IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10561

A Project Work on Computer Vision Based Indoor Navigation System

Dr. Chethan Chandra S Basavaraddi[1], Prof. Sapna S Basavaraddi[2], Prof. Shashidhara M S[3], Rajath S (1KI19CS074)[4], Arpitha M (1KI19CS010)[4], Prathima K P(1KI19CS066)[4], Swarna N T (1KI19CS0113)[4]

Associate Professor, Department of Computer Science and Engineering, Kalpataru Institute of Technology, Tiptur-572201[1]

Assistant Professor, Department of Computer Science and Engineering, Kalpataru Institute of Technology, Tiptur-572201[2]

HOD, Department of Computer Science and Engineering, Kalpataru Institute of Technology, Tiptur- 572201[3]

Student, Department of Computer Science and Engineering, Kalpataru Institute of Technology, Tiptur- 572201[4]

Abstract : The navigation system plays a crucial role in determining the pedestrian's position and providing the optimal route to a specific destination. While outdoor navigation systems rely on GPS signals and line of sight with orbiting satellites for localization, these methods are not applicable for indoor environments. In indoor settings, various technologies such as Infrared (IR), Wi-Fi, and RFID have been employed for localization. However, this research project focuses on computer vision-based techniques for indoor positioning. To achieve indoor positioning, a visual indoor positioning system utilizing a CNN-Based image retrieval method has been developed. The system's database comprises images representing each signboard within the room or store, along with their corresponding CNN features. Furthermore, the system provides absolute coordinates with respect to a given local coordinate system and scene labels. In this study, we propose the utilization of the pre-trained ResNet-18 deep learning model to determine the pedestrian's position and destination. The user inputs their desired destination, and the system employs Dijkstra's algorithm to calculate the optimal route. The ResNet-18 model has demonstrated an impressive accuracy of 92% in preliminary testing. This research project aims to enhance indoor navigation by leveraging computer vision techniques and deep learning models. By providing accurate positioning and optimal routes, the system holds great potential for improving navigation experiences within indoor environments. The findings of this study contribute to the growing field of indoor localization and pave the way for further advancements in this domain.

I. INTRODUCTION

1.1 OVERVIEW

The term ‘navigation’ collectively represents tasks that include tracking the user’s position, planning possible routes and guiding the user through the routes to reach the desired destination. In the past, considerable number of navigation systems were developed for accessing outdoor and indoor environments. Most of the outdoor navigation systems adopt GPS and Global Navigation Satellite System (GNSS) to track the user’s position. Important applications of outdoor navigation systems include wayfinding for vehicles, pedestrians, and blind people [1]. In indoor environments, the GPS cannot provide fair accuracy in tracking due to non-line of sight issues [2]. This limitation hinders the implementation of GPS in indoor navigation systems, although it can be solved by using “high-sensitivity GPS receivers or GPS pseudo Lites” [3]. However, the cost of implementation can be a barrier to applying this system in real world scenarios.

Indoor navigation systems have broad number of applications. The certain applications are wayfinding for humans in railway stations, bus stations, shopping malls, museums, airports, and libraries. Visually impaired people also benefit from indoor navigation systems. Unlike outdoor areas, navigation through indoor areas is more difficult. The indoor areas have several types of obstacles, which increases the difficulty of implementing navigation systems. General block diagram of a human indoor navigation system is illustrated in Fig. 1.1

A human indoor navigation system mainly consists of the following three modules: (1) Indoor positioning system module, (2) Navigation module, and (3) Human– machine interaction (HMI) module. The indoor positioning system estimates the user’s position, the navigation module calculates routes to the destination from current location of the user, and the HMI module helps the user to interact with the system and provide instructions to user.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁞ Impact Factor 8.066 ⁞ Peer-reviewed / Refereed journal ⁞ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10561

