Intelligent Student Feedback System for Online Education

Hari Krishnan
Dept. of Computer Science and Engineering
Christ College of Engineering,
Irinjalakuda, Thrissur, India

Liya K.V
Dept. of Computer Science and Engineering
Christ College of Engineering, Irinjalakuda
Thrissur, India

Lazar Tony
D Dept. of Computer Science and Engineering
Christ College of Engineering, Irinjalakuda
Thrissur, India

Nova Mary Thomas
Dept. of Computer Science and Engineering
Christ College of Engineering,
Irinjalakuda, Thrissur, India

Remya K Sasi
Dept. of Computer Science and Engineering
Christ College of Engineering,
Irinjalakuda Thrissur, India

***Abstract*—Nowadays, deep learning techniques are gaining big success in various fields including computer vision. Indeed, a convolutional neural networks (CNN) model can be trained to analyze images and identify face emotion. Our project aims to create a system that recognizes students' emotions from their faces. Our system consists of four phases: face detection using MTCNN, normalization, emotion recognition using CNN on FER 2013 database and calculation of concentration metric with seven types of expressions. Obtained results show that face emotion recognition is feasible in education, consequently, it can help teachers to modify their presentation according to the students' emotions.**

Keywords— Student facial expression; Emotion recognition; Convolutional neural networks (CNN); Deep learning; Intelligent Student feedback system.

I. INTRODUCTION

Learning is an exciting adventure during which both the teacher and the students participate. The participation of the student is inevitable in improving the quality of education, and it is very important to receive live feedback from students to adjust their pedagogy and proceed with their classes in an effective manner. In a conventional scenario, a teacher takes this input from the facial and body expressions of their studentsbutthisisnotpossibleinanonlinescenario.

The face is the most expressive and communicative part of a person's being. Facial expression recognition identifies emotion from a face image, it's a manifestation of the activity and personality of a person . According to diverse research, emotion plays an important role in education. The establish- ment of a learning-emotion recognition model to guideonline education can improve not only the quality of teaching but also the real-time nature of information transmission, which is of great significance for the construction of learner-oriented teaching to create personalized learning and so on. Moreover, the general learning mood of learners also reflects theteaching quality of the instructors. Learning the emotions of learners in online education has also become an important indicator for assessingtheteachingqualityofinstructors.

With the development of computer vision in recent years, the accuracy of facial expression based on face detection has continuously improved and It has become easy to observe the students' reaction on a particular topic which is being taught by theinstructor.

The purpose of our project is to implement emotion recognition in education by realizing an automatic system that analyzes students' facial expressions based on Facial Emotion Recognition, which is a deep learning algorithm that is widely used in facial emotion detection and classification. It is a Convolutional Neural Network model that consists of a multi- stage image processing to extract feature representations. Our system includes four phases: face detection, normalization, emotion recognition and calculation of concentration metric. There are seven emotions under consideration: neutral, anger, fear, sadness, happiness, surprise and disgust.

II. EXISTING WORKS

The majority of the student feedback systems use FER. FER systems can be classified mainly into two main categories:

Emotion prediction from extracted facial features and facial emotion recognition directly from the facial images

Currently, CNN is the most widely used method for FER, followed by SVM, FNN, HMM, Binary Classifiers ,and other forms of Neural Networks.CNN has the flexibility of modification and raises some opportunities for further researchers to develop a new method of recognition from CNN modification.

III. PROPOSEDWORK

To implement a proof of concept ”Intelligentstudent feed- back system ” consisting of 2 interfaces: student and faculty interfaces The student interface deals with the contentdelivery, emotion recognition and the calculation of concentration metric The faculty interface enables users to upload thecontent
, integrate the individual metric of each student and most importantly provides a user-friendly data visualization.

