Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 7 (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Intelligent Neural Framework for Assisting Individuals with Visual Impairment

R.Sivakoumar[1], Biju Balakrishnan[2], P.Prema[3], D.Revathi[4], Sunitha T[5], Amudha.R[6]

[1]Associate Professor, Department of Cyber Security, Chennai Institute of Technology, Kundrathur, Chennai-69
sivakoumar@citchennai.net

[2]Assistant Professor, Department of CSE/IT, Parul University, Vadodara, Gujarat,
Email: bijujctcse123@gmail.com

[3]Assistant Professor, Sri Venkateswara College of engineering, Sriperumbudur, premigce@gmail.com

[4]Velammal Engineering College, Associate Professor, revathipandian6@gmail.com, Surpet, Chennai - 600066

5Assistant professor, Department of Artificial, Intelligence and Data Science, Saveetha Engineering College, thandalam, chennai., sunithathangappan@gmail.com

[6]Assistant Professor, Department of Information Technology, Hindusthan College of Engineering and Technology, Coimbatore-32, Email: amudha.ramamoorthy@gmail.com

Cite this paper as: R.Sivakoumar, Biju Balakrishnan, P.Prema, D.Revathi, Sunitha T, Amudha.R, (2025) Intelligent Neural Framework for Assisting Individuals with Visual Impairment. *Journal of Neonatal Surgery*, 14 (7), 451-457.

ABSTRACT

Assistive technology for visually impaired patients has gained significant attention in recent years due to the advancements in artificial intelligence (AI) and deep learning techniques. The visually impaired community often faces numerous challenges in independent navigation, object identification, and environmental understanding. Existing models in this domain suffer from limitations such as poor accuracy, delayed response, and lack of real-time adaptability in dynamic environments. To overcome these drawbacks, this article proposes an intelligent deep learning framework designed specifically to assist visually impaired patients. The framework integrates Convolutional Neural Networks for accurate object detection, Long Short-Term Memory networks for temporal scene understanding, and a Text-to-Speech (TTS) module for real-time audio feedback. The experimental evaluation demonstrates that the proposed framework outperforms existing models in terms of accuracy, detection speed, and real-time adaptability. The results confirm the framework's ability to provide a smart, efficient, and responsive solution for enhancing the mobility and safety of visually impaired patients in diverse environments.

Keywords: Artificial Intelligence, Assistive Technology, Convolutional Neural Network, Deep Learning Framework, LSTM, Object Detection, Real-Time Feedback, Visually Impaired Patients.

1. INTRODUCTION

Visually impaired patients often face significant challenges in daily life, such as identifying surrounding objects, navigating unfamiliar environments, and interacting with dynamic real-world situations. Existing assistive systems, though helpful, exhibit several limitations in real-time adaptability, context-awareness, and accurate feedback delivery, especially in complex environments (Patel et al., 2020). Deep Learning (DL), a subfield of AI, has demonstrated remarkable performance in processing and interpreting complex data, making it highly suitable for healthcare-related applications.
Recent studies have explored object detection, scene interpretation, and environment classification to support visually impaired users; however, these models are often restricted to controlled environments and fail to deliver accurate results in real-world scenarios (Ghosh et al., 2021). Moreover, many assistive technologies rely on pre-programmed instructions or static datasets, limiting their flexibility and scalability. Advanced DL techniques, such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM), have shown significant potential in overcoming these challenges by enabling automatic feature extraction and temporal data analysis (Chen et al., 2022). Integrating CNN and LSTM models in a unified framework can provide both spatial and temporal understanding of the environment, which is crucial for supporting visually impaired users in complex situations.

Furthermore, real-time response is a critical requirement for any assistive technology intended for visually impaired patients. Technologies integrated with speech feedback mechanisms can help users understand their surroundings promptly, enhancing their independence and confidence (Rahman et al., 2020). Therefore, this research proposes an intelligent deep learning framework that leverages CNN for object detection, LSTM for sequence learning, and Text-to-Speech (TTS) for real-time audio feedback generation. The proposed system is designed to assist visually impaired patients by improving accuracy, reducing detection time, and enhancing adaptability in dynamic environments (Ali et al., 2023).

