A Project Report

on

Find attentiveness of students present in class using AI and ML.

Submitted in partial fulfillment of the
requirement for the award of the degree of

Bachelor of Technology.

GALGOTIAS UNIVERSITY
(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

Under The Supervision of--
Dr. K.M Baalamurugan :

Submitted By

Vishal	Vishal Das
18SCSE1010006	18SCSE1010324

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA
INDIA

Dec,2021

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING

GALGOTIAS UNIVERSITY, GREATER NOIDA

CANDIDATE'S DECLARATION

I/We hereby certify that the work which is being presented in the thesis/project/dissertation, entitled **"Find attentiveness of students present in class using AI and ML"** in partial fulfillment of the requirements for the award of the Bachelors of Technology submitted in the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of month, Year to Month and Year, under the supervision of Name **"Dr. K.M Baalamurugan"** Designation, Department of Computer Science and Engineering/Computer Application and Information and Science, of School of Computing Science and Engineering , Galgotias University, Greater Noida. The matter presented in the thesis/project/dissertation has not been submitted by us for the award of any other degree of this or any other places.

Vishal(18SCSE1010006)

Vishal Das(18SCSE1010324)

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Dr. K.M Baalamurugan

CERTIFICATE

The Final Thesis/Project/ Dissertation Viva-Voce examination of **“Vishal(18SCSE1010006) and Vishal Das(18SCSE1010324) ”** has been held on __________________ and their work is recommended for the award of B.Tech(CSE).

Signature of Examiner(s) **Signature of Supervisor(s)**

Signature of Project Coordinator **Signature of Dean**

Date: December, 2021

Place: Greater Noida

ACKNOWLEDGEMENT

Apart from the efforts of our, the success of any project depends largely on the encouragement and guidelines of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project.

We would like to show my greatest appreciation to "**Dr. K.M Baalamurugan "** and also our reviewer "**Dr. Kavita"** and also our dean "**Dr. Munish Sabarwal"**. We can't say thank you enough for his tremendous support and help. We feel motivated and encouraged every time, We attend his meeting. Without his encouragement and guidance this project would not have materialized.

The guidance and support received from all the members who contributed and who are contributing to this project, was vital for the success of the project. We are grateful for their constant support and help.

Abstract

Now days due to pandemic everything got shut down and every one has to switch to online meeting for the jobs , education and for business purposes . In the school class , the teacher can’t observe the childe closely that’s why some students faces lack of education even paying the same fees as in offline classes . So to improve this thing we proposed a model in which we observe each and every child or employ closely so that we can make some changes as much as we can do . Our aim is to give better education to the child just with the help of small things . when a teacher makes a video call or a zoom meeting for class, every time there is a lot of students are present at same time. To observe everyone of them we have to assign this work to a observer who will do this work for our . he will go to each and every student window and observe what they are doing . If it’s find anything inappropriate he will report to the teacher so that the teacher can take some strict action on student. Teachers must be able to monitor students’ behavior and identify valid cues in order to draw conclusions about student's actual engagement in learning activities. Teacher training can support (inexperienced) teachers in developing these skills by using videotaped teaching to highlight which indicators should be considered. However, this supposes that (a) valid indicators of students’ engagement in learning are known and (b) work with videos is designed as effectively as possible to reduce the effort involved in manual coding procedures and in examining videos

Table of Contents

CHAPTER-1

Introduction

1.1 Introduction:

One avenue for addressing these issues is to utilize the technological advances made in recent years in fields such as machine learning to improve the analysis of classroom videos. Assessing students' attention-related processes through visible indicators of (dis)engagement in learning might become more effective if automated analyses can be employed. Thus, in the present study, we validated a new manual rating approach and provided a proof of concept for a machine vision-based approach evaluated on pilot classroom recordings of three lessons with university students. The manual rating system was significantly correlated with self-reported cognitive engagement, involvement, and situational interest and predicted performance on a subsequent knowledge test. The machine vision-based approach, which was based on gaze, head pose, and facial expressions, provided good estimations of the manual ratings. Adding a synchrony feature to the automated analysis improved correlations with the manual ratings as well as the prediction of posttest variables. The discussion focuses on challenges and important next steps in bringing the automated analysis of engagement to the classroom.

Cognitive activation, classroom management, and teacher support are the three central tenants of teaching quality (Klieme et al. 2006; Praetorius et al. 2018). The level of students' (dis)engagement in learning activities can be considered a major indicator of both cognitive activation and classroom management because it signals students' engagement in the deep processing of learning content and reveals the time on task (Carroll 1963) provided by the teachers for students' learning. To this end, teachers are required to take note of their students' attentional focus and make sure the students are engaging in the desired learning activities. Thus, the ability to monitor students' attention and to keep it at a high level is part of the competencies that novice teachers need to acquire. However, research has indicated that teachers might not always be aware of their students' attentional focus, and this may be particularly true for novice teachers.

In general, beginning teachers have trouble monitoring all students in the classroom evenly and noticing events that are relevant for student learning (Berliner 2001; Cortina et al. 2015; Star and Strickland 2008; Stürmer et al. 2017). Therefore, teacher training needs to support future teachers in developing the necessary knowledge structures that underlie these abilities (e.g., Lachner et al. 2016). Consequently, providing an improved measurement approach for student attention will be beneficial for research and can potentially contribute to teacher training.

Research has already demonstrated that both inexperienced and experienced teachers' ability to notice relevant cues in the classroom benefits from observing and reflecting on their own videotaped teaching (Kleinknecht and Gröschner 2016; Sherin and van Es 2009). Until now, however, instructors have typically had to watch hours of video material to select the most crucial phases of lessons. Similarly, when it comes to research on teaching effectiveness and the development of teachers' ability to notice relevant cues in classroom instruction (i.e., professional vision skills), researchers typically have to invest considerable resources, especially coding resources, to examine the association between teacher behavior and classroom processes (Erickson 2007). The required effort further increases when investigating students' attention across an entire lesson and analyzing attention at the group level instead of among individuals. In this vein, attention- and engagement-related behavior during classroom instruction has rarely been studied due to the difficulty of data collection and labeling. However, learners might behave differently in naturalistic settings and show versatile behavior that cannot be found in a lab.

One potentially valuable avenue for addressing these issues is to utilize the technological advances made in recent years in fields such as computer vision and machine learning. Therefore, in an ongoing research project (Trautwein et al. 2017), we have been investigating whether and how the automated assessment of students' attention levels can be used as an indicator of their active engagement in learning. This automated assessment can in turn be used to report relevant cues back to the teacher, either simultaneously or by identifying and discussing the most relevant classroom situations (e.g., a situation where students' attention increases or decreases significantly) after a lesson.

Student attention is a key construct in research on both teaching and learning. However, definitions vary widely and are discussed from multiple perspectives. Here, we focus on describing three lines of research that inspired our research program: cognitive psychology models that describe attention as part of information processing, engagement models in which attention makes up part of a behavioral component, and teaching quality models in which student attention is a crucial factor.

In current models in the psychology of learning, *attention* denotes a filtering mechanism that determines the kind and amount of information that enters working memory. This mechanism is crucial for preventing working memory overload and allows the learner to focus on the right kind of information. Only sensory information that enters working memory is encoded, organized, and

linked to already existing knowledge. Thus, attention serves as a selection process for all incoming sensory information as it dictates which pieces of information will be processed further and will get the chance to be learned. Thus, attention determines the success of knowledge construction . Engle (2002) further proposed that executive attention, which actively maintains or suppresses current representations in working memory, is part of working memory. Certain instructional situations strongly depend on executive processes such as shifting, inhibition, or updating (Miyake et al. 2000) and thus necessitate top-down attentional control. Although information processing occurs in a covert manner, some aspects of attentional processes are likely to be observed from the outside: for example, visually orienting toward a certain stimulus, which improves processing efficiency (Posner 1988).