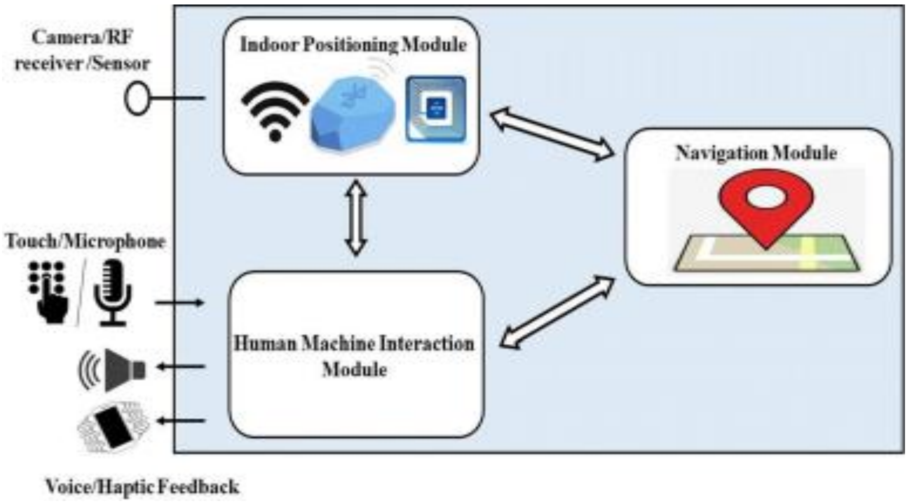

Figure 1.1 Human indoor navigation system: a general block diagram

1.2 PROBLEM STATEMENTS

People use navigation system for access unfamiliar environments. Navigation is the task of finding a path from one place to another place. The navigation system can be classified into two main categories: outdoor navigation and indoor navigation. For outdoor navigation, people currently rely on GPS (Global Positioning System). But in an indoor environment GPS system fails because of non-reception of GPS signals in indoor setting. The system we present could find potential use in a variety of application areas like: accessibility aids for the visually impaired people, shopping malls including hypermarket, hospitals, hotels, airports, railway stations, indoor robotics, tourism etc.

1.3 PROPOSED SYSTEM

The system operates as follows:

1. Data Collection and Database Creation:
 Images of signboards within the indoor environment are collected and stored in a database.
 Each image is associated with its corresponding CNN features, obtained by extracting the high-level features using the ResNet-18 model.

2. User Input:
 The user inputs their desired destination through an interface provided by the system.

3. Pedestrian Positioning:
 The system captures an image of the current environment using a camera or retrieves the image from an existing dataset.
 The pre-trained ResNet-18 model analyzes the image and extracts its CNN features.
 The system matches the extracted features with the stored features in the database to identify the current signboard and its associated coordinates.

4. Route Generation:
 Once the user's destination is determined and the current position is identified, the system employs Dijkstra's algorithm to calculate the optimal route.
 The algorithm takes into account the coordinates of the current position, the destination, and the known connectivity between signboards.
 The generated route guides the pedestrian through the indoor environment, providing step-by-step instructions.

5. Navigation and Guidance:
 The system presents the calculated route to the user through a user-friendly interface, which may include maps, visual cues, or textual instructions.
 The user can follow the provided instructions to reach their destination efficiently.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁒ Impact Factor 8.066 ⁒ Peer-reviewed / Refereed journal ⁒ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10561

II. LITERATURE SURVEY

The proposed work relates to many fields, such as visual localization, image retrieval, and visual pose estimation. At present, visual localization systems can be roughly divided into three categories.

Structure-based localization methods are the most common visual localization methods that utilize local features to estimate 2D-to-3D matches between features in a query image and points in 3D models, or employ 3D-to-3D matches between RGB-D images and 3D models. Then camera pose will be estimated from the correspondence. Similarly, Torsten [4] compared 2D image-based localization with 3D structure-based localization, and they drew a conclusion that purely 2D-based methods achieve the lowest localization and 3D-based methods offer more precise pose estimation with more complex model construction and maintenance. They proposed a combination of 2D-based methods with local structure-from-motion (SfM) reconstruction which has both a simple database construction procedure and accurate pose estimation. However, the drawback of their method is significantly longer run-time during the location process.

Image-based localization methods were pushed by massive repositories of public geo-labelled images. These methods employ an image retrieval-based strategy [5, 6], which match the query image with images from the database. Owing to the prosperity of social network and street view photos, quantity of images with geo-tags has emerged which can be used for reference to these data-driven image-based localization methods. Conventional methods retrieve images based on local descriptor matching and reorder with elaborate spatial verification [7-9]. Content based image retrieval search for images relies on visual content such as edges, colors, textures, and shape [10]. Recent works leverage deep convolution neural networks for image retrieval, the majority of them use a pre-trained network as local feature extractor. Moreover, some work even can address the problem of geometric invariance of CNN features [11, 12], and to accurately represent images of different sizes and aspects ratios [13, 14].

