SSN COLLEGE OF ENGINEERING

KALAVAKKAM-603110

INTERNALLY FUNDED STUDENT PROJECT - 2022

Project Sanjaya - Live Visual Assistant in Physics lab

R Abhinav (205001001)
Aravind R (205001013).
Bhuvnesh Magotra (205001021)
III year, Department of CSE

Dr. B. Bharathi, Associate Professor
Dr. T. T. Mirnalinee, Professor & Head
Dr. J. Bhuvana, Associate Professor
Department of Computer Science and Engineering

Budget (in thousands): Rs. 26,450

Duration: 12 Months

Signature of the Project Students **Signature of the Project Guide(s)**

Signature of the HOD

1. Project Title: Project Sanjaya - Live Visual Assistant in Physics lab

2. Broad Subject: Computer Vision; Image Recognition

3. Project Duration (in months): 12 Months

4. Budget (in thousands): Rs. 26,450

5. Project Abstract

The braille based systems are very much helpful in conveying all the theoretical knowledge. So we found many living legends with visual impairment in fields of arts, specifically music, and other fields such as sports, national services and educational sectors in India but not in science based professional sectors. From this one can infer that the braille based system is not contributing much in advancing the visually impaired in technology oriented scientific fields. To promote the visually impaired people in science, it is necessary to inculcate scientific practical knowledge during their school education for them to take up science related courses in their higher education. People with vision disability may need a human assistant to carry out scientific lab experiments. Allowing a person to help each of the visually impaired students in the lab is practically not workable, and it is tedious. Because of these challenges, it does not encourage students to take up subjects with practical courses, namely science major subjects. This motivated us to develop a device that bridges the gaps that enable students to take up the science courses in higher secondary education.

The basic modules involved in this project are image acquisition, Object detection and recognition, image captioning, text to speech system. For image acquisition, we fixed a camera with a processor. For proposed system development, we will capture the image fromthe physics lab. We will use these images for training the object detection and recognition modules. There are many pre-trained models available in the literature to perform the object detection task, but not on the equipment used in the laboratories. So, the pre-trained models are fine-tuned by using the images gained in the labs. Even though successful techniques are available for scene understanding or image captioning for some languages, a full-fledged system is yet to be developed for Vernacular Languages. Once the image captioning is available in the Tamil text, the Text to Speech module will convert the text into audio output. Presently for the proposed work, we collect speech data from a new speaker and build a text-to-speech system. We can use the existing text-to-speech available for other languages, that may not be comfortable for the visually impaired students, because the pronunciation / accent of Tamil utterances by the speaker trained on English text may not be correct.

6. Keywords :

Live video processing, Computer vision, Text to Speech Translation, Object detection

7. Objectives

I. Real-Time video interpretation

Project Sanjaya aims to understand the scenario of the lab captured by the camera component using computer vision and image recognition and tries to give meaningful information regarding surrounding equipment such as measurement tools, electrical devices, optical devices and various other apparatus.

II. Dual-Language Translation

Project Sanjaya aims to be a bilingual operation with support for both English and Vernacular Language Chosen.

8. Introduction

The purpose of science education, based on school institutions, is to prepare individuals to develop fundamental levels of scientific understanding and basic applied science. Practical education is essential to get hands-on experience on various developed classical and modern scientific methodologies. Thus, practical work is an important aspect of scientific research and education, and good quality practical work improves students' engagement and interest, as well as help to develop basic skills, scientific knowledge and conceptual understanding. Doing practical science, however, requires intensive use of the senses, especially the eyes to be an excellent observer. This is a problem for particular learners as they might have difficulty using their eyes because of visual impairment. This project facilitates the learners to make use of latest computer vision and image recognition technology to help identify the various aspects of their practical science environment, i.e., a Physics Lab in particular, such as measurement tools, electrical devices, optical devices and various other apparatus, and guide them in order to make practical science learning a much pleasurable experience

9. Definition of the Problem

The problems the visually impaired students are facing in taking up the science related courses at school and the problem of doing scientific lab experiments, serve as a motivation to develop an

intelligence-enabled system for doing the same. This proposal aims to develop an lab assistive tool that helps the visually impaired people to carry out the scientific physics lab experiments.

