Parking Lot Spot Object Detection

Kasra Abdi
kma190003@utdallas.edu

Jeremy Alexander
jja210005@utdallas.edu

Omar Hussain
omh190000@utdallas.edu

Jesse Incoom
jxi210002@utdallas.edu

1. Abstract

Our project aims to address the parking challenges faced by students at the University of Texas at Dallas, a predominantly commuter school. With a significant number of students driving to campus daily, finding an available parking spot often becomes time-consuming and frustrating. To alleviate this issue, we developed a real-time parking space detection system leveraging the YOLO (You Only Look Once) object detection framework.

We utilized the PKLot dataset, which contains labeled images of parking spaces under various weather conditions and camera angles, as the foundation for training our model. To enhance model robustness, we applied data augmentation techniques such as adjusting brightness, contrast, and rotation using Albumentations. The YOLOv11n model was then fine-tuned for detecting parking spaces as either "empty" or "occupied."

Our testing process included evaluating the system on validation images and real-world data from Google Maps of Lot T on the UT Dallas campus. We measured performance using standard metrics, including precision, recall, F1-score, and accuracy. The results demonstrated high reliability, with our model effectively identifying parking availability while maintaining a balanced trade-off between precision and recall.

This project not only highlights the application of advanced object detection models to real-world problems but also underscores the potential for improving parking efficiency, reducing search times, and fostering sustainable practices by minimizing vehicle idling. Through this work, we aim to enhance the overall commuter experience and pave the way for scalable and responsive parking management systems.

2. Introduction

UT Dallas is known as a commuter school, which means a large proportion of students commute to school. These students don't live on campus, and therefore need a place to park when they come to campus. However, the parking pass system at UT Dallas means there is no guarantee of getting a parking spot in any given lot. A commuter may need to search through many different parking lots and structures to find a spot that corresponds to their specific pass. This can take a significant amount of time, and may lead these students to being late for class. The lack of an efficient parking management system exacerbates the issue of time wastage for UT Dallas commuters. This challenge not only affects students' punctuality but also contributes to unnecessary fuel consumption and environmental pollution caused by prolonged vehicle idling. By introducing a real-time parking space detection system, we aim to alleviate these issues, improve the overall commuter experience, and foster sustainable practices.

3. Related Work

We took advantage of a few related works to design our project. Redmon et al. [4] was the landmark paper that most of our project was based on, which established the YOLO model. Ogawa et al. [3] provided a good example of what our work could look like, and we took inspiration from the design and used some of the architecture in our own project. Sharma et al. [6] helped us further refine our architecture and helped us with some issues we ran into during the design. Vina et al. [7] and Xie et al. [8] allowed us to compare our results with some existing works. Carion et al. [2] and Ren et al. [5] were used during our research phase while deciding what model architecture to implement.

4. Method

Our solution leverages the YOLO (You Only Look Once) object detection framework for real-time identification of parking spaces. We felt YOLO is well-suited for this task due to its ability to perform fast and accurate object detection. Our model classifies parking spots into two categories: "empty" or "occupied." We designed the solution to be robust and accurate, incorporating several key steps. Those steps are Data Preprocessing and Augmentation, which includes labeled images of parking spots across three different parking lots under varying weather conditions (sunny, overcast, and rainy) and from diverse camera angles. And also, to improve model performance under different conditions, we applied data augmentation, introducing variations in brightness, contrast, and rotation. In terms of the model training, we employed the YOLOv11n configuration, fine-tuning it specifically for this task to optimize the detection of parking spots. In addition, we implemented the solution in a collaborative coding environment to handle data, train the model, and evaluate results. And to measure the effectiveness of our solution, we used standard evaluation metrics, including precision, recall, F1-Score, and accuracy.

Image 1. An output from the trained YOLO model on Lot T at UT Dallas.

