Interdisciplinary Journal of Information, Knowledge, and Management

An Official Publication of the Informing Science Institute InformingScience.org

IJIKM.org

Volume 19, 2024

REVOLUTIONIZING AUTONOMOUS PARKING: GNN-POWERED SLOT DETECTION FOR ENHANCED EFFICIENCY

U. Vignesh*	Vellore Institute of Technology, Chennai, Tamil Nadu	vignesh.u@vit.ac.in
Tushar Moolchandani	Vellore Institute of Technology, Chennai, Tamil Nadu	Tushar.moolchandani2020@vitstudent.ac.in

* Corresponding author

ABSTRACT

Aim/Purpose	Accurate detection of vacant parking spaces is crucial for autonomous parking. Deep learning, particularly Graph Neural Networks (GNNs), holds promise for addressing the challenges of diverse parking lot appearances and complex visual environments. Our GNN-based approach leverages the spatial layout of detected marking points in around-view images to learn robust feature representations that are resilient to occlusions and lighting variations. We demonstrate significant accuracy improvements on benchmark datasets compared to existing methods, showcasing the effectiveness of our GNN-based solution. Further research is needed to explore the scalability and generalizability of this approach in real-world scenarios and to consider the potential ethical implications of autonomous parking technologies.
Background	GNNs offer a number of advantages over traditional parking spot detection methods. Unlike methods that treat objects as discrete entities, GNNs may leverage the inherent connections among parking markers (lines, dots) inside an image. This ability to exploit spatial connections leads to more accurate parking space detection, even in challenging scenarios with shifting illumination. Real-time applications are another area where GNNs exhibit promise, which is critical for autonomous vehicles. Their ability to intuitively understand linkages across marking sites may further simplify the process compared to traditional deep-learning approaches that need complex feature development. Furthermore, the proposed GNN model streamlines parking space recognition by potentially combining slot inference and marking point recognition in a single

Accepting Editor Narongsak Sukma │Received: February 28, 2024│ Revised: April 24, May 10, June 1, June 11, June 13, 2024 │ Accepted: June 28, 2024.

Cite as: Vignesh, U., & Moolchandani, T. (2024). Revolutionizing autonomous parking: GNN-powered slot detection for enhanced efficiency. *Interdisciplinary Journal of Information, Knowledge, and Management, 19,* Article 19. https://doi.org/10.28945/5334

step. All things considered, GNNs present a viable method for obtaining stronger and more precise parking slot recognition, opening the door for autonomous car self-parking technology developments.

Methodology	The proposed research introduces a novel, end-to-end trainable method for parking slot detection using bird's-eye images and GNNs. The approach involves a two-stage process. First, a marking-point detector network is employed to identify potential parking markers, extracting features such as confidence scores and positions. After refining these detections, a marking-point encoder network extracts and embeds location and appearance information. The enhanced data is then loaded into a fully linked network, with each node representing a marker. An attentional GNN is then utilized to leverage the spatial relationships between neighbors, allowing for selective information aggregation and capturing intricate interactions. Finally, a dedicated entrance line discriminator network, trained on GNN outputs, classifies pairs of markers as potential entry lines based on learned node attributes. This multi-stage approach, evaluated on benchmark datasets, aims to achieve robust and accurate parking slot detection even in diverse and challenging environments.
Contribution	The present study makes a significant contribution to the parking slot detection domain by introducing an attentional GNN-based approach that capitalizes on the spatial relationships between marking points for enhanced robustness. Additionally, the paper offers a fully trainable end-to-end model that eliminates the need for manual post-processing, thereby streamlining the process. Furthermore, the study reduces training costs by dispensing with the need for detailed annotations of marking point properties, thereby making it more accessible and cost-effective.
Findings	The goal of this research is to present a unique approach to parking space recognition using GNNs and bird's-eye photos. The study's findings demonstrated significant improvements over earlier algorithms, with accuracy on par with the state-of-the-art DMPR-PS method. Moreover, the suggested method provides a fully trainable solution with less reliance on manually specified rules and more economical training needs. One crucial component of this approach is the GNN's performance. By making use of the spatial correlations between marking locations, the GNN delivers greater accuracy and recall than a completely linked baseline. The GNN successfully learns discriminative features by separating paired marking points (creating parking spots) from unpaired ones, according to further analysis using cosine similarity. There are restrictions, though, especially where there are unclear markings. Successful parking slot identification in various circumstances proves the recommended method's usefulness, with occasional failures in poor visibility conditions. Future work addresses these limitations and explores adapting the model to different image formats (e.g., side-view) and scenarios without relying on prior entry line information. An ablation study is conducted to investigate the impact of different backbone architectures on image feature extraction. The results reveal that VGG16 is optimal for balancing accuracy and real-time processing requirements.
Recommendations for Practitioners	Developers of parking systems are encouraged to incorporate GNN-based techniques into their autonomous parking systems, as these methods exhibit enhanced accuracy and robustness when handling a wide range of parking scenarios. Furthermore, attention mechanisms within deep learning models can

	provide significant advantages for tasks that involve spatial relationships and contextual information in other vision-based applications.
Recommendations for Researchers	Further research is necessary to assess the effectiveness of GNN-based methods in real-world situations. To obtain accurate results, it is important to employ large-scale datasets that include diverse lighting conditions, parking layouts, and vehicle types. Incorporating semantic information such as parking signs and lane markings into GNN models can enhance their ability to interpret and understand context. Moreover, it is crucial to address ethical concerns, including privacy, potential biases, and responsible deployment, in the development of autonomous parking technologies.
Impact on Society	Optimized utilization of parking spaces can help cities manage parking resources efficiently, thereby reducing traffic congestion and fuel consumption. Automating parking processes can also enhance accessibility and provide safer and more convenient parking experiences, especially for individuals with disabilities. The development of dependable parking capabilities for autonomous vehicles can also contribute to smoother traffic flow, potentially reducing accidents and positively impacting society.
Future Research	Developing and optimizing graph neural network-based models for real-time deployment in autonomous vehicles with limited resources is a critical objective. Investigating the integration of GNNs with other deep learning techniques for multi-modal parking slot detection, radar, and other sensors is essential for enhancing the understanding of the environment. Lastly, it is crucial to develop explainable AI methods to elucidate the decision-making processes of GNN models in parking slot detection, ensuring fairness, transparency, and responsible utilization of this technology.

Keywords	parking slot detection, bird's-eye images, graph neural networks, attention mechanism, autonomous vehicles, deep learning, real-time processing

INTRODUCTION

The growing number of passenger cars in major cities, fuelled by societal advancements, presents a significant challenge: finding parking spaces in congested neighborhoods. Studies indicate that over half of drivers experience frustration during parking searches. Furthermore, parking maneuvers contribute to a substantial portion of vehicle accidents. A number of converging reasons are driving up the requirement for autonomous parking. First, the growing number of automobiles, especially in cities, makes it difficult to find parking spots, which causes traffic jams and annoyance. One possible answer is autonomous parking technology, which allows cars to park themselves more accurately and effectively while making the most of available space. Second, parking can be dangerous and increase the risk of accidents, particularly in confined locations. With sensors and sophisticated algorithms at their disposal, autonomous parking systems have the potential to drastically lower these hazards by assuming full control of the parking procedure and guaranteeing precise and safe operations. Furthermore, drivers can benefit greatly from automated parking in terms of ease.

Envision not having to deal with the anxiety and annoyance of finding a place to park and navigating through confined places; this is particularly helpful for people who are physically unable to do these things or for inclement weather. Future-focused, autonomous parking is essential to the idea of smart cities and self-driving transit networks. It facilitates the smooth integration of automobiles with ride-sharing programs and other mobility options, encouraging economical transportation and resource management. Lastly, automated parking may help reduce emissions and enhance air quality in metropolitan areas by increasing parking efficiency and reducing time spent looking for a spot. Essentially, the need for autonomous parking arises from its ability to tackle several issues, opening the door to a future of mobility that is safer, more convenient, and environmentally friendly.

To address this issue, the parking assist system (PAS) has been developed with three key components: object identification designation, route planning, and parking guiding. The most crucial component of PAS is the object position designation, which is responsible for precisely identifying available parking spaces. Accurate parking-slot identification is crucial for Autonomous Valet (AV) parking, which is a significant application for autonomous vehicles. Compared to manual parking, AV parking offers several advantages. Precise parking maneuvers and safer control minimize the risk of scratches and collisions caused by human error. Parking, even for experienced drivers, can be a demanding task. It requires repeated back-and-forth movements for precise positioning while maintaining constant situational awareness to avoid collisions with objects and pedestrians. This demands high driving skills and reaction times. Autonomous vehicles necessitate a robust parking system. Traditionally, vision-based parking slot recognition techniques aimed to precisely identify a slot from its surroundings. However, these methods often serve as supplementary features in mass-produced cars, only functional until the vehicle is physically driven near a potential parking space. Modern AVs require the ability to independently locate parking opportunities. Accurate parking slot detection plays a critical role in this process. False positives in parking space identification must be avoided, as they can lead to accidents or parking violations.

