RoomSense: AR-Enhanced Classroom Engagement System

Luke Liu
University of Rochester
Rochester, New York, USA
lliu88@ur.rochester.edu

Ertugrul Senturk
University of Rochester
Rochester, New York, USA
esent@ur.rochester.edu

Tasin Khan
University of Rochester
Rochester, New York, USA
tkhan17@ur.rochester.edu

Pranav Suryakiran Waghanna
University of Rochester
Rochester, New York, USA
pwaghanna@ur.rochester.edu

Abstract

In a classroom setting, lecturers try their best to teach students, but often, due to the large size of the class, it is difficult to gauge students' engagement and stay on track with the lecture topics. Moreover, students feel shy to ask questions, fear their peers' judgment, or do not wish to stand out. We propose RoomSense, an AR solution along with a web application to assist professors and improve classroom engagement. The system uses Snap Spectacles to provide real-time engagement visualization through color-coded health bars above students' heads, while a companion web application enables anonymous question submission. Our formative user study with 8 participants demonstrated that RoomSense is highly usable even to users with little immersive-technology experience.

Keywords

Augmented Reality, Classroom Engagement, Human-Computer Interaction, Snap Spectacles, Educational Technology

ACM Reference Format:
Luke Liu, Ertugrul Senturk, Tasin Khan, and Pranav Suryakiran Waghanna. 2025. RoomSense: AR-Enhanced Classroom Engagement System. In *CSC 412 Final Project Report*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

In-person classroom teaching still depends on a presenter's ability to "read the room," but presenters often have limited real-time awareness of the audience's attention, confusion, or engagement. Subtle nonverbal cues, such as students losing focus, silently getting confused, or disengaging, are easy to miss, especially in larger or fast-paced classrooms. During a live lecture, the presenter is focused on explaining content, managing time, advancing slides, and answering questions, which makes it difficult to continuously monitor the room's cognitive state. Improving situational awareness in the moment could make teaching more adaptive, inclusive, and effective.

1.1 Current State of the Art

Existing classroom support tools (polls, clicker questions, chat backchannels) require the presenter to stop and explicitly ask for feedback, which interrupts lecture flow and depends on voluntary self-reporting from students [2]. Lecture capture and analytics tools focus primarily on after-class review rather than providing instructors with immediate, real-time awareness of audience engagement. Meanwhile, AR collaboration systems build immersive virtual environments and emphasize presence and embodiment, but they are generally designed for remote or simulated meetings rather than physical classrooms, and they do not summarize live audience engagement for an in-room speaker [1].

1.2 Proposed Solution

To allow professors and lecturers to improve interaction and understand audience engagement, we introduce an AR-based approach that allows the lecturer to gauge the engagement of the people they are presenting to and accept queries through a web application.

2 Architecture and Workflow

2.1 System Architecture

The system is divided into multiple components as shown in Figure 1. The Web Application, hosted on Vercel, serves as the primary interface for professors to schedule lectures and manage content, and for students to view materials and submit queries. The Snap Spectacles application runs locally on the device, allowing professors to gauge student engagement and display notes, topics, and queries. The Backend API, hosted on AWS Elastic Beanstalk, handles data flow between all components. MongoDB provides a flexible database structure that supports future customization and improvements.

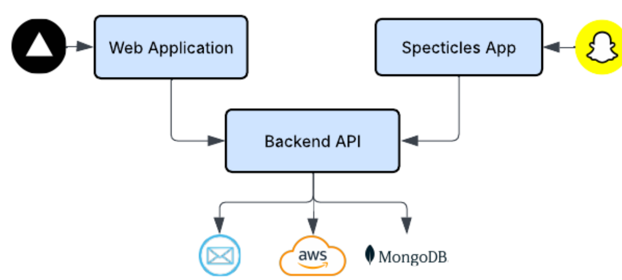

Figure 1: RoomSense System Architecture

2.2 User Workflow

Initially, a professor goes to our website and logs in through their email. The professor then creates a class for the semester, and the system automatically generates all the individual lectures throughout that period. Before each lecture begins, the professor can add notes, set timelines, and define the topic for that specific day through the system. A unique lecture code is generated and shared with students. For each lecture, the professor can also view students' engagement and select whether to accept queries from students.

The professor then connects to the Snap Spectacles and starts the lecture. Students can log in through the same website by entering the lecture code under the "Student" tab. The students do not have to provide their email's nor their names.

