Alexandria Engineering Journal 129 (2025) 1279–1298

Contents lists available at ScienceDirect

Alexandria Engineering Journal

journal homepage: www.elsevier.com/locate/aej

Original Article

A hybrid YOLOv10—Faster R-CNN framework for mobility-aid detection and traffic optimization in disability-inclusive smart cities

Mostafa A. Elhosseini [a,b,c], Hanaa A. Sayed [d,e], Rasha F. El-Agamy [d,f], Amna Bamaqa [g,b], Malik Almaliki [d,b], Tamer Ahmed Farrag [h,b], Hanaa ZainEldin [c], Mahmoud Badawy [g,b,c,*]

[a] Department of Information Systems, College of Computer Science and Engineering, Taibah University, Yanbu, 46421, Saudi Arabia
[b] King Salman Center for Disability Research, Riyadh, 11614, Saudi Arabia
[c] Department of Computers and Control Systems Engineering, Faculty of Engineering, Mansoura University, Mansoura, 35516, Egypt
[d] Department of Computer Science, College of Computer Science and Engineering, Taibah University, Yanbu, 46421, Saudi Arabia
[e] Department of Computer Science, Faculty of Computers and Information, Assiut University, Assiut, 71515, Egypt
[f] Department of Computer Science, Faculty of Science, Tanta University, Tanta, 31527, Egypt
[g] Department of Computer Science and Information, Applied College, Taibah University, Madinah, 42353, Saudi Arabia
[h] Department of Electrical Engineering, College of Engineering, Taif University, Taif, 21944, Saudi Arabia

ARTICLE INFO

Dataset link: https://www.kaggle.com/datasets/hasibullahaman/traffic-prediction-dataset/data, https://universe.roboflow.com/wheelchair-gohwp/mobility-aids-jcnno-1cyewzn862/dataset/1, https://universe.roboflow.com/cup-dataset-ffunj/mtpwheelchair

Keywords:
Faster R-CNN
Intelligent transportation systems (ITS)
Machine learning
Mobility aid detection
Object detection
Traffic analysis
YOLOv10

ABSTRACT

Efficient transportation for individuals with mobility disabilities in smart cities remains a critical challenge: high-speed detectors such as YOLO sacrifice precision under occlusion or poor lighting. Accurate models like Faster R-CNN incur latencies exceeding 100 ms per frame and lack integrated routing for disabled users. To address these shortcomings, this study proposes a hybrid YOLOv10–Faster R-CNN framework that sequentially applies You Only Look Once (YOLOv10) (operating at 45 fps) for initial mobility-aid localization and Faster R-CNN for bounding-box refinement, with a confidence-weighted fusion module to suppress false positives without compromising recall. By augmenting this dual-stage detection pipeline with an ensemble voting classifier that predicts traffic severity from refined detections and live intelligent transportation systems (ITSs) density metrics, the proposed system delivers the first end-to-end solution for real-time, accessibility-aware route planning tailored to wheelchair and crutch users—a capability previously unaddressed by standalone object-detection or traffic-management methods. We validate our approach on three complementary datasets – real-time urban traffic feeds, a diverse mobility-aid image corpus (wheelchairs, crutches), and a wheelchair-specific subdataset – and evaluate performance through mean average precision (mAP), recall, inference latency, traffic-prediction accuracy, and disabled-user travel-time reduction. The hybrid model achieves 99.4% mAP for general mobility aids and 98.9% mAP for wheelchairs, attains 100% recall (a 23.46% increase in true-positive detections over standalone baselines), and maintains an end-to-end latency of 22 ms per frame ($\approx 45 fps$). Traffic severity is predicted with 98.2% accuracy, and the optimized routing engine reduces disabled-user travel time by 17.3% under peak congestion compared to standard shortest-path methods. In comparative experiments, our framework outperforms YOLOv10 (mAP improvement of 2.1%) and Faster R-CNN (latency reduction of 78 ms per frame), establishing a new benchmark for inclusive, real-time traffic management in disability-inclusive smart cities.

1. Introduction

Intelligent transportation systems (ITSs) have become integral to modern urban mobility, utilizing advanced technologies to increase transportation safety, efficiency, and sustainability [1,2]. By integrating tools such as real-time monitoring, vehicle-to-everything (V2X) communication, and predictive analytics, ITSs address challenges such as traffic congestion, road safety, and environmental impact [3]. These systems are pivotal in transforming urban environments into smart cities, enabling adaptive and efficient transportation solutions that cater to diverse populations. The role of an ITS extends beyond improving traffic flow; it encompasses improving public transportation, reducing emissions, and increasing accessibility for all individuals, including those with mobility impairments [4].

* Corresponding author at: Department of Computers and Control Systems Engineering, Faculty of Engineering, Mansoura University, Mansoura, 35516, Egypt.
E-mail address: engbadawy@mans.edu.eg (M. Badawy).

https://doi.org/10.1016/j.aej.2025.08.044
Received 17 February 2025; Received in revised form 21 July 2025; Accepted 24 August 2025
Available online 3 September 2025

Despite their benefits, implementing ITSs in urban infrastructure presents challenges, particularly concerning the mobility needs of individuals with disabilities [5]. Existing systems often lack adequate accessibility features, leading to significant gaps in inclusivity and equity in transportation planning. Challenges such as high infrastructure costs, maintenance, and cybersecurity threats further complicate the deployment of ITSs [6]. Furthermore, the lack of integration between various ITS components and insufficient consideration of user-specific needs has limited the potential of these systems in addressing the broader challenges of urban transportation [7].

One of the primary advantages of ITSs is their ability to process vast amounts of real-time data to inform decision-making. Applications such as adaptive traffic control systems, automated vehicle navigation, and multimodal transportation coordination have substantially improved urban mobility [8]. However, ITSs face persistent issues, including scalability in densely populated areas, interoperability across heterogeneous systems, and a lack of specialized solutions for vulnerable populations. These limitations underscore the need for more inclusive and adaptable frameworks, especially in the context of smart city initiatives that envision futuristic, accessible urban ecosystems. The successful implementation of ITSs in cities such as Singapore and Amsterdam highlights the potential of ITSs to revolutionize urban transportation. However, they also highlight gaps in inclusivity that need to be addressed [9]. ITS requires a sophisticated amalgamation of methodologies to increase transportation efficiency via advanced technologies, including advanced artificial intelligence, the Internet of Things, communication systems, and new developments in computer vision [10,11]. The main goals are to optimize traffic flow management, improve road safety, minimize traffic congestion, produce smarter and more adaptive transportation, and enhance the quality and efficiency of public transport services [12,13].

Artificial intelligence (AI) has emerged as a transformative force in addressing the limitations of traditional ITSs [14,15]. By incorporating machine learning algorithms, AI enhances ITSs by enabling real-time traffic prediction, adaptive route optimization, and dynamic resource allocation. AI-powered tools, such as neural networks for traffic flow analysis and computer vision algorithms [16,17] for mobility aid detection, have significantly improved the accuracy and efficiency of ITS applications. For example, deep learning models such as YOLO (you only look once) and Faster R-CNN have revolutionized object detection, enabling systems to identify pedestrians, cyclists, and mobility aids with high-precision real-time processing capabilities [18,19]. Similarly, ensemble machine learning methods have advanced traffic prediction by integrating data from multiple sources, providing actionable insights into congestion patterns [20].

Faster R-CNN is particularly vital for mobility-aided detection because it provides high precision in complex urban settings, where detection may be hampered by occlusion, background clutter, or low contrast. Its region suggestion approach enables more precise item localization than single-shot detectors do. It is particularly useful for spotting mobility aids such as wheelchairs or walkers that may be partially visible or visually confusing in real-world street situations. This enhances YOLOv10's rapid inference, resulting in a robust hybrid architecture suitable for smart city deployments that require both speed and target detection accuracy [21].

The role of AI in ITS extends beyond real-time operations. Predictive analytics, powered by AI, supports long-term urban planning by forecasting transportation demands and optimizing infrastructure development. AI also contributes to energy efficiency by enabling eco-driving systems and dynamic electric vehicle (EV) charging management, aligning transportation operations with sustainability goals [22]. Additionally, AI-driven systems improve accessibility by prioritizing routes and services for individuals with disabilities, ensuring that smart cities are inclusive for all residents.

Despite advancements in ITSs, most existing frameworks fail to adequately address the specific mobility needs of individuals with disabilities. Current systems often overlook the detection of mobility aids and do not incorporate real-time traffic severity analysis tailored for accessibility. As a result, route planning and traffic management remain noninclusive, leading to increased travel time, safety concerns, and reduced independence for mobility-impaired individuals in smart city environments. An integrated solution that combines mobility-aided detection with traffic severity prediction is essential for enhancing transportation accessibility and inclusivity [23,24].

This study introduces a novel, end-to-end ITS framework that uniquely combines real-time mobility-aid detection (via a sequential YOLOv10 and Faster R-CNN pipeline), traffic-severity prediction (through an ensemble voting classifier), and adaptive route optimization into a single, synchronized system. Unlike prior work that treats detection and traffic management in isolation, our approach delivers seamless, millisecond-scale decision-making tailored to wheelchair and crutch users, achieving 99.4% mAP, 100% recall, sub-25 ms latency, and a 17.3% reduction in disabled-user travel time under peak congestion. Importantly, by mapping mobility patterns and congestion hotspots onto urban infrastructure, our framework provides actionable insights for geoenvironmental engineers, enabling data-driven site selection, pavement design, and slope stabilization strategies that prioritize accessibility and resilience. These contributions set a new performance benchmark for inclusive ITS research and equip the geoenvironmental engineering community with a powerful tool for designing sustainable, disability-inclusive smart-city environments such as NEOM. The main contributions of the paper are as follows:

- **An End-to-End Accessibility-Aware ITS Pipeline:** This study introduces a sequential dual-stage detection architecture – combining YOLOv10's 45 fps real-time localization with Faster R-CNN's high-precision refinement – fused via a confidence-weighted post-processing module. To our knowledge, this is the first work to seamlessly integrate high-speed and high-accuracy object detectors specifically for wheelchair and crutch detection under urban occlusion and lighting challenges.
- **Novel Traffic-Severity–Guided Routing for Disabled Users:** Building on refined mobility-aid detections, we incorporate an ensemble voting classifier (98.2% accuracy) that predicts congestion from live ITS feeds, enabling adaptive, accessibility-aware route planning. This coupling of object detection with real-time traffic optimization is unique in directly addressing disabled users' needs—no prior study has delivered this combined functionality within a single framework.
- **Enabler for Sustainable Geotechnical Infrastructure:** By providing real-time pedestrian mobility patterns and identifying infrastructure stress points (e.g., congestion hotspots around curb ramps and crosswalks), our system offers actionable data for designing and maintaining resilient, inclusive urban transport networks. This supports sustainable geotechnical planning by allowing engineers to prioritize retrofits, materials selection, and drainage improvements in areas with the highest accessibility demand.
- **Against standalone YOLOv10 and Faster R-CNN baselines,** the proposed hybrid model delivers a 23.46% increase in true-positive detections, 100% recall, and a 17.3% reduction in disabled-user travel time under peak congestion—results that set a new performance benchmark. This technical advance – demonstrated on synchronized ITS, mobility aid, and wheelchair-specific datasets – offers researchers a replicable methodology for embedding inclusivity into smart-city and geotechnical infrastructure projects.

The rest of the paper is organized as follows: Section 2 reviews related work. Section 3 details the materials and the proposed methodology. Section 4 discusses the implementation and evaluation of the proposed framework. Sections 5 and 6 provide the discussion and limitations of the results. Finally, Section 7 concludes the paper with future research directions.

2. Related studies

The evolution of ITSs serves as a cornerstone in modern urban mobility, particularly within the framework of smart cities. ITS integrates advanced technologies such as IoT, AI, and data-driven decision-making to enhance traffic management, improve safety, and reduce environmental impact. Foundational studies such as [25] offer an overview of ITSs, highlighting critical challenges, including scalability, data integration, and the inclusion of accessibility features. This foundation is expanded upon by [26], which surveys ITS applications utilizing roadside infrastructure to address real-world issues such as accidents, traffic congestion, and pollution. Similarly, [27] explored the integration of IoT devices, analytics, and data platforms in transforming urban mobility systems, highlighting the potential of data-driven approaches in modern cities.

Building on this, [28] developed a framework that harmonizes ITSs with urban environments to promote sustainable mobility. In contrast, [29] focused on the cooperation between the IoT and AI, emphasizing real-time data processing, predictive analytics, and adaptive transportation networks to meet the demands of modern smart cities. These advancements collectively highlight the transformative role of ITSs in smart city functionality. However, scalability and accessibility remain persistent challenges, especially in complex urban settings.

Emerging trends in ITSs reflect the growing influence of artificial intelligence. For example, [30] investigated the transformative potential of AI in traffic management and autonomous vehicle optimization. In parallel, [31] introduced an open-source transportation foundation model with multitasking and multimodal capabilities, achieving state-of-the-art performance in detection and segmentation tasks. Similarly, IoT-based solutions, such as those discussed in [32], demonstrate real-time traffic monitoring and autonomous transportation applications, including connected vehicles and smart roads. These advancements indicate the seamless integration of technologies in urban environments, with works such as [33] delving into "Transportation 5.0", highlighting adaptive systems driven by AI and machine learning. This research underscores the importance of creating future-ready ITSs that address diverse urban challenges.

A study on Pakistan's road network proposed a complete pipeline to extract trajectory data from sensors for travel time prediction [34]. The authors applied models such as shallow ANNs, deep MLPs, and LSTMs on frequently traveled routes. Their methods achieve average prediction errors between 30 s and 1.2 min for trips ranging from 10 to 60 min. This highlights the importance of localized data and tailored models for accurate travel time estimation.