The detection of vacant parking spaces using surrounding view imagery must fulfill two essential criteria to be deemed practical and useful: it must distinguish between different parking spaces and remain resilient in complex visual environments. To achieve this objective, Suhr and Jung (2012, 2013, 2014, 2018) introduced several marking point-based parking spot recognition techniques. These methods utilize inherent characteristics, which are sensitive to illumination changes, to identify marking points. Due to the exceptional feature extraction capabilities of the Deep Convolutional Neural Network (DCNN) (Zhang et al., 2018), the accuracy of parking space identification is significantly enhanced by this technique. The traditional method of parking spot detection involves several time-consuming procedures that do not provide information about the occupancy status of the parking space. To overcome this limitation, a comprehensive DCNN was introduced as a complementary solution (Zinelli et al., 2019), allowing for simultaneous automated identification and categorization of parking spaces. However, this approach relies on the Faster R-CNN baseline and is therefore unable to meet real-time requirements.

The majority of object recognition techniques currently in use assume that each object in an image operates independently of the others. However, this assumption is inaccurate in situations where the objects, or nodes, in a picture are connected. For instance, in an Around-View Monitor (AWM) image, marking points are connected naturally to one another, and these connections can be used to detect parking spaces more accurately. The interactions between marking points can be represented using Graph Neural Networks (GNN) (Kipf & Welling, 2016), which take into account the marking points' graphical structure. The success of the Transformer (Vaswani et al., 2017) has inspired us to propose incorporating an attention mechanism into GNN to detect parking spaces. By using the attention mechanism, graphical nodes can assign varying weights to different nodes within a neighborhood, allowing for the implicit learning of connection information between marking points. The graphical neural network-based parking-slot detection model proposed by us combines parking-slot inference and marking-point detection in a single step, unlike earlier CNN-based approaches (Heimberger et al., 2017; Huang et al., 2019).

Images of parking spaces with different characteristics and outside circumstances are included in our collection. We also include severe data sets when a portion of the parking space is blocked by different impediments. To minimize the likelihood of a false-positive issue with parking spot recognition, the dataset also includes data samples of a non-parking area category with an appearance similar to the parking spot. Advances in material science are required to assure component longevity as autonomous cars continue to evolve. Research is underway to increase the fracture toughness of thermosetting polyester polymers used in vehicle components (Nusyirwan et al., 2023).

The overall architecture of the suggested parking-spot identification model based on GNNs is shown in Figure 1. The three main components of the proposed method are the entry line discriminator, graph feature aggregation, and graph feature encoder. In particular, the detector for the marking points initially detects the markings before extracting the deep properties associated with all of them. The location data and deep features are then fused using the feature aggregation network, and the surrounding data is aggregated using a graphical neural network. The entryway line distinction network processes the encoded discriminative features present in the marking point combinations to determine whether or not they constitute an entry line.

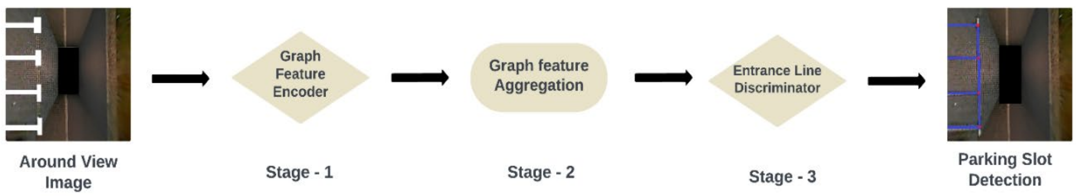

Figure 1. Proposed architecture

Figure 2 showcases seven examples of surround-view images that illustrate the substantial variation in visual patterns of parking lots in real life. In (a) and (b), the parking spots are perpendicular and photographed in indoor parking lots; (c) and (d) display parallel parking spots, with (d) showing visible damages to the parking lines; (e) was taken on a wet day, while (f) was captured at night under a streetlight. In (g), the parking lines are obscured by the intense shadow cast by neighboring trees. The varied ground materials are also evident in these images. In (h), the shot was obtained when a person was strolling alongside a parking lot.

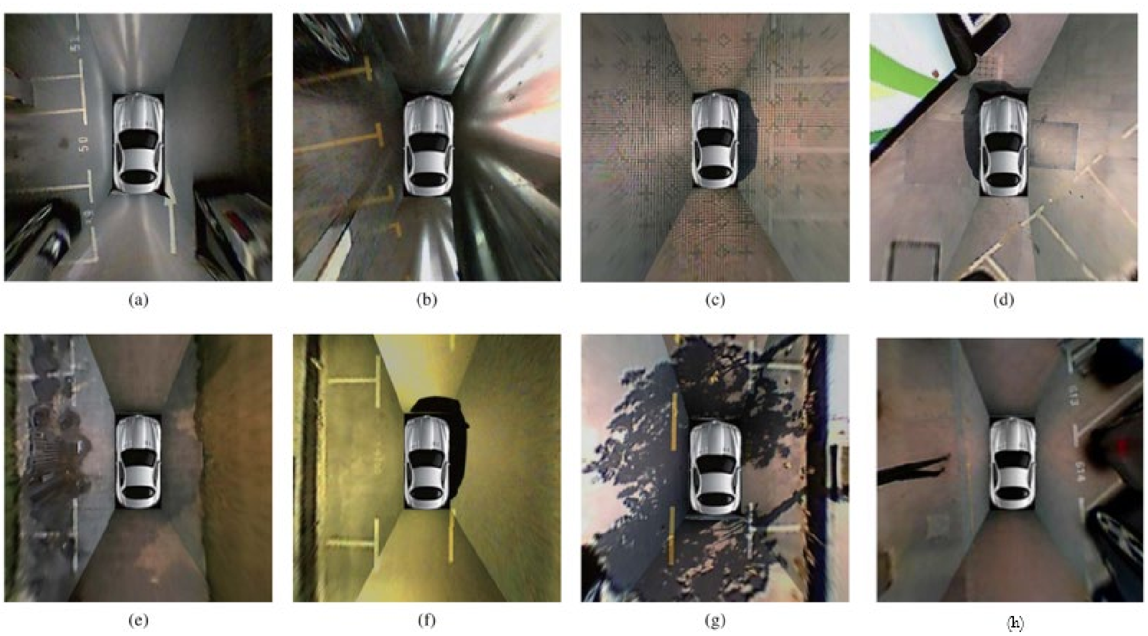

Figure 2. Examples of different types of parking slots

The parking space identification techniques used today frequently have drawbacks. Conventional methods ignore the vital relationships between parking marks that define a space, treating each item separately. GNNs can overcome this issue by making use of these correlations to achieve more accurate recognition, particularly in intricate layouts. Furthermore, the performance of existing systems may be affected by differences in sunlight or dirt. GNNs may be more resistant to these external influences since they concentrate on the underlying structure of marks using a graph representation. Moreover, whilst GNNs show promise for implicit learning of connections, possibly streamlining the process, certain deep learning algorithms need sophisticated feature building. Finally, the detection of markers and the inference of the parking spot may need distinct processes in classical approaches. In certain situations, the suggested GNN model could simplify this procedure by condensing them into a single phase. Ultimately, by taking advantage of geographical correlations, perhaps improving resilience, and simplifying processing, GNNs overcome the drawbacks of existing techniques and open the door to more precise and dependable parking space identification, especially for autonomous cars (Figure 3).

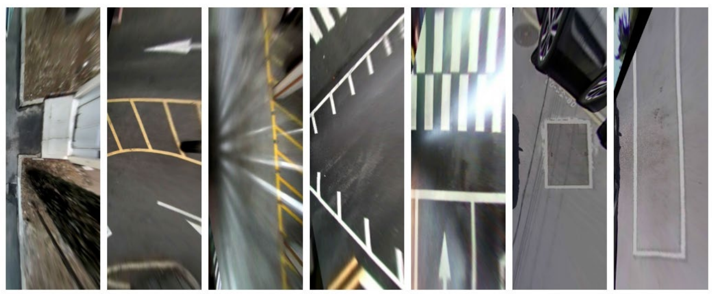

Figure 3. Example of non-parking space data samples Commons

The following are some of the highlights of the work:

- The purpose of the method is to improve the accuracy of parking-slot recognition by utilizing a graphical neural network to process the surrounding information between marking points in the 360-degree image. This information is represented in a graph-structured format.
- The suggested technique decreases training costs by eliminating the requirement for extensive remarks, including identifying point direction and form.
- We provide a fully trainable end-to-end parking-slot recognition technique.

