202211004_Plagiarism_Report

by Jigar Shekhat

General metrics

40,617	6,135	547	24 min 32 sec	47 min 11 sec
characters	words	sentences	reading time	speaking time

Score

98

This text scores better than 98% of all texts checked by Grammarly

Writing Issues

102	10	92
Issues left	Critical	Advanced

Plagiarism

This text seems 100% original. Grammarly found no matching text on the Internet or in ProQuest's databases.

Writing Issues

87	**Clarity**	
40	Passive voice misuse	
15	Unclear sentences	
25	Wordy sentences	
4	Intricate text	
3	Hard-to-read text	
2	**Engagement**	
2	Word choice	
13	**Correctness**	
2	Misplaced words or phrases	
1	Faulty subject-verb agreement	
2	Wrong or missing prepositions	
1	Improper formatting	
4	Unknown words	
1	Misspelled words	
1	Confused words	
1	Conjunction use	

Unique Words

Measures vocabulary diversity by calculating the percentage of words used only once in your document

21%

unique words

Rare Words

Measures depth of vocabulary by identifying words that are not among the 5,000 most common English words.

46%

rare words

Word Length

Measures average word length

5.3

characters per word

Sentence Length

Measures average sentence length

11.2

words per sentence

202211004_Plagiarism_Report

Leveraging Semantics Information to Recognizing Seen and Unseen Locations Using a Mobile Rover

Abstract—The complicated nature of rapidly changing environments is not well captured by traditional static mapping
techniques. To solve this problem, we describe a novel initiative
that utilizes cutting-edge technology and innovative approaches
to provide mobile robots with the ability to acquire information
about their surroundings. First, We examine Graph-Structured
Sum-product networks (GraphSPNs) are a probabilistic method for
structured prediction in circumstances where arbitrary, dynamic
graphs represent latent variable dependencies. While many structured
prediction methods impose strict limits on the relationships
between inferred variables, many real-world Problems can only
be described[1] using complicated graph structures of various sizes,
frequently tainted with noise when derived from
actual data. We show how noisy topological relations found
by a robot navigating extensive office areas can be used[2] to support
an inference about semantic, conceptual place descriptions. We
implement a graph neural network to address these problems.
However, because this data contains spatial data on the neighboring node, we cannot get reliable conclusions from it.
Based entirely on spatial information, we are unable to estimate
location. To work on the natural, significant aspects of

place. Our project focuses on integrating YOLOv8 object detection, WiFi measurements, and GPS coordinates to create a complete understanding of indoor spaces. The key objective is to enable mobile robots to identify their location, even when GPS signals are weak or unavailable, providing relevant information about the places they navigate. This project exceeds the limitations of static mapping by offering dynamic, real-time insights into constantly changing environments. To achieve robust location identification, we employed Support Vector Machines (SVM), Deep Neural Networks (DNN), Naive Bayes, and Random Forest. These models effectively categorize place names and types, offering mobile robots a comprehensive understanding of their surroundings. Our project represents a significant advancement in location identification for mobile robots in indoor environments. Combining state-of-the-art object detection, WiFi measurements, and GPS coordinates with a machine-learning framework equips robots with the knowledge and adaptability needed to excel in dynamic surroundings.

Index Terms—Object Detection, WiFi Measurements, GPS Coordinates

I. I INTRODUCTION

The field of robotics continues to develop, and mobile robots must be able to see and understand their surroundings, especially in dynamic interior situations. For a mobile robot to remain aware of its surroundings, it needs a representation of spatial information, a framework that organizes the understanding of the environment. Constantly changing Environments provide challenges for traditional mapping and localization approaches. A representation[3] should take advantage of established spatial links to integrate knowledge across geographical scales and levels of abstraction. Topological maps are a well-established framework for characterizing the spatial connections between neighboring locations.

They make it easier to establish high-level conceptual data and provide planning algorithms easy access. [4]To integrate the acquired geographical knowledge, resolve ambiguities, and provide predictions[5], such frameworks often use structured prediction algorithms[4]. Unfortunately, a robot exploring a real-world environment often[6] finds complicated and noisy relations, which poses[7] complex inference problems. As the robot explores its surroundings, the number of nodes and connections in topological maps increases and varies based on the surroundings. However, robots cannot obtain reliable results if they have[8] spatial knowledge about their surroundings. We describe an innovative approach that uses WiFi measurements, GPS coordinates, and YOLOv8 object identification to overcome that challenge and enable mobile robots to understand their surroundings thoroughly. This initiative has the potential to completely change how mobile robots react and interact with indoor environments. The immense complexity and dynamic of indoor surroundings are unique features. The dynamic nature of indoor places is frequently tricky for traditional static mapping systems to adapt to because they depend on pre-existing maps or infrastructure. In tackling this problem, our research uses innovative strategies to allow mobile robots to navigate well despite the lack of dependable GPS signals. This[9] will enable them to recognize their location and comprehend the environments they navigate. Our project's main element is a comprehensive dataset carefully gathered to represent various indoor features. Several types of information are included[10] in this collection, such as GPS coordinates, WiFi signal measurements, place categories, building names, and photos of various inside spaces.

We added YOLOv8 object identification, a state-of-the-art deep learning model to this dataset to enhance its ability to identify things unique to each location. By applying semantic understanding to the dataset, this method helps the system identify and classify areas according to the stuff in them.

