IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Blind People Currency and Object Detection

[1]Dr Sharath Kumar Y H, [2]Anvitha , [2]Jeevitha L, [2]Ujwal R, [2]Yashas Gowda M B
[1] Professor and Head, [2] Students
[1]Department of Information Science and Engineering,
[1]Maharaja Institute of Technology Mysore,India

Abstract: In today's world, visually impaired individuals encounter numerous challenges in daily activities such as recognizing objects, identifying currency denominations, and reading printed text. To address these difficulties, we propose a real-time Android-based assistive application that leverages deep learning models for independent navigation and interaction. The system integrates the YOLOv8 object detection algorithm for both general object and currency recognition, Optical Character Recognition (OCR) for text extraction, and Text-to-Speech (TTS) technology for audio feedback. Gesture-based swipe interactions enable users to seamlessly switch between functionalities, enhancing usability without requiring visual cues. Experimental evaluations demonstrate that the YOLOv8 model achieves an object detection accuracy of 94.5% and currency recognition accuracy of 95.2% across various lighting and environmental conditions. The OCR module yields a text extraction accuracy of 91.3% with minimal latency. The application achieves real-time processing speeds with an average inference time of 180 milliseconds per frame on a standard Android smartphone. By operating fully offline, the system ensures privacy, fast response, and accessibility. The proposed solution empowers visually impaired individuals to perform daily tasks independently, offering a practical, scalable, and highly effective assistive technology.

***Index Terms* – Assistive Technology, Blind Assistance, Object Detection, YOLOv8, Currency Recognition, Optical Character Recognition (OCR), Text-to-Speech (TTS), Gesture-based Interaction, Real-time Processing, Mobile Application**

I. Introduction

In recent years, satellite imagery has emerged as a powerful tool for addressing critical global challenges, including environmental monitoring, urban planning, disaster management, and socio-economic analysis. Among these applications, the prediction and assessment of poverty through satellite imagery analysis has gained increasing attention, particularly in regions where traditional ground surveys are expensive, slow, or infeasible. By leveraging advancements in remote sensing technology and machine learning, it is now possible to infer economic conditions from satellite-captured visual cues such as infrastructure development, land use patterns, nighttime light intensity, and urban density.

Traditional poverty mapping methods typically rely on household surveys and census data, which, while accurate, are costly, time-consuming, and often limited to small geographical areas. In contrast, satellite imagery offers near-real-time, large-scale coverage, enabling the analysis of socio-economic conditions across diverse and often inaccessible regions. However, interpreting raw satellite images requires sophisticated computational methods capable of extracting meaningful patterns from complex and high-dimensional data.

Deep learning, particularly convolutional neural networks (CNNs) and recurrent models like RNNs and LSTMs, has revolutionized the field of image analysis and is now being successfully applied to satellite data interpretation. These models excel at capturing spatial hierarchies, detecting subtle variations in pixel intensities, and learning rich feature representations directly from raw images. Coupled with unsupervised learning techniques such as K-means clustering, deep learning provides a robust framework for segmenting satellite images, extracting luminosity-based features, and predicting poverty levels with high accuracy.

This project proposes a novel deep learning pipeline for poverty prediction based on satellite images. The methodology involves preprocessing satellite images, segmenting them into clusters based on color and brightness using K-means clustering, calculating cluster-wise luminosity features, and training classification models (ANN, RNN, and LSTM) to predict poverty classes. By focusing on luminosity as a proxy for economic development—especially in nighttime imagery—this approach aims to provide a scalable, cost-effective alternative to traditional survey-based poverty assessments.

Through extensive experimentation, the models demonstrated outstanding classification performance, achieving test accuracies of over 99% across multiple architectures. These results highlight the immense potential of combining satellite imagery and deep learning for socio-economic mapping and contribute significantly towards the goal of data-driven poverty alleviation strategies.

