SER 594: HCI	**Phase 2**
Title:	
Part 1 Reviving Fantasy into Reality	**Part 2** Projecting narration with Virtual Reality
Team Name	
EEE (Explore Enhance Execute)	
Team Members	
A, 1225248336 VISHWANATH REDDY YASA **B**, 1226212496 SAISH VEMULAPALLI **C**, 1226092805 PAVITHRA MORAVANENI	
Contribution:	
Conceptualization: A, B, C Methodology: A, B Software/ Prototype: A, B, C Validation: B, C Formal Analysis: A, B, C Investigation: A, C Resources: A Writing Report: A, B, C	
Abstract/Overview:	
This research paper investigates the use of AI technology and deep learning algorithms in visualizing real-time video projection for narrating stories. The data is collected with the proposal	

design on one hand and on the other hand, having the video projection which is pre-recorded. Two design approaches were compared through user experience testing to identify the most effective method in creating a realistic and engaging experience. The study provides a technical explanation and real-world examples of the potential of AI in enhancing digital storytelling. The proposed design has the potential to revolutionize the way we interact with digital content, creating new possibilities for immersive user experiences.

Overview:

Proposed Design:

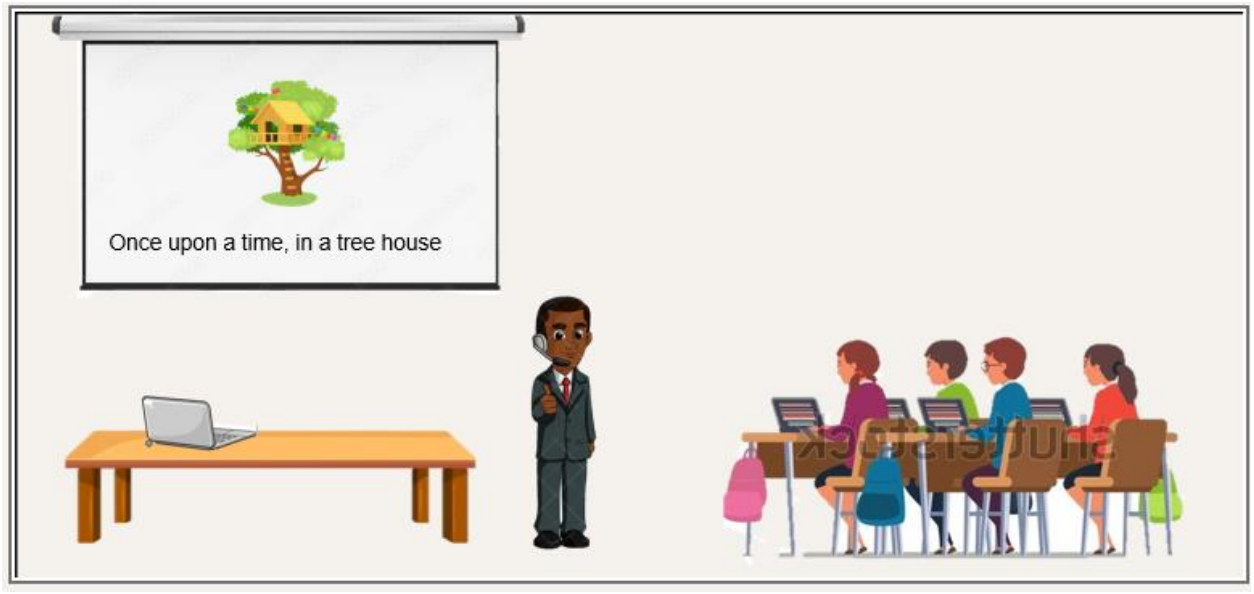

Fig 1(a): Professor narrating story to the students.

The proposed design is a cutting-edge approach to digital storytelling that utilizes voice commands and AI technology to create an immersive user experience. The Voice User Interface allows users to interact with the interface seamlessly, enhancing the overall user experience. The design features a screen, microphone, and projector that work together to create a dynamic and engaging display.

With the help of Speech Conversion API, the system converts the user's voice commands into text, which is then presented as captions on the screen. The DALL-E API is then used to transform the text into corresponding cartoon images, which appear on the screen. To further enhance the user

experience, the system projects images of the exact scenes mentioned in the narration, giving the user a realistic and immersive experience.

For example, as depicted in the picture above, if a user narrates a story like "Once upon a time in a tree house." To give the user a realistic experience, the exact house will be projected onto the screen.

Enhanced Design:

Fig 1(b): User can narrate a story using Mobile Application.

The pursuit of smaller, faster, and more effective technology has been a driving force in the evolution of digital interfaces. This study explores an alternate design for the proposed topic, which involves a mobile or web interface.