Our approach to improving position recognition accuracy and resilience heavily relies on this semantic understanding, even in scenarios with poor or inconsistent GPS

signals. We used

2

Machine learning models like Support Vector Machines (SVM), Deep Neural Networks (DNN), Naive Bayes, and Random Forest to further improve the system's capabilities. These

models act as the system's brain, allowing location names and categories to be accurately classified. Mobile robots can

adapt to their environment and carry out duties successfully by efficiently organizing these features. This project has extensive potential effects that extend methods in addition to robots.
This technology can benefit mobile robots in
various industries, including security, logistics, and healthcare.
They are flexible advantages since they can move around and navigate effectively in many environments. Our approach also presents a novel function for recognizing places that have never been observed[11]. When the robot visits a new location for the first time, it gathers information and allocates a particular tag. Subsequent visits to the same location result in rapid and accurate recognition. Our study significantly advances mobile robot location recognition in dynamic indoor situations. By integrating advanced object identification, WiFi
measurements, and GPS coordinates with a robust machine-learning framework, we equip robots with the intelligence and flexibility required to perform well in dynamic situations. The technology has the strength to improve indoor
navigation and enable a broad range of applications, changing the capabilities of mobile robots in various industries. As the project progresses, we will examine the particulars of each part and show how the system could impact indoor robots in the future.

II. BACKGROUND

We give a thorough overview of the project's background in this section. Initially, we provide a detailed explanation of the

characteristics of GraphSPN and GCN, the process of obtaining inferences, and the techniques used for learning the network's parameters and structure. The spatial knowledge model
that we utilize to facilitate environmental awareness for mobile robots is called Deep Affordance Spatial Hierarchy (DASH), and it is available in the COLD dataset. We discuss topological maps in general and the DASH topological map generation process. After that, We describe that This Project focuses on the challenges of indoor localization, the significance of YOLOv8 object detection, WiFi frequency analysis, and GPS signal utilization. The difficulties of signal attenuation and multipath effects in indoor environments are highlighted[12], which impact traditional GPS-based methods. YOLOv8's real-time object detection capabilities and WiFi frequency analysis's ability to capture spatial signal distributions are explained[13]. While GPS signals are less reliable indoors, their potential for contributing to indoor context is acknowledged. The concept of multimodal data fusion, combining visual, wireless, and satellite-based data is introduced[14] to enhance indoor localization accuracy.
A. Localization:
One essential prerequisite for mobile robots functioning in interior environments is precise localization. Finding a robot's location about[15] its environment is what it involves.
Traditional GPS-based localization is sometimes insufficient in

dynamic[16] interior environments due to the vulnerability of GPS signals to multipath interference and signal blockages, which

can cause the movements to become nonexistent or inaccurate.
Other techniques, such as WiFi-based location, have become more popular to overcome this difficulty. WiFi positioning estimates a device's location by utilizing the signals from wireless access points. Robots can now traverse within buildings without GPS signals because of the effectiveness of this technology's indoor localization solutions.

B. Object Detection:

For mobile robots, object recognition is essential to contextual comprehension. It involves identifying and recognizing things in the robot's surroundings. The cutting-edge deep learning-based object identification system YOLO (You Only Look Once) is excellent at recognizing objects in real time. YOLO's integration into the project gives the system the capacity [17] to recognize and classify things that are [17] exclusive to particular interior spaces. By identifying things typical of specific locations, [18] this semantic knowledge improves the system's ability to recognize and navigate interior surroundings.

III. LITERATURE S URVEY

Research by Guvenc and Chong provides a comprehensive survey of Time of Arrival (TOA) based wireless localization and non-line-of-sight (NLOS) mitigation techniques. Understanding WiFi signal strengths and distributions is crucial for our project's localization algorithm.[4]

• Autonomous indoor mobile robots must have WiFi localization and navigation. Examining the latest techniques in WiFi-based localization offers insightful information that helps us improve the strategy of our project.[2]

• Satellites facilitate accurate positioning and navigation

across outdoor environments. This integration not only

enhances the accuracy of the overall system but also ensures a continuous and

reliable localization experience as

mobile robots traverse between diverse environments.[3]

•

IV. T TECHNOLOGIES

A. SPN

SPN, proposed by Poon and Domingos, is a new class

of probabilistic graphical models with built-in properties that

allow tractable inference, a significant advantage over traditional

graphical models such as Bayesian networks. The idea is built [19]

upon Darwiche's work on network polynomials and differentials in arithmetic

circuit representation of the polynomial [5].

Here, we provide the definition of SPN and several of its fundamental

properties[6].

Definition (SPN) Let X = {X1, · · · Xn } be a set of variables.

A Sum-Product Network(SPN) defined over X is a rooted

directed acyclic graph. The leaves are indicators. The internal

nodes are sum nodes and product nodes. Each edge (i, j) from

the sum node

Pi has a nonnegative weight wi j. The importance of a sum

node is j ∈ Ch(i), wij vj, where Ch(i) is s the children of i.

The value of a product node is the product of the importance of its

children. The value of an SPN is the value of its root. We use

3

S to denote an SPN as a function of the indicator variables (i.e., the leaves). Let x be an instantiation of the indicator variables, an entire state. Let e be an evidence (partial instantiation). For a given node i, we use Si to denote the sub-SPN rooted at i. Also, we use xap[20] to mean [Xp = a] is true, and use x–

pa

to mean the negation, for simplicity. We define the following properties of SPN.

