INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 04 | APRIL - 2024 SJIF RATING: 8.448 ISSN: 2582-3930

CNN-Based Object Recognition and Tracking System to Assist Visually Impaired People

Karthiga C
Department of Computer Science and Engineering,
Sethu Institute of Technology Virudhunagar, India.
karthigach@gmail.com

Radhakrishnan S
Department of Computer Science and Engineering,
Sethu Institute of Technology Virudhunagar, India.
radhakrish@sethu.ac.in

***Abstract*—Visually impaired peoples (VIPs) are a significant population segment worldwide. A system has been developed to aid mobility and ensure their safety. The system uses artificial voice for real-time navigation, allowing VIPs to perceive their surroundings. A web application is also created for their safety, allowing them to share their location with family. The project focuses on object recognition, converting detected object names (such as person1, person2, Dog, cat) into audio for easier understanding.**

***Keywords*— Visually impaired peoples (VIPs), artificial voice, object recognition, object name, audio**

I. INTRODUCTION

Visually impaired peoples (VIPs) are a significant segment of the population, and technology has significantly improved their daily lives. A system has been developed to aid mobility and ensure their safety, using an artificial voice for real-time navigation. The system also includes a web application for their safety, which can be shared by the user while preserving privacy. The project focuses on object recognition for VIPs, identifying objects and converting detected object names into audio for easier understanding. Image processing is a rapidly expanding technology that involves importing images, analyzing and altering them, and producing reports or changed images. Digital image processing involves pre-processing, augmentation, presentation, and information extraction. Object detection algorithms combine image classification and object localization, producing bounding boxes with class labels. These algorithms can handle multi-class classification and multiple occurrences of objects. However, they cannot accurately estimate measurements such as the area or perimeter of an object. Fig. 1 depicts the image for object detection.

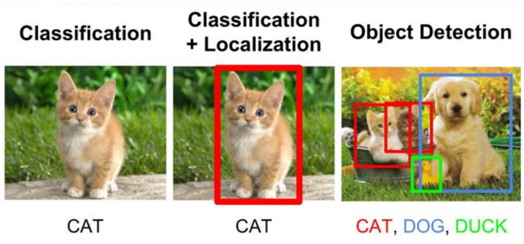

Fig.1. Object detection image

A. Project objective

The study aims to develop a real-time visual surveillance system with object detection, classification, tracking, activity analysis, and recognition for visually impaired individuals using CNN-LSTM for feature selection.

II. LITERATURE REVIEW

This study introduces a novel application of the combined multi-objective tracking technique using graph neural networks, which is crucial for detecting and analyzing unique characteristics in spatial and temporal domains [1]. The study introduces a system that integrates object detection tracking algorithms with low-end computing for real-time video surveillance, utilizing deep learning techniques like region-based convolutional networks for improved object identification accuracy [2]. The study proposes an automated assistive device using artificial intelligence to enhance the understanding of the visually impaired by identifying and providing real-time audio inputs [3]. The study aims to create a novel indoor item identification system using the deep convolutional neural network, RetinaNet, as its foundation [4]. The study presents a system that uses two algorithms, you only look once (Yolo) and Yolo_v3, to identify daily items and generate voice alarms, assessing accuracy and performance [5]. Information technology has significantly impacted data management, particularly in data acquisition, processing, and prediction. Interdisciplinary efforts, including image processing and AI-based models, have enabled computational experiments. This paper studies object detection methodologies, analyzing their strengths and weaknesses, and highlighting their potential for prediction and future scope [6]. The paper explores the importance of part elements in scene images through data processing. It uses a top-down tree structure with annotations and bounding boxes to represent objects in images. The LSTM network is used to train the representation. The object detection method is efficient and accurate, with improved accuracy using Google open images datasets. However, it has limitations, such as limited datasets and a focus on prediction [7].

III. SYSTEM PROPOSAL

A. Existing system

The existing system uses traditional feature extraction for object recognition, resulting in time-consuming feature selection and classification. The disadvantages are model building and fitting time is high, with more data loss compared to conventional methods. Preprocessing is complicated, response time is low, accuracy is low, and data loss is high

B. Proposed system

The proposed solution involves developing and training a specialized Convolutional Neural Network (CNN) architecture addressing challenges in low-light conditions for object detection. This system can accurately detect humans in diverse scenarios, contributing to surveillance, safety, and robotics. Traditional feature extraction techniques currently take time for feature selection, enhancing automated object detection capabilities. This system offers limited time taken, easy data retrieval, and low data loss during pre-processing, and higher accuracy compared to existing systems. Fig. 2 represents the architecture figure which depicts the sequence of suggested technique.

