IoT Empowered AI: Transforming Object Recognition and NLP Summarization with Generative AI

Samuel Joseph[1], Dr. Bhagavathi Priya S[2], Poorvaja R[3], Santhosh Kumaran M[4], Shivaraj S[5], Jeyanth V[6], Shivesh P R[7]
[2)]Associate Professor, Department of Computer Science and Engineering Amrita School of Computing, Amrita Vishwa Vidhyapeetham, Chennai.
[1,3,4,5,6,7)] Department of Artificial Intelligence and Machine Learning, Rajalakshmi Engineering College, Chennai.

1. sjoseph2552@gmail.com , 2. s_baghavathipriya@ch.amirta.edu, 3. rpoorvaja195@gmail.com,
4. 10santhoshkumaran12@gmail.com, 5. s.shivaraj25.ss@gmail.com, 6. ramanjeyanth3@gmail.com,
7. shiveshraguraman@gmail.com.

***Abstract* —** In anticipation of the widespread adoption of augmented reality in the future, this paper introduces an advanced mobile application that seamlessly integrates AR and IoT technologies. The application aims to make these cutting-edge technologies more affordable and accessible to users while highlighting their immense benefits in assisting with household appliance control, as well as providing interactive and educational experiences. The app employs advanced algorithms such as object detection, Natural Language Processing (NLP), and Optical Character Recognition (OCR) to scan the smartphone's camera feed. Upon identification, AR controls for appliances, their power consumption, and electric bill tracking are displayed. Additionally, the application makes use of APIs to access the internet, retrieving relevant 3D generative models, 360-degree videos, 2D images, and textual information based on user interactions with detected objects. Users can effortlessly explore and interact with the 3D generative models using intuitive hand gestures, providing an immersive experience without the need for additional hardware or dedicated VR headsets. Beyond home automation, the app offers valuable educational benefits, serving as a unique learning tool for students to gain hands-on experience. Medical practitioners can quickly reference organ anatomy and utilize its feature-rich functionalities. Its cost-effectiveness, requiring only installation, ensures accessibility to a wide audience. The app's functionality is both intuitive and efficient, detecting objects in the camera feed and prompting user interactions. Users can select objects through simple hand gestures, choosing desired content like 3D generative models, 2D images, textual information, 360-degree videos, or shopping-related details. The app then retrieves and overlays the requested information onto the real-world view in AR. In conclusion, this groundbreaking AR and IoT-powered app revolutionizes home automation and learning experiences, leveraging only a smartphone's camera, without the need for additional hardware or expensive installations. Its potential applications extend to education, industries, and healthcare, making it a versatile and valuable tool for a broad range of users.

***Keywords* — Augmented Reality(AR), Internet of Things (IoT) technologies, Object Detection, Natural Language Processing, Optical Character Recognition (OCR), Application Program Interface (API),3D, 2D, GenAI.**

I. INTRODUCTION

In recent years, the technological landscape has witnessed remarkable advancements, paving the way for innovative solutions that promise to shape our daily lives. One such exciting prospect on the horizon is the widespread adoption of augmented reality (AR) and Internet of Things (IoT) technologies. These cutting-edge technologies have the potential to revolutionize the way we interact with our surroundings, from enhancing home automation to transforming education and healthcare experiences. In anticipation of this transformative future, this paper introduces a groundbreaking mobile application that seamlessly integrates AR and IoT, bringing these powerful technologies to the fingertips of users in a cost-effective and accessible manner.

The primary focus of this advanced mobile application is to empower users with greater control over household appliances through the utilization of AR and IoT capabilities. The integration of object detection, Natural Language Processing (NLP), and Optical Character Recognition (OCR) algorithms allows the smartphone's camera feed to identify objects and appliances accurately. Upon identification, AR controls are displayed, enabling users to manage their appliances efficiently and monitor power consumption, leading to more informed decisions and better control over electric bills.

