International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 26 August 2023 **Revised:** 20 October 2023 **Accepted:** 02 November 2023*

Virtual Eye – Revolutionizing Vision Assistance For People With Disabilities

[1]Dr. Venkatesh . S, [2]Mrs.Jeevitha.D, [3]Mrs.G.Keerthana
[1]Assistant Professor, Department of Data Science and Business Systems,
School of Computing, SRM Institute of Science and Technology, Kattankulathur Chennai, India
venkates9@srmist.edu.in
[2]Assistant Professor, Department of Computer Science and Engineering,
Jeppiaar Engineering College, Chennai, India
jeevithadamotharanjpr@gmail.com
[3]Assistant Professor, Department of Computer Science and Engineering,
KCG College of Technology, Chennai, India
keerthanagsv@gmail.com

Abstract

Visually challenged individuals have faced numerous challenges in their daily lives. These challenges include: Visually challenged individuals have difficulty reading printed materials, including books, magazines, and newspapers. This limitation can significantly impact their education, as they may not have access to all the materials they need to learn. Moving around in unfamiliar places can be a daunting task for the visually impaired. They may struggle to access digital or printed materials, as these are often not available in accessible formats. It might be difficult for those who are blind to identify objects. This can be frustrating, especially in situations where they are alone and need to identify objects.

To address this issue, we are developing a mobile application for visually challenged individuals by providing a range of features such as text-to-speech, speech-to-text, image-to-audio, and PDF-to-audio. It enables visually challenged individuals to access information, read books, identify objects, communicate, and navigate with ease and independence. The app's user-friendly interface can be accessed both manually and by voice command, making it easy to use for people with varying levels of technical expertise. Overall, the Virtual Eye app helps visually challenged individuals lead more fulfilling and independent lives.

Overall, Virtual Eye application is an essential tool for visually challenged individuals, helping them navigate their daily lives with ease and independence. With this app, they can access information, communicate, and identify objects without the need for a third party, enhancing their quality of life and sense of autonomy**.**

Keyword: Virtual Eye , Vision Assistance, Revolutionizing.

1. INTRODUCTION

Virtual Eye is an innovative mobile application designed to assist visually challenged individuals in their daily lives. For those who are visually Impaired, simple tasks such as reading a book or identifying an object can be incredibly difficult. This app aims to bridge that gap by providing a wide range of features that can help them perform these tasks with ease.

Without this application, visually challenged people face numerous challenges that hinder their daily activities Reading a book, for example, can be incredibly difficult and time-consuming as they need to rely on braille or someone to read it out for them. Identifying an object or a person's face can also be face difficulties because they can't rely on visible cues. This can lead to a feeling of isolation, frustration, and even anxiety, ultimately impacting their quality of life. With this application, these challenges can be overcome.

The app provides a text-to-speech feature that can convert any written material, including books, documents, and web pages, into audio format. This means that visually challenged individuals can easily access information without the need for a third party. In addition, the speech-to-text feature allows them to dictate their responses, messages, and notes without relying on someone else to write them down.

The image-to-audio feature is another valuable addition to the app, as it allows the visually impaired to identify objects or people's faces with ease. They can simply take a photo of the object or person and let the app describe it to them. This feature can be especially useful in situations where the visually impaired are alone and need assistance in identifying objects.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: 26 August 2023 Revised: 20 October 2023 Accepted: 02 November 2023

Furthermore, the PDF-to-audio feature enables them to access digital documents with ease, eliminating the need for special software or equipment. The app's user-friendly interface can be accessed both manually and by voice command, making it easy to use for people with varying levels of technical expertise.

2.System Analysis

Performing a system analysis is "the process of looking into a process or business to identify its goals and purposes and develop systems and procedures that will successfully achieve them." From a different angle, system analysis is a method of problem-solving that involves disassembling a system into its component parts and analyzing how well each part works both separately and jointly to fulfill the system's objectives. Data collection and analysis, problem detection, and dissection of a system into its component elements are all part of it.

A system's objectives are determined by doing a system analysis on it or one of its components. The system is strengthened and every component is present thanks to this approach to problem- solving. Virtual Eye is a mobile application that provides a range of features such as text-to-speech, speech-to-text, image-to-audio, and PDF-to-audio, making it a comprehensive assistive technology for visually challenged individuals. The app's ability to be accessed both manually and by voice command makes it accessible to individuals with varying levels of technical expertise.