Figure 1: SPN implementing a naive Bayes mixture model (three components, two variables).

We call such nodes placeholders and assume that the robot still needs to obtain evidence about their semantics.

The topological graph is assembled incrementally based on a dynamically expanding 2D occupancy map. The 2D map is built[21] from laser range data captured by the robot using a grid mapping approach based on Rao-Blackwellized particle filters. Placeholders are added at neighboring, reachable but unexplored locations and connected to existing places. Then, once the robot performs an exploration action, a placeholder is converted into a home, and local evidence captured by the robot about the semantic place category is anchored to the graph node [14]. An example of such semantic-topological the map is shown[22] in Fig. 2

Fig. 1: Example of SPNs

Fig. 1: A simple SPN for a naive Bayes mixture model

P(X1, X2), with three components over two binary variables.

The bottom layer consists of indicators for X1 and X2.

Weighted sum nodes, with weights attached to inputs, are

marked with +, while product nodes are labeled[23] with ×. We

will assume that sums and products

are arranged[24] in alternating layers, i.e., all children of an aggregate

are products or leaves, and vice-versa.

For example, for the SPN in Figure 1, S(x1 , x2 , x1 , x2) =

0.5(0.6x1 + 0.4x1)(0.3x2 + 0.7x2) + 0.2(0.6x1 +

0.4x1)(0.2x2 + 0.8x2) + 0.3(0.9x1 + 0.1x1)(0.2x2 + 0.8x2)

The network polynomial S(x) = (0.5 × 0.6 × 0.3 + 0.2 ×

0.6 × 0.2 + 0.3 × 0.9 × 0.2)x1 x2 + · · · If a complete state

x is X1 = 1, X2 = 0, then S(x) = S(1, 0, 0, 1) If The

evidence e is X1 = 1, then S(e) = S(1, 1, 0, 1). Finally,

S(∗) = S(1, 1, 1, 1).

B. Topological Graphs

GraphSPNs are applicable[25] to arbitrary graphs. However,

here, our dataset is specifically[26] a topological graph built by a

mobile robot exploring a large-scale environment. The primary

purpose of our topological graph is to support the behavior of

the robot.[27] As a result, nodes in the graph represent places the

robot can visit, and the edges represent navigability. The graph

nodes are associated with latent variables representing semantics, and the

edges can be seen[28] as spatial relations forming a

global semantic map. Local evidence about the semantics of a place might be available, and such evidence is inherently uncertain and noisy. Additional nodes in the graph are created[29] to represent exploration frontiers, possible places the robot has not yet visited but can navigate to.

Fig. 2: Topological graph (Stockholm-floor4)

C. GraphSPNs

GraphSPNs learn a template model over arbitrary graph-structured data, with local evidence X_i and latent variables $Y_i = \{Y_{i1}, \cdots, Y_{iM}\}$ for each graph node or edge i, with dependencies between the latent variables expressed in terms of the graph structure. Then, an instance of GraphSPN distribution $P(X_1, Y_1, \cdots, X_N, Y_N)$ is assembled for a specific graph to perform inference. To this end, we define a set S of template sub-graphs, and associate each template sub-graph S ∈ S with a separate template SPN modeling the distribution over variables X_i and Y_i corresponding to the nodes and edges of the template sub-graph. The structure and parameters of each template SPN can be learned directly from data obtained by decomposing training graphs into sub-graphs corresponding to S.

Given a set of trained template SPNs and a specific graph to be modeled, an instance of GraphSPN is assembled[30] as illustrated in Fig. 2. First, the chart is decomposed[31] into multiple times, each time differently, into sub-graphs using sub-graph templates S in descending order of the template size (i.e., more complex templates have priority). The subgraphs should not.

4

overlap in each decomposition and the corresponding template SPNs should cover all variables of Xi Yi in the model. This condition guarantees completeness and decomposability, resulting in a valid instance of GraphSPN. For each decomposition and each sub-graph,[32] we instantiate the corresponding

template SPN, resulting in multiple SPNs sharing weights and structure. The instantiations for a single graph decomposition are combined[33] with a product node, and the product nodes for all decompositions become children of a root sum node, realizing the complete mixture model.

In order to[34] incorporate the latent variables Yij, we include an an intermediate layer of product nodes into the template SPNs. Each such product node combines arbitrary distribution of Di kj[35] (Xi) with an indicator λYi j = ckj[36] for a specific value ckj[37] of Yij. The template SPN is built[38] on top of the product nodes and can be learned from data and the distributions Di kj[39] (Xi) can be arbitrary, potentially also realized with an SPN with a data-driven structure[15].

In our experiments, we assumed only one latent variable (semantic place category) Yi per graph node i, with V al(Yi) = {c1, · · ·, cL }, and we defined Dik (Xi) for a single hypothetical binary observation xi, which we assumed to be observed[40]:

(

αik

if Xi = xi

k

Di (Xi) =

k

1 − αi if Xi = xi

Such simplification allows us to thoroughly evaluate GraphSPNs for the problem of learning topological semantic maps

by directly simulating hypothetical evidence about the semantic category of varying uncertainty and under various noise

conditions. Furthermore, it allows us to compare GraphSPNs

with Markov Random Fields using the same αik as the value

of local potentials, i.e. ϕi (Yi = ck) = αik. The proposed

approach naturally extends to the case where a more complex

distribution is used [41] to model semantic place categories based

on robot observations.