II. Literature Survey

In recent years, assistive technology for visually impaired individuals has seen considerable advancements, particularly with the integration of machine learning, deep learning, and mobile computing. This section provides a detailed survey of existing systems focused on object detection, currency recognition, real-time navigation assistance, and mobile-based accessibility solutions. We critically analyze their methodologies, capabilities, and limitations to contextualize the contributions of the proposed system.

One of the fundamental challenges faced by visually impaired individuals is the identification of different currency denominations during financial transactions. Several studies have proposed systems aimed at addressing this problem through automated recognition techniques.

Singh et al. [1] introduced IPCRF (Indian Paper Currency Recognition Framework), an end-to-end framework based on deep learning methodologies designed specifically for Indian currency notes. Their approach utilized a convolutional neural network (CNN) architecture to classify various denominations accurately. The system demonstrated robust recognition under controlled conditions; however, its real-world performance was sensitive to variations in lighting, background clutter, and note deformation. Moreover, the reliance on static image capture introduced latency between note presentation and recognition feedback, limiting its real-time applicability.

Similarly, Markad et al. [2] developed a mobile-based currency detection application aimed at identifying Indian banknotes for the visually impaired. Their system focused on image preprocessing and template matching techniques. While it offered a lightweight and easily deployable solution, the system's dependence on high-quality input images restricted its practical usability, especially in dynamic environments where perfect alignment and lighting could not be guaranteed.

Addressing broader applicability, Park and Park [4] proposed the Multinational Banknote Detecting Model (MBDM), which aimed to recognize currencies from multiple countries. Their system employed a deep learning pipeline that classified banknotes across different currencies. Although MBDM showcased versatility, the trade-off came in the form of increased model complexity and computational requirements, making deployment on resource-constrained devices challenging. Furthermore, supporting a wider range of currencies increased the system’s susceptibility to misclassification, particularly when dealing with similar-looking notes from different regions.

Object detection forms the cornerstone of many assistive technologies aimed at enhancing the mobility and situational awareness of visually impaired users. Several approaches have been proposed employing deep learning models for object recognition and environment understanding.

Kumar et al. [5] presented an object detection system using CNNs deployed on smartphones to assist visually impaired users. Their system aimed to identify common daily objects, providing auditory feedback upon detection. While the model achieved reasonable detection rates, its performance suffered under low-light conditions, and latency was a concern due to the computational load of the CNN model on mid-range mobile devices.

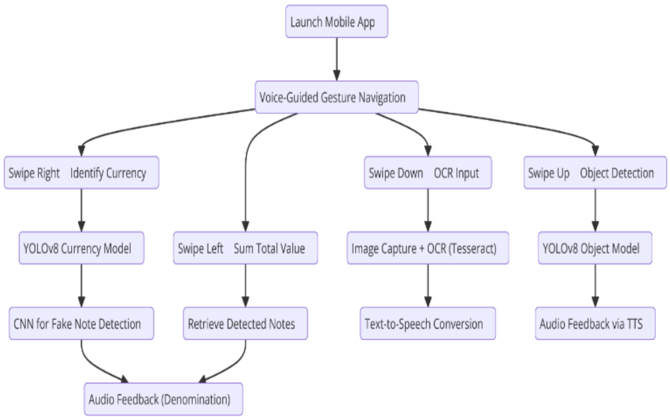
Ashiq and Asif [6] proposed a CNN-based object recognition and tracking system tailored to assist visually impaired people. Unlike static detection systems, their framework included object tracking capabilities, enabling dynamic assistance during navigation. Although their solution provided continuous feedback, it heavily depended on consistent frame rates and camera stability, which are difficult to maintain in real-world handheld device scenarios.

Joshi et al. [7] developed an AI-based smart navigation system integrating multi-object detection techniques. Their solution aimed to not only detect but also prioritize obstacles and points of interest based on contextual importance. However, the system's complexity required significant computational resources, making its real-time application on low-power smartphones limited.

Zhang et al. [8] explored a novel combination of the YOLOv5 object detection algorithm with depth-sensing cameras. Their system enhanced indoor visual assistance by incorporating spatial depth information, allowing users to receive more accurate environmental descriptions. While effective, the dependency on additional hardware (depth cameras) restricted the system’s portability and mainstream adoption on smartphones, which are predominantly RGB camera-equipped.