LITERATURE SURVEY

PARKING-SLOT DETECTION USING TRADITIONAL METHOD

Traditional approaches for detecting parking spaces have relied on sensor-based techniques, with ultrasonic sensors being the most popular. However, Suhr and Jung (2014, 2016, 2018) previously noted the following. First, they often require a specific sensor setup, which may not be feasible or cost-effective in all parking lot configurations. Second, they struggle with environmental variations like rain, snow, or bright sunlight, which can affect sensor readings and result in inaccurate occupancy detection. Additionally, ultrasonic sensors may have difficulty distinguishing between a parked vehicle and other objects within the parking space, such as shopping carts or low-hanging foliage.

Another sensor-based approach utilizes computer vision techniques, such as L. Li, et al. (2017) research that classified the occupancy state of parking spaces using a gray histogram, which analyses the distribution of pixel intensities within the image. Although effective for basic detection, this method may not be robust enough to handle complex parking lot environments with cluttered backgrounds or variations in lighting. Similarly, studies like Lee and Seo's (2016) explored stereo vision, which utilizes two cameras to capture depth information and detect small objects like bicycles or motorcycles within a parking space. However, this approach requires specialized hardware and can be computationally expensive.

Machine learning algorithms have also been employed in sensor-based parking space detection, as demonstrated by research by Amarappa and Sathyanarayana (2010) and Dalal and Triggs (2005). These methods utilized the Histogram of Gradients (HOG) and a Support Vector Machine (SVM) for feature extraction and classification, respectively. HOG captures the local variations in pixel intensity within an image, while an SVM learns a classification model based on labeled training data. This approach can be effective for identifying parking spaces with clear lane markings. Unlike sensor-based methods, vision-based techniques rely on cameras to capture images of the parking space for detection. These techniques are particularly well-suited for Advanced Valet Parking (AVP) systems, where cameras are often already integrated into the vehicle. However, these cameras typically capture a perspective view of the parking environment, which can be susceptible to variations in lighting, perspective, and background clutter. To address this issue, additional image processing is required when dealing with bird's-eye view imagery obtained from strategically placed high-resolution cameras or drones. These distortions are often caused by the fisheye lens commonly used in bird's-eye view cameras. Once corrected, the image becomes suitable for further analysis using various algorithms like those discussed previously.

Vignesh and Elakya (2024) present a framework for a blockchain-driven management system that may be used for a decentralized parking management system; in contrast, a system powered by a GNN offers more accuracy and uses fewer resources. Parking space markings play a crucial role in vision-based detection methods. Typically, four painted lines define a standard parking space, although curbs or fences may replace some of the lines in specific scenarios. The line separating adjacent parking spaces is referred to as the separating line, while the area where a vehicle enters or exits the parking space is called the entry line. These lines are essential for algorithms to identify the location and dimensions of individual parking spaces within the image. It is worth noting that while vision-based techniques offer some advantages in terms of computational efficiency, they still suffer from sensitivity to environmental changes and are heavily reliant on the quality of the training data. Previous research has explored Local Binary Pattern (LBP) (Ojala et al., 2002) for extracting features and a Support Vector Machine (SVM) for classification (Rianto et al., 2018). However, both methods fall prey to the limitations of conventional feature extraction and classification techniques, which can be susceptible to variations in lighting, perspective, and background clutter.

The objective of this research endeavor is to develop a novel approach to parking space detection by drawing upon the unique advantages of aerial bird's-eye views while simultaneously addressing the limitations of existing sensor-based and vision-based methods. This innovative system aims to achieve accurate, real-time identification of available parking spaces through the integration of advanced image processing techniques and powerful machine learning algorithms specifically designed for bird's-eye view imagery. The proposed method seeks to leverage the strengths of existing methods while overcoming their challenges in order to provide a more effective solution for parking space detection.

Vision-based methods offer an alternative, exploiting camera images, particularly from around-view systems (AVM). These techniques fall into three primary categories:

a) Segmentation-based approaches, such as DFNet (Jiang et al., 2019), have been developed to address the challenges of imbalanced classes and blurred boundary segmentation in panoramic images for semantic segmentation. These approaches focus on detecting lane markings and parking slots for self-driving and automatic parking applications. The proposed unified structure for parking space detection using semantic segmentation (Jang & Sunwoo, 2018) and vertical grid encoding aims to address the challenges of precise and reliable parking space detection in various conditions. Another approach, VH-HFCN (Wu et al., 2018), is based on a classification technique to identify marker lines and parking spaces in a panoramic surround image. By presenting a highly fused convolutional network (HFCN) model for accurately and reliably extracting linear characteristics, this technique tackles the difficulties associated with autonomous parking. The model is tested on a publicly available PSV dataset. These techniques can identify parking spots by segmenting the picture, but they can be computationally costly and noise-sensitive.

b) DeepPS, an example of a point-based approach, is a parking-slot detection approach that employs a DCNN (Zhang et al., 2018) to accurately and efficiently detect and locate parking spaces close to the car that is indicated by regular line segments. The findings show that DeepPS performs better in parking-slot recognition than other cutting-edge techniques, with high recall and accuracy rates, which qualifies it for application in time-sensitive autonomous parking systems. With sensors installed on off-the-shelf cars, the system can automatically identify different kinds of parking space marks according to different light conditions. Its recall and precision are higher than those of earlier approaches. A complete off-the-shelf vehicle-mounted sensor system has been created (Suhr & Jung, 2018) that is capable of identifying various parking space markers under all lighting circumstances.

c) The method of Parking Slot Detection (Q. Li et al., 2018) consists of two primary stages: recognition of lines and entry detection of parking spaces. A feature-clustering technique is used in the initial step of splitting line identification to identify separating lines in a variety of ground and lighting situations. In the second step, the parking space entry identification, typical parallel and perpendicular parking spaces are identified using the Multiview interface training technique. A parking-slot detection algorithm utilizes the Directional-DBSCAN (Lee et al., 2016) to address the challenges of detecting parking slots in long-range surround-view images. This method proposes a novel clustering algorithm for line-segment detection and slot pattern recognition.

To create lines that serve as parking space markers, the Sobel filter and probabilistic Hough transform in Surround view-based parking lot detection and tracking (Hamada et al., 2015) are used. However, the Hough transform is susceptible to the effects of bright illumination and deep shadows. To address noise challenges in Autonomous Parking Vision System (C. Wang et al., 2014), line-based techniques based on methods like the Radon transform have been proposed. The Radon transform is utilized for straight-line detection, and clustering and filtering based on parking slot characteristics are incorporated to enhance noise tolerance and robustness. This approach aims to improve overall parking slot detection performance in noisy environments.

PARKING-SLOT DETECTION USING DEEP-LEARNING

Our research on existing literature indicates a growing interest in applying deep learning techniques to parking slot detection, particularly in the context of autonomous valet parking (AVP) systems. One of the pioneering works in this domain is DeepPS (Heimberger et al., 2017), which utilizes Convolutional Neural Networks (CNNs) for marking point detection and matching. This approach achieves a high degree of accuracy in identifying parking spaces. However, it requires the deployment of two separate CNNs – one for marking point detection and another for matching – which can introduce additional complexity to the system. DeepPS contrasts with other deep learning-based methods, such as YOLO (Redmon et al., 2016), SSD (Liu et al., 2016), and their successors, YOLOv2 to v4 (Liu et al., 2016). These techniques typically rely on object detection neural networks (Ma et al., 2022) to directly identify marking spots or other critical elements within the parking area, such as the junction points of parking lines. This approach streamlines the process compared to DeepPS' two-step approach.

Building upon this foundation, L. Wang et al. (2023) propose a technique for parking slot identification that leverages the YOLOv2 object detection architecture. This method demonstrates the effectiveness of YOLOv2 in directly identifying parking spaces within an image. Similarly, VPS-Net (W.

Li et al., 2020) employs the YOLOv3 object detection algorithm to achieve empty parking space detection. These single-stage methods offer a potential advantage in terms of computational efficiency compared to the two-stage approach used by DeepPS.

While DeepPS has demonstrated exceptional performance in diverse environmental conditions, showcasing its robustness, another two-step method known as DMPR-PS (Huang et al., 2019) warrants examination. This approach utilizes a CNN-based technique to initially predict the direction, location, and shape of identifying marks within the parking area. Subsequently, the method applies manually defined geometric criteria to filter and match paired marking points. Although DMPR-PS achieves reasonable accuracy, it requires time-consuming post-processing steps to refine the identified marking points. This drawback is addressed by proposing a method that eliminates the need for such post-processing. This approach achieves direct marking point coordinate prediction using a dedicated CNN, demonstrating superior efficiency and improved generalization capabilities compared to DMPR-PS. This advancement highlights the ongoing efforts within the research community to optimize parking slot detection methods for accuracy and efficiency.

The aforementioned studies provide valuable insights into the potential of deep learning for parking slot detection. However, it is crucial to acknowledge certain limitations. DeepPS, while achieving high accuracy, suffers from its two-stage architecture, which can be computationally expensive. Techniques like YOLOv2 and VPS-Net offer a more streamlined approach but may be susceptible to performance degradation in complex or poorly lit environments where accurate marking point identification becomes challenging. Furthermore, methods like DMPR-PS, despite their initial success, introduce an additional layer of complexity through the need for manual geometric criteria in the matching stage.