Attention is often mistaken for engagement, even though it constitutes only part of it. Engagement is defined as a multidimensional meta-construct and represents one of the key elements for learning and academic success (Fredricks et al. 2004). It includes observable behaviors, internal cognitions, and emotions. Covert processes such as investment in learning, the effort expended to comprehend complex information, and information processing form part of cognitive engagement (Fredricks et al. 2004; Pintrich and De Groot 1990). Emotional engagement in the classroom includes affective reactions such as excitement, boredom, curiosity, and anger (Connell 1990; Fredricks et al. 2004). Attention is considered a component of behavioral engagement alongside overt participation, positive conduct, and persistence .Per definition, cognitive engagement refers to internal processes, whereas only the emotional and behavioral components are manifested in visible cues. Nevertheless, all engagement elements are highly interrelated and do not occur in isolation (Fredricks et al. 2004). Thus, attention plays a crucial role because it may signal certain learning-related processes that should become salient in students’ behavior to some extent.

Learners’ attention also plays a crucial role in research on teaching. Teachers must determine whether their students are attentive by considering visible cues, continually monitoring the course of events in order to manage the classroom successfully (Wolff et al. 2016) and providing ambitious learning opportunities. A student’s attention or lack thereof (e.g., when distracted or engaging in mind wandering) can signal whether she or he is on-task or off-task. This in turn can provide hints about instructional quality and the teacher’s ability to engage his or her students in the required learning activities. Thus, it is important to help teachers develop the skills needed to monitor and support student attention and engagement and adapt their teaching methods. Consequently, accounting for student attention and more broadly student engagement in teaching is considered crucial for ensuring teaching quality, including classroom management, cognitive activation, and instructional support .

In sum, the definitions, theoretical backgrounds, and terminology used in various lines of research to describe observable aspects of students' cognitive, affective, or behavioral attention/engagement in learning are diverse, but experts agree on their importance and key role in learning. As teachers must rely on visible cues to judge their students' current attention levels, we focused on observable aspects of attention and inferences that were based on visible indicators. In the remainder of the article, we use the term *visible indicators of (dis)engagement in learning* to describe these aspects. These visible indicators are highly likely to be associated with learning, but this assumption needs to be validated.

1.2 Formulation of problems :

The difficulty in assessing students' engagement-related processes in real-world classroom settings consists of externalizing learners' internal (covert) states through visible overt aspects to the greatest extent possible. In psychology, affective states and cognitive processes such as attentional control are usually determined from physiological signals, such as heart rate, electrodermal activity, eye tracking, or electroencephalography. Using this kind of

psychologically sound measurements makes it possible to detect covert aspects of learningrelated processes; however, these measures are hardly feasible in classroom instruction, especially when teachers must be equipped with knowledge about what indicators to look for in students. Furthermore, these approaches are useful for answering very specific research questions. However, they are not sufficient for determining whether students' ongoing processes are

actually the most appropriate for the situation. By contrast, overt behavior can provide visible indicators of appropriate learning-related processes in students.

Overt classroom behavior is an important determinant of academic achievement (Lahaderne 1968; McKinney et al. 1975). Although overt behavior does not always represent a reliable indicator of covert mental processes, previous findings have demonstrated a link between cognitive activity and behavioral activity (Mayer 2004). Previous studies have analyzed students' behavior and have determined its relation to achievement (Helmke and Renkl 1992; Hommel 2012; Karweit and Slavin 1981; Stipek 2002). Furthermore, in research on engagement, correlations between student engagement and academic achievement have been found (Lei et al. 2018). Other studies have found opposing results (e.g., Pauli and Lipowsky 2007); however, these studies either relied on self-reports as opposed to observer ratings or only focused on certain facets of engagement-related behavior (e.g., only active on-task behavior).

There have been various attempts to systematically assess visible indicators of engagement in classroom learning, for example, Helmke and Renkl (1992) based their research on an idea by Ehrhardt et al. (1981) and related observable student behavior to internal processes using timeon-task as an indicator of whether a student was paying attention to classroom-related content.

Assessing observable content-related behavior is essential to this operationalization of higher order attention. Hommel (2012) modified this approach and applied it to the video-based analysis of instructional situations. Rating behavior as either on- or off-task with varying subcategories demonstrated the interrelation between visual cues and achievement or reduced learning (Baker et al. 2004; Helmke and Renkl 1992).

However, learners can differ in their learning activities but still be engaged in a certain task. The ICAP framework proposed by Chi and Wylie (2014) distinguishes between passive, active, constructive, and interactive overt behavior, which differ across various cognitive engagement activities. This framework focuses on the amount of cognitive engagement, which can be detected from the way students engage with learning materials and tasks (Chi and Wylie 2014). This theoretical model provides a promising approach for further expanding the different types of on-task behavior so that variations in student behavior can be accounted for.

In sum, considering learning content has been shown to be useful; however, there is a lack of research involving the continuous analysis of attention or engagement over the course of one or more lessons. A unique feature of the present study is that we aimed to acquire a continuous assessment (i.e., a score for every student in the classroom for every second of instruction time) of students' visible indicators of (dis)engagement in learning. This temporal resolution was crucial in our approach because we aimed to provide comparable data that could be used to train a machine-learning algorithm. To reach this high level of temporal resolution, we decided to annotate learners' behavior continuously. The free software CARMA (Girard 2014) enables the continuous interpersonal behavior annotation by using joysticks (see Lizdek et al. 2012). However, this new approach limited us in terms of using already existing rating instruments because existing instruments do not allow for a high enough level of temporal resolution.

Furthermore, the CARMA software requires annotations on a scale rather than rating the behavior in terms of categories as already existing instruments do. When developing the new instrument, we mainly oriented on the MAI (Helmke and Renkl 1992; Hommel 2012). However, we needed to define more fine-grained indicators of student behavior to make annotations along a continuous scale possible. Therefore, we added indicators from various established instruments to extend our rating scale. We assumed that the manual observer annotations would serve only as approximations of the actual cognitive states of the students and that the averaged (i.e., intersubjective) manual annotations would reflect the "true score" of the visible indicators of (dis)engagement in learning better than a single rater could. Subsequent to the ratings, we thus calculated the mean of the raters for every second. The mean values for each second and student were used as the ground truth to train a machine-learning approach.

1.2.1 Tools and Technology Used:

- ❖ SOFTWARE REQUIREMENT :

  As the project is developed in python, we have used Anaconda for Python 3.6.5 and PyCharm.

1. **Anaconda**

It is a free and open source distribution of the Python and R programming languages for data science and machine learning related applications (large-scale data processing, predictive analytics, scientific computing), that aims to simplify package management and deployment. Package versions are managed by the package management system conda. The Anaconda distribution is used by over 6 million users, and it includes more than 250 popular data science packages suitable for Windows, Linux, and MacOS.

2.PyCharm:-

It is an IDE i.e. Integrated Development Environment which has many features like it supports scientific tools(like matplotlib, numpy, scipy etc) web frameworks (example Django,web2py and Flask) refactoring in Python, integrated python debugger, code completion, code and project navigation etc. It also provides Data Science when used with Anaconda.