IV. TECHNOLOGY

The project is implemented in Python 3.8.9, versions above
3.8.9 cannot be used as the current version of tensorflow framework does not support higher versions of Python and versions below 3.5 cannot be used as the PySide2 module does not support those versions. The GUI for the project is implemented using PySide2. Frameworks like OpenCV, TensorFlow and Keras are used to handle the web camera andtheneuralnetworks.SQLite3isusedasthedatabase

A. Python

Python is an interpreted high-level general-purpose programming language. Python’s design philosophy emphasizes code readability with its notable use of significant indentation. Its language constructs as well as its object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python is dynamically- typed and garbage-collected. It supports multiple programming paradigms, including structured (particularly, procedural), object-oriented and functional programming. Python is often described as a ”batteries included” language due to its comprehensive standard library. All Python releases are Open Source

B. PySide2

PySide2 is the official Python module from the Qt for Python project, which provides access to the complete Qt 5.12+ framework. Qt for Python is available under LGPLv3/GPLv2 and commercial license. Qt is a cross- platform application development framework for desktop, embedded and mobile. Supported Pl Linux, OS X, Windows, VxWorks, QNX, Android, iOS, BlackBerry, Sailfish OS and others. Qt is not a programming language on its own. It is a framework written in C++. A preprocessor, the MOC (Meta- Object Compiler), is used to extend the C++ language with features like signals and slots. Before the compilation step, the MOC parses the source files written in Qt-extended C++ and generates standard compliant C++ sources from them. Thustheframeworkitselfandapplications/librariesusingitcanbe compiled by any standard compliant C++ compiler like Clang, GCC, ICC, MinGW and MSVC. Qt is available under various licenses: The Qt Company sells commercial licenses, but Qt is also available as free software under several versions of the GPL and the LGPL The latest version of PySide, PySide6 for Qt6, was not used as several features essential for this project are not available init.

C. OpenCV-Python

OpenCV (Open Source Computer Vision Library: http://opencv.org) is an open source library that includes several hundreds of computer vision algorithms. It is a C++ API. OpenCV 4.5.0 and higher versions are licensed under the Apache 2 License. OpenCV 4.4.0 and lower versions, including OpenCV 3.x, OpenCV 2.x, and OpenCV 1.x, are licensed under the 3-clause BSD license. OpenCV-Python is the Python API for OpenCV, combining the best qualities of the OpenCV C++ API and the Python language. Compared to languages like C/C++, Python is slower. That said, Python can be easily extended with C/C++, which allows us to write computationally intensive code in C/C++ and create Python wrappers that can be used as Python modules. This givesus two advantages: first, the code is as fast as the original C/C++ code (since it is the actual C++ code working in the background) and second, it is easier to code in Python than C/C++. OpenCV-Python is a Python wrapper for the original OpenCV C++ implementation. OpenCV-Python makes use of Numpy, which is a highly optimized library for numerical operations with a MATLAB-style syntax. All the OpenCV array structures are converted to and from Numpyarrays. This also makes it easier to integrate with other libraries that use Numpy such as SciPy and Matplotlib. Opencv-python package is available under MIT license.

D. TensorFlow

TensorFlow is an end-to-end open source platform for ma- chine learning. Its flexible architecture allows easydeployment of computation across a variety of platforms (CPUs, GPUs, TPUs), and from desktops to clusters of servers to mobile and edge devices. Originally developed by researchers and engineers from the Google Brain team within Google’s AI organization, it comes with strong support for machine learning and deep learning and the flexible numerical computation core is used across many other scientific domains. TensorFlow pro- vides stable Python and C++ APIs, as well as non-guaranteed backward compatible API for other languages. Tensor Flowis cross-platform. It runs on nearly everything: GPUs and CPUs—including mobile and embedded platforms—and even tensor processing units (TPUs), which are specialized hard-ware to do tensor math on. The TensorFlow distributed execution engine abstracts away the many supported devices and provides a high performance-core implemented in C++ for the TensorFlow platform. On top of that sit the Python and C++ frontends. The Layers API provides a simpler interface for commonly used layers in deep learning models. On top of that sit higher level APIs, including Keras more on the Keras.io site) and the Estimator API, which makes training and evaluating distributed models easier. It was released under the Apache License2.0.

E. Keras

Keras is an open-source software library that provides a Python interface for artificial neural networks. Keras acts asan interface for the TensorFlow library. It provides essential abstractions and building blocks for developing and shipping machine learning solutions with high iteration velocity. Keras contains numerous implementations of commonly used neural network building blocks such as layers, objectives, activation functions, optimizers, and a host of tools to make working with image and text data easier to simplify the coding necessary for writing deep neural network code. In addition to standard neural networks, Keras has support for convolutional and recurrent neural networks. It supports other common utility layers like dropout, batch normalization, and pooling. Keras allows users to productize deep models on smartphones (iOS and Android), on the web, or on the Java Virtual Machine. It also allows use of distributed training of deep-learning models on clusters of Graphics processing units (GPU) and tensor processing units(TPU).