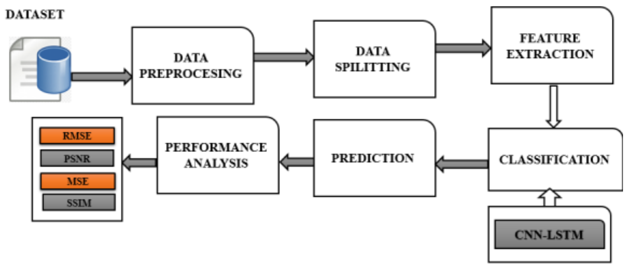

Fig.2. Architecture figure

A flow diagram illustrates a sequence of occurrences, system movements, and choices. Despite of the procedures degree of complexity, they provide a thorough explanation of each phase. Fig. 3 represents the flow drawing.

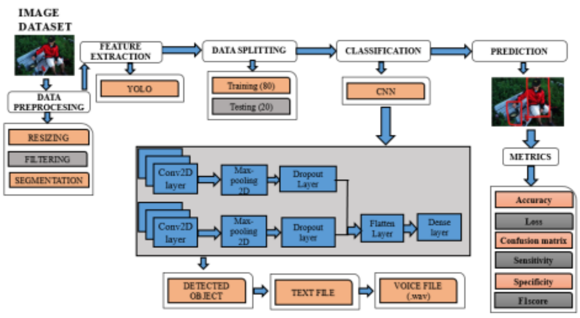

Fig.3. Flow drawing

A series figure represents a large number of continuously arising objects or processes in parallel lines that are vertical and the communications that are sent between them in the form of horizontal arrows, all in the chronological sequence. Fig. 4 represents the series figure.

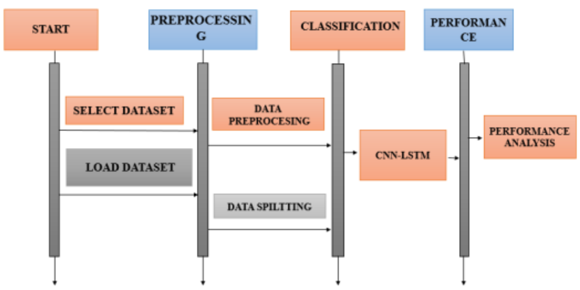

Fig.4. Series figure

C. Data collection and preprocessing

An image based dataset is used in the object detection and tracking system. Preprocessing involves importing the images, scaling them to a desired size (e.g., 224x224 pixels), and then using OpenCV (GaussianBlur) to apply Gaussian blur for preprocessing the data.

D. Feature extraction

The design of the YOLO model performs feature extraction automatically. YOLO is a deep learning system that uses its CNN layers to conduct feature extraction as part of its object detection tasks. During the YOLO model's forward pass, features related to object existence, locations, dimensions, and classes are obtained.

E. Classification using CNN-LSTM

The suggested CNN-LSTM based object detection system has the capability to reliably and effectively identify people in various settings, hence supporting fields such as automation, security, and monitoring. This method can increase the capacity of automated object recognition and boost overall image understanding by utilizing CNN and LSTM.

IV. SYSTEM REQUIREMENTS

A. Hardware requirements

The system includes a Pentium IV 2.4 GHz processor, 200GB hard disk, Logitech mouse, enhanced keyboard with 110 keys, and 4GB RAM.

B. Software requirements

The Python-based tool, Spyder 3.9, is compatible with Windows 7 or Windows 10, with front-end Anaconda Navigator – Spyder and back-end Anaconda Navigator – Spyder Console.

C. Software description

Spyder is an accessible Python Integrated Development Environment for scientific coding, published using the MIT authorization. Pierre Raybaut established it in 2009. It integrates with popular packages like pandas, IPython, SymPy, Cython, NumPy, SciPy, and Matplotlib. Spyder is extensible with plugins, encourages the use of data-driven tools for inspection, and includes Python-specific tools for script inspection. It is available on various platforms and uses Qt for its GUI.

The Spyder IDE is a powerful tool for Python development, offering syntax highlighting, introspection, and code completion. It supports multiple IPython consoles, allows variable exploration, provides rich text documentation, and supports static code analysis. It also features a run-time profiler for benchmarking code, project support, a built-in file explorer, a "Find in Files" feature, an online help browser, and an internal console for introspection.

V. INSTALLATION PROCESS

A. Python Spyder IDE Installation

The Python Spyder IDE is a default implementation with Anaconda Python distribution, available for easy installation via the official Anaconda website which is shown in Fig. 5 (https://www.anaconda.com).