❖ Features include:

- editor with syntax highlighting and introspection for code completion

- support for multiple Python consoles (including IPython)
- the ability to explore and edit variables from a GUI

❖ Available plugins include:

- Static Code Analysis with Pylint
- Code Profiling
- Conda Package Manager with Conda

❖ Hardware Interfaces

1. Processor : Intel CORE i5 processor with minimum 2.9 GHz speed.
2. RAM : Minimum 4 GB.
3. Hard Disk : Minimum 500 GB

❖ Software Interfaces

1. Microsoft Word 2007
2. Database Storage : Microsoft Excel
3. Operating System : Windows10

1.4 Motivation:

Since time immemorial, the Indian education system has been the envy of the world. Right since the time of the Indus Valley Civilisation, knowledge had been passed on from one generation to the other. What initially had the spoken form, soon became written texts. Scientific and mathematical marvels were achieved. Somewhere within this period, Aryabhatta discovered the number 0, a discovery that completely changed the world of mathematics. As millennia passed by, centres of learning were established and universities like Nalanda passed on the torch of education from one generation to the other.   Overtime, the Indian education system moved from strength to strength. Today half the software engineers in the United States are those of Indian origin. However, the other side of the story here is the fact that the highest number of academics-related suicides happen in this country. With every passing year, more and more students are falling prey to depression and anxiety issues.

1.5 Objective :

In the present study, we present a proof of concept for such a machine vision-based approach by using manual .Our objective is to improve online education system so that we don't have to hire anyone to observe the students in the class. With this system , the higher authorities will get to know who works better and what improvements we can do for those students who got depressed. With help of this system we can easily knows that which student are suffering from mental conditions and who got depressed . we can arrange the councils for these students to help them for being a good society.

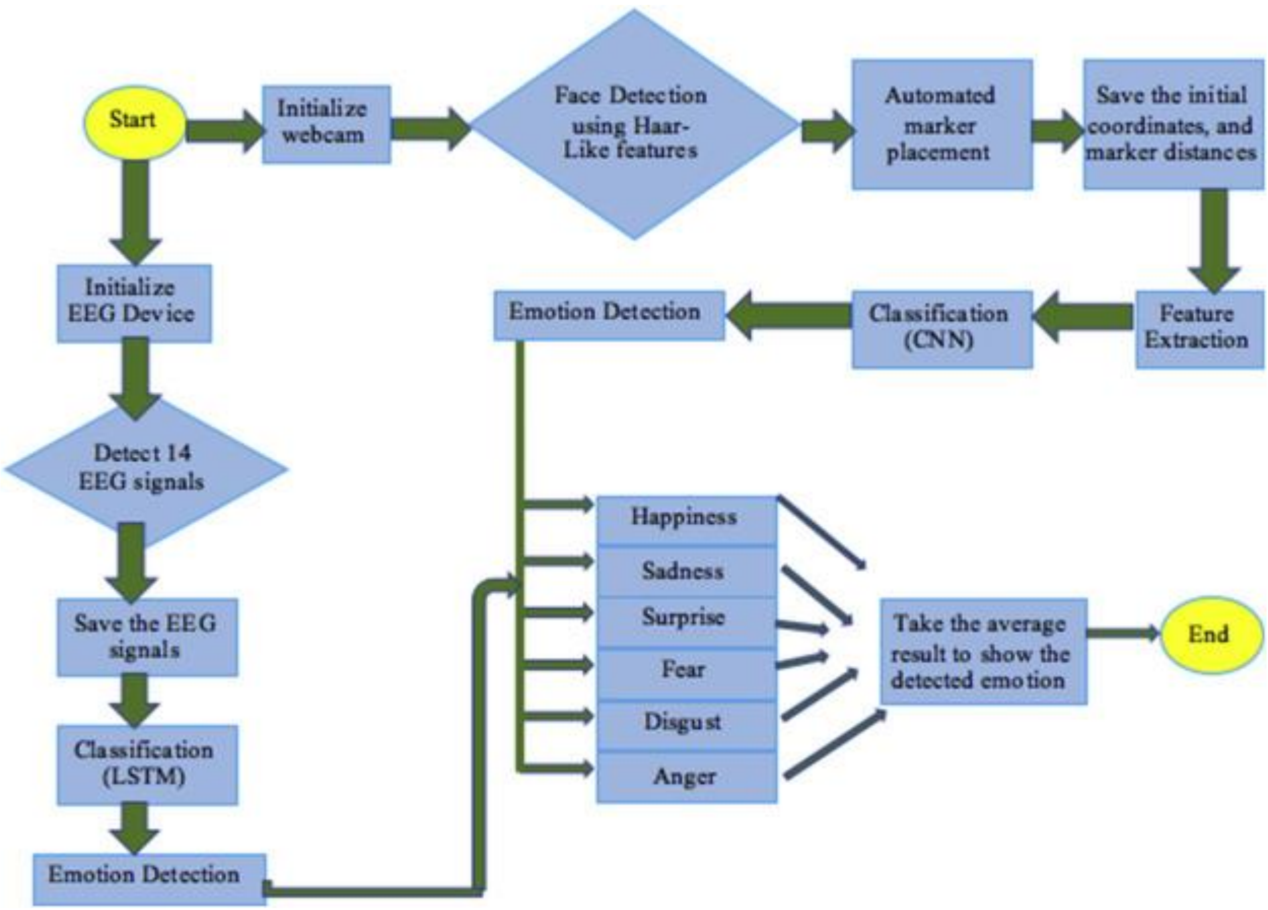
1.6 features :

In the present study, we present a proof of concept for such a machine vision-based approach by using manual ratings of visible indicators of students' (dis)engagement in learning as a basis for the automated analysis of pilot classroom recordings of three lessons with university students. More specifically, by combining multiple indicators from previous research (i.e., Chi and Wylie 2014; Helmke and Renkl 1992; Hommel 2012), we developed a manual rating instrument to continuously measure students' observable behavior. In addition, we performed an automated analysis of the video recordings to extract features of the students' head pose, gaze direction, and facial expressions using modern computer vision techniques. Using these automatically extracted features, we aimed to estimate manually annotated attention levels for each student. Because we had continuous labeling, this could be done by training a regressor between the visible features and the manual labels. We investigated the predictive power of both the manual and automatic analyses for learning (i.e., performance on a subsequent knowledge test). To account for complexity within classrooms and enrich the automated analysis, we also considered synchronous behavior among neighboring students. In the present article, we report initial empirical evidence on the reliability and validity of our automated assessments and their association with student performance.

CHAPTER-2

Literature Survey

2.1 Data Flow



2.2 **Present System:**

In sum, considering learning content has been shown to be useful; however, there is a lack of research involving the continuous analysis of attention or engagement over the course of one or more lessons. A unique feature of the present study is that we aimed to acquire a continuous assessment (i.e., a score for every student in the classroom for every second of instruction time) of students' visible indicators of (dis)engagement in learning. This temporal resolution was crucial in our approach because we aimed to provide comparable data that could be used to train a machine-learning algorithm. To reach this high level of temporal resolution, we decided to annotate learners' behavior continuously. The free software CARMA (Girard 2014) enables the continuous interpersonal behavior annotation by using joysticks (see Lizdek et al. 2012). However, this new approach limited us in terms of using already existing rating instruments because existing instruments do not allow for a high enough level of temporal resolution. Furthermore, the CARMA software requires annotations on a scale rather than rating the behavior in terms of categories as already existing instruments do. When developing the new instrument, we mainly oriented on the MAI (Helmke and Renkl 1992; Hommel 2012). However, we needed to define more fine-grained indicators of student behavior to make annotations along a continuous scale possible. Therefore, we added indicators from various established instruments to extend our rating scale. We assumed that the manual observer annotations would serve only as approximations of the actual cognitive states of the students and that the averaged (i.e., intersubjective) manual annotations would reflect the “true score” of the visible indicators of (dis)engagement in learning better than a single rater could. Subsequent to the ratings, we thus calculated the mean of the raters for every second. The mean values for each second and student were used as the ground truth to train a machine-learning approach.

