is created. When a healthcare task (that is implemented by a process) is finished, a documentation record is appended into the process-instance record table.

To support the conceptual architecture and API modeling for the KM-EP HIS's Emotion Detection system, the activities of the SenseCare Emotional Monitoring Use Case Scenario are:

- 1- The offline video analysis pipeline of the KM-EP HIS's Emotion Detection uses pre-recorded patient videos. These files are stored offline to be pre-trained with CNN models and classifiers in order to detect emotion from facial expression.
- 2- The real-time video analysis of the KM-EP HIS's Emotion Detection uses data input from webcams for detection and recognition of facial expression in real time, this process is the main focus of the remainder of this paper.

4 Prototype Implementation of Emotion Detection

A prototype solution for the SenseCare KM-EP Emotion Detection has been developed with the Model-View-Controller (MVC) architecture paradigm. Hence, the software prototypes's source code is divided into three layers. On the model layer, data storage, integrity, consistency, querying, and access support is allocated. The global neural network models that are exported on this level from faceapi.js are AgeGenderNet, FaceExpressionNet, FaceLandmark68Net, FaceLandmark68TinyNet, FaceRecognitionNet, SsdMobilenetv1, TinyFaceDetector, Mtcnn, and TinyYolov2. On the Controller level, the operations receive, interpret & validate input, create & update are specified and implemented. On the View level, the query & modify models are specified and implemented. In our case the clinical user or the patient interacts with the interface by means of a webcam on the view layer.


The implementation of a corresponding REST API requires these elements:

1. Identify the objects that will be presented as a resource is the very first step in designing a REST API-based application.
2. Create model URIs by designing the resource URIs – focus on the relationship between resources and its sub-resources. These resources URIs are endpoints for RESTful services.
3. Determine Representations: Mostly representations are defined in either XML or JSON format. For example:

emotions: {angry: number, disgusted: number, fearful: number, happy: number, neutral: number, sad: number, surprised: number}

Number in our case is the percentage of the security of the model that the detectives have a particular emotion, each face element has expressions attribute.



Example:  => surprised: 0.990011861078746733256

In the initial prototype implementation the following base technologies are employed:

- Tensorflow.js [25] is a library for machine learning in JavaScript, develop ML models in JavaScript, and use ML directly in the browser or in Node.js.
- Face-api.js [26] is a JavaScript module, built on top of tensorflow.js core, and it implements several CNNs for face detections and recognition, and it has been optimized to work on web and mobile devices.
- Node.js [27] for synchronous or real-time communication in the web application, it is employed to produce highly accurate face recognition and detection.
- MongoDB [28]: is an open source NoSQL database, it is a popular choice for handling big data.
- Mongoose and NodeJSExpress [29] for transactions written in real-time and db connectivity to MongoDB in order to store results of real-time video analysis of facial expression.

The overall distribution and operational deployment of the system within client server distribution architecture is shown in Figure 2 below.

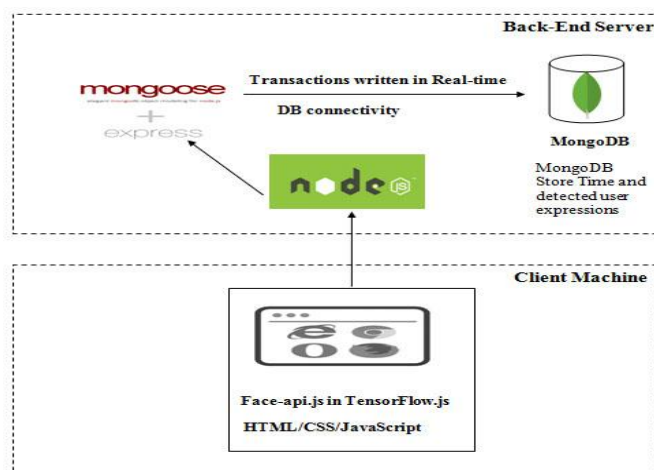


Fig. 2. Client-Server architecture of the initial Emotion Detection prototype.

The system is divided into the frontend and the backend. The frontend in the clients' machine is combined of Face-api.js in TensorFlow.js, HTML/CSS/JavaScript, and browser to display the front-end. The backend server is developed using NodeJS Express, mongoose.Database, and MongoDB. The implementation allows both offline video and stream video to be uploaded and processed. We can input an HTML element like images or offline video using the id of the element, and input stream video with function startVideo() to start webcam in the browser.