10. **Review of status of Research and Development in the subject**

10.1 National Status

The authors of [1], have developed a real-time web application, which uses image processing to capture day to day objects and provide a textual output. They have used Gttx (A google API) for their TTS system.The idea implemented by [2] is through an Android mobile app that focuses on voice assistant, image recognition, currency recognition, e-book, chat bot etc. The app is capable of using voice commands to recognize objects in the surrounding, do text analysis to recognize the text in the hard copy document. It will be an efficient way in which blind people can also interact with the environment with the help of technology and utilise the facilities of the technology.

10.2 International Status

The authors of [3], have used object recognition techniques to assist visually impaired people in Chemistry laboratories which improves the accessibility by providing non-visual cues. The authors of [4] have designed DAVID, a digital assistant application aimed to help the visually impaired in recognizing text on real-world objects and provide audio feedback in real-time. It utilises voice user interface technology such as speech recognizing and speech synthesis as the means of interaction through voice input.

11. Novelty / Importance of the proposed project in the context of current status

With the advancements in technologies, several assistive devices for mobility and hearing have been developed for the differently abled. Particularly for the visually impaired people, assistive devices such as smart cane, smart readers, dictation software, smart glasses, OCR systems were developed and made available. All these smart assistive devices have been evolved to assist them to lead a quality life.

To promote the visually impaired people in the field of science, it is necessary to inculcate scientific practical knowledge during their school education for them to take up science related courses in their higher education. In general, people with vision disability may need a human assistant to carry out scientific lab experiments. Allowing a person to help each of the visually impaired students in the lab is practically not feasible and tedious. Due to these challenges students are not encouraged to take up subjects with practical courses namely science major subjects. This motivated us to develop a device that bridges the gaps that enable students to take

up the science courses in higher secondary education. The challenges associated with these devices are many, most primarily their affordability, potability, ease of use, etc.