However, the application's functionality extends beyond mere home automation. Leveraging APIs to access the vast resources of the internet, the app offers a wealth of interactive and educational experiences for users. Upon interacting with detected objects, users can effortlessly explore 3D generative models, view 360-degree videos, access 2D images, and retrieve textual information relevant to their interests. What sets this application apart is its intuitive interface, which enables users to interact with virtual content through simple hand gestures, negating the need for additional hardware or dedicated virtual reality (VR) headsets. This seamless integration of AR and IoT on a ubiquitous device like a smartphone makes these transformative technologies affordable and accessible to a wide audience.

The educational potential of this application is profound, providing students with a unique and immersive learning tool. By enabling students to interact with 3D generative models and access relevant information in real time, the app fosters hands-on learning experiences that complement traditional classroom teachings. Additionally, medical practitioners stand to benefit from the app's feature-rich functionalities, allowing quick reference to organ anatomy and aiding in diagnosis and treatment decisions.

The cost-effectiveness of this AR and IoT-powered application further reinforces its potential to reach a broad spectrum of users. Unlike other AR solutions that often require expensive installations or specialized hardware, this app leverages the existing smartphone camera, making it a viable option for individuals from various socio-economic backgrounds.

This advanced mobile application marks a significant step forward in the convergence of AR and IoT technologies, promising to revolutionize home automation, education, and healthcare experiences. By making these transformative technologies accessible to users through their smartphones, it opens up a world of possibilities for industries, educators, medical professionals, and everyday users. The app's potential applications are vast, and its versatility is underscored by its capacity to cater to multiple sectors and audiences. As we embrace a future shaped by augmented reality and the Internet of Things, this application stands at the forefront, heralding a new era of interactive, intuitive, and accessible technological experiences.

II. LITERATURE REVIEW

"Smart Digital Education Enhanced by AR And IoT Data" [1], discusses the potential of integrating various technologies like the Internet of Things (IoT), Augmented Reality (AR), Virtual Reality (VR), Artificial Intelligence (AI), and more, to enhance smart digital education. The paper emphasizes the importance of adapting education to new technology paradigms and creating an interconnected learning environment. The use of AR enhanced by IoT Data is explored, blurring the line between reality and computer-generated information. The paper also reviews the current state of AR in smart digital education, and identifies opportunities, technical limitations, and challenges in building AR applications enriched by IoT Data and AI features. It discusses the advantages, disadvantages, and challenges of employing augmented reality for education.

The research paper titled "Internet of Things-Based Intelligent Smart Home Control System" [2] presents a cloud-based intelligent home automation system using IoT technology. The system allows interactive control and monitoring of appliances and environmental factors through an Android app. It comprises two modules: one controls appliances and environmental factors, while the other focuses on home security by detecting motion and capturing images of potential intruders. Machine learning, specifically the between regular home occupants and intruders to avoid false alarms. The mobile app serves as the interface for users to interact with the system remotely. The prototype was implemented using ESP8266 and ESP32-CAM boards, a relay module, and sensors. This research demonstrates the potential of machine learning to enhance home automation functionality and security, contributing to the growing field of smart homes and IoT applications.

The research titled “Smart Home Interaction Using Augmented Reality with Internet of Things” [3] by LV Abarajithan explores the integration of Augmented Reality (AR) technology as a solution for interacting with Internet of Things (IoT) devices. By incorporating AR into smartphones, the interaction with IoT is enhanced, providing perceptually enriched user experiences. AR technology enhances the natural environments or situations and makes the interaction with physical reality more responsive. Rather than relying solely on Graphical User Interface (GUI) interactions like buttons or sliders, this research implements real-world touch or gestures on virtually placed objects to control IoT devices.