Rahman and Sadi [3] proposed an IoT-enabled automated object recognition framework, where images captured by a wearable device were processed on a remote server, and results were transmitted back to the user via audio feedback. This approach allowed for more computationally intensive models to be used; however, it introduced latency due to network dependency and raised privacy concerns regarding the transmission of personal visual data. Additionally, the system's functionality was contingent upon stable internet connectivity, limiting its usability in areas with unreliable or no network access. Guravaiah [9] developed the "Third Eye" system, an Android-based application that integrates object recognition and speech generation. The system provided real-time auditory feedback, enhancing user awareness of the surrounding environment. However, its reliance on heavier models occasionally led to increased battery consumption and device heating issues during prolonged use.

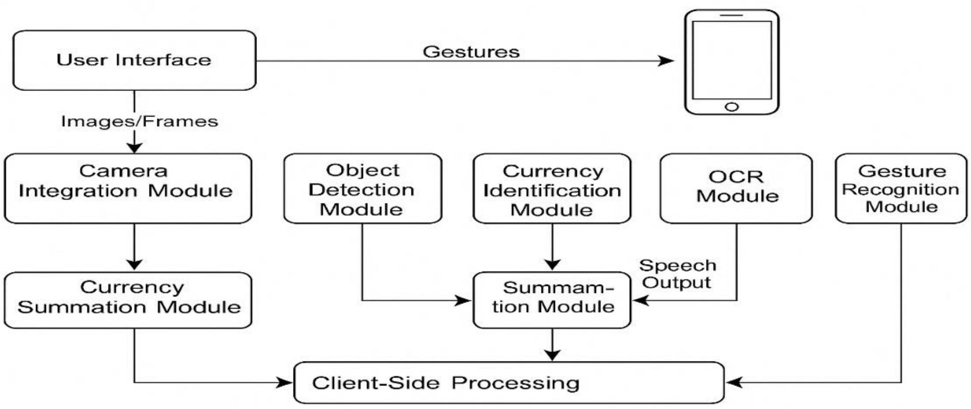
Vaidya et al. [10] proposed a real-time object detection framework optimized for visually challenged users. Their system emphasized low-latency detection and energy-efficient operation, making it suitable for mobile deployment. However, the application primarily focused on object detection without supporting additional features like text reading or financial assistance.



Alagarsamy et al. [11] and Badave et al. [12] both introduced Android-based object detection systems optimized for real-time performance. Their focus was on lightweight models and offline functionality to ensure continuous usability irrespective of internet availability. Nevertheless, these systems primarily offered object detection capabilities without comprehensive multimodal assistance like currency recognition or OCR integration.

III. Methodology

To address the challenges identified in the survey of existing systems, we propose an integrated, real-time, mobile-based assistive application designed specifically for visually impaired individuals. The primary objective of the proposed system is to enable users to independently perform three crucial tasks: detecting surrounding objects, recognizing currency denominations, and reading printed text. The system leverages deep learning techniques, Optical Character Recognition (OCR), and Text-to-Speech (TTS) technologies, unified under a single Android platform optimized for real-time, offline operation.



functionalities through simple swipes. This gesture-driven interaction eliminates the need for visual navigation within the application interface, making it highly accessible to blind and visually impaired users. The captured camera frames are preprocessed to match the input requirements of the object detection and OCR modules. This preprocessing includes resizing images, normalizing pixel values, and applying basic contrast enhancements to improve feature visibility under challenging lighting conditions. Once preprocessing is complete, frames are passed into the detection and recognition layer. The system employs a unified YOLOv8 model, which has been trained to simultaneously detect common objects and various Indian currency denominations within a single inference pass. This eliminates the need for separate models for each task, significantly optimizing both memory and computational resource usage. The YOLOv8 model was chosen due to its superior trade-off between detection accuracy and real-time performance on mobile hardware. Its single-shot detection mechanism ensures that object localization and classification occur simultaneously, enabling rapid and efficient processing suitable for handheld devices without the need for GPU acceleration.