12. Patent details (domestic and international), if applicable

13. Work plan and Detailed technical information

13.1 Methodology

The steps involved in the proposal is given in the architecture Fig. 1

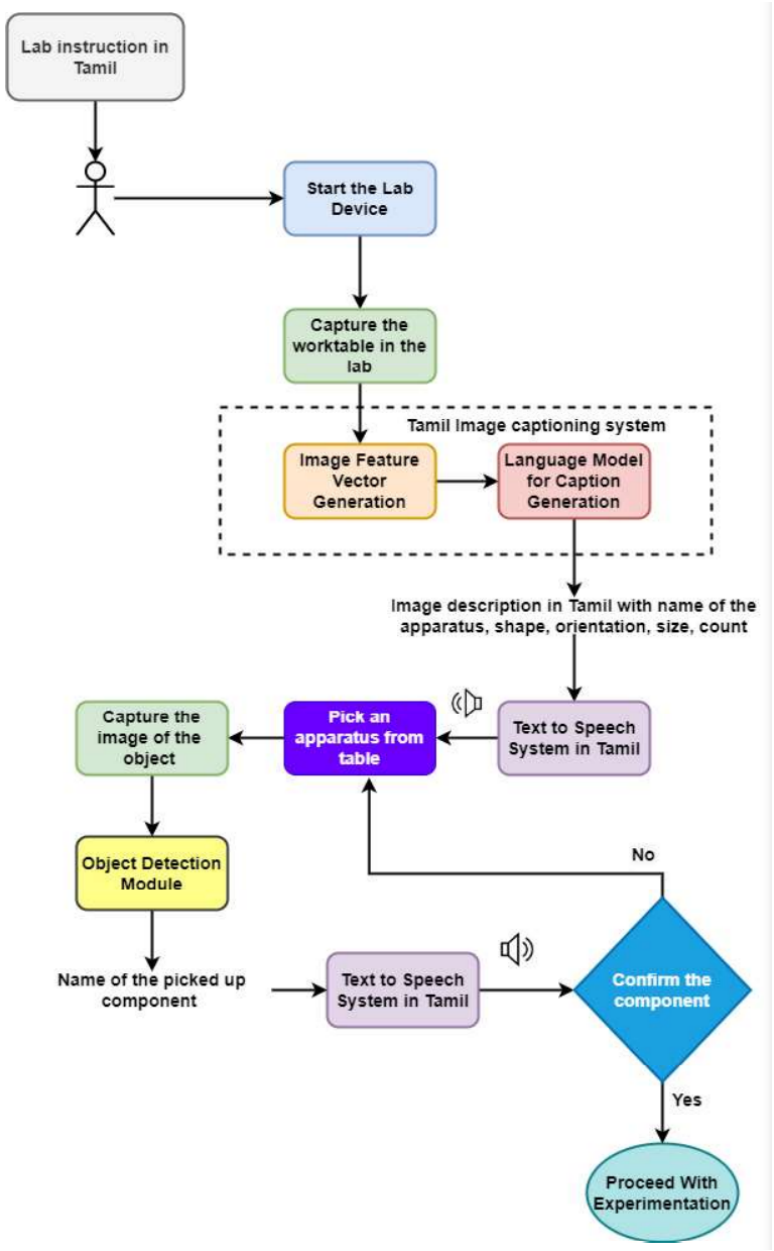

Fig. 1 Architecture of the proposed system

The proposed AI- based smart lab assistive device requires the image, text and speech data in Tamil for the development of the device. The following are the data requirements:-

1. **Image data collection**: In the data collection phase, we capture the images of the laboratories that are necessary for training the image captioning network.
2. **Text data collection**: Text data describing the images captured from the science laboratories of the higher secondary school will be collected. The text data will have the name of the equipment available in the lab, their sizes, their colour, numbers and so on. This text data will be used as annotation for building the image captioning network and also training the text-to-speech model.
3. **Speech data collection**: Speech data is collected from sample human voice to provide a basis for the text-to-speech system.

With requirements collected, we identified the three architectures using the deep learning algorithms as follows:

1. **Object detection and recognition**: The scene understanding of the captured images is done using an deep learning based object detection network, and the name of the recognized object will be sent to the text to speech module.
2. **Image captioning**: The text data along with the image features are used to generate the captioning using a Long Short Term Memory based network.
3. **Text to speech**: The text generated from the captioning model will be split into sentences. Each sentence will be recorded separately to maintain the quality. An unrestricted-domain text-to-speech synthesis system has to be developed.

The development of the device will be carried out in two stages. At the first level, the software modules required for the proposed device will be implemented using the following software tools.

- Operating system: Linux
- Software tools: Pytorch, OpenCV, WavRNN, Yolo
- Programming languages: Python
- Techniques and methods: YOLOV5, EfficientNet-B7, LSTM

At the second level, the trained models will be loaded as mobile applications. The hardware components are camera and speaker, where the camera will be connected with a spectacle.

14. Time schedule of activities giving milestones

14.1 Time Schedule of Activities through BAR Diagram

1. Procurement of equipment - 1 Month
2. Data collection - image, speech, text - 2 Months
3. Design and implementation of Object detection and recognition -2 Months
4. Design and implementation of Image captioning- 2 Months
5. Design and implementation of Text to speech module -2 Months
6. Integration of modules and testing - 2 Months
7. Documentation - 1 Month