GCNs are one of the most popular GNN architectures and are

designed to perform convolution operations on graphs[12].

E. GCNs

Graph convolutional networks (GCNs) (Kipf & Welling,

2017), generalize convolutional neural networks (CNNs) (LeCun et al., 1995) to graph-structured data. To learn the graph

representations, the "graph convolution" operation applies the

same linear transformation to all the neighbors of a node

followed by a nonlinear activation function. GCNs have become one of the most popular GNN architectures due to their simplicity, effectiveness, and scalability.

GCNs work by propagating[43] information from a node's neighbors to update its own[42,43] representation. This[44] is achieved[45] through a convolution operation in the graph domain, where the filter is applied[46] to the node and its neighbors. The output of the convolution operation is a new feature representation of the node, which incorporates information from its local neighborhood.

The mathematical formula for a single layer of GCN can be written as follows:

H (l+1) = σ(D̂– 2 ÂD̂– 2 H (l) W (l))

where H (l) is the hidden state matrix of the nodes in layer l, Â is the adjacency matrix of the above graph with self-loops, D̂ is a diagonal node degree matrix of Â, W (l) is the weight matrix of layer l, and σ is the activation function.

The D̂– 2 ÂD̂– 2 term is known as the normalized adjacency matrix, which ensures that the weight of the connections between nodes is scaled based on their degree. The output of the convolution operation is then passed[47] via the activation function to produce the updated hidden state matrix H (l+1).

Multiple layers of GCN can be stacked to learn increasingly complex representations of the nodes in the graph. The output of the last layer can be used for different downstream tasks such as node classification and graph classification[13].

D. GraphNNs

Graph Neural Networks are type of neural network that are designed[48] to process graph-structured data. They
are used to perform various tasks such as node identification, graph classification, and link prediction. GNNs have become popular because to[49] their ability to handle complex relationships and dependencies in graph data.
GNNs work by using[50] a message-passing algorithm to update the hidden states of nodes in a graph based on the hidden forms of their neighbors. The algorithm involves passing messages between nodes and aggregating the messages to update the hidden[51] records. The process is repeated for multiple iterations until the hidden states converge to a stable condition. GNNs can be designed[52] with different types of[53] message-passing functions, such as convolutional or attention-based functions[53].
There are various architectures of GNNs, such as graph Convolutional Networks (GCNs), Graph Attention Networks (GATs), GraphSAGE and Graph Isomorphism Networks (GINs). Each of these architectures has its own[54] advantages and disadvantages, and the choice of structure depends on the specific work and the characteristics of the graph data.

F. YOLOv8

YOLOv8 was released in January 2023 by Ultralytics, the company that developed YOLOv5. YOLOv8 provided five scaled versions: YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l and YOLOv8x. YOLOv8 supports multiple vision uses such as object

detection, segmentation, pose estimation, tracking, and classification [8].
YOLOv8 is the latest iteration in the YOLO family of detection models, which are known for their capabilities for [55] joint detection and segmentation. Similar to [56] YOLOv5, the architecture consists of a backbone, head, and neck. It boasts a new architecture, improved convolutional layers (backbone), and a more advanced detection head, making it a top choice for real-time object detection. YOLOv8 also offers support for [57] the latest computer vision algorithms, such as instance segmentation, enabling multiple object detection in an image or video. The model uses the Darknet-53 backbone network, which is faster and more precise than the previous YOLOv7 [10] network. YOLOv8 predicts bounding boxes through an anchor-free. [1]

5

G. WiFi Frequency Analysis

WiFi frequency analysis involves the collection and analysis of WiFi signal strength measurements within indoor environments. WiFi networks emit radio waves at specific frequencies, and the strength of these signals can be measured [58] in decibels relative to 1 milliwatt (dBm). By analyzing the distribution of WiFi signal strengths across an indoor space, it's possible to gain insights into the spatial characteristics of the environment, including signal attenuation, interference, and the presence of obstacles.

Understanding WiFi signal distribution is crucial for indoor localization projects. WiFi signals interact with walls, furniture, and other objects, leading to variations in signal strength across different locations. By mapping these variations, it's possible to create location fingerprints that help identify and differentiate between various [59] indoor spaces based on their unique WiFi signal patterns.

WiFi frequency analysis provides a [60] valuable layer of [60] information that contributes to the overall indoor localization process. By combining WiFi signal strengths with other data sources like YOLOv8 object detection and GPS signals, the project aims to create a comprehensive understanding of indoor spaces. This integrated approach enhances the accuracy of indoor positioning and enables the determination of the best-matching room [61] based on a combination of modalities.

H. GPS Signal Utilization

Global Positioning System (GPS) signals, initially designed for outdoor navigation, [62] can still provide valuable context for indoor localization projects despite their limitations within enclosed spaces. GPS signals are transmitted [63] by a network of [64] satellites orbiting the Earth and contain information about satellite positions and timestamps. Although GPS signals weaken and experience multipath effects indoors because of signal attenuation caused by walls and structures, they can still contribute to indoor localization in specific scenarios.

In indoor environments, where traditional GPS signals might be weak or entirely [65] absent, utilizing available GPS signals or data derived from weak signals can provide additional contextual information. This context can be integrated with[66] other localization methods, such as WiFi frequency analysis and object detection to enhance the accuracy of indoor positioning.