F. SQLite

SQLite is a relational database management system (RDBMS) contained in a C library that implements a self-contained, serverless, zero-configuration, transactional SQL databaseengine.SQLiteisanembeddedSQLdatabaseengine. Unlike most other SQL databases, SQLite does not have a separate server process. SQLite reads and writes directly to ordinary disk files. A complete SQL database with multiple tables,indices,triggers,andviews,iscontainedinasingledisk file. The database file format is cross-platform - one can freely copy a database between 32-bit and 64-bit systems or between big-endian and little-endian architectures. These features make SQLiteapopularchoiceasanApplicationFileFormat.SQLite database files are a recommended storage format by the US Library of Congress. SQLite3 can be integrated with Python using the sqlite3 module, which was written by Gerhard Haring. It provides an SQL interface compliant with the DB- API2.0specificationdescribedbyPEP249.

V.IMPLEMENTATION

The project consists of two graphical interfaces: a student interface and a dashboard, a database and two neuralnetworks: MTCNN and a custom keras model. The student interface allows a student to login and attend the lectures. While the lecture is being delivered the webcam captures the images of the student from which the face of the student is recognized and cropped with the help of the MTCNN model. This face is given as input to the custom keras model which predicts the probability scores for seven emotions, namely sad, happy, disgust, surprise, fear, neutral, anger. A weighted average of these values is obtained by multiplying the corresponding confidence scores with the predefined weight for thatemotion.

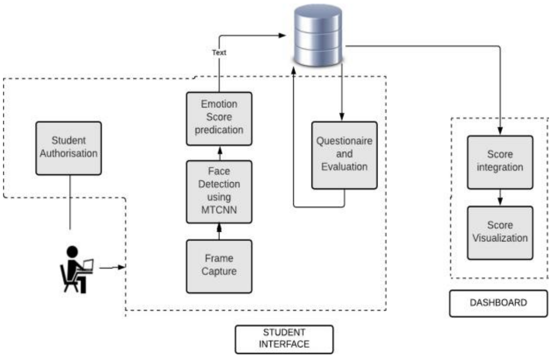


Fig. 1. Flowchart

These predefined weights signifies there lation of that emotion towards the level of concentration. These scores are stored in the database. At the end of the lecture a set of predefined questions are popped on the screen and the student responses are also stored in the database. The dashboard visualizes the corresponding analysis of each lecture. For a given lecture the whole of the lecture is divided into segments of 2 seconds. The average scores of every student for these segments are taken and these values are in turn averaged to get a single score for everysegment.

A. Graphical Interfaces

1) Student Interface: The student interface consists of a QFrame for login purposes, a QListView for video list, a QVideoWidget for displaying the lecture, a QLabel for camera preview, two QButtons for play and pause and a QListWidget for displaying the frame details. The login frame has two QLineEdit objects and two QPushButtons. Only after success- ful login will the other widgets be activated. On clicking an item from the video list the corresponding lecture is displayed on the video widget and the live preview of the camera is displayed on theQLabel.

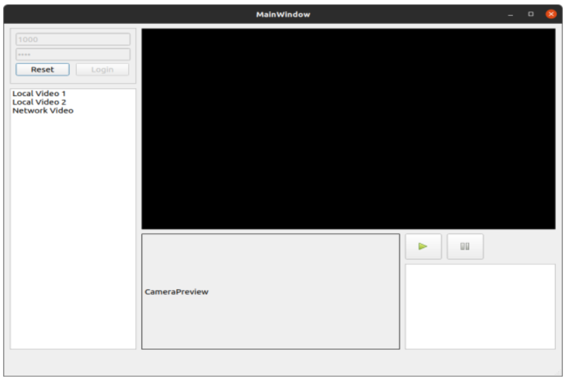
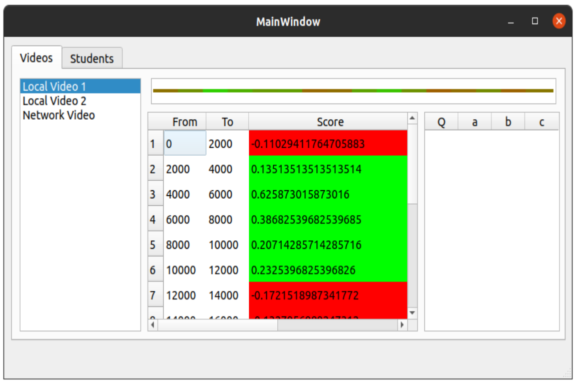


Fig. 2. Student interface

2) Dashboard: It is a graphical interface for faculty which provides at - a glance views of student understanding level relevanttoeachvideoinauserfriendlymanner.