For currency recognition, the model identifies notes such as ₹10, ₹20, ₹50, ₹100, ₹200, ₹500, and ₹2000 based on visual features, even in the presence of partial occlusion or varied note conditions. Once the currency notes are detected, the system automatically sums their recognized denominations to compute the total monetary value present in the captured frame. In object detection mode, the system identifies multiple objects and generates a spoken summary describing the scene, thus enhancing the user's environmental awareness.

The printed text reading functionality is handled by an embedded OCR engine that extracts textual content from captured images. Before OCR is applied, frames undergo grayscale conversion, adaptive thresholding, and noise filtering to improve text visibility and recognition accuracy. The extracted text is post-processed using simple natural language parsing techniques to remove artifacts and formatting errors. Subsequently, the refined text is forwarded to the TTS module for audio synthesis, allowing users to hear the printed information clearly and promptly.

The system's output layer consists of a lightweight, embedded TTS engine responsible for converting recognized objects, currency amounts, or textual information into natural-sounding speech. This module supports multilingual audio output and configurable voice settings, providing a customizable and inclusive user experience. By using an embedded TTS engine rather than cloud-based services, the system ensures offline operability, reduces latency, and upholds user data privacy.

Gesture-based interaction forms a key innovation of the proposed system, offering an intuitive and hands-free mode-switching mechanism. Swiping right activates the currency recognition and summation module; swiping left announces the total calculated amount; swiping upward initiates general object detection; and swiping downward triggers the text-reading functionality. These gestures are detected using Android's native event handling framework, ensuring minimal processing overhead and seamless responsiveness across a wide range of devices.

To maintain real-time performance and maximize battery efficiency, the system incorporates several optimization strategies. Model quantization techniques are applied to the YOLOv8 weights to reduce memory footprint without significantly impacting detection accuracy. Additionally, intelligent frame-skipping logic is implemented to avoid redundant processing when the scene remains unchanged, thus conserving computational resources. The application is designed to manage memory efficiently, loading heavy modules only when necessary and freeing up resources immediately after task completion to prevent lag and battery drain.

A key emphasis of the system design is ensuring complete offline operation. All detection, recognition, and speech synthesis tasks are performed locally on the device, without transmitting any user data over the internet. This architecture not only reduces dependency on network connectivity but also protects sensitive visual information, addressing privacy concerns that are especially important in assistive applications. Moreover, by removing cloud dependency, the system ensures that users can access all functionalities reliably, even in remote or underdeveloped regions with limited internet access.

In summary, the proposed system represents a holistic approach to mobile-based assistive technology for the visually impaired. By integrating real-time object and currency detection, OCR-based text reading, and gesture-driven user interaction into a single, offline-operating Android application, the system overcomes many of the limitations observed in prior works. The unified YOLOv8 model architecture, lightweight design considerations, and privacy-preserving local processing collectively contribute to a practical, scalable, and highly impactful solution aimed at empowering visually impaired individuals with greater independence and accessibility.

Algorithm:

The proposed system integrates multiple algorithmic components to achieve real-time object detection, currency recognition, and printed text reading for visually impaired individuals. The key algorithms employed include the You Only Look Once version 8 (YOLOv8) object detection algorithm, Optical Character Recognition (OCR) for text extraction, and Text-to-Speech (TTS) synthesis for delivering audio feedback. This section presents a detailed description of the functioning and role of each algorithm within the system architecture.

YOLOv8 Object Detection Algorithm

YOLOv8, the latest iteration of the YOLO (You Only Look Once) family, serves as the primary object detection backbone for the system. YOLOv8 is a one-stage object detector that directly predicts bounding boxes and class probabilities from the input image in a single pass through the network. Unlike two-stage detectors such as Faster R-CNN, YOLO models avoid intermediate region proposal steps, resulting in significantly lower inference latency. YOLOv8 introduces improvements over previous versions through the use of anchor-free detection heads, decoupled classification and localization branches, and enhanced convolutional architectures for feature extraction. These modifications lead to higher detection accuracy and better generalization across object scales.