The research paper titled "Smart Learning Based on Augmented Reality with Android Platform and Its Applicability"[4] by Yogita Bahuguna, Aashish Verma, and Kunal Raj from the Department of Electronics and Communication at Tula’s Institute in Dehradun, India, explores the application of Augmented Reality (AR) technology on Android devices. The paper provides an overview of AR and its main concepts, highlighting its growth and use in various fields, including education and medicine. AR enhances the learning experience by enabling virtual content to be experienced in the real world. The paper proposes using AR on the widely used Android platform to improve teaching methods and make learning more engaging and effective for students.

III. PROPOSED SYSTEM

This study has explored the concept and implementation to create an advanced mobile application that seamlessly integrates augmented reality (AR) and Internet of Things (IoT) technologies to provide users with an affordable and accessible solution for home automation and interactive learning experiences. This innovative application leverages cutting-edge algorithms like object detection, Natural Language Processing (NLP), and Optical Character Recognition (OCR) to harness the power of a smartphone's camera. By combining these technologies, the app empowers users to control household appliances, monitor power consumption, and track electric bills using AR overlays. Additionally, it offers a unique educational platform that allows students to gain hands-on experience and assists medical practitioners in referencing organ anatomy efficiently. The system's cost-effectiveness ensures widespread accessibility, as it requires only simple installation and relies solely on a smartphone's camera without the need for costly additional hardware or dedicated VR headsets.

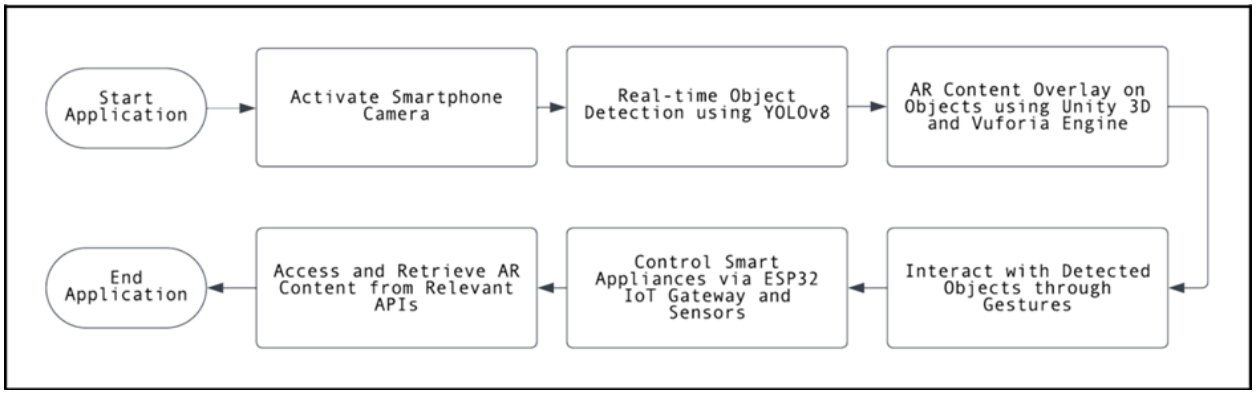

FIGURE 1. Concept Flow Diagram

Key Features and Functionalities:

1. Object Detection and AR Controls:
The heart of the proposed system lies in its sophisticated object detection capabilities. Using advanced AI algorithms, the app can identify various household appliances and objects through the smartphone's camera feed. Once identified, AR controls are seamlessly superimposed on these objects, enabling users to interact with and manage their appliances in an intuitive and user-friendly manner. For instance, users can turn on/off lights, adjust thermostats, or control other smart devices by simply tapping on the AR overlays.

2. Power Consumption and Electric Bill Tracking:
Through the integration of IoT technology, the app can gather real-time data on power consumption from connected appliances. Users can access detailed insights into their electricity usage, empowering them to make informed decisions and adopt energy-efficient practices. The application can also track electric bills, providing users with a clear understanding of their consumption patterns and aiding in cost-saving efforts.