5. Experiments

To train our model, we relied on the PKLot dataset [1]. This dataset contains nearly 700,000 images captured from three different parking lots with two different camera views. It contains samples from multiple times of day with great granularity, and features multiple weather conditions, such as sunny, rainy, and overcast. We took a subset of 12,000 images from this dataset that were labeled by the publishers. To test our model, we took images from Google Maps of Lot T on the UT Dallas campus and evaluated our model using them.

We have used accuracy, precision, recall, and f1-score as our evaluation metrics. Precision measures the accuracy of the positive predictions made by the model. For detecting parking spaces, high precision means that when the model predicts a space is occupied, it is often correct. Low precision indicates that the model makes many false positive errors such as predicting an empty space to be occupied when it is actually empty. Recall measures the ability of the model to find all the relevant positive cases. In the context of parking spaces, high recall means the model successfully identifies most of the occupied spaces. Low recall indicates that the model misses many actual occupied spaces (false negatives). The f1-score is the harmonic mean of precision and recall. It provides a single metric that balances both the precision and recall of the model. For parking detection, a high f1-score means that the model has a good balance between precision and recall, effectively identifying occupied spaces while minimizing false positives and false negatives. Accuracy measures the overall correctness of the model. It is defined as the ratio of correctly predicted observations to the total observations. In the context of YOLO's parking space detection, high accuracy means that most of the predictions for both empty and occupied spaces are correct. However, accuracy can sometimes be misleading if the dataset is imbalanced, especially in cases where there are many more empty spaces than occupied ones.

Image 2. Some images from the PKLot dataset with their bounding boxes.

	precision	recall	f1-score	support
0	0.91	0.77	0.83	13
1	0.75	0.90	0.82	10
accuracy			0.83	23
macro avg	0.83	0.83	0.83	23
weighted avg	0.84	0.83	0.83	23

Accuracy: 0.8260869565217391

Figure 1. Quantitative results of evaluation metrics after testing the model.

In figure 1, we saw impressive results with our evaluation metrics based on how the model performed on the given test data. The labels used are 0 which represents empty parking spaces and 1 which represents occupied parking spaces. The model has a precision value of 0.91 for detecting empty spaces which is very high. This means that the model is very good at correctly identifying empty parking spaces with only a small number of false positives. This is particularly beneficial as it reduces the chances of the system incorrectly informing students that a space is available when it is actually occupied. The model's recall value for occupied spaces is 0.90, indicating that it successfully identifies a large majority of the occupied parking spaces. This high recall reduces the likelihood of occupied spaces being missed by the system, ensuring that students are less likely to be directed to already occupied spots. The f1-score is 0.83 for empty spaces and 0.82 for occupied spaces which is relatively high, and this shows a balanced performance between precision and recall. This balance ensures that the model is consistently reliable in various scenarios, whether the parking lot is mostly full or mostly empty. With an overall accuracy of 83%, the model performs well in correctly identifying both occupied and empty spaces. This high level of accuracy translates to improved reliability for real-time parking management. The precision for occupied spaces is 0.75, which indicates some room for improvement in reducing false positives. Enhancing precision for occupied spaces could involve refining the model to be more discerning, ensuring fewer empty spaces are incorrectly labeled as occupied. The recall for empty spaces is 0.77, which suggests that the model could miss some empty spaces. Improving this recall rate would help ensure the system does not overlook available parking spots. By reducing false positives and increasing true negatives, we believe that we can further refine our model, which will make it more robust and reliable for real-time parking management. This will ultimately improve the overall parking experience for students.

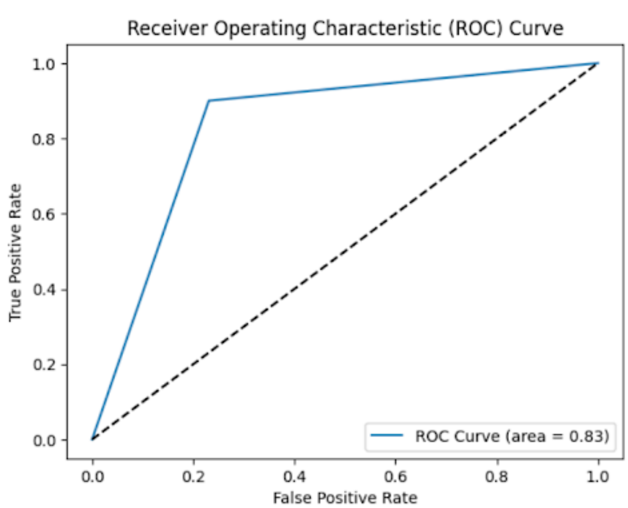

Figure 2. ROC curve plotting true positive rate against false positive rate.