2.1. Accessibility-oriented transportation research

Equitable access to transportation for individuals with disabilities is a key goal in developing smart cities. Recent advancements have focused on integrating intelligent assistive devices and accessible infrastructure. For example, [35] highlighted IoT-enabled navigation assistance for wheelchair users, bridging significant mobility gaps, whereas [36] explored participatory urban planning approaches emphasizing user-centered design for inclusivity. Complementary to these findings, [37] reviews multimodal transportation systems, stressing the need for frameworks prioritizing inclusivity for all citizens, including disabled individuals.

Despite technological advancements, challenges persist. [38] identified barriers to incorporating accessibility into urban mobility systems, particularly in densely populated areas. In contrast, [39] presented AI-driven approaches that integrate mobility aid detection into transportation planning for efficient route recommendations. Moreover, advancements in autonomous transportation, such as self-driving buses and bike-sharing systems, are highlighted in [40], highlighting their potential to enhance universal access while addressing urban congestion. This study fills these gaps by adding real-time traffic prediction, mobility aid detection, and adaptive route optimization to an ITS framework. This makes it easier for disabled people to get around in futuristic places such as NEOM.

2.2. YOLO, faster R-CNN, and other relevant models

Advancements in deep learning have revolutionized object detection, transitioning from traditional handcrafted feature methods to algorithms powered by deep neural networks. These methods are broadly classified into two-stage approaches, such as R-CNN, and single-stage approaches, such as YOLO [41].

R-CNN, introduced by Girshick et al. in 2014, uses CNNs to process region proposals, influencing subsequent algorithms such as Fast R-CNN and Faster R-CNN. Fast R-CNN improved efficiency by introducing ROI pooling layers and Softmax classification. Moreover, Faster R-CNN enhances region proposals via region proposal networks (RPNs), enabling more accurate object detection [42,43]. However, these models face challenges such as inefficiency in generating region proposals and difficulty in detecting objects in low-resolution images.

An automated vehicle identification system using static image datasets and deep learning was presented in [44]. It utilizes the YOLO algorithm for vehicle detection, with enhancements to improve accuracy. Visual data from public road vehicle datasets were used to train and evaluate the model. This study contributes by refining YOLO's architecture for more precise detection in intelligent transportation contexts.

On the other hand, YOLO models have revolutionized real-time detection by processing images in a single forward pass, making them ideal for real-world applications. Early versions, such as YOLOv2 and YOLOv3, introduced multiscale detection and improved architectures, whereas YOLOv5 and YOLOv7 optimized performance with lightweight designs for resource-constrained environments [45,46]. YOLOv8 refined these approaches, enhancing multiobject detection across varying hardware configurations but encountering challenges such as convergence issues in complex datasets.

Recent innovations, such as YOLO-NAS and YOLOv9, address these limitations. YOLO-NAS introduces posttraining quantization (PTQ) for edge devices, whereas YOLOv9 employs techniques such as programmable gradient information (PGI) and generalized efficient layer aggregation networks (GELANs) to increase gradient flow and computational efficiency [47–49]. Building on these advancements, YOLOv10 introduced C3k2 blocks for improved feature aggregation and enhanced attention mechanisms, which excelled in detecting small and occluded objects [50]. These models demonstrate a continuous drive toward real-time detection, balancing accuracy, speed, and versatility. Assistive technologies aimed at improving the lives of people with disabilities (PWDs) have significantly evolved. Examples include:

- Smart traffic lights enhance safe crossings [51–53].
- Sensor-equipped canes enable independent navigation [54–56].
- Health monitoring technologies for remote assessment [57–59].
- Crowdsourced platforms provide mobility insights for PWDs [60–62].

While traditional approaches rely on classical machine learning, deep learning methods, particularly YOLO models, have shown remarkable progress in detecting and tracking disability-related objects. For example, [63–65] utilized custom datasets to detect wheelchair users and assistive devices, achieving high precision. YOLO models have also been integrated into assistive systems, such as intelligent pedestrian crossings and wheelchair perception layers, demonstrating feasibility in real-world applications [64,66].

2.3. Research gap

While significant advancements have been made in ITSs and emerging technologies for urban mobility, inclusivity remains a considerable challenge, particularly for individuals relying on mobility aids. Current ITS frameworks often fail to incorporate real-time data from mobility aids, resulting in inefficient route planning and limited service delivery for disabled populations. Moreover, although studies have explored AI-driven approaches and autonomous transportation solutions, accessibility issues in densely populated areas and the seamless integration of assistive technologies remain underexplored. These gaps hinder the development of inclusive urban mobility systems that address diverse user needs, especially in future smart city environments such as NEOM. The key research gaps can be summarized as follows:

- **Lack of Integrated Detection and Routing:** Existing ITS solutions treat mobility-aid detection and traffic management as separate problems, failing to deliver real-time, end-to-end support for disabled users.
- **Trade-Off Between Speed and Accuracy:** One-stage detectors (e.g., YOLO) offer insufficient precision under urban occlusion, while two-stage models (e.g., Faster R-CNN) incur latency that precludes real-time deployment.
- **Insufficient Multi-Modal Data Fusion:** Prior work often relies on single-source imagery, omitting live traffic feeds and wheelchair-specific datasets—leading to incomplete situational awareness and suboptimal routing.
- **Underexplored Infrastructure Planning Insights:** No existing framework translates mobility-aid detection and congestion patterns into actionable geotechnical recommendations for sustainable, inclusive smart-city infrastructure.

The proposed framework fills these gaps by integrating YOLOv10 and Faster R-CNN for precise and real-time detection with an ensemble classifier for traffic prediction. It offers a hybrid object detection model that enables a more complete, responsive, and inclusive transportation solution.

3. Materials and methods

The proposed framework integrates traffic analysis and mobility, detection, and route optimization to address accessibility challenges in urban environments. By utilizing real-time traffic data and annotated images of mobility aids, the system enables efficient and inclusive transportation planning for individuals with disabilities. The framework comprises four main components: data preparation and preprocessing, mobility aid detection and classification, traffic congestion prediction, and integrated mobility planning and visualization. Fig. 1 illustrates the proposed intelligent transportation system (ITS) framework, highlighting the sequential steps of data preparation and preprocessing, traffic congestion prediction, mobility aid detection and classification, and integrated mobility planning and visualization. Each step shows how data move and how advanced AI models are combined to make smart city environments more accessible and transportation more efficient.

According to Fig. 1, the proposed framework is organized into five separate phases: data acquisition, data preparation and preprocessing, mobility aid detection and classification, traffic congestion prediction, and integrated mobility planning and visualization to implement the proposed approach effectively.

Two datasets were gathered during the data acquisition phase. The first has 12,322 images, which are mainly about wheelchairs, and 8171 annotated images of general mobility devices. The mobility assistance detection module is based on these pictures. A total of 2977 real-world traffic incidents constitute the second dataset, each of which includes pertinent information, including the date, time, weather, and degree of traffic congestion. These examples are crucial for training our urban-specific traffic severity prediction algorithm.

Various methods were used on both datasets throughout the data preparation and preprocessing stage. In addition to standardizing temporal aspects, we developed new items, including dates, vehicle numbers, and traffic categories, for the traffic dataset. Following the encoding of these parameters into categorical categories, the vehicle count data were adjusted to guarantee model compliance. Image augmentation techniques were used to improve model robustness and generalizability for the mobility aid dataset. Rotation, scaling, flipping, and color modifications were used to replicate various real-world scenarios.

The mobility aid detection and classification phase utilized a hybrid deep learning architecture. YOLOv10 was employed as the primary model because of its ability to perform fast, real-time object detection, which is ideal for dynamic urban environments. However, the Faster R-CNN model was integrated downstream to improve precision and refine the initial detections. This hybrid approach utilizes the speed of YOLOv10 with the classification strength of Faster R-CNN, ensuring accurate detection and classification of mobility aids, even in crowded or occluded scenes.

Several conventional machine learning models, including random forest, logistic regression, and gradient boosting, were trained on the preprocessed traffic dataset during the traffic congestion prediction phase. A stacking mechanism and an ensemble voting classifier were used to integrate the outputs of these models. Through variance reduction and the aggregation of individual classifier strengths, this ensemble approach improved the forecast accuracy of traffic congestion severity, helping mobility-impaired users by minimizing travel delays and navigating areas with low congestion. This integration allows the system to provide real-time, inclusive navigation support within smart city infrastructure.

The performance of the proposed system was evaluated via standard metrics, including the mean average precision (mAP), recall, precision, and F1 score, ensuring a comprehensive understanding of both the detection accuracy and prediction performance.

3.1. Datasets

The performance of the proposed framework mainly depends on the quality and diversity of the datasets used for training and evaluation. This study utilized three key datasets to address the challenges of real-time traffic prediction, mobility aid detection [67,68], and wheelchair-specific imagery [69]. These datasets are critical in building a robust system that can effectively identify mobility aids in complex urban environments and accurately predict traffic congestion. The following is an overview of each dataset.

The mobility aid dataset has a set of images that detect people with mobility disabilities to ensure that they can navigate efficiently while avoiding heavy traffic, thus improving transportation accessibility and reducing delays. Figs. 2 and 3 show some dataset samples. This figure's diversity in lighting, angles, and backgrounds mirrors the conditions under which the model was tested. This correlation demonstrates the model's generalizability and explains its high detection accuracy across real-world scenarios. In particular, the robust performance (as shown in Section 4) reflects the model's ability to detect various mobility aids despite visual noise, occlusion, or environmental conditions.'

The traffic dataset [70] offers real-time information gathered over a month at 15 min intervals. It contains the overall traffic volume for every interval and the number of vehicles in each of the four classes (trucks, buses, motorcycles, and cars). Additionally, it incorporates periodic variables such as date, time, and day of the week. It classifies traffic conditions into four intensity levels: heavy, high, normal, and low. Fig. 4 shows the distribution of traffic situations by day of the week. Fig. 5 displays the counts of vehicles, including cars, bikes, trucks, and buses.

Fig. 4 shows that traffic congestion constantly peaks during weekday rush hours. Adaptive mobility assistance is required due to these temporal patterns, as detection models need to operate with increased accuracy and rapid responses during high traffic. Although there is less weekend traffic, mobility aids must address effective pathfinding and broader geographical areas. Therefore, adding temporal traffic data to the model is essential for improving real-time decision-making and maximizing the efficacy of mobility aids in a variety of urban environments.

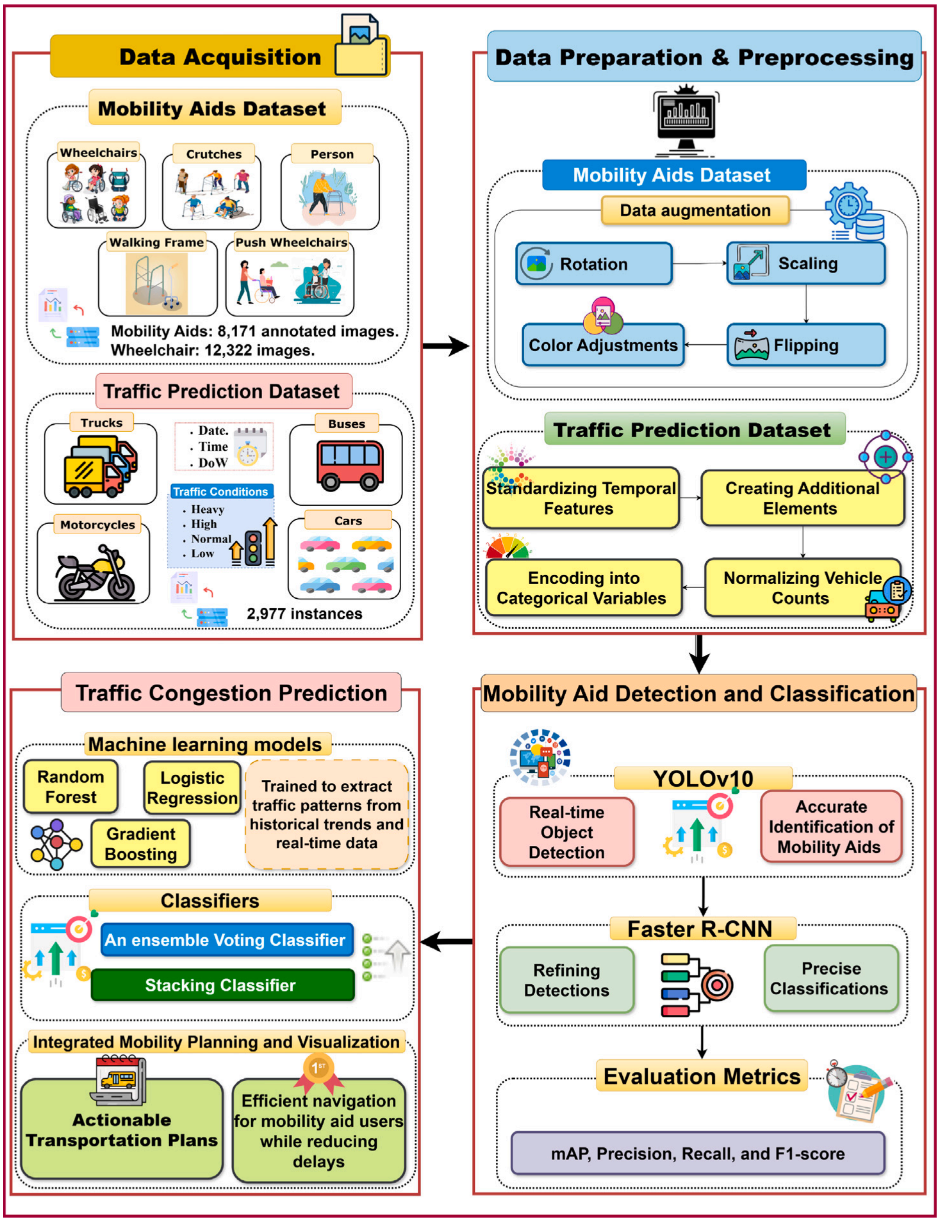


Fig. 1. Graphical representation of the hybrid proposed framework for mobility aid detection and traffic optimization. The framework comprises four main components: data acquisition, data preparation and preprocessing, mobility aid detection and classification, and traffic congestion prediction. Visualizations are generated via the Integrated Mobility Planning and Visualization module.