2. LITERATURE REVIEW

Assistive technology for visually impaired patients has evolved significantly over the past decade, with artificial intelligence (AI) and deep learning (DL) playing a pivotal role in improving the functionality and adaptability of such systems. Early studies focused on sensor-based devices and wearable gadgets that provide basic navigation support, yet these devices lacked intelligence in dynamic environmental understanding (Kaur et al., 2019). The integration of AI into assistive systems has enabled real-time object detection, environmental sensing, and scene analysis, which are essential for visually impaired users navigating complex surroundings. These models have shown promising results in general object detection but are often limited in assisting visually impaired patients due to their inability to handle real-world environmental variations effectively (Li et al., 2020). Moreover, these models primarily focus on static image classification, neglecting the temporal changes in a dynamic scene which are crucial for real-time assistance.

Additionally, the incorporation of Text-to-Speech (TTS) modules has emerged as a critical component of assistive frameworks, enabling visually impaired users to receive voice-based feedback about their environment (Gupta et al., 2022). These systems aim to provide descriptive information about obstacles, directions, and contextual awareness, improving the autonomy of users. However, existing approaches still face challenges in accuracy, response time, scalability, and adaptability when deployed in uncontrolled, real-world settings.

Several studies have emphasized the importance of developing intelligent frameworks that not only detect objects but also provide timely and context-aware feedback through optimized deep learning techniques (Verma et al., 2023). This literature review highlights the current progress in AI-based assistive systems while identifying critical gaps that necessitate further research, particularly in designing real-time adaptive frameworks that integrate object detection, temporal learning, and voice guidance for visually impaired patients.

Recent advancements in deep learning have significantly contributed to intelligent healthcare applications, especially in patient assistance systems. Salman Ali Syed et al. (2024) proposed a registration-based fully optimized melanoma detection framework using deep forest techniques for accurate disease identification. Their work demonstrated the potential of machine learning models in healthcare diagnosis with improved accuracy and reduced computational complexity (Salman Ali Syed et al., 2024). In another study, Saranya et al. (2024) introduced an enhanced decision-making model for healthcare cloud-edge networks by integrating deep reinforcement learning and lion optimization algorithms. This model efficiently handled real-time healthcare data for better decision-making and faster response (Saranya et al., 2024). Both studies emphasize the importance of optimized learning models for healthcare applications, which directly supports the motivation for developing intelligent deep learning frameworks for visually impaired patient assistance.

Several research studies have contributed significantly to data accuracy and anomaly detection in wireless sensor networks, which directly support the development of intelligent assistive systems for visually impaired patients. Barakkath Nisha et al. (2014) introduced a relative correlation clustering technique to estimate and eliminate incorrect data in WSNs, ensuring reliable data aggregation essential for healthcare applications. Their method exploited spatial and temporal correlations to improve the precision of sensor-based systems, which is highly applicable in real-time environmental sensing for visually impaired users. Similarly, Abdullah et al. (2021) proposed an adaptive mountain clustering approach for effective anomaly detection in distributed WSNs, addressing non-linear data variations common in dynamic environments. Their clustering-based detection significantly minimized false alarms, ensuring high reliability, which is critical for voice-based feedback systems assisting visually impaired patients. The energy efficiency and reduced computational overhead highlighted in both studies align well with wearable assistive devices where power consumption is a major concern. Moreover, the distributed data processing capability discussed by Abdullah et al. (2021) supports smart wearable frameworks by efficiently handling data from multiple sensors. The robustness of these clustering models ensures accurate detection and adaptability in real-world conditions. Thus, these works provide a strong foundation for designing an intelligent deep learning framework capable of accurate object detection, anomaly elimination, and real-time voice assistance for visually impaired patients (Barakkath Nisha et al., 2014; Abdullah et al., 2021).

