Research Article
Towards a Low-Cost Teacher Orchestration Using Ubiquitous Computing Devices for Detecting Student’s Engagement

Ibrar Ahmad,1 Shah Khusro1, Iftikhar Alam2, Inayat Khan3, and Badam Niazi4

1Department of Computer Science, University of Peshawar, Peshawar 25000, Pakistan
2Department of Computer Science, City University of Science and Information Technology, Peshawar 25010, Pakistan
3Department of Computer Science, University of Buner, Buner 19290, Pakistan
4Department of Computer Science, University of Nangarhar, 2600, Afghanistan

Correspondence should be addressed to Badam Niazi; niazi5.48@gmail.com

Received 7 April 2022; Accepted 13 May 2022; Published 18 June 2022

Academic Editor: Zahid Khan

Copyright © 2022 Ibrar Ahmad et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The ubiquitous devices and technologies to support teachers and students in a learning environment include the Internet of things (IoT), learning analytics (LA), augmented or virtual reality (AR/VR), ubiquitous learning environment (ULE), and wearables. However, most of these solutions are obtrusive, with substantial infrastructure costs and pseudo-real-time results. Real-time detection of students’ activeness, participation, and activity monitoring is important, especially during a pandemic. This research study provides a low-cost teacher orchestration solution with real-time results using off-the-shelf devices. The proposed solution determines a teacher’s activeness using multimodal data (MMD) from both teacher and student’s devices. The MMD extracts different features from data, decodes them, and displays them to the instructor in real time. It allows the instructor to update their teaching methodology in real time to get more students on board and provide a more engaging learning experience. Our experimental results show that real-time feedback about the classroom’s current status helped improve learning outcomes by about 45%. Also, we investigated a 50% increase in classroom engaging experience.

1. Introduction

This pervasive development in technology has made computers more robust and smaller. The computers successfully made their way from giant PCs to small portable mobile devices, which created a new era of ubiquitous computing that made accessible computers anywhere with more excellent perception and understanding of the surrounding environment through sensors [1]. Mark Weiser coined the term “ubiquitous” in the 90s [1], resulting in several smart devices including smartphones, smartwatches, and smart TV [2] being used in various scenarios. For example, wearable devices (smartwatches, bands, etc.) are advantageous in health-related applications and others. Users’ gestures are required as they are constantly connected to the skin and fixed on the human body [3]. Like other fields, including business, health, and entertainment, these devices have more potential to be efficiently and effectively exploited to improve education quality. One such prominent example is smart wearable devices for teacher orchestration [4]. These technologies are utilized for teaching, learning, and orchestration in a learning environment.

Teacher orchestration refers to managing different classroom activities encompassing individual, small group, and whole class in a face-to-face classroom by a teacher [5]. The word orchestration came from orchestra, which means carefully organizing a complicated event [6]. In the context of a smart classroom, teacher orchestration is the careful arrangement of a technologically more prosperous classroom environment and activities to achieve the required learning outcomes [6]. The main focus is facilitating a teacher in monitoring and healthier students’ performance in a ubiquitous learning environment (ULE) [7]. In ULE, small and handheld devices perform various monitoring and data visualization tasks to support teachers’ and students’ learning pedagogies. A typical classroom contains multiple kinds of activities. A teacher must manually manage several paper-based activities, such as taking attendance
by calling students’ names and marking them present or absent. It can be technology-assisted by using different sensors and other devices [8]. Traditional orchestration is more ubiquitous in most institutes because it is easy to use and requires less training.

Still, time consumption and resource wastage are among the expected downsides of this approach. Technological devices were employed to assist teachers and students in the learning process to overcome these issues, which created the smart classroom concept. Smart classrooms are technologically rich learning spaces where computers and other devices are exploited to help teachers and students [9]. Although this approach helped overcome these issues, there are other complications regarding user acceptance and development costs due to giant infrastructures. Several custom-built hardware with multiple sensors are used [10], increasing the setup cost and affecting user experience and social acceptance. Therefore, smartphones were deployed instead of using custom-built hardware and equipped with different sensors [11–13]. Using handheld devices, a new era of ULEs evolved, transforming the educational context into complex social and technological ecologies by expanding the scope of education beyond the classroom [7].

Several studies proposed numerous approaches to perform the orchestration process using multimodal data from multiple sources. These studies leverage different technologies, including the Internet of things (IoT) [10, 14, 15], intelligent tutoring systems (ITSs), learning dashboards [16], augmented and virtual reality (AR/VR) [17, 18], smart wearables [19], different sensors, and ubiquitous computing devices [7]. However, most of the proposed solutions are either not real time or expensive and less user-friendly because they need technical assistance to be used.

Most studies focused on custom-made hardware, which provided good results in some circumstances, such as lab environments [20]. Still, the setup cost and acceptability in the real-world classroom are a matter of concern. It requires technical assistance from paid experts or specially hired employees to deploy and use these clumsy infrastructures in a learning space. Using ubiquitous computing devices reduces setup costs, but they are not real time and provide results after the classroom session. This delay is becoming the main reason for time wastage for teachers and learners as they cannot adjust their behaviors in that specific session. To mitigate these issues, it is required to provide a real-time teacher orchestration solution using off-the-shelf and low-end devices, which is the paper’s main aim. The proposed solution is low-cost, easy-to-use, with a real-time feedback facility about the class’s current status. The main goal of this research work is to leverage the potential of off-the-shelf smart devices, including smartphones and smartwatches, in teacher orchestration to reduce the use of custom-made and specialized hardware, which increases the setup cost and requires technical assistance for deployment and usage.