The difficulty in assessing students’ engagement-related processes in real-world classroom settings consists of externalizing learners’ internal (covert) states through visible overt aspects to the greatest extent possible. In psychology, affective states and cognitive processes such as attentional control are usually determined from physiological signals, such as heart rate, electrodermal activity, eye tracking, or electroencephalography. Using this kind of psychologically sound measurements makes it possible to detect covert aspects of learningrelated processes; however, these measures are hardly feasible in classroom instruction, especially when teachers must be equipped with knowledge about what indicators to look for in students. Furthermore, these approaches are useful for answering very specific research questions. However, they are not sufficient for determining whether students’ ongoing processes are actually the most appropriate for the situation. By contrast, overt behavior can provide visible indicators of appropriate learning-related processes in students.

2.3 **PROPOSED SYSTEM:-**

Here we will detect the person and find what activity they are doing on sitting in class. Here we are going to mark the each class and each student with different marks which are based on the activity they are doing .

- If the student are just listening then we will give him 0 marks or we can say the student is concentrated .
- If the student are raising their hands on teacher call or to ask their doubt to the teacher we will give them +1 marks.
- If the student are talking with teacher that means the student is fully engaged with his study so we will give him +2 marks.
- If the students are not looking forward then we will give him -1 marks , because it is a sign that student are not concentrated on its study.
- If the student leaves the meeting or walk around from the front camera, that means he is doing another work and not focused for study.

<u>On bases of that we will mark each and every class every day and find out which student is less concentrated for studies and which teacher gives its best so that maximum class is engaged with it . with this model of education, teachers can do the teaching experiments very easily and get the results very efficiently and easily also.</u>

CHAPTER-3

System Requirement

3.1 OpenCV OpenCV is the library we will be using for image transformation functions such as converting the image to grayscale. It is an open source library and can be used for many image functions and has a wide variety of algorithm implementations. C++ and Python are the languages supported by OpenCV. It is a complete package which can be used with other libraries to form a pipeline for any image extraction or detection framework. The range of functions it supports is enormous, and it also includes algorithms to extract feature descriptors.

3.2 Dlib Dlib is another powerful image-processing library which can be used in conjunction with Python, C++ and other tools. The main function this library provides is of detecting faces, extracting features, matching features etc. It has also support for other domains like machine learning, threading, GUI and networking.

3.3 Python Python is a powerful scripting language and is very useful for solving statistical problems involving machine learning algorithms. It has various utility functions which help in preprocessing. Processing is fast and it is supported on almost all platforms. Integration with C++ and other image libraries is very easy, and it has in-built functions and libraries to store and manipulate data of all types. It provides the pandas and numpy framework which helps in manipulation of data as per our need. A good feature set can be created using the numpy arrays which can have n-dimensional data.

3.4 Scikit-learn Scikit-learn is the machine learning library in python. It comprises of matplotlib, numpy and a wide array of machine learning algorithms. The API is very easy to use and understand. It has many functions to analyze and plot the data. A good feature set can be formed using many of its feature reduction, feature importance and feature selection functions. The algorithm it provides can be used for classification and regression problems and their sub-types.

3.5 Jupyter Notebook Jupyter Notebook is the IDE to combine python with all the libraries we will be using in our implementation. It is interactive, although some complex computations require time to complete. Plots and images are displayed instantly. It can be used as a one stop for all our requirements, and most of the libraries like Dlib, OpenCV, Scikit-learn can be integrated easily.

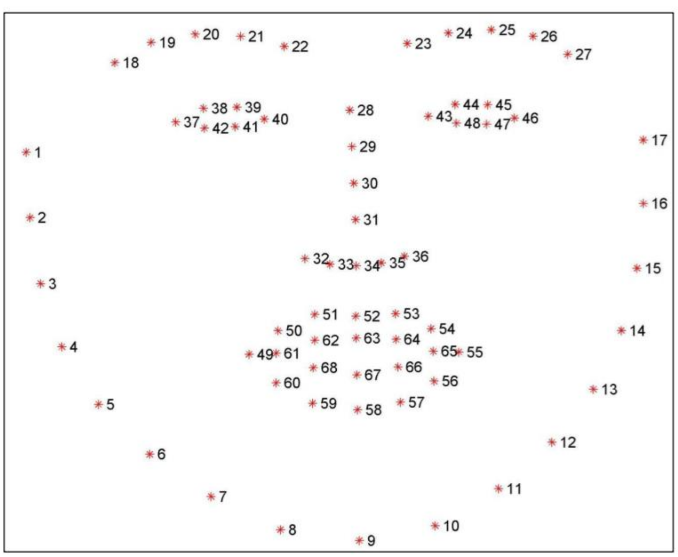
3.6 Database We have used the extended Cohn-Kanade database (CK+) and Radbound Faces database(RaFD). CK+ has around 593 images for 123 subjects. Only 327 files have labeled/identified emotions. It covers all the basic human emotions displayed by the face. The emotions and codes are as follows: 1 – Angry, 2 – Contempt, 4 – Fear, 5 – Happy, 6 – Sadness, 7 – Surprise. The database is widely used for emotion detection research and analysis. There are 3 more folders along with the images.

CHAPTER-4

Project Design:

Facial Action Coding System is used to give a number to facial moment. Each such number is called as action unit. Combination of action units result in a facial expression. The micro changes in the muscles of the face can be defined by an action unit. For example, a smiling face can be defined in terms of action units as 6 + 12, which simply means movement of AU6 muscle and AU12 muscle results in a happy face. Here Action Unit 6 is cheek raiser and Action Unit 12 is lip corner puller. Facial action coding system based on action units is a good system to determine which facial muscles are involved in which expression. Real time face models can be generated based on them.

Landmarks on the face are very crucial and can be used for face detection and recognition. The same landmarks can also be used in the case of expressions. The Dlib library has a 68 facial landmark detector which gives the position of 68 landmarks on the face.



Good features are those which help in identifying the object properly. Usually the images are identified on the basis of corners and edges. For finding corners and edges in images, we have many feature detector algorithms in the OpenCV library such as Harris corner detector. These feature detectors take into account many more factors such as contours, hull and convex. The Key-points are corner points or edges detected by the feature detector algorithm. The feature descriptor describes the area surrounding the key-point. The description can be anything including raw pixel intensities or co-ordinates of the surrounding area. The key-point and descriptor together form a local feature. One example of a feature descriptor is a histogram of oriented gradients. ORB (based on BRIEF), SURF, SIFT etc. are some of the feature descriptor algorithms

This method uses cascaded regression trees and finds the important positions on the face using images. Pixel intensities are used to distinguish between different parts of the face, identifying 68 facial landmarks . Based on a current estimate of shape, parameter estimation is done by transforming the image in the normal co-ordinate system instead of global. Extracted features are used to re-estimate the shape parameter vectors and are recalculated until convergence.