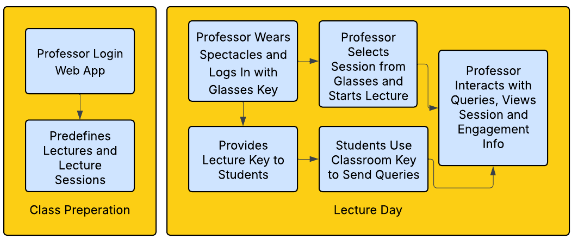

Figure 2: User Workflow showing the interaction between professors, students, and the system

Once the lecture starts, students can send their queries through the web application and choose whether to identify themselves. When the professor wears the Snap Spectacles, they can view:

- Lecture notes and timeline will be displayed as transparent floating bubbles that can be repositioned freely.
- Each student's individual engagement level will be shown as a bar above their head.
- The summary bubble will display the remaining time, the number of students detected, and the average engagement level throughout the class.
- Each query will be displayed for a minimum of 30 seconds to ensure the lecturer has adequate time to read it, even when consecutive queries are submitted. Any additional queries will appear after this period.
- Engagement levels are indicated through color-coded bars: green for strong engagement and red for poor engagement.

3 Implementation

The proposed solution is built in 3 parts: Frontend Web Application, Backend API, and Snap Spectacles. The frontend and Snap Spectacles interact with the backend through the Flask API.

3.1 Frontend Web Application

The frontend of RoomSense is built using Next.js with TypeScript to create a responsive, component-based user interface for both professors and students.

3.1.1 Professor Interface. The interface for professors includes an email-verification based login page (Figure 3), Dashboard (Figure 4), Lecture Manager Page, and an AR Glasses connection page. Professors authenticate using an email-based verification flow that triggers the backend's Mail API to send a verification link.

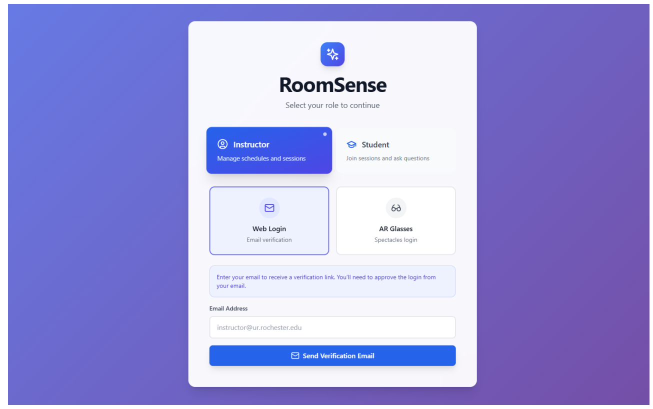

Figure 3: Professor Login Interface with email verification

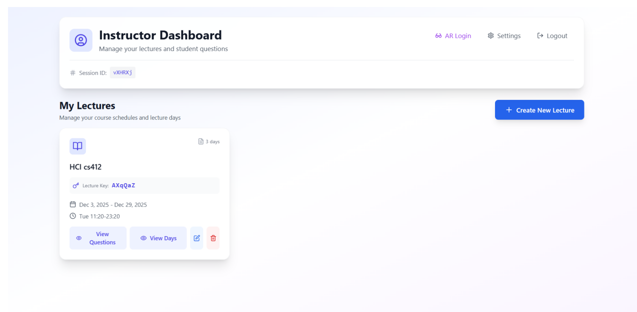

Figure 4: Professor Dashboard showing lecture management options

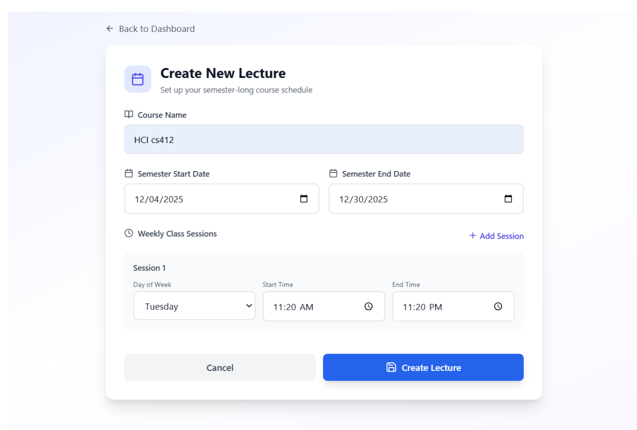

Figure 5: Create Lecture interface with scheduling options

3.1.2 Student Interface. Students do not have to provide personal details. The interface for them is simple and has only 2 pages: Connect-to-Session Page where students enter the lecture key (Figure 6), and the Query Page where they can ask questions and choose to identify themselves (Figure 7).