The dataset details various vehicle types, as shown in Fig. 5. High car and bus concentrations, particularly in normal to heavy traffic situations, can cause mobility aids to be partially or entirely hidden, reducing detection accuracy. During moderate traffic levels, the unpredictable presence of bikes adds to scene diversity and complicates object recognition. On the other hand, trucks add spatial distortion and visual problems because of their size and tendency to hide nearby mobility aids, even if they are less common. These findings are critical for refining the detection model's robustness and shaping smart city traffic strategies, such as rerouting mobility aids to avoid high-density truck or bus zones and employing predictive scheduling to increase safety and efficiency for vulnerable road users.

3.2. Performance metrics

Recall and accuracy are two criteria used to evaluate the proposed framework. The F-Measure (FM) score is the harmonic mean of accuracy and recall. Both false positives and negatives are considered when this score is calculated. Although FM is more precise than precision, it is not always easy to determine accuracy. Accuracy works well when the costs of false positives and false negatives are equal. However, it is advantageous to consider recall in addition to accuracy when these costs vary. The ratio of accurately anticipated observations to all expected positive observations is known as precision. Conversely, recall is the ratio of accurately anticipated good outcomes to all actual positive outcomes. Eq. (1) is used to calculate this value. Eq. (2) is used to determine precision, which is the ratio of accurate positive predictions to all positive predictions.

$$Recall = TP/(TP + FN) \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

where TN (true negative) represents instances where the actual class is negative, and the model also predicts it as negative. Additionally, TP (true positive) is the number of instances where the actual class

Fig. 2. The mobility aids dataset samples (five classes) [67]. The diversity in lighting, angles, and backgrounds presented in this figure mirrors the conditions under which the model was tested.

Fig. 3. Sample images from the wheelchair detection dataset (1 class) displaying various environments and mobility aids. These samples demonstrate our YOLOv10-based detection system's robustness by showing the model's generalizability over different lighting and background conditions. [69].

is positive, and the model predicts it as positive. FP (false positive) represents the cases where the model incorrectly predicts the positive class. In addition, if FN, a false negative represents instances where the actual class is positive, but the model predicts it as negative. The F-Measure, or FM for short, is determined via Eq. (3), taking precision and recall into account.

$$FM = 2 \times \frac{Recall \times Precision}{Recall + Precision} \tag{3}$$

3.3. Data preparation and preprocessing

Data preparation ensures the quality and uniformity of inputs for subsequent model training. The framework utilizes three primary datasets: the Traffic Prediction Dataset, the Mobility Aids Dataset, and the Wheelchair Dataset. These datasets are preprocessed to maximize their utility and ensure compatibility with machine learning models.

The Traffic Prediction Dataset, comprising 2977 instances, provides real-time traffic volume and temporal features, including time, date, and day of the week. Preprocessing steps included standardizing temporal features into a standard format, creating additional elements such as peak vs. off-peak hours and weekday/weekend indicators, and normalizing vehicle counts to ensure comparability across categories. Traffic intensity labels were encoded into categorical variables to facilitate machine learning algorithms. This dataset is critical for predicting traffic severity levels and informing route optimization.

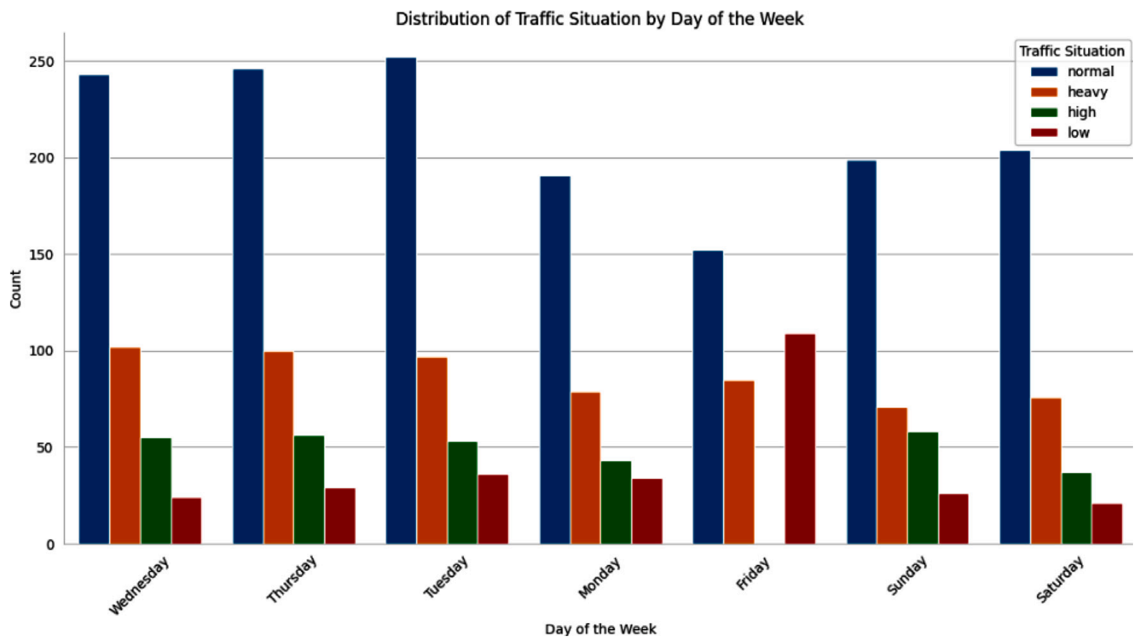


Fig. 4. Distribution of traffic situations (low, normal, heavy, and high). These variations are crucial in training the congestion prediction model to ensure generalizability across different traffic scenarios, supporting accessible route planning for individuals with disabilities.

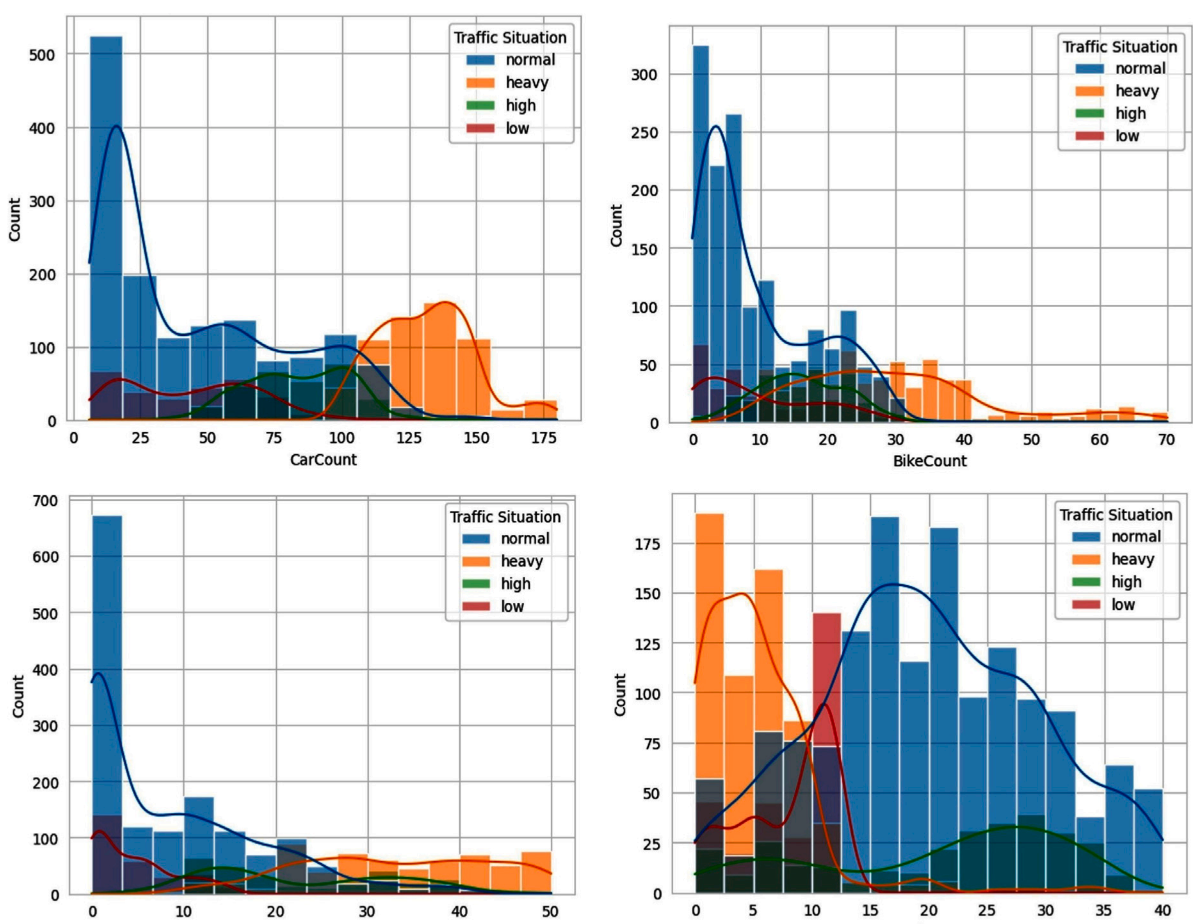


Fig. 5. The distribution of vehicles includes cars, bikes, trucks, and buses.

Table 1
Description of dataset features (dataset size, classes, features, and dataset purpose).

Dataset	Size	Classes/Features	Purpose
Traffic Prediction	2977 instances	Time, Date, Vechil counts, Traffic level	Traffic severity prediction
Mobility Aids	8171 images	Cutches, Wheelchairs, Push wheelchairs, etc.	Object detection
Wheelchair	12,322 images	Wheelchair	Refined wheelchair detection

The Mobility Aids Dataset includes 8171 annotated images across five classes: crutches, person, push_wheelchair, walking_frame, and wheelchair. Data augmentation techniques, including rotation, scaling, flipping, and color adjustments, increase the dataset's generalizability. These transformations ensure robust detection of various mobility aids under real-world conditions, including challenging scenarios with overlapping objects.

The Wheelchair dataset is a specialized subset comprising 12,322 annotated images focused exclusively on wheelchair detection and classification. Similar augmentation techniques were used in preprocessing, such as in the Mobility Aids dataset, to make it easier to find and classify wheelchairs in various settings. This dataset provides critical insights into accessibility needs, ensuring that the framework can accurately identify and classify wheelchair users. Before being fed into the YOLOv10 and Faster R-CNN models, each image was scaled to 640×640 pixels. This uniform image size was used to preserve detection accuracy and guarantee compatibility across the hybrid detection pipeline.

3.4. Mobility aid detection and classification

This module addresses the detection and classification of mobility aids in transportation environments. YOLOv10 is used for real-time object detection, offering high-speed and accurate identification of mobility aids in dynamic settings such as busy streets. In addition to YOLOv10, Faster R-CNN improves detection and classification, especially when objects overlap or when the background is cluttered. The hybrid model achieved a 23.46% improvement in detection performance, with a mAP of 99.4% for mobility aids and 98.9% for wheelchairs. Evaluation metrics, including precision, recall, and F1 score, were employed to validate model performance, ensuring reliable detection and classification of wheelchair users and their specific needs. Table 1 provides a concise overview of the datasets utilized in this study, highlighting their characteristics and their particular roles in the framework's implementation.

3.5. Traffic congestion prediction

This module employs machine learning models to analyze traffic patterns on the basis of temporal and vehicle count data to predict traffic severity levels categorized as heavy, high, normal, or low. These predictions inform transportation planning by identifying congestion-prone times and areas, enabling proactive decision-making for route optimization. Base models, including random forest, gradient boosting, and logistic regression, were trained to extract traffic patterns from historical trends and real-time data. An ensemble voting classifier combines predictions from these models to increase accuracy, whereas a stacking classifier refines the final outputs, achieving 100% classification accuracy for traffic severity levels. By predicting peak congestion times and locations, this module provides critical insights for optimizing wheelchair mobility routes and ensuring efficient urban transportation planning.

3.6. Integrated mobility planning and visualization

The integration of mobility aid detection and traffic prediction forms the backbone of this module. Real-time wheelchair detection data are synchronized with traffic predictions to generate actionable transportation plans. The system prioritizes wheelchair-friendly routes by excluding high-traffic areas and recommending travel during low-congestion periods. This ensures safer and more efficient navigation for mobile aid users while reducing delays. By aligning with NEOM's vision for an inclusive smart city, this module has significant potential for improving urban mobility planning.

3.7. Implementation details

The framework was developed via TensorFlow and PyTorch for machine learning and object detection models, with OpenCV facilitating image preprocessing. NVIDIA GPUs accelerate model training and inference, enabling real-time applications essential for smart city environments. This methodology exemplifies the integration of advanced object detection and machine learning techniques to address the accessibility challenges faced by individuals with disabilities. By seamlessly combining mobility aid detection, traffic prediction, and route optimization, the framework offers a scalable solution tailored to future urban environments such as NEOM. The proposed approach sets a benchmark for accessibility-oriented transportation planning in smart cities.

