JEECS (Journal of Electrical Engineering and Computer Sciences)
Vol. 10, No. 2, December 2025, pp. 152-159
e-ISSN: 2579-5392 p-ISSN: 2528-0260

Development of Augmented Reality-Based English Learning Media for Vocabulary and Pronunciation

Said Hirzi Hadi

Universitas Pignatelli Triputra, Duwet Raya No. I Karangasem, Surakarta, Indonesia

Article Info

Article history:
Received: 8 October 2025
Revised: 9 November 2025
Accepted: 16 November 2025

Keyword:
Augmented Reality
English
Pronunciation
Vocabulary
Learning Media

Abstract

Digital note-taking applications serve as essential tools for personal information management, presenting opportunities for enhanced security mechanisms to protect sensitive data. Most current solutions depend on server-side processing, creating potential vulnerabilities and privacy concerns. However, a robust solution that fully executes hybrid encryption on the client-side to seamlessly protect both text and image data within a single application remains unexplored. This study introduces SmartNotes, a web-based application safeguarding text and image notes through an innovative hybrid encryption system synergistically combining RC4 and XChaCha20 algorithms. A key contribution is the full client-side execution of encryption–decryption processes, eliminating server dependencies and significantly reinforcing data confidentiality. The hybrid design strategically utilizes RC4 for rapid data processing and XChaCha20 for robust cryptographic protection, creating an optimal balance between performance and security. System performance was rigorously evaluated using seven private datasets under diverse key conditions. Testing methodology included comprehensive assessment of processing speed, data integrity verification, and resistance against unauthorized access attempts. Results demonstrated flawless data restoration across all test cases, validating robustness and reliability. Encryption averaged 1.5 seconds, while decryption required 20.20 seconds metrics well-suited for practical web environments. These findings affirm SmartNotes delivers a secure, autonomous, user-centric solution for digital note management, advancing applied cryptography through a novel client-side hybrid encryption paradigm. This approach successfully balances strong security with practical performance, making it suitable for securing data in everyday web applications.

Corresponding author:
Said Hirzi Hadi, hirzihadi01@gmail.com

DOI: https://doi.org/10.54732/jeecs.v10i2.6

1. Introduction

English is an international language that plays an important role in global communication, education, and technology. One of the fundamental skills in learning English is mastering vocabulary and pronunciation. Incorrect pronunciation can lead to misunderstandings of meaning and hinder communication ability [1]. However, in learning practice, pronunciation skills often receive insufficient attention due to limitations in media, methods, and the time available in class. [2]. Recent studies have emphasized the importance of integrating interactive technology into English language learning to improve both vocabulary and pronunciation. Emilia et al. [3] showed that the use of Augmented Reality (AR) can make English vocabulary learning more effective and interactive for secondary students, while 'Aini et al [4] revealed that AR-based methods increased learners' motivation and vocabulary retention . Similarly, Umar et al. [5] demonstrated that AR could enhance both vocabulary and speaking skills among tourism employees. Rizkina developed an AR Application for teaching ESP Vocabulary [6]. Kurnia et al. [7] confirmed that AR was highly effective in teaching English vocabulary to seventh-grade students. These findings suggest that AR has great potential to overcome the challenges faced by students in vocabulary and pronunciation learning by providing immersive and interactive learning environments.

Along with technological developments, various innovative learning media have been developed to support the language learning process. Mayer[8] explains that multimedia-based learning can enhance understanding by integrating text, audio, and visuals simultaneously. One potential technology is Augmented Reality (AR). AR allows the integration of virtual objects into the real world in real time, creating a more immersive and interactive learning experience.[9]. Several studies have shown the effectiveness of AR in language learning. Nurjanah[10] proved that the use of AR in English learning at the elementary school level can significantly improve students' vocabulary mastery. Basri [11] also reported that AR can enhance vocabulary retention through contextual 3D object visualization. In addition, study [12] integrated AR with voice recognition for children's English learning, allowing students to practice pronunciation while receiving voice feedback. In the context of pronunciation, Fadly [1] tested an AR application for improving vowel sound pronunciation among university students, and the results showed a significant improvement in pronunciation clarity. These findings confirm that AR is not only beneficial for enriching vocabulary but also effective in training phonological aspects in language learning. However, most previous AR-based English learning applications have concentrated on vocabulary recognition or object visualization, with limited integration of pronunciation training and contextual vocabulary practice. Although the use of AR in English learning continues to grow, few studies have specifically developed AR media designed to help students practice the pronunciation of everyday words, particularly those related to household items. Vocabulary such as refrigerator, microwave, and scissors is often unfamiliar to students at the junior high school level, highlighting the need for contextual and interactive learning tools that also provide accurate pronunciation guidance.