This study attempts to avoid using any external server for data acquisition, processing, or result generation. Thus, it significantly reduces the cost and effort required for setting up and using the system in real-world classroom scenarios. The proposed solution needs a smartwatch, i.e., a wearable device worn by the teacher in their dominant hand and connected to a smartphone placed in front of the teacher. The connected smartwatch sends its sensor’s data to the connected smartphone, processing and analyzing for final result generation on the smartphone. The application collects data from both teachers and students. The facial orientation of the teacher is used to measure her activeness or tiredness level. A server application is developed and deployed on the teacher’s smartphone to collect and process the multimodal data from the teacher’s and students’ devices. The processed data is displayed on the teacher’s smartphone showing statistics about the current state of the class, e.g., how many students are active and inactive and what is his voice quality during the lecture. Results show a significant increase in learning outcomes, i.e., a 45% increase. Also, we investigated a 50% increase in classroom engagement. The gathered data shows that this solution is less intrusive and has no serious issues for students and teachers. Also, the system can be applied in other lecture-demonstration methods.

The rest of the paper is divided into six sections. Section 2 is a comprehensive yet concrete literature review. Section 3 is the proposed methodology that elaborates the technical aspects. Section 4 is the implementation of the system. Section 5 is the experimental setup section. Section 6 shows results and analysis that further discusses the obtained results. The last section is the conclusion section. The references are numbered in the last section.

2. Literature Review

Mobile and computer technology have been introduced into educational contexts over the past two decades [21]. Access to computers and large-scale one-to-one computing programs have been implemented in several countries globally [22–24], such that elementary and middle school student(s) and teacher(s) have their electronics and mobile devices. In terms of encouraging and promoting innovation and modernization in education through mobile and information technology (IT), it also supports traditional lecture-style teaching, convenient information gathering, and information sharing and promotes innovative teaching methods such as cooperative learning [25, 26], exploratory learning outside the classroom, and game-based learning [27]. On the flip side, the marvelous expansion of sensor technology in smartphone(s), along with their sensing capabilities for accurate capturing, monitoring, and analysis of information, helps us know about traffic conditions [28–30], road conditions [31–33], environmental impacts of noise level [34], air quality and pollution level [35–39], humidity and temperature [40], understand patterns of objects movements [41–43], alerting and monitoring disaster [44, 45], weather information [46], etc.

The IT and mobile technologies can facilitate and enable innovative educational methods. Simultaneously, these patterns in educational practices will likely help subject content learning and facilitate the development of communication, problem-solving, creativity, and other high-level skills among students [20]. Also, it will support teachers in orchestrating different classroom activities and increase the
Assessments during Class. Student monitoring and engagement are positively linked with the required learning outcome. For instance, good grades in curricular and extracurricular activities are directly linked to critical thinking and the efficiency of the subject(s) [61]. Being a teacher is one of the most important factors for student(s) engagement and attention [62]. Teachers’ coordination and proper communication facilitated by a verbal, gestural, and written connection with their student(s) can benefit the student(s) memorization and attention. Classroom monitoring can be considered a powerful tool to determine the quantity and quality of active learning in the classrooms [63]. Monitoring activities lead to many engagement improvements, e.g., to improve learning [64], engagement to improve throughput rates and retention [65, 66], engagement for equality/social justice [67], and curricular relevance [68]. Submissive to the important monitoring and engagement, different tools [69], technologies [54], algorithm(s) [70], and strategies have been used to measure and estimate the attention level of both student(s) and teacher(s).

According to [71], only 46% to 67% of the students pay positive attention to the class during lecture delivery. It means half of the students could never be productive. With this information in hand, both the teachers and researchers have examined potential problems that arise during their classes, and efforts have been made to eradicate and correct them, which may have a long-term benefit on the learning efficiency of the learner and students. The study also showed that students’ engagement and focus are positively linked with good grades and critical thinking [61]. It is only possible with full attention and focus, which depends on numerous elements and factors, including the teacher [62]. According to [72], a classroom’s size influences student attention and engagement. In large classes, the teacher needs to use more time to draw students’ attention, which is sometimes emotionally exhausting.

Face detection, face recognition, facial features, pose estimation, etc. techniques have been used for student monitoring, for instance, student attendance monitoring system based on deep learning [73, 74], tracking through eye tracking [75], monitoring meeting through head orientation, and gaze direction [76], assessing and monitoring classroom attention [77], and estimation of activeness, transcribing, unavailing, distracted and transition, automatic recognition of engagement from students’ facial expressions [78].

State-of-the-Art Orchestration Solutions. According to Chan, “orchestration” is derived from orchestra in teacher orchestration [79]. Each student interacts with a digital device in a smart classroom to support them in the learning process. A smart classroom is an intelligent learning space equipped with different devices, sensors, and custom software agents [19]. Leeuwen and Rummel [80] reviewed various orchestration tools for teachers to help them understand students’ collaboration in their groups. Smart wearables were also analyzed in a pedagogical context, like [81, 82], to explore wearable technologies in the educational aspect and discuss different approaches to using smart wearable and smartphones for m-learning [10] and teacher orchestration. Suárez et al. [82] discussed using smartphones in education using inquiry-based learning by examining multiple approaches and their strengths and limitations.

The IoT was extensively used in the classroom to support both teachers and students [17]. Subbarao et al. [83] analyzed different IoT-based approaches providing solutions for several learning pedagogies using devices and sensors. Also, different augmented and virtual reality (AR/VR) solutions for supporting learning activities are discussed in [10, 84]. These approaches are categorized based on their technology stack and used infrastructure in the following subsections.