3. Access to Rich AR Content:
The proposed system utilizes APIs to access the internet's vast repository of AR content. When users interact with detected objects, the application can retrieve relevant 3D generative models, 360-degree videos, 2D images, and textual information from the web. This feature enriches users' experiences, enabling them to explore and learn about objects and subjects in a highly interactive and immersive manner.

4. Intuitive Hand Gesture Interactions:
To further enhance user experience, the application employs intuitive hand gesture interactions. Users can effortlessly explore and interact with the retrieved AR content using simple hand movements, eliminating the need for complex touch or button controls. This hands-free approach not only makes the app more accessible but also adds to the overall immersion and engagement.

5. Educational Learning Tool:
The proposed system serves as an invaluable educational tool, particularly for students. By providing access to a wealth of AR-based learning materials, including 3D generative models and textual information, the app fosters interactive and experiential learning. Students can explore complex concepts, visualize abstract ideas, and gain practical insights into various subjects, making learning more enjoyable and effective.

6. Medical Reference Tool:
Medical practitioners can benefit significantly from the proposed system. With its feature-rich functionalities, the app allows quick referencing of organ anatomy in AR. Doctors and medical students can overlay detailed 3D generative models of organs onto real-world views, aiding in diagnosis, treatment planning, and medical education.

IV. WORKING MECHANISM

The AR-IoT Home Assistant and Learning Application operate through the seamless integration of various technologies, including Unity 3D, Vuforia Engine, YOLOv8 concept, ESP32 microcontroller, relays, current sensors, temperature sensors, and relevant APIs. This integrated system allows for accurate object detection, appliance control, real-time power consumption monitoring, educational content retrieval, and an immersive AR experience using intuitive hand gestures.

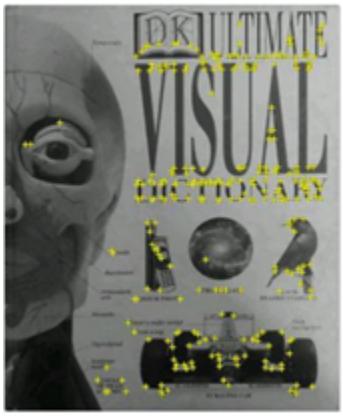

FIGURE 2. Object Detection and Recognition Points

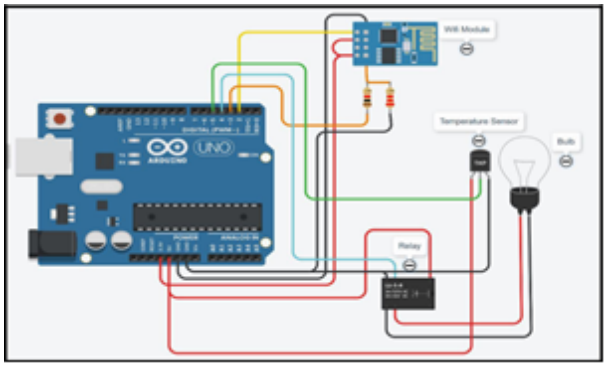

FIGURE 3. IoT Architecture diagram

FIGURE 4. AR controlled appliances

→ The application utilizes the YOLOv8 concept, a state-of-the-art deep learning algorithm, to perform real-time object detection through the smartphone's camera feed.
→ The Unity 3D engine, combined with Vuforia Engine, enables AR content overlays on the detected objects.
→ When a user points their smartphone's camera towards a household appliance, the YOLOv8 model identifies the object, and the Vuforia Engine anchors AR controls, such as power buttons and temperature settings, directly onto the appliance.
→ To achieve seamless home automation, the application connects to the ESP32 microcontroller, acting as the IoT gateway.
→ The ESP32 communicates with smart appliances using wireless protocols like Wi-Fi or Bluetooth. Additionally, current sensors are integrated into the electrical circuits of appliances to monitor real-time power consumption.
→ Temperature sensors are also incorporated into devices like thermostats or refrigerators to provide users with relevant data.
→ The application processes the real-time power consumption data received from the current sensors via the ESP32.
→ By analyzing this data over time, the app calculates and presents users with detailed insights into their electricity usage patterns.
→ Users can track their electric bills and identify areas for potential energy-saving measures, promoting energy efficiency and cost savings