This study focuses on developing AR-based learning media designed to help junior high school students learn English vocabulary and pronunciation within the context of household objects. The application displays 3D objects along with their pronunciation sounds when the marker is recognized by the device's camera. The evaluation is conducted using Black Box Testing to examine system functionality and the System Usability Scale (SUS) to assess user acceptance of the developed application.

2. Research Methodology

2.1 Research Flowchar

This study uses the Multimedia Development Life Cycle (MDLC) development model, which consists of six stages: Concept, Design, Material Collecting, Assembly, Testing, and Distribution[13].

1. Concept
   The concept stage sets the foundation for the project, focusing on defining the primary goal and objectives. In this phase, the aim was clear: to develop an interactive AR-based learning application to improve English vocabulary and pronunciation among junior high school students. The authors identified the need for a tool that not only teaches vocabulary but also provides an engaging and immersive way to practice correct pronunciation. This involved initial research into the current gaps in existing learning media, particularly the lack of AR applications focused on pronunciation and real-life vocabulary objects.
2. Design
   The design phase involves planning the structure and flow of the application. For this, the authors created detailed flowcharts and storyboards. The flowchart specifically outlines the entire user journey, starting from the opening of the app to interacting with 3D objects and listening to their pronunciation. This step ensured that all system interactions were logically structured and user-friendly. Additionally, storyboards were created to visualize the user experience, showcasing how users would engage with the AR content and how the app's visual elements would interact with physical markers.
3. Material Collecting
   This phase focuses on gathering the necessary assets for the application. For this study, the authors collected 3D models of household objects such as a refrigerator, microwave, and scissors, which were selected as the key vocabulary items for pronunciation practice. The authors also recorded pronunciation audio in MP3 format, ensuring each word was pronounced clearly and accurately. These materials were sourced from both open libraries for the 3D models and professional recordings to maintain high-quality audio. The 3D assets were carefully prepared for integration with the AR system, ensuring compatibility with the Unity engine and Vuforia SDK.
4. Assembly
   The assembly phase involves integrating the collected materials into the application. The authors used Unity in combination with the Vuforia SDK to develop the AR application. Unity provided the development environment, while Vuforia was utilized to manage the marker recognition system. During this phase, the

authors integrated the 3D objects, set up the camera functionality for marker scanning, and embedded the audio files for pronunciation playback. Special attention was given to ensuring smooth performance across different devices, optimizing the 3D models to prevent lag during AR rendering. This stage required a lot of iterative work, particularly in adjusting the spatial positioning of objects to align correctly with the markers.

5. Testing
   Testing was carried out in two phases:
   Black Box Testing: This functional testing method was employed to ensure that all features worked according to the application design. The authors tested features such as the marker recognition system, 3D object rendering, and audio playback. The focus was entirely on the app's output in response to user inputs, without delving into the underlying code structure. Through this method, the authors were able to confirm that the system was operating as expected, and identified minor bugs that were promptly fixed.
   System Usability Scale (SUS): To evaluate the application's user experience, a SUS questionnaire was distributed to 20 junior high school students. The feedback helped assess the ease of use, intuitiveness, and overall satisfaction with the application. The authors analyzed the results and used the feedback to make final improvements to the app's user interface and interaction design.

6. Distribution
   The final phase involved releasing the application for broader testing and feedback. The application was uploaded to a testing platform and made available for use by a small group of students and English teachers. These users were encouraged to explore the app, interact with its various features, and provide feedback on its functionality and usability. This phase allowed the authors to gather critical data on the app's effectiveness in real classroom settings and make further refinements based on the insights provided by users.