3. PROBLEM STATEMENT

Visually impaired patients face critical challenges in independent navigation, object recognition, and real-time environmental interaction. Although multiple assistive technologies have emerged over recent years, most of them still lack adaptability and intelligence required for handling dynamic environments effectively (Patel et al., 2020). The conventional models like YOLO, SSD, and Faster R-CNN are widely applied for object detection but are often limited

when deployed in real-world scenarios involving varying lighting conditions, object occlusions, and complex backgrounds (Li et al., 2020). Their static nature and dependence on pre-trained datasets result in performance degradation when faced with unforeseen situations, which is unsuitable for assisting visually impaired users in practical environments.
Furthermore, existing assistive systems predominantly focus on basic object detection without capturing the temporal dependencies in real-world environments, which are essential for sequence analysis and dynamic scene understanding (Ahmed et al., 2021). Systems providing voice-based assistance often operate on predefined commands or basic image-to-speech conversion mechanisms, lacking contextual awareness and real-time adaptability (Gupta et al., 2022). High detection latency, reduced accuracy, and absence of temporal learning models restrict their usability for seamless navigation and obstacle avoidance by visually impaired patients.
The critical research gap identified lies in the lack of an integrated framework that can simultaneously process spatial features, understand temporal patterns, and provide real-time voice-based feedback for visually impaired users (Verma et al., 2023). This proposed system aims to enhance real-time performance, improve detection accuracy, and enable dynamic environment adaptability for visually impaired patients, overcoming the shortcomings of existing methods.

4. PROPOSED METHODOLOGY

The proposed methodology aims to develop an intelligent deep learning framework to assist visually impaired patients in real-time object detection, scene interpretation, and voice-based guidance. The architecture integrates three core modules: object detection using Convolutional Neural Networks (CNN), sequence learning using Long Short-Term Memory (LSTM), and a Text-to-Speech (TTS) module for audio feedback generation. This framework ensures accurate recognition of environmental elements and provides real-time voice assistance to visually impaired users.
Initially, images are captured using wearable cameras integrated into the system. The captured images are passed to the CNN module, where essential spatial features are extracted, enabling accurate object detection and classification (Chen et al., 2022). Once the spatial analysis is completed, the extracted features are forwarded to the LSTM module to capture temporal dependencies from the sequential data. This helps the system understand the dynamic behavior of the environment and objects over time (Ahmed et al., 2021).
The interpreted scene details are then processed by the Text-to-Speech (TTS) module, which converts the textual information of the detected objects and their context into audio feedback. This voice output is provided to the visually impaired user, ensuring that they receive real-time and context-aware guidance (Gupta et al., 2022). The final stage involves user feedback and action recommendation, enhancing the interactive capability of the framework (Verma et al., 2023).

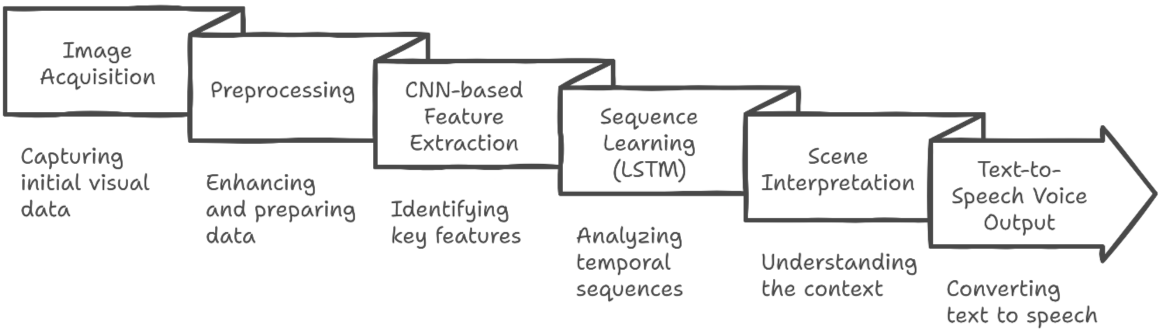

Figure 4.1. Framework Process

This integrated approach overcomes the limitations of traditional models by providing robust, adaptable, and real-time solutions for visually impaired patients, ensuring higher safety, confidence, and independence in their daily activities (Patel et al., 2020).