The ROC curve in figure 2 assesses how the model classifies a parking space as empty or occupied. The false positive rate represents the proportion of actual negatives (empty parking spaces) that are incorrectly identified as positives (occupied parking spaces). The true positive rate measures the proportion of actual positives (occupied parking spaces) that are correctly identified. The area under the curve value is a single scalar that summarizes the overall performance of the model. An AUC of 0.83, as indicated on the curve, suggests that the yolo model has good discriminative ability. The closer the AUC is to 1, the better the model is at distinguishing between positives and negatives. The ROC curve also shows that the model performs particularly well at lower FPR values, where the TPR increases rapidly. This suggests that the model is capable of detecting most parking spaces accurately with minimal false positives, which is a desirable characteristic for a parking space detection system. We have kept track of the class labels based on the model's predictions in a bar graph as shown in figure 3.

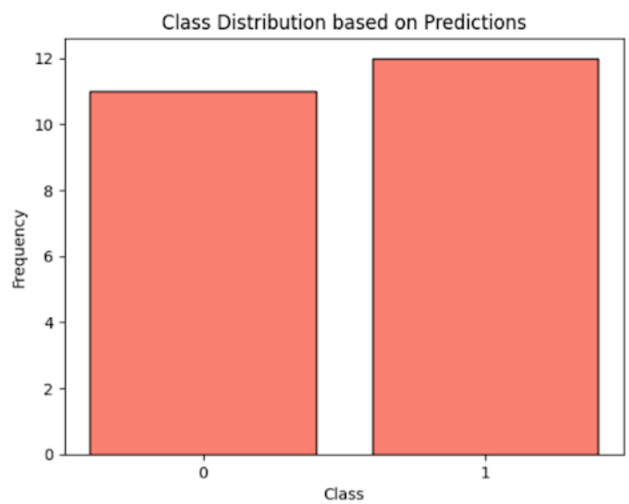

Figure 3. Model predictions for each class label.

The x-axis shows the class labels of 0 and 1 to represent their respective parking space identification. The y-axis shows the number of times each class was predicted by the model. This is the count of the model's predictions for each class. The graph shows a relatively balanced number of predictions for empty and occupied spaces. This balance suggests that the model is not biased towards predicting more empty or occupied spaces, which is important for providing reliable information about the parking lot status. Based on the data gathered, the model can easily optimize space utilization, direct drivers to empty spots, and reduce congestion. We analyzed the trends of each evaluation metric such as the graph shown in Figure 4. The x-axis represents the confidence threshold, ranging from 0.0 to 1.0. The confidence threshold determines the minimum score that the model must assign to a prediction for it to be considered as part of a certain class. The y-axis represents the f1-score, also ranging from 0.0 to 1.0. The f1-score is still the regular metric used to represent the harmonic mean of precision and recall. The light blue curve represents empty spaces and it shows how the f1-score for detecting empty parking spaces changes with different confidence levels. The orange curve represents space occupied and it shows how the f1-score for detecting occupied parking spaces changes with different confidence levels. The dark blue curve represents all classes and it highlights the overall f1-score across all classes, with a specific point of interest where the f1-score reaches 0.97 at a confidence level of 0.550. For the light blue curve, as the confidence threshold changes, the model's ability to correctly identify empty parking spaces fluctuates. A high f1-score means the model accurately identifies most empty spaces while minimizing false positives and false negatives. For the orange curve, changes in the confidence threshold affect the model's performance in detecting occupied spaces. A peak in the curve indicates an optimal balance where the model performs best in detecting occupied parking spaces. For the dark blue curve, the point where the f1-score is 0.97 at a confidence level of 0.550 is significant, indicating an optimal threshold where the model achieves its best overall accuracy. At this threshold, the model balances detecting both empty and occupied spaces effectively, minimizing errors in both categories. Overall, this helps visualize the trade-offs between precision and recall. A high f1-score at certain thresholds implies that the model is effective at correctly identifying both empty and occupied spaces, thus providing reliable and accurate information to users.