2.2 System Architecture and Flow

The application workflow includes the following stages: the camera captures the marker → the system recognizes the marker → a 3D object is displayed → the pronunciation audio is played. The AR working mechanism can be seen in Figure 2.1.

As shown in the Figure 1, the AR used in this study is marker-based AR, where the 3D image appears when the AR camera is directed at the logo or marker to be displayed in 3D. After that, the 3D image along with the sound button will appear on the smartphone screen.

2.3 Application Flowchart

This augmented reality (AR) application aims to enhance English pronunciation learning by visualizing 3D objects and corresponding audio. The system consists of three primary options: Play, Settings, and Exit, all accessible from the main menu as seen in Figure 2.

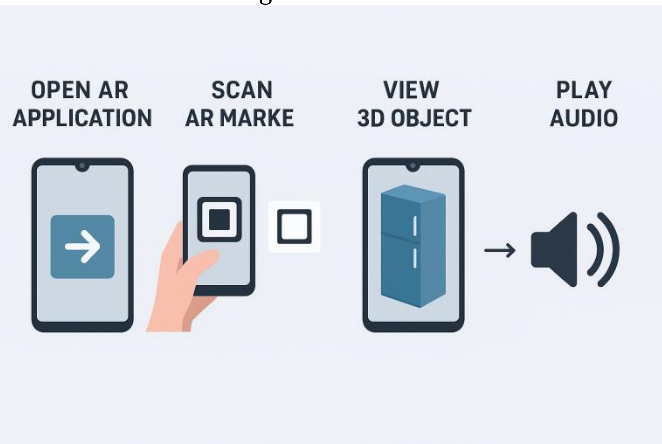

Figure 1. AR Working Process

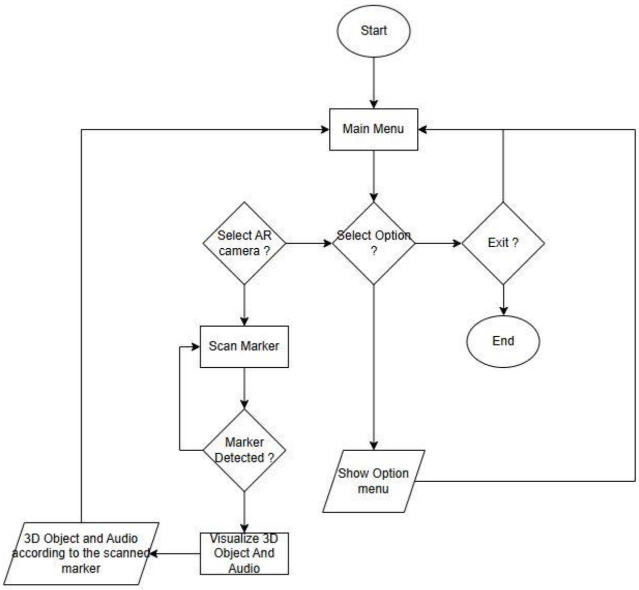

Figure 2. Flowchart

The process flow is as follows:

1. Start Screen:
   The user begins by launching the application, which leads to the main menu.
2. Main Menu:
   Upon accessing the main menu, the user is presented with three choices:
   - Play: Initiates the AR camera to visualize 3D objects for learning.
   - Option: Allows the user to modify the app's configuration (such as sound settings or camera preferences).
   - Exit: Closes the application.
3. Selecting Play:
   When the Play option is selected, the user is prompted to choose an AR camera. This step ensures the correct camera view for scanning the markers, which are essential for displaying the 3D objects.
4. Scan Marker:
   The AR system requires the user to scan a designated marker (e.g., a printed QR code or image). The marker serves as a trigger to display the appropriate 3D object.
5. Marker Detection:
   If the marker is detected, the system proceeds to visualize the selected 3D object. In this case, the options include objects like a refrigerator, microwave, or scissors, each representing a specific English vocabulary term for pronunciation practice.
6. 3D Object and Audio Visualization:
   Upon successful detection of the marker, the system displays the 3D object in augmented reality. Additionally, a corresponding audio file is played, allowing the user to hear the correct pronunciation of the English word associated with the object.
7. Exit Option:
   If the user selects Exit from the main menu or decides to stop using the app, the system will close the application, ending the process.
8. Show Option Menu:
   If the user opts for any other option besides Exit, the system displays additional settings or options based on the previous choice.

