



Augmented Reality with Speech Recognition for Animal Recognition Applications

Aldia Deponsa Judistira ^a, Agus Pratondo ^{a,*}, Anang Sularsa ^a

^a School of Applied Science, Telkom University, Indonesia

aldidi@student.telkomuniversity.ac.id, pratondo@telkomuniversity.ac.id, ananks@telkomuniversity.ac.id

ARTICLE INFO

Received December 25th, 2021
Revised January 10th, 2022
Accepted April 13th, 2023
Available online June 9th, 2023

Keywords

Augmented reality, speech recognition, learning activity, PGRI Sumedang Kindergarten

ABSTRACT

The advance of educational technology is significantly changing and is very effective in supporting education nowadays. However, technology to support education in Indonesia is not maximal; one of the cases occurred in the PGRI Sumedang Kindergarten. Teachers at PGRI Sumedang Kindergarten have difficulties providing material that can build students' interest, especially general knowledge about animals. This study aims to design Augmented Reality (AR) technology features with Speech Recognition for an Android-based interactive multimedia application of animal knowledge to assist learning activities in PGRI Sumedang Kindergarten. AR is a computer technology that combines virtual objects with the real world. This AR feature uses a card-shaped marker with an image and speech recognition as a medium for interaction with animals. Speech recognition is a computer technology to process human speech into digital letters or words. The resulting interaction is that if the user pronounces the animal's name displayed in AR, the animal animation will move. This AR feature can help to learn activities at Kindergarten PGRI Sumedang.

* Corresponding author at:
School of Applied Science, Telkom University.
Jl. Telekomunikasi No. 1, Terusan Buah Batu, Bandung, 40257.
Indonesia.
E-mail address: pratondo@telkomuniversity.ac.id

ORCID ID:

- First Author: 0000-0003-1181-4213
- Second Author: 0000-0002-6976-7459

<https://doi.org/10.25124/ijait.v6i02.4351>

Paper_reg_number IJAIT000060202 2022 © The Authors. Published by School of Applied Science, Telkom University.
This is an open access article under the CC BY-NC 4.0 license (<https://creativecommons.org/licenses/by-nc/4.0/>).

1. Introduction

The development of technology to support education has been significantly advanced. However, technology to support education in Indonesia is currently not maximized [1]. One of the cases that occurred in this problem was PGRI Sumedang Kindergarten. Based on the results of interviews with the principal of the PGRI Sumedang Kindergarten, there are teaching staff who have difficulty conveying material so students lack interest in learning general knowledge, one of which is the introduction of animals. Interest in learning is the interest of students in the studying process they are undergoing; with interest in education, students can quickly achieve success in their knowledge [2].

Based on the results of interviews with the principal of the PGRI Sumedang Kindergarten, the principal wanted an interactive multimedia feature for children as a learning medium. This research is a proposal to help learning activities and foster interest in learning for PGRI Sumedang Kindergarten students with augmented reality (AR) technology. AR is a technology that inserts virtual objects processed by a computer into the real world in real-time [3][4]. This AR feature can interact by using speech recognition as an interaction. Speech recognition is a computer technology that can convert voice signals from human speech into digital letters [5][6]. Implementation of this technology can make learning activities more interactive [7-12].

Based on the previous description, the study entitled AR feature design with speech recognition for interactive multimedia animal recognition applications aims to build AR features for Android-based ZooFun applications based on the PGRI Sumedang Kindergarten curriculum. The AR feature uses an illustrated marker in the form of a card or book to be scanned to bring up 3D objects on the screen of an Android smartphone device. The interaction that occurs is when the user mentions the name of the animal displayed on AR in English, and the animation of the 3D animal moves. The AR interaction provided will help increase interest in learning in TK PGRI Sumedang students.

The remainder of this journal is structured as follows: Section I explains the context and issue of this journal, the results of previous studies, the works related to the method similar to the method proposed, and presents the works proposed for this journal; section II describe MDLC method and it's steps; section III describe the technical implementation, testing, and result; and the conclusions of this paper is described in section IV

2. Methods

AR features in this interactive multimedia application use the Multimedia Development Life Cycle (MDLC) method. The diagram can be seen in Figure 1 below. This method is used because it is suitable for building an interactive multimedia application for quality education. MDLC has six stages: concept, design, material collecting, assembly, testing, and distribution [13][14][15][16].