Methods of Utilization:

1) Coarse Location Estimation: Weak GPS signals might still be received indoors, albeit with reduced accuracy. These signals can provide a rough estimate of the user's location within a building or floor.

2) Contextual Information: While weak GPS signals may not provide precise indoor positioning, they can offer additional context. For example, they can help validate or corroborate indoor location estimates derived from other sources.

3) Hybrid Localization: Hybrid localization methods combine data from multiple sources to enhance accuracy. GPS signals can be integrated[67] with WiFi frequency analysis, object detection, and other techniques

To create a more comprehensive localization solution. These hybrid methods aim to mitigate the limitations of individual techniques and leverage the strengths of each data source[68].

Challenges:

1) Weak Signals: Indoor environments typically receive weaker GPS signals due to signal blockage from structures. These signals might not be as accurate as outdoor signals.

2) Multipath Effects: Multipath interference occurs when GPS signals bounce off walls and structures, leading to signal distortions[69] and inaccuracies. This effect can complicate the use of GPS data for indoor positioning.

3) Coarse Accuracy: The accuracy of GPS signals indoors is generally lower than in open outdoor spaces. However, even coarse location estimates can be valuable for specific applications.

4) Integrity Monitoring: GPS signal integrity monitoring is essential, as weak or inaccurate signals can lead to erroneous location estimates. Implementing measures to assess the quality of received signals is necessary.

V. DATASETS

A. COLD Dataset

Cognitive rObot Localization Database (COLD), a large database of localization, sensory information, and local environment representations (place scans), as well as topological maps, thoroughly annotated with semantic place categories. There are 100 sequences in total, collected by a mobile robot in three buildings at three different locations: Stockholm, SE (40 seqs), Freiburg, DE (26 seqs), and Saarbrucken, DE (34 seqs).

The COLD-TopoMaps Dataset consists of topological maps collected as the robot explores the environment. Each topological map is an undirected graph (also called ”topological graph”), where vertices are places that the robot could access, and edges indicate navigability. Each area is annotated by a semantic place category, such as corridor, kitchen, or doorway. [70]

nodes.dat A CSV file where each row is the data for a particular node on the topological map. The column values (and types) are below:

Each place has eight views (as shown in the figure below). For every edge connected to the node through an idea, there are Three entries of data in the row for this node:

If the neighbor node ID is -1 for a view number, then this the view has no connected edge. If we think of a node as a disk, then, the view numbers correspond to regions as annotated in the following illustration[9].

B. New Dataset

The dataset includes a comprehensive collection of photos illustrating various indoor spaces and rooms. A broad representation of the interior areas under examination is ensured [71] by the [72] different angles and views from which these photographs are captured [73]. Essentially, they function as visual representations of the settings in which our mobile robots work.

6

Fig. 3: Single Floor Graph

Fig. 4: Data Attributes

Fig. 7: Schema of Dataset

Fig. 5: 3 Entries for Each view

Notification outcomes produced by the advanced deep learning algorithm YOLOv8. By correlating identified items to particular interior places, our results provide a layer of semantic knowledge to the dataset. By[74] identifying things typical of specific locations, object detection improves accuracy and adaptability[74] and provides semantic comprehension that helps the system identify and cope with interior environments.

VI. W ORKING S TEPS

Fig. 6: 8 views

work, capturing the constantly changing arrangements, light, and item presence that defines dynamic interior spaces. Apart from the images, the collection also includes WiFi signal strength measurements taken from well-placed wireless access points in the target areas. The complex geographical distribution of WiFi signals is captured[75] by this data, which is an invaluable asset for position determination. We obtain a thorough grasp of the signal strength variations by gathering WiFi measurements from various locations within each interior environment, which[76] is an essential component in our system’s capability to locate a robot precisely. Pywifi is the module we used. For dealing with wireless interfaces, it provides a cross-platform Python package. With the help of this module, we can get a list of WiFi networks that are available[77], measure them, and save the results in a dataset.

In order to[78] provide our system with ground truth and spatial reference, GPS coordinates are also captured[79] at certain spots within the interior areas. In order to[80] provide smooth transitions when robots navigate between indoor and outdoor regions, these coordinates serve as a crucial connection between outdoor and inside localization. In situations when GPS signals are weak or unstable, this integration is essential since it enables continuous navigation and precise location determination.

Furthermore, the dataset has been modified[81] with object identity-

A. Previous Method Working Steps

The project described above involves node classification in the first step of semantic mapping. Here are the general steps that we followed for the previous method:

1) First, the COLD (Cognitive rObot[82] Localization Database) dataset was used[83] to launch the project. Working with this dataset proved challenging. Topological maps are acquired and saved in the COLD-TopoMaps Dataset while the robot explores its surroundings. Each map is an undirected graph, with vertices serving[84] as potential access places for the robot and edges denoting navigability. Each area is annotated[85] with a semantic place category.

2) We attempted to create GraphSPN while working with the COLD dataset, but as[86] a result, we encountered several library dependency issues, which we later identified

as deprecated versions of the same.

3) In the following phase, we choose to implement GNN.

Graph Neural Networks (GNNs) are built[87] to work with graph data, representing a network of connected nodes and edges.