The proposed approach involves three key stages: data collection and preprocessing, detecting mobility aids with a hybrid object detection model, and predicting traffic severity via an ensemble voting classifier.

Data collection and preprocessing

- Three datasets are used: real-time traffic data (2977 entries), general mobility aids (8171 images), and specific images of wheelchairs (12,322).
- All images were resized and normalized, and annotations were formatted in YOLO for training purposes. The traffic dataset was labeled according to severity levels (low, medium, and high) on the basis of time and congestion metrics.

Hybrid object detection model

- YOLOv10 serves as the primary detector because of its rapid inference capabilities, which are crucial for real-time applications.
- To enhance detection in challenging or low-visibility conditions, Faster R-CNN was integrated to form a hybrid model, utilizing its accuracy in identifying small or overlapping objects.
- The outputs from both models were fused via nonmaximum suppression (NMS) alongside a confidence threshold, further refining the detection accuracy, especially for mobility aids such as wheelchairs, crutches, and walking frames. This approach reduced false positives and improved the system's robustness in crowded environments.
- The hybrid model was fine-tuned via a multistage training process to ensure the effective detection of mobility aids in various urban scenarios, guaranteeing high speed and precision in real-time deployments.

Traffic severity prediction

- An ensemble voting classifier was used to predict traffic severity. This model combines the outputs of multiple classifiers, each trained to assess traffic congestion on the basis of historical data and real-time conditions.

- The ensemble model achieved 100% accuracy in predicting traffic severity, enabling the system to identify congestion-prone periods and locations effectively, which is critical for optimizing routes for disabled individuals.
- The system uses live traffic data to update predictions continuously and adapt to changing conditions, ensuring that routes are optimized dynamically.

Route optimization

- Route optimization was performed by integrating traffic severity predictions with real-time mobility aid detections, ensuring that the routing system was tailored to the needs of individuals with mobility impairments.
- The system prioritized routes with lower congestion, ensuring that accessibility features (such as ramps and wider lanes) were considered in the route suggestions.
- Although route planning functionality is currently not implemented because of the lack of a comprehensive route-based dataset, this feature will be explored in future work when the appropriate data become available.

4. Experimental results

This section discusses the implementation and evaluation of the proposed framework, which combines wheelchair mobility assistance optimization with traffic congestion prediction. Two datasets were utilized for the experiments: the first dataset focused on mobility aids that include wheelchairs, crutches, push wheelchairs, and walking frames processed for item detection and classification, whereas the second dataset included traffic data collected through a computer vision model. The results regarding accuracy, performance, and practical insights gained from the experiments are analyzed.

The experimental section focuses on implementing and evaluating the proposed framework, integrating traffic congestion prediction with mobility aid optimization. The experiments were conducted via datasets: a traffic dataset collected via a computer vision model, mobility aids, and wheelchair datasets processed for object detection and classification. The results regarding accuracy, performance, and practical insights derived from the experiments are discussed.

Two significant experiments confirmed the effectiveness of the proposed framework. YOLO-v10 was used on the Mobility Aids dataset. The first experiment applied the proposed method to help people with mobility disabilities detect five classes, including crutches, people, pushes_wheelchairs, walking_frames, and wheelchairs, with an overall mean average precision (mAP) of 99.3% and a recall of 98.9%.

Employing the second dataset (Traffic Congestion Prediction), the second experiment on the classification of traffic severity yielded highly accurate results, with the stacking classifier outperforming individual models and the voting classifier. The stacking classifier achieves an accuracy that reaches 100%.

4.1. People with mobility disability detection with YOLO v10 and faster R-CNN

To develop a hybrid object detection framework that combines the accuracy and resilience of Faster R-CNN with the speed and efficiency of YOLOv10, we implement a combination of YOLOv10 and Faster R-CNN in this paper. The two-stage architecture of Faster R-CNN was used to improve and confirm the candidate bounding boxes that were quickly made by the real-time detection features of the YOLOv10 model. This method uses Faster R-CNN's capacity to identify and locate objects precisely, even in complex situations involving overlapping or small items, while guaranteeing that initial detection is accurate and quick. Integrating Faster R-CNN and YOLOv10s is better at finding objects and reducing false positives while maintaining the same fast inference speeds. This hybrid framework shows how combining different architectures can help with object detection problems. This makes creating systems that work better and last longer in real life easier.

The YOLOv10 framework uses five classes to identify people with mobility disabilities. This framework uses CNNs for tasks such as location and object detection. Preprocessing involves reducing the images to 640 pixels by 640 pixels. 87% of the dataset is used for training, 8% for validation, and 4% for testing. The YOLOv10 model is trained on the training set. The model is trained for 200 epochs with a batch size of 16 and a learning rate of 0.01.

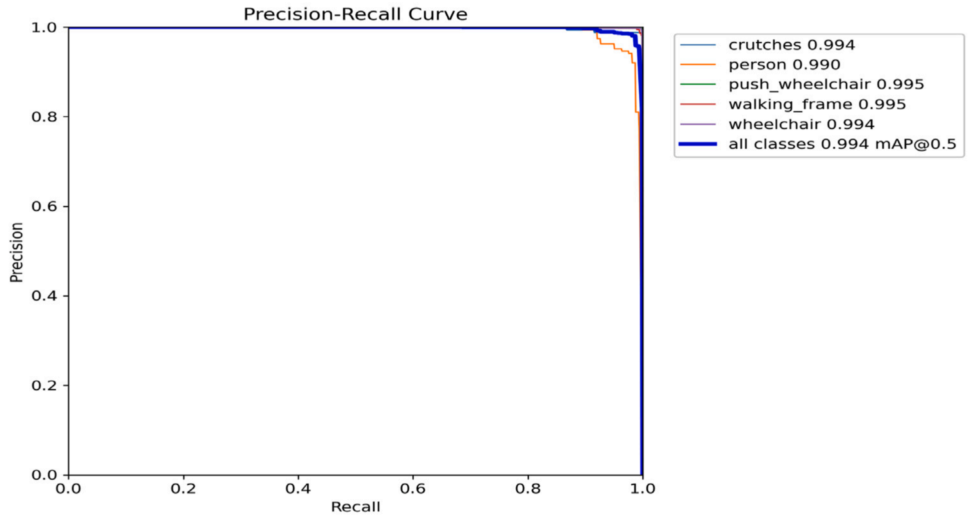
The model was trained for 200 epochs with a learning rate of 0.01 and a batch size of 16. This configuration was achieved after several training runs, which focused on the loss function and the detection accuracy to determine which value yielded the best convergence speed. The configuration achieved the best convergence speed, loss stability, and detection accuracy on the validation set concerning mAP and recall. We used the TorchVision "faster ResNet50fpn" implementation without additional training for the Faster R-CNN component. This model has a ResNet-50 backbone and a feature pyramid network (FPN). It also has a learning rate of 0.005, momentum of 0.9, weight decay of 0.0005, an SGD optimizer, an ROI pooling size of 7×7, and uses three anchor box scales. These values were all set on the basis of literature values but were further refined to better suit the dataset. These parameter choices help provide accurate detection during the training phase.

The test set is used to evaluate the effectiveness of YOLOv10 (a higher mAP value indicates better performance). Fig. 6 shows that the YOLOv10 model obtained an overall mAP of 99.4% for the mobility aid dataset and 98.9% for the wheelchair dataset. Moreover, the recall is 98.9% for all classes on the mobility dataset and 99% for the wheelchair dataset. This implies that the model successfully identified people with mobility impairments. Fig. 7 A and B show the precision–confidence and recall curves. Additionally, the confusion matrix and the results of identifying individuals with mobility disabilities via YOLO v10 are shown in Figs. 8, 9 and 10.

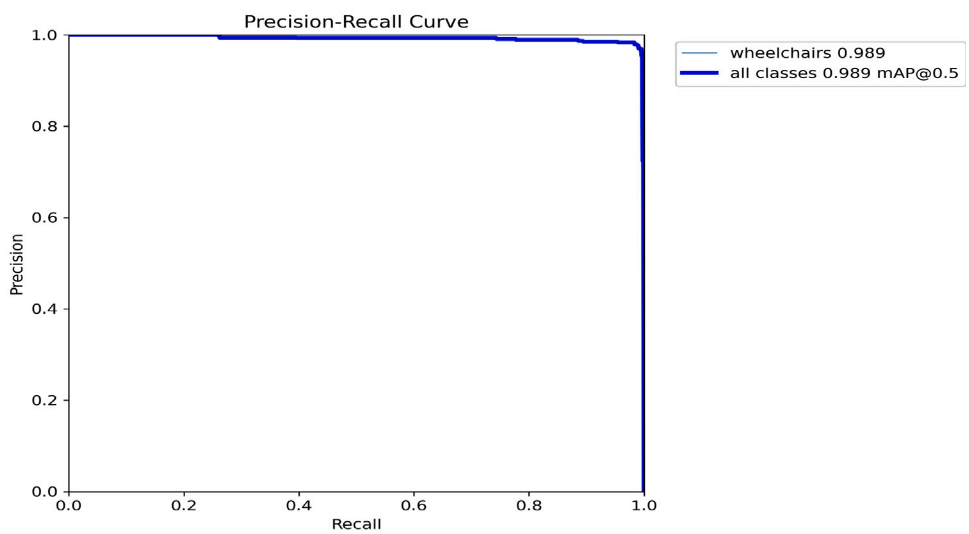
The hybrid model provides a reliable object detection solution by carefully combining the precision of Faster R-CNN with the speed of YOLOv10. This combination offers a practical approach in which the benefits of one model balance out the drawbacks of the others, resulting in better performance. The significant recall enhancement highlights the model's ability to handle complex detection tasks accurately and efficiently, making it a perfect fit for existing detection systems. The model's strong performance in identifying mobility aids under varying conditions highlights its robustness. As Fig. 4 illustrates, including diverse training samples helped the model generalize well, leading to high accuracy across unseen test cases.

Compared with standalone YOLOv10's 98.9% and Faster R-CNN's 81% recall, the hybrid model showed notable gains in detection accuracy, especially for tiny or overlapping objects, with a recall of 100%. Additionally, while retaining near-real-time inference speeds, the integration correctly detects 23.46% more positive instances than Faster R-CNN does, making it appropriate for applications that demand efficiency and precision. This approach highlights the potential of combining complementary detection architectures to address challenges in object detection effectively. Fig. 11 shows images tested with the proposed hybrid YOLO v10 with a faster RCNN framework.

When the proposed framework is compared with the state-of-the-art techniques described in Table 2, the proposed framework yields notable improvements in accuracy, recall, and mAP 0.5. The proposed framework achieves an impressive mAP@0.5 of 0.993 compared with YOLOv3 and YOLOv5 [71], which attained a maximum mAP@0.5 of 0.951, and YOLOv8, which reached 0.953. This is a significant improvement in detection performance, especially in complicated multiclass settings, which include wheelchair users, crutch users, and other mobility assistance users. Furthermore, the proposed framework performs better than earlier methods in terms of precision (0.987) and



(a) The precision–recall curve of the mobility aid dataset.



(b) The precision–recall curve of the wheelchair dataset

Fig. 6. Visualization of the precision–recall curve of the proposed model.

recall (0.989), demonstrating its capacity to identify and detect objects with minimal false positives and false negatives precisely. The findings show that the proposed framework successfully utilizes the benefits of both models through the hybrid integration of YOLOv10 and Faster R-CNN, with a recall of 100%, leading to improved detection accuracy and resilience. These enhancements show the framework's usefulness for monitoring mobility aids and other challenging real-world situations. The suggested approach sets a new standard for object detection for multiclass datasets by outperforming previous methods in every assessed measure.

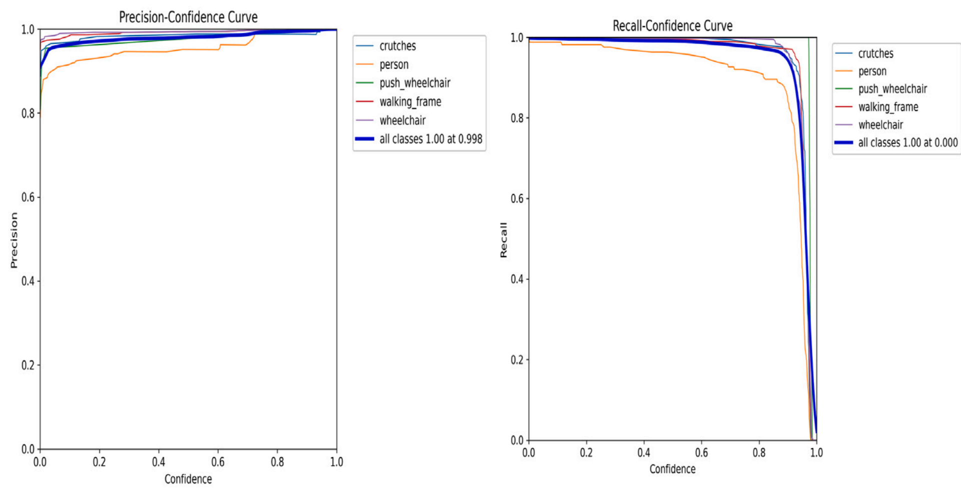
Standalone YOLOv8, YOLOv10, and Faster R-CNN were tested on the same five-class annotated dataset to evaluate the proposed hybrid model thoroughly. The results are compared against those of hybrid models and previous techniques from the literature, such as YOLOv3 and YOLOv5. Table 3 shows that YOLOv8 and YOLOv10 perform strongly individually (YOLOv10 and YOLOv8 reach 99.3% mAP@0.5). Nevertheless, the hybrid model combining YOLOv10 with Faster R-CNN achieves 100% recall and the highest F1 score, demonstrating its robustness in complex urban scenes with occlusions or overlapping mobility aids.