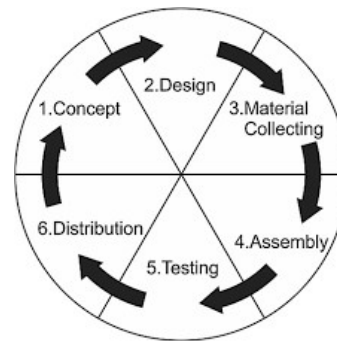


Figure 1 MDLC Methods

2.1. Concept

This process carries out making to collect the needs in making AR features. The purpose and results of the features created must be appropriate and meet partners' needs.

2.2. Design

At the design stage, the needs that will be made to build AR features are described. The AR feature contains two main functions for interacting with 3D objects: virtual buttons and speech recognition. The interaction that occurs is running animated 3D objects on the smartphone screen.

2.3. Material Collecting

The third stage is making the need for animal recognition AR features. This stage includes making 3D objects, animation, and AR markers according to the needs obtained at the concept stage. The software used in making the need for animal recognition AR features that includes Blender 2.90 and Adobe Photoshop CC.

2.4. Assembly

This stage is the compositing of the component requirements that have been prepared—making animal recognition AR features must be following the concepts that have been designed and meet the criteria and needs of these AR features. The assembly process is carried out in Unity Engine software.

2.5. Testing

This stage is the testing stage of the features to be created. Tests are carried out on PC in Unity software, and real-time tests will be installed on Android devices. Tests are carried out to avoid errors or defects in the AR feature. Testing is done with two types of testing: alpha testing and beta testing.

2.6. Distribution

The finished AR feature will be installed in the ZooFun animal interactive multimedia application at this stage. After that, the ZooFun application will be submitted to PGRI Sumedang Kindergarten.

3. Implementation

3.1. Concept

The results of interviews with the principal of PGRI Sumedang Kindergarten, designing the AR features of this interactive multimedia application is to help teachers deliver material and foster student interest in learning in PGRI Sumedang Kindergarten interactively.

Based on the results of the interviews, the conclusions can be seen in Table 1 below:

Table 1 Interview Results with The Principal of PGRI Sumedang Kindergarten

No.	Needs
1	The animal objects created are domestic animals and wild animals with their habitats both on land and at sea.
2	Interesting medium of interaction for children in a colored cards form which also functions as an AR marker
3	It was necessary to provide voice/speech recognition and interaction in English.

3.2. Design

The design of the AR features can be seen in the description below:

3.2.1. 3D Object

The Object displayed is a 3D virtual object in the form of an animal with a low-poly style and is animated. The number of 3D animal objects created is 20 consisting of 12 land animals and eight marine animals. Each Object has two animations, which are idle when the animal is still and trigger when the animal interacts with the users.

3.2.2. Markers

Making AR markers is given a different color theme. Land animals are given an orange color according to the theme color of the application, and the accent color of land animals is also distinguished. Herbivores are colored green, while carnivores are colored red. Marine animals are given a blue color showing the nuances of the sea.

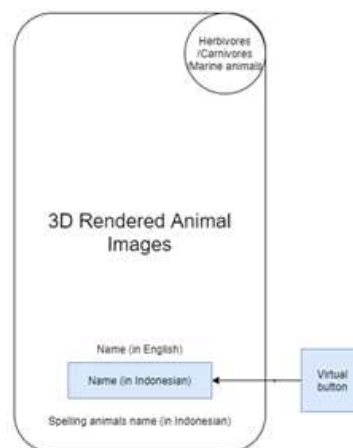


Figure 2 Mock-Up for AR Marker

The color composition uses pastel colors to not be too contrast and soft. The marker has a virtual button located on the animal's name in Indonesian. If the writing is blocked from in front of the smartphone camera, the 3D animal object will move. The layout can be seen in Figure 2.

3.3. Material Collecting

Material collecting is the initial stage of implementing the animal recognition AR feature.

3.3.1. 3D Object

3D objects are made entirely from modeling, rigging, weight painting, and painting materials. Creating 3D objects using Blender 2.90 software. The creation of 3D objects uses a low-poly style not to take up too much of the implemented device's resources. The total number of 3D animal objects is 20 species.

After rigging, the animation uses the bone rigging method, merging the bone with the Object using weights. Blender has a feature of auto-weighting with auto-weight, but because it's not perfect, weight-painting must be done again to adjust the 3D Object with the bone. The animation speed used in units of frames per second (FPS) is 60FPS, so that the animation looks smoother. The coloring process uses painting material with a solid color by adjusting the color of the original animal. The process of making 3D objects can be seen in Figure 3 below.

