Article

Application of Sustainable Blockchain Technology in the Internet of Vehicles: Innovation in Traffic Sign Detection Systems

Yanli Liu 1, Qiang Qian 1, Heng Zhang 1,*, Jingchao Li 1, Yikai Zhong 1 and Neal N. Xiong 2,3

1 School of Electronic Information, Shanghai Dianji University, Shanghai 201306, China; liuyl@sdju.edu.cn (Y.L.); lijc@sdju.edu.cn (J.L.)
2 Department of Mathematics and Computer Science, Sul Ross State University, Alpine, TX 79830, USA; xiongnaixue@gmail.com
3 School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411201, China
* Correspondence: zhangheng@sdju.edu.cn

Abstract: With the rapid development of the Internet of Vehicles (IoV), traffic sign detection plays an indispensable role in advancing autonomous driving and intelligent transportation. However, current road traffic sign detection technologies face challenges in terms of information privacy protection, model accuracy verification, and result sharing. To enhance system sustainability, this paper introduces blockchain technology. The decentralized, tamper-proof, and consensus-based features of blockchain ensure data privacy and security among vehicles while facilitating trustworthy validation of traffic sign detection algorithms and result sharing. Storing model training data on distributed nodes reduces the system computational resources, thereby lowering energy consumption and improving system stability, enhancing the sustainability of the model. This paper introduces an enhanced GGS-YOLO model, optimized based on YOLOv5. The model strengthens the feature extraction capability of the original network by introducing a coordinate attention mechanism and incorporates a BiFPN feature fusion network to enhance detection accuracy. Additionally, the newly designed GGS convolutional module not only improves accuracy but also makes the model more lightweight. The model achieves an enhanced detection accuracy rate of 85.6%, with a reduced parameter count of $0.34 \times 10^7$. In a bid to broaden its application scope, we integrate the model with blockchain technology for traffic sign detection in the IoV. This method demonstrates outstanding performance in traffic sign detection tasks within the IoV, confirming its feasibility and sustainability in practical applications.

Keywords: Internet of Vehicles; traffic sign detection; sustainability; blockchain technology; GGS-YOLO; model accuracy verification

1. Introduction

Blockchain is a decentralized ledger technology employed to securely record and share all network transactions, bolstering data integrity [1]. Each node maintains the same ledger, ensuring that all operations are visible to all nodes, providing a secure data-sharing mechanism. However, this technology requires high computing costs and power consumption. According to relevant data, Bitcoin consumed 0.55% of global electricity in 2020 [2], while the monthly storage cost of Ethereum blockchain is approximately USD 20 million [3]. Therefore, ensuring the sustainability of blockchain is crucial for the long-term operation of such systems. In the domain of autonomous driving, the timely detection of traffic signs is of utmost importance. It helps vehicles comply with road rules, improves traffic safety, and reduces the risk of accidents. At the same time, autonomous driving systems must possess the capability to promptly identify and interpret road signs for...
precise decision making and route planning. The real-time monitoring and analysis of traffic flow, coupled with measures to optimize vehicle movement and reduce congestion and greenhouse gas emissions, are crucial for the advancement of autonomous driving technology. Accurate detection of traffic signs is a key foundation for achieving this goal. Additionally, the detection and analysis of traffic signs contribute to acquiring road traffic information, supporting traffic management, and monitoring and optimizing urban traffic congestion. The distributed storage provided by blockchain also contributes to enhancing the sustainability of such systems. In this context, achieving sustainability involves optimizing the energy efficiency of blockchain in the IoV and considering the overall environmental impact. This ensures that the benefits of this technology align with long-term ecological considerations.

In industrial settings, the traffic sign detection model designed in this study exhibits various potential advantages. Firstly, with real-time monitoring in traffic management, we can effectively ensure driver compliance with regulations, significantly reducing the potential risk of traffic accidents. This enhances overall safety in industrial facilities, minimizing the risks of personnel injuries and equipment damage. In the realm of industrial park surveillance, the application of this model enables efficient detection of signs within industrial areas, ensuring the safe movement of equipment and transport vehicles in complex environments. This positively impacts the prevention of accidents among pieces of equipment, optimizes transport routes, and enhances overall industrial productivity. Through the safety monitoring of vehicles and equipment, we can preemptively prevent potential workplace accidents, thereby reducing costs associated with maintenance and replacement and increasing the reliability and lifespan of equipment. Regarding process efficiency, the application of an intelligent parking system allows industrial facilities to utilize parking space more effectively, mitigating congestion and providing real-time parking information. This not only improves efficiency in the management of vehicles and equipment but also contributes to reducing wait time in operations, comprehensively enhancing overall workflow efficiency.