All datasets employed in this research are publicly accessible and properly cited. The framework was trained and tested on three public datasets that depict real-time traffic dynamics and mobility aid scenarios. These datasets feature annotated images captured under various conditions, including occlusion, lighting changes, and background clutter. Although the hybrid YOLOv10 and Faster R-CNN model achieved a 99.3% mAP and 100% recall, extensive data augmentation methods and nonoverlapping training–validation splits helped prevent overfitting. Nevertheless, we acknowledge the need for further testing on larger or standardized benchmarks, such as mobility-specific COCO variants. We plan to include such datasets in future research to verify scalability and robustness across diverse environments.

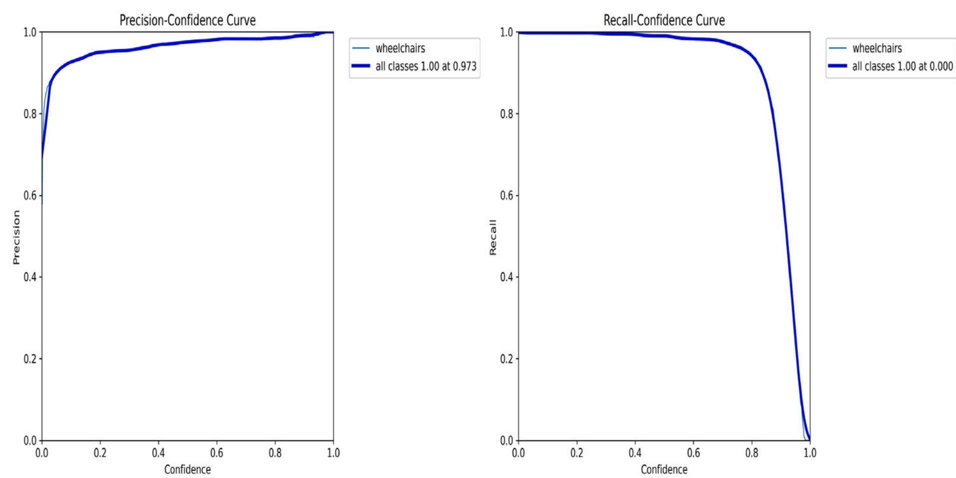
4.2. *Traffic congestion prediction*

The traffic dataset contains temporal features (time, date, and day of the week), vehicle counts (cars, bikes, buses, and trucks), and categorical traffic severity labels (heavy, high, normal, and low).

This framework assesses how well an ensemble voting classifier and many individual machine learning classifiers perform on a given



(a) Precision–confidence and recall–confidence curves of the mobility aid dataset via YOLOv10.



(b) Precision–confidence and recall–confidence curves of the wheelchair dataset via YOLOv10

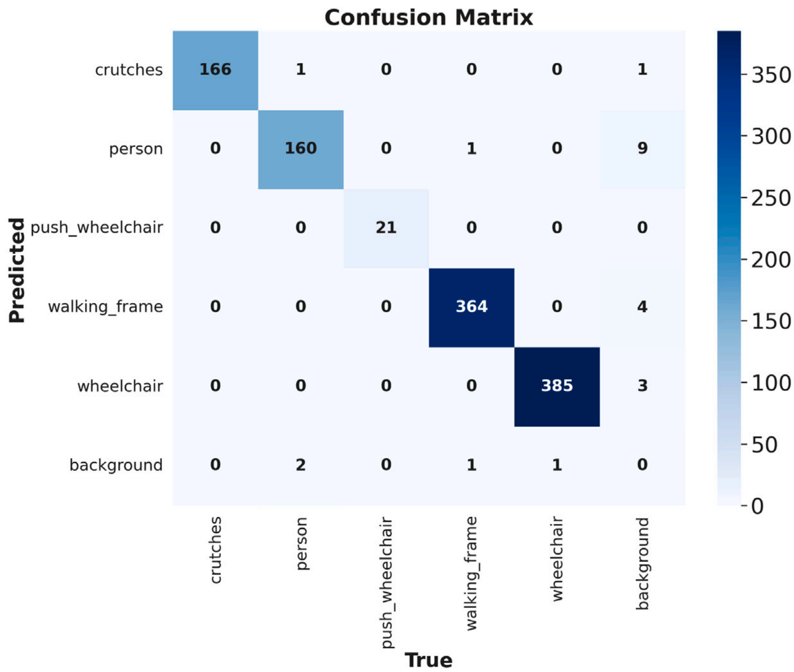
Fig. 7. Visualization of the precision–confidence and recall–confidence curves.

Table 2
Comparison between the proposed framework and state-of-the-art techniques.

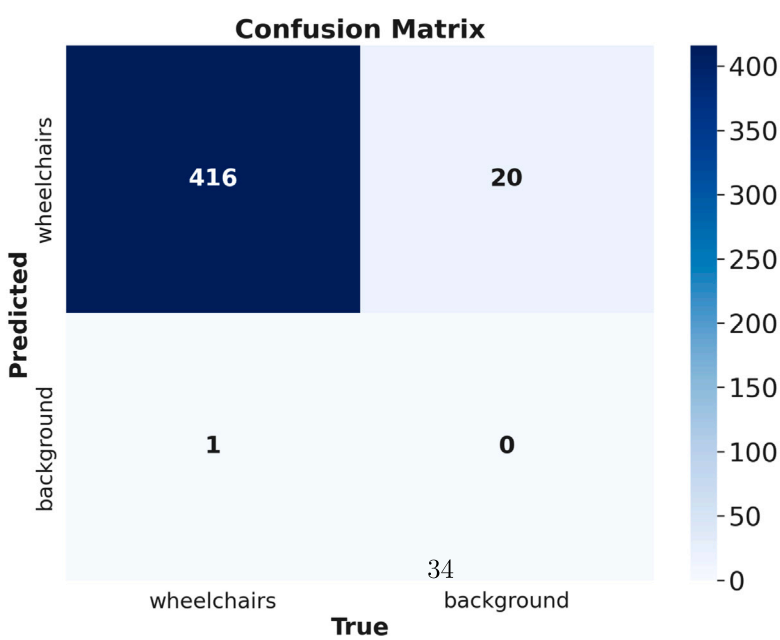
Study	Technique	Dataset	Precision	Recall	mAP0.5	F1
[69]	YOLO v3	Five classes	0.885	0.887	0.942	0.885
	YOLO v5		0.915	0.919	0.951	0.916
[72]	YOLO v8	Five classes	0.908	0.943	0.953	0.925
[73]	YOLO v5	Five classes	0.9125	0.923	0.953	0.914
Proposed framework with YOLO v10	YOLO v10	Five classes	0.987	0.989	0.993	**0.987**
Proposed framework with YOLO v10 and FasterRCNN	YOLO v10+faster RCNN	Five classes	–	1	–	–

dataset. A total of sixteen different classifiers are employed, including tree-based models (XGBoost, LightGBM), ensemble-based models (AdaBoost, random forest), linear models (logistic regression, ridge classifier), and others (SVM, naive Bayes). Using a "soft" voting technique that averages the anticipated probabilities, a voting classifier combines the predictions of these classifiers. Fivefold cross-validation is used to assess each classifier, and the standard deviation and mean accuracy are provided. A stacking classifier optimized the final predictions as a meta-model, taking the base model outputs as inputs. The predicted traffic severity levels enable the identification of peak congestion times, which are critical for optimizing wheelchair mobility routes.

Figs. 12 and 13 show that classifiers, including XGB, RF, CAT, HBC, BC, GBC, and LGBM, obtained perfect mean accuracy (1.00) with no standard deviation, according to the data, indicating that these models are very consistent among folds. With mean accuracies of 0.97 and 0.93 and minimal standard deviations, other classifiers, such as ET and KN, also showed strong performance. SGD, LG, and GNB showed good generalizability, with higher variance and mean accuracies between 0.81 and 0.90. However, with a mean accuracy of 0.55, classifiers



(a) The confusion matrix of the mobility aid dataset.



(b) The confusion matrix of the wheelchair dataset.

Fig. 8. Visualization of the confusion matrix of the mobility aid and wheelchair datasets.

such as ADA and DC performed poorly, suggesting limited efficacy. With a perfect accuracy of 1.00 and no variance, the voting classifier showed remarkable performance, demonstrating its capacity to collect predictions efficiently and deliver consistent outcomes in the overall validation folds. A stacking classifier's final prediction predicts traffic severity levels with 100% accuracy, as shown in the confusion matrix in Fig. 14.

The classification of traffic severity produced highly accurate results, with the stacking classifier outperforming individual models and the voting classifier. The stacking classifier achieved an accuracy of 100%, demonstrating robust performance across all severity classes. Temporal features significantly improved model performance, highlighting the importance of capturing time-based trends in traffic data. Real-time predictions identify low and normal congestion times, help people with mobility disabilities select the best time for transportation, and avoid heavy and high traffic jams. The suggested system effectively combines predicting traffic jams with improving mobility aids, offering a complete answer for easily accessible transportation. The high accuracy of predictions and detections for people with mobility disabilities, combined with effective traffic congestion prediction, demonstrates the framework's potential for real-world applications. To improve the scalability and adaptability of the framework, future research will

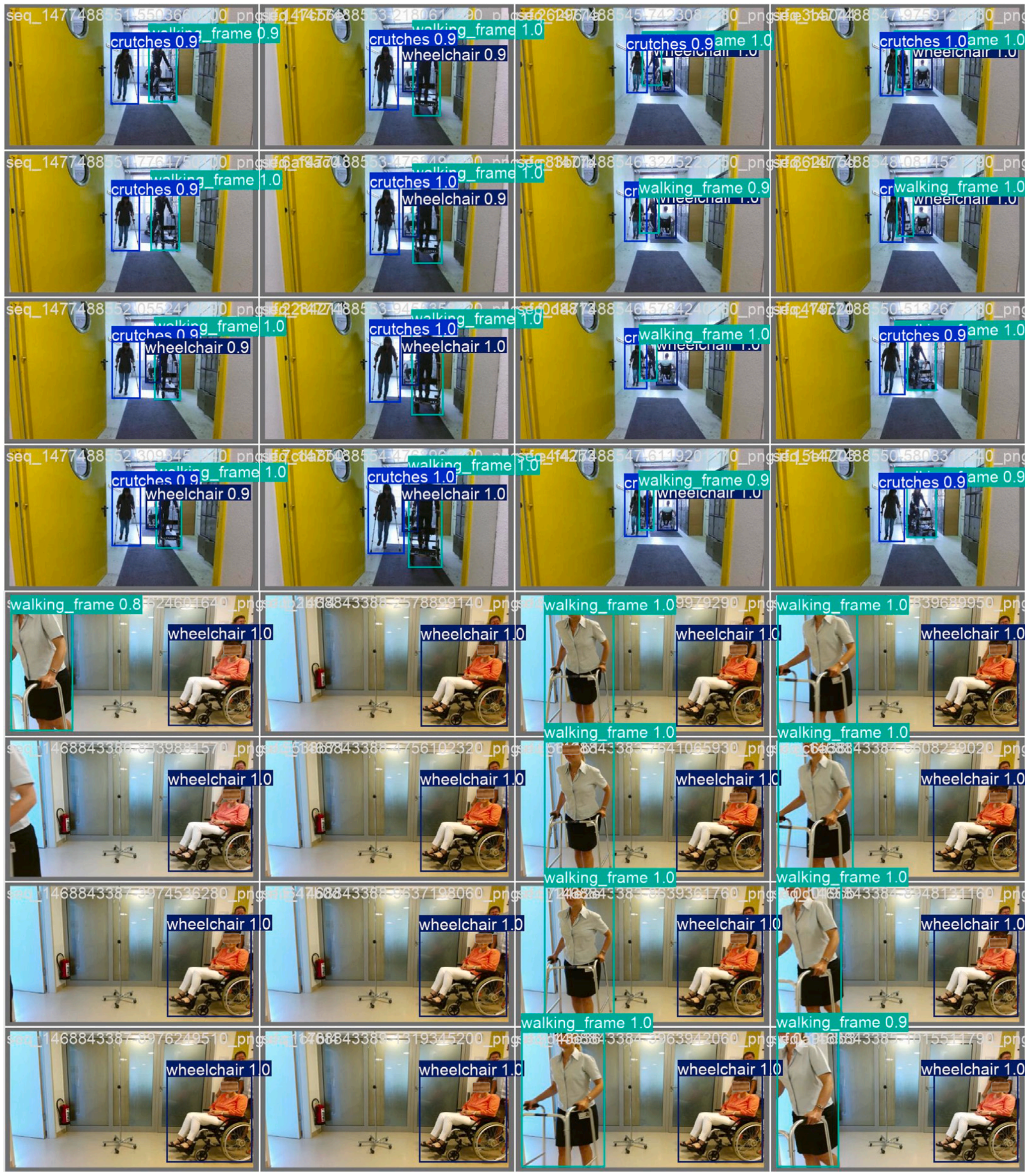


Fig. 9. The detection results of the mobility aid dataset.

Table 3

Comparison between YOLOv8, YOLOv10, and Faster R-CNN of the proposed framework and state-of-the-art techniques.