In summary, the app provides an interactive and engaging learning environment by integrating AR to visualize real-world objects while playing pronunciation audio. This approach enhances the learning experience, making it more immersive and effective for language learners.

2.4 Dataset

The dataset used in this research consists of three household objects—refrigerator, microwave, and scissors—which were selected to represent contextual daily vocabulary commonly taught at the junior high school level. These items were specifically chosen to maintain system simplicity and focus during user testing while still reflecting essential household vocabulary. The selection aligns with prior AR-based English learning studies that also employed a limited set of contextual objects to ensure usability and accuracy. Each object was modeled in 3D format and paired with pronunciation audio recorded by a native English speaker to ensure consistency in articulation and clarity.

2.5 Testing

Black Box Testing: Black Box Testing is a functional testing method used to verify whether the software performs according to its specifications without considering the internal code structure or program logic. In this approach, the tester focuses solely on the input and output of the system to ensure that all features function as intended.

This method was applied in the present study to test the main components of the AR learning application, including marker detection, 3D object visualization, and audio pronunciation playback. Each feature was tested by providing input (e.g., scanning a marker) and observing the system's response (e.g., displaying the correct 3D object and playing its pronunciation). The results indicated that all features performed as expected, and no functional errors occurred during the testing process.

System Usability Scale (SUS)

The System Usability Scale (SUS) is a standardized questionnaire developed by John Brooke [14] to evaluate the usability of a system through user perceptions of ease of use, efficiency, and satisfaction. The SUS consists of ten statements, each rated on a five-point Likert scale ranging from Strongly Disagree (1) to Strongly Agree (5). The final SUS score is calculated by summing the adjusted item scores and multiplying the total by 2.5, yielding a value between 0 and 100.

A SUS score above 68 is generally considered above average, while scores above 80 indicate an excellent level of usability [15]. In this study, the SUS was employed to assess the ease, attractiveness, and overall usability of the AR application among 20 junior high school students.

The final SUS score of 82.5 places the application in the "Excellent" category, indicating that users found the system intuitive, engaging, and effective for learning English vocabulary and pronunciation.

3. Results and Discussions

3.1. Application Implementation

The developed application consists of a menu with three buttons: the camera button, the option button, and the exit button. When a button is pressed, it directs the user to a different page. The menu page can be seen in Figure 3. When the camera button is pressed, the application opens the AR camera, which can be used to scan the marker. After the marker is scanned, a 3D image along with a button labeled with the object's name appears on the camera view, positioned precisely above the provided marker. When the button with the object's name is pressed, the pronunciation sound of that object is played. The complete display can be seen in Figure 4.

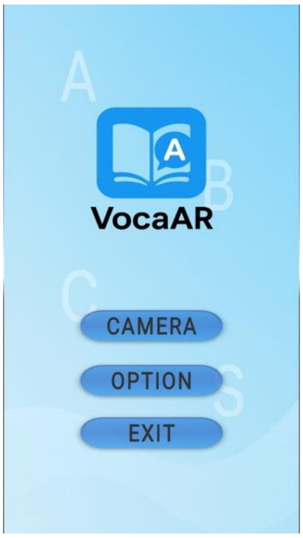

Figure 3. Main Menu

Figure 4. AR Camera

3.2. Testing result

The black box testing phase was conducted to evaluate the functionality of the software by executing various input data and observing the corresponding outputs. During this process, data were collected from 10 high school students who participated in the application testing to ensure the accuracy and reliability of the results.

The aspects evaluated during the testing phase were as follows:

1. Page View
   This test aimed to verify that every page of the application was displayed correctly and functioned smoothly without any visual or navigation issues.
2. Application Sound
   This aspect assessed whether all audio components—including pronunciation voices, button click sounds, and background music—could be heard clearly without distortion or interference.
3. Menu Buttons
   The functionality of all menu buttons was examined to ensure that each responded accurately to user input and executed the intended commands properly.
4. Augmented Reality Camera
   Testing in this area focused on ensuring that the AR camera operated correctly, detected markers efficiently, and displayed the corresponding 3D animations without lag or malfunction.
5. Augmented Reality Markers
   This test aimed to determine whether the AR markers could be accurately recognized by the camera and whether the correct 3D objects appeared according to the selected category or menu option.