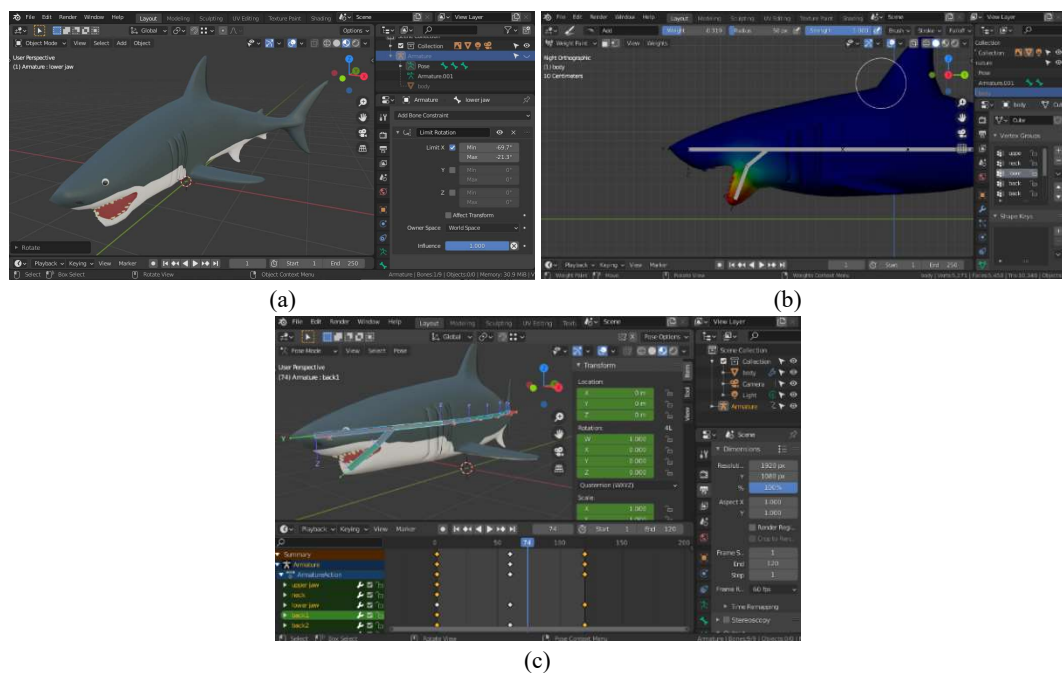
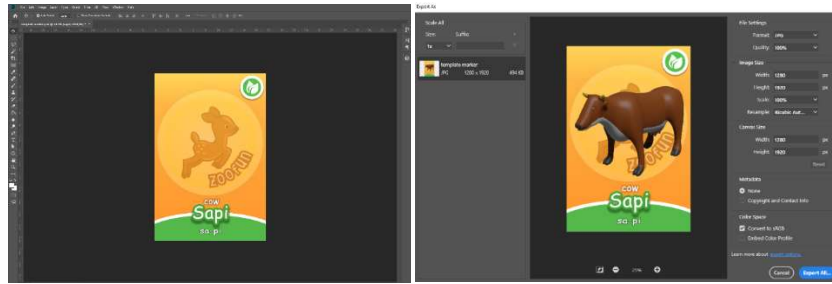


Figure 3 Modeling and Painting 3D Object (a), Rigging and Weight Painting (b), and Animating (c).

3.3.2. Marker

The marker is made in 1920x1280 pixels, making it sharper and more accessible for smartphones to scan later. Then the marker will be submitted in the form of a print-ready file.

1. Design the AR marker layout, and insert the 3D rendered image. The file is exported in a .jpg file as seen in Figure 4 below.

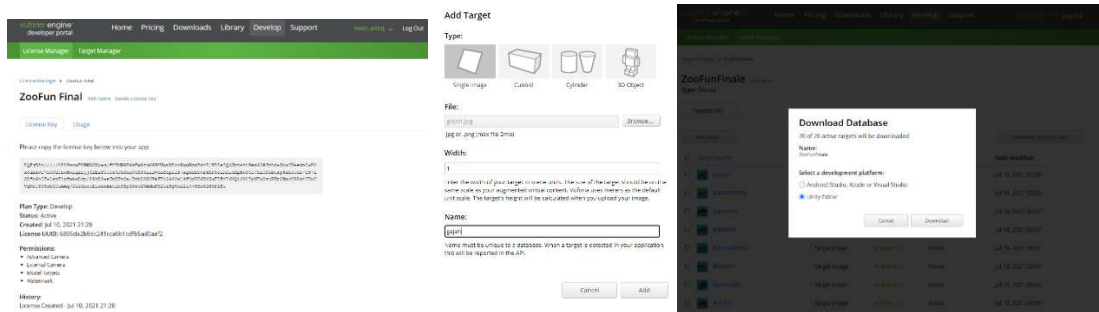


(a)

(b)

Figure 4 Designing Marker Layout (a), Insert 3D Rendered Animal, and Export File (b).