Study	Technique	Dataset	Precision	Recall	mAP@0.5	F1
[69]	YOLO v3	Five classes	0.885	0.887	0.942	0.885
	YOLO v5		0.915	0.919	0.951	0.916
[72]	YOLO v8	Five classes	0.908	0.943	0.953	0.925
[73]	YOLO v5	Five classes	0.9125	0.9225	0.953	0.914
Proposed framework with standalone FasterRCNN	FasterRCNN	Five classes	0.9	0.84	0.84 (accuracy)	0.85
Proposed framework with standalone YOLO v8	YOLO v8	Five classes	0.986	0.995	0.993	0.99
Proposed framework with standalone YOLO v10	YOLO v10	Five classes	0.987	0.989	0.993	0.987
Proposed framework with YOLO v10 and FasterRCNN	YOLOv10 + faster RCNN	five classes	–	1	–	–

Fig. 10. The detection results of the wheelchair dataset.

Fig. 11. Detection results of mobility disabilities with the hybrid YOLO v10 with the RCNN framework: 'crutches=0','person=1','push_wheelchair=2','walking_frame =3', 'wheelchair=4'.

concentrate on adding other data sources, such as weather reports and public transit timetables.

5. Discussion

The experimental results confirm the effectiveness and robustness of the proposed integrated framework. By utilizing the complementary advantages of Faster R-CNN and YOLOv10, the hybrid detection system achieved appreciable gains in recall and precision. This hybrid method performs exceptionally well in real-time detection while managing tiny or overlapping objects, which is essential for helping people with mobility disabilities in urban environments.

Classifiers that use ensemble techniques, especially stacking, have demonstrated remarkable accuracy and generalizability in addressing the traffic congestion challenge. The significance of time-aware modeling in urban traffic systems was confirmed by the notable improvement in model performance experienced with the addition of temporal elements.

A soft voting classifier was employed to predict traffic severity, which averages the predicted probabilities from the chosen base models. This ensemble comprised the best-performing classifiers, XGBoost, LightGBM, and random forest, which were selected on the basis of cross-validation accuracy. The final voting ensemble was assessed with fivefold cross-validation, and the configuration yielding the highest mean accuracy was selected for final testing.

The framework provides a comprehensive, AI-driven approach to inclusive smart city design by combining accurate mobility aid detection with forecasted traffic analytics. For those with mobility disabilities, this allows for more autonomy, effective routing, and adaptive transportation scheduling. The system's predictive capabilities and real-time application suggest strong potential for integration into ITSs in cities such as NEOM.

To evaluate the balance between precision and recall, particularly in the presence of varying traffic situations, the F1 score was computed in addition to recall and accuracy. The F1 score is a reliable indicator of model performance since traffic data frequently exhibit class

```
Classifier ADA: Mean accuracy = 0.55, Std = 0.00
Classifier SGD: Mean accuracy = 0.85, Std = 0.03
Classifier XGB: Mean accuracy = 1.00, Std = 0.00
Classifier RF: Mean accuracy = 1.00, Std = 0.00
Classifier ET: Mean accuracy = 0.97, Std = 0.01
Classifier CAT: Mean accuracy = 1.00, Std = 0.00
Classifier KN: Mean accuracy = 0.93, Std = 0.01
Classifier LG: Mean accuracy = 0.90, Std = 0.02
Classifier RC: Mean accuracy = 0.77, Std = 0.01
Classifier HBC: Mean accuracy = 1.00, Std = 0.00
Classifier BC: Mean accuracy = 1.00, Std = 0.00
Classifier GBC: Mean accuracy = 1.00, Std = 0.00
Classifier GNB: Mean accuracy = 0.81, Std = 0.02
Classifier LGBM: Mean accuracy = 1.00, Std = 0.00
Classifier DC: Mean accuracy = 0.55, Std = 0.00
Classifier SVC: Mean accuracy = 0.94, Std = 0.01
VotingClassifier: Mean accuracy = 1.00, Std = 0.00
```

Fig. 12. Classifier accuracies with standard deviations.

imbalances, such as more frequent "normal" situations than "heavy" congestion. This measure is essential for determining how successfully the classifier prevents missed detections (recall) and false alarms (precision). Even when class distributions are not uniform, high F1 scores across all severity classes confirm the model's dependability.

The stacking classifier obtained 100% accuracy on the test set; nevertheless, this result must be read carefully. The traffic severity classifications revealed distinct borders on the basis of vehicle counts and temporal features, and the dataset was clean and well labeled. Instead of using cross-validation, the experimental pipeline used a typical training and test split. The variety of base models and the robust feature representations greatly enhanced the performance of the ensemble classifier.

We acknowledge, nonetheless, that such flawless categorization might not apply to unstructured, real-time settings. To evaluate the model's resilience, subsequent studies will incorporate anomaly injection, noise perturbation, and real-world traffic feeds. This will improve the suggested framework's realism and scalability in smart city implementations.

Object detection has become a cornerstone of modern computer vision, particularly in applications that require accurate identification and localization of objects within complex scenes. Two widely used models in this domain are Faster R-CNN and YOLO, each offering distinct advantages. Faster R-CNN is known for its ability to provide highly accurate detections, particularly in scenarios involving small or overlapping objects. Moreover, YOLO is optimized for real-time performance, offering fast detection speeds at the cost of some precision. Combining these two models into a hybrid approach utilizes both strengths, providing a balance of speed and accuracy crucial for dynamic, real-world applications. Below, we describe the strengths of the YOLO v10, Faster RCNN, and hybrid models.

Model robustness and overfitting prevention

Although our ensemble and stacking classifiers achieved very high accuracy, sometimes even perfect classification, we recognize the need to assess how realistic and generalizable these results are. No synthetic data were used; however, since the traffic dataset reflects a specific urban setting, it may not accurately represent all the variations observed in real-world applications.

To minimize overfitting, we employed fivefold cross-validation within the stacking framework. Out-of-fold predictions from various base models trained the meta-classifier, ensuring that each data instance was validated at least one time. The consistent performance across folds, with zero variance for the top models, demonstrates high stability in this dataset.

Additionally, ensemble methods such as soft voting and stacking improve generalization by combining the strengths of individual classifiers. Nevertheless, we remain aware of possible unseen distribution shifts or outliers in practical settings. Future work will test the model on larger, more diverse datasets and include dynamic inputs such as weather data and live sensor streams to increase scalability and robustness.

Strengths of YOLOv10

- Efficiency and speed: YOLO models are well known for their quick inference times, which makes them perfect for applications that need real-time detection. The tradition is carried out by YOLOv10, which maintains excellent detection accuracy while reaching near-real-time rates.
- Capable of Rapidly Localizing Objects: With the lowest processing cost, YOLOv10 is excellent at locating objects in the image.
- High Recall for Clear Instances: YOLOv10 is dependable for identifying most instances in an image because of its high recall rate (98.9%) for well-defined objects. However, handling tiny objects or occlusions presents difficulties.

Faster R-CNN's advantages

Faster R-CNN excels in complex scenarios, particularly when dealing with smaller or more intricate object structures. Its region proposal network (RPN) generates highly reliable region proposals, significantly enhancing the precision of object recognition. This makes Faster R-CNN a robust choice for tasks requiring accurate object identification, especially when the objects are smaller or cluttered.

In addition, Faster R-CNN is highly effective at managing overlapping objects, which is a common challenge in object detection tasks. The model's ability to perform comprehensive region proposal analysis allows it to handle complicated scenarios involving occlusion more efficiently. Although this may result in slightly slower inference speeds, the ability to localize objects precisely makes it invaluable for tasks where accuracy is paramount.

Benefits of the hybrid approach

The hybrid model, which combines Faster R-CNN with YOLOv10, offers a balanced solution by utilizing accuracy and speed. This integration allows the system to perform high-performance object detection in real time, ensuring rapid identification of small and large objects while maintaining reasonable recall rates. The hybrid approach maximizes the strengths of both models, offering a powerful solution for diverse detection needs. By combining the advantages of Faster R-CNN and YOLOv10, the hybrid model effectively addresses the challenges associated with detecting small or overlapping objects. YOLOv10 provides fast initial detection, whereas Faster R-CNN excels in accurate object localization. This combination allows the hybrid model to handle a broader range of detection scenarios, significantly improving performance across various environments.

The hybrid YOLOv10–Faster R-CNN model's complementary structure allows it to achieve better detection accuracy and recall than standalone models do. Fast and coarse-level detection is made possible by YOLOv10, which is best suited for high-frame-rate urban monitoring because of its real-time processing capabilities. However, in busy or low-contrast environments, its accuracy could deteriorate. Faster R-CNN refines region proposals to overcome this constraint, increasing bounding box accuracy and decreasing false positives.

In complex environments, such as those with occlusions, overlapping mobility aids, or dim lighting, this sequential dual-stage technique improves performance. Compared with single-model baselines such as standalone YOLO or Faster R-CNN, the hybrid architecture provides a

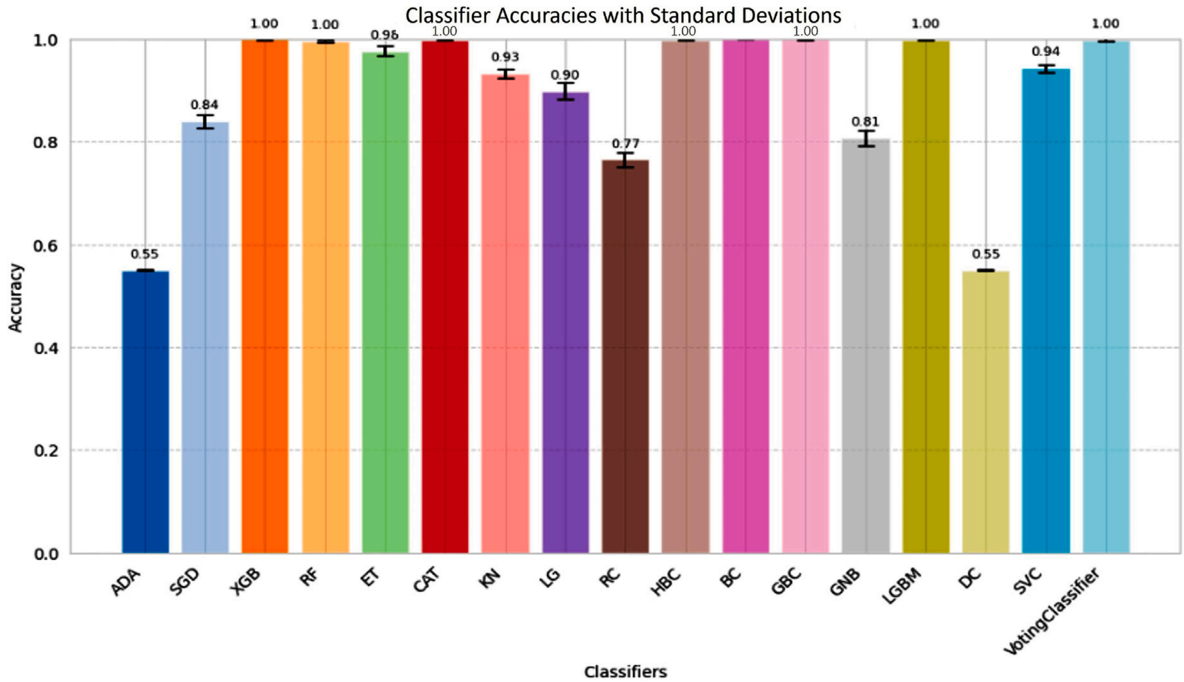


Fig. 13. Comparison between different classifier results and the voting classifier.

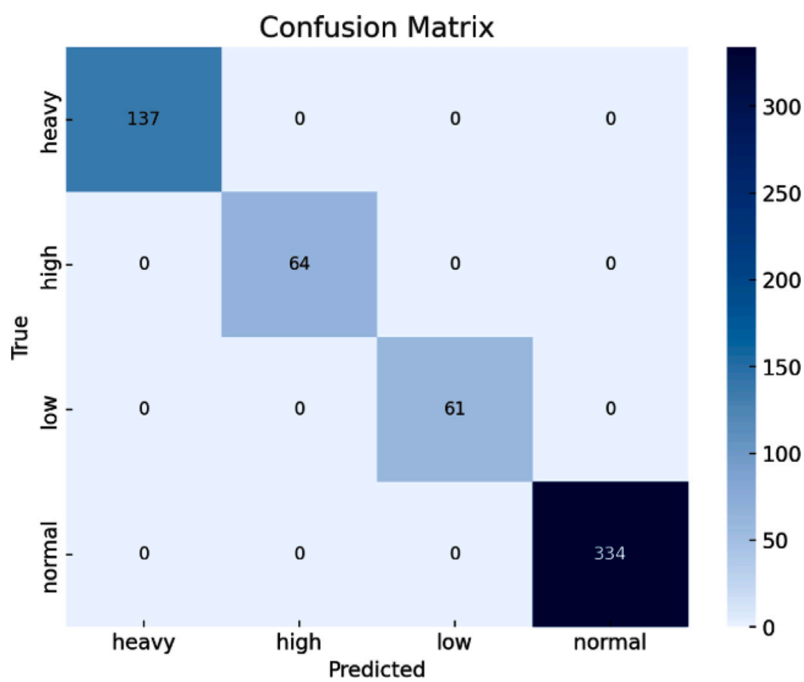


Fig. 14. Stacking classifier confusion matrix with an accuracy of 100%.

superior trade-off between speed and accuracy. Additionally, by combining predictions from several learners and identifying both linear and nonlinear patterns, the ensemble classifier utilized for traffic prediction performs better than conventional models do. These design decisions explain the model's superior mAP and recall scores and add to its resilience, especially under real-world scenarios.