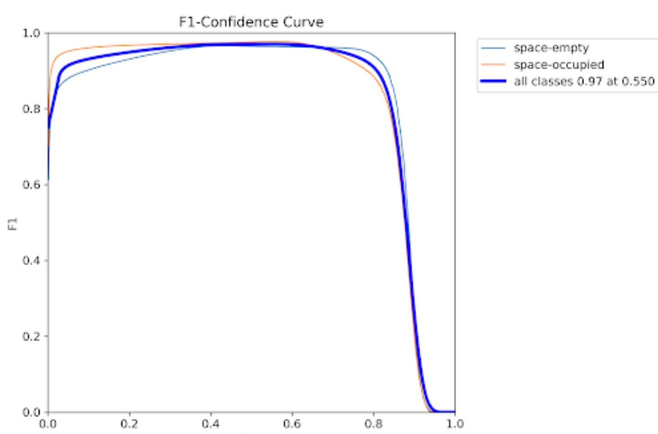

Figure 4. F1-confidence curve based on given class labels.

After analyzing the trends for the f1-confidence curve, we moved onto seeing the progression of the model's precision over time. We generated a precision-confidence curve to help us visualize the results of this metric as shown in Figure 5. The light-blue curve illustrates how the precision for detecting empty parking spaces varies with different confidence thresholds. The orange curve shows how the precision for detecting occupied parking spaces changes with different confidence thresholds. Finally, the dark blue curve highlights the overall precision across all classes, with a specific point of interest where the precision reaches 1.00 at a confidence level of 0.954. The graph shows that the precision for detecting empty parking spaces increases as the confidence threshold rises, which eventually reaches very high precision at higher confidence levels. This indicates that at lower confidence thresholds, the model makes more false positive errors by incorrectly predicting empty spaces. As the threshold increases, the model becomes more conservative, and the precision improves because it makes fewer false positive errors. However, this could come at the cost of lower recall, meaning some true empty spaces might be missed. In the case of occupied parking spaces, the precision shows a similar trend where it increases as the confidence threshold rises, eventually reaching high precision. Higher precision at increased confidence levels means the model is better at accurately identifying occupied spaces without mistakenly classifying empty spaces as occupied. As with the empty spaces, higher precision might reduce recall, which could potentially miss some actual occupied spaces. The overall precision across all classes reaches 1.00 at a specific confidence level of 0.954. This point of perfect precision indicates no false positives at this threshold. It signifies that the model is making highly accurate predictions at this threshold, correctly classifying all detected parking spaces without any false positives. This optimal threshold is essential where the cost of false positives is high, ensuring the reliability of the system. While high precision is desirable, it is essential to balance it with recall. High precision might lead to lower recall, where the model misses some true positives. The curve helps visualize this trade-off, guiding the selection of a threshold that offers a good balance between precision and recall.

We moved onto the final metric which was recall to analyze its growth over time. We plotted the recall-confidence curve as shown in Figure 6. The light blue and orange curves represent the same classes as shown in the previous curves. The dark blue curve highlights the overall recall across all classes, with a specific point where the recall is 1.00 at a confidence level of 0.000. The light blue curve starts high and remains high, indicating that the recall for detecting empty spaces is consistently strong across most confidence levels. This is crucial for ensuring that users can reliably find available parking spots. The curve suggests that even at lower confidence thresholds, the model does a good job of detecting empty spaces, making it very effective at not missing available spaces. The orange curve is