5 Discussion of the findings

The conducted experiment showed us that the developed module functionally meets basic requirements, and it is important to implement additional functionality in order to increase research study benefits. In the case of real-time video emotion recognition, the SenseCare KM-EP HIS's Emotion Detection API stores the best emotion detected from the webcam in every 500 milliseconds, this choice of timing can be changed in the API. The stored data has the following format:

AllExpressiondetected: {date + time, label of best expression}

A part of the stored data in MongoDB can be seen in the table below.

Table 1. Data stored in the MongoDB

```
# db.emotionsave.find()
{ "_id" : ObjectId("5f399f182232afa8b58f96ab"), "dateTime" : "2020-8-16 22:2:0",
```

```
"expression" : "neutral", "__v" : 0 }
{ "_id" : ObjectId("5f399f1e2232afa8b58f96ac"), "dateTime" : "2020-8-16 22:2:6",
"expression" : "neutral", "__v" : 0 }
```

A demonstration of face expression Recognition of images from “FACES A database of facial expressions in younger, middle-aged, and older women and men” [30] is shown in Figure 3.

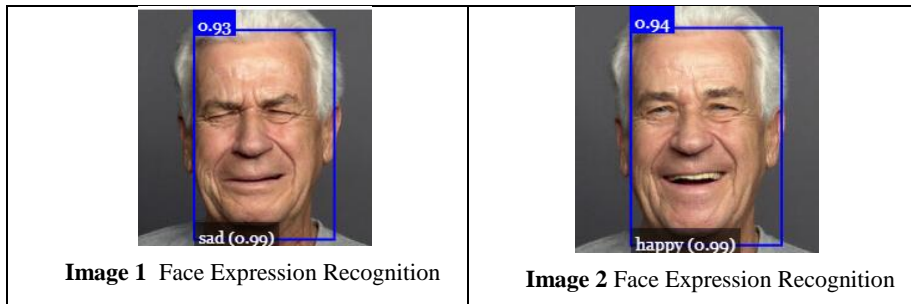


Fig. 3. Demonstration of an Emotion Detection output.

The prototype is under development and in our first observation during a test on machines with different OSs (e.g. Windows, Ubuntu, MacOS), the values of the results of real-time video emotion analysis, and the response time changes according to the capacity and the hardware performance of the web server. Hence, a real-time detection of emotions requires powerful hardware, e.g. Memory of the server must be greater than 6 GB. And high quality images in the input stream are required to identify a face (descriptor). We also observed that SSD Mobilenet V1 neural network gives better accuracy than Tiny Face detector and MTCNN, and the accurate detection of emotions based on facial expressions decreases when the light quality in the experiment site decreases. Finally, the challenge is how we can recognize video facial expressions with increased accuracy and in a quick inference time.

6 Conclusion and Future Work

In this paper, we describe the implementation of a video-based automated emotional monitoring prototype consisting of two new subsystems of the SenseCare KM-EP. The first subsystem that is a prototype implementation of the Carina Patient data management and information system for the area managing healthcare service processes. The second subsystem is the Emotion Detection subsystem that is implemented prototypical to detect emotions based on analyzing facial expressions in videos.

We discuss the use of CNNs in an initial prototype implementation to support face detection and expression recognition supporting deriving corresponding emotion as AC results. To establish a REST API we employ the face-api.js package, Node.js, TensorFlow.js core, and MongoDB to store patient detected expressions with the date and time in real-time.

We also have presented an initial conceptual architecture as well as an initial information model of our system and have specified the technical software architecture of the API and discussed our first findings during the implementation of the API. Future work includes:

- Integration of the video-based automated emotional monitoring module in the carna.dmg/KM-EP, and evaluation of the solution in a real HIS (Hospital Information System, e.g. GNU Health).
- Visualization and perception of all stored expressions or Graphical representation of Emotions/Time, in order to make optimal decisions in healthcare.
- Implementation of additional support processes in carna.dmg, and integration of real sources such as video/audio data.