As shown in the above image, the user can simply turn on the microphone and speak, and the Speech Conversion API will convert their speech into text. The system will then generate corresponding animations based on the text and display them on the screen.

This alternate design has several advantages, including portability and accessibility. Users can access the interface from anywhere, at any time, and the design is optimized for use on mobile devices.

By examining the potential of this alternate design and comparing it to the original proposed design, this study seeks to provide valuable insights into the future of digital storytelling and interactive interfaces.

Interfaces Design:

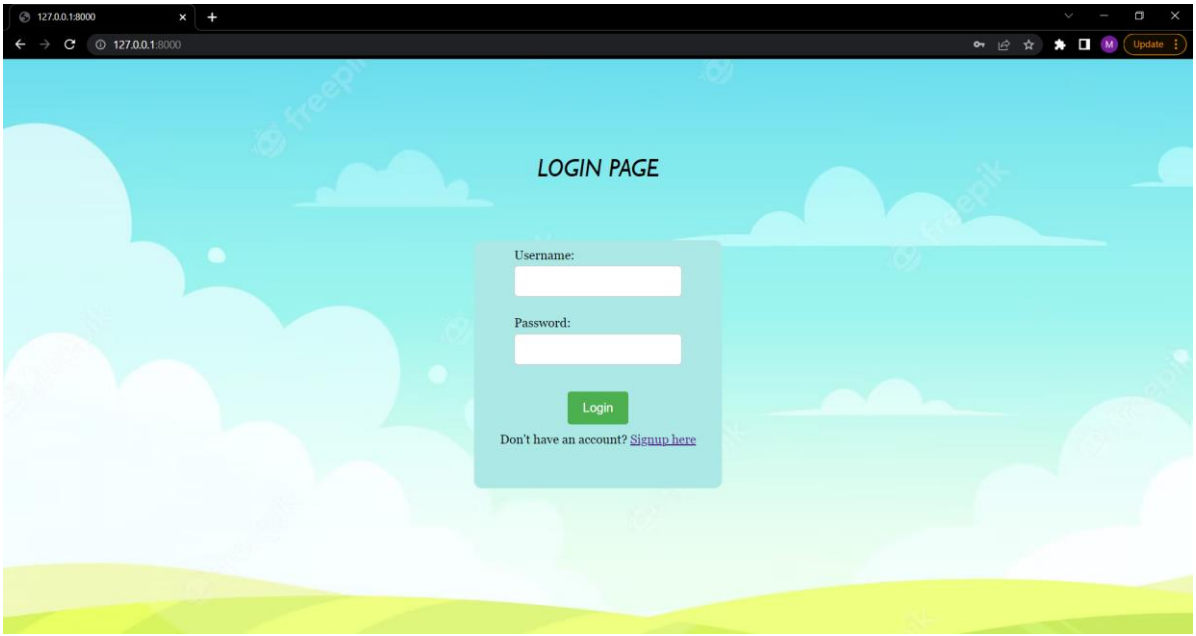

Fig 2(a): Login Page

Using the credentials, the user will be logged in if they have an account.

User will be redirected to the home page where he can invoke the microphone using the start button as shown in below Figure 2(b).

Once the user starts speaking after the microphone is enabled, the text will be displayed in the first text box and parallelly the image will be displayed as shown in below Figure 2(c).