Internet of Things (IoT). The connection of different devices (things) with the Internet is known as the Internet of things [83, 85]. A smart classroom contains multiple intelligent devices, which eventually need to communicate to enrich the learning experience. IoT is one of the widely used approaches in different solutions; unlike other fields of life, it also evolved in teaching and learning pedagogies. Most of the solutions found in the literature, which use sensors for getting data from learning space, are based on the IoT paradigm. Rico et al. in [86] and Subbarao et al. in [9] review different IoT-based approaches providing multiplicity solutions for several learning pedagogies using a combination of devices and sensors.

Gligorić et al. [87] determine lecture quality using different sensors like PIR and sound sensors and a video camera. Similarly, another study [84] finds the student’s satisfaction from a classroom session using physical environment parameters. The student uses their smartphones to input their feedback as satisfied or not satisfied [88]. In another study, Gligorić et al. [8] designed an LED lamp to show students’ interest or satisfaction levels using Raspberry Pi (https://www.raspberrypi.org). They record 30 lectures using cameras and microphones and annotate students’ data using their smartphones. Students click exciting or not interesting and something satisfactory or unsatisfactory. A 30-second window was labelled when more than 90% of votes were received.

Mahmood et al. [14, 84] used a camera connected with Raspberry Pi to calculate students’ interest levels from their facial expressions and notify the teacher about their current status. Besides getting data about the lecture, IoT is also used for classroom attendance; in [89], Atabekov designed a smart chair for getting classroom attendance and time spent by a student in the classroom.

Near-Field Communication. The Near-Field Communication (NFC) technology is also used for automatic student attendance, indoor classroom location, and real-time feedback [90]. In [91], an RFID-based campus security system is proposed by Mirza and Brohi, which monitors and
tracks different resources, including students’ records, exam papers, and student certificates, using cloud computing. Another similar approach [92] used PIR and RFID sensors with Arduino to monitor classrooms and parking lots and determine which occupied or empty classroom or parking space. Furthermore, they used a video camera with a cloud to control smart classrooms by getting different data and decreasing energy conservation.

### 2.2.3. Augmented and Virtual Reality

Augmented and virtual reality (AR/VR) allows users to be physically involved in different blended scenarios and create a hybrid learning environment by combining physical and digital objects [18]. As students learn 50% of what they hear and read while 90% of what they do [95], AR/VR for learning purposes might significantly provide positive results and help students grasp more helpful information. Herpich et al. [96] discussed different mobile-based augmented reality solutions for supporting learners.

Elkoubaiti et al. [97] explore AR/VR in education and smart classrooms. They describe the technical requirements including latency, field of view, resolution, frame rate, network requirements, and measurements for the privacy and security of AR and VR applications. Similarly, Munoz et al. [98] represent a case study using an AR-based tool named GLUEPS-AR and a VR game (Game of Blazons). The study conducted different VR/AR-based activities for students and showed how VR/AR tools help teachers create different learning situations. Also, Kosmas et al. [93] evaluate the effect of the motion-based game on student performance during language learning classes.

Khan et al. [99] developed an augmented reality mobile application to examine their learning motivation. They used the ARCS (attention, relevance, confidence, and satisfaction) model to find the significance of AR technology on students’ learning performance. Although the available literature has extensive studies focused on AR/VR, according to Murat and Gokce [16], many students cannot arrange AR/VR headsets. Also, it distracts students’ attention, and undoubtedly, it is expensive as well.

### 2.2.4. Learning Dashboards

A learning dashboard is a visualization tool supporting teachers and learners in different learning scenarios for better decision-making [100]. It is a

---

### Table 1: Mobile sensor technologies and their usages in teacher orchestration and smart classroom.

<table>
<thead>
<tr>
<th>S. no.</th>
<th>Mobile sensor</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Camera sensors</td>
<td>Mostly used to capture images of both user(s) and its surrounding with numerous applications—such as user recognition, recognizing user(s) inclinations and its surrounding context [48]. Can be exploited in teacher orchestration, for example, student head-movement detection, body pose, and facial expression estimation for better performance [15, 49–51]. Used to capture voice(s) generated in the surrounding, either by the user itself or other objects with numerous applications—such as voice identification system, accident detection, and spying system. A microphone sensor has been used in the classroom to detect voice level created by student and teacher and for discourse analysis among student(s) or between teacher(s) and student(s). In education and teacher orchestration, microphone sensor has also been used in affective student modeling [52], teaching and learning [47], teacher-student dialogue recognition [53], etc.</td>
</tr>
<tr>
<td>b</td>
<td>Microphone sensor</td>
<td>Used to measure the light intensity of the surrounding atmosphere, with numerous applications—such as environmental pollution monitoring systems, weather forecasting systems, and other similar systems.</td>
</tr>
<tr>
<td>c</td>
<td>Ambient light sensor</td>
<td>Mostly used to capture images of both user(s) and its surrounding with numerous applications—such as user recognition, recognizing user(s) inclinations and its surrounding context [48]. Can be exploited in teacher orchestration, ambient light sensors can be exploited to measure the light level of the classroom for the healthy projection of light [54, 55] and students-to-board focus. Used to measure movements, acceleration information of users, angles, and inclination with numerous applications—such as old people healthcare systems, automatic traffic accident detection systems, and games [56, 57].</td>
</tr>
<tr>
<td>d</td>
<td>Accelerometer sensor + gyroscope sensor</td>
<td>In teacher orchestration, an accelerometer sensor has been used in measuring student-teacher communication [58], assessment of voice quality among college students [59], etc.</td>
</tr>
<tr>
<td>e</td>
<td>Proximity sensor</td>
<td>Used to detect objects with numerous applications—such as blind people guidance systems for navigating them through the pattern of their steps.</td>
</tr>
<tr>
<td>f</td>
<td>GPS sensor</td>
<td>GPS sensor has been used in orchestration too; for instance, Sun et al. [60] showed how GPS can be used to increase self-efficacy, self-regulation, and student achievement. Also, GPS can be used for indoor location or seat mapping.</td>
</tr>
<tr>
<td>g</td>
<td>Compass sensor</td>
<td>Usually used to measure phones or user’s location and their direction.</td>
</tr>
</tbody>
</table>
specific intervention of learning analytics used to identify meaningful data for various stakeholders (like teachers, students, and administrators) and how data representation can be helpful in sense-making [16]. Korozi et al. [5] developed LECTOR—a web-based tool for students’ reengaging systems in smart classrooms using multimodal data from different sources, including an eye tracker, depth camera, microphone, and other embedded sensors.