CHAPTER-5

Implementation/working of project

Sample and Procedure

We decided to conduct a study involving university students in order to validate our approach before administering it in school classrooms. A total of $N = 52$ university students (89.5% women, 8.8% men, mean age = 22.33, $SD = 3.66$) at a German university volunteered to take part in the study. The study was conducted during regular university seminar sessions on quantitative data analysis (90 min). A total of three different seminar groups were assessed. The topics of the sessions were either *t tests for independent samples* (sessions 1 and 2) or *regressions* (session 3) and ranged from 30 to 45 min. The sessions were videotaped with three cameras (one teacher camera, two cameras filming the students). If students refused to be videotaped, they were either seated outside the scope of the cameras or switched to a parallel seminar. Participants were informed in advance of the study's purpose, procedure, and ethical considerations such as data protection and anonymization. To avoid confounding effects of the teacher, the same person taught all sessions in a teacher-centered manner. Before the session started, students filled out a questionnaire on background variables (age, gender, final high school examination [Abitur] grade, school type) and individual learning prerequisites. After the session, participants completed a knowledge test on the specific topic of the session and completed another questionnaire about learning activities during the seminar.

Instruments

Individual Learning Prerequisites

We used established questionnaire measures to assess three individual learning prerequisites: Dispositional interest in the session's topic was captured with four items ($\alpha = .93$) adapted from Gaspard et al. (2017). Self-concept in quantitative data analysis was assessed with five items ($\alpha = .80$; adapted from Marsh et al. 2006), and 13 items were used to test for self-control capacity ($\alpha = .83$; Bertrams and Dickhäuser 2009). Moreover, we administered the short version of the quantitative subscale (Q3) of the cognitive abilities test (Heller and Perleth 2000). Measuring these learning prerequisites allowed us to control for potential confounding variables in the analyses.

Learning Outcomes

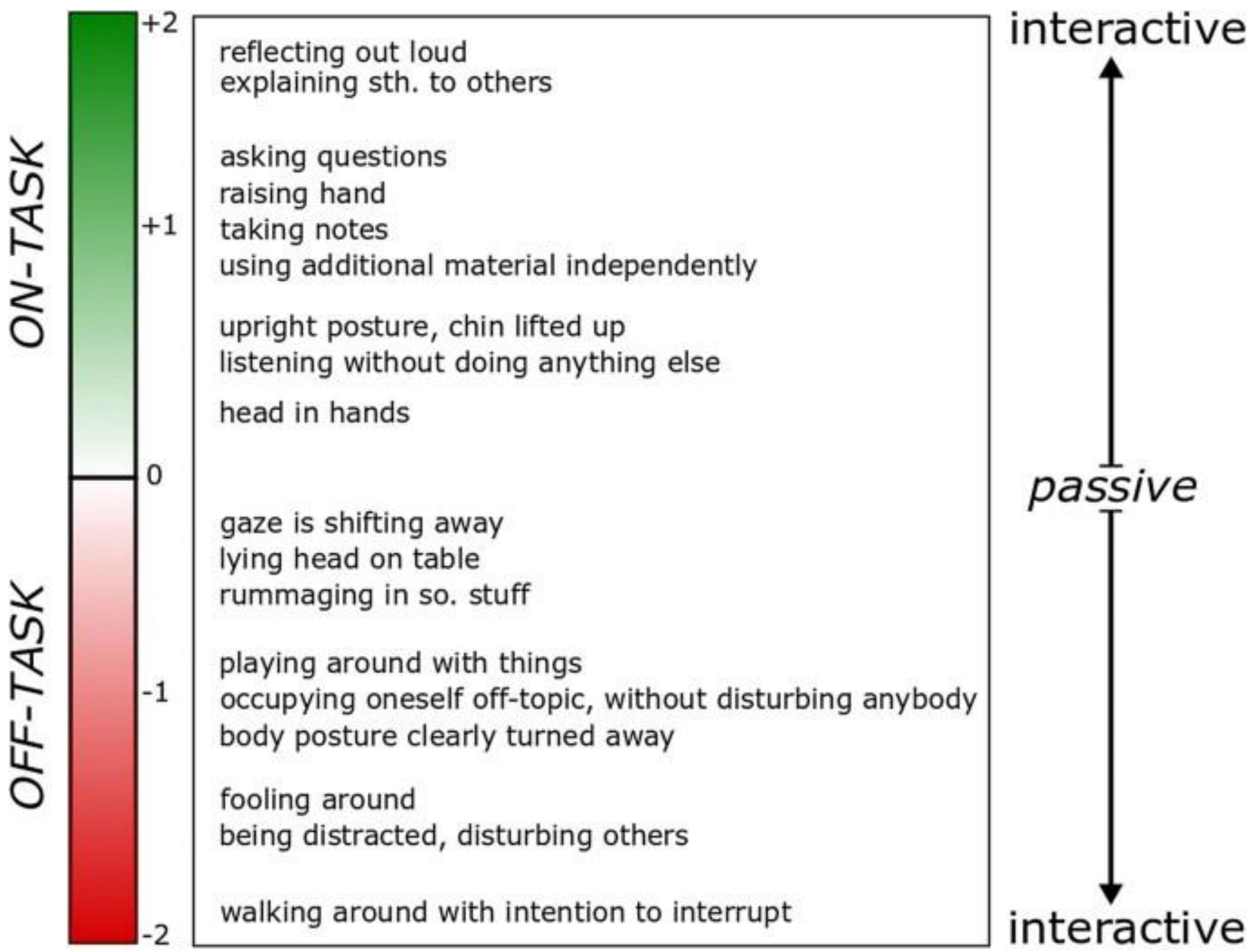
The knowledge test consisted of 12 and 11 items that referred to participants' declarative and conceptual knowledge of the session topic, respectively. We *z*-standardized the knowledge test scores within each group for subsequent analysis.

Self-Reported Learning Activities

After the session, we assessed students' involvement (four items, α = .61; Frank 2014), cognitive engagement (six items, α = .79; Rimm-Kaufman et al. 2015), and situational interest (six items, α = .89; Knogler et al. 2015) duringthe seminar session.