Figure 6: Student Login interface - simple and privacy-focused

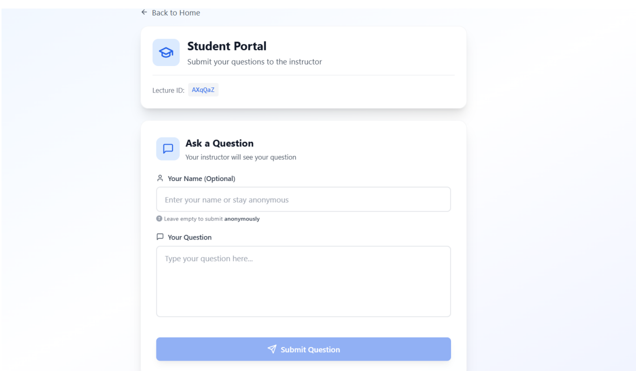

Figure 7: Student Query Page for submitting questions during lectures

3.2 Snap Spectacles Engagement Visualization

On the Snap Spectacles, we implement a Lens Studio **TypeScript** component that estimates per-student engagement from facial landmarks and visualizes it as floating 3D health bars above each detected face. The script runs on-device and supports up to three faces simultaneously, using Snap's Head Binding component to access landmarks around the eyebrows and mouth (Figure 8).

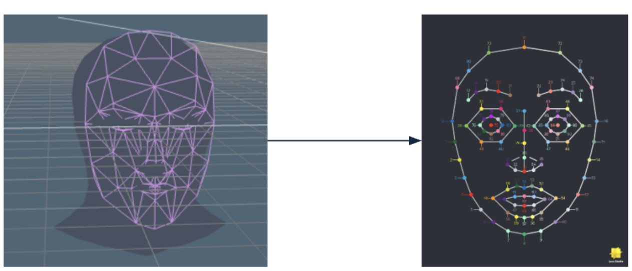

Figure 8: Extracted Head Binding and Landmarks

For each detected face, the script retrieves landmarks for the left eyebrow, right eyebrow, and mouth center. From these, it computes a set of normalized metrics:

- **Furrow ratio:** horizontal distance between eyebrows, normalized by a per-face baseline, used to detect brow furrowing (possible confusion or strain).
- **Tilt ratio:** vertical difference between eyebrow heights, normalized by eyebrow distance, used to capture head tilt.
- **Yaw metric:** asymmetry between left and right eyebrow–mouth distances, indicating whether the person is turning their head away from the lecturer.
- **Pitch metric:** change in average eyebrow–mouth distance relative to baseline, approximating nodding or looking up/down.

These metrics are converted into penalty scores with configurable thresholds and weights. A weighted sum of penalties is subtracted from full 'life" to produce a scalar engagement value in $[0.1, 1.0]$ for each face. High engagement corresponds to a higher life fraction, while strong penalties from looking away or excessive tilt/furrow reduce the value. The life fraction is then mapped to a color by interpolating between red (low engagement) and green (high engagement). For visualization, each face is associated with a 3D bar object (a box mesh) that is positioned above the head using the world position of the Head Binding object plus a fixed vertical offset. The bar's local scale on the x-axis encodes the current life fraction, and its material color is updated each frame based on the interpolated engagement color. Both the life fraction and color are smoothed over time using linear interpolation, preventing jittery changes when the face or landmarks move slightly.

Figure 9: Engagement Bar Penalties as student attention decreases

3.3 Backend API

The backend of RoomSense is implemented as a **Flask API** which acts as a communication hub between the web application and Snap Spectacles. It is responsible for authenticating professors, managing lecture sessions, and handling real-time queries received by students. It is deployed on AWS Elastic Beanstalk and connected to MongoDB for data storage.

The API supports CRUD operations including endpoints for creating lectures, retrieving sessions, updating lecture days, managing questions, and marking questions as answered.

3.4 Database Schema

The backend API uses MongoDB Compass for data storage. MongoDB's flexible schema allows new fields to be added without migration. The database holds 3 collections:

- **Users:** Stores professor information including email, login timestamps, and session IDs (active, pending, inactive)
- **Lectures:** Holds lecture data including key, lecturer ID, course name, semester dates, class sessions, and lecture days
- **Student Questions:** Contains queries with lecture key, student name (optional), question text, and delivery status

There is no data stored regarding students unless they choose to identify themselves when they send a query. Only the professor's information is stored in the "users" collection.