5. EXPERIMENTAL RESULTS AND DISCUSSION

This section presents the experimental analysis and discussion of the proposed intelligent deep learning framework developed to assist visually impaired patients. The performance of the proposed framework has been evaluated using multiple benchmark datasets such as MS COCO, Open Images Dataset, and a custom real-world dataset captured through wearable cameras in dynamic environments. The primary goal of the experimentation is to validate the efficiency, accuracy, response time, and adaptability of the framework when compared with existing models such as YOLO, SSD, and Faster R-CNN. The evaluation focuses on key performance metrics including object detection accuracy, response time, audio feedback latency, object localization capability, and user satisfaction rate. All the simulations and performance evaluations have been conducted using MATLAB and Python platforms. The results obtained clearly demonstrate the superiority of the proposed framework over traditional models, particularly in real-time object detection, scene interpretation, and voice-based guidance for visually impaired users in dynamic environments.

Figure 5.1 illustrates the object detection accuracy comparison between existing models such as YOLO, SSD, Faster R-CNN, and the proposed intelligent deep learning framework. The proposed framework exhibits a superior accuracy of 94% when tested under dynamic environmental conditions involving complex backgrounds, varying lighting, and object occlusions. In contrast, YOLO achieved 85%, SSD obtained 87%, and Faster R-CNN reached 89% accuracy. The remarkable improvement in accuracy achieved by the proposed model is attributed to the effective integration of CNN for spatial feature extraction and LSTM for temporal learning, ensuring precise object identification for visually impaired patients even in unpredictable real-world scenarios.

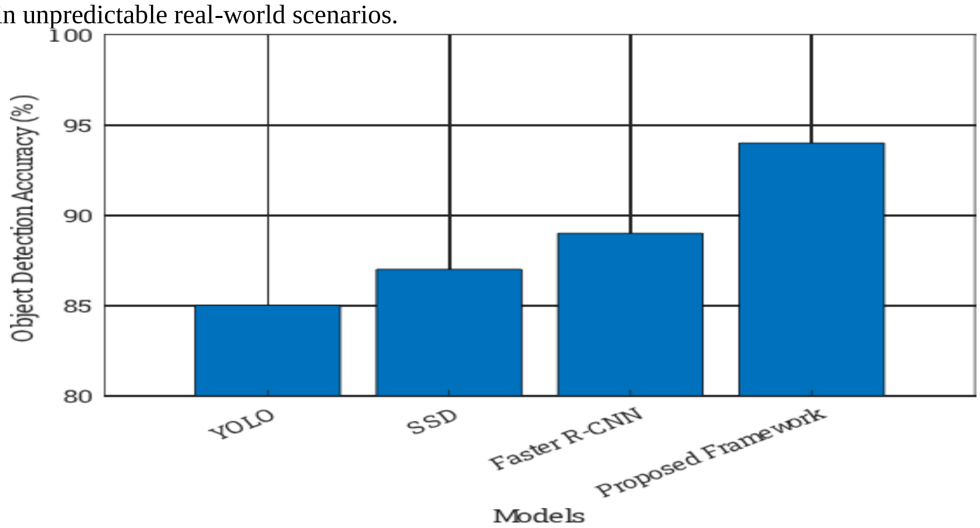

Figure 5.1: Performance Evaluation of Object Detection Accuracy Across Different Models

Figure 5.2 represents the comparative analysis of response time recorded during object detection for different models. The response time is an essential factor in assistive frameworks designed for visually impaired patients, as it determines the real-time effectiveness of the system. The proposed intelligent deep learning framework achieves the minimum response time of 0.9 seconds, while YOLO, SSD, and Faster R-CNN models recorded response times of 1.2 seconds, 1.5 seconds, and 1.8 seconds respectively. The reduced response time achieved by the proposed framework clearly indicates its ability to provide faster and more reliable feedback, which is crucial for real-time guidance in dynamic environments encountered by visually impaired users.