Moreover, the app's speech-to-text feature allows users to send WhatsApp messages by converting their speech to text, enabling visually challenged individuals to communicate with greater ease. This feature enhances the app's overall functionality and usability for visually challenged individuals.

S. Venkatesh, (2023)“ An Improved coyote optimization algorithm- based clustering for Extending network lifetime in Wireless Sensor network” an Improved Coyote Optimization Algorithm- based Clustering Technique (ICOACT) is proposed for spreading the generation for making competent adoptions for cluster heads while sustaining a unswerving balance between manipulation and study.

Dr. Venkatesh. S (2022) discussed a image Enhancement and Implementation OF CLAHE Algorithm and Bilinear Interpolation. In summary, the literature suggests that assistive technologies have a significant impact on the quality of life of visually challenged individuals. Virtual Eye provides a comprehensive solution to many of the challenges faced by visually challenged individuals, and its speech-to-text feature enables them to communicate with greater ease, enhancing their independence and quality of life.

3. System Requirements.

A document or collection of documents known as system requirements contains descriptions of the characteristics and behavior of a system or software application. The customer's assertions about the system must be both general and detailed in order to be able to meet their wants. In the statement, the customer should be clear about what they want and how they want it. Any number of project requirements, including those related to contracts, problems, goals, and standards, may be demanded by a customer. The requirements for each system will be unique and will depend on the project.

4. Hardware Specification:

The physical computer hardware is the most typical set of specifications listed by any operating system or software program. It may be a desktop computer or a laptop with Windows 7 or later. processors consist of Any x86-64 CPU from Intel or AMD that supports the AVX2 instruction set and has four cores is advised. Simulink requires 4 GB of RAM, Polyspace 8 GB, with 4 GB suggested for each core. No particular graphics card is necessary. It is advised to use a hardware-accelerated graphics card with 1GB of GPU memory that supports OpenGL 3.3.

5. User System Requirements:

- Any personal Computer or Mobile
- Any Python Interpreter

6. Software Requirements

- OS :Windows 8 Or Windows 10
- Language: Python
- Package – Kivy, Open CV, NumPy, etc.,
- Data set: YOLO datasets.
- Tools: VS code.

7. Module Traceability

7.1. Workflow of the Application

The main screen of the application will display four icon buttons: text to speech, image to audio, speech to text, and PDF to audio. If the user wants to access any of the features using voice commands, they simply click on the microphone symbol at the bottom of the display. and speak the name of the feature they want to use. The app will then take them to the corresponding screen and perform the requested action.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: 26 August 2023 ***Revised:*** 20 October 2023 ***Accepted:*** 02 November 2023

8. Text to Speech Section

To use the text to speech feature, the user can tap on the text to speech icon or say "Open Text to speech" using the voice command option. This will take them to a new screen where the app will use OCR technology to recognize any text in the image and convert it into audio. The audio will then be played back to the user.

9. Image to Audio Section

To use the image to audio feature, the user can tap on the image to audio icon or say "Open Image to audio" using the voice command option. This will take them to a new screen where they can select an image from their device's or take a fresh photo with their device's camera. This app will then use YOLO algorithm to recognize any text in the image and convert it into audio. The audio will then be played back to the user.

10. Speech to Text Section

To use the speech to text feature, the user can tap on the speech to text icon or say "Open Speech to text" using the voice command option. This will take them to a new screen where they can speak into their device's microphone. Their speech will be converted into text by the app using speech recognition technology. Afterwards, the screen will display the text. To send a WhatsApp message using speech to text, the user can say "Send this message to [receiver name]" followed by the message they want to send using the voice command option. The app will then redirect to Whatsapp, it will use speech recognition technology to convert their speech into text, and then use Selenium technology to open WhatsApp and send the message to the desired contact.

11. PDF to Audio Section

To use the PDF to audio feature, the user can tap on the PDF to audio icon or say "Open PDF to audio" using the voice command option. This will take them to a new screen where they can select a PDF file from their device's storage. The app will then use modules like PyPDF2 to extract the text from PDF and pyttsx3 to recognize the text and convert it into audio. The audio will then be played back to the user. Once the user is done with any of the above tasks, they can go back to the main screen of the app and select another task to perform.