Similarly, another approach used LECTOR [100] and a smartwatch app called NotifEye [101], which shows a teacher’s smartphone notification with different information regarding students’ current learning status, activeness, and other positive interventions. Holstein et al. [102] developed a real-time dashboard for the intelligent tutoring system (ITS), which assists students during their programming course for learning http://ASP.net (https://dotnet.microsoft

---

**Algorithm 1: Detect facial feature and head direction.**

**Input:** camera image \( \text{img} \) in Bitmap//Google Vision APIs only accept bitmap images

**Output:** facial features (head rotation/title)

1. Detect faces in \( \text{img} \) using Google APIs FaceDetector and store them in list< FirebaseVisionFace > object faces
2. For \( i = 0 \rightarrow \)length of faces list
   i. Set face to faces \( [i] \)
   ii. Set headRotation to angle \( Y \) of face
   iii. Set headTilt to angle \( Z \) of face
   //Now using this data for decision-making
   iv. If headRotation > 20 and headRotation < −20 //means the student is not looking straight
      a. If warnTeacher is false//if last time he was looking straight, then wait for the next iteration before marking him inactive
         i. Set warnTeacher to true
      b. Else//it means he was also looking somewhere else last time
         i. Mark this student inactive
      c. Else//means the student is looking straight
         d. Set warnTeacher to false//clear previous state
3. End

---

**Figure 1:** Activity diagram of the proposed solution.
VanLehn et al. developed a FACT multimedia system [7]—a web-based AI tool that records students’ collaborative activities of arranging paper cards on the math class poster. Wetzel et al. [103] analyze the same FACT system with a traditional paper-pen-based approach to evaluate the time wastage factor of both conventional and electronic systems in learning pedagogies. Although learning dashboards better visualize students’ data, most systems require extra hardware and sensors.

2.2.5. Ubiquitous Computing and Other Sensors. The educational contexts have evolved into complex technological and social ecologies using different ubiquitous devices to transform the traditional learning space in ubiquitous learning environments (ULEs) [104]. Iqbal [46] represented a mobile application for teachers to mark quiz and exam papers and input feedback about students’ performance. Viswanathan and VanLehn in [105] and Tissenbaum et al. in [106] used students’ interaction logs with web app and tablet apps, 

![Head pose example](https://developers.google.com/vision/face-detection-concepts).
respectively, to identify their collaboration in a classroom session. In [107], Yu-Gang et al. proposed a mobile-based learning model, enhancing smartphones’ traditional learning. Smartphones are also used for automating the attendance process in ULE to facilitate teachers. Budi et al. [108] used image processing to take students’ attendance by using a mobile camera and a trained machine learning model running on the server for face recognition to identify different individuals in the uploaded image. Yang et al. [20] used voice print to mark students’ attendance and detect their indoor location in the classroom. In [109], Gligoric et al. measure the level of interest of a lecture by detecting student movements using a video camera, classroom sound (with microphone), and teacher’s movement from his smartphone accelerometer.

Prieto et al. [110] used the teacher’s smartphone’s accelerometer with other devices like a camera, microphone, and electroencephalogram (EEG) sensor (for capturing brain activities) to identify different classroom activities like an explanation, questioning, and monitoring. They identify teachers’ actions in a classroom session from multimodal data and build an “orchestration graph.” And while the orchestration graph defines who does what and when [111], it is a time-series graph plotting different activities with a given time and duration. Similarly, other approaches [112–114] reduce the infrastructure and use low-end devices; they used microphones to capture audio data and segment the lecture into different subactivities like question-answering. But these approaches require training the system for each teacher individually because of the change in voice tone and different speaking styles.

Recommendation techniques recommend tailored items to a user [115–118]. Liu et al. [119] proposed a smart learning recommendation system, which captures data from different sources to determine students’ current learning state and then suggests or reinforces different learning strategies (like quiz). In another approach, Bdiwi et al. [109] investigated the impact of teachers’ positions on students’ performance in higher education. Wang et al. [19] used an eye tracker to determine how much the teacher’s gaze guidance affects the students learning performance in video lectures. Similarly, Viilo et al. [120] perform teacher orchestration video data recorded in the classroom.