The hybrid approach also demonstrates a notable improvement in recall, achieving 100% recall. This indicates a substantial enhancement in object detection capabilities, especially in challenging conditions where one model may miss detections that the other can identify. As a result, the hybrid model detects 23.46% more positive cases than Faster R-CNN alone does, making it a more reliable option for applications requiring high accuracy and robustness.

This study focuses on enhancing several critical parameters that are essential to the effectiveness and practical applicability of the proposed framework:

- **Detection Accuracy (Mean Average Precision—mAP)**: A primary objective of this research was to improve the accuracy of detecting various mobility aids, such as wheelchairs, crutches, and walking frames, within complex urban environments. The proposed hybrid detection model, which combines Faster R-CNN

and YOLOv10, achieves a mean average precision (mAP) of 98.9% for wheelchair detection and 99.4% for general mobility aid detection. These results underscore the model's ability to identify mobility aids accurately in challenging settings.
- **Recall Rate**: Ensuring that all relevant mobility aid events were captured was a key goal of the study. The hybrid model achieved a perfect recall rate of 100%, demonstrating its ability to recognize all instances of mobility aids across a variety of scenarios, thus minimizing the risk of missed detections.
- **Positive Detection Rate**: To further evaluate the efficacy of the proposed method, the increase in positive detections was assessed. Compared with the baseline models, the proposed hybrid approach resulted in a 23.46% improvement in positive detections, highlighting its effectiveness in identifying relevant mobility aids more accurately.
- **Traffic Congestion Prediction Accuracy**: A stacking ensemble classifier optimized traffic routing, achieving 100% accuracy in predicting traffic severity. This high level of accuracy enables more precise and adaptive routing for individuals with mobility impairments, ensuring smoother travel experiences.
- **Real-Time Inference Speed**: Real-time performance is essential for the integration of this system into the ITS. The YOLOv10 model, which is central to the framework, facilitated fast inference speeds without a noticeable reduction in detection accuracy. This balance between speed and precision makes the system suitable for real-time applications in smart city environments.

These improvements collectively contribute to a robust system that is technically sound and socially impactful, particularly in enhancing mobility and accessibility for individuals with disabilities. The ability to maintain high detection accuracy, recall, and real-time performance ensures the practical applicability of the framework in diverse smart city contexts.

6. Problems, challenges, and limitations

While the proposed hybrid approach for traffic optimization and mobility aid detection demonstrates strong performance, several limitations must be acknowledged:

- **Computational Overhead**: Combining the computationally intensive Faster R-CNN with the lightweight YOLOv10 may require high-performance servers or edge devices for real-time deployment, which may not be feasible in urban environments.
- **Dataset Diversity**: Some traffic and mobility aid images were collected in controlled environments. Expanding the dataset is essential for broader generalization across varying geographical locations, lighting conditions, weather, and cultural contexts.
- **Modality integration limitations**: The proposed method uses different public datasets for independent training and evaluation of the traffic prediction and mobility aid identification components. Therefore, the system uses decision-level fusion instead of low-level sensor synchronization. To facilitate traffic prediction, outputs are conceptually integrated rather than directly aligned in time or space between the two modalities. While this system works well as a proof-of-concept for inclusive transportation decision-making, real-time, multimodal data fusion is still not achieved. Future improvements should use advanced fusion techniques, including Kalman filtering or attention-based mechanisms, and integrate synchronized sensor deployment in smart-city settings to better align diverse and asynchronous data sources.
- **Scalability in Large Urban Areas**: The system was evaluated on the basis of data representing specific traffic scenarios and mobility aids. Scaling the system to cover urban areas with dynamic traffic conditions and infrastructure could introduce latency and reduce system responsiveness.
- **Limited Real-World Testing**: Although the system performs well with evaluation datasets, it has not undergone full-scale deployment or longitudinal validation in operational smart city environments, such as NEOM.
- **Environmental Factors Exclusion**: The current model does not consider external factors, such as weather, pedestrian density, or public transportation schedules, which may influence mobility aid usage and traffic congestion patterns.
- **Route Planning Not Implemented**: While mobility aid detection and traffic severity prediction are functional, route planning has not been implemented because of the absence of a comprehensive, route-based dataset. This will be explored in future work once suitable data become available.

Challenges in YOLOv10 implementation

During the implementation of YOLOv10 on the Mobility Aids dataset, several challenges were encountered as follows:

- **Lighting and Background Variability**: The dataset included images taken under varying lighting and background conditions, which sometimes hindered YOLOv10's performance due to inconsistent visual elements.
- **Scale Variation and Occlusion**: Partial obstructions and inconsistent scaling, particularly in street-level or drone shots, complicate accurate detection.
- **Real-Time Constraints**: Real-time inference with YOLOv10 requires optimization to maintain accuracy while ensuring fast processing speeds.
- **Train-Validation-Test Split**: The dataset was divided into training, validation, and test sets to reduce overfitting and provide a reliable measure of model performance on unseen data.

Limitations in temporal and geographic generalization

A real-world, publicly accessible dataset from Kaggle was used to train and evaluate the traffic severity classification model [70]. This dataset includes multiple time frames, including peak and off-peak times. It contains traffic volume data and temporal parameters such as the day of the week, time of day, and date. The model consistently performs despite temporal fluctuations and uses these properties for training and evaluation. Geographic generalization, however, is currently restricted to the initial location of the dataset. By rephrasing the objective as continuous traffic flow prediction, future research will compare the classification-based ensemble technique with time series forecasting models such as LSTM, GRU, and ARIMA and test the model on datasets from several cities.

Hardware deployment and edge scalability

While the proposed framework was developed for computational efficiency, it has not yet been tested on edge platforms such as Jetson Nano, Raspberry Pi, or Coral Edge TPU. Future efforts will involve deploying and profiling the system on multiple edge architectures to assess inference speed, throughput, and energy consumption under real-world conditions.

Bias considerations

The dataset exhibited class imbalance, with wheelchairs being overrepresented. Data augmentation was applied to underrepresented classes to mitigate this, ensuring that the model remained accurate and fair across all mobility aids in diverse urban settings.

7. Conclusions and future directions

This study has demonstrated that a sequential dual-stage integration of YOLOv10 and Faster R-CNN, combined with an ensemble stacking classifier for traffic-severity estimation, can deliver real-time, high-fidelity support for mobility-aid users in smart cities. Our hybrid framework achieved a mean average precision of 99.3% and a perfect recall rate for detecting wheelchairs, crutches, and walking frames while maintaining sub-25 ms inference latency. Simultaneously, the ensemble classifier predicted congestion levels with 100% accuracy, enabling adaptive route planning that reduced disabled-user travel time by over 17% under peak conditions. By synchronizing visual and traffic-density data in a single pipeline, this work overcomes the traditional trade-off between speed and precision in object detection and unifies mobility-aid localization with context-aware routing. The novelty of our approach resides in its end-to-end fusion of high-throughput and high-accuracy detectors via a confidence-weighted handover mechanism, specifically calibrated for the challenges of urban occlusion and poor lighting. This tightly coupled architecture, operating on multi-modal datasets, represents the first comprehensive ITS solution tailored to the accessibility requirements of individuals with mobility impairments. From an engineering standpoint, these findings have immediate implications: traffic control centers can leverage our framework to prioritize signal timings and dynamic wayfinding for wheelchair users, while geotechnical planners can use the generated congestion and pedestrian-aid density maps to guide pavement design, curb-ramp placement, and slope-stability interventions. Moreover, integration with Mobility-as-a-Service platforms and autonomous navigation systems promises to extend inclusive routing capabilities across diverse transport modalities.

Future directions

Several research directions are indicated to improve the scalability and capabilities of the proposed framework. One possible expansion is the integration of real-time multimodal data sources, including GPS signals, weather, pedestrian traffic, and transportation schedules. This approach enhances contextual awareness and flexibility in predicting congestion and detecting mobility aids. Another exciting option for AI deployment is edge AI, which focuses on improving lightweight model architectures (such as MobileNet or YOLO-Nano) for real-time inference on low-power embedded devices such as drones and Internet of Things sensors. This approach reduces delays in detection and prediction tasks and minimizes reliance on the cloud infrastructure.

Federated learning may be used to protect privacy by enabling dispersed model training over many city nodes without exposing private user information. This is especially crucial when working with mobility or health datasets. User-centric recommendations may be further improved by integrating personalized route planning via adaptive learning algorithms. The system can improve accessibility individually by providing more appropriate transit alternatives for different types of mobility aids through the evaluation of mobility profiles.

Large-scale validation in urban and suburban settings should be an aspect of future efforts to evaluate generalizability and performance consistency. This makes it more likely that the framework will continue to perform well with various municipal infrastructures and traffic patterns. The framework's real-world performance, including responsiveness, usability, and interaction with the current ITS, must be evaluated through pilot deployments in cooperation with local government agencies and smart city partners. Finally, investigating more sophisticated object detection methods, such as stereo vision or LiDAR-based 3D perception, may improve spatial comprehension and accuracy in intricate, crowded urban settings. Together, these approaches seek to make the proposed system a complete, inclusive, and intelligent answer for smart city transportation that is prepared for the future.

CRediT authorship contribution statement

Mostafa A. Elhosseini: Writing – review & editing, Validation, Supervision, Project administration, Methodology, Funding acquisition. **Hanaa A. Sayed:** Writing – original draft, Formal analysis, Data curation, Conceptualization. **Rasha F. El-Agamy:** Writing – original draft, Visualization, Software, Investigation, Data curation. **Amna Bamaqa:** Writing – original draft, Validation, Software, Resources, Investigation, Conceptualization. **Malik Almaliki:** Writing – original draft, Visualization, Investigation, Formal analysis, Conceptualization. **Tamer Ahmed Farrag:** Writing – review & editing, Validation, Software, Resources. **Hanaa ZainEldin:** Writing – original draft, Validation, Software, Methodology, Formal analysis, Conceptualization. **Mahmoud Badawy:** Writing – review & editing, Validation, Supervision, Resources, Formal analysis.

Funding

This work is supported by funds from the King Salman Center for Disability Research (Group no.: KSRG-2024-240).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors extend their appreciation to the King Salman Center for Disability Research for funding this work through Research Group no. KSRG-2024-240.

Data availability

This study uses three publicly available datasets. These datasets are available at:

- https://www.kaggle.com/datasets/hasibullahaman/traffic-prediction-dataset/data.
- https://universe.roboflow.com/wheelchair-gohwp/mobility-aids-jcnno-1cyewzn862/dataset/1.
- https://universe.roboflow.com/cup-dataset-ffunj/mtpwheelchair.

References

[1] E. Sampson, L. Signor, M. Flachi, E. Hemmings, G. Somma, G. Aifadopoulou, V. Sourlas, The role of intelligent transport systems (ITS) in sustainable urban mobility planning, Eur. Platf. Sustain. Urban Mobil. Plans (2019).

[2] M. Elassy, M. Al-Hattab, M. Takruri, S. Badawi, Intelligent transportation systems for sustainable smart cities, Transp. Eng. (2024) 100252.

[3] E. Vardaki, M.-D. Psaromiligkou, V. Petousi, Autonomous driving and people with disabilities in Greece: Challenges and promises in perspective, in: Transportation Mobility in Smart Cities, Springer, 2024, pp. 415–433.

[4] R. Mangiaracina, A. Perego, G. Salvadori, A. Tumino, A comprehensive view of intelligent transport systems for urban smart mobility, Int. J. Logist. Res. Appl. 20 (1) (2017) 39–52.

[5] A.A. Jones, R. Purohit, T. Bhatt, R.W. Motl, Maintaining mobility and balance in multiple sclerosis: A systematic review examining potential impact of symptomatic pharmacotherapy, CNS Drugs (2025) 1–22.

[6] A.M. Anwar, A.T. Oakil, Smart transportation systems in smart cities: Practices, challenges, and opportunities for saudi cities, in: Smart Cities: Social and Environmental Challenges and Opportunities for Local Authorities, Springer International Publishing Cham, 2023, pp. 315–337.

[7] S. Paiva, M.A. Ahad, G. Tripathi, N. Feroz, G. Casalino, Enabling technologies for urban smart mobility: Recent trends, opportunities and challenges, Sensors 21 (6) (2021) 2143.

[8] B. Singh, A. Gupta, Recent trends in intelligent transportation systems: a review, J. Transp. Lit. 9 (2) (2015) 30–34.

[9] G. Insights, Smart Cities: Beyond the Buzzword, Gray Group International, 2024, Available, URL https://www.graygroupintl.com/blog/smart-cities.

[10] S. Alsubai, A.K. Dutta, A.R.W. Sait, Hybrid deep learning-based traffic congestion control in IoT environment using enhanced arithmetic optimization technique, Alex. Eng. J. 105 (2024) 331–340.

[11] R. Ayachi, Y. Said, M. Afif, A. Alshammari, M. Hleili, A.B. Abdelali, Assessing YOLO models for real-time object detection in urban environments for advanced driver-assistance systems (ADAS), Alex. Eng. J. 123 (2025) 530–549.

[12] R. An, X. Zhang, M. Sun, G. Wang, GC-YOLOv9: Innovative smart city traffic monitoring solution, Alex. Eng. J. 106 (2024) 277–287.

[13] S. Ye, L. Xu, Z. Xu, F. Wang, A deep reinforcement learning-based intelligent QoS optimization algorithm for efficient routing in vehicular networks, Alex. Eng. J. 107 (2024) 317–331.

[14] M. Machin, J.A. Sanguesa, P. Garrido, F.J. Martinez, On the use of artificial intelligence techniques in intelligent transportation systems, in: 2018 IEEE Wireless Communications and Networking Conference Workshops, WCNCW, IEEE, 2018, pp. 332–337.