Table 1 presents the results based on the tests conducted on ten respondents on those five aspects.

Table 1. Blackbox Result

System Feature	Expected Result	Actual Result
Page View	The page loads smoothly without any interruptions.	Pass
Sound	Pronunciation audio, button sounds, and background music play clearly.	Pass
Button	Buttons function correctly.	Pass
AR Camera	The AR camera operates properly, successfully scanning markers and displaying 3D objects.	Pass
Marker	Markers are detected and recognized by the AR camera without issues.	Pass

Table 2. SUS Rating

SUS Score	Grade	Descriptive Rating
> 80.3	A	Excellent
68%—80.3	B	Good
68	C	Okay
51—68	D	Poor
< 51	E	Awful

Table 4. SUS scores

No	Questions	Average SUS Score (Pre-test)	Average SUS Score (Post-test)
1	I am willing to use this application again.	74.00	84.50
2	Using this app feels complicated.	75.50	87.00
3	This application is straightforward and user-friendly.	65.00	75.00
4	I require technical support to operate this application.	73.00	85.00
5	The features of this application are well-coordinated.	75.50	84.00
6	There are several inconsistencies in the application's functionalities.	72.00	78.00
7	I believe many people would quickly learn to use this application.	69.00	80.00
8	This application is challenging to operate.	85.00	87.00
9	I feel confident when navigating this application.	72.00	82.00
10	I need to learn many things before I can use this application effectively.	76.00	92.00
	Average SUS Score	73.70	83.45

3.3. System Usability Testing

The application's usability was assessed using the System Usability Scale (SUS). This survey contains ten items and was completed by the same group of twenty-one junior high school students during both the pre-test and post-test stages. Once the responses were collected, the data were analyzed to determine the overall usability score. The total SUS score for each participant was summed and then divided by the number of respondents to obtain the average. The formula applied to compute the mean SUS score is as follows:

$$\tilde{n} = \frac{\Sigma x}{n} \tag{1}$$

where :

x Average score
Σ Sum of the total SUS scores
n Number of participants

The grade and adjective rating of the average SUS score is presented in Table 2.

Table 4 presents the SUS scores obtained from the pre-test and post-test evaluations. The pre-test score represents the SUS value measured when the application was initially released, while the post-test score reflects the value after improvements were made to the application. Based on the questionnaire results, the application achieved a SUS score of 73.70 during the pre-test, indicating that it was already good for use. After the enhancements, the post-test SUS score increased to 83.45, signifying that the application reached an excellent level of usability.

3.4 Discussions

The findings indicate that the AR-based application effectively supports interactive English learning by combining 3D visualization and auditory cues (Chen et al., 2022; Li & Chen, 2023). The improvement in SUS scores suggests that the system refinements successfully enhanced usability and user satisfaction. Students reported increased motivation and engagement during learning sessions, emphasizing the potential of AR for immersive vocabulary and pronunciation learning (Lee & Kim, 2023). However, the current study is limited to a small set of vocabulary items. Future work should expand object categories and integrate automatic speech recognition for pronunciation evaluation.

4. Conclusion

This study successfully developed an Augmented Reality (AR)-based learning application designed to assist junior high school students in learning English vocabulary and pronunciation within the context of household objects. The application integrates 3D visualization and pronunciation audio to create an interactive and engaging learning experience

Based on the Black Box Testing results, all system components—including page display, audio playback, button functionality, AR camera performance, and marker recognition—functioned correctly without any technical issues. This indicates that the application operates in accordance with its intended design and specifications.

Furthermore, the System Usability Scale (SUS) evaluation showed a significant improvement in user satisfaction, with the SUS score increasing from 73.70 (Good) in the pre-test to 83.45 (Excellent) in the post-test after system enhancements. These findings demonstrate that the AR-based learning media is not only functional but also highly usable and well-received by students.

In conclusion, the developed AR application proves to be an effective and user-friendly tool for enhancing students' English vocabulary and pronunciation learning. Future research could expand the content to include more diverse vocabulary categories and incorporate interactive feedback features to further improve pronunciation accuracy and learning engagement.