4) An adjacency matrix and a feature matrix are used to[88] represent a graph as the input data for the GCN node classification algorithm. While the feature matrix comprises feature vectors for each and every[89] node in the chart, the adjacency matrix encodes the connections between the nodes in the network. Subsequently, we turned the dataset into a feature matrix and an adjacency matrix for our model's input.

7

5) The first step in GCN is to perform a graph convolution operation. This operation involves multiplying the feature matrix with the adjacency matrix to capture the relationship between nodes in the graph.

6) After performing the graph convolution operation, an activation function is applied to introduce non-linearity to the model. The Rectified Linear Unit (ReLU) function is the most commonly used activation function.

7) The output of the final layer of GCN is then fed into a softmax function to obtain a probability distribution over the different classes of nodes.

8) The GCN model is trained using labeled data to minimize the cross-entropy loss between the predicted probability distribution and the proper class labels.

B. Main Method Working Steps

Here are the general steps that we followed to create this project:

1) Data Collection and Preprocessing: We collected Images, Record GPS Signals Readings, and Measure the WiFi Signal Strengths of Different Rooms. We annotated images with bounding boxes around objects of interest using the annotation tool Roboflow. Each bounding box should correspond to a thing, and the annotation file should contain object labels and coordinates. We calculate the elevation(height) of the location from its coordinates of GPS and store them in the dataset. After that, We divide our dataset into training, validation, and testing sets.

2) Model Training: We used the preprocessed dataset and annotations to train YOLOv8 using the training script provided by the framework. After that, We monitored training progress, analyzed loss curves, and made adjustments to [90] the learning rate and batch size as needed.

We examine the trained model on the validation set and fine-tune the model parameters based on validation performance. Signal strength maps to each room using the recorded WiFi signal strength measurements, and we use the recorded GPS signals to obtain coarse location estimates for each room. In this step, we can predict

'spaceType' from objects of locations. And[91] GPS coordinates, WiFi readings, and space type help collectively[92]
indicate actual places.

a) Logistic Regression for Indoor Localization Logistic Regression is a statistical technique primarily
employed for binary classification tasks[7]. However, it
can be extended[93] to handle multi-class classification
tasks through a technique known as "One-vs-All."
or "One-vs-Rest." This approach involves training
multiple binary classifiers, each focused on distinguishing one class from the rest. In this project, We
use this technique to predict the likelihood of each
room's presence based on WiFi and GPS data.

• Training Multiple Classifiers(One-vs-All): For
each room in your dataset, train a separate
Logistic Regression classifier. Each classifier
learns to distinguish the signals and features

specific to its assigned room from signals in all
other rooms [5]. For each room, create a binary
target variable 1 for the particular room and 0
for all other rooms combined and train separate
Logistic Regression classifiers, each focusing on
distinguishing a specific space.

b) Naive Bayes Classification: The probabilistic
classifier Naive Bayes is straightforward and effective. It is used in our project[94]
to classify location

kinds according to input parameters such as GPS
locations, WiFi signal intensities, and object detection outcomes. Despite being more straightforward than other
classifiers, Naive Bayes works effectively in multi-class classification applications. Its probabilistic
methodology adds variation to the ensemble of
classifiers and increases the project's adaptability
by offering a simple and understandable way to
categorize locations.

c) Random Forest: Another crucial classifier for
our project is Random Forest. It is well-known for
its ensemble learning methodology, which combines many decision trees to provide predictions
as a group. Random Forest is a location identification system that can handle a broad range
of input information, such as GPS coordinates,
WiFi signal intensities and object detection results,
with high accuracy. Random Forest's ensemble
approach lessens overfitting and increases the resilience of the system[95].

3) Testing: We evaluate Localization Accuracy in this
phase. In this, We assess the accuracy of your indoor
localization system by comparing the predicted room
labels with ground truth. And also evaluate for an unseen
location. If the area does not exist in the dataset,
then[96] it can expect location type and neighborhood.

4) Web App: We also created a Web Application for
this project to collect data automatically. This web

The application stores images of GPS coordinates, calculates
elevation, and takes readings of WiFi on its own. This
web application also has a train and predict feature.
So that users can easily use this thing without any
difficulties.

VII. RESULTS

In a previous experiment, we compared the performance of
two different models for node classification: a graph convolutional network
(GCN) and an artificial neural network (ANN).
We used the COLD-topo map dataset with 22 classes and
trained both models to classify nodes into their respective
classes.
After training and testing the models, we found that the GCN
achieved an accuracy of 39.5%, while the ANN achieved an
accuracy of 30%. This [97] suggests that the GCN model was
better suited to this particular task of node classification and
was able to capture the relationships between nodes in the
graph more effectively than the ANN.
One possible reason for the superior performance of the GCN

8

Fig. 8: Dashboard

It was able to leverage the structure of the graph to
improve its predictions. The GCN uses the adjacency matrix
of the chart to propagate information between neighboring

nodes, allowing it to capture higher-order interactions between nodes. In contrast, the ANN treats each node as an independent data point and does not explicitly model the relationships between nodes. After that, in a new experiment,

Fig. 10: Accuracy Comparison

VIII. CONCLUSION

Fig. 9: Example of YOLOv8 Object Detection

YOLOv8 performs well at recognizing things inside photos of indoor space type. The accuracy of 84.62% The accuracy for indoor localization was achieved [98] by using different models for WiFi and GPS signals are 80-90%. Currently, for this dataset, Logistic Regression has 76%, Naive Bayes has 85%, and Random Forest has 92%. This level of accuracy indicates the system's ability to forecast the proper room based on WiFi signal strength and GPS context with the help of object detection. The end [99] result combines the results of YOLOv8 Object Recognition with multi-class classification of WiFi and GPS localization. This comprehensive strategy depends on the benefits of both modalities, improving the overall accuracy of indoor localization [11].