Analysis Continuous Manual Annotation

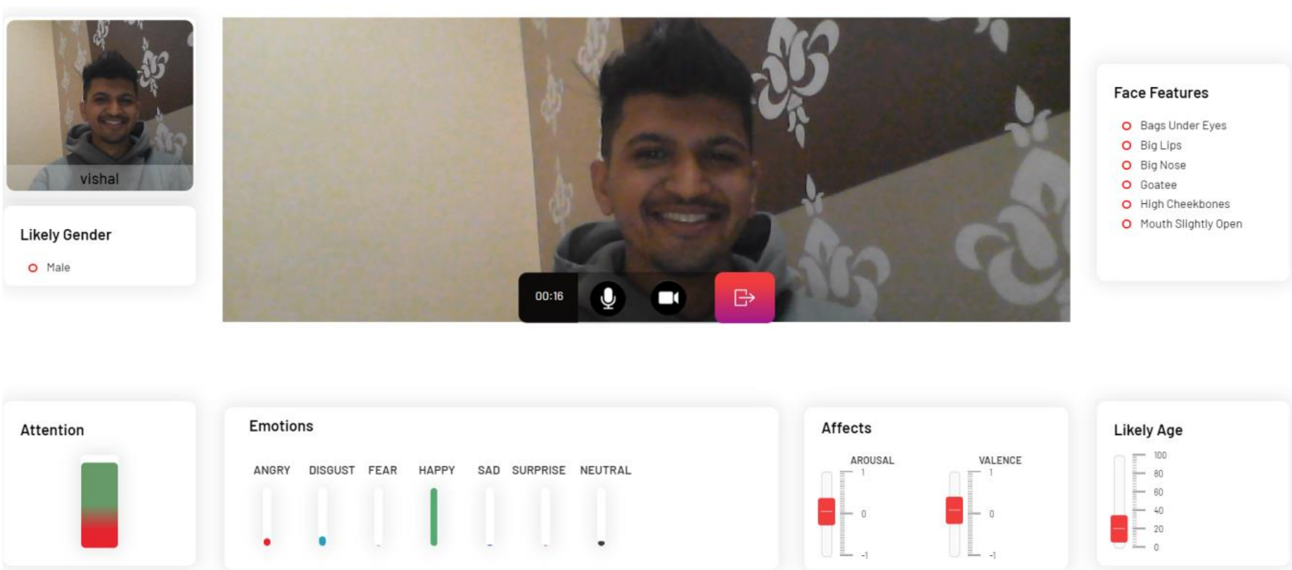
Todevelop a continuous manual annotation that included potential valid indicators of students' visible (dis)engagement in learning, we used the instruments developed by Helmke and Renkl (1992) and Hommel (2012) as a basis. However, these instruments label behavior in categories and thus cannot be used as a continuous scale. Therefore, we combined the idea of on-/off-task behavior and active/passive subcategories with existing scales from the engagement literature. Furthermore, we used the theoretical assumptions about students' learning processes and related activities in classrooms pointed out by the ICAP framework (Chi and Wylie 2014) as an inspiration to define more fine-grained differentiations within the possible behavioral spectrum. The distinction into passive, active, constructive, and interactive behavior allowed us to make subtler distinctions between the different modes of on-task behavior, and this concept could be transferred to off-task behavior (i.e., passive, active, deconstructive, and interactive) as well. By combining different approaches, we could define visible indicators of (dis)engagement in learning on a continuous scale. The resulting scale ranged from − 2, indicating interruptive and disturbing off-task behavior, to +2, indicating highly engaged on-task behavior where, for example, learners ask questions and try to explain the content to fellow learners (see Fig. 1). When a person could not be seen or was not present in the classroom, the respective time points were coded as missing values in subsequent analyses.

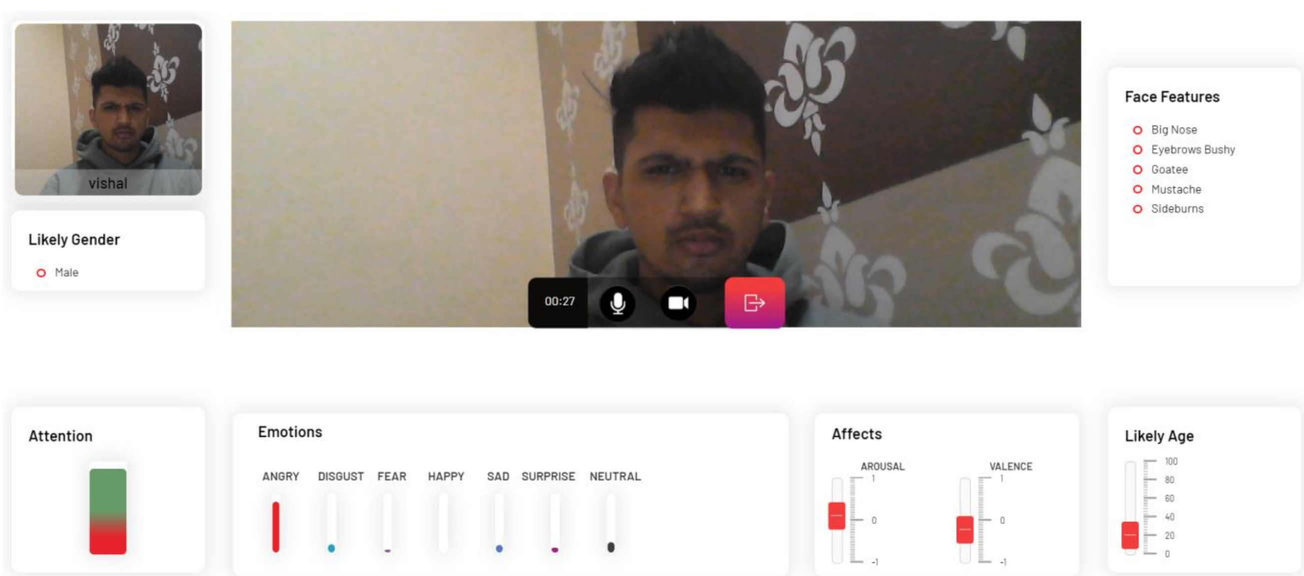


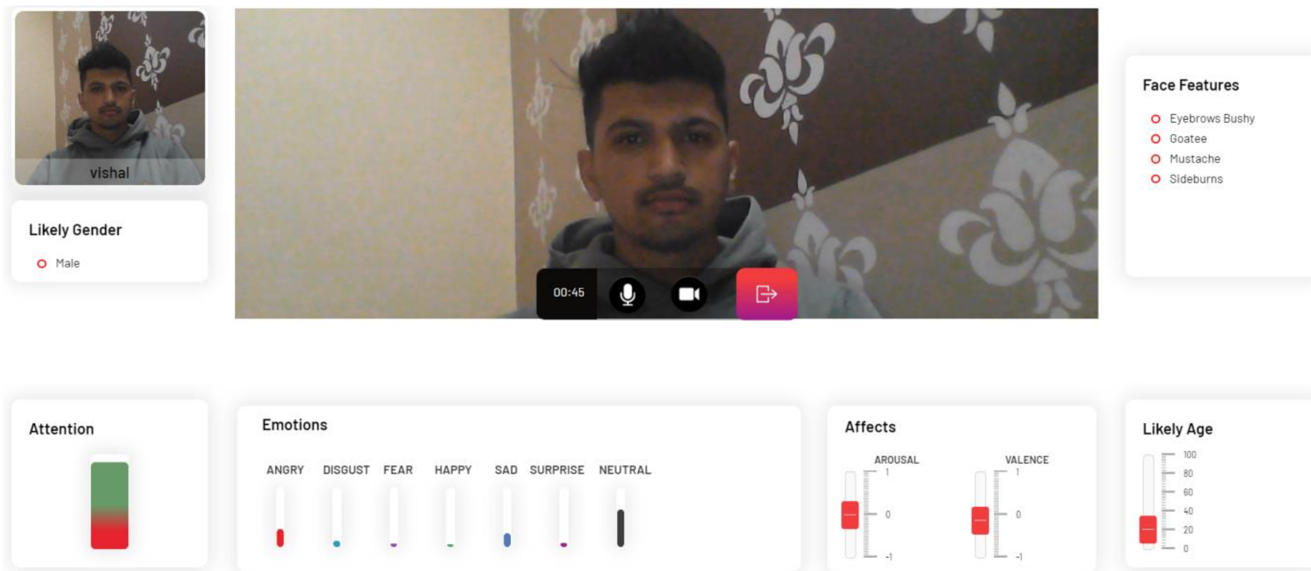
he behavior of each observed person throughout the instructional session was coded in 1-s steps using the CARMA software (Girard 2014) and a joystick. A total of six raters annotated the videotaped seminar sessions, and each session was annotated by a total of three raters. The raters consisted of student assistants and one researcher, all of whom were trained carefully before annotating the videos. First, raters were introduced to the conceptual idea of the rating and the rating manual. They were told to concentrate on observable behavior to avoid making inferences and considering information from previous ratings. The raters focused on one student at a time in a random order. Every rater had to code one of two specific sections of the video for training, and the raters had to annotate special students who showed different types of behavior. To ensure that we could use all the video material for our analysis, raters who used video section A for training annotated video section B later and vice versa. The respective video sections used for training purposes were not included in the analysis. Only after their annotations reached an interrater reliability with an expert rating of at least ICC(2,1) = .60 were raters allowed to annotate the study material. We report the ICC(2,1) here as an indicator of interrater reliability because our data were coded on a metric scale level, and we had more than two raters per participant. We calculated the ICC(2,1) for every student, indicating the interrater reliability averaged across all time points, whereby values between .60 and .74 indicated good interrater reliability (Hallgren 2012); the ICC (2,1) for each student was .65 on average (absolute agreement). When the annotations between the raters deviated strongly, critical situations were discussed among the raters and recoded following consensus. The raters were not informed about the students' individual prerequisites, their learning outcomes, or their self-reported learning activities.