2.2.6. Wearables. The advantage of wearables over mobile devices is that they can be available most of the time, unlike mobile technology, mainly in pockets or bags [121]. In a study, Garcia [122] proposed a smartwatch app named “ScienceStories,” where students can record their science concepts. They find that the gamification mode has the highest use among the students. Quintana et al. [123] evaluate the acceptability of wearables in education by using the smartwatch to remind different tasks to the teacher during the classroom session. Also, Lu et al. [124] used a smartwatch for learning analytics to predict various activities using the hand gestures of a particular student. Another study designed, developed, and evaluated a wearable application for students with intellectual and developmental disabilities (IDDs) to assist them in the educational environment [19]. Wearables like smartwatches and smart bands are another common type of wearables named optical head-mounted

**Algorithm 3: Get sensor data from smartwatch.**

**Figure 3: Accelerometer axes (http://developer.android.com/reference/android/hardware/SensorEvent.html).**

```java
Input: application context to create SensorsManager object
Output: raw values from sensors
1. Start
   a. Create SensorsManagers object sensorManager using application context
   b. Get sensorsList from SensorsManagers
c. For eachSensor in sensorsList:
   i. Set sensor X to value of x-axis of eachSensor
   ii. Set sensor Y to value of y-axis of each sensor
   iii. Set sensor Z to value of z-axis of eachSensor
   iv. Wait for 300 milliseconds
   v. If application is not closed
      i. Go back to step c/to continuously capture sensor data
d. Release sensorManager/to avoid resource leakage
2. End
```
displays (OHMDs) or simply head-mounted displays (HMDs). They are usually worn over the eyes, which can either be utterly immersive like VR headset (Oculus (https://www.oculus.com)) or nonimmersive such as smart glasses (Google Glass [124] or Microsoft HoloLens (https://www.microsoft.com/en-us/hololens/)) [110]. In [112], the

Figure 4: (a) Choose to start the application in either teacher or student mode. (b) Ask the teacher to enter the number of rows and seats in each row for classroom structure.

Figure 5: (a) Teacher application showing class status and activeness level in real time. (b) The final report was presented at the end of a class.
teacher wore Google Glass to view the emotional status of each student in the classroom.

Patrick [114] used audio data from the microphone for different segment activities in a learning session. The author used a machine learning approach to train a classifier and then predict various activities from the given audio data, like answering, supervising students, and lecturing. Similarly, Donnelly et al. [86] also used audio data from the microphone to detect teacher questions from a live classroom session. Finally, Bdiwi et al. [108] used RFIDs to find the impact of the teacher’s position on students’ performance using an IoT-based approach.

Gligorić et al. [87] also used IoT devices, including PIR and sound sensors, to detect the lecture quality. Finding the lecture quality in real time is a positive approach, but using extra hardware raises costs and acceptability-related issues. In another study, Gligoric et al. [84] used a video camera, mic, and Android smartphone to detect the level of interest a lecture created. The author also proposes another IoT-based solution to show students’ satisfaction levels [84]. Finally, Mahmood and Salman [125] used a video camera and Raspberry Pi to find students’ attentiveness levels using their facial expressions and assist teachers in improving their teaching methodology.

The materials and methods should contain sufficient detail so that all procedures can be repeated. It may be divided into headed subsections if several methods are described.

3. Proposed Methodology

The proposed solution needs a smartwatch worn by the teacher in their dominant hand and connected to a smartphone placed in front of the teacher. First, it helps collect the teacher’s hand and foot movement to identify if the teacher is moving during the lecture or remains static. Then, the smartwatch sends its sensor data to the connected
smartphone, processing and analyzing for final result generation on the smartphone.

If it is only the teacher’s voice, it is classified as a lecturing event. If there is a combination of students’ and teachers’ voices, it is counted as a question-answer session or discussion. The facial orientation of the teacher is used to measure her activeness or tiredness level. A server application is developed and deployed on the teacher’s smartphone to collect and process the multimodal (different sources) data from the teacher’s and students’ devices. The processed data is displayed on the teacher’s smartphone showing statistics about the current state of the class, e.g., how many students are active and inactive and what is his voice quality during the lecture.

The application shows the current status of the classroom after collecting multimodal data in real time. It also provides a short glimpse of different activities at the end of a classroom session, for example, how much time the teacher spent lecturing, question answering (discussion), and writing on board. The application can mark students as active and inactive by processing the head and voice-related data discussed later in sections. The teacher’s activeness (Equation (3)) is calculated from two factors, i.e., classroom current status and voice level of individual students. The classroom’s current status can be found using

\[
CS = \frac{\text{no. of active students}}{n},
\]

where \( n \) is the total number of connected students, i.e., both active and inactive in that specific learning session, at the same time, and \( CS \) stands for classroom status, which will be a decimal value between 0 and 1. Similarly, the voice level can be calculated using

\[
VL = \frac{\sum_{i=1}^{n} V_i}{n \times \text{max-voice-level}}.
\]

Here, \( V_i \) represents voice level for an individual student, max-voice-level is the maximum threshold set for voice level, i.e., 90 decibels (dB) for our experiment, and \( n \) is the total number of connected students. The resulted value of voice level (VL) will be a decimal number between 0 and 1. And finally, Equation (3) uses these CS values, and VL can compute the teacher’s activeness level, which will be again a decimal number from 0 to 1.

\[
\text{Teacher’s activeness} = \frac{CS + VL}{2}.
\]

Finding the value of the teacher’s activeness fulfills our first object of this research work. Now to meet the second objective, i.e., finding the contribution of each modality, we analyze the kind of data captured from these modalities and then find the use of that captured data.

4. Implementation

The system works in a local area network to get data from different stakeholders. The teacher’s application acts as a
server to collect data from connected students. The student’s application running on different students’ smartphones is responsible for collecting and processing the data and then sending that processed data to the teacher’s smartphone for final representation and results in a generation. This section discusses how the application captures and processes this multimodal data in real time.

4.1. Data Acquisition and Processing. The following data is collected from both teachers and students, analyzed and used to find the classroom status and voice level as stated in Equations (1) and (2).

