International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Product identification and text recognition for blind peoples using deep learning algorithm

Mr. NOOR AHAMED J[1] M.C.A., M.Phil., SARAVANA KUMAR.S[2]

1Assistant Professor (SG), Department of Computer Applications, Nehru college of management, Coimbatore, Tamil Nadu, India.

jnamca@gmail.com

2II MCA, Department of Computer Applications, Nehru college of management, Coimbatore, Tamil Nadu, India.

ABSTRACT:

Blind and visually impaired individuals face significant challenges in identifying products and reading text in everyday life. This project proposes a deep learning-based system that enables product identification and text recognition to assist blind people in their daily activities. The system utilizes Convolutional Neural Networks (CNNs) and Optical Character Recognition (OCR) techniques to recognize products and extract text from labels, packaging, and documents. A pre-trained CNN model (such as MobileNet, ResNet, or EfficientNet) is employed for product classification, while Tesseract OCR or deep learning-based EAST (Efficient and Accurate Scene Text Detection) is used for text recognition. The system can be integrated into a mobile application or a wearable device equipped with a camera to capture images of products and text .Once the product is identified or text is extracted, the system converts the information into speech output using Text-to-Speech (TTS) technology, enabling blind users to receive real-time audio feedback. The model is trained on a diverse dataset of product images and text samples to enhance accuracy in different environments, including supermarkets, kitchens, and workspaces. This deep learning-powered assistive technology aims to enhance independence and accessibility for visually impaired individuals, reducing their reliance on external help for product identification and text reading. The proposed system provides a fast, efficient, and user-friendly solution to improve the quality of life for blind people.

KEYWORDS:

Deep Learning, Product Identification, Text Recognition, Blind Assistance, Convolutional Neural Networks, OCR, Text-to-Speech

1.INTRODUCTION:

Blind and visually impaired individuals face numerous challenges in their daily lives, especially in recognizing products and reading text from labels, packaging, and documents. Tasks such as identifying grocery items, distinguishing between similar-looking objects, and reading instructions become difficult without external assistance. To address these challenges, advancements in deep learning and computer vision offer promising solutions by enabling real-time product identification and text recognition.This project aims to develop an AI-powered assistive system that helps blind individuals identify products and read text using deep learning algorithms. The system utilizes a Convolutional Neural Network (CNN) for product classification and Optical Character Recognition (OCR) for text extraction. By integrating these technologies with Text-to-Speech (TTS) conversion, the system provides audio feedback, allowing blind users to understand product details and read important text.The proposed system operates by capturing images of objects or product labels through a smartphone camera or a wearable device. The deep learning model processes the image, identifies the product, and extracts relevant text. The recognized information is then converted into speech output, enabling visually impaired users to interact with their environment more independently.This project has the potential to significantly enhance accessibility, allowing blind individuals to perform essential daily tasks without relying on others. By leveraging deep learning and artificial intelligence, the system aims to improve the quality of life for visually impaired users, fostering greater independence and inclusion in society.

2. RELATED WORKS:

Several research studies and technological advancements have been made in the field of assistive technologies for the visually impaired, particularly in product identification and text recognition. Various approaches, including deep learning, computer vision, and optical character recognition (OCR), have been explored to enhance accessibility for blind individuals. Existing solutions such as Microsoft's Seeing AI, Google Lookout, and OrCam MyEye utilize Convolutional Neural Networks (CNNs) and OCR models to assist users in identifying objects and reading text aloud. These applications leverage pre-trained deep learning models for object detection, text extraction, and speech synthesis, enabling blind users to navigate their environment more independently.Product identification has been widely explored using CNN-based models such as MobileNet, ResNet, and EfficientNet,which classify objects based on image recognition. Object detection models like YOLO (You Only Look Once) and SSD (Single Shot Multibox Detector) have also been employed in real-time assistive applications to recognize multiple products simultaneously. Some research works have explored RFID and barcode-based product identification, but these methods require physical modifications such as barcode scanning or RFID tags, limiting their scalability in real-world applications. Deep learning-based object recognition eliminates the need for such modifications, making it a more practical solution.For text recognition, Tesseract OCR, Google Vision API, and deep learning-based EAST (Efficient and Accurate Scene Text Detector) have been widely used. While traditional OCR methods struggle with distorted, low-contrast, or handwritten text, deep learning-based models offer improved accuracy in recognizing complex text structures. Recent studies have combined OCR with Natural Language Processing (NLP) to enhance text readability and correct errors in extracted text. However, many existing text recognition solutions require internet connectivity for cloud-based processing, making them less effective in offline environments.Additionally, assistive technologies integrate Text-to-Speech (TTS) engines, such as Google TTS, Amazon Polly, and Microsoft Azure Speech, to convert extracted text into audio feedback. While commercial solutions like OrCam MyEye provide real-time text reading, their high cost makes them inaccessible to many users. In contrast, mobile applications such as Seeing AI and Google Lookout offer free solutions but may lack offline capabilities or accurate product identification in complex environments.Despite advancements in assistive technology, existing solutions still face limitations, including high computational costs, dependency on internet connectivity, difficulties in recognizing cluttered objects, and inaccurate OCR performance under challenging conditions. This project aims to

address these gaps by developing a deep learning-powered system that integrates product recognition, text extraction, and real-

learning models optimized for real-time processing, this solution enhances accessibility for blind individuals and empowers them to navigate their daily lives with greater independence.