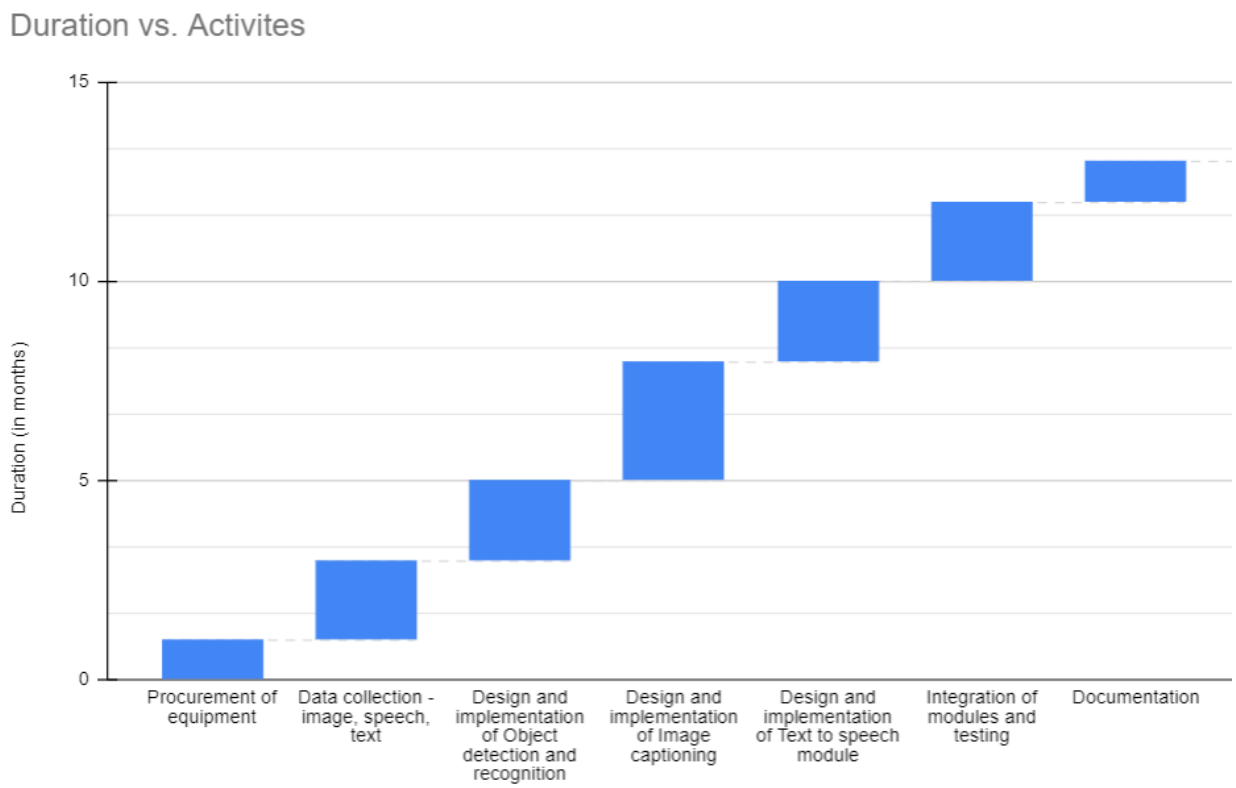

Fig. 2: Timeline chart

15. Deliverables

- Image corpus with captions in Vernacular Language
- Image Captioning System (ICS)
- Text - to - speech system (TTS) for the lab device for visually impaired
- Working Manual with description of working and other essential details

16. Target beneficiaries of the proposed work

- Visually impaired school students

17. Suggested plan of action for utilisation of research outcome expected from the project

17.1 As journal publication

We aim to publish our proposed work as a journal publication

17.2 Patent filing

New discoveries will be patented

17.3 Project preparation for submission to external funding

We will submit the research work for external funding

18. References

1. Sharma, Vipul and Singh, Vishal Mahendra and Thanneeru, Sharan, Virtual Assistant for Visually Impaired (April 19, 2020). Available at SSRN: https://ssrn.com/abstract=3580035 or http://dx.doi.org/10.2139/ssrn.3580035
2. S. M. Felix, S. Kumar and A. Veeramuthu, "A Smart Personal AI Assistant for Visually Impaired People," 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI), 2018, pp. 1245-1250, doi: 10.1109/ICOEI.2018.8553750.
3. The value of safety and practicality: Recommendations for training disabled students in the sciences with a focus on blind and visually impaired students in chemistry laboratories Gabriella M. Nepomuceno, Debbie M. Decker, Julian D. Shaw, Lee Boyes, Dean J. Tantillo, and Henry B. Wedler, “Journal of Chemical Health & Safety 2016 *23* (1), 5-11, DOI: 10.1016/j.jchas.2015.02.003
4. E. Marvin, "Digital Assistant for the Visually Impaired," 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), 2020, pp. 723-728, doi: 10.1109/ICAIIC48513.2020.9065191.

19. List of facilities and Equipments available with Department for the project

20. Budget Estimates

Sl. No.	Product Name	Foreign or Indigenous	Quantity	Price (in Rs.)
1.	**Spectacle**	**I**	1	5000
2.	**Camera**	**I**	1	6200
3.	**Speaker**	**I**	1	2750
4.	**Integration cost**	**-**	-	12500
Total				26,450

21. Budget Justification

S. No.	Product Name	Justification
1	**Spectacle**	Used as a mount for the camera
2	**Camera**	Used to capture the images of equipment in the lab
3	**Speaker**	Used to send audio cues to the user based on the scenario
4	**Integration cost**	Assembling all the components and feeding the model into the mobile app