Vignesh and Ratnakumar (2024) proposed analyses of the influenza virus by comparing different deep-learning algorithms to bring out the best in terms of accuracy for detection and prediction. Beyond these limitations, there's a growing recognition of the potential shortcomings associated with relying solely on marking point detection for parking slot identification. In scenarios where parking spaces are not clearly delineated by markings, such as in unpaved lots or temporary parking arrangements, these methods may struggle to achieve accurate detection. Additionally, factors like weather conditions, shadows, or parked vehicles partially obscuring markings can further hinder performance.

Researchers are actively exploring alternative approaches to address these limitations. One promising avenue involves leveraging semantic segmentation techniques. Semantic segmentation networks excel at classifying each pixel within an image, assigning it a specific label based on the content it represents. In the context of parking slot detection, a semantic segmentation network could be trained to classify pixels belonging to a parking space, a vehicle occupying a space, or the background area. This approach offers the potential to overcome the dependence on clearly defined marking points by enabling the system to identify parking spaces based on their overall shape and context within the image. Hamada et al. (2015) exemplify this approach, proposing a deep learning framework that incorporates both semantic segmentation and object detection for robust parking slot detection.

Another emerging area of exploration involves utilizing 3D point cloud data for parking slot detection. Light Detection and Ranging (LiDAR) sensors are capable of generating 3D point cloud representations of the surrounding environment. By processing this data, it is possible to extract features relevant to parking spaces, such as the height and location of objects within the parking lot. This approach offers potential advantages in situations where visual information is limited due to poor lighting or occlusions. Vaswani et al. (2017) demonstrate the feasibility of this approach by proposing a method that leverages 3D point clouds captured by LiDAR sensors for parking space detection.

Graphical Convolutional Networks

Graph Neural Networks (GNN) (Scarselli et al., 2009) have gained prominence for their ability to model global relationships. There are two types of GNN models:

a) The spectral model is a new method for applying CNNs to high-dimensional irregular environments represented by graphs (Defferrard et al., 2016). The model can acquire regional, stable, and structural graph characteristics by designing rapid specialized convolutional filtering techniques based on spectral graph theory. This method generalizes CNNs to signals defined on non-grid domains (Bruna et al., 2014) and proposes two constructions for deep neural networks on graphs: spatial construction and spectral construction. These constructions aim to achieve efficient architectures with a reduced number of parameters independent of the input size.

Though spectral CNNs provide a potential way to get around the drawbacks of conventional CNNs in a top-down approach, it's vital to recognize that research in this field is still in its early stages. More research is needed to improve graph-building methods, create reliable pipelines for data collecting and preprocessing, and maximize Spectral CNN topologies designed for parking space recognition from an ocular perspective.

The literature review concludes by showing a thriving field of study centered on using deep learning to find parking spaces. Although current approaches provide insightful information, they have drawbacks. Among the alternatives, spectral CNNs seem like a strong contender to get beyond these drawbacks and accomplish reliable parking spot identification in the framework of the bird's eye view. With its emphasis on the intrinsic distortions of the bird's eye view and its capacity to extract relational characteristics from the parking lot graph, spectral CNNs provide a special chance to progress the area of automated valet parking (AVP) systems.

In addition to the technical aspects, investigating Spectral CNNs for parking space identification from aerial photography is consistent with a wider movement in deep learning models toward interpretability and explainability. Conventional deep learning models are sometimes referred to as "black boxes," as it is unclear how decisions are made within. Concerns about safety may arise from this lack of transparency in AVP systems and other sensitive applications. Spectral CNNs provide a possible route toward better interpretable models by utilizing the well-defined mathematical framework of spectral graph theory. Developers may learn a lot about how the system makes decisions by discovering how the model uses the connections between nodes in the graph structure for parking spot recognition. This improved comprehension can help foster trust and guarantee the dependable and safe operation of AVP systems.

Additionally, Spectral CNNs may be used for purposes other than only detecting parking slots. More functionality inside AVP systems becomes possible with the capacity to analyze linkages and connectivity patterns within a graph structure. Envision an AVP system that maximizes its navigation path within the parking lot and detects open spots. Through an examination of the graph structure, the system was able to determine the best path to an open area while accounting for variables such as pedestrian traffic, obstructions, and flow. This would improve AVP systems' effectiveness and security even further.

In summary, a promising new direction for the development of AVP systems is the investigation of Spectral CNNs for parking spot recognition from an ocular perspective. By utilizing the intrinsic benefits of Spectral CNNs – their resistance to distortions, capacity to capture relationship information, and potential for interpretability – this strategy presents a chance to get beyond the drawbacks of current techniques. Additionally, examining relationships in the parking lot graph more broadly opens the door to features like improved navigation path planning in AVP systems. Spectral CNN research is still in its early stages, but when combined with bird's eye view technology, it might completely change parking and provide a more practical, effective, and ultimately safe commuting environment.

b) In non-spectral approaches, one alternative is the Graph Attention Networks (GATs) (Veličković et al., 2018), a cutting-edge neural network architecture specifically designed for graph-structured data. By adding disguised self-attentional levels, our strategy overcomes the constraints of previous graph-convolution approaches and delivers state-of-the-art outcomes in several graph-based benchmark tasks. Another possibility is a comprehensive technique for semi-supervised training on graph-structured databases (Kipf & Welling, 2016), which employs a powerful form of CNN. Experiments conducted on reference systems and an information graph database have demonstrated better classification efficiency and accuracy using this technique.

Although parking slot recognition from an ocular perspective may be achieved with Spectral CNNs and GATs, a comparison study is required to identify the best method. The relative benefits and drawbacks of each strategy should be investigated in research under various parking lot layouts, lighting circumstances, and meteorological variables. The comparison will help choose the best deep learning architecture for this particular use case.

Beyond Spectral CNNs and GATs, the literature also explores alternative strategies for graph-structured data analysis. Kipf and Welling (2016) propose a comprehensive technique for semi-supervised training of a powerful form of CNN on graph-structured databases. This approach demonstrates promising results for classification tasks on information graph databases. While not directly focused on parking slot detection, it highlights the potential for further exploration of various deep learning architectures specifically designed for graph data within the context of AVP systems.

In conclusion, deep learning research for bird's-eye view parking spot recognition goes beyond Spectral CNNs. Another interesting strategy is to use network Attention Networks (GATs), which can efficiently manage sparse network topologies and prioritize relevant interactions within the parking lot graph. Although Spectral CNNs and GATs are promising research directions, further investigation is needed to fine-tune their use for this particular job. This entails evaluating the effectiveness of the two methods, refining GAT attention mechanisms, and investigating the possibilities of further deep learning architectures made for graph data. As research advances, the best deep-learning architecture for robust parking slots will be chosen based on a thorough knowledge of the advantages and disadvantages of these different approaches.

When considering the broader landscape of deep learning for AVP systems, it is important to acknowledge the potential benefits of combining different deep learning architectures. Ensemble learning, a machine learning technique that combines the predictions of multiple models, offers an intriguing possibility. By leveraging the complementary strengths of Spectral CNNs, GATs, or potentially other graph-based deep learning architectures, an ensemble model could achieve superior performance compared to any single approach. For instance, a combination of a Spectral CNN, adept at capturing global structural features, and a GAT, excelling at prioritizing informative relationships, could potentially lead to more robust and accurate parking slot detection, even in challenging scenarios.

Beyond the previously explored methods like Spectral CNNs and GATs, GNNs emerge as another powerful category of deep learning architectures specifically designed for processing graph-structured data. Similar to Spectral CNNs and GATs, GNNs can effectively analyze the relationships and connections between various elements within the parking lot, represented as a graph in the bird's eye view framework. This approach offers significant potential for robust parking slot detection by considering the spatial context of the parking environment.

The core strength of GNNs lies in their ability to iteratively aggregate information from neighboring nodes within the graph. Imagine a parking space represented as a node in the graph. GNNs can "learn" from the features associated with neighboring nodes, such as curbs, lane markings, and potentially even other parked vehicles. Through this iterative process, GNNs can progressively build a comprehensive understanding of the relationships between elements within the parking lot, ultimately enabling them to identify valid parking spaces. The proposed approach for parking slot detection using GNNs aligns with the principles outlined in previous works like Veličković et al. (2018) and Kipf and Welling (2016). These studies highlight the effectiveness of spatial convolution within GNNs for aggregating node features in local regions of a graph. This resonates with the task of parking slot detection, where understanding the spatial relationships between a potential parking space and its surrounding elements is crucial for accurate identification.

The suitability of GNNs for visual applications has already been demonstrated in various domains, including segmentation of point clouds (Y. Wang et al., 2019), picture matching (Sarlin et al., 2020), and human posture prediction (Jin et al., 2020). Adapting these successful applications to the visual domain of parking lot analysis through a bird's eye view is a promising avenue for research. By leveraging GNNs' ability to process spatial relationships within images, the system can effectively identify parking spaces based on their context and location relative to other elements in the parking lot.