Architecture Diagram

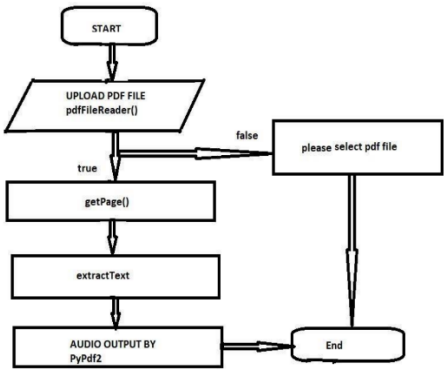

Figure1 .PDF to Audio Convertor

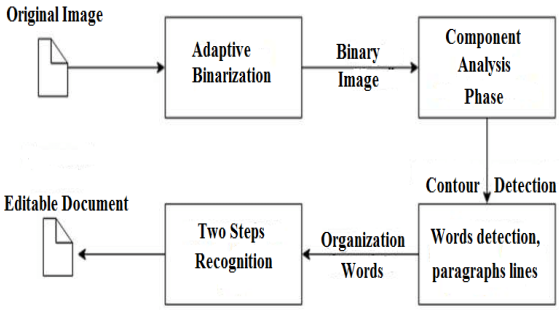

Figure 2. Text To Speech Convertor

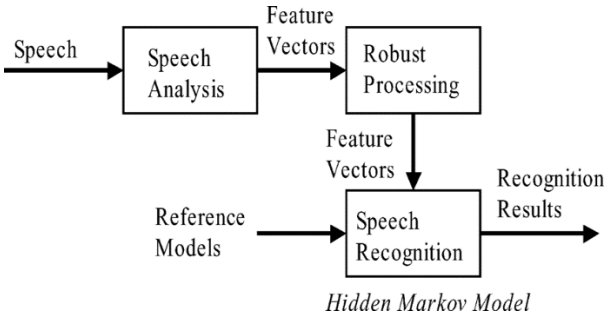

Figure 3. Speech to Text Convertor

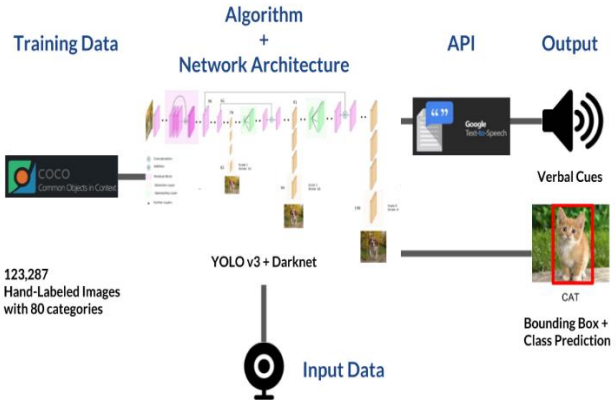

Figure 4. Image to Audio Convertor

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: 26 August 2023 Revised: 20 October 2023 Accepted: 02 November 2023

12. Non-Efficient Necessities

A. Serviceability: The Virtual Eye platform must be informal to use and accessible for users with different levels of technical expertise, including those with physical or cognitive impairments.

B. Performance: The platform should be fast and responsive, providing real-time feedback and minimizing any lag or delay in processing requests.

C. Safety: The platform must guarantee the privacy and security of user data, including sensitive information such as transcribed speech or personal preferences.

D. Scalability: The platform should be able to easily accommodate an increase in users and content without sacrificing usability or speed.

E. Consistency: The platform should be able to easily accommodate an increase in users and content without sacrificing usability or speed.

F. Compatibility: The platform should be compatible with different operating systems, devices, and browsers, ensuring that users can access it on their preferred device and platform.

G. Maintainability: The platform should have comprehensive documentation, support for future features, and be simple to manage and update.

H. Localization: The platform should support localization, including the ability to display content in different languages and recognize speech in different accents and dialects.