On the COLD dataset, which has 22 distinct labels, we examined the effectiveness of graph convolutional networks (GCNs) in this study to classify nodes. According to our findings, the GCN model ranked the nodes into their respective labels with an accuracy of 39.5%.

These findings indicate that GCNs are a potential method for

node classification tasks in challenging datasets like COLD,

where typical machine learning models have difficulty capturing connections

between nodes.

Our work underlines the need for more research to improve the

performance of GCNs and solve their weaknesses while also

demonstrating the promise of GCNs for node classification

tasks in challenging datasets like COLD. The project effectively integrates

YOLOv8's object identification capabilities

with Logistic Regression-based WiFi and GPS localization to

obtain overall accuracy that improves indoor placement. The

9

3) Handling Dynamic and Undefined Objects

We will investigate dynamic object-tracking methods to

account for objects that change position or appearance.

This [100] could involve employing techniques like object-tracking algorithms and

considering ways to track and localize these objects while maintaining overall [101]

accuracy.

R REFERENCES

Fig. 11: Different model's accuracy

Findings provide a spotlight on the potential and constraints

of multimodal indoor localization. YOLOv8's accuracy of

84.62% demonstrates its effectiveness in recognizing things within indoor room pictures. Accurate object identification improves spatial comprehension and contributes to indoor localization precision. The combined accuracy of 76.92% achieved through WiFi and GPS localization using Logistic Regression highlights their importance in indoor locations. Despite WiFi signal loss and the limitations of indoor GPS signals, these Logistic Regression approaches give useful information.

IX. F UTURE SCOPE

The project has a lot of potential for future development. Here are some possible areas for improvement:

1) Refining Model and Semantic Understanding

• Fine-Tuning Object Detection: We have investigated techniques to improve object detection accuracy, especially for objects currently being misidentified, [102] and fine-tune the YOLOv8 model on misclassified objects and update bounding box annotations.

• Semantic Object Relationships: We need to explore methods to establish semantic connections between detected objects and develop algorithms to infer relationships, positions, or interactions between objects for a richer spatial understanding.

2) WiFi Frequency Analysis and Contextual Insights

We need to analyze these frequencies to gain insights into areas with higher or lower WiFi activity. We have to develop techniques to create contextual WiFi signatures for rooms, combine signal strength with frequency count, and use these signatures to enhance the accuracy of WiFi-based localization.

1.	*many real-world Problems can only be described*	Passive voice misuse	Clarity
2.	*topological relations found by a robot navigating extensive office areas can be used*	Passive voice misuse	Clarity
3.	~~representation~~ → word, model, picture, presentation	Word choice	Engagement
4.	*To integrate the acquired geographical knowledge, resolve ambiguities, and provide predictions, such frameworks often use structured prediction algorithms.*	Unclear sentences	Clarity
5.	*To integrate the acquired geographical knowledge, resolve ambiguities, and provide predictions*	Misplaced words or phrases	Correctness
6.	*often*	Misplaced words or phrases	Correctness
7.	~~poses~~ → pose	Faulty subject-verb agreement	Correctness
8.	~~if they have~~ → without	Wordy sentences	Clarity
9.	*This*	Intricate text	Clarity
10.	*are included*	Passive voice misuse	Clarity
11.	*been observed*	Passive voice misuse	Clarity
12.	*are highlighted*	Passive voice misuse	Clarity
13.	*are explained*	Passive voice misuse	Clarity
14.	*is introduced*	Passive voice misuse	Clarity
15.	~~about~~ → in	Wrong or missing prepositions	Correctness
16.	~~dynamic~~ → Dynamic	Improper formatting	Correctness

17.	*YOLO's integration into the project gives the system the capacity to recognize and classify things that are exclusive to particular interior spaces.*	Unclear sentences	Clarity
18.	*By identifying things typical of specific locations, this semantic knowledge improves the system's ability to recognize and navigate interior surroundings.*	Unclear sentences	Clarity
19.	*is built*	Passive voice misuse	Clarity
20.	*xap*	Unknown words	Correctness
21.	*is built*	Passive voice misuse	Clarity
22.	*is shown*	Passive voice misuse	Clarity
23.	*are labeled*	Passive voice misuse	Clarity
24.	*are arranged*	Passive voice misuse	Clarity
25.	~~are applicable~~ → apply	Wordy sentences	Clarity
26.	~~specifically~~	Wordy sentences	Clarity
27.	robot's behavior	Wordy sentences	Clarity
28.	*be seen*	Passive voice misuse	Clarity
29.	*are created*	Passive voice misuse	Clarity
30.	*is assembled*	Passive voice misuse	Clarity
31.	*is decomposed*	Passive voice misuse	Clarity
32.	*For each decomposition and each sub-graph, we instantiate the corresponding template SPN, resulting in multiple SPNs sharing weights and structure.*	Unclear sentences	Clarity
33.	*are combined*	Passive voice misuse	Clarity