In conclusion, the exploration of deep learning for parking slot detection using a bird's eye view extends beyond Spectral CNNs and GATs. GNNs offer a compelling alternative approach with their ability to iteratively learn and aggregate information from neighboring nodes within the parking lot graph. This approach aligns with existing research on applying spatial convolutions within GNNs for graph analysis tasks. Furthermore, the success of GNNs in various visual applications like point cloud segmentation paves the way for their adaptation to parking slot detection using bird's eye view data. By addressing challenges related to GNN architecture selection, feature engineering, and data collection, GNNs hold significant promise for robust parking slot detection within AVP systems, ultimately contributing to a more efficient and intelligent transportation future.

The method we proposed leverages graph neural networks (GNNs), offering several advantages. Unlike traditional methods, GNNs inherently capture the spatial relationships between marking points, which is crucial for identifying valid parking spaces. Compared to DeepPS, the proposed single GNN-based approach eliminates the need for separate steps, improving efficiency. Furthermore, GNNs demonstrate superior performance in modeling intricate relationships compared to simpler deep-learning object detection approaches, potentially leading to higher accuracy and robustness.

METHODOLOGY

OVERVIEW OF THE PROPOSED METHOD

The suggested technique enhances the accuracy of parking slot detection by incorporating GNNs into the model. GNNs excel at capturing the inherent relationships between data points, which makes them ideal for tasks where understanding the connections between elements is essential. In parking slot detection, these connections exist between the marking points that define the boundaries of each parking space. Conventional parking slot identification relied on manually defined geometric rules to establish these connections. However, these rules can be inflexible and struggle to adapt to variations in parking lot layouts. The proposed method aims to address this limitation by utilizing GNNs. This method takes into consideration four key marking locations (P1, P2, P3, and P4) within a parking lot. A critical challenge is identifying the correct pairing of marking points (P1 and P2) that define the entry line for a specific parking space (as illustrated in Figure 1). The order of these points is crucial, as they define the anticlockwise orientation of the four-point polygon representing the entry line.

The proposed method uses a two-step process to tackle this challenge. The first step involves a "marked-point encoder" system that extracts the unique characteristics of each marking point from the image features. These characteristics could include color, size, and location within the image. The second step involves a "marking-point analyzer" that leverages a GNN to analyze the relationships between these marking points based on their extracted characteristics and identified locations within the image. This GNN essentially "learns" the inherent connections between the points, effectively determining which point pairs likely represent the entry line for a parking space. The proposed method is designed to utilize only American English, adhering strictly to its spelling, specific terms, and phrases.

The following text has been rephrased to use a formal tone while maintaining the original content and structure. No changes have been made to the citations, references, or in-line citations. The numbers in the text remain unchanged. The "entryway line discriminator network" acts as the final judge in determining whether the identified pairs constitute a valid entry line. Each node in the fully connected graph utilized by the GNN represents a potential labeling location. This approach offers several key advantages. It eliminates the need for manually defined geometric rules, making the model more adaptable to diverse parking lot configurations. Additionally, the entire process is "end-to-end trainable," meaning the model learns all stages – from feature extraction to relationship analysis and entry line identification – simultaneously during training on a large dataset of labeled parking lot images.

This allows the model to continuously refine its understanding of allocating the vehicle dynamically and space allocation for the vehicle. In the case of parking the vehicle on a narrow or wide path, it is a complex task using the GNN technique; extracting the pathway of the neighboring nodes makes parking clearly labeled to ignore the space complexity. To make in all parking lot layouts and relationships between marking points, leading to superior parking slot detection accuracy.

GRAPH FEATURE ENCODER

The system incorporates a crucial element known as the marking-point detector, which identifies possible marking points within an image. This detector produces a feature map with dimensions S×S×3, comprising three channels. The first channel, denoted by "c," signifies the confidence score for each pixel as a potential marking point. A higher value in this channel indicates a greater probability of a marking point existing at that location. The remaining two channels, "m" and "n," represent the predicted horizontal and vertical coordinates of the marking point, respectively. These channels provide a vital initial estimate for the marking point's position within the image.

One potential issue with the raw output from the marking-point detector is that it may contain redundant information. To address this, the technique of non-maximum suppression (NMS) is implemented. NMS operates as a filtering mechanism, evaluating the confidence scores and selecting only the most likely marking points. This process helps eliminate duplicate detections and yields a more focused set of potential marking points, typically represented by N. These N points serve as the most promising candidates for further analysis. After the initial detection and filtering stages, the system employs a marking-point feature encoder. This encoder consists of four convolutional layers arranged sequentially. These layers successively extract high-level features from the image regions surrounding the identified marking points. The convolutional layers work by applying learnable filters to the image data, ultimately generating a refined feature map of size S×S×64. This feature map captures the characteristics of the areas surrounding the potential marking points.

The system employs a precise method for identifying N marking points within an image. This involves determining the coordinates (m, n) of each potential marking point. These coordinates represent the location of the marking point within the image. To extract the most relevant features, a bilinear interpolation technique is utilized on a previously generated S×S×64 feature map. Bilinear interpolation estimates the value at a specific point within the feature map based on the values of its surrounding pixels, which allows for a more accurate representation of the features associated with each individual marking point. The system typically fixes the size of the feature maps to a specific value, often set to 16 (S = 16), to ensure consistency and simplify processing steps. To further enrich the representation of each marking point, a Multi-Layer Perceptron (MLP) is employed. This MLP acts as a dedicated network tasked with learning a compressed representation of the positional information for each marking point. It takes the coordinates (m, n) of a marking point as input and transforms them into a high-dimensional vector. This vector captures the relative position of the marking point within the image.

$$v_i = f_i + MLP(m_i, n_i)$$

The system then merges these positionally encoded vectors with the appearance features extracted earlier (F) using an element-wise addition operation. This process combines the spatial information (where the marking point is located) with the visual characteristics (what the marking point looks like) to create a richer and more informative representation v_i for each marking point. Here, v_i represents the enhanced characteristic vector for the ith marking point. This combined representation allows the system to more effectively reason about the relationships between different marking points later within the processing pipeline.

GRAPH FEATURE AGGREGATION

The proposed approach efficiently integrates marking-point characteristics by utilizing GNNs, which are composed of a hierarchical layered architecture and a fully connected graph representation. Each layer refreshes the image representation by accumulating inputs from every edge, including vertices. By incorporating an attention mechanism in message aggregation, nodes can focus on relevant information from neighboring nodes, which enhances the model's ability to represent intricate interactions in the graph.

Additionally, the inclusion of a multi-head attention mechanism allows for the concurrent evaluation of several neighboring information aspects, which increases the model's expressivity. This technology demonstrates its effectiveness and flexibility in processing graph-structured data and shows promise for tasks such as node classification, edge prediction, and graph construction.

To improve the performance of the GNN forecasting in a time series-based prediction for extra complicity, the prediction of neighboring nodes on time should be improved. Time evaluation for every edge interaction with its neighboring gives information on how to prolong the edge occupancy of the space, like a history data for each edge node that has to be maintained. It also helps to enhance the model's availability in the future.

ENTRANCE LINE DISCRIMINATOR

This research proposes a novel approach for improving the accuracy of entrance line detection within parking lot layouts. This method utilizes a combination of attentional graph neural networks (GNNs) and discriminative models. GNNs are highly effective at learning complex relationships within graph-structured data. In this context, the parking lot layout can be represented as a graph, with nodes corresponding to marking points (such as corners or lane dividers) and edges representing the connections between these points. By processing this graph structure, the GNN can effectively capture the spatial relationships between different marking points. However, for the specific task of entrance line detection, the GNN requires additional guidance to distinguish between various types of connections. This is where the discriminator comes in, as it is trained to specifically identify entrance lines.

To achieve this, the system utilizes learned node attributes extracted from the attentional GNN. These attributes capture the information gleaned by the GNN for each marking point within the parking lot layout. To construct meaningful input features for the discriminator, the system employs a creative technique that concatenates the marking-point feature pairs. Essentially, it forms pairs of marking points based on their connections within the graph and combines the corresponding learned node attributes from the GNN for each pair. These marking-point feature pairs are then fed into the discriminator network for analysis.

The discriminator network is designed to be resistant to overfitting, a common challenge in machine learning, where the model performs well on training data but struggles with unseen examples. To address this issue, the discriminator network includes multilayer perceptron (MLP) layers. These layers function as building blocks for the network, allowing it to learn increasingly complex relationships within the data. However, to prevent the network from simply memorizing the training data and sacrificing its ability to generalize to new scenarios, a technique called dropout regularization was utilized during the training process. Dropout regularization operates by randomly dropping a specified percentage of neurons within MLP layers during training. This approach forces the network to learn from a broader range of features within the data and prevents it from relying excessively heavily on specific connections.