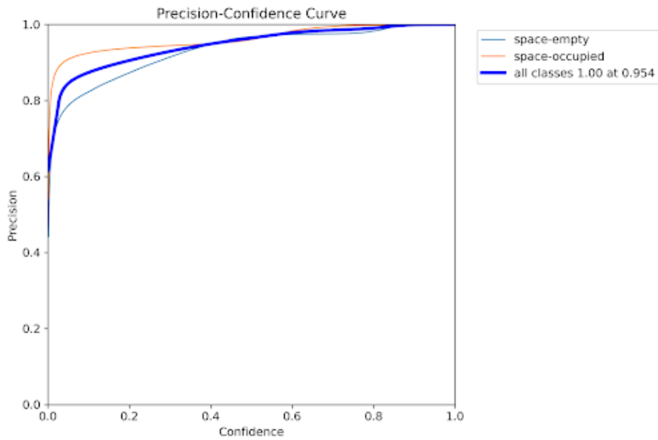

Figure 5. Precision-Confidence Curve based on given class labels.

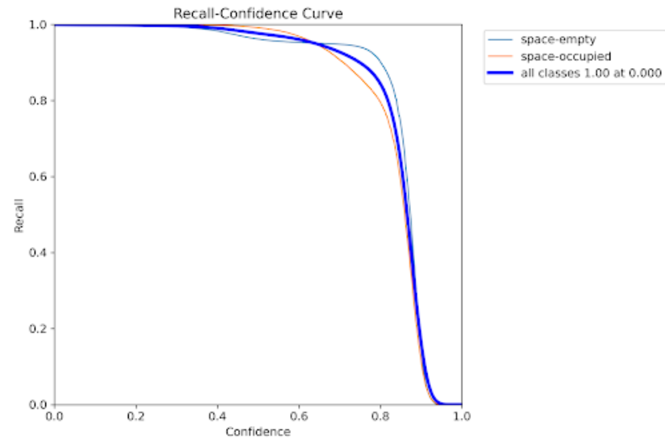

Figure 6. Recall-confidence curve based on given class labels.

similar to the light blue curve because this curve also starts high and remains high, indicating that the recall for detecting occupied spaces is consistently strong across most confidence levels. High recall for occupied spaces means that the model accurately identifies most of the occupied spaces, which is essential for avoiding false negatives where an occupied space is missed and mistakenly identified as empty. The consistency across confidence levels indicates robust performance in identifying occupied spaces. The dark blue curve shows that the overall recall across all classes is perfect with a value of 1.00 at a confidence threshold of 0.000. A recall of 1.00 at a confidence level of 0.000 indicates that at this threshold, the model correctly identifies every true positive, without missing any. While this is ideal in terms of recall, such a low confidence threshold might lead to many false positives, as the model would classify almost all instances as positive. Therefore, we believe that recall needs to be balanced with precision to ensure optimal performance.

Upon analyzing the recall-confidence curve, this led us to comparing the growth of precision and recall over time. We constructed a precision-recall curve as shown in Figure 7. The light blue and orange lines represent the same class labels in the previous curves. The dark blue line in this curve is unique since it represents the mean average precision for all classes, with a value of 0.991 at an intersection over union threshold of 0.5. The light blue line shows that the model is highly accurate when predicting that a parking space is empty. For every prediction that a space is empty, 98.9% of the time this prediction is correct. High recall indicates that the model can identify most of the empty spaces correctly. The curve's behavior reflects how well the model performs at different recall levels, maintaining high precision. The orange line shows that the model is highly accurate when predicting that a parking space is occupied. For every prediction that a space is occupied, 99.2% of the time this prediction is correct. High recall with respect to this line indicates that the model can identify most of the occupied spaces correctly. The mean average precision is a single metric that summarizes the precision-recall performance across all classes. For the dark blue line, an mAP of 0.991 indicates that the model performs exceptionally well in distinguishing between empty and occupied spaces, achieving high precision and recall on average across all classes. Intersection over union is a threshold that determines how much overlap there must be between the predicted bounding box and the ground truth bounding box for a prediction to be considered correct. As shown in the figure, we have achieved a threshold of 0.5 which is a common value that balances strictness and leniency. Overall, the high precision values for both empty and occupied spaces indicate that the model makes very few false positive errors. It reliably distinguishes between empty and occupied spaces, ensuring that the predictions are accurate. The shape of the precision-recall curve suggests that the model maintains a good balance between precision and recall across different thresholds. This balance is crucial as it ensures that the model does not miss many occupied or empty spaces while also avoiding false alarms. By analyzing this curve we can see how the system can accurately inform users about the availability of parking spaces, reducing the time spent searching for a spot and improving overall user satisfaction.