4.1.1. Facial Data. According to Mahmood et al. [84], the understanding of student interest level is allied with the quality of the lecture. Therefore, the application captures face-related data from teachers and students to get their level of interest and activeness in the current classroom session. This study focuses on head movement to analyze how much head direction helps identify the current attention level of the student. For this purpose, Google Vision APIs (https://cloud.google.com/vision/) detect users’ faces from images captured using a smartphone’s camera. These APIs provide a framework for detecting and tracking objects in images and videos. It supports face detection, barcode reading, and text recognition. For example, the head left to right movement represents head rotation, with a value between −60 and +60 and represented with \( \gamma \). Similarly, it also gives clockwise rotation, representing head tilt angle from −45 to +45 annotated as \( r \). The application takes a picture every 5 seconds and passes the captured bitmap image to Algorithm 1 to detect different face-related features.

The APIs offer different face-related data, including the number of faces detected, head rotation, head tilt (in degrees), smiling probability, eye-opening probability, and facial landmarks. Shown in Figure 2 is how these APIs consider head rotation tilt angle. Since the APIs provide head rotation and tilt, the rotation exceeds 20 degrees, i.e., +20 degrees on the right side, while −20 for looking at the left side (Step vii). Then, the application checks whether he exceeded this limit last time; if this is the first time he was noted, the application will wait for the next cycle/iteration; otherwise, it marks him as inactive. So, for example, if a student is not looking straight in the first cycle, the system will set a flag value warnTeacher to true, but in the next process, if the student is found looking straight, the application will mark him active and set warnTeacher back to false.

4.1.2. Voice Data. The application also collects voice data to infer classroom activities like lecturing or question-answer session. The microphone is used from existing smartphone devices in front of the teacher and students. The application collects audio data and performs preprocessing for noise removal on the student side. This cleaned data is used to measure the voice level of teachers and students in the classroom environment. If it detects only the teacher’s voice, it is marked as a lecture. But if there is a combination of both teacher and student’s voices within a defined threshold, then the system considers it a discussion or question-answering session. It uses standard Android APIs to collect and extract features from audio data for audio processing. As the application measures the voice level, we used the MediaRecorder class from Android APIs to get the maximum amplitude of audio data. The student application sends this amplitude value to the teacher’s smartphone, and the teacher application compares these values captured from different students. As given in Algorithm 2, if the voice difference between the two nearest students is noticeable, i.e., a value from student A is 35 dB, while the next student (student B) sends a value of 60 dB, then the application checks whether voice amplitude is captured on the teacher’s smartphone if the teacher’s voice is around 50 to 60 dB. Thus, the application infers that the teacher is lecturing while student B talks with someone. But suppose the teacher’s voice amplitude is less than 30 dB. In that case, the application considers that the student is asking a question and therefore marks that session as “discussion” or “question answering,” as shown in Algorithm 2.

4.1.3. Hand and Foot Movement Data. To find the teacher’s mobility and interaction in the classroom, the system captures her hand and feet to infer whether the teacher is standing still or moving. The system includes an off-the-shelf Android Wear-OS (https://wearos.google.com/) available smartwatch worn by the teacher on the dominant hand. In addition, it captures data from IMU (Inertial Measurement Units) sensors, including accelerometer, gyroscope, and pedometer mainly. The application uses Android APIs to interact with sensors and captures data at the rate of 40 samples per second to correctly recognize gestures from raw data [126]. Further details of these sensors are given below. Algorithm 3 shows steps getting sensory data from smartwatches.

<table>
<thead>
<tr>
<th>S. no.</th>
<th>Type</th>
<th>Min age</th>
<th>Max age</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Teacher</td>
<td>34</td>
<td>51</td>
<td>12</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>Student</td>
<td>24</td>
<td>28</td>
<td>17</td>
<td>5</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 2: Demographics of participants.
of icons representing student setting in the classroom. The application in server mode, the teacher presents a grid in each row in the classroom (Figure 4(a)). Then, starting from there, he is prompted to input the number of rows and seats. Therefore, when the teacher starts the application to monitor the classroom, use the seat map on the screen to mimic the real classroom structure. To reduce cognitive overload, the application shows a seating arrangement with a colored icon. To decide which icon will be updated in the grid of the application, use the seat number.

After collecting multimodal data from several connected devices, all the data is combined on the teacher’s smartphone for final calculation and result generation. The system contains different features regarding face and voice data from the student’s side. The application continuously updates the seat-map grid to show the latest data on the screen. For example, if the voice level is less than 40 dB (see Algorithm 2, Step 3). Similarly, suppose the user’s face is not detected or their head direction exceeded by 20 degrees (see Algorithm 1); in that case, the application provides real-time feedback to the teacher.

On the teacher side, after getting this multimodal data from all students, the application first calculates the class status CS using the number of active students (marked using Algorithm 1) and total students using Equation (1). Similarly, the overall classroom voice level VL is also calculated using Equation (2). And finally, by substituting the values of CS and VL in Equation (3), the teacher’s current activity level can be calculated. The application continuously calculates the activity value and updates a progress bar on the teacher’s smartphone to provide real-time feedback, as shown in Figure 5(a).

The system also included a smartwatch (Asus Zenwatch 2) worn by the teacher to capture his hand and foot movement. The application captures sensor data of a five-second window and processes that data on the teacher’s smartphone to get the number of steps taken and process hand gesture data. If the number of steps in three consecutive time windows is less than 1 or greater than 3, the system’s foot movement is less efficient for better lecture quality. In addition, it counts the number of steps during a particular classroom session, shown in the final report presented at the end of the class and a detailed summary of a learning session (Figure 5(b)).

5. Experimental Setup

This section describes the environment setup used for our experiments during actual classroom sessions.

5.1. Classroom/Environment Setup. Figure 6 depicts the layout and management of teacher and students’ positions in the classroom during the experiment. The smartphone was placed in front of a student using the specialized smartphone jacket installed on the back of the student’s chair in front of them. Figure 7 shows a chair with a smartphone jacket installed at the back to get students’ faces and audio data. Five positions were selected to sit a student with a smartphone, whereas a teacher is equipped with a smartphone and smartwatch (Figure 6). The teacher is standing and moving during the classroom session. Therefore, his smartphone is placed in a neck holder to make it easier to move.