The aim of this method is to produce a comprehensive matrix that includes all marking-point pairs and their associated probabilities of being entrance lines, which serves as a crucial component of the parking-slot analysis framework. The proposed method benefits from the integration of learned node attributes from the GNN and the discriminatory power of the trained network, resulting in a marked improvement in the accuracy of entrance line detection within parking slot layouts. This approach effectively combines the strengths of both attentional GNNs and discriminative models, thereby providing a more robust and dependable system for examining parking layouts and identifying accessible parking spaces.

LOSS FUNCTION

We consider the losses for the marking-point and entry-line forecasts. We use the mean square error for the marking-point forecast and the binary cross-entropy loss (bi-entropy loss) for the entry-line forecast.

$$loss = \mu_1 \, . \, loss_{point} + \mu_2 \, . \, loss_{line}$$

where the prediction loss for the marking point is denoted by $loss_{point}$, & $loss_{line}$, the entry line forecast loss μ_1 & μ_2, the balancing weights for both the losses.

The marking-point prediction of the loss function is the squared errors combined sum between the forecasted of the $S \times S$ cells and associated ground-truths since the marking-point detector's output layer is divided into a grid of $S \times S$.

$$loss_{point} = \frac{1}{S^2} \sum_{i=1}^{S^2} \{ \left(d_i - \hat{d}_i \right)^2 + \mathbb{1}_i [(m_i - \widehat{m}_i)^2 + (n_i - \hat{n}_i)^2] \}$$

where (m_i, n_i) signifies the forecast placements of marking points on the layout and $\hat{d}_i$ is the confidence of cell i. The sign $(\hat{})$ represents the ground-truths of the labeling spots. $\mathbb{1}_i = 1$ if (m_i, n_i) is the forecasted ground-truth marking point and 0 otherwise.

The loss of the entry line detector is as follows:

$$loss_{line} = \frac{1}{P^2} \left(- \sum_{m=1}^{P} \sum_{n=1}^{P} \hat{l}_{mn} \log l_{mn} \right)$$

where l_{mn}, the expected likelihood of the m-th and n-th marking points that make up the parking space's entry. The marking points' related ground-truths are indicated by symbols designated with$(\hat{})$.

EXPERIMENTAL RESULTS AND DISCUSSIONS

DATASETS DESCRIPTION

This study investigates the usefulness of using Around View Monitor (AWM) pictures for parking space detection. The chosen dataset, ps2.0, has a comprehensive collection of photographs created expressly for this purpose. It contains 14,077 photographs, divided into 9,827 training shots and 4,250 testing images. To provide a rigorous evaluation, we use a subset of this data, including 7,780 photos for training and 2,200 images for testing. The ps2.0 dataset covers a wide range of climatic variables, including parking scenarios in both indoor and outdoor locations. This variation is critical for training and assessing the parking slot identification system's effectiveness in real-world scenarios.

It is critical to understand the underlying technology that powers AVM systems. These systems generally comprise four to six wide-angle cameras placed strategically around the vehicle, each capturing a unique perspective. A surround-view synthesis module processes the acquired wide-angle photos, intelligently stitching them together to provide a full 360° panoramic vision of the automobile. This technology, which includes the development of the surround-view system and the calibration of the wide-angle optics, adheres to well-established automotive industry standards. Our project intends to construct a strong parking slot identification system capable of properly detecting unoccupied slots in varied parking situations using AVM pictures acquired from numerous angles. The data in Tables 1 and 2 illustrate the comparison results.

Table 1. Precision and recall on the ps2.0 dataset

Method	Precision	Recall
L. Wang et al. (2023)	98.49%	60.33%
Suhr and Jung (2018)	96.75%	76.22%
PSD_L	98.80%	90.96%
DMPR-PS	99.42%	99.37%
Hamada et al. (2015)	98.45%	61.37%
FCN-baseline	97.70%	97.85%
GNN	**99.68%**	**99.51%**

$$Precision = \frac{true\ positive}{true\ positive + false\ positives}$$

$$Recall = \frac{true\ positive}{true\ positive + false\ negatives}$$

The frequency of accurate parking spot detections is known as the true positive, whereas the number of erroneous positives is known as the false positive. Conversely, the false negative indicates how frequently detections go unnoticed. The results of the experiments reveal that the proposed approach is very robust and reliable, especially in the presence of obstacles like other cars and shadows.

IMPLEMENTATIONS DESCRIPTION

Our methodology for training focuses on optimizing the proposed system for precise parking slot detection. We employ the Adam optimizer, renowned for its effectiveness in handling intricate datasets such as the one used for training. The training process extends over 200 epochs, signifying the number of times the entire dataset is passed through the network. This iterative approach enables the model to progressively enhance its capacity to recognize parking spaces.

To expedite the training process, we harness the computational power of a single Nvidia RTX 3060 12 GB GPU, a powerful graphics card well-suited for deep learning tasks. Additionally, we utilize PyTorch, a prominent deep learning framework that provides the necessary tools and functionalities for constructing and training the model. The training initiates with a learning rate of 0.001 for a batch size of 16 images. The learning rate essentially governs the tempo at which the model updates its internal parameters based on the training data.

A lower learning rate, like the one selected here, ensures a more gradual learning process, minimizing the likelihood of the model overfitting the training data. To ensure consistency across training runs, we set the momentum parameters μ_1 & μ_2 to 100.0 and 1.0, respectively. These parameters influence how the optimizer integrates past learning updates when adjusting the model's internal weights.

Lastly, the training process incorporates the VGG16 architecture for feature extraction. This pre-trained convolutional neural network serves as a potent tool for extracting pertinent features from the bird's eye view images employed for training. The overall model architecture consists of a graphical neural network with four distinct heads and three layers (Figure 4). These components collaborate to analyze the extracted features and ultimately classify whether a parking space in the image is vacant or occupied. This in-depth look at the training process highlights the meticulous approach taken to develop a robust and reliable system for autonomous valet parking slot detection.

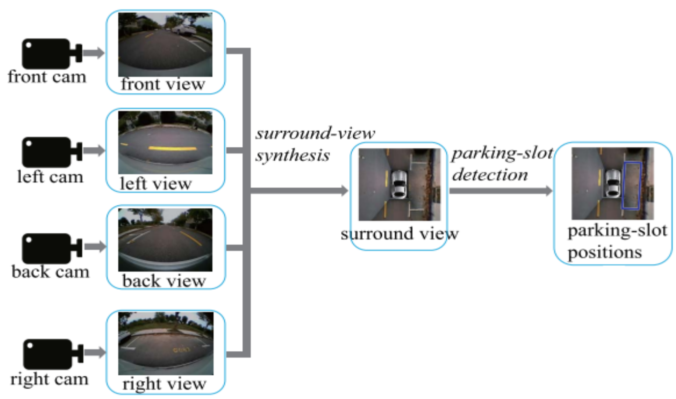

Figure 4. The concluded picture has an aspect ratio of 1000 x 1000, and the four source pictures (in an automobile frame, a single pixel symbolizes one centimeter by one centimeter, giving a 10m×10m real plane area)

RESULTS AND DISCUSSION

The statistical findings of the ps2.0 test set are shown in Table 1, which clearly shows how much better the suggested approach operates than the three traditional parking-slot detection algorithms. This finding suggests that learning-based systems are capable of learning more discriminative features, particularly in complex visual settings. Additionally, the method achieves comparable accuracy and recall to DMPR-PS, a cutting-edge learning-based methodology. Using a two-stage approach, DMPR-PS first predicts the location, orientation, and form of marking points using a convolutional neural network, and then it recognizes parking places by using geometric criteria that are manually established. In contrast, the suggested method uses a graphical neural network to instantly and completely infer the parking spot by treating the marking points as data arranged in graphs and only forecasting their positions. Moreover, the proposed technique needs less accurate observations of the marking spots' form and position, which lowers training costs and improves practical applicability.

Effectiveness of graphical neural network

This study aims to evaluate the effectiveness of the graphical neural network developed for parking-slot identification performance. To achieve this, a comparison is made with a fully connected neural network, represented by FCN-baseline, which utilizes naive complete connection layers. Table 1 shows that the suggested technique performs 0.74% and 0.58% better in accuracy and recall rate, respectively than the FCN baseline. The FCN baseline fails to account for the interrelationship between different marking points and treats them as independent data points. In contrast, the graphical neural network takes advantage of the connections between multiple marking points to achieve precise parking slot recognition.