13. Results and Discussion

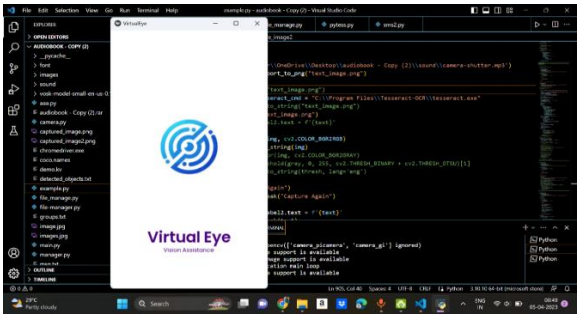

Figure 5. Splash Screen Of the Application

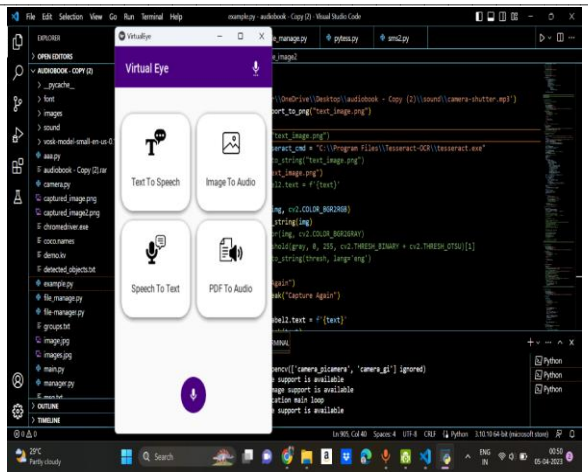

Figure 6. Home Screen Of the Application

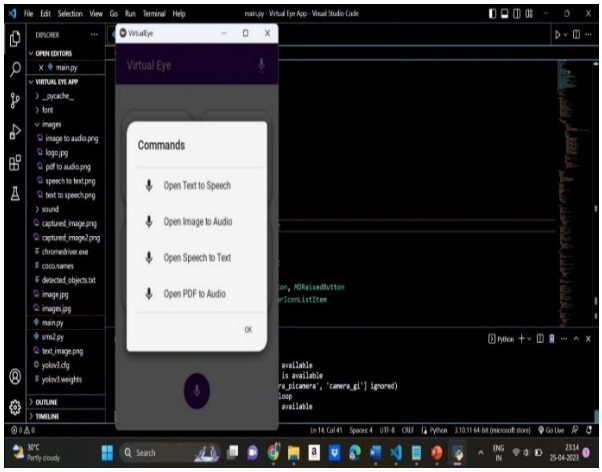

Figure 7. Commands Dialog Box

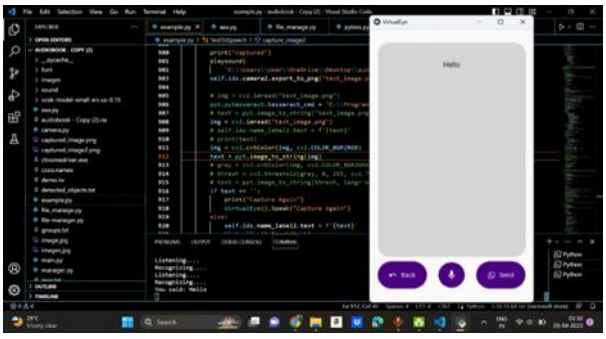

Figure 8. Speech to text Section.