	From	To	Score
1	0	2000	🟥 -0.11029411764705883
2	2000	4000	🟩 0.13513513513513514
3	4000	6000	🟩 0.625873015873016
4	6000	8000	🟩 0.38682539682539685
5	8000	10000	🟩 0.20714285714285716
6	10000	12000	🟩 0.2325396825396826
7	12000	14000	🟥 -0.1721518987341772

Fig. 3. faculty dashboard.

3) *Database:*The database consists of five tables:

a) *Student:* To store the details ofstudents

b) *Video:* To store the details oflectures

c) *Question:* To store the questions for eachlecture

d) *Frame data:* To store the emotion scores of students withrespecttoframesofavideolecture

e) *Answer:* To store the student response a question

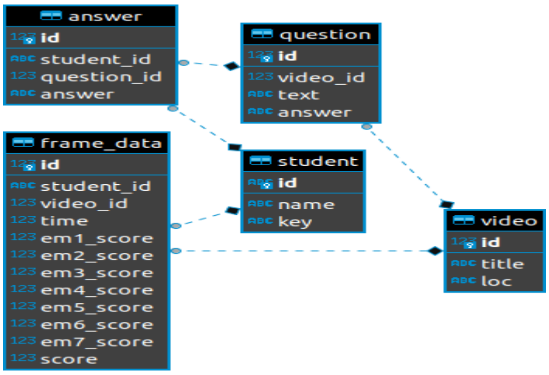


Fig. 4. Database.

B. NeuralNetworks

4) *MTCNN:* Facial detection is a technique used by computer algorithms to detect a person's face through images. Accordingly, the objective of facial detection is to get different features of human faces from images. Even thoughthere are many Face detection classifiers we have used MTCNN MTCNN (Multi-task Cascaded Convolutional Neural Net- works) is an algorithm consisting of 3 stages, which detects the bounding boxes of faces in an image along with their 5 Point FaceLandmarks

a) Stage 1: The Proposal Network (P-Net) This first stage is a fully convolutional network (FCN). The difference between a CNN and a FCN is that a fully convolutional network does not use a dense layer as part of the architecture. This Proposal Network is used to obtain candidate windows and their bounding box regression vectors. Bounding box regression is a popular technique to predict the localization of boxes when the goal is detecting an object of some predefined class, in this case faces. After obtaining the bounding box vectors, some refinement is done to combine overlapping regions. The final output of this stage is all candidate windows after refinement to downsize the volume of candidates.

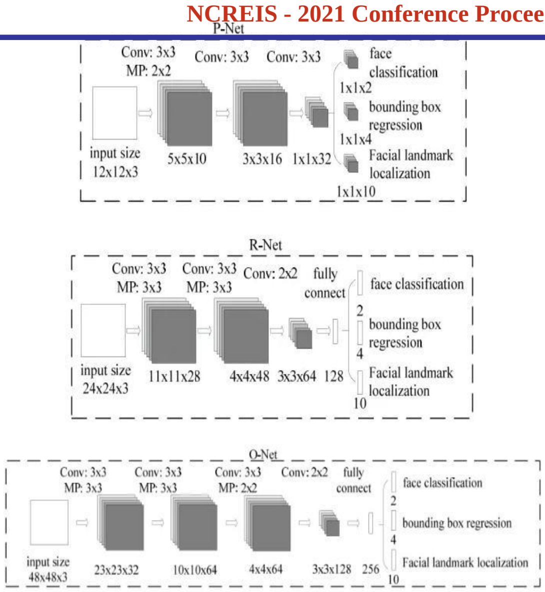


Fig. 5. MTCNN

b) Stage 2: The Refine Network (R-Net) All candidates from the P-Net are fed into the Refine Network. Notice that this network is a CNN, not a FCN like the one before since there is a dense layer at the last stage of the network architecture. The R-Net further reduces the number of candidates, performs calibration with bounding box regression and employs non-maximum suppression (NMS) to merge overlapping candidates. The R-Net outputs whether the input is a face or not, a 4 element vector which is the bounding box for the face, and a 10 element vector for facial landmark localization.

c) Stage 3: The Output Network (O-Net) This stage is similar to the R-Net, but this Output Network aims to describe the face in more detail and output the five facial landmarks' positions for eyes, nose and mouth. The detector returns a list of JSON objects. Each JSON object contains three main keys: 'box', 'confidence' and' keypoints':

The bounding box is formatted as [x, y, width, height] under the key 'box'.