3.5 System Performance

RoomSense was designed to operate smoothly during live lectures without interrupting the instructor's teaching flow. In practice, the system maintained real-time responsiveness across all three components: Snap Spectacles, backend communication, and the web interface.

On-device processing allowed the glasses to compute engagement levels locally without relying on server inference, avoiding network-related delays. Engagement bars updated at approximately 20fps using a throttled update loop, which provided smooth feedback while remaining lightweight for wearable hardware. Notifications, lecture notes, and timelines rendered without noticeable delay, and engagement changes appeared within a second when a student shifted posture or looked away.

Overall, the system performed reliably during live trial sessions. All eight user study participants were able to receive feedback, send questions, and visualize engagement without technical intervention—supporting our goal of enabling instructors to monitor attention passively rather than relying on manual check-ins. This suggests that RoomSense is operationally viable for real-classroom use, meeting our core objective of providing real-time engagement awareness to lecturers.

4 Evaluation Study

4.1 Study Design

Our goal in the evaluation was to understand (1) whether RoomSense is usable for instructors and students, and (2) how the AR overlays (floating lecture notes, timeline, engagement bars, and question notifications) affect the lecture experience. We conducted a formative user study followed by a short post-test survey.

4.1.1 Participants. For the in-person study, we recruited two teaching assistants (TAs) who regularly teach recitations and six students who alternated between instructor and student roles. The post-test survey received 8 responses in total. Most respondents identified as students (75%), with the remaining 25% as teaching assistants. The majority (87.5%) reported never using AR in their teaching/learning environment, and 12.5% had only used it rarely.

4.1.2 Procedure. The study consisted of three parts:

1. **Instructor flow:** Participants logged in via email verification, created or edited a course, and configured lecture-level settings. The TA tested delivering a mini-lecture while wearing the glasses, monitoring engagement bars, glancing at floating notes and timeline, and responding to questions.
2. **Student flow:** Students worked in small groups of 3 participants, with each participant trying the lecturer dashboard while others acted as students.
3. **Post-test survey:** Participants rated usability of individual features on 5-point Likert scales and overall satisfaction on a 0–10 scale.

4.2 Result Analysis

Even though 7 out of 8 survey respondents reported having never used AR in their classroom context, participants were able to complete all main tasks in both the instructor and student interfaces with minimal guidance. This suggests that RoomSense is accessible even to users with little immersive-technology experience.

4.2.1 Usability of the Web Interface. Survey ratings indicate that the web application was perceived as easy to use. Navigation was rated with a mean of 4.1/5, and logging in via email verification was rated very positively (mean $\approx$ 4.6/5). All participants successfully logged in and created a lecture. Students reported that joining a lecture by code was straightforward.

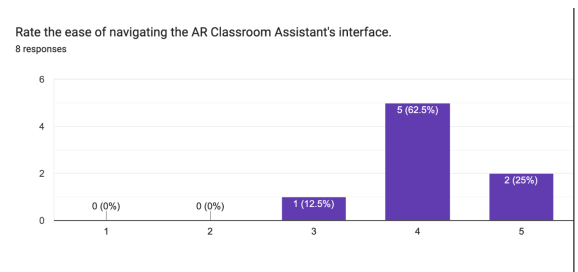

Figure 10: User Rating of Ease of Navigation

4.2.2 Usability of AR Features. Participants rated the AR overlays as both usable and helpful. Floating features (timeline, notes, time remaining) were rated highly (mean $\approx$ 4.5/5). Engagement bars were rated 4 or 5 by all respondents. The settings panel ease of use was rated 4 by 3 participants and 5 by 5 participants.

In qualitative feedback, the TA described the engagement bars as "immediately understandable," and the traffic-light-style bars above students' heads made it easy to scan the room and quickly see who might be disengaged. However, participants noted the potential for visual overload when all overlays are visible at once, though the toggle options in settings helped mitigate this.

4.2.3 Perceived Usefulness and Satisfaction. Overall satisfaction with RoomSense was high. On a 0–10 scale, no participant rated the system below 8; 3 participants chose 8, 4 chose 9, and 1 chose 10 (mean $\approx$ 8.75/10). Many participants described the system as "cool" or "futuristic" and reported that they could imagine using it in real classroom settings, especially large lectures where hand-raising is intimidating and engagement is hard to gauge.