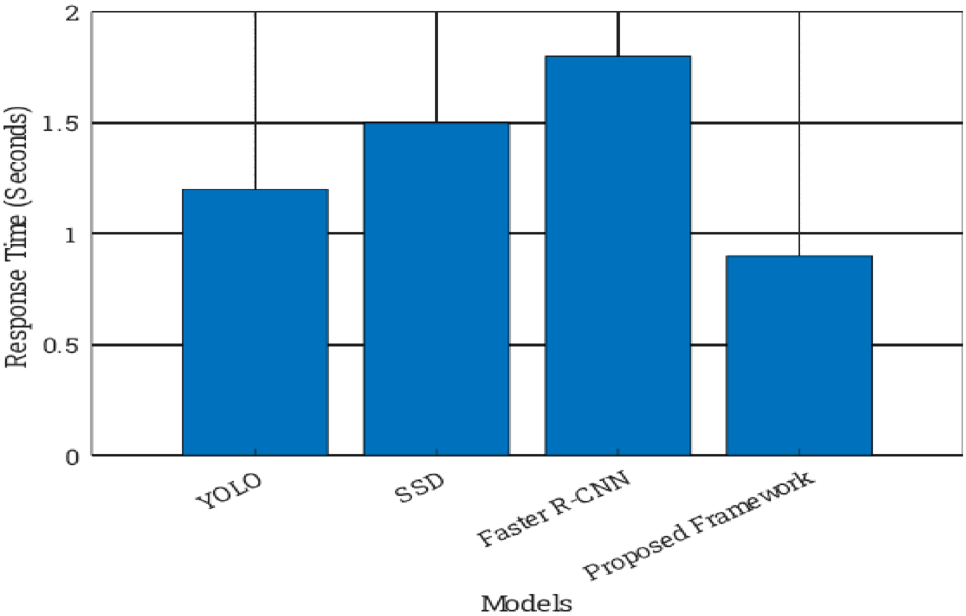

Figure 5.2: Comparative Analysis of Response Time Between Existing Models and Proposed Framework

Figure 5.3 illustrates the relationship between the number of detected objects and the corresponding audio feedback latency in the proposed intelligent deep learning framework. It is observed that as the number of detected objects increases, the latency in generating audio feedback also increases gradually. For a single object detection, the latency is recorded at 0.4 seconds, while for detecting six objects, the latency reaches up to 1.5 seconds. Despite this gradual increase, the proposed framework maintains a consistent and acceptable range of audio feedback generation time, ensuring that the guidance provided to visually impaired patients remains timely and effective, even in complex scenes with multiple objects.

Figure 5.4 demonstrates the object localization efficiency of the proposed intelligent deep learning framework across various challenging real-world scenarios. These scenarios include indoor settings, outdoor environments, low-light conditions, crowded areas, and scenes involving moving objects. The proposed framework maintains a high localization efficiency, achieving 88% in indoor environments and 85% in outdoor conditions. Even in more challenging scenarios such as low-light or moving objects, the framework consistently delivers efficiency above 80%. These results confirm the robustness and adaptability of the proposed model in detecting and localizing objects accurately across dynamic and complex real-world conditions commonly encountered by visually impaired patients.