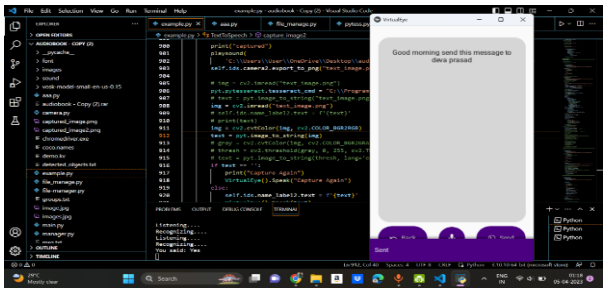

Figure 9. Sending Whatsapp Message By Using Voice Command

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

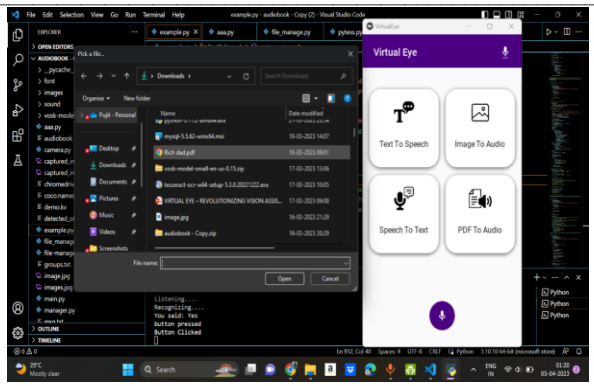

Figure 10. Selecting PDF To Audio Section

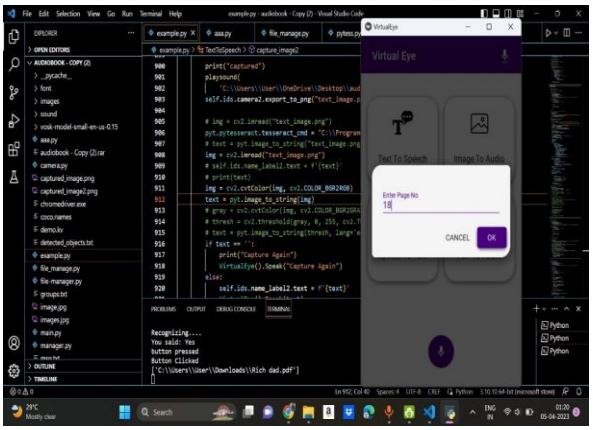

Figure 11. Select the Page Number

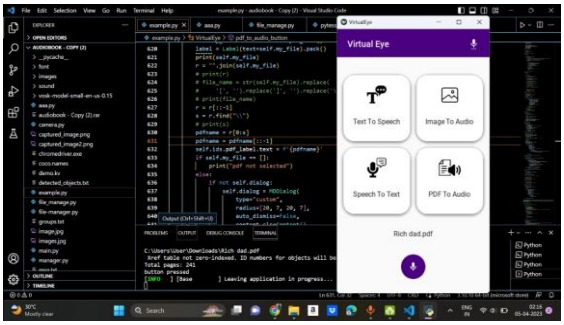

Figure 12. Reads the Selected Page

14. Conclusion & Future Works

Conclusion

The Virtual Eye mobile application is a powerful tool for individuals with visual or reading impairments, as well as those who simply prefer an audio-based interface. The app's text to speech, image to audio, speech to text, and PDF to audio features provide a variety of different ways to access and interact with information, making it more convenient and accessible for users. The user interface is designed to be user-friendly and accessible, with the option to use voice commands for hands-free operation. Additionally, the integration of pytesseract, pyttsx3, PyPDF2 libraries and YOLO algorithm allows for powerful text recognition and object identifivation, text to speech, PDF to audio and image to audio functionalities.

In addition to aiding individuals with visual and reading impairments, the Virtual Eye application can also be useful for those with blurred vision who may have difficulty reading small text, and for those with speaking disabilities who want to practice and improve their speaking ability. By converting text to speech and providing speech to text functionality, the application can facilitate communication for individuals with speaking disabilities and provide a platform for practicing and improving their spoken language skills.

Overall, Virtual Eye is a promising tool for improving accessibility and inclusivity for individuals with visual or reading impairments, and has the possible to greatly enhance the eminence of life for many users.

15. Future Works

The addition of a personalized content recommendation system based on user preferences and behaviour. This system could use mechanism learning algorithms to examine a user's interactions with the platform, such as the types of content they listen to, transcribe, or summarize, and then use this data to suggest new content that the user might be interested in .For example, if a user frequently listens to news articles about technology and innovation, the platform could suggest other articles or podcasts in the same subject area. This would allow users to discover new content that is tailored to their interests and preferences, while also increasing engagement and usage of the Virtual Eye platform.

The ability to read and transcribe text in multiple languages. By expanding its language recognition and translation capabilities, Virtual Eye could become a powerful tool for users who need to access and understand content in various languages. With this enhancement, Virtual Eye would be able to analyze and process text, images, and PDFs in languages other than the user's primary language. It would have the ability to accurately transcribe spoken words and convert them into text in multiple languages.

Additionally, Virtual Eye could incorporate machine translation capabilities, enabling users to listen to and read content in their preferred language. This feature would greatly benefit users who need to interact with multilingual documents or communicate with individuals who speak different languages. The implementation of advanced noise reduction techniques to minimize disturbances during voice command input. By integrating noise reduction algorithms, Virtual Eye can enhance the accuracy and reliability of voice commands, even in noisy environments.

This enhancement would improve the user experience by ensuring that voice commands are accurately captured and interpreted, regardless of background noise or environmental factors. It would enable users to interact with the Virtual Eye platform seamlessly, even in busy or noisy settings.