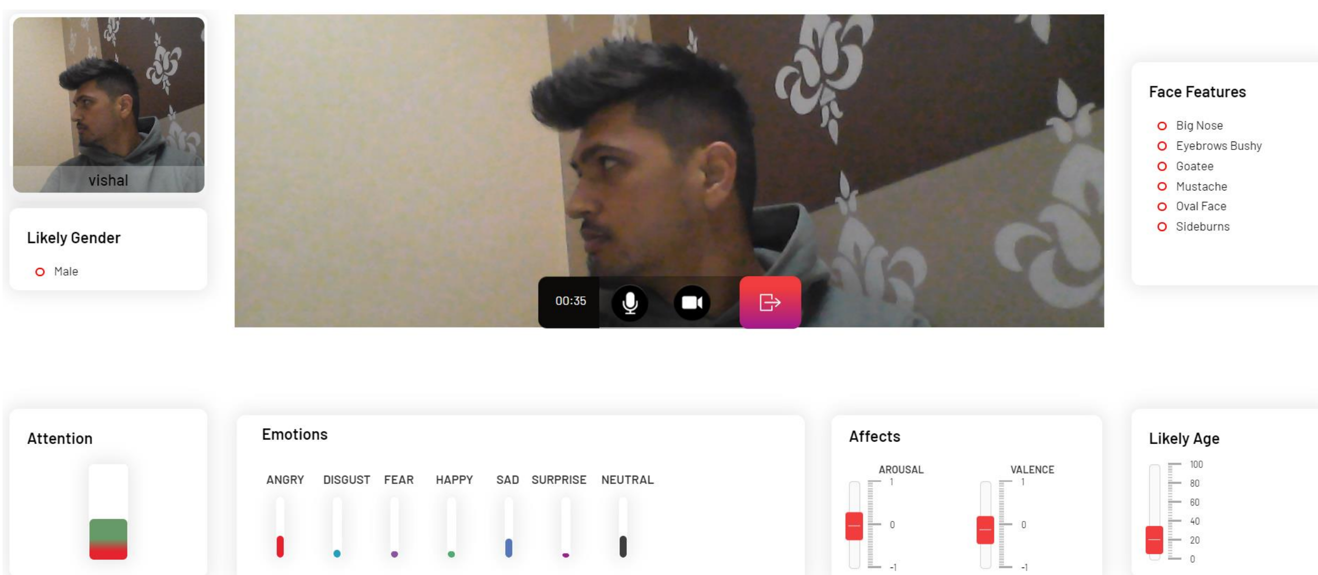
CHAPTER-6

Result/output






vishal
Likely Gender
Male
00:45
Face Features
Eyebrows Bushy
Goatee
Mustache
Sideburns
Attention
Emotions
ANGRY
DISGUST
FEAR
HAPPY
SAD
SURPRISE
NEUTRAL
Affects
AROUSAL
VALENCE
Likely Age


vishal
Likely Gender
Male
00:35
Face Features
Big Nose
Eyebrows Bushy
Goatee
Mustache
Oval Face
Sideburns
Attention
Emotions
ANGRY
DISGUST
FEAR
HAPPY
SAD
SURPRISE
NEUTRAL
Affects
AROUSAL
VALENCE
Likely Age

CHAPTER-7

Future scope

We tested the validity of our manual rating instrument in two steps. First, we investigated construct validity by correlating the manual ratings with the self-reported learning activities. The manual annotations were significantly correlated with students' self-reported cognitive engagement, situational interest, and involvement ($.49 \leq r < .62$).

Additionally, we calculated a multiple linear regression with the three self-reported learning activities as regressors. Together, they explained 42.9% of the variance in the manual ratings.

This corresponds to a multiple correlation of $r = .66$. Second, we examined the predictive validity of our new instrument. We inspected the intercorrelations between all variables with the knowledge test . The knowledge test scores (the dependent variable in this study) were significantly correlated with the manual ratings, cognitive abilities, and situational interest ($.30 \leq r < .42$). To test for effects of possible confounding variables, we calculated two additional linear regression models in which we added background variables (model 2) and learning prerequisites (model 3) into the regression and compared them with the prediction that involved only manual ratings . The effect of the manual ratings remained robust and still explained a significant proportion of the variance in the knowledge test results.

We applied our trained regression to test subjects at 1-s intervals and applied mean pooling to create a final estimation that summarized participants' engagement. Table 4 shows the performance of different modalities for estimating (dis)engagement in learning. The performance measures were mean squared errors in the regression and the Pearson correlation coefficient between the manual annotations' mean level and our models' prediction during the instructional session. As shown in Table 4, the head pose modality exhibited a lower correlation with the manual ratings ($r = .29$) than the other features. By contrast, gaze information and facial expressions (AU intensities) were more strongly correlated with the manual annotations ($r = .44$). Combining head pose and gaze ($r = .61$) or all three modalities ($r = .61$) also led to substantial correlations with the manual annotations. In addition, we tested the correlations between the posttest variables (i.e., the knowledge test and self-reported learning activities) and the different models for estimating the manual ratings .

According to these results, regression models, which perform better with respect to MSE and lead to higher correlations with the manual ratings, seem to contain more information that is relevant for the posttest variables, particularly with respect to involvement and cognitive engagement. The cosine similarities of the manual annotations between neighboring students were strongly

correlated with each neighbor's mean engagement level throughout the recording ($r = .78$). More specifically, taking the synchronization into consideration improved the correlation with the manual ratings by 9%, thus showing that synchronization information is helpful for understanding (dis)engagement in learning.

The correlations between the different models for estimating the manual ratings and students' self-reported learning activities and outcomes revealed that the best models were those in which head pose and gaze features were combined with neighbor synchrony ($r = .08$, .43, .39, and .26 for the knowledge test, involvement, cognitive engagement, and situational interest, respectively; Table 5). We calculated the mean correlation (based on Fisher's z-transformed correlations) of the three manual annotations (average $r = .74$) and the mean correlation of each rater and the scores from a model combining head pose, gaze features, and neighbor synchrony (average $r = .64$) for the subsample.