6. Conclusion

In conclusion, we have successfully developed and evaluated a YOLO-based object detection model for identifying empty and occupied parking spaces. The detailed analysis of evaluation metrics, including precision, recall, f1-score, and accuracy, provided a comprehensive understanding of the model's performance. Although our model may not be

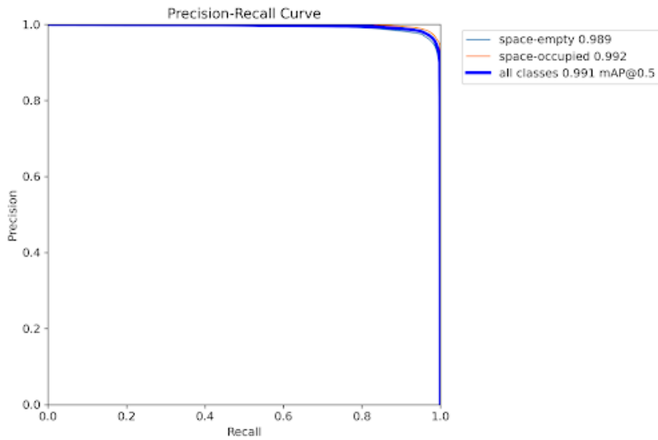

Figure 7. Precision-recall curve based on given class labels.

entirely perfect, we are confident in making any necessary improvements so that we can ensure it provides an efficient solution for parking space detection. We believe that more exposure to various forms of training data will help us improve the model's performance so that it is both optimal and reliable.

Integration of our model with campus parking systems presents a promising next step. For instance, embedding the model into a UTD parking app could provide students with real-time updates on available parking spaces, reducing search times and improving the campus commuting experience. This would require collaboration with campus IT teams to ensure compatibility with existing infrastructure and seamless user experience. A key direction for future work involves testing the model on live camera feeds from UTD parking lots to assess its performance under varying real-world conditions, such as lighting changes, weather variations, and different camera angles. This step would provide insights into the model's robustness and adaptability. Lastly, incorporating feedback from students and campus administrators could help refine the system further, ensuring it addresses the unique challenges faced by the UTD community. These steps will bridge the gap between our experimental results and practical deployment which paves the way for a scalable, real-world solution.

References

[1] P. Almeida, L. S. Oliveira, E. Silva Jr, A. Britto Jr, and A. Koerich. Pklot – a robust dataset for parking lot classification, expert systems with applications. 42(11):4937–4949, 2015. 2

[2] Nicolas Carion, Francisco Massa, Gabrial Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. End-to-end object detection with transformers. *Computer Vision–ECCV, 16th European Conference, Glasgow, UK*, pages 213–229, 2020. 1

[3] Manato Ogawa, Tomer Arnon, and Edward Gruber. Identifying parking lot occupancy with yolov5. *Journal of Student Research*, 12, 2023. 1

[4] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016. 1

[5] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 2015. 1

[6] Nabin Sharma, Sushish Baral, May Phu Paing, and Rathachai Chawuthai. Parking time violation tracking using yolov8 and tracking algorithms. *Sensors*, 23:5843, 2023. 1

[7] Abirami Vina. Ultralytics yolov8 for smarter parking management systems. 1

[8] Zheng Xie and Xing Wei. Automatic parking space detection system based on improved yolo algorithm. In *2021 2nd International Conference on Computer Science and Management Technology (ICCSMT)*, pages 279–285, 2021. 1