Learning-based localization methods emerged in the past few years, which benefited from the dramatic progress made in a variety of computer vision tasks. By training models from given images with pose information, scenes can be represented by these learned models. These learning-based localization methods either predict matches for pose estimation [15–17] or directly regress the camera pose such as PoseNet [18], PoseNet2 [19], and VlocNet [20]. PoseNet was the first approach to use DCNNs to solve the metric localization problem, and then Bayesian CNN implementation was utilized to address the pose uncertainty [21]. After that, architectures such as long-short term memory (LSTM) [22, 23] and symmetric encoder-decoder [24] were utilized to facilitate the performance of DCNNs.

III. SOFTWARE ENVIRONMENT

To be utilized effectively, all computer programming wants bound instrumentation segments or alternative programming assets to be offered on a computer. These necessities area unit noted as (computer) framework wants and area unit typically utilized a lot of unremarkably in situ of a hell for leather run the show. With increasing notice for bigger handling pressure and belongings in the additional energizing variants of programming, structure necessities have an inclination to increment when a moment. Trade investigators propose that this sample features a lot of impact in victimization movements up to current laptop computer frameworks than revolutionary head ways.

The proposed system architecture for the indoor navigation system based on computer Vision

1. User Interface: The user interface serves as the primary means for users to interact with the system. Users can input their desired destination and receive navigation instructions. The interface provides visual cues, maps, or textual directions to guide the user throughout the navigation process.

2. Image Capture and Preprocessing:
The system utilizes cameras or other image-capturing devices to capture images of the indoor environment. These images undergo preprocessing to enhance their quality and extract meaningful information. Techniques such as image filtering, edge detection, and feature extraction are applied to prepare the images for further analysis.

3. CNN-Based Image Retrieval:
A crucial component of the proposed system architecture is the utilization of a pre-trained ResNet-18 model for image analysis and feature extraction. The ResNet-18 model, which has been trained on a large-scale dataset, is capable of learning high-level visual representations. By applying the ResNet-18 model to the preprocessed images, the system extracts relevant features from the images, enabling accurate identification and recognition of signboards within the indoor environment.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁒ Impact Factor 8.066 ⁒ Peer-reviewed / Refereed journal ⁒ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10561

4. System Database:
The system's database plays a vital role in storing and managing the necessary information. It contains a collection of images representing each signboard within the indoor environment, along with their corresponding CNN features. The database also records the absolute coordinates of each signboard with respect to a given local coordinate system and assigns scene labels for navigation and route generation purposes.

5. Route Generation:
To generate the optimal route from the user's current position to their desired destination, the system employs Dijkstra's algorithm. This algorithm considers the coordinates of the current position, destination, and the connectivity between signboards in the indoor environment. By calculating the most efficient path, the system generates a step-by-step route that guides the user along the sequence of signboards and directions to follow.

6. Navigation and Guidance:
The system provides real-time navigation instructions to the user through the user interface. It continuously updates the user's position as they navigate through the indoor environment, recalculating the route if necessary. The navigation instructions can include visual cues, maps, or textual directions, ensuring accurate guidance throughout the navigation process.

By incorporating the pre-trained ResNet-18 model into the system architecture, the indoor navigation system benefits from its powerful image analysis capabilities. The ResNet-18 model has learned to recognize and classify various visual features from a large dataset, which enhances the system's ability to accurately identify signboards within the indoor environment. This, in turn, contributes to the overall effectiveness and reliability of the system in providing accurate positioning and navigation guidance to users.