Are you satisfied with existing teacher orchestration solution?

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>30%</td>
<td>70%</td>
</tr>
</tbody>
</table>

**Figure 10:** Satisfaction level after using existing teacher orchestration solutions.

**Table 3:** Questions asked after a test run.

<table>
<thead>
<tr>
<th>S. no</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Is the application easy to use?</td>
</tr>
<tr>
<td>2</td>
<td>Did the application improve teacher performance?</td>
</tr>
<tr>
<td>3</td>
<td>Did the application produce disturbance during the classroom session?</td>
</tr>
<tr>
<td>4</td>
<td>Did smartphone-based orchestration provide a more engaging experience in the classroom?</td>
</tr>
<tr>
<td>5</td>
<td>Are you satisfied with smartphone-based teacher orchestration?</td>
</tr>
<tr>
<td>6</td>
<td>Should this application be used in other classrooms as well?</td>
</tr>
<tr>
<td>7</td>
<td>Any comments or suggestions.</td>
</tr>
</tbody>
</table>

(1) **Accelerometer.** An accelerometer is used to measure the acceleration (change of velocity) in three axes (x, y, and z) [127]; see Figure 3. It reads these acceleration values from the smartwatch accelerometer to find hand gestures.

(2) **Gyroscope.** A gyroscope is used to measure the angular velocity (orientation/tilt) of a device’s three dimensions (x, y, and z) [8]. Therefore, it correctly identifies hand gestures by combining them with the accelerometer data.

(3) **Pedometer.** A pedometer is an electromechanical sensor used to detect and count each person’s step [127]. The application uses several steps to identify whether the teacher is standing still or moving toward the students in the classroom.

4.1.4. **Data Representation.** To better user experience and reduce cognitive overload, the application shows a seating map on the screen to mimic the real classroom structure. Therefore, when the teacher starts the application to monitor, he is prompted to input the number of rows and seats in each row in the classroom (Figure 4(a)). Then, starting the application in server mode, the teacher presents a grid of icons representing student setting in the classroom (Figure 4(b)). This icon changes according to the current student status; for example, when a student is not connected, the white icon, but when a new student gets connected, the application captures his seat number from the connection request packet and updates their status from the white icon to a colored icon. To decide which icon will be updated in the grid of the application, use the seat number.

The system also included a smartwatch (Asus Zenwatch 2) worn by the teacher to capture his hand and foot movement. The application captures sensor data of a five-second window and processes that data on the teacher’s smartphone to get the number of steps taken and process hand gesture data. If the number of steps in three consecutive time windows is less than 1 or greater than 3, the system’s foot movement is less efficient for better lecture quality. In addition, it counts the number of steps during a particular classroom session, shown in the final report presented at the end of the class and a detailed summary of a learning session (Figure 5(b)).
and provide real-time statistics on his smartphone screen. In addition, the teacher wears a smartwatch on his dominant hand to capture their hand movements and count their steps during the classroom using the built-in pedometer of the smartwatch.

5.2. Display Seating Map. Real-world classroom size is not fixed, and the system must show a student’s exact position in the classroom. Therefore, to offer the exact indoor location, the application uses QR codes to recognize a student’s accurate seat map, unlike some existing solutions that use RFID [8] for indoor location, which is costly and requires technical assistance. The QR code is placed in front of each seat to get the seat number and position in the classroom, as shown in Figure 8.

5.3. Evaluation. For the evaluation of the proposed system, we conduct questionnaire-based surveys. We first take a pre-task study from participating teachers during the experiment to know how many teachers had used an orchestration solution before. After that, we conduct experiments in several classroom sessions to try our Android application in real classroom scenarios. Finally, we take a posttask questionnaire to get participants’ responses after using the Android application. The statistical data from both questionnaires are gathered and coded in SPSS version 21 for further analysis and significance testing.

6. Results and Discussion

After implementing the proposed system, we conducted several experiments in different classroom sessions for one month to better understand and impact our developed Android application. This section discusses the results and findings obtained from pre- and posttask questionnaires.

6.1. The Demographics of Participants. For the experiments, we asked several teachers and students to voluntarily participate and use the Android application on their smartphones during classroom sessions. First, we explain how the system works to all participants and provide a more engaging user experience using low-cost off-the-shelf devices. By requesting approximately 30 teachers, 18 teachers (12 males and 6 females) agreed to use this application and contribute their feedback voluntarily. Similarly, by asking 40 students, 22 agreed to participate, where 17 were male, and 5 were female students between 24 and 28 years (see Table 2).

6.2. The Pretask Findings. We asked the participants whether they had used any teacher orchestration solution before and their experience with those solutions/tools in the pretask questionnaire. As shown in Figure 9, around 80% of participants did not use any orchestration tool before, and they were not familiar with teacher orchestration. The other 20% were mostly teachers, who were also unfamiliar with teacher orchestration, but they used MOOCs to assist their students in the learning process.

6.3. The Posttask Findings. After the experimental classroom sessions, we conducted a posttask questionnaire-based survey. The participants were asked about their experience...
and observations after using the Android application. In addition, they were asked whether they feel any improvement and how much the smartphone-based orchestration solution will help create a more engaging learning experience. These questions are given in Table 3.

After collecting their responses, we coded all the recorded data in SPSS version 21 and performed a paired sample t-test for these different questions and variables. The first question in our survey was about knowing how the user felt in terms of easiness regarding the proposed solution. As shown in Figure 11, around 50% of the participants strongly agreed that the application was easy to use because the user could join and start with only 2 to 3 clicks. In contrast, the rest of the 10% and 5% mark the easiness as neutral and disagree.