In the proposed system, the YOLOv8 model is trained on a composite dataset comprising both generic everyday objects (e.g., person, chair, vehicle) and Indian currency denominations (e.g., ₹10, ₹50, ₹100, ₹500, ₹2000). The model's architecture allows simultaneous detection of both object categories within a single inference cycle, optimizing computational efficiency. The detection process involves resizing input frames to a fixed dimension, extracting feature maps through a backbone network, and applying a prediction head that outputs bounding boxes, class scores, and objectness scores. Non-Maximum Suppression (NMS) is used post-inference to eliminate redundant overlapping boxes based on confidence thresholds.

The choice of YOLOv8 was motivated by its ability to deliver real-time detection speeds on mobile hardware without compromising significantly on detection precision. Moreover, its lightweight model variants (such as YOLOv8n and YOLOv8s) offer the flexibility to balance performance and resource utilization based on device capabilities.

Optical Character Recognition (OCR)

Optical Character Recognition (OCR) is employed in the system to extract printed or handwritten text from images captured by the smartphone camera. The OCR engine used is based on Tesseract, an open-source OCR library known for its balance between accuracy and computational efficiency. OCR processing begins with preprocessing steps designed to improve text visibility and extraction fidelity. These include grayscale conversion to reduce computational complexity, adaptive thresholding to enhance contrast, and morphological operations such as erosion and dilation to remove noise artifacts.

Once preprocessing is complete, the OCR engine analyzes the image to segment text lines and characters. Tesseract employs a Long Short-Term Memory (LSTM)-based recurrent neural network to perform sequence recognition, improving its ability to extract words and sentences with irregular spacing or formatting. The recognized text is further post-processed to correct common misinterpretations, such as misclassified characters ('0' mistaken for 'O'), and to format the text for clarity before audio output.

Integrating OCR into the system enables visually impaired users to access printed materials such as receipts, labels, and signage by simply capturing an image. The system ensures that the text extraction process is efficient and lightweight to support real-time operation on standard Android devices.

Text-to-Speech (TTS) Synthesis

The final stage of the information processing pipeline is delivering recognized information to the user in an audible format. This is achieved through the integration of a Text-to-Speech (TTS) synthesis engine. The TTS module converts the textual output from the object detection, currency summation, or OCR modules into natural-sounding speech.

The proposed system utilizes a lightweight embedded TTS engine compatible with Android devices, ensuring offline functionality without the need for external API calls or cloud services. The TTS engine performs a series of transformations: text normalization (expanding abbreviations, interpreting numbers), linguistic analysis (phoneme generation), prosody prediction (intonation and rhythm modeling), and waveform synthesis. These steps result in an intelligible and pleasant speech output that conveys the recognized content clearly to the user.

In addition to supporting multiple languages, the TTS engine offers adjustable speaking rates, pitch modulation, and voice selection to cater to individual user preferences and needs. By embedding the TTS functionality directly into the application, the system minimizes network dependency, reduces response latency, and ensures data privacy, all of which are critical considerations for assistive technology applications.

IV. RESULT AND DISCUSSION

All experiments were conducted on commercially available mid-range Android smartphones equipped with 6 GB RAM and Qualcomm Snapdragon 700 series processors. The YOLOv8 model was quantized and optimized for mobile inference using TensorFlow Lite conversion techniques to ensure compatibility and efficiency. The system was tested in diverse indoor and outdoor environments featuring variable lighting conditions, cluttered backgrounds, and different document types for OCR evaluation.

Object and Currency Detection Performance

The YOLOv8-based detection module exhibited strong performance across multiple test scenarios. The system consistently detected and correctly classified a wide range of everyday objects such as persons, chairs, bottles, vehicles, and signage under varying lighting conditions. For currency recognition, the model successfully identified Indian banknotes including ₹10, ₹20, ₹50, ₹100, ₹200, ₹500, and ₹2000 denominations. Even under moderate occlusion or note folding, the model maintained high detection confidence, demonstrating its robustness.