→ The application integrates with relevant APIs to access a vast repository of 3D generative models, 360-degree videos, 2D images, and textual information.
→ When users interact with detected objects, the app retrieves AR content from these APIs, enriching the AR experience.
→ Using gesture recognition libraries within Unity 3D, the application interprets hand gestures captured by the smartphone's front-facing camera.
→ Users can select objects, navigate AR content, and control appliances through simple hand movements.
→ The application serves as a valuable educational tool for students.
→ When students interact with specific objects, the app retrieves relevant educational content from APIs, providing them with 3D generative models, textual explanations, and immersive learning experiences.
→ Medical practitioners can utilize the app's AR capabilities to reference organ anatomy by overlaying detailed 3D generative models of organs in real-world views, facilitating medical education and decision-making.
→ The AR-IoT Home Assistant and Learning Application's cost-effectiveness stems from its reliance solely on the smartphone's camera, without the need for expensive dedicated hardware.
→ This accessibility helps a broad audience benefit from the app's features without extra financial burdens.

Aspect	Marker Based AR using OpenCV& Python/ MATLAB	Unity AR with Vuforia Engine and IoT
AR Development Platform	Reference image with good feature feature points needed	Unity
AR Recognition Technology		Vuforia Engine
IoT Integration		Yes
Requirement for Reference Image	Script is needed to detect feature points and to render virtual objects on target images.	No (Depends on IoT data source)
Feature Point Detection Algorithm		Not applicable (IoT-driven)
Digital Content Augmentation		Yes
Asset Import and Management		Yes (Unity-based)
Ease of Use	Use of this methodology fails in augmenting complex virtual objects.	Unity's user-friendly interface
IoT Data Utilization		Real-time data display and control
Limitations		None specified

TABLE 1. Comparative Analysis between Marker based AR using OpenCV and Unity AR with Vuforia Engine and IoT.

V. RESULT AND DISCUSSIONS

The integration of AR and IoT technologies in the application results in a powerful and cost-effective solution for controlling household appliances. By employing object detection, Natural Language Processing (NLP), and Optical Character Recognition (OCR) algorithms, the app can efficiently scan the smartphone's camera feed to identify various objects in the environment. This ability to recognize household appliances allows users to access AR controls for these devices. As a result, users can remotely operate and monitor their appliances, promoting convenience and energy efficiency. The app's feature to display power consumption and electric bill tracking for appliances is a valuable addition. By providing users with real-time information on energy usage, the application empowers them to make informed decisions about energy consumption, leading to potential cost savings and a reduced carbon footprint.

Beyond home automation, the app offers exciting educational benefits. Its ability to retrieve relevant 3D generative models, 360-degree videos, 2D images, and textual information based on user interactions with detected objects makes it a unique learning tool. Students can benefit from hands-on experiences, enabling them to explore and interact with 3D generative models using intuitive hand gestures. This immersive learning experience enhances understanding and retention of complex concepts, making education more engaging and effective. Medical practitioners can also leverage the app's educational functionalities to quickly reference organ anatomy. The availability of feature-rich functionalities provides healthcare professionals with a convenient tool to enhance their knowledge and improve patient care

One of the app's most significant strengths is its accessibility and cost-effectiveness. Unlike dedicated AR or VR headsets which can be expensive, the application requires only installation on a smartphone. Users can enjoy the benefits of AR without investing in additional hardware, making it accessible to a broader audience. Moreover, the app's efficiency in detecting objects in the camera feed and prompting user interactions through intuitive hand gestures contributes to its user-friendliness. This ease of use ensures that users can quickly and effortlessly access the desired information and functionalities.