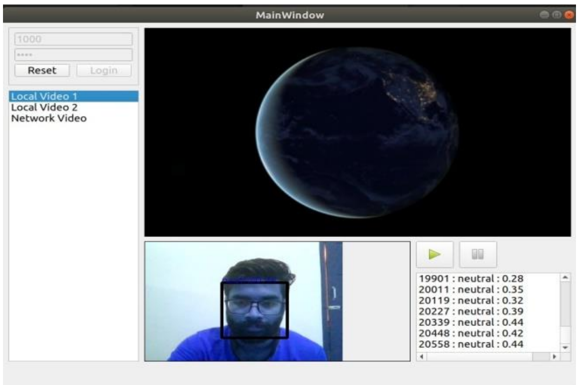
The confidence is the probability for a bounding box to be matching a face.

The keypoints are formatted into a JSON object with the keys 'left eye', 'right eye', 'nose', 'mouth left', 'mouth right'. Each keypoint is identified by a pixel position (x, y). We have tested 4 algorithms(MTCNN, Dlib, OpenCV DNN, OpenCV Haar) using the same video and compared. After analyzing we could observe greatest number of correct face detection than others.Greatestaccuracyprovidedwasthereasonforselecting MTCNN

5) Custom Keras Model: We use a keras model for facial emotion recognition. The faces from the MTCNN model are usedasinputtothisnetwork.Themodelconsistsofseveral2D convolutional layers with ReLU activation function and max pooling. Batch normalisation is used to stabilise the learning process and dramatically reduce the number of training epochs required to train the deep networks. Softmax function is used as the last activation function of the network to normalize the output to a probability distribution over the 7 predicted output classes.Thisnetworkgivesasetofpredictedconfidencescores for the seven emotionclasses.

VI. EXPERIMENTALRESULT

The project was evaluated by 12 different volunteers. 7 out of the 12 volunteers affirmed the prediction of our system. The major factor that accounted for the inaccuracy of the system in other students was that of different baseline emotions. That is the effect of different emotions on the level of concentration was different for different students. A student with generally a sad face always received a low concentration score no matter how concentrated he was. In order to overcome this errorwe can compare the scores of the questionnaire and the predicted score and then make slight changes in the weights of the emotions accordingly for that particular student. If there is significant change in both these scores we adjust the weights of the emotion to produce accurate results. The alterations in weightsaredoneseparatelyforeachindividualstudent.



```
19901 : neutral : 0.28
20011 : neutral : 0.35
20119 : neutral : 0.32
20227 : neutral : 0.39
20339 : neutral : 0.44
20448 : neutral : 0.42
20558 : neutral : 0.44
```

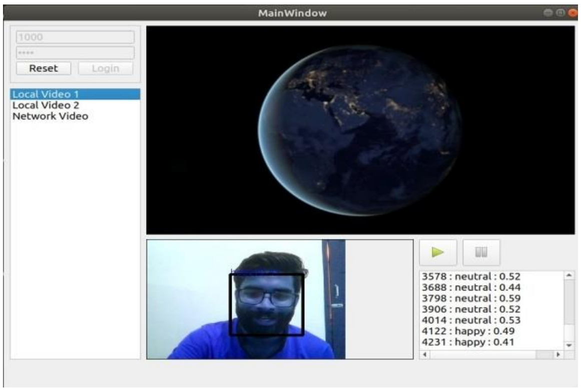
Fig. 6. Student interface:neutral

VII. ACKNOWLEDGMENT

This project is realized as part of the B-TECH. project. The authors acknowledge the support from Dr. Remya K .Sasi, HOD, Department of Computer Science and Engineering, Christ College of Engineering.