Because the model in which head pose and gaze were combined with neighbor's synchrony had the highest correlation with the manual rating, we calculated a linear regression to predict the posttest variables (Table 6). In order to understand the contribution of neighbor's synchrony, we trained our regression models using the same features with and without synchronization information. Adding neighbor's synchrony improved the prediction of all posttest variables and explained at least 2% more variance. However, the manual rating remained superior.

CHAPTER-7

Conclusion:

Remote approaches from the field of computer vision have the potential to support research and teacher training. For this to be achieved, valid visible indicators of students' (dis)engagement in learning are needed. The present study provides a promising contribution in this direction and offers a valid starting point for further research in this area.

The present study reported key initial results from the development of a machine vision-based approach for assessing (dis)engagement in the classroom. We were able to find empirical support for the validity of our newly developed manual rating instrument. Furthermore, the machinelearning approach proved to be effective, as shown by its correlation with the manual annotations as well as its ability to predict self-reported learning activities. Finally, as expected, including an indicator of synchrony in the automated analyses further improved its predictive power. Next, we discuss our main results in more detail before turning to the limitations of the present study and the crucial next steps.

The manual rating of visible indicators for (dis)engagement in learning predicted achievement on a knowledge test following a university seminar session. This prediction was robust when we controlled for individual characteristics (research question 1). In terms of validity, self-reported cognitive engagement, involvement, and situational interest were strongly correlated with the manual rating. As these self-reported learning activities reflect students' cognitive processes during the seminar session, we concluded that our manual ratings capture visible indicators that are actually related to (dis)engagement in learning. Therefore, we inferred that it is reasonable to use these manual ratings as a ground truth for our machine vision-based approach.

In the automated analyses of engagement, we used several visible features (head pose, gaze, facial expressions). More specifically, we compared their contribution with visible indicators of (dis)engagement in learning separately and in combination. Our results showed that facial expressions were more strongly correlated with the manual rating than head pose or gaze alone; however, combining the engagement-related features and combining all three visible indicators improved the correlation with the manual annotations substantially, thus emphasizing the complexity of human rating processes. However, we were not able to replicate the prediction of the knowledge test scores by considering these visible features alone

CHAPTER-9

Future scope

The present study lays the basis for achieving these goals by developing and testing an automated approach to assessing visible indicators of students' (dis)engagement in learning. Such a remote approach requires comparable data (generated by human raters) that can be used as the ground truth in order to train a classifier. However, existing instruments (Helmke and Renkl 1992; Hommel 2012) for measuring engagement-related processes in learning (a) require human observers to make a huge number of inferences and (b) require data to be collected in 30-s or 5-min intervals. This is problematic for our context because an automated analysis can only rely on visible indicators, does not consider content-specific information at all, and operates at a more fine-grained temporal resolution. Therefore, we developed a new instrument to annotate student behavior manually by applying a rating method with visible indicators over time. This manual rating served as the starting point from which to train an algorithm by applying methods from machine learning and computer vision.

1. Is the new manual annotation of visible indicators of (dis)engagement in learning related to students' learning processes and outcomes? To validate our instrument, we examined how the manual ratings were correlated with students' self-reported cognitive engagement, involvement, and situational interest. We expected these self-reported

learning activities to cover different facets of (dis)engagement in learning, and when combined, we expected them to account for cognitive parts of the construct. Furthermore, we tested whether the scores resulting from the manual annotation would predict students' performance on a knowledge test at the end of an instructional session.

2. Is it possible to adequately replicate the relation to students' learning processes and outcomes by using visible indicators of (dis)engagement in learning based on the machine-learning techniques that estimated the manual ratings? We used gaze, head posture, and facial expressions to estimate the manual ratings. To test the quality of our machine vision-based approach, we examined the associations between the scores generated from the automated approach and the manual ratings and students' self-report data regarding their learning processes, and we used the machine-learning scores to predict achievement on the knowledge test.
3. How do adding synchrony aspects of student behavior affect the automated estimations of the manual ratings? The results of previous studies have indicated that immediate neighbors have a significant influence on a student's engagement (Raca and Dillenbourg 2013; Raca et al. 2013). As a first step toward including indicators of synchrony in our project, we added students' synchrony with the person sitting next to them as an additional variable to our prediction models, which were based on the automated assessment of student engagement.

CHAPTER-6

References

1. Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). Off-task behavior in the cognitive tutor classroom: when students "game the system". Conference on Human Factors in Computing Systems, 383–390.

2. Baltrušaitis, T., Zadeh, A., Lim, Y. C., & Morency, L.-P. (2018). OpenFace 2.0: facial behavior analysis toolkit. Paper presented at the 13th IEEE International Conference on Automatic Face & Gesture Recognition, ake Placid, NY.

https://doi.org/10.1109/WACV.2016.7477553.

3. Berliner, D. C. (2001). Learning about and learning from expert teachers. International Journal of Educational Research, 35, 463–482.

4. Bertrams, A., & Dickhäuser, O. (2009). Messung dispositioneller SelbstkontrollKapazität. Diagnostica, 55(1), 2–10. https://doi.org/10.1026/0012-1924.55.1.2.

5. Article Google Scholar

6. Bosch, N., D'Mello, S. K., Baker, R. S., Ocumpaugh, J., Shute, V., Ventura, M., . . . Zhao, W. (2016a). Detecting student emotions in computer-enabled classrooms. The Twenty-Fifth International Joint Conference on Artificial Intelligence, 4125-4129.

7. Bosch, N., D'Mello, S. K., Ocumpaugh, J., Baker, R. S., & Shute, V. (2016b). Using video to automatically detect learner affect in computer-enabled classrooms. ACM

Transactions on Interactive Intelligent Systems, 6(2), 1–26.

https://doi.org/10.1145/2946837.

8. Brünken, R., & Seufert, T. (2006). Aufmerksamkeit, Lernen, Lernstrategien. In H. Mandl & H. F. Friedrich (Eds.), Handbuch Lernstrategien (pp. 27–37). Göttingen: Hogrefe.

9. Google Scholar

10. Büttner, G., & Schmidt-Atzert, L. (2004). Diagnostik von Konzentration und Aufmerksamkeit. Hogrefe Verlag.

11. Carroll, J. B. (1963). A model for school learning. Teachers College Record, 64(8), 723– 733.

12. Google Scholar

13. Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: linking cognitive engagement to active learning outcomes. Educational Psychologist, 49(4), 219–243.

https://doi.org/10.1080/00461520.2014.965823.

14. Connell, J. P. (1990). Context, self, and action: a motivational analysis of self-system processes across the life span. In D. Cicchetti & M. Beeghly (Eds.), The self in transition:

infancy to childhood (Vol. 8, pp. 61–97). Chicago: The University of Chicago Press.

15. Cortina, K. S., Miller, K. F., McKenzie, R., & Epstein, E. (2015). Where low and high inference data converge: validation of CLASS assessment of mathematics instruction using mobile eye tracking with experts and novice teachers. International Journal of Science and Mathematics Education., 13(2), 389–403.

https://doi.org/10.1007/s10763014-9610-5.

16. Cumming, G. (2014). The New Statistics: Why and how. Psychological Science, 25(1), 7–29. https://doi.org/10.1177/0956797613504966.

17. Driver, J. (2001). A selective review of selective attention research from the past century. British Journal of Psychology, 92(1), 53–78. https://doi.org/10.1348/000712601162103.

****