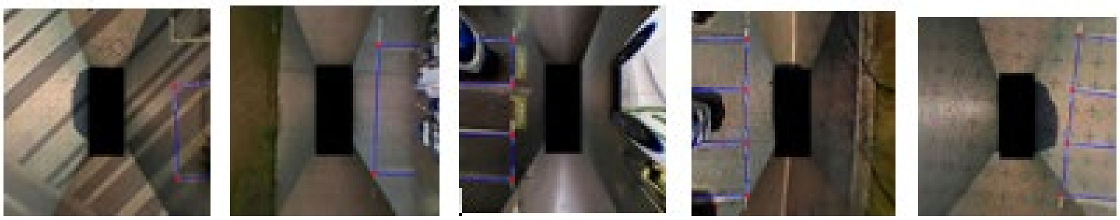

Figure 5. The positions of the recognized markings are depicted in red, while the entryway and lateral lines appear in blue

To corroborate the effectiveness of the graphical neural network, the FCN baseline is compared with the standard cosine correspondence that is computed over the characteristics of every pair of marked markers. The average cosine similarity score of the graphical neural network model is 0.6671, which is higher than that of the FCN baseline, which is 0.3168. This means that the characteristics between the two combined marking points in the graphical neural network are more comparable compared to those of the FCN baseline. The comparison of GNN-before and GNN-after shows that after the graphical neural network, the properties of coupled marking points do not change. This implies that the matched marking-point similarity is retained by the graphical neural network. Following the graphical neural network, there is a drop in resemblance between the unmatched marking-point characteristics from 0.0936 to -0.1527, indicating that the graphical neural network makes the unattached marking points more different. This proves that the designed graphical neural network can detect link information, which is the knowledge that the properties of unmatched marking points differ from those of paired marking points, which constitute parking spots. This improves parking lot recognition by enabling the neural network to learn more discriminative characteristics.

Table 2. Precision and recall with different backbones

Backbone	Precision	Recall	Speed (ms)
RessNet50	97.54%	97.54%	20.1
DarkNet19	98.30%	98.71%	19.1
VGG16	**99.68%**	**99.51%**	24.3

In Figure 5, we provide evidence of the effectiveness of the proposed approach in various parking-slot detection scenarios. Regardless of the conditions, the method demonstrates exceptional performance, achieving enhanced levels of accuracy and robustness, particularly in complex situations. Yet there have been reports of failure, especially if the marked places are not visible. This is due to the fact that the strategy we proposed is based on detecting marking positions. Inferring parking places from partially visible marker points is a tough and exciting topic that we intend to address in future improvements. Working with different image formats, such as side-view photos, may require additional effort. Additionally, the current model requires entry line information and annotations of the marking-point placements as it is designed for around-view photos.

CASE STUDY

In urban areas, the automated parking system is applicable for availability and on-time efficiency as the service user is able to book the parking slot in a fully connected network (FCN) baseline. By getting the update on the parking slot before the user reaches the desired location, the update message containing the time series-based availability can also be fetched to the user message. Each and every time, the edge node's marking point is updated to ensure effectiveness in the space availability slot. If a slot is unavailable in the user-selected location, it suggests some other space availability based on nearby node accessibility.

ABLATION STUDY

Different backbone

This study examines the influence of backbone architectures on image feature extraction. As shown in Table 2, the VGG16-based backbone outperforms other backbones in terms of precision-recall ratings. With an average detection speed of 40 Hz and a processing time of 24.3 milliseconds per image on an Nvidia RTX 3060 12GB GPU, the VGG16 backbone meets the real-time requirement. The ResNet18 backbone has the fastest inference speed, at 15.2 milliseconds per image. The multi-state models have parking-slot detection speeds of 59.50 ms and 41.29 ms, respectively.

Loss weights

Moreover, this research evaluates various weight combinations for entry-line prediction loss and marking-point prediction loss. The proposed model uses a higher weight for marking-point prediction loss. This is due to the fact that the entrance line prediction problem is not as difficult as the marking-point prediction problem.

CONCLUSION

One promising area of future research is the development and optimization of GNN-based models that are specifically designed for real-time deployment in autonomous vehicles with limited computational resources. While CNNs excel at image recognition tasks, their high computational demands may not be suitable for resource-constrained environments. GNNs, on the other hand, offer a promising alternative for modeling relationships between objects in a scene, which aligns well with the task of identifying parking spaces and their relationships to each other within the bird's eye view. By exploring GNN architectures specifically tailored for parking slot detection, research can pave the way for a lightweight and efficient system suitable for onboard processing in resource-constrained autonomous vehicles. This would be a significant step towards practical implementation of the Autonomous Vehicle Parking (AVP) system in real-world scenarios.

Future research could explore the potential benefits of combining GNNs with other deep-learning methodologies for multi-modal parking slot detection. While the bird's eye view offers valuable information, it may be advantageous to integrate it with data collected from other sensors on the autonomous vehicle. For instance, light detection and ranging (LiDAR) sensors can provide detailed three-dimensional information about the environment, while radar sensors can excel at detecting objects in low-visibility conditions. By combining GNNs with deep learning models trained on data from these additional sensors, the system can gain a more comprehensive understanding of the parking environment. For example, a GNN could leverage the bird's eye view to identify potential parking spaces, while LiDAR data could be used to confirm the dimensions and absence of obstacles within those spaces. Similarly, radar data could be incorporated to enhance detection accuracy in poor weather conditions. A multi-modal approach has the potential to significantly improve the robustness and reliability of the parking slot detection system, ensuring accurate identification of vacant spaces across diverse environmental conditions.

The development of explainable AI (XAI) methods is of paramount importance for ensuring the ethical, responsible, and transparent use of this technology. Considering the complexity of GNNs for parking slot detection, it is crucial to understand their decision-making processes. By integrating XAI techniques, researchers can develop methods that explain the rationale behind the GNN model's identification of a specific space as vacant or occupied. This level of transparency is essential for building trust in the AVP system and ensuring that it operates without any form of bias. Moreover, XAI can be employed to pinpoint potential vulnerabilities in the model, allowing researchers to address these issues and enhance the system's dependability. In summary, investigating GNN optimization, multi-modal sensor integration, and the development of XAI holds immense potential for advancing the bird's eye parking slot detection system toward a robust, reliable, and trustworthy solution for autonomous valet parking and beyond.

Furthermore, future research could investigate the integration of GNNs with other deep-learning techniques for multi-modal parking slot detection. While the bird's eye view provides valuable information, it can be advantageous to combine it with data from other sensors on the autonomous vehicle. LiDAR sensors can provide detailed 3D information about the environment, while radar sensors can excel at detecting objects in low-visibility conditions. By integrating GNNs with deep learning models trained on data from these additional sensors, the system can gain a richer understanding of the parking environment. For example, a GNN could leverage the bird's eye view to identify potential parking spaces, while LiDAR data could be used to confirm the dimensions and absence of obstacles within those spaces. Similarly, radar data could be incorporated to enhance detection accuracy in poor weather conditions. A multi-modal approach has the potential to significantly improve the robustness and reliability of the parking slot detection system, ensuring accurate identification of vacant spaces across diverse environmental conditions.

It also helps to conduct research on other domains like GIS (Geographic Information System). It offers increased use space potential and large landscape service for end-to-end node connections. Logistics can use a route map with a GNN technique to identify the potential parking space for the purposes of planning vehicles in logistics, and it also increases the efficiency of using the space effectively. Technically, the high optimization of parking slot detection is applicable in a precise way in logistic planning.

REFERENCES

Amarappa, S., & Sathyanarayana, S. (2010). Data classification using Support Vector Machine (SVM), a simplified approach. *International Journal of Electronics and Computer Science Engineering*, *3*(4), 435-445. https://pzs.dstu.dp.ua/DataMining/svm/bibl/10.1.1.683.210.pdf

Bruna, J., Zaremba, W., Szlam, A., & LeCun, Y. (2014). *Spectral networks and locally connected networks on graphs*. ArXiv. https://doi.org/10.48550/arxiv.1312.6203

Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA*, 886-893. https://doi.org/10.1109/cvpr.2005.177

Defferrard, M., Bresson, X., & Vandergheynst, P. (2016). *Convolutional neural networks on graphs with fast localized spectral filtering*. ArXiv. https://doi.org/10.48550/arXiv.1606.09375

Hamada, K., Hu, Z., Fan, M., & Chen, H. (2015, June). Surround view based parking lot detection and tracking. *Proceedings of the IEEE Intelligent Vehicles Symposium, Seoul, Korea (South)*, 1106-1111. https://ieeexplore.ieee.org/document/7225832

Heimberger, M., Horgan, J., Hughes, C., McDonald, J., & Yogamani, S. (2017). Computer vision in automated parking systems: Design, implementation and challenges. *Image and Vision Computing*, *68*, 88-101. https://doi.org/10.1016/j.imavis.2017.07.002

Huang, J., Zhang, L., Shen, Y., Zhang, H., Zhao, S., & Yang, Y. (2019, July). DMPR-PS: A novel approach for parking-slot detection using directional marking-point regression. *Proceedings of the IEEE International Conference on Multimedia and Expo, Shanghai, China*, 212-217. https://doi.org/10.1109/icme.2019.00045

Jang, C., & Sunwoo, M. (2018). Semantic segmentation-based parking space detection with standalone around view monitoring system. *Machine Vision and Applications*, *30*, 309-319. https://doi.org/10.1007/s00138-018-0986-z