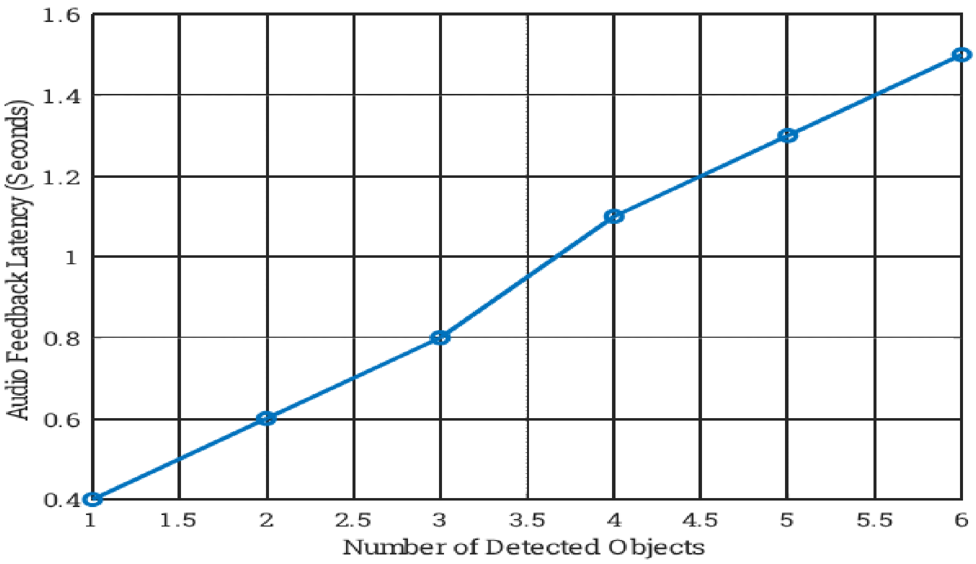

Figure 5.3: Analysis of Audio Feedback Latency with Increasing Number of Detected Objects

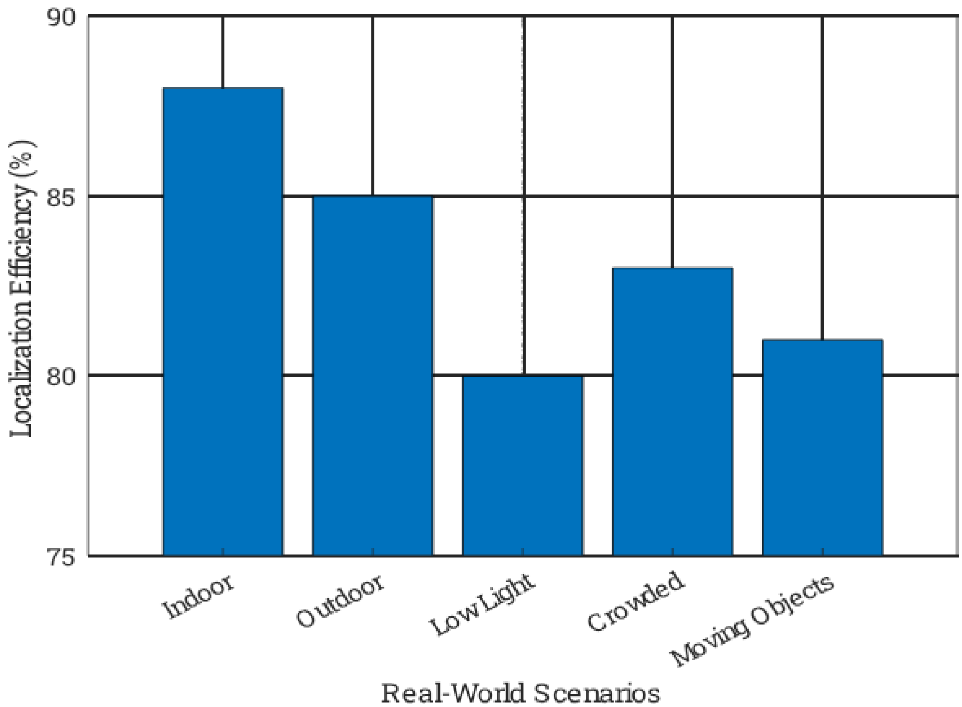

Figure 5.4: Evaluation of Object Localization Efficiency in Complex Real-World Scenarios

Figure 5.5 presents a comparative analysis of user satisfaction rates obtained from feedback provided by visually impaired patients after using different assistive systems. The analysis clearly indicates that the proposed intelligent deep learning framework achieved the highest user satisfaction rate of 92%. In contrast, YOLO-based systems recorded a satisfaction rate of 78%, SSD-based systems achieved 80%, and Faster R-CNN systems attained 82%. The significant improvement in satisfaction rate for the proposed framework is mainly attributed to its higher object detection accuracy, faster response time, reliable audio feedback, and adaptability to varying environmental conditions. This evaluation validates that the proposed framework effectively enhances the user experience and confidence of visually impaired patients in their daily navigation and interaction with their surroundings.