2. Create a license key on the Vuforia developer portal for use on Unity Engine. Upload the markers to make the target object. The database will be downloaded to be imported into Unity Engine. The steps can be seen in Figure 5 below.



(a)

(b)

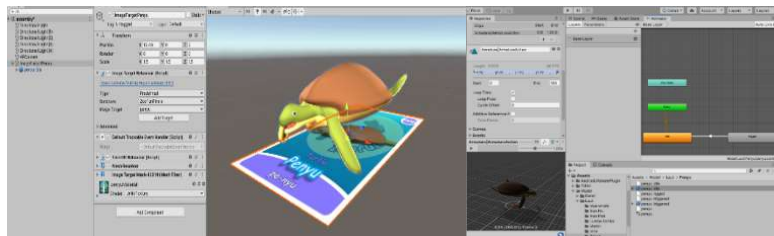
(c)

Figure 5 Create License Key (a), Upload The Marker File (b), and Download The Database (c).

3.4. Assembly

The following stage is the assembly process of all the materials that have been collected in the previous stage. Assembly process using Unity Engine software.

1. Use the Vuforia tools provided in Unity Engine. Put the object 3D on the upside of the marker. 3D animation must be arranged first in animator controller depiction can be seen in Figure 6 below.



(a)

(b)

Figure 6 Assembly of the 3D objects with a marker (a), and arrangement of The Animation Controller (b)

2. Add the virtual button and speech recognition features. Both features use Zetcil Framework. The process can be seen in Figure 7 below.

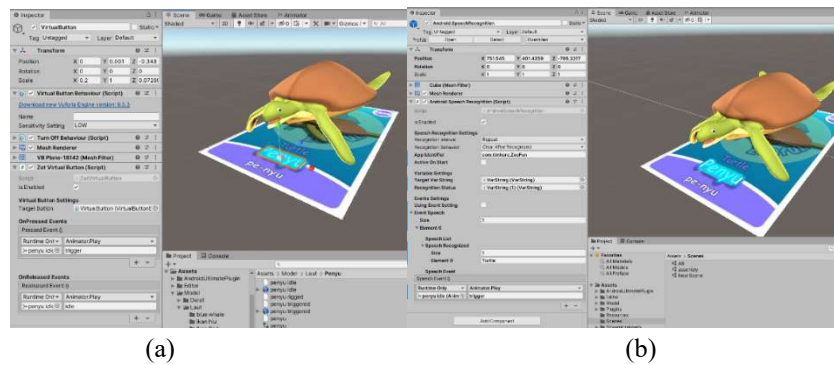


Figure 7 Add Virtual Button Interaction (a) and Speech Recognition (b).

3. Create a user interface (UI) with a portrait layer orientation. The microphone button is used to start voice recording for interaction, the Back button is used to return to the main menu, and the Help button is used to open instructions for using the feature. The UI can be seen in Figure 8.

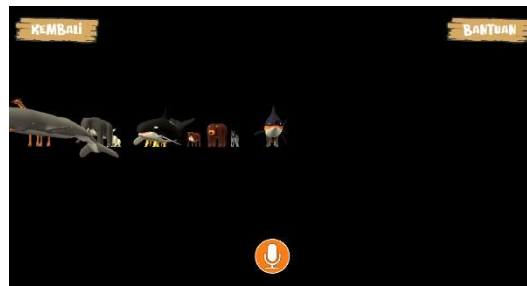


Figure 8 Create User Interface

3.5. Testing

At this stage, the aim is to determine whether the functionality of the AR feature is suitable for use by users. The testing process is carried out in two stages: Alpha Testing and Beta Testing.

3.5.1. Alpha Testing

Alpha testing using the black box testing method. The testing phase is carried out using an Android smartphone. Table 2 below is a scenario and the results of black-box testing.