References

[1] R. Fadli, "Using augmented reality application toward university students to improve pronunciation on vowel sounds," *Indonesian Journal of Research and Service Studies*, vol. 1, no. 1, pp. 1–19, 2024.

[2] A. S. Susanto, S. N. Ardini, and R. Sukmaningrum, "Dictionary Speech Assistant to Improve Students' Pronunciation," *Allure Journal*, vol. 4, no. 1, pp. 14–22, 2024, doi: 10.26877/allure.v4i1.17267.

[3] E. Emilia, I. Ismail, and H. Hasan, "Utilizing Augmented Reality Technology in English Vocabulary," *Seltics Journal: Scope of English Language Teaching Literature and Linguistics*, vol. 8, no. 1, pp. 44–55, 2025, doi: 10.46918/seltics.v8i1.2602.

[4] A. F. ʻaini, N. Tarihoran, and Kheryadi, "Can Augmented Reality Revive Vocabulary Learning ? A Study on Motivation Among Secondary EFL Learners," *Voices of English Language Education Society*, vol. 9, no. 2, pp. 243–257, 2025, doi: 10.29408/veles.v9i2.30339.

[5] U. Umar, Y. Yuliadi, M. Hidayatullah, A. Rahman, and Z. Hassan, "Augmented Reality for Vocabulary Learning And Speaking Skills of Tourism Employees," *JOLLT Journal of Languages and Language Teaching*, vol. 13, no. 2, pp. 811–823, 2025, doi: 10.33394/jollt.v13i2.13860.

[6] P. A. Rizkina, T. Suwartono, S. Nurhayati, O. Wijayanti, and F. F. Dalimarta, "Creating AR Application for Advanced ESP Vocabulary Teaching in Indonesia," *International Journal of Learning, Teaching and Educational Research*, vol. 24, no. 5, pp. 679–705, May 2025, doi: 10.26803/IJLTER.24.5.35.

[7] I. Kurnia, Riswanto, and R. Serasi, "The Effectiveness of Augmented Reality in Teaching English Vocabulary to Seventh-Grades Students," *Journal of Innovation in Teaching and Instructional Media*, vol. 6, no. 1, pp. 372–378, 2025, doi: 10.52690/jitim.v6i1.1291.

[8] C. I. Johnson and R. E. Mayer, "A Testing Effect With Multimedia Learning," *Journal of Educational Psychology*, vol. 101, no. 3, pp. 621–629, 2009, doi: 10.1037/a0015183.

[9] R. T. Azuma, "Making Augmented Reality a Reality," *Optics InfoBase Conference Papers*, 2017, doi: 10.1364/3D.2017.JTu1F.1.

[10] A. Nurjanah, "International Journal of Pedagogy and Teacher Education The Impact of Augmented Reality (Ar) on Vocabulary Acquisition in Elementary School Students," *International Journal of Pedagogy and Teacher Education*, vol. 9, no. 1, pp. 16–29, 2025, doi: 10.20961/ijpte.v9i1.97639.

[11] H. Basri, J. Badriyah, M. Roda, and N. Hidayati, "Pemanfaatan Augmented Reality untuk Meningkatkan Retensi Kosakata Bahasa Inggris pada Siswa Sekolah Dasar," *Jurnal Krisnadana*, vol. 4, no. 1, pp. 46–51, 2024, doi: 10.58982/krisnadana.v4i1.786.

[12] T. Ari Wibowo and N. Nasrullah, "the Effect of Pronunciation Training on English Speaking Skill Among English Non-Native Speakers," *Exposure : Jurnal Pendidikan Bahasa Inggris*, vol. 13, no. 1, pp. 92–100, 2024, doi: 10.26618/exposure.v13i1.10740.

[13] A. C. Luther, "Authoring interactive multimedia," *Press Professional, Inc*, 1994.

[14] J. Brooke, "SUS-A quick and dirty usability scale," *Usability evaluation in industry*, 1996.

[15] Bangor, Aaron, P. T. Kortum, and J. T. Miller., "An empirical evaluation of the system usability scale," *Intl. Journal of Human–Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008, doi: 10.1080/10447310802205776.