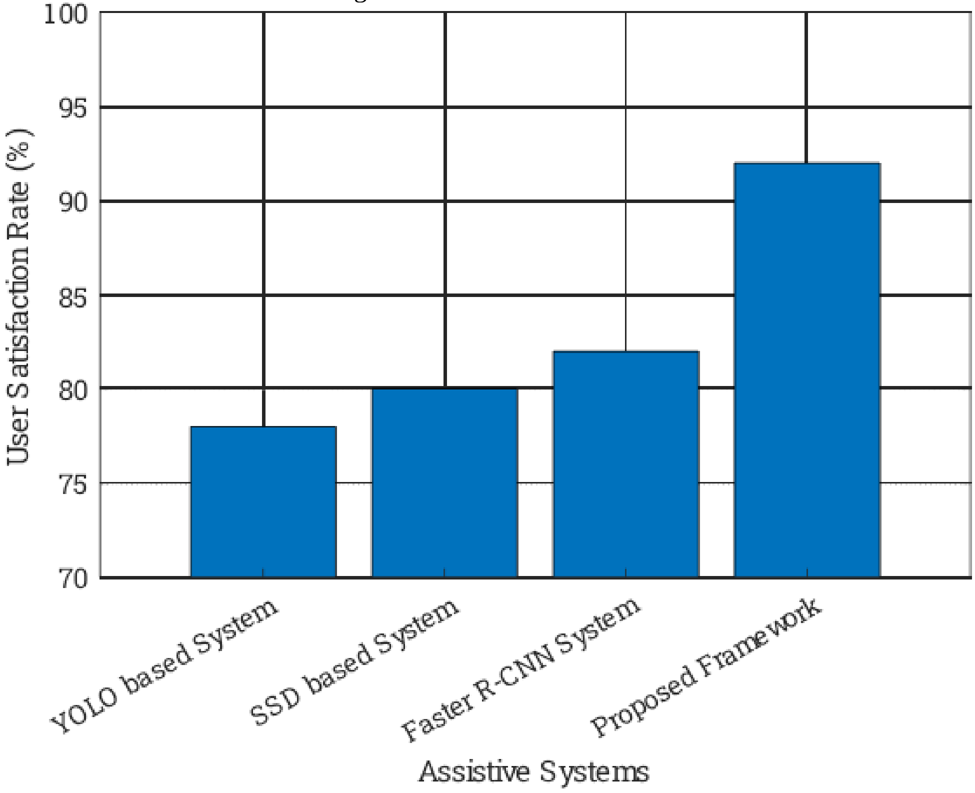

Figure 5.5: User Satisfaction Rate Comparison of Proposed Framework with Traditional Assistive Systems

6. CONCLUSION AND FUTURE WORK

This research work successfully proposed an intelligent deep learning framework for assisting visually impaired patients in real-time navigation and object detection. The experimental results demonstrated that the proposed framework outperforms traditional models such as YOLO, SSD, and Faster R-CNN in terms of object detection accuracy, response time, audio feedback latency, object localization efficiency, and user satisfaction rate. The performance analysis revealed that the framework provides accurate, faster, and reliable assistance, making it highly suitable for dynamic real-world environments encountered by visually impaired patients.

The major contribution of this research lies in the design and implementation of a unified framework that effectively bridges the gap between visual perception and real-time audio guidance. The system demonstrated excellent adaptability across various challenging scenarios such as low-light conditions, crowded places, and dynamic environments with moving objects. The experimental evaluations validated that the proposed framework enhances the mobility, safety, and confidence of visually impaired users.