Table 2 Black Box Testing Result

No.	Testing activity	Expected results	Testing Results	Conclusion
1	Test marker scan	The 3D animal appears in an idle state after successfully scanning the marker.	3D animals appear in an idle state.	Accepted
2	Test microphone button	When pressed, the system starts scanning the sound pattern.	The system started scanning the sound pattern after the button was pressed.	Accepted
3	Testing speech recognition interaction	If the animal's name matches the marker, then the 3D animal moves.	3D animals move while correctly pronouncing animal names.	Accepted
		Pronouncing a different vocabulary beyond the systems will cause no interaction.	No interaction occurs when mentioning other vocabulary words.	Accepted
4	Testing virtual button interactions	When an animal's name in Indonesian is blocked from the camera's view, the 3D animal moves.	The 3D animal moves when pressing the animal's name in the marker.	Accepted

3.5.2. Beta Testing

The Beta-testing phase involved 21 school residents from PGRI Sumedang Kindergarten, parents, and the general public. Testing is done by giving a questionnaire that contains six questions related to the AR feature. The test results are calculated from each question and answer using a 5-point Likert scale, resulting in a percentage of the feasibility of the AR feature. Score interpretation criteria based on interval as shown in Table 3.

Table 3 Score Interpretation

Interval	Category
0 % - 19.99 %	Very dissatisfied
20 % - 39.99 %	Dissatisfied
30 % - 59.99 %	Neutral
60 % - 79.99 %	Satisfied
80 % - 100 %	Very satisfied

The formula of the Likert scale is as aequation (1)

$$P = \frac{S}{Ideal\ score} \times 100 \tag{1}$$

Information:

P: Searched percentage

S: The number of answer frequencies multiplied by the answer scale.

Ideal score: The highest answer value multiplied by the number of samples.

The following are the test results of each question that has been calculated using a Likert scale.

1. How is the learning activities experience after using the AR feature?

Table 4 First Question Results

Question	Answer	Score	Frequency	Total
1	Very satisfying	5	4	20
	Satisfied	4	15	60
	Neutral	3	2	6
	Dissatisfied	2	0	0
	Very dissatisfied	1	0	0
Total			105	86

$$P = \frac{86}{105} \times 100 = 81,9\%$$

Based on Table 4 above, it can be concluded that the experience of learning activities is very satisfied, getting better with the AR feature of animal recognition, proof that the material is conveyed well.

2. How is the quality of interaction in using AR features?

Table 5 Second Question Results

Question	Answer	Score	Frequency	Total
2	Very satisfied	5	3	15
	Satisfied	4	10	40
	Neutral	3	8	24
	Dissatisfied	2	0	0
	Very dissatisfied	1	0	0
Total			105	79

$$P = \frac{79}{105} \times 100 = 75,23\%$$

Based on Table 5 above, it can be concluded that users are satisfied with interacting with the animal recognition AR feature.

3. How difficult is it to use AR features?

Table 6 Third Question Results

Question	Answer	Score	Frequency	Total
3	Very satisfied	5	5	25
	Satisfied	4	8	32
	Neutral	3	7	21
	Dissatisfied	2	1	2
	Very dissatisfied	1	0	0
Total			105	80

$$P = \frac{80}{105} \times 100 = 76.19\%$$

Based on Table 6 above, it can be concluded that the AR feature of animal recognition is not difficult to use.

4. How long does it take to get used to using AR features?

Table 7 Fourth Question Results

Question	Answer	Score	Frequency	Total
4	Very satisfied	5	7	35
	Satisfied	4	10	40
	Neutral	3	3	9
	Dissatisfied	2	1	2
	Very dissatisfied	1	0	0
Total			105	86

$$P = \frac{86}{105} \times 100 = 81,9\%$$

Based on Table 7 above, it is easy for users to use the animal recognition AR feature.

5. What do you think about assessing 3D animals in AR both in terms of objects and animations?

Table 8 Fifth Question Results

Question	Answer	Score	Frequency	Total
5	Very satisfied	5	5	25
	Satisfied	4	13	52
	Neutral	3	3	9
	Dissatisfied	2	0	0
	Very dissatisfied	1	0	0
Total			105	86

$$P = \frac{86}{105} \times 100 = 81,9\%$$

Based on Table 8 above, the 3D animal objects displayed are good in terms of object appearance and animation.