The proposed solution’s primary purpose is to improve teacher performance and increase learning outcomes. Table 4 shows the statistical data gathered from participating students presenting the improvements made after using the proposed solution. About 45% of the students strongly agreed, and 35% agreed that the application improved performance by presenting valuable data to the teacher, which supported him in understanding the entire classroom’s current status. The same data is also represented in Figure 12 using a bar graph.

Along with improving teacher performance, we were also fascinated by the proposed system’s negative factor or downside. Therefore, we asked the participants whether the application produced any disturbance or distracted them during the classroom session. Only 35% of the participants marked a slight annoyance (Figure 13) because the teacher was wearing a neck holder stand to hold his smartphone, and the majority of participants in this 30% were teachers. In contrast, most students, around 35%, disagree with the
disturbance, and only 20% mark it as neutral. Of course, a neck holder in the classroom might create a slightly negative impact, which was only used to allow the teacher to view data easily on his smartphone. But it can be replaced with a monitor screen installed behind the students, which provides the teacher with a freer environment to move. Still, on the other hand, it will add some extra cost to the proposed solution because the primary purpose was to use the existing devices to create a low-cost solution.

We also investigate how much the proposed smartphone-based orchestration solution helped create an engaging experience in the classroom. The majority of the

![Smartphone based orchestration provides a more engaging experience in classroom](image1)

**Figure 14:** Smartphone-based orchestration provides a more engaging experience in the classroom.

![Participants’ satisfaction level after using smartphones for teacher orchestration](image2)

**Figure 15:** Participants’ satisfaction level after using smartphones for teacher orchestration.

**Table 5:** Paired sample t-test of the proposed solution with existing teacher orchestration solutions.

<table>
<thead>
<tr>
<th>Pair</th>
<th>You are satisfied with existing T.O. solutions—you are satisfied with smartphone-based T.O. solution</th>
<th>Lower</th>
<th>Upper</th>
<th>95% confidence interval of the difference</th>
<th>Std. error mean</th>
<th>Std. deviation</th>
<th>50%</th>
<th>0%</th>
<th>10%</th>
<th>40%</th>
<th>20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>You are satisfied with existing T.O. solutions—you are satisfied with smartphone-based T.O. solution</td>
<td></td>
<td></td>
<td>-51.88743 - 15.71257</td>
<td>6.51460</td>
<td>14.56709</td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>35</td>
<td>50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pair</th>
<th>You are satisfied with existing T.O. solutions—you are satisfied with smartphone-based T.O. solution</th>
<th>Lower</th>
<th>Upper</th>
<th>95% confidence interval of the difference</th>
<th>Std. error mean</th>
<th>Std. deviation</th>
<th>50%</th>
<th>0%</th>
<th>10%</th>
<th>40%</th>
<th>20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>You are satisfied with existing T.O. solutions—you are satisfied with smartphone-based T.O. solution</td>
<td></td>
<td></td>
<td>-51.88743 - 15.71257</td>
<td>6.51460</td>
<td>14.56709</td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>35</td>
<td>50</td>
</tr>
</tbody>
</table>

![Paired sample test](image3)
participants, i.e., 90%, accepted that the proposed solution successfully made an engaging experience in their learning environments, while only 10% answered this question as neutral but none of the participants disagreed with the engaging impact created by our proposed solution (Figure 14).

Similarly, to know the impact of using low-cost smartphone devices rather than huge and expensive infrastructures, we asked the participants how satisfied they were with using smartphones for teacher orchestration; 35% strongly agreed and 50% agreed that they were satisfied with using off-the-shelf smartphone devices (Figure 15). While 10% responded neutral, only 5% disagreed that using their smartphones is a good idea because of the privacy concerns.

Now, we compare this satisfaction result with the post-task results. We asked the participants about their satisfaction level after using the existing teacher orchestration solutions. Therefore, we perform a paired sample t-test and use this hypothesis and alternate hypothesis:

- \( H_0: \) the satisfaction level of participants is not significant.
- \( H_1: \) the difference between these satisfaction levels is significant.

A confidence interval value of 95% shows the generated results in Table 5, where the \( p \) value is calculated as 0.007. This is less than 0.05. Therefore, we can drop the null hypothesis and accept the alternative hypothesis as valid. The participants are more satisfied with the proposed smartphone-based teacher orchestration solution than the available solutions.

Lastly, we asked whether this application should be used in their other classrooms. After getting the satisfaction level, the response to this question was also very encouraging. Around 70% of the students recommend using this application in other classrooms for teacher orchestration; see Figure 16. And 15% mark this question as neutral, while only 10% disagree with utilizing this application.

7. Conclusion

This study presented state of the art in teacher orchestration and provided a more engaging student experience in a smart classroom. It evaluated several learning pedagogies and their effect on different stakeholders, including students, teachers, and administrators. This study proposed a solution that used off-the-shelf devices for teacher orchestration in a smart learning environment. The solution captures data from teacher and students and processes it, where each device processes its data and sends the results to the teacher’s smartphone to provide real-time results. We also evaluate the significance of the proposed solution by using the application in real classrooms and get participants’ feedback using a brief questionnaire survey. The results were significantly positive and also encouraged smartphone-based orchestration solutions. Pose recognition significantly impacts studying body language [128]; therefore, processing a teacher’s pose in a learning session can open numerous opportunities in a teacher’s orchestration.

Data Availability

The data that support the findings of this study are available upon request from the first author.

Conflicts of Interest

The authors claim no conflict of interests.

References