REFERENCES

1. Chaw Su Thu Thu and TheingiZin "Implementation of Text to Speech Conversion" International Journal of Engineering Research & Technology (IJERT) vol. 3 no. 3 March 2014 ISSN 2278-0181.
2. Joseph Redmon and Ali Farhadi "YOLOv3: An Incremental Improvement" 2018.
3. RushikeshLaxmikant Kulkarni "Handwritten Character Recognition Using HOG COM by OpenCV& Python" International Journal of Advance Research in Computer Science and Management Studies vol. 5 no. 4 April 2017.
4. Q. Ye and D. Doermann "Text detection and recognition in imagery: A survey" IEEE transactions on pattern analysis and machine intelligence vol. 37 no. 7 pp. 1480-1500 2014.
5. S. Azenkot and N. B. Lee "Exploring the use of speech input by blind people on mobile devices" 15th International ACM SIGACCESS Conference on Computers and Accessibility. pp. 11 2013.
6. T. Chakraborty and D. Samanta "EasyTap: An Easy Text Entry Method for Blind Mobile Phone Users" Australian Journal of Intelligent Information Processing Systems vol. 13 no. 4 2014.
7. A. Kumar T. Paek and B. Lee "Voice typing: a new speech interaction model for dictation on touchscreen devices" SIGCHI Conference on Human Factors in Computing Systems. pp. 2277-2286 2012.
8. AbdelhamidBenhocine, LamriLaouamer, Laurent Tchamnda Nana, Anca Christine Pascu Improving extraction of watermarks in color attacked watermarked images Journal of Communication and Computer, ISSN 1548-7709, USA
9. Abraham, J., & Paul, V. (2016). An imperceptible spatial domain color image watermarking scheme. Journal of King Saud University - Computer and Information Sciences. doi:10.1016/j.jksuci.2016.12.004
10. Abuturab, M. R. (2017). Multiple color-image fusion and watermarking based on optical interference and wavelet transform. Optics and Lasers in Engineering, 89, 47–58. doi:10.1016/j.optlaseng.2016.02.014
11. Ahmad, F., & Cheng, L.-M. (2018). Authenticity and copyright verification of printed images. Signal Processing, 148, 322–335. doi:10.1016/j.sigpro.2018.02.029
12. Albalawi, U., Mohanty, S. P., &Kougianos, E. (2017). A new region aware invisible robust blind watermarking approach. Multimedia Tools and Applications, 76(20), 21303-21337.
13. Ammar, M., Mitrea, M., Hasnaoui, M., & Le Callet, P. (2018). MPEG-4 AVC stream-based saliency detection. Application to robust watermarking. Signal Processing: Image Communication, 60, 116–130. doi:10.1016/j.image.2017.09.007
14. Ancuti, Codruta O.; Ancuti, Cosmin; De Vleeschouwer, Christophe; Bovik, Alan C.. Single-Scale Fusion: An Effective Approach to Merging Images. In: IEEE Transactions on Image Processing, Vol. 26, no.1, p. 65-78 (2017)
15. Ansari, A., Hong, S., Saavedra, G., Javidi, B., & Martinez-Corral, M. (2018). Ownership protection of plenoptic images by robust and reversible watermarking. Optics and Lasers in Engineering, 107, 325–334. doi:10.1016/j.optlaseng.2018.03.028
16. Araghi, T. K., Manaf, A. A., &Araghi, S. K. (2018). A secure blind discrete wavelet transform based watermarking scheme using two-level singular value decomposition. Expert Systems with Applications, 112, 208–228. doi:10.1016/j.eswa.2018.06.024
17. Asikuzzaman, M., & Pickering, M. R. (2017). An Overview of Digital Video Watermarking. IEEE Transactions on Circuits and Systems for Video Technology, 1–1. doi:10.1109/tcsvt.2017.2712162
18. Bai, Y., Bai, S., Zhu, G., You, C., & Liu, B. (2010, October). A blind audio watermarking algorithm based on FFT coefficients quantization. In Artificial Intelligence and Education (ICAIE), 2010 International Conference on (pp. 529-533). IEEE.
19. Barni, M., &Bartolini, F. (Eds.). (2004). Watermarking systems engineering: enabling digital assets security and other applications. CRC Press.
20. Barni, M., Bartolini, F., Caldelli, R., De Rosa, A., & Piva, A. (2000, May). A robust watermarking approach for raw video. In Proceedings of the 10th international packet video workshop.
21. S. Venkatesh, (2022) “Image Enhancement and Implementation OF CLAHE Algorithm and Bilinear Interpolation”, International Journal of Cybernetics and System Annexure 1, ISSN: 0196-9722, Vol. 6, Issue 12, November 2022.
22. S. Venkatesh, (2023)“An Improved coyote optimization algorithm- based clustering for Extending network lifetime in Wireless Sensor network”, KSII Transactions on Internet and Information System , Annexure . vol.17 no.7 July 2023.