3. EXISTING SYSTEM:

The current assistive technologies available for blind individuals rely on image recognition, optical character recognition (OCR), and text-to-speech (TTS) systems to help with product identification and text recognition. Applications such as Microsoft Seeing AI, Google Lookout, and TapTapSee use deep learning-based object detection to identify products and read text aloud. However, these solutions often face challenges in complex environments, poor lighting conditions, and cluttered backgrounds, which reduce their accuracy. Some systems, like OrCam MyEye, provide real-time recognition through wearable devices but are expensive and inaccessible to many users.For text recognition, most existing systems utilize Tesseract OCR, Google Vision API, or EAST (Efficient and Accurate Scene Text Detector). While these technologies can extract printed text effectively, they struggle with distorted fonts, curved surfaces, and handwritten text. Many OCR-based solutions require internet connectivity for cloud processing, making them unreliable in offline scenarios. Additionally, some product identification systems depend on barcode scanning or RFID- based methods, which require manual interaction and are limited to products with pre-registered labels.Despite the advancements in AI-driven assistive tools, existing systems still have limitations in accuracy, real-time processing, affordability, and ease of use. Many applications either lack offline functionality, require high-end hardware, or have slow recognition speeds, making them inconvenient for visually impaired individuals.time speech synthesis into a user-friendly and cost-effective mobile or wearable application. By leveraging lightweight deep

These limitations highlight the need for a more efficient, offline-capable, and cost-effective solution that integrates deep learning-based product identification and text recognition with real-time speech output. The proposed system aims to enhance accuracy, improve accessibility, and provide a seamless user experience for blind individuals.

4.PROPOSTED SYSTEM:

The proposed system aims to develop a deep learning-based product identification and text recognition system to assist visually impaired individuals in their daily lives. Unlike existing solutions that rely on barcode scanning, RFID tags, or cloud-based OCR, this system leverages real-time computer vision and AI-driven text recognition to provide an efficient and offline-capable assistive tool. The system will use Convolutional Neural Networks (CNNs), such as MobileNet, ResNet, or EfficientNet, for product identification, ensuring fast and accurate recognition even in cluttered backgrounds and varying lighting conditions. For text recognition, the system will integrate deep learning-based OCR techniques, such as EAST (Efficient and Accurate Scene Text Detector) and CRNN (Convolutional Recurrent Neural Network), to accurately extract printed and handwritten text from product labels, packaging, and documents.Once the product is identified or text is extracted, the system will use Text-to-Speech (TTS) conversion with AI- based speech synthesis models to provide real- time audio feedback to the user. The solution will be designed to work offline, eliminating the need for continuous internet connectivity, which is a major limitation in existing cloud- dependent assistive tools. Additionally, the system will feature a user-friendly interface with voice commands and touch-based navigation, making it accessible for blind users. The application can be implemented on

smartphones, wearable devices, or embedded systems, ensuring portability and ease of use.This proposed system aims to overcome the limitations of existing solutions by offerina cost-effective, highly accurate, and real-time assistive technology for blind individuals. By integrating advanced deep learning, and speech synthesis, this project enhances accessibility and independence, enabling visually impaired users to identify products and read text effortlessly in any environment.

5.DATA FLOW DIAGRAM:

A data flow diagram shows the way information flows through a process or system. It includes data inputs and outputs, data stores, and the various sub processes the data moves through. DFDs are built using standardized symbols and notation to describe various entities and their relationships. Data flow diagrams visually represent systems and processes that would be hard to describe in a chunk of text. You can use these diagrams to map out an existing system and make it better or to plan out a new system for implementation. Visualizing each element makes it easy to identify inefficiencies and produce the best possible system.

5.1 LEVEL-0:

The Level 0 DFD shows how the system is divided into 'sub-systems' (processes), each of which deals with one or more of the data flows to or from an external agent, and which together provide all of the functionality of the system as a whole. It also identifies internal data stores that must be present in order for the system to do its job, and shows the flow of data between the various parts of the system.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