The blockchain-based real-time traffic sign detection described in this article faces multiple challenges, and achieving sustainability further adds to the complexity. Factors such as weather conditions, time, and changes in lighting can impact the appearance of traffic signs, thereby increasing the difficulty of detection. These challenges necessitate detection algorithms with high adaptability, capable of reliably identifying traffic signs under various environmental conditions. Additionally, vehicles, pedestrians, or other obstacles on the road may obstruct the signs or even render them completely invisible, posing challenges to the accuracy and robustness of detection algorithms. High energy consumption is a significant concern in blockchain technology. To mitigate the energy demand, optimizing data storage and transmission processes can reduce overall computational costs and energy consumption, thereby enhancing the sustainability of such systems. Overall, achieving sustainability in blockchain-based real-time traffic sign detection requires the comprehensive consideration of algorithm flexibility, perception capabilities, and energy efficiency optimization in blockchain technology. This ensures that the system operates robustly in different environments while keeping resource consumption within acceptable limits. The contributions of this paper are as follows:

- To achieve the effective integration of blockchain and object detection models, we propose an innovative architecture designed to optimize data storage and transmission processes, ensuring sustainable resource utilization throughout the system’s long-term operation. This contributes to reducing the overall energy consumption of the system, aligning it more closely with the principles of environmentally sustainable development;
- The YOLOv5 model is enhanced with the incorporation of a coordinate attention mechanism to improve its feature extraction capabilities. Additionally, the integration of BiFPN further bolsters the model’s effectiveness in fusing multi-scale features;
• We introduce the GGS module in conjunction with GSconv to replace the original CBS module, resulting in a more lightweight model.

The rest of this paper is organized as follows: The second section reviews relevant literature and existing work, providing a theoretical foundation for the study. The third section details the overall structural framework of this study and the content of the GGS-YOLO model. The fourth section presents the experimental results and analyzes the performance of the improved model. Finally, the fifth section summarizes the main research of this paper and points out the shortcomings of the study.

2. Related Work

In this chapter, we divide the literature into two main research directions. The first type of research focuses on the application and exploration of object detection technology in the field of traffic sign detection. The second type of research focuses on blockchain research in data exchange and data protection in the IoV.

2.1. Overview of Traffic Sign Detection Methods

Currently, traffic sign detection methods can be primarily categorized into two groups: traditional methods and deep learning-based methods. Traditional methods typically rely on features like color and shape for detection. Liu and his team classified traditional methods, including various methods such as color and shape [4]. Yang and his team proposed a probabilistic model based on the Ohta color space to detect traffic signs by drawing color probability maps [5]. Zaklouta et al. [6] used the Hough transform to detect the shape of signs. While shape detection methods are resilient to the effects of natural lighting conditions, they exhibit limited capability in addressing issues such as sign aging and deformation, making them prone to false positives and false negatives. Although traditional methods perform reasonably well in daily detection, their performance in complex environments is not ideal.

In the realm of deep learning methods, convolutional neural networks (CNNs) are prevalent technologies extensively employed in computer vision. These methods are typically categorized into one-stage and two-stage detection approaches. In two-stage detection methods, candidate regions are first generated and then classified and positioned. Girshick et al. [7] proposed R-CNN, which is considered the pioneering work in the field of two-stage detection methods. This method first extracts candidate regions and then uses CNN and SVM classifiers for classification. Although this method achieves good results in classifying target areas, it wastes a lot of storage and computing resources due to the repeated selection of candidate boxes. On this basis, Zhang et al. [8] proposed a cascade R-CNN, which obtains multi-scale information using feature pyramids and uses product-weighted operations to refine traffic sign features, thus enhancing detection accuracy and robustness. Wang et al. [9] improved Faster R-CNN using the Res2net network backbone and proposed a resampling post-processing scheme, resulting in improved detection performance for small targets. In addition, Redmon et al. [10] proposed the YOLO target detection algorithm, which eliminates the need to generate candidate regions and instead directly performs classification and localization, leading to higher detection speed. The SSD algorithm is also one of the one-stage methods [11]. Wang J. et al. [12] presented an enhanced approach built upon the YOLOv5s network. This method incorporates the AF-FPN feature pyramid network and an automatic learning data enhancement strategy to improve multi-scale recognition and model robustness. Despite improvements, there are still issues with model complexity and insufficient detection accuracy. Chen et al. [13] proposed a strategy of adding receptive field cross modules and attention mechanisms to the backbone network for YOLOv4 and YOLOv5 to improve the detection of occluded targets. However, this leads to a reduction in frame rate, impacting real-time performance.

In summary, one-stage methods have notably enhanced detection speed compared with two-stage methods, making them suitable for a wide range of applications. In addition, in the field of traffic sign recognition, researchers continue to explore new methods and
technologies. For example, Zhang et al. [14] considered the unique characteristics of traffic signs and put forward a traffic sign detection algorithm that relies on deep convolutional networks. To enhance detection speed and accuracy, they optimized the network by either adding or removing convolutional layers in suitable regions. After data set verification, it was proved that this improved detection algorithm is more effective. Lin et al. [15] proposed a traffic sign detection method using a lightweight multi-scale feature fusion network, which speeds up object detection by simplifying the model. Additionally, it enhances generalization and recognition accuracy through multi-scale feature fusion. Xiong et al. [