References

1. Crist TM, Kaufman SB, Crampton KR. Home telemedicine: a home health care agency strategy for maximizing resources. *Home Health Care Management Practice*. 1996 Jun 1; 8(4):1-9.
2. Sensor Enabled Affective Computing for Enhancing Medical Care, link : <http://www.sensecare.eu/>, (viewed 24 July 2020).
3. Engel, F., Bond, R., Keary, A., Mulvenna, M., Walsh, P., Hiuru, Z., Kowohl, U., Hemmje, M.L.: *Sensecare: Towards an experimental platform for home-based, visualisation of emotional states of people with dementia*. Computer Science, Springer, 2016.
4. Duan, L., & Zhu, G. (2020). Psychological interventions for people affected by the COVID-19 epidemic. *Lancet Psychiatry*, 7(4), 300-302. doi:10.1016/s2215-0366(20)30073-0.
5. Torous, J., Jän Myrick, K., Rauseo-Ricupero, N., & Firth, J. (2020). Digital Mental Health and COVID-19: Using Technology Today to Accelerate the Curve on Access and Quality Tomorrow. *JMIR Ment Health*, 7(3), e18848. doi:10.2196/18848.
6. Gerke, S., Stern, A. D., & Minssen, T. (2020). Germany's digital health reforms in the COVID-19 era: lessons and opportunities for other countries. *NPJ Digit Med*, 3, 94. doi:10.1038/s41746-020-0306-7.
7. Celano, C. M., Villegas, A. C., Albanese, A. M., Gaggin, H. K., & Huffman, J. C. (2018). Depression and Anxiety in Heart Failure: A Review. *Harv Rev Psychiatry*, 26(4), 175-184. doi:10.1097/hrp.000000000000162.
8. Ismail, Z., Gatchel, J., Bateman, D. R., Barcelos-Ferreira, R., Cantillon, M., Jaeger, J., . . . Mortby, M. E. (2018). Affective and emotional dysregulation as pre-dementia risk markers: exploring the mild behavioral impairment symptoms of depression, anxiety, irritability, and euphoria. *Int Psychogeriatr*, 30(2), 185-196. doi:10.1017/s1041610217001880.
9. M. X. Bornschlegl, K. Berwind, M. Kaufmann, F. C. Engel, P. Walsh, M. L. Hemmje, and R. Riestra, "IVIS4BigData: A reference model for advanced visual interfaces supporting big data analysis in virtual research environments", Lecture

Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 10084 LNCS, pp. 1-18, 2016.

10. B. Vu, 'Realizing an Applied Gaming Ecosystem - Extending an Education Portal Suite towards an Ecosystem Portal', Technische Universität Darmstadt, 2015.
11. OpenID Connect, link: <https://openid.net/connect/> (viewed 24 July 2020).
12. IndivoHealth , <http://indivohealth.org/>.
13. Tolven, <http://tolvenhealth.com> .
14. Ekman, P., & Yamey, G. (2004). Emotions revealed: recognising facial expressions: in the first of two articles on how recognising faces and feelings can help you communicate, Paul Ekman discusses how recognising emotions can benefit you in your professional life. Student BMJ, 12, 140-142.
15. J. Ko, C. Lu, , M. Srivastava, J. Stankovic, A. Terzis, and M. Welsh. Wireless sensor networks for healthcare. In Proceedings of the IEEE, 2010.
16. Subhas Chandra Mukhopadhyay, "Wearable Sensors for Human Activity Monitoring: A Review," IEEE Sensors Journal 15, no. 3 (2015): 1321–30.
17. Mingmin Zhao, Fadel Adib, Dina Katabi,"Emotion Recognition using Wireless Signals", 2016 , <http://equadro.csail.mit.edu/files/equadro-paper.pdf>.
18. Valossa Video AI | Video Recognition | Image Analysis | Content Intelligence, link: <https://valossa.com/>, (viewed 24 July 2020).
19. Facial expression recognition software | FaceReader, link: <https://www.noldus.com/facereader>, (viewed 24 July 2020).
20. SHORE@, link: <https://www.iis.fraunhofer.de/en/ff/sse/imaging-and-analysis/ils/tech/shore-facedetection.html>, (viewed 24 July 2020).
21. T. Guo, J. Dong, H. Li and Y. Gao, "Simple convolutional neural network on image classification," 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), Beijing, 2017, pp. 721-724.
22. Machine Intelligence and Signal Processing, Ebook, Proceedings of International Conference, Springer, Singapore, MISP 2019, ISBN 978-981-13-0923-6.
23. OpenCV (Open Source Computer Vision Library), link: <https://opencv.org/> , (viewed 23 June 2020).
24. Lange, Julian: Use of a data management system to support the Treatment of patient in the psychological field. 2019.
25. TensorFlow.js, JavaScript library for machine learning, link: <https://www.tensorflow.org/js>, (viewed 23 June 2020).]
26. Face-api.js, JavaScript API for face detection and face recognition in the browser and nodejs with tensorflow.js, link:<https://github.com/justadudewhohacks/face-api.js/>, (viewed 23 June 2020).
27. Node.js, link: <https://nodejs.org/en/>, (viewed 24 July 2020).
28. The most popular database for modern apps | MongoDB link: <https://www.mongodb.com/> , (viewed 24 July 2020).
29. Express - Node.js web application framework, <https://expressjs.com/>, (viewed 24 July 2020).
30. FACES A database of facial expressions in younger, middle-aged, and older women and men <https://faces.mpdl.mpg.de/imeji/collection/IXTdg721TwZwyZ8e?q=#> , (viewed 24 July 2020).