	Acc Sensitivity		Rot Rate		Stab Index	
	Mean	SD	Mean	SD	Mean	SD
Accelerometer	3.42	1.23	3.65	1.34	2.50	2.02
Gyroscope	2.88	0.85	1.34	1.90	2.11	2.50

TABLE 2. Mean and Standard Deviation of Accelerometer and Gyroscope.

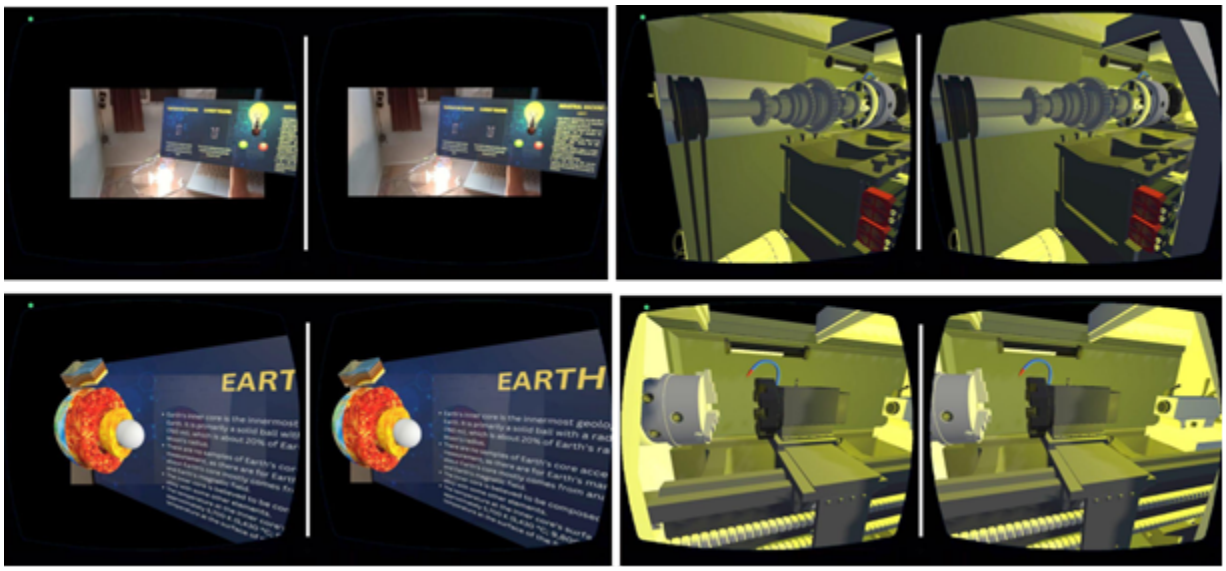

FIGURE 5. Usage of App in Education and Automation

VI. FUTURE ENHANCEMENT

1. Enhanced Object Recognition and Tracking:
To further improve the user experience, future enhancements should focus on refining the object detection and tracking capabilities of the application. Advancements in computer vision and machine learning algorithms could lead to faster and more accurate recognition of various household appliances and objects in the environment. This improvement would result in smoother interactions and reduce the chances of misidentification.

2. Expanded Educational Content:
To bolster its educational applications, the app could collaborate with educational institutions, content creators, and experts to expand its database of 3D generative models, videos, and textual information. Offering a wide range of educational content across diverse subjects, from science and history to engineering and art, would make the app a go-to resource for learners of all ages.

3. Social Interaction and Sharing:
Implementing social features within the application could foster a community of users who share their AR experiences and knowledge. Users could interact with each other, exchange educational content, and even collaborate on educational projects. This social aspect would add a new dimension to the app, promoting learning through peer-to-peer engagement.

4. Virtual Collaboration and Remote Assistance:
Enhancing the app's capabilities for virtual collaboration and remote assistance would benefit industries and professionals. For example, remote experts could guide technicians or students through complex tasks using AR annotations and live video feeds, improving learning outcomes and problem-solving efficiency.