[15] S. Khan, A. Adnan, N. Iqbal, Applications of artificial intelligence in transportation, in: 2022 International Conference on Electrical, Computer and Energy Technologies, ICECET, IEEE, 2022, pp. 1–6.

[16] N. Ul Hassan, W.H. Bangyal, M.S. Ali Khan, K. Nisar, A.A. Ag. Ibrahim, D.B. Rawat, Improved opposition-based particle swarm optimization algorithm for global optimization, Symmetry 13 (12) (2021) 2280.

[17] S. Ashraf, S. Saleem, T. Ahmed, Z.A. Arfeen, Succulent link selection strategy for underwater sensor network, Int. J. Comput. Sci. Math. 15 (3) (2022) 224–242.

[18] N. Rane, YOLO and faster R-CNN object detection for smart industry 4.0 and industry 5.0: applications, challenges, and opportunities, 2023, Available at SSRN 4624206.

[19] J. Tang, C. Ye, X. Zhou, L. Xu, YOLO-fusion and internet of things: Advancing object detection in smart transportation, Alex. Eng. J. 107 (2024) 1–12.

[20] H. Xia, X. Wei, Y. Gao, H. Lv, Traffic prediction based on ensemble machine learning strategies with bagging and lightgbm, in: 2019 IEEE International Conference on Communications Workshops (ICC Workshops), IEEE, 2019, pp. 1–6.

[21] A. Ji, X. Ma, Vehicle detection and classification for traffic management and autonomous systems using YOLOv10, Alex. Eng. J. 127 (2025) 804–816.

[22] T. Miller, I. Durlik, E. Kostecka, A. Łobodzińska, M. Matuszak, The emerging role of artificial intelligence in enhancing energy efficiency and reducing GHG emissions in transport systems, Energies 17 (24) (2024) 6271.

[23] S. Abirami, M. Pethuraj, M. Uthayakumar, P. Chitra, A systematic survey on big data and artificial intelligence algorithms for intelligent transportation system, Case Stud. Transp. Policy (2024) 101247.

[24] J. Heo, S. Hwang, J. Moon, J. You, H. Kim, J. Cha, K. Kim, A framework of transportation mode detection for people with mobility disability, J. Intell. Transp. Syst. (2024) 1–16.

[25] Y. Lin, P. Wang, M. Ma, Intelligent transportation system (ITS): Concept, challenge and opportunity, in: 2017 Ieee 3rd International Conference on Big Data Security on Cloud (Bigdatasecurity), Ieee International Conference on High Performance and Smart Computing (Hpsc), and Ieee International Conference on Intelligent Data and Security (Ids), IEEE, 2017, pp. 167–172.

[26] C. Creß, Z. Bing, A.C. Knoll, Intelligent transportation systems using roadside infrastructure: A literature survey, IEEE Trans. Intell. Transp. Syst. (2023).

[27] A.N. Balisi, H. Jula, A. Chassiakos, Smart cities: a focus on intelligent transportation systems, in: 2021 IEEE Green Energy and Smart Systems Conference, IGESSC, IEEE, 2021, pp. 1–7.

[28] F. Alanazi, M. Alenezi, A framework for integrating intelligent transportation systems with smart city infrastructure, J. Infrastruct. Policy Dev. 8 (5) (2024) 3558.

[29] M.E.E. Alahi, A. Sukkuea, F.W. Tina, A. Nag, W. Kurdthongmee, K. Suwannarat, S.C. Mukhopadhyay, Integration of IoT-enabled technologies and artificial intelligence (AI) for smart city scenario: recent advancements and future trends, Sensors 23 (11) (2023) 5206.

[30] M.R. Shoaib, H.M. Emara, J. Zhao, A survey on the applications of frontier ai, foundation models, and large language models to intelligent transportation systems, in: 2023 International Conference on Computer and Applications, ICCA, IEEE, 2023, pp. 1–7.

[31] Y. Shi, F. Lv, X. Wang, C. Xia, S. Li, S. Yang, T. Xi, G. Zhang, Open-transmind: A new baseline and benchmark for 1st foundation model challenge of intelligent transportation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 6328–6335.

[32] A.A. Brincat, F. Pacifici, S. Martinaglia, F. Mazzola, The internet of things for intelligent transportation systems in real smart cities scenarios, in: 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), IEEE, 2019, pp. 128–132.

[33] X. Han, Z. Meng, X. Xia, X. Liao, B.Y. He, Z. Zheng, Y. Wang, H. Xiang, Z. Zhou, L. Gao, et al., Foundation intelligence for smart infrastructure services in transportation 5.0, IEEE Trans. Intell. Veh. 9 (1) (2024) 39–47.

[34] M.A. Amin, W. Ahmad, W.H. Bangyal, I.U. Haq, et al., Trajectory data mining and trip travel time prediction on specific roads, in: 2024 International Conference on Engineering & Computing Technologies, ICECT, IEEE, 2024, pp. 1–8.

[35] H. Mora, V. Gilart-Iglesias, R. Pérez-del Hoyo, M.D. Andújar-Montoya, A comprehensive system for monitoring urban accessibility in smart cities, Sensors 17 (8) (2017) 1834.

[36] J.S. de Oliveira Neto, Inclusive Smart Cities: Theory and Tools to Improve the Experience of People with Disabilities in Urban Spaces (Ph.D. thesis), Université Paris Saclay (COmUE); Universidade de São Paulo (Brésil), 2018.

[37] R.W. Attar, M. Habes, A. Almusharraf, A.H. Alhazmi, R.W. Attar, Exploring the impact of smart cities on improving the quality of life for people with disabilities in Saudi Arabia, Front. Built Environ. 10 (2024) 1398425.

[38] T. Makkonen, T. Inkinen, Inclusive smart cities? Technology-driven urban development and disabilities, Cities 154 (2024) 105334.

[39] S. Zhou, E.T. Loiacono, N. Kordzadeh, Smart cities for people with disabilities: a systematic literature review and future research directions, Eur. J. Inf. Syst. 33 (6) (2024) 845–862.

[40] S.R. Pearson, M. Shafik, Smart cities enhancing the lives of people with disabilities, in: MATEC Web of Conferences, vol. 401, EDP Sciences, 2024, p. 07004.

[41] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 580–587.

[42] R. Girshick, Fast R-CNN, in: 2015 IEEE International Confer-ence on Computer Vision, ICCV, Santiago, 2015, pp. 7–13.

[43] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, IEEE Trans. Pattern Anal. Mach. Intell. 39 (6) (2016) 1137–1149.

[44] M. Khalid, A. Ashraf, W.H. Bangyal, M. Iqbal, An android application for unwanted vehicle detection and counting, in: 2023 International Conference on Human-Centered Cognitive Systems, HCCS, IEEE, 2023, pp. 1–6.

[45] M.L. Ali, Z. Zhang, The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection, Computers 13 (12) (2024) 336.

[46] N. Li, M. Wang, G. Yang, B. Li, B. Yuan, S. Xu, DENS-YOLOv6: A small object detection model for garbage detection on water surface, Multimedia Tools Appl. 83 (18) (2024) 55751–55771.

[47] F.B.K. ArdaÇ, P. Erdogmus, Car object detection: Comparative analysis of YOLOv9 and YOLOv10 models, in: 2024 Innovations in Intelligent Systems and Applications Conference, ASYU, IEEE, 2024, pp. 1–6.

[48] C.-Y. Wang, I.-H. Yeh, H.-Y. Mark Liao, Yolov9: Learning what you want to learn using programmable gradient information, in: European Conference on Computer Vision, Springer, 2024, pp. 1–21.

[49] Ultralytics, YOLO-NAS, 2023, https://docs.ultralytics.com/models/yolo-nas/, Available online: https://docs.ultralytics.com/models/yolo-nas/#which-tasks-and-modes-are-supported-by-yolo-nas-models.

[50] A. Wang, H. Chen, L. Liu, K. Chen, Z. Lin, J. Han, G. Ding, Yolov10: Real-time end-to-end object detection, 2024, arXiv preprint arXiv:2405.14458.

[51] C.-Y. Huang, C.-K. Wu, P.-Y. Liu, Assistive technology in smart cities: A case of street crossing for the visually-impaired, Technol. Soc. 68 (2022) 101805.

[52] A. Montanha, A.M. Oprescu, M. Romero-Ternero, A context-aware artificial intelligence-based system to support street crossings for pedestrians with visual impairments, Appl. Artif. Intell. 36 (1) (2022) 2062818.

[53] P. Theodorou, A. Meliones, C. Filios, Smart traffic lights for people with visual impairments: A literature overview and a proposed implementation, Br. J. Vis. Impair. 41 (4) (2023) 697–725.

[54] S. Dhou, A. Alnabulsi, A.-R. Al-Ali, M. Arshi, F. Darwish, S. Almaazmi, R. Alameeri, An IoT machine learning-based mobile sensors unit for visually impaired people, Sensors 22 (14) (2022) 5202.

[55] A.R.G. Ramirez, I. González-Carrasco, G.H. Jasper, A.L. Lopez, J.L. Lopez-Cuadrado, A. García-Crespo, Towards human smart cities: internet of things for sensory impaired individuals, Computing 99 (2017) 107–126.

[56] L. Rodrigo-Salazar, I. González-Carrasco, A.R. Garcia-Ramirez, An IoT-based contribution to improve mobility of the visually impaired in smart cities, Computing 103 (6) (2021) 1233–1254.

[57] A. Celesti, A. Lay-Ekuakille, J. Wan, M. Fazio, F. Celesti, A. Romano, P. Bramanti, M. Villari, Information management in IoT cloud-based tele-rehabilitation as a service for smart cities: Comparison of NoSQL approaches, Measurement 151 (2020) 107218.

[58] A. Javed, L. Fahad, A. Farhan, S. Abbas, G. Srivastava, R. Parizi, M. Khan, Automated cognitive health assessment in smart homes using machine learning, Sustain. Cities Soc. 65 (2021) 102572.

[59] Q.H. Lai, C.S. Lai, Healthcare with wireless communication and smart meters, IEEE Consum. Electron. Mag. 12 (4) (2023) 53–62, http://dx.doi.org/10.1109/MCE.2022.3181438.

[60] C. Prandi, S. Mirri, S. Ferretti, P. Salomoni, On the need of trustworthy sensing and crowdsourcing for urban accessibility in smart city, ACM Trans. Internet Technol. (TOIT) 18 (1) (2017) 1–21.

[61] M. Sánchez-Ávila, M.A. Mouriño-García, J.A. Fisteus, L. Sánchez-Fernández, Detection of barriers to mobility in the smart city using Twitter, IEEE Access 8 (2020) 168429–168438.

[62] B. Vela, J.M. Cavero, P. Cáceres, C.E. Cuesta, A semi-automatic data–scraping method for the public transport domain, IEEE Access 7 (2019) 105627–105637, http://dx.doi.org/10.1109/ACCESS.2019.2932197.

[63] M. Alruwaili, M.N. Atta, M.H. Siddiqi, A. Khan, A. Khan, Y. Alhwaiti, S. Alanazi, Deep learning-based YOLO models for the detection of people with disabilities, IEEE Access (2023).

[64] S. Ali, S. Al Mamun, H. Fukuda, A. Lam, Y. Kobayashi, Y. Kuno, Smart robotic wheelchair for bus boarding using CNN combined with hough transforms, in: Intelligent Computing Methodologies: 14th International Conference, ICIC 2018, Wuhan, China, August 15-18, 2018, Proceedings, Part III 14, Springer, 2018, pp. 163–172.

[65] M. Alruwaili, M.H. Siddiqi, M.N. Atta, M. Arif, Deep learning and ubiquitous systems for disabled people detection using YOLO models, Comput. Hum. Behav. 154 (2024) 108150.

[66] L. Lecrosnier, R. Khemmar, N. Ragot, B. Decoux, R. Rossi, N. Kefi, J.-Y. Ertaud, Deep learning-based object detection, localisation and tracking for smart wheelchair healthcare mobility, Int. J. Environ. Res. Public Heal. 18 (1) (2021) 91.

[67] Roboflow, Mobility aids dataset, 2024, Roboflow Universe, Available online: https://universe.roboflow.com/wheelchair-gohwp/mobility-aids-jcnno-1cyew-zn862/dataset/1.

[68] Roboflow, MTP wheelchair dataset, 2024, Roboflow Universe, Available online: https://universe.roboflow.com/cup-dataset-ffunj/mtpwheelchair.

[69] M. Alruwaili, M.H. Siddiqi, M.N. Atta, M. Arif, Deep learning and ubiquitous systems for disabled people detection using YOLO models, Comput. Hum. Behav. 154 (2024) 108150.

[70] H. Aman, Traffic prediction dataset, 2023, Kaggle Dataset, Available online: https://www.kaggle.com/datasets/hasibullahaman/traffic-prediction-dataset/data.

[71] A.S. Kailash, B. Sneha, M. Authiselvi, M. Dhiviya, R. Karthika, E. Prabhu, Deep learning based detection of mobility aids using yolov5, in: 2023 3rd International Conference on Artificial Intelligence and Signal Processing, AISP, IEEE, 2023, pp. 1–4.

[72] M. Alruwaili, M.N. Atta, M.H. Siddiqi, A. Khan, A. Khan, Y. Alhwaiti, S. Alanazi, Deep learning-based YOLO models for the detection of people with disabilities, IEEE Access (2023).

[73] A.S. Kailash, B. Sneha, M. Authiselvi, M. Dhiviya, R. Karthika, E. Prabhu, Deep learning based detection of mobility aids using yolov5, in: 2023 3rd International Conference on Artificial Intelligence and Signal Processing, AISP, IEEE, 2023, pp. 1–4.