34.	~~In order to~~ → To	Wordy sentences	Clarity
35.	*kj*	Unknown words	Correctness
36.	*ckj*	Unknown words	Correctness
37.	~~ckj~~ → can	Misspelled words	Correctness
38.	*is built*	Passive voice misuse	Clarity
39.	*kj*	Unknown words	Correctness
40.	*be observed*	Passive voice misuse	Clarity
41.	*is used*	Passive voice misuse	Clarity
42.	~~own~~	Wordy sentences	Clarity
43.	*GCNs work by propagating information from a node's neighbors to update its own representation.*	Unclear sentences	Clarity
44.	*This*	Intricate text	Clarity
45.	*is achieved*	Passive voice misuse	Clarity
46.	*is applied*	Passive voice misuse	Clarity
47.	*is then passed*	Passive voice misuse	Clarity
48.	*are designed*	Passive voice misuse	Clarity
49.	~~to~~ → of	Wrong or missing prepositions	Correctness
50.	~~work by using~~ → use	Wordy sentences	Clarity
51.	~~hidden~~ → confidential, secret	Word choice	Engagement
52.	*be designed*	Passive voice misuse	Clarity
53.	*GNNs can be designed with different types of*	Unclear sentences	Clarity

	message-passing functions, such as convolutional or attention-based functions.		
54.	~~own~~	Wordy sentences	Clarity
55.	*YOLOv8 is the latest iteration in the YOLO family of detection models, which are known for their capabilities for joint detection and segmentation.*	Unclear sentences	Clarity
56.	~~Similar to~~ → Like	Wordy sentences	Clarity
57.	~~offers support for~~ → supports	Wordy sentences	Clarity
58.	*the strength of these signals can be measured*	Passive voice misuse	Clarity
59.	~~various~~	Wordy sentences	Clarity
60.	*WiFi frequency analysis provides a valuable layer of information that contributes to the overall indoor localization process.*	Unclear sentences	Clarity
61.	*This integrated approach enhances the accuracy of indoor positioning and enables the determination of the best-matching room based on a combination of modalities.*	Unclear sentences	Clarity
62.	*Global Positioning System (GPS) signals, initially designed for outdoor navigation, can still provide valuable context for indoor localization projects despite their limitations within enclosed spaces.*	Unclear sentences	Clarity
63.	*are transmitted*	Passive voice misuse	Clarity
64.	~~a network of~~	Wordy sentences	Clarity
65.	~~entirely~~	Wordy sentences	Clarity
66.	*This context can be integrated with*	Passive voice misuse	Clarity
67.	*be integrated*	Passive voice misuse	Clarity
68.	*These hybrid methods aim to mitigate the*	Unclear sentences	Clarity

	limitations of individual techniques and leverage the strengths of each data source.		
69.	~~signal distortions~~ → distortion	Wordy sentences	Clarity
70.	*Each area is annotated by a semantic place category, such as corridor, kitchen, or doorway.*	Passive voice misuse	Clarity
71.	*is ensured*	Passive voice misuse	Clarity
72.	*A broad representation of the interior areas under examination is ensured by the different angles and views from which these photographs are captured.*	Unclear sentences	Clarity
73.	*are captured*	Passive voice misuse	Clarity
74.	*By identifying things typical of specific locations, object detection improves accuracy and adaptability and provides semantic comprehension that helps the system identify and cope with interior environments.*	Unclear sentences	Clarity
75.	*is captured*	Passive voice misuse	Clarity
76.	~~, which~~ → . This	Hard-to-read text	Clarity
77.	*With the help of this module, we can get a list of WiFi networks that are available, measure them, and save the results in a dataset.*	Unclear sentences	Clarity
78.	~~In order to~~ → To	Wordy sentences	Clarity
79.	*are also captured*	Passive voice misuse	Clarity
80.	~~In order to~~ → To	Wordy sentences	Clarity
81.	*been modified*	Passive voice misuse	Clarity
82.	~~rObot~~ → robot	Confused words	Correctness
83.	*was used*	Passive voice misuse	Clarity

84.	~~serving~~	Wordy sentences	Clarity
85.	*is annotated*	Passive voice misuse	Clarity
86.	~~, but as~~ → . Still, as	Hard-to-read text	Clarity
87.	*are built*	Passive voice misuse	Clarity
88.	~~are used to~~	Wordy sentences	Clarity
89.	~~each and every~~ → every, each	Wordy sentences	Clarity
90.	~~made adjustments to~~ → adjusted	Wordy sentences	Clarity
91.	~~And~~	Conjunction use	Correctness
92.	~~collectively~~	Wordy sentences	Clarity
93.	*be extended*	Passive voice misuse	Clarity
94.	*It is used in our project to classify location kinds according to input parameters such as GPS locations, WiFi signal intensities, and object detection outcomes.*	Unclear sentences	Clarity
95.	system's resilience	Wordy sentences	Clarity
96.	~~then~~	Wordy sentences	Clarity
97.	*This*	Intricate text	Clarity
98.	*The accuracy for indoor localization was achieved*	Passive voice misuse	Clarity
99.	~~end~~	Wordy sentences	Clarity
100.	*This*	Intricate text	Clarity
101.	~~overall~~	Wordy sentences	Clarity
102.	~~, and~~ → . We	Hard-to-read text	Clarity