VIII. CONCLUSION

Emotions of learning and acquiring knowledge are inextricably intertwined. The establishment of a learning-emotion- recognition model to guide online education can improve the quality of teaching and leads to the construction of learner- oriented teaching to create personalized learning. We have modelled a system that has a wide scope in mimicking classroom feedback system to virtual classrooms. Though this is not a foolproof solution we hope that this project will serve as a pioneer and guiding line for future development in this field.



```
3578 : neutral : 0.52
3688 : neutral : 0.44
3798 : neutral : 0.59
3906 : neutral : 0.52
4014 : neutral : 0.53
4122 : happy : 0.49
4231 : happy : 0.41
```

Fig. 7. Student interface:
happy.

REFERENCES

[1] Sahla K. S, T. Senthil Kumar "Classroom Teaching Assessment Based on Student Emotion," 2019 9th International Conference on Education and Socail Science(ICESS2019).

[2] Chao Ma, Chongliang Sun, Donglei Song, Xuan Li, Hao Xu," A Deep Learning Approach for Online Learning Emotion Recognition" The 13th International Conference on Computer Science & Education (ICCSE 2018)

[3] ImaneLasri, Anouar Riad Solh, Mourad El Belkacemi, "Facial Emotion Recognition of Students using Convolutional Neural Network," 2019 IEEE.

[4] JielongTang,Xiaotian Zhou, Jiawei Zheng , "Design of Intelli- gent classroom facial recognition based on Deep Learning",Journal of Physics: Conf. Series 1168 (2019) 022043, doi:10.1088/1742- 6596/1168/2/022043

[5] Archana Sharma, VibhakarMansotra," Deep Learning based Student Emotion Recognition from Facial Expressions inClassrooms",International Journal of Engineering and Advanced Technology (IJEAT),ISSN: 2249 – 8958, Volume-8 Issue-6, August 2019

[6] OussamaElHammoumi, FatimaezzahraBenmarrakchi, Nihal Ouherrou, JamalElKafi,AliElHore,"UseoffacialemotionrecognitioninE - learning systems", Information Technologies and Learning Tools · September 2017, DOI: 10.33407/itlt.v60i4.1743

[7] Andreas Savva, VassoStylianou, Kyriacos Kyriacou, Florent Domenach "Recognizing Student Facial Expressions:AWeb Application", 2018 IEEEGlobalEngineeringEducationConference(EDUCON)

[8] ChinunBoonroungrut, Toe ToeOo, Kim One, "Exploring Classroom Emotion with Cloud-Based Facial Recognizer in the Chinese Beginning Class:APreliminaryStudy1",InternationalJournalofInstruction, January 2019

[9] Krithika L.B., Lakshmi Priya GG, Student Emotion Recognition System (SERS) for e-learning improvement based on learner concentrationmetric 2016, doi:10.1016/j.procs.2016.05.264

[10] Sheng Chen, Jianbang Dai, YizheYan, Classroom Teaching Feedback SystemBasedonEmotionDetection2019DOI.2526/icess.2019.179

[11] ChuangaoTang, Pengfei Xu, Zuying Luo, GuoxingZhao,Tian Zou, Automatic Facial Expression Analysis of Students in Teaching Envi- ronments 2015, DOI:10.1007/978-3-319-25417-352

[12] D.Yang, AbeerAlsadoon, P.W.C. Prasad, A. K. Singh, A. Elchouemi, An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment, 2018, 10.1016j.procs.2017.12.003

[13]Ug˘urAyvaz,Hu¨seyinGu¨ru¨ler,MehmetOsmanDevrim,USEOF FACIAL EMOTION RECOGNITION IN E-LEARNING SYSTEMS, 2017, ISSN:2076-8184.

[14]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,

[15] Facial Expression Recognition for Learning Status Analysis. 2011, Human-Computer Interaction, Part IV,HCII 2011, LNCS 6764, pp. 131–138,2011.

[16] Abdulkareem Al-Alwani,Mood Extraction Using Facial Features to Improve Learning Curves of Students in E-Learning Systems, (IJACSA) InternationalJournalofAdvancedComputerScienceandApplications,Vol. 7, No. 11, 2016

[17] Pan Xiang, Facial Expression Recognition for Learning Status Analysis, 2011,DOI10.1109/ICDMA.2011.255

[18] Kenichi Takahashi, Masahiro Ueno, Improvement of detection for warning students in e-learning using web cameras, 2014, doi:10.1016/j.procs.2014.08.157