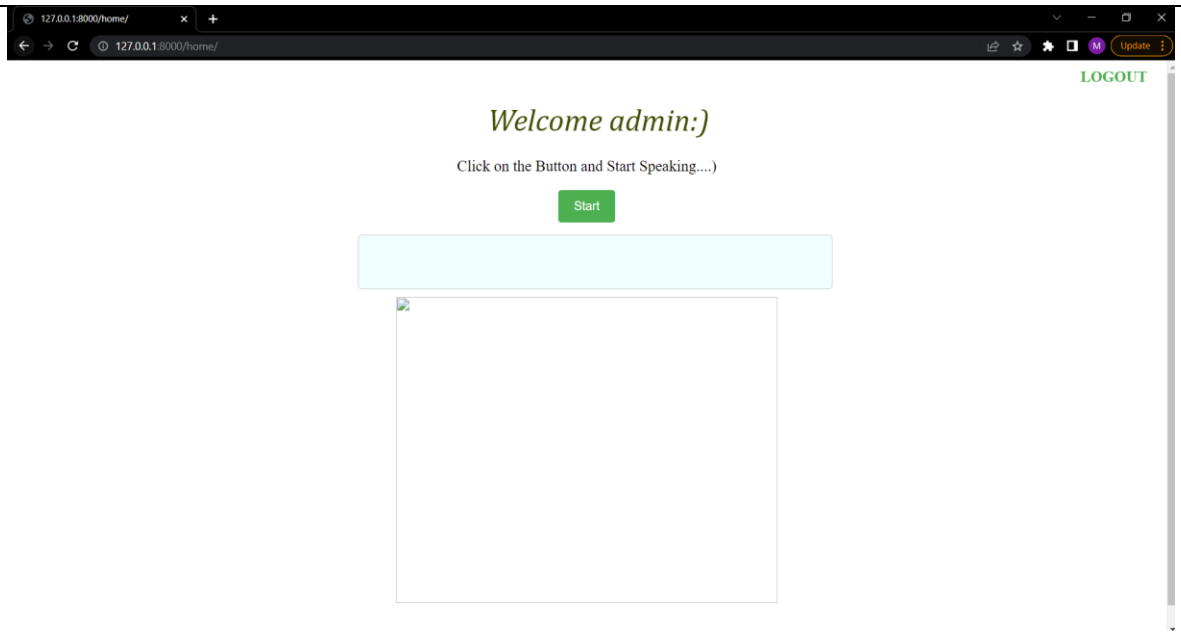

Fig 2(b): Home Page

For example, User gave a sentence as "*Once upon a time when Tigers used to smoke*", the image is displayed as shown below with Tiger in it.

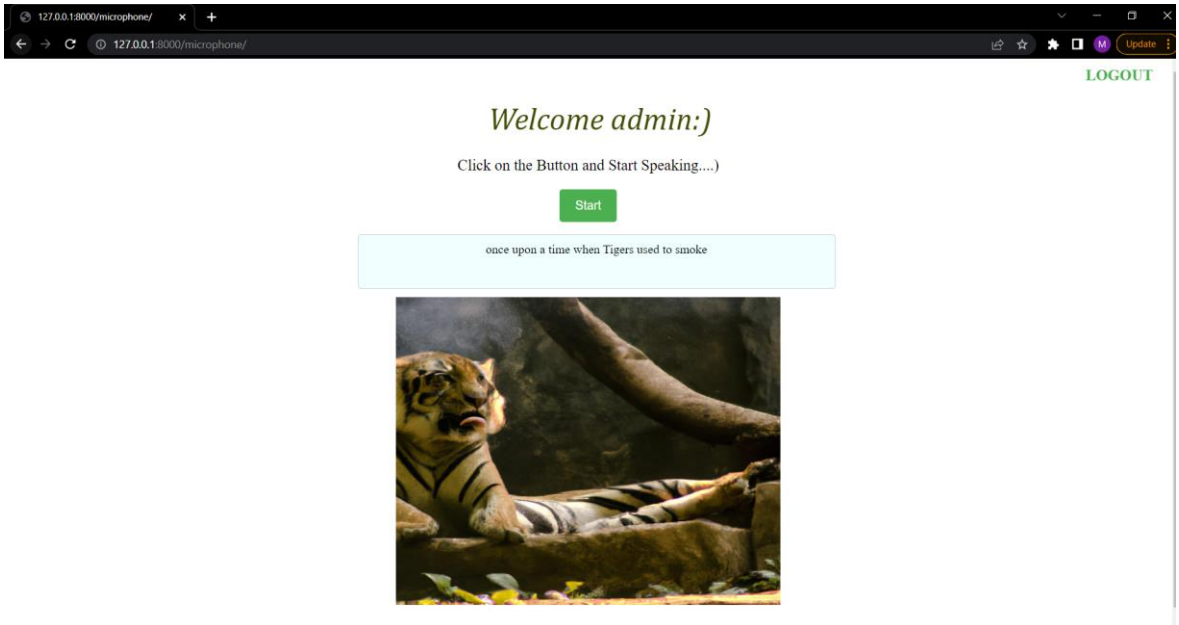

Fig 2(c): Output for given voice command by the user.