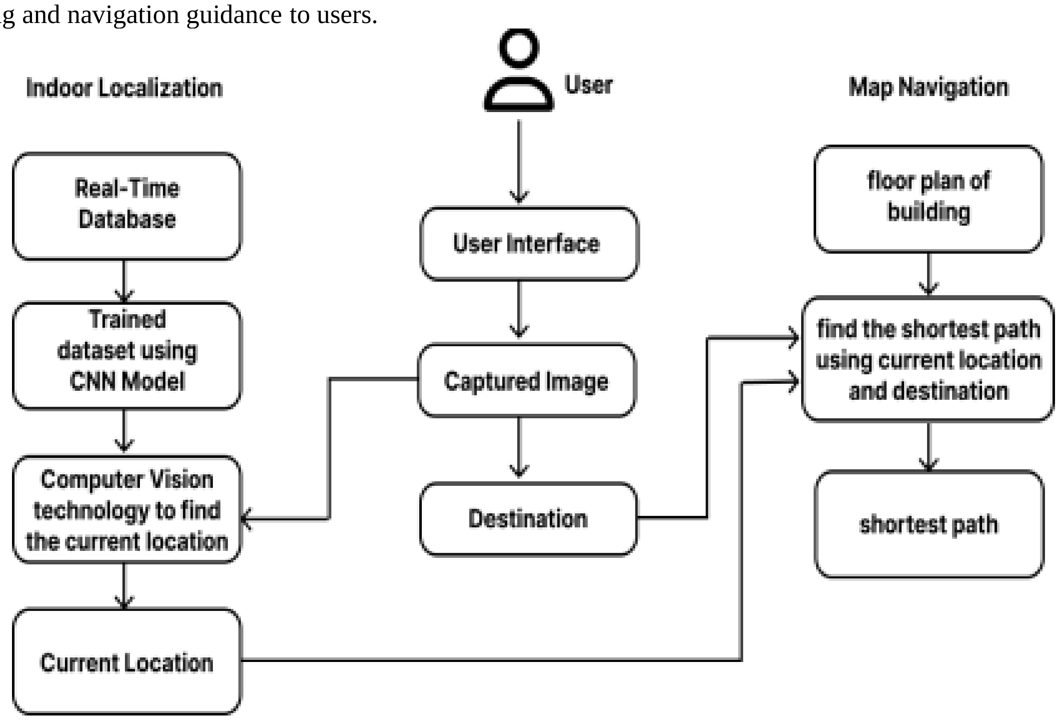

Figure 4.1 Overview of our Indoor Navigation System

The flowchart in Figure 4.1 illustrates the step-by-step process of the proposed system architecture. It visually represents the interactions between the different components, including image capture and pre-processing, CNN-based image retrieval using the ResNet-18 model, the system database, route generation with Dijkstra's algorithm, and navigation and guidance provided to the user. Through the integration of computer vision techniques and deep learning models, particularly the pre-trained ResNet-18 model, the proposed indoor navigation system offers accurate positioning, route generation, and navigation guidance within complex indoor environments. The flowchart provides a comprehensive overview of the system's functionality and demonstrates the seamless interaction between its various components to deliver a reliable and efficient navigation experience.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁒ Impact Factor 8.066 ⁒ Peer-reviewed / Refereed journal ⁒ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10561

IV. PROPOSED ALGORITHMS

The proposed indoor navigation system incorporates two key algorithms: CNN based image retrieval for signboard recognition and Dijkstra's algorithm for route generation.

1. CNN-Based Image Retrieval:

The system employs a Convolutional Neural Network (CNN) for image analysis and feature extraction. Specifically, the pre-trained ResNet-18 model is utilized to extract high-level features from the captured and pre-processed images of the indoor environment. The ResNet-18 model has been trained on a large dataset, enabling it to learn intricate visual representations and features.

By applying the ResNet-18 model to the pre-processed images, the system extracts elevant features that are compared with the features stored in the system's database. This comparison allows the system to accurately identify the current signboard within the indoor environment. The CNN-based image retrieval technique provides robust and accurate recognition of signboards, facilitating precise positioning within the indoor environment.

The ResNet-18 architecture consists of several layers and modules designed to effectively capture and process features in an image. Here is a breakdown of the ResNet 18 architecture:

• Input Layer: The network takes as input an RGB image of size 224x224 pixels.
• Convolutional Layers: The initial layer performs a convolution operation on the input image with a kernel size of 7x7 and a stride of 2. This is followed by a max pooling layer with a kernel size of 3x3 and a stride of 2. These operations help reduce the spatial dimensions of the input image.
• Residual Stages: ResNet-18 consists of four residual stages, each comprising multiple residual blocks. Each residual block consists of two or three convolutional layers, followed by batch normalization and ReLU activation. The residual blocks allow the network to learn residual functions and enable the flow of information through skip connections. The number of residual blocks per stage varies, with the first stage having two blocks, and subsequent stages having two or three blocks each. 13
• Global Average Pooling: After the last residual stage, a global average pooling layer is applied, which reduces the spatial dimensions of the feature maps to a single value per channel. This pooling operation helps in summarizing the learned features and reducing the total number of parameters.
• Fully Connected Layer: The final layer is a fully connected layer that performs the classification task. It maps the learned features to the corresponding class labels. For ResNet-18, the fully connected layer consists of a single linear layer followed by a softmax activation function, producing a probability distribution over the classes.