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 04 | APRIL - 2024 SJIF RATING: 8.448 ISSN: 2582-3930

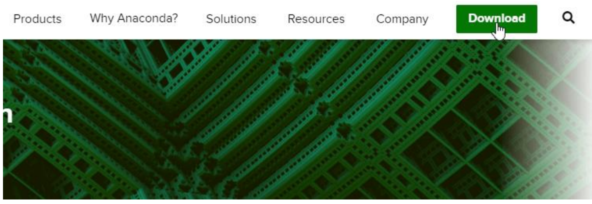

Fig.5. Installation page

Select the version that works best for your OS and click Download After downloading and finishing the setup, click Finish. Open Spyder and type Anaconda Navigator into the system search box. When it launches, a screen like that is shown in Fig. 6.

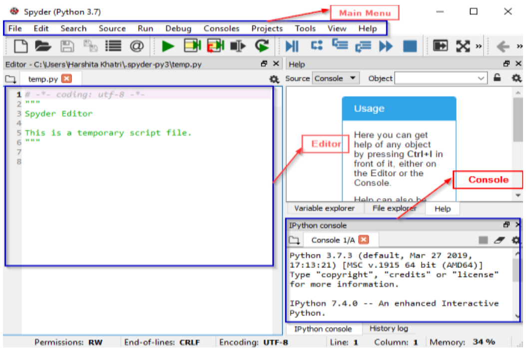

Fig.6. Visualization of navigators

To create a new file or start a project, navigation for the file are File–>New File.

B. Writing the code

Spyder's multi-language code editor offers powerful tools like structural emphasis, quick evaluation, style analysis, and instant fulfillment. It provides a well-organized dispatch stacks for techniques, offering every potential point of assertion. Take a look at the example shown as Fig. 7.

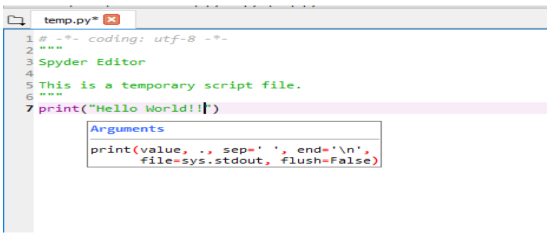

Fig.7. Coding page

The editor displays the complete syntax of the print function and alerts users of any errors precede the line quantity, giving a thorough explanation of the problem. Check out the Fig. 8.

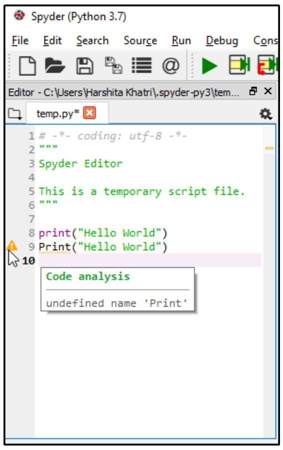

Fig.8. Error alert

To execute a folder, click on the Run decision and the output will appear on the Console as shown in Fig. 9.

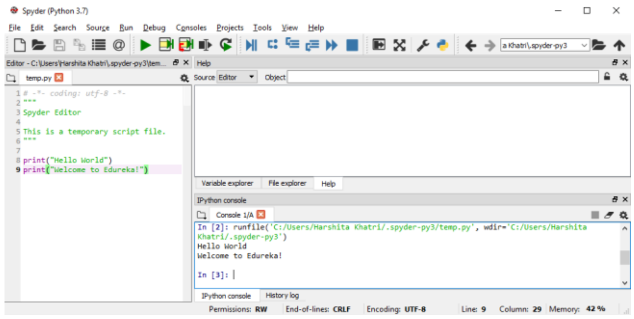

Fig.9. Output page

C. Testing of product

System testing is the implementation stage of a system, ensuring its accuracy and efficiency before live operation. It involves executing programs to find errors, and candidates undergo various tests, including online response, Volume Street, recovery, security, and usability tests. Unit testing integrates modules, ensuring each model works according to expected output. Integration testing evaluates system performance, addressing data loss and module effects. White box testing ensures independent paths are exercised, while black box testing fixes errors in applications. Validation testing ensures software functions reasonably expected by the customer. User acceptance testing is crucial for system success. Output testing confirms the correct screen format.