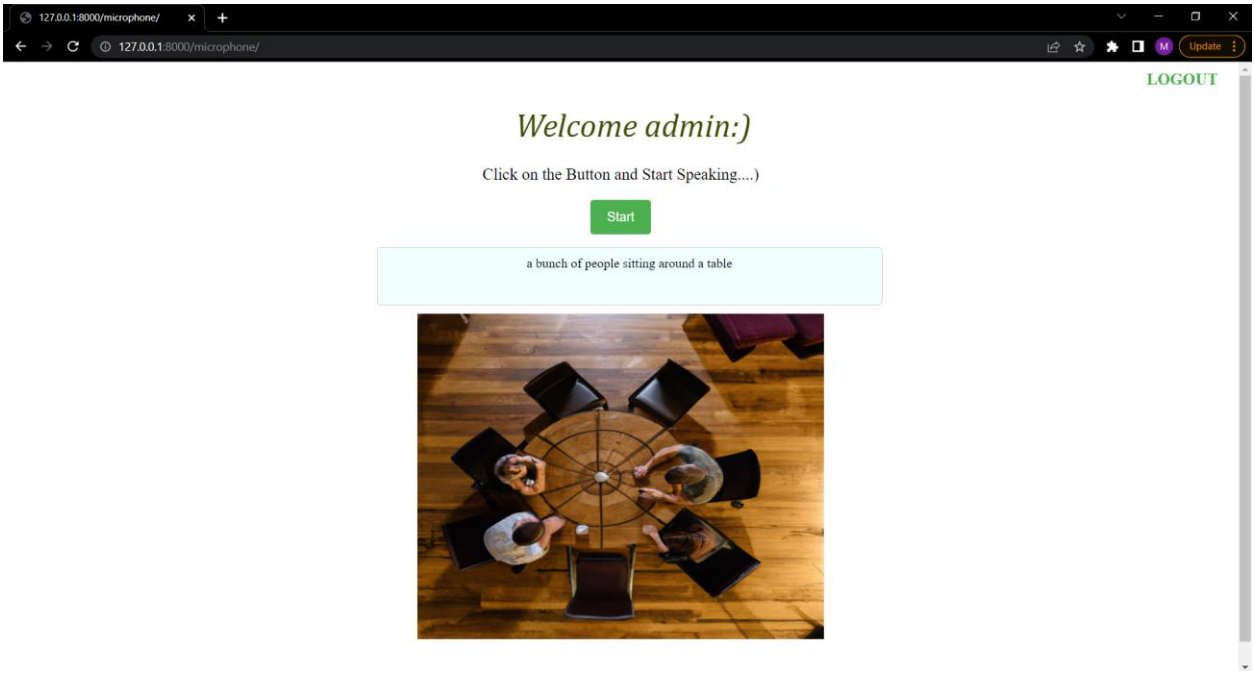

Fig 2(d): Output for given voice command by the user.

Second example is "*A bunch of people sitting around a table*" and the image is displayed.

Implementation Details

Proposed Design:

This design consists of Microphone to speak for narrating a story and Projector to display the cartoons or images related to the story.

This is divided into two stages:

- The first stage where the microphone takes the speech as input from the user and then it converts it into text which is displayed on the projector screen. To achieve this, numerous speech recognition algorithms with accurate results are available in machine learning and AI technologies.
- Next stage is to convert the received text into cartoons which are pre-trained and are invoked when the text is called.

For example, as specified above in Proposed Design figure (a), When the professor speak using the headset microphone say as "Once upon a time, there is a tree house", where the same is displayed on the projector screen which gives the students an unimaginable visual experience.

Enhanced Design:

Another design is using Mobile or Web Interface, where user can directly speak by enabling the Microphone and the video of the cartoons displayed on the mobile screen itself. This design can also be further enhanced by using text input from the users which is then converted to images that are to be displayed on the screen or interface.

Prototype Design:

The prototype design for our interactive web interface allows users to interact with voice commands, making it easy and convenient for anyone to use. By enabling the microphone, users can speak to the system and have their speech converted into text using the latest Speech Conversion API. The text is then processed through the cutting-edge DALL-E API, which uses artificial intelligence and machine learning to generate stunning, realistic images that are directly related to the text input.

The resulting text and images are then projected onto the interface, creating an immersive and visually captivating experience for the user. By combining the power of speech recognition and deep learning algorithms, this prototype design offers a unique and innovative way to interact with digital content, opening up new possibilities for storytelling and engagement.

This type of interface has exciting applications across a range of industries, from education and entertainment to marketing and advertising. By making it easy for users to interact with digital content using voice commands and advanced AI, we can create more engaging and personalized experiences that capture and hold the attention of users.

System Evaluation

After the user experience study and the analysis of the application performance, functionality, below are few observations:

- Users were very excited to use this application as many of them felt having the kind of visual understanding of any learning would make them remember for a long time.
- Many research has shown that visualizing any image or video can enhance learning and withholding in memory can be easier to process.

- System was slow in displaying the images and text and this can be enhanced for getting a quick response.
- Many users have recommended enhancing this application by projecting video response rather than images.
- Voice commands given by the user were displaying accurate text information and the users were satisfied.

To conclude, users were excited to have these applications as the world is getting digitalized and making it easier to get things done in less time. Moreover, this was recommended by many of them to be developed in future.