XenoLingo

An AR-based application for Indigenous Language and Culture Preservation

Aavash Chhetri[1*], Ashim Bahadur Karki[2*], Aayusha Odari[3*], and Mamata Maharjan[4*]

[1,2,3,4]Department of Electronics and Computer Engineering, Institute of Engineering, Pulchowk Campus, Nepal

[1,2,3,4](077bct004.aavash, 077bct006.aayusha, 077bct013.ashim, 077bct043.mamata)@pcampus.edu.np

XenoLingo is an innovative project with a dual mission: preserving indigenous languages and fostering language learning. We achieve this through the integration of the Vuforia Engine and Unity Game Engine, along with a Text-to-Speech (TTS) system based on Tacotron2 for language articulation. When users point their device's camera at an object, XenoLingo identifies and presents the object's name in their native language scripts and speech. This interactive approach makes language acquisition a dynamic and enjoyable experience.

1 Introduction

In an increasingly interconnected world, the preservation and promotion of indigenous languages have become vital. XenoLingo, an innovative Augmented Reality (AR) application, is dedicated to supporting the preservation and learning of indigenous languages. Our application is built on the belief that observation is the most effective way to learn a language. By harnessing the device's camera, XenoLingo empowers users to recognize real-world objects in real-time. When users point their devices at objects, XenoLingo identifies them and provides labels in both the user's native language script and English, facilitating better comprehension. Additionally, it offers pronunciation through speech, ensuring that language acquisition seamlessly integrates into users' daily experiences.

2 Methodology

2.1 Object Detection

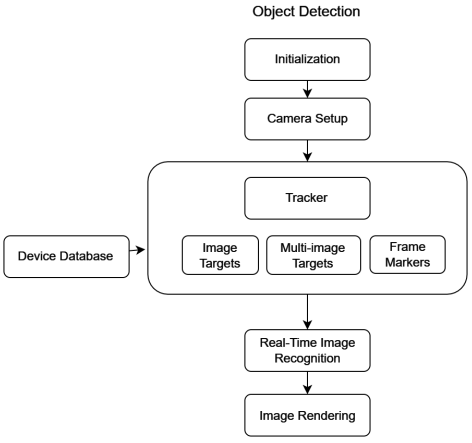

Figure 1: Block diagram for Object Detection

XenoLingo employs the Vuforia Engine, a robust Augmented Reality (AR) framework, as a fundamental component of its methodology. This engine identifies and tracks real-world objects by analyzing the incoming frames, comparing them with a predefined database of target images, and accurately determining the position and orientation of these objects within the real-world environment.To facilitate precise and reliable object recognition, XenoLingo maintains a dedicated device database. The Unity Game Engine is employed to integrate this AR functionality into the application. The Unity

engine acts as the bridge that enables XenoLingo to overlay digital content onto the live camera feed in real-time. This rendering process is executed with precision, ensuring that the digital objects seamlessly align with the real-world objects detected by the Vuforia Engine.

2.2 Text-to-Speech

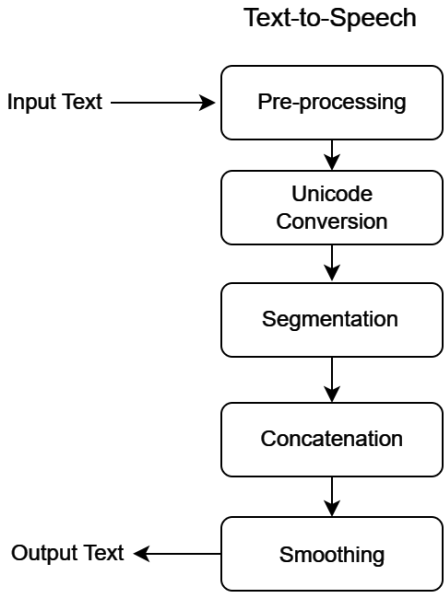

Figure 2: Text to Speech Formulation

In the development process, we utilized TensorFlow and Flask for handling the aspects of the Text-to-Speech functionality in XenoLingo. The textual data is subjected to preprocessing to ready it for synthesis. The text is transformed into Unicode for character representation, with segmentation, concatenation, and smoothing applied as needed to ensure natural-sounding speech. Xenolingo is supported by a Tacotron2-based Text-To-Speech Synthesis Model for Speech Synthesis.
Additionally, audio resources were obtained from the existing project "Nepali Text-to-Speech Synthesis using Tacotron2 for Melspectrogram Generation".

3 Features

- Immersive Learning: XenoLingo creates an immersive language learning experience by integrating AR technology, allowing users to interact with real-world objects.
- Multilingual Object Labeling: Objects are labeled in both the user's native language script and English for enhanced understanding.
- Seamless Digital Overlay: Digital content overlays real-world objects seamlessly, enhancing the user experience.
- Pronunciation Assistance: It includes a Text-to-Speech (TTS) system, enabling users to hear accurate pronunciation of object names in their native language.
- Educational Tool: It can be used in schools, cultural centers, and language preservation initiatives for educational purposes.
- Offline Mode: XenoLingo works seamlessly even without an internet connection, ensuring accessibility in various environments.

4 Conclusion

XenoLingo is an application that leverages Augmented Reality (AR) technology to combat the critical challenges of indigenous language erosion and cultural loss in Nepal. Indigenous communities and speakers of endangered languages stand to gain the most from the app. XenoLingo serves as a potent educational instrument, allowing language educators and teachers to transform language learning into an engaging and interactive experience. We can further expand our reach by collaborating with cultural museums, providing virtual tours of these locations, and describing exhibits in the native tongue. Thus, through XenoLingo, we aim to bridge linguistic gaps, preserve cultural heritage, and celebrate linguistic diversity. Our ultimate vision is to ensure that indigenous languages not only survive but flourish.

References

[1] S. Khadka, R. G.C., P. Paudel, and R. Shah. Nepali text-to-speech synthesis using tacotron2 for melspectrogram generation. Available: https://gitlab.com/shrutiaudio/shrutiaudio.

[2] T. L. Manandhar, *Modern Language Of Kathmandu Valley*, D. A. Vergati, Ed. Agam Kala Prakashan, 1986.

[3] M. B. Shakya, *Prachalit, Ranjana and Bhujinmol Scripts*.

[1] [2] [3]