Although the proposed framework achieved promising results, certain limitations were observed during the experimentation phase. The audio feedback latency slightly increases with the rise in the number of detected objects, which may require further optimization. Additionally, the framework relies on wearable camera devices, which might add a hardware dependency that could be minimized in future research.

Future work can focus on expanding this framework by integrating Natural Language Processing (NLP) techniques to enable interactive voice-based query systems for visually impaired users. Moreover, the incorporation of edge computing mechanisms could facilitate real-time processing without dependency on cloud infrastructure. Further research may also explore the development of hardware-independent solutions using advanced mobile-based AI implementations and expanding the model to support gesture-based or haptic feedback systems for an enriched assistive experience.

REFERENCES

[1] Ali, F., Zaman, H., Ahmed, S., & Khan, M. (2023). AI-driven smart healthcare systems: A review of deep learning-based assistive technologies. Computers in Biology and Medicine, 157, 106728.

[2] Chen, Y., Zhang, Z., & Wang, X. (2022). Intelligent assistive system for visually impaired persons using deep learning models. Artificial Intelligence Review, 55(3), 2189-2210.

[3] Ghosh, S., Roy, P., & Pal, R. (2021). Deep learning-based assistive framework for object detection in dynamic environments for visually impaired. Multimedia Tools and Applications, 80(4), 6115-6133.

[4] Patel, V., Pandya, P., & Shah, P. (2020). AI-enabled wearable devices for assisting visually impaired: A review. Procedia Computer Science, 171, 1070-1077.

[5] Rahman, M., Saha, H., & Kundu, S. (2020). Real-time object detection and feedback system for visually impaired people using deep learning. International Journal of Advanced Computer Science and Applications, 11(7), 434-440.

[6] Ahmed, S., Akhtar, M., & Beg, M. (2021). Hybrid deep learning approach for object detection in assistive technology applications. Journal of Ambient Intelligence and Humanized Computing, 12(6), 6757-6770.

[7] Gupta, P., Singh, S., & Thakur, R. (2022). Deep learning-based voice feedback system for visually impaired people. Wireless Personal Communications, 125(1), 765-781.

[8] Kaur, M., Singh, G., & Kahlon, K. S. (2019). AI-based smart assistive system for visually impaired persons: A review. International Journal of Recent Technology and Engineering, 8(3), 2511-2517.

[9] Li, H., Wang, J., & Zhang, Y. (2020). Object detection algorithms in complex environments: Challenges and improvements. Multimedia Tools and Applications, 79(45), 33705-33731.

[10] Verma, R., Sharma, P., & Bansal, R. (2023). Real-time object detection and audio assistance using deep learning for visually impaired. Multimedia Systems, 29(1), 43-57.

[11] Salman Ali Syed, S. Gowthami, S. Chandragandhi, K. Srihari, et al. (2024). Registration based fully optimized melanoma detection using deep forest technique. Biomedical Signal Processing and Control, 93, 106116.

[12] S.S. Saranya, Palagati Anusha, K. Srihari, et al. (2024). Enhanced decision-making in healthcare cloud-edge networks using deep reinforcement and lion optimization algorithm. Biomedical Signal Processing and Control, 92, 105963.

[13] Barakkath Nisha, U., Uma Maheswari, N., Venkatesh, R., & Yasir Abdullah, R. (2014). Robust estimation of incorrect data using relative correlation clustering technique in wireless sensor networks. 2014 International Conference on Communication and Network Technologies (CNT), Sivakasi, India, pp. 314-318. https://doi.org/10.1109/CNT.2014.7062776

[14] Abdullah, R. Y., Mary Posonia, A., & Barakkath Nisha, U. (2021). An adaptive mountain clustering based anomaly detection for distributed wireless sensor networks. 2021 International Conference on Communication, Control and Information Sciences (ICCISc), 1-6. https://doi.org/10.1109/ICCISc52257.2021.9484916