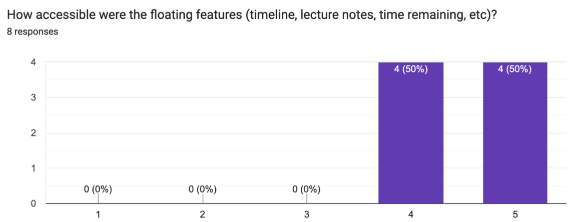

Figure 11: User Rating of Usability of Features

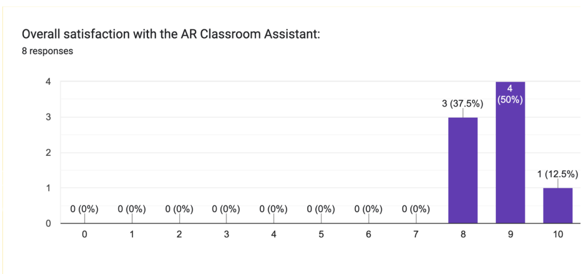

Figure 12: Overall User Rating of Satisfaction

5 Discussion and Reflection

5.1 Discussion

Through our study, various interesting insights were highlighted. One notable observation was that some features participants initially considered as limitations were actually seen as useful. The 30-second display duration of the query notification was said to be helpful for pacing and would prevent dwelling on a topic for too long. Similarly, although the engagement bars lagged when students arrived late, participants said this delay was similar to an "initialization period."

Participants mentioned that the floating notes prevented them from looking away from the students, helping them speak more fluidly. This suggests that AR overlays for notes can improve teaching behavior by maintaining the connection between the professor and students.

The participants also enjoyed the engagement bars over students' heads, which was similar to health bars in video games. The color scheme helped them understand overall class engagement without checking all bars individually. Initially, a few participants felt overwhelmed when all overlays were toggled on simultaneously, but the presence of toggles and settings helped mitigate this—demonstrating the importance of user-controlled customization in AR environments.

5.2 Reflection

If we were to start the project again, we would add several improvements:

(1) **Real-time Communication:** We would invest more time into real-time communication. A lightweight WebSocket should reduce latency and allow smoother updates especially for query notifications.
(2) **AR Overlay Ergonomics:** We assumed more information equaled higher utility, but it led to higher visual load. A better layout with adaptive transparency, hierarchical grouping and context-aware display would improve usability.
(3) **Participant Diversity:** Most participants were peers and roommates, limiting exposure to varied teaching styles and accessibility needs. Testing with multiple professors across different departments would provide richer feedback.
(4) **Student Detection:** Implementing persistent face tracking or recalibration shortcuts would address the main technical limitation of handling late arrivals.

5.3 Limitations

Although RoomSense showed promising usability, several limitations constrain the current prototype and inform directions for future development.

First, the system can reliably track and visualize engagement for up to three students simultaneously. When more than three faces are detected, the frame rate drops and responsiveness decreases due to on-device computation costs for landmark extraction and bar rendering. Scaling to larger classrooms will require optimization or offloading parts of the pipeline to an external device.

Second, the engagement metric itself is heuristic rather than empirically validated. The current implementation infers engagement from eyebrow distance, tilt, yaw, and pitch, but these cues do not guarantee cognitive engagement or understanding. A controlled study would be necessary to determine whether these signals correlate with real learning outcomes, and whether additional metrics (such as gaze, hand movement, or posture) improve reliability.

Finally, the hardware used (Snap Spectacles) introduces physical constraints. The limited field of view means that only students directly within the lecturer's forward-facing view are analyzed, requiring the instructor to rotate their head to observe the entire room. This is sufficient for small group teaching but may be restrictive for wide lecture halls.

6 Conclusion

Overall, RoomSense demonstrated strong potential as an AR-enhanced classroom engagement system. The high satisfaction scores (mean 8.75/10) and positive qualitative feedback indicate that our approach of combining AR overlays with a web-based query system addresses real pain points in classroom interaction. With more time, a deeper technical foundation would allow us to develop a more refined and scalable version of the system that could transform the passive lecture into a truly interactive experience.

Acknowledgments

We would like to thank Professor Yukang Yan for guidance throughout this project, and all participants who took part in our user study.

References

[1] Mark Billinghurst, Adrian Clark, and Gun Lee. 2015. A survey of augmented reality. *Foundations and Trends in Human–Computer Interaction* 8, 2–3 (2015), 73–272.

[2] Jane E. Caldwell. 2007. Clickers in the large classroom: Current research and best-practice tips. *CBE—Life Sciences Education* 6, 1 (2007), 9–20.