6. How are the learning outcomes after using the AR feature?

Table 9 Sixth Question Results

Question	Answer	Score	Frequency	Total
6	Very satisfied	5	6	30
	Satisfied	4	9	36
	Neutral	3	6	18
	Dissatisfied	2	0	0
	Very dissatisfied	1	0	0
Total			105	84

$$P = \frac{84}{105} \times 100 = 80\%$$

Based on Table 9 above, it can be concluded that the learning outcomes after using the AR feature are very satisfying, proof that the material is conveyed well.

3.6. Distributions

The following are the results of the animal recognition AR feature that has been installed in the ZooFun application on Android smartphones. The application will be distributed to PGRI Sumedang Kindergarten, including the marker in the form of a print-ready file.

4. Conclusions

The results of the implementation and testing can be concluded that the animal recognition AR feature has met the requirements and met the initial objectives of the study. Based on the test results, the animal recognition AR feature can help to learn activities at PGRI Sumedang Kindergarten to become more interactive for students.

Bibliography

- [1] A. Akbar and N. Noviani, "Tantangan dan Solusi dalam Perkembangan Teknologi Pendidikan di Indonesia," *Pros. Semin. Nas. Pendidik. Progr. Pascasarj. Univ. PGRI Palembang*, vol. 2, no. 1, pp. 18–25, 2019.
- [2] S. Larasati Lubis and Rohmani, "Makalah Minat Belajar." p. 13, 2020.
- [3] R. Indra Borman, "Implementasi Augmented Reality Pada Aplikasi Android Pegenalan Gedung Pemerintahan Kota Bandar Lampung," *J. Teknoinfo*, vol. 11, no. 1, p. 10, 2017.
- [4] R. T. Azuma, "A survey of Augmented Reality," *Teleoperators Virtual Environ.*, vol. 6, no. 4, pp. 355–385, 1997.

- [5] A. Kumar and V. Mittal, "Speech recognition: A complete perspective," *Int. J. Recent Technol. Eng.*, vol. 7, no. 6, pp. 78–83, 2019.
- [6] M. M. Mustaquim, "Automatic speech recognition- an approach for designing inclusive games," *Multimed. Tools Appl.*, vol. 66, no. 1, pp. 131–146, 2013.
- [7] B. Setyawan, Nf. Ruffi, and A. N. Fatirul, "Augmented Reality Dalam Pembelajaran Ipa Bagi Siswa Sd," *Kwangsan J. Teknol. Pendidik.*, vol. 7, no. 1, pp. 78–90, 2019.
- [8] F. N. Rohmah, "Media Pembelajaran Pengenalan Buah Dengan Teknologi Augmented Reality Untuk Anak Usia Dini Berbasis Android," pp. 1–102, 2019.
- [9] S. Prayoga, R. Yusup, and A. F. Rahman, "Penerapan Augmented Reality sebagai Metode Pembelajaran Taman Kanak-Kanak untuk Mengenal Binatang," *J. Tek. Inform.*, no. Juni, 2019.
- [10] F. Hadary, D. Yuniarni, and S. Wulandari, "Multimedia Interaktif dalam Pembelajaran di Taman Kanak-kanak Kota Pontianak Kalimantan Barat," *Al-Khidmah*, vol. 2, no. 2, pp. 42–48, 2019.
- [11] D. F. Afandi, "Aplikasi Pengenalan Binatang Dengan Augmented Reality Berbasis Android (Studi Kasus: Raudlotul Athfal Muslimat NU Kemasari)," 2018.
- [12] Y. C. Buana and F. A. P., "Augmented Reality For Anatomy Study With Speech Recognition," pp. 1–7, 2013.
- [13] R. I. Borman and Y. Purwanto, "Implementasi Multimedia Development Life Cycle pada Pengembangan Game Edukasi," *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 2, pp. 119–124, 2019.
- [14] Azim Muhammad Fauzan, Hidayat Eka Wahyu, Rachman Andi Nur "Android Battle Game Based on Augmented Reality with 2D Object Marker", *Jurnal Online Informatika*, Volume 3 No. 2 pp. : 116-122, 2018
- [15] Afrianto, Irawan, and Rifan Muhammad Furqon. "The Herbalist Game Edukasi Pengobatan Herbal Berbasis Android." *Jurnal Sistem Informasi Bisnis* 8.2 (2018): 141-148.
- [16] Sugiarto, Hari. "Penerapan Multimedia Development Life Cycle Pada Aplikasi Pengenalan Abjad Dan Angka." *IJCIT (Indonesian Journal on Computer and Information Technology)* 3.1 (2018).