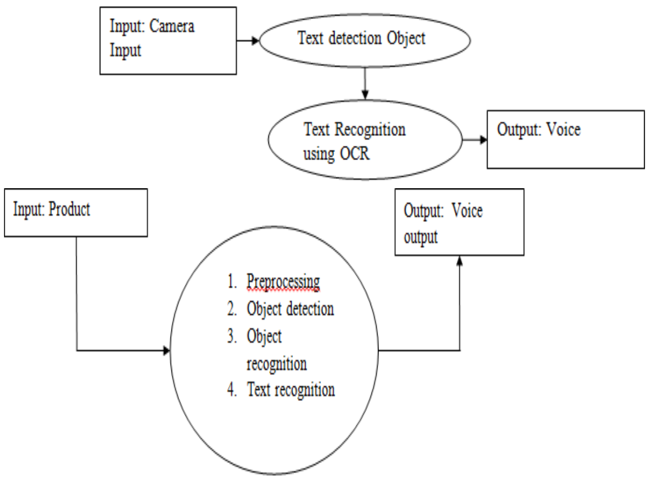


Figure 5.1.1 Data Flow Diagram

5.2 LEVEL-1

The next stage is to create the Level 1 Data Flow Diagram. This highlights the main functions carried out by the system. As a rule, to describe the system was using between two and seven functions - two being a simple system and seven being a complicated system. This enables us to keep the model manageable on screen or paper

5.3 LEVEL-2

A Data Flow Diagram (DFD) tracks processes and their data paths within the business or system boundary under investigation. A DFD defines each domain boundary and illustrates the logical movement and transformation of data within the defined boundary. The diagram shows 'what' input data enters the domain, 'what' logical processes the domain applies to that data, and 'what' output da

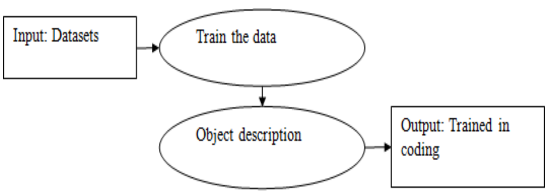


Figure 5.3.1 Data Flow Diagram

ta leaves the domain. Essentially, a DFD is a tool for process modeling and one of the oldest.

5.4 LEVEL-3

A data flow diagram (DFD) is a graphical representation of the flow of data through an information system. A DFD shows the flow of data from data sources and data stores to processes and from processes to data stores and data sinks. DFDs are used for modelling and analyzing the flow of data in data processing systems, and are usually accompanied by a data dictionary, an entity- relationship model, and a number of process descriptions

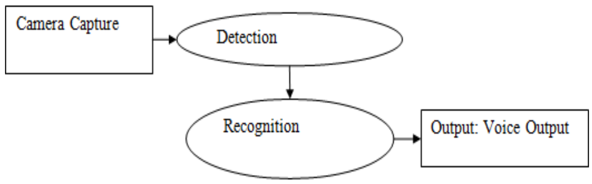
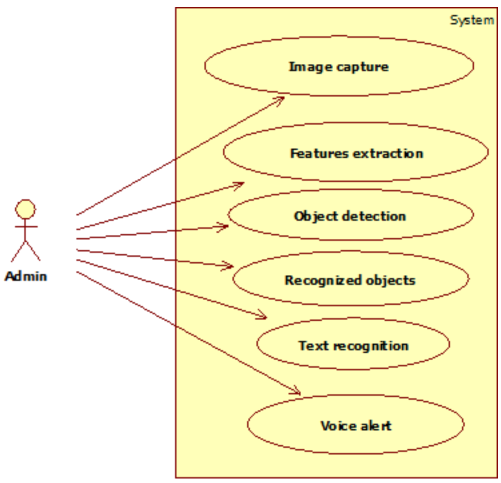


Figure 5.4.1 Data Flow Diagram

6. USE CASE DIAGRAM:

In its most basic form, a use case diagram is a depiction of a user's interaction with the system that illustrates the connection between the user and the many use cases that the user is involved in. A "system" in this sense refers to something that is being created

Figure 6.1 Use Case Diagram



7. CONCLUSION:

The Product Identification and Text Recognition System using deep learning algorithms is a significant step toward improving accessibility for blind and visually impaired individuals. By leveraging computer vision, OCR (Optical Character Recognition), and text-to-speech (TTS) technologies, this system enables users to identify products and read text from labels, packaging.

8. REFERENCES:

[1] Ten World Health Organization facts regarding blindness and visual impairment

Image Processing, natural scenes pages. 87-99, Vol. 13, No. 1, 2004.

[6] N. G. Bourbakis and D. Dakopoulos, Wearable obstacle A survey on blind people's electronic travel aids for avoidance, IEEE Systems, Man, and Cybernetics Transactions, Vol. 40, No. 1, pp. 25-35, 2010.

Impairment, 2009.

[2] National Agency for Advance Data Reports Survey of Health Interviews, 2008.

[3] International Camera-Based Workshop Recognition and Analysis of Documents (CBDAR 2005, 2007,2009, 2011) .

[4] A. L. Yuille and X. Chen, Identifying and reading a material in a natural setting, Vol. 2, pp., CVPR.II-366 II-373, 2004.

[5] A. Waibel, J. Zhang, J. Yang, and X. Chen Automatic identification and detection of indicators from In IEEE Transactions on

[7] Y. Wexler, E. Ofek, and B. Epshtein, Identifying text with a change in stroke width in natural scenes, In CVPR,pp. 2963-2970, 2010.

[8] R. Schapire and Y. Freund, Tests with A novel method for boosting, Int. Conf. on Machine Learning, 1996, pp. 148156.