VII. CONCLUSION

The advanced mobile application discussed in this paper combines augmented reality (AR) and Internet of Things (IoT) technologies, offering potential benefits in home automation and education. It integrates AR controls for household appliances, power consumption monitoring, and electric bill tracking, providing an affordable solution for improved energy efficiency and convenience at home. Additionally, it utilizes object detection, Natural Language Processing (NLP), and Optical Character Recognition (OCR) algorithms to enable users to explore 3D models, videos, images, and textual information for immersive learning.

This application's versatility extends beyond the home, with applications in healthcare, education, and product visualization. It's cost-effective, user-friendly, and accessible, requiring only a smartphone camera. Moreover, the inclusion of AI chatbots, smart home devices, and social features offers potential for future improvements, enhancing functionality and connectivity.

As AR and IoT technologies advance, this application stands at the forefront of innovation, empowering individuals and professionals. Ongoing research and development, along with expanded educational content and improved object recognition, promise to reshape how we interact with our surroundings and acquire knowledge. Embracing this technology allows users to access augmented experiences, enriched learning, and enhanced control through their smartphones.

VIII. REFERENCES

1. M. Cabanillas-Carbonell, A. Canchaya-Ramos, and R. Gómez-Osorio, "Mobile application with augmented reality as a tool to reinforce learning in pre-Inca cultures," 2020 IEEE Engineering International Research Conference (EIRCON), Lima, Peru, 2020, pp. 1-4, doi: 10.1109/EIRCON51178.2020.9254018.

2. M. Hu, D. Weng, F. Chen, and Y. Wang, "Object Detecting Augmented Reality System," 2020 IEEE 20th International Conference on Communication Technology (ICCT), Nanning, China, 2020, pp. 1432-1438, doi: 10.1109/ICCT50939.2020.9295761.

3. R. Osorio-Comparan, F. G. Osorio, H. Kaschel, C. Ahumada, S. Cordero and G. Lefranc, "Virtual Monitoring of the 3D Movement of a Mobile Object, using UNITY," 2022 IEEE International Conference on Automation/XXV Congress of the Chilean Association of Automatic Control (ICA-ACCA), Curicó, Chile, 2022, pp. 1-6, doi: 10.1109/ICA-ACCA56767.2022.10006241.

4. C. Montellanos, J. Luis, M. Vásquez, C. Alberto, H. Salazar, and J. Luis, "Augmented reality mobile application and its influence in Quechua language learning," 2019 IEEE Sciences and Humanities International Research Conference (SHIRCON), Lima, Peru, 2019, pp. 1-4, doi: 10.1109/SHIRCON48091.2019.9024860.

5. J. Kim, M. Lorenz, S. Knopp, and P. Klimant, "Industrial Augmented Reality: Concepts and User Interface Designs for Augmented Reality Maintenance Worker Support Systems," 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Recife,

Brazil, 2020, pp. 67-69, doi: 10.1109/ISMAR-Adjunct51615.2020.00032.

6. J. Motejlek and E. Alpay, "Taxonomy of Virtual and Augmented Reality Applications in Education," in IEEE Transactions on Learning Technologies, vol. 14, no. 3, pp. 415-429, 1 June 2021, doi: 10.1109/TLT.2021.3092964.

7. J. P. Cruzado, N. De La Cruz, J. L. H. Salazar, and J. Tarazona, "ARAnimals: Mobile Application with Augmented Reality for the Learning of Vertebrate Animals," 2020 IEEE Congreso Bienal de Argentina (ARGENCON), Resistencia, Argentina, 2020, pp. 1-4, doi: 10.1109/ARGENCON49523.2020.9505519.

8. S. Imran Hussain, S. Deepalakshmi, R. J. Benilla, and V. Charu Nivetha, "Automation of Smart Home using Smart Phone via Google Assistant," 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2023, pp. 499-505, doi: 10.1109/ICSSIT55814.2023.10060979.