D. System Implementation

Software implementation includes the last install of the downloaded file in its actual setting to guarantee system functionality and user satisfaction. It is crucial for users to be aware of the benefits, build faith in the program, and receive appropriate direction to be comfortable using the app. For the suggested system to fulfill its goals and provide the desired results, user training is necessary. Education complements training by providing background information, creating an atmosphere, and motivating user staff. Users must be trained on the new application software, including screen flow, design, help, errors, validation checks, and data correction methods. Operational documentation is also crucial to make the system user-friendly and familiar.

E. System maintenance

For a system to respond to developments to its environment, such as social, technological, and environmental changes, the maintenance portion of the software cycle is essential. The system can adapt to these changes with the help of proper maintenance methods. Software maintenance includes corrective maintenance, adaptive maintenance, perceptive maintenance, and preventive maintenance. Corrective maintenance involves diagnosing and correcting errors, while adaptive maintenance modifies software to interfere with rapid changes in computing. Perceptive maintenance addresses user recommendations for new capabilities and functions, while preventive maintenance involves modifications to increase dependability or maintainability in the future, frequently using re- and reverse-engineering methods.

F. Feasibility study

A feasibility study evaluates the worth of implementing a proposed system, focusing on three sections: Economic, Technical, and Behavioral. Economic analysis determines the benefits and savings expected from the system, while Technical feasibility examines the hardware and software required for system development. Behavioral feasibility considers the resistance to change and the need for adequate training, while Behavioral feasibility considers the system's ability to generate daily reports at user request.

VI. RESULTS

A. Input image

A file upload component is usually used by the user to upload the input image. After an image is uploaded, the Image module opens and displays it (in JPG, JPEG, or PNG formats). Utilizing the YOLO (You Only Look Once) object identification approach, predictions are made based on this input image. Fig. 10 depicts the input image.

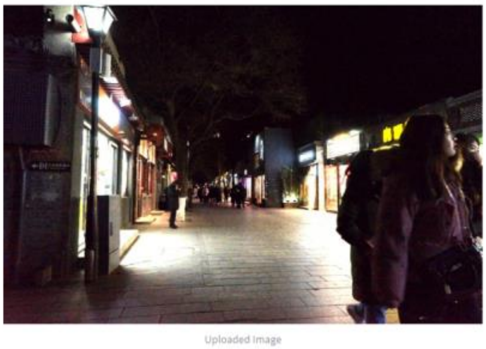

Fig. 10. Input image

B. Object detected image

The object-detected image is shown once the uploaded image has been subjected to the YOLO object detection model,. Bounding boxes are used to emphasize the identified items, and the anticipated class of each object is indicated on the box. Fig. 11 depicts the detected image.

Fig.11. Object detected Image

VII. CONCLUSION

The study proposes using a CNN model for human identification in low vision conditions, deploying it on a small website, and using the YOLO-V8x model for object classification, enhancing accuracy through additional layers. The project's main goal is accuracy and time, but the preprocessing method has some limitations, despite being user-friendly and straightforward. The project will be enhanced with videos, future IoT security software testing, and a camera for low vision and nighttime human identification.

REFERENCES

REFERENCES

[1] Y. Wang, K. Kitani, and X. Weng, "Joint object detection and multi-object tracking with graph neural networks", In 2021 IEEE International Conference on Robotics and Automation (ICRA) 2021, (pp. 13708-13715). IEEE.

[2] S. Jha, C. Seo, E. Yang, and G.P. Joshi, "Real time object detection and trackingsystem for video surveillance system", Multimedia Tools and Applications, 80(3), 2021, pp.3981-3996.

[3] R.C. Joshi, S. Yadav, M.K. Dutta and C.M, "Travieso-Gonzalez, Efficient multi-object detection and smart navigation using artificial intelligence for visually impaired people", Entropy, 22(9), 2020, p.941.

[4] M. Afif, R. Ayachi, Y. Said, E. Pissaloux and M. Atri, "An evaluation of retinanet on indoor object detection for blind and visually impaired persons assistance navigation", Neural Processing Letters, 51, 2020, pp.2265-2279.

[5] M. Mahendru and S.K. Dubey, "Real time objects detection with audio feedback using Yolo vs. Yolo_v3", In 2021 11th international conference on cloud computing, data science & engineering (confluence) 2021, (pp. 734-740). IEEE.

[6] M. Vashisht and B. Kumar, "A survey paper on object detection methods in image processing", In 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA) 2020, (pp. 1-4). IEEE.

[7] K. Das and A.K. Baruah, "A Study of Data Processing for Object Recognition in Scene Image using FRCNN: A Smart Grid Technology", In 2021 Innovations in Energy Management and Renewable Resources (52042) 2021, (pp. 1-5). IEEE.