The ResNet-18 architecture is characterized by its residual connections, which allow for the training of deeper networks and address the vanishing gradient problem. These skip connections facilitate the propagation of gradients during training, leading to improved accuracy and performance.

ResNet-18 has a total of approximately 11 million trainable parameters, making it relatively lightweight compared to deeper versions of ResNet, while still achieving impressive performance on various image recognition tasks.

2. Dijkstra's Algorithm for Route Generation:

Once the user's current position and desired destination are known, the system employs Dijkstra's algorithm to calculate the optimal route between the two points. Dijkstra's algorithm is a graph-based algorithm that finds the shortest path in a graph with weighted edges. In the context of indoor navigation, the signboards within the indoor environment act as nodes in the graph, and the connectivity between signboards represents the weighted edges.

By considering the coordinates of the user's current position, destination, and the connectivity between signboards, Dijkstra's algorithm determines the most efficient path. The algorithm calculates the shortest route by iteratively exploring neighboring signboards and updating the distances to reach each node. The resulting route guides the user through the indoor environment, providing step-by-step instructions to follow.

Here's an explanation of how Dijkstra's Algorithm works:
• Initialization: Start by assigning a tentative distance value to every node in the graph. Set the distance of the source node to 0 and mark it as the current node.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁒ Impact Factor 8.066 ⁒ Peer-reviewed / Refereed journal ⁒ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10561

• Selection of the Next Node: Among the set of unvisited nodes, choose the node with the smallest tentative distance and make it the current node. Initially, the source node will be selected in the first iteration.
• Update Neighboring Nodes: For each neighbor of the current node, calculate the distance from the source node through the current node. If this distance is smaller than the previously assigned tentative distance, update the tentative distance of the neighbor.
• Mark Current Node as Visited: Once all the neighboring nodes have been processed, mark the current node as visited. This ensures that its tentative distance will not be further modified.
• Repeat Steps 2-4: Repeat the above steps until all nodes in the graph have been visited or until the destination node (if specified) has been reached.
• Termination: Once the algorithm terminates, the shortest path from the source node to each node in the graph will have been determined.
• Path Reconstruction: To obtain the actual shortest path from the source node to any given node, follow the chain of nodes with the minimum tentative distance backward until reaching the source node. This process allows the sequence of nodes forming the shortest path to be obtained.

Dijkstra's Algorithm uses a greedy approach by continuously selecting the node with the smallest tentative distance. It guarantees finding the shortest path as long as the graph does not contain negative-weight edges.

V. IMPLEMENTATION

The implementation of the proposed indoor navigation system involves several key steps to bring the system architecture to life. Here, we outline the main aspects of the implementation process:

5.1 Indoor Localization:
In the first step, the system focuses on indoor localization, which aims to determine the user's current position within the indoor environment. This is achieved using computer vision techniques and the pre-trained ResNet-18 model.

The system captures an image of the user's surroundings using a camera or image capturing device. The captured image is pre-processed, including resizing, normalization, and other necessary transformations, to ensure consistency and compatibility with the ResNet-18 model. The pre-processed image is then passed through the ResNet-18 model, which extracts relevant features from the image. The model has been trained on a dataset containing images of various signboards or landmarks within the indoor environment. By comparing the extracted features with the features stored in the model's database, the system can determine the user's current location.