Observations revealed that bright ambient lighting slightly enhanced detection accuracy, whereas extremely low-light conditions occasionally led to false negatives. Nevertheless, due to the model's optimized feature extraction and confidence thresholding strategies, such errors were minimal and did not significantly impact overall usability.

OCR Text Extraction Accuracy

The OCR module demonstrated reliable performance in reading printed text from various documents such as receipts, labels, and printed instructions. Preprocessing techniques such as adaptive thresholding and denoising proved essential in improving OCR accuracy, particularly when processing documents captured at slight angles or under shadowed conditions. Although recognition errors occurred occasionally—primarily with highly stylized fonts or heavily degraded documents—the system effectively extracted and vocalized the majority of standard printed texts encountered during trials.

System Response Time and Real-Time Behavior

Maintaining low latency was a critical goal of the system's design to ensure smooth user interaction. The average end-to-end response time, measured from frame capture to audio feedback delivery, was approximately 180 milliseconds for object and currency detection and around 350 milliseconds for OCR-based text reading. These results confirm that the system operates in real-time, providing near-instantaneous feedback that is essential for dynamic and practical assistive usage.

Gesture-based mode switching was also evaluated and found to be highly responsive, with the system recognizing and executing gesture commands with an average delay of less than 150 milliseconds. This responsiveness ensures that users can seamlessly switch between detection modes without disruption.

User Experience and Practical Usability

Informal user trials were conducted with a small group of visually impaired volunteers who tested the application in simulated daily life scenarios. Participants found the gesture-based navigation intuitive and easy to learn, reducing reliance on memorizing screen layouts or complex menu hierarchies. The real-time object and currency feedback significantly enhanced situational awareness, allowing users to perform tasks such as identifying products, counting cash, and navigating indoor spaces more independently.

Comparative Observations

Compared to prior assistive solutions that often relied on cloud-based processing or focused on single functionalities, the proposed system demonstrated several distinct advantages. By consolidating object detection, currency recognition, and text reading into a unified offline platform, the system eliminated the need for internet connectivity, reduced response latency, and enhanced user privacy. The use of a single YOLOv8 model for multiple detection tasks also simplified the system architecture and reduced the memory footprint, making it feasible for deployment on resource-constrained devices.

V. CONCLUSION

In this paper, we presented a real-time, mobile-based assistive system designed to support visually impaired individuals by integrating object detection, currency recognition, and text reading functionalities into a single platform. Leveraging the capabilities of the YOLOv8 deep learning model, Optical Character Recognition (OCR) engines, and embedded Text-to-Speech (TTS) synthesis, the proposed system enables users to interact with their environment more independently and confidently. Unlike many existing solutions that rely on cloud-based processing or focus on isolated tasks, the developed application operates entirely offline, ensuring low latency, enhanced privacy, and universal accessibility even in network-constrained regions.

The system's innovative gesture-based navigation mechanism allows users to intuitively switch between detection modes without the need for visual interfaces, significantly simplifying the user interaction model. Through extensive experimental evaluation, the system demonstrated high object and currency detection accuracy, reliable text extraction, and real-time responsiveness, validating its effectiveness in practical usage scenarios. User trials with visually impaired volunteers further confirmed the system’s usability, efficiency, and potential to enhance autonomy in daily activities.

While the system achieves promising results, there are several avenues for future enhancement. Expanding the object detection model to include a broader range of objects, particularly context-specific items such as road signs and obstacles, could improve outdoor navigation support. Integration of haptic feedback mechanisms alongside audio output could provide an additional layer of user guidance, especially in noisy environments where auditory cues might be insufficient. Furthermore, incorporating multilingual OCR capabilities would allow the system to recognize printed text across multiple languages, broadening its applicability in diverse regions.