Jiang, W., Wu, Y., Guan, L., & Zhao, J. (2019, May). DFNet: Semantic segmentation on panoramic images with dynamic loss weights and residual fusion block. *Proceedings of the International Conference on Robotics and Automation, Montreal, QC, Canada*, 5887-5892. https://doi.org/10.1109/icra.2019.8794476

Jin, S., Liu, W., Xie, E., Wang, W., Qian, C., Ouyang, W., & Luo, P. (2020). Differentiable hierarchical graph grouping for multi-person pose estimation. In A. Vedaldi, H. Bischof, T. Brox, & J. M. Frahm (Eds.), *Computer vision – ECCV 2020* (pp. 718-734). Springer. https://doi.org/10.1007/978-3-030-58571-6_42

Kipf, T. N., & Welling, M. (2016). *Semi-supervised classification with graph convolutional networks.* ArXiv. https://doi.org/10.48550/arxiv.1609.02907

Lee, S., Hyeon, D., Park, G., Baek, I.-J., Kim, S.-W., & Seo, S.-W. (2016, June). Directional-DBSCAN: Parking-slot detection using a clustering method in around-view monitoring system. *Proceedings of the IEEE Intelligent Vehicles Symposium, Gothenburg, Sweden*, 349-354. https://doi.org/10.1109/ivs.2016.7535409

Lee, S., & Seo, S.-W. (2016). Available parking slot recognition based on slot context analysis. *IET Intelligent Transport Systems*, *10*(9), 594-604. https://doi.org/10.1049/iet-its.2015.0226

Li, L., Li, C., Zhang, T., Guo, T., & Miao, Z. (2017). Automatic parking slot detection based on around view monitor (AVM) systems. *Proceedings of the 9th International Conference on Wireless Communications and Signal Processing, Nanjing, China*. https://doi.org/10.1109/WCSP.2017.8170903

Li, Q., Lin, C., & Zhao, Y. (2018). Geometric features-based parking slot detection. *Sensors*, *18*(9), 2821. https://doi.org/10.3390/s18092821

Li, W., Cao, L., Yan, L., Li, C., Feng, X., & Zhao, P. (2020). Vacant parking slot detection in the around view image based on deep learning. *Sensors*, *20*(7), 2138. https://doi.org/10.3390/s20072138

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Computer vision – ECCV 2016* (pp. 21-37). Springer. https://doi.org/10.1007/978-3-319-46448-0_2

Ma, Y., Liu, Y., Shao, S., Zhao, J., & Tang, J. (2022). Review of research on vision-based parking space detection method. *International Journal of Web Services Research*, *19*(1), 1-25. https://doi.org/10.4018/ijwsr.304061

Nusyirwan, N., Abiem, P., Hairul, A., Hendery, D., Eka, S., Eka, S., & Nanda, I. P. (2023). Methods for increasing fracture toughness of thermosetting polyester polymers with vinyl ester mixtures as raw materials for automotive components. *Indian Journal of Engineering*, *20*, e20ije1648. https://doi.org/10.54905/disssi/v20i53/e20ije1648

Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(7), 971-987. https://doi.org/10.1109/tpami.2002.1017623

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016, June). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA*, 779-788. https://doi.org/10.1109/cvpr.2016.91

Rianto, D., Erwin, I. M., Prakasa, E., & Herlan, H. (2018, November). Parking slot identification using local binary pattern and support vector machine. *Proceedings of the International Conference on Computer, Control, Informatics and its Applications, Tangerang, Indonesia,* 129-133. https://doi.org/10.1109/IC3INA.2018.8629530

Sarlin, P.-E., Detone, D., Malisiewicz, T., & Rabinovich, A. (2020). *SuperGlue: Learning feature matching with graph neural networks.* ArXiv. https://doi.org/10.48550/arXiv.1911.11763

Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2009). The Graph Neural Network Model. *IEEE Transactions on Neural Networks*, *20*(1), 61-80. https://doi.org/10.1109/tnn.2008.2005605

Suhr, J. K., & Jung, H. G. (2012, September). Fully-automatic recognition of various parking slot markings in Around View Monitor (AVM) image sequences. *Proceedings of the 15th International IEEE Conference on Intelligent Transportation Systems, Anchorage, AK, USA,* 1294-1299 https://doi.org/10.1109/ITSC.2012.6338615

Suhr, J. K., & Jung, H. G. (2013). Full-automatic recognition of various parking slot markings using a hierarchical tree structure. *Optical Engineering*, *52*(3), 37203. https://doi.org/10.1117/1.oe.52.3.037203

Suhr, J. K., & Jung, H. G. (2014). Sensor fusion-based vacant parking slot detection and tracking. *IEEE Transactions on Intelligent Transportation Systems*, *15*(1), 21-36. https://doi.org/10.1109/tits.2013.2272100

Suhr, J. K., & Jung, H. G. (2016). Automatic parking space detection and tracking for underground and indoor environments. *IEEE Transactions on Industrial Electronics*, *63*(9), 5687-5698. https://doi.org/10.1109/tie.2016.2558480

Suhr, J. K., & Jung, H. G. (2018). A universal vacant parking slot recognition system using sensors mounted on off-the-shelf vehicles. *Sensors*, *18*(4), 1213. https://doi.org/10.3390/s18041213

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention is all you need.* ArXiv. https://doi.org/10.48550/arxiv.1706.03762

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2018). *Graph attention networks.* ArXiv. https://doi.org/10.48550/arxiv.1710.10903

Vignesh, U., & Elakya, R. (2024). Blockchain-enabled supply chain management for revolutionizing transparency, security, and efficiency. In U. Vignesh, M. M., & R. Doshi (Eds.), *Bio-inspired optimization techniques in blockchain systems* (pp. 101-113). IGI Global. https://doi.org/10.4018/979-8-3693-1131-8.ch005

Vignesh, U., & Ratnakumar, R. (2024). Optimization techniques for influenza prediction in biological expert systems. In U. Vignesh, M. M., & R. Doshi (Eds.), *Bio-inspired optimization techniques in blockchain systems* (pp. 192-206). IGI Global. https://doi.org/10.4018/979-8-3693-1131-8.ch010

Wang, C., Zhang, H., Yang, M., Wang, X., Ye, L., & Guo, C. (2014). Automatic parking based on a bird's eye view vision system. *Advances in Mechanical Engineering*, *6*. https://doi.org/10.1155/2014/847406

Wang, L., Musabini, A., Leonet, C., Benmokhtar, R., Breheret, A., Yedes, C., Bürger, F., Boulay, T., & Perrotton, X. (2023, October). Holistic parking slot detection with polygon-shaped representations. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Detroit, MI, USA*, 5797-5803. https://doi.org/10.1109/IROS55552.2023.10342486

Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., & Solomon, J. M. (2019). Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics*, *38*(5), Article 146. https://doi.org/10.1145/3326362

Wu, Y., Yang, T., Zhao, J., Guan, L., & Jiang, W. (2018). *VH-HFCN based parking slot and lane markings segmentation on panoramic surround view.* arXiv. https://doi.org/10.48550/arxiv.1804.07027

Zhang, L., Huang, J., Li, X., & Xiong, L. (2018). Vision-based parking-slot detection: A DCNN-based approach and a large-scale benchmark dataset. *IEEE Transactions on Image Processing*, *27*(11), 5350-5364. https://doi.org/10.1109/tip.2018.2857407

Zinelli, A., Musto, L., & Pizzati, F. (2019, June). A deep-learning approach for parking slot detection on surround-view images. *Proceedings of the IEEE Intelligent Vehicles Symposium, Paris, France*, 683-688. https://doi.org/10.1109/IVS.2019.8813777

AUTHORS

Dr. U. Vignesh is currently an Assistant Professor, Senior Grade 2 in the School of Computer Science and Engineering, Vellore Institute of Technology (VIT) - Chennai. Prior to his recent appointment at the VIT, he was a Post-Doctoral Fellow at the National Institute of Technology (NIT), Trichy – India. Dr. Vignesh received his undergraduate degree in B.Tech (IT), his M.Tech (IT) degree from Anna University, Chennai, and his PhD in Computer Science and Engineering from VIT University - Chennai. Dr. Vignesh published several papers in preferred Journals, patents, and book chapters and participated in a range of forums on computer science, social science, etc. He also presented various academic and research-based papers at several national and international conferences. His research activities are currently twofold: while the first research activity is set to explore the developmental role that society needs with technology such as Artificial Intelligence, the second major research theme that he is pursuing is focused on bioinformatics and data mining.

Tushar Moolchandani is currently pursuing a Bachelor of Technology in Computer Science and Engineering with a specialization in Artificial Intelligence and Machine Learning. His enthusiasm for artificial intelligence has led him to actively work in the field. This research builds upon his interest in exploring the applications of Graph Neural Networks for real-world tasks.