VI. OBSERVATIONS AND RESULTS

In the evaluation of the proposed indoor navigation system, rigorous testing was conducted within the college administration block, consisting of eight distinct rooms: 'MBA block', 'Auditorium', 'Entry', 'Exit', 'Health Centre', 'Library', 'Office', 'President Cell', and 'Principal Cell'. To train the ResNet-18 model for accurate indoor localization, a comprehensive dataset was meticulously collected for each of these rooms. For the training the ResNet-18 sample image datasets are shown in the below fig[6.1-6.8]

Figure 6.1 MBA block

Figure 6.2 Auditorium

DOI: 10.17148/IARJSET.2023.10561

Figure 6.3 Entry

Figure 6.4 Exit

Figure 6.5 Health Centre

Figure 6.6 Library 27

Figure 6.7 President Cell

Figure 6.8 Principal Cell

The trained ResNet-18 model exhibited remarkable performance, achieving an impressive accuracy rate of 92% in accurately identifying and localizing the user's current position within the indoor environment. This substantiates the efficacy of the model in recognizing and distinguishing between the diverse visual features present in the collected dataset, enabling precise room identification. To facilitate seamless navigation, a graph was meticulously constructed, comprising eleven essential nodes representing prominent locations within the college administration block.

The nodes were labeled as follows: 'Entry', 'Lobby', 'library', 'principal cell', 'stairs', 'exit', 'MBA block', 'office', 'auditorium', 'health centre', and 'president cell'. By leveraging Dijkstra's algorithm, the system efficiently determined the shortest and most optimal path from the user's current location to the desired destination. The system's output proved to be highly effective in providing step-by-step instructions, meticulously guiding users throughout their indoor navigation journey. The instructions were precisely tailored to the user's specific route, incorporating accurate distances and directional guidance. This level of detail ensured that users could confidently navigate the college administration block, relying on clear and concise instructions for each step of their journey.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10561

For instance, consider the following sample instructions provided by the system:

- from Entry
- 10.3 feets
- slight rigth from Lobby
- 20.6 feets
- straight towards principal cell
- 27.7 feets
- turn left towards stairs
- 28.7 feets
- straight towards office
- 20 feets
- reached Auditorium

The meticulous and comprehensive instructions provided by the system empower users to navigate the complex indoor environment effortlessly, ensuring a seamless and efficient experience. By meticulously adhering to the step-by-step instructions, users can confidently traverse the college administration block, arriving at their desired destination promptly and without unnecessary detours.

In conclusion, the implemented indoor navigation system, leveraging the power of the ResNet-18 model for accurate indoor localization and incorporating Dijkstra's algorithm for optimal route calculation, has exhibited exceptional performance. The provision of meticulous and precise step-by-step instructions ensures users can effortlessly navigate the college administration block, reflecting the immense potential of computer vision based techniques and algorithmic approaches in enhancing indoor navigation systems.

VII. CONCLUSION

In conclusion, the proposed indoor navigation system offers a robust and efficient solution for navigating complex indoor environments. With its high accuracy in localization and precise step-by-step instructions, the system provides users with a seamless navigation experience.

Further advancements and refinements in this field hold great potential for revolutionizing indoor navigation and improving the quality of life for individuals in various settings.

REFERENCES

[1] Godha, Saurabh, and Gerard Lachapelle. "Foot mounted inertial system for pedestrian navigation." Measurement Science and Technology 19, no. 7 (2008): 075202.
[2] Koyuncu, Hakan, and Shuang Hua Yang. "A survey of indoor positioning and object locating systems." IJCSNS International Journal of Computer Science and Network Security 10, no. 5 (2010): 121-128.
[3] Tolias, G.; Avrithis, Y. Image Search with Selective Match Kernels: Aggregation across Single and Multiple Images. Int. J. Comput. Vis. 2016, 116, 262.
[4] Tolias, G.; Jégou, H. Visual query expansion with or without geometry: Refining local descriptors by feature aggregation. Pattern Recognit. 2014, 47, 3466–3476.
[5] Liew,W.C.; Law, N.F. Content-Based Image Retrieval. Comput. Syst. Appl. 2009, 43, 85.
[6] Ge, S., Wu, B., Shen, H. T., & Cao, X. (2021). DeepIMU: Deep Learning-based Indoor Localization and Navigation with IMU Sensors. ACM Transactions on Sensor Networks (TOSN), 17(1), 1-23.
[7] Sun, R., Lin, Y., Zhou, F., Liu, L., & Gao, Y. (2020). Vision-based Indoor Navigation for Mobile Robots Using Deep Reinforcement Learning. IEEE Transactions on Industrial Electronics, 68(3), 2463-2473.
[8] Xu, L., Deng, Z., Liu, K., & Zhang, B. (2021). Indoor Visual Navigation for Mobile Robots Using Hierarchical Deep Reinforcement Learning. IEEE Transactions on Robotics, 37(4), 1071-1087.