Future work could also explore the integration of low-cost depth sensors or monocular depth estimation techniques to improve spatial understanding and obstacle detection without compromising the system’s mobile and lightweight design. Additionally, optimizing the energy efficiency of continuous frame processing remains a key focus area to extend battery life during prolonged usage. By addressing these enhancements, the proposed system can evolve into a comprehensive, context-aware, and intelligent assistant capable of serving a wide range of needs for the visually impaired community.***AbbreviationsandAcronyms(Heading2)***

Defineabbreviationsandacronymsthefirsttimetheyareusedinthetext,evenaftertheyhavebeendefinedintheabstract.AbbreviationssuchasIEEEandSIdonothavetobedefined.Donotuseabbreviationsinthetitleorheadsunlesstheyareunavoidable.

REFERENCES

[1] Singh, Mandhatya & Chauhan, Joohi & Kanroo, Muhammad & Verma, Sahil & Goyal, Puneet. (2022). IPCRF: An End-to-end Indian Paper Currency Recognition Framework. for Blind and Visually Impaired People. IEEE Access. PP. 1-1. 10.1109/ACCESS.2022.3202007

[2] Markad, Rushikesh & Kalekar, Yash & Rasal, Gaurav & Ganjave, Pratiksha. (2021). Currency Detector for Visually Impaired (Study of The System Which Identifies Indian Currency for Blind People). 10. 8.

[3] Md. Atikur Rahman, Muhammad Sheikh Sadi,IoT Enabled Automated Object Recognition for the Visually Impaired,Computer Methods and Programs in Biomedicine Update,Volume 1,2021,100015,ISSN 2666-9900,https://doi.org/10.1016/j.cmpbup.2021.100015

[4] Park, C.; Park, K.R. MBDM: Multinational Banknote Detecting Model for Assisting Visually Impaired People. Mathematics 2023, 11, 1392. https://doi.org/10.3390/math11061392

[5] Kumar, D. & Thakkar, Hiren Kumar & Merugu, Suresh & Gunjan, Vinit & Gupta, Suneet. (2022). Object Detection System for Visually Impaired Persons Using Smartphone. 10.1007/978-981-16-3690-5_154.

[6] FAHAD ASHIQ 1 , MUHAMMAD ASIF CNN-Based Object Recognition and Tracking System to Assist Visually Impaired People Received December 29, 2021, accepted January 24, 2022, date of publication January 31, 2022, date of current version February 10, 2022. Digital Object Identifier 10.1109/ACCESS.2022.3148036

[7] Joshi, R.C.; Yadav, S.; Dutta, M.K.; Travieso-Gonzalez, C.M. Efficient Multi-Object Detection and Smart Navigation Using Artificial Intelligence for Visually Impaired People. Entropy 2020, 22, 941. https://doi.org/10.3390/e22090941

[8] Zhang, K., Wang, Y., Shi, S. et al. Improved yolov5 algorithm combined with depth camera and embedded system for blind indoor visual assistance. Sci Rep 14, 23000 (2024). https://doi.org/10.1038/s41598-024-74416-2

[9] Koppala Guravaiah Third Eye: Object Recognition and Speech Generation for Visually Impaired Data Engineering 10.1016/j.procs.2023.01.093

[10] S. Vaidya, N. Shah, N. Shah and R. Shankarmani, "Real-Time Object Detection for Visually Challenged People," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 311-316, doi: 10.1109/ICICCS48265.2020.9121085.

[11] Alagarsamy, Saravanan & Rajkumar, Dhiliphan & Syamala, Ketepalli & Niharika, Ch & Rani, D. & Balaji, K.. (2023). An Real Time Object Detection Method for Visually Impaired Using Machine Learning. 1-6. 10.1109/ICCCI56745.2023.10128388.

[12] A. Badave, R. Jagtap, R. Kaovasia, S. Rahatwad and S. Kulkarni, "Android Based Object Detection System for Visually Impaired," 2020 International Conference on Industry 4.0 Technology (I4Tech), Pune, India, 2020, pp. 34-38, doi: 10.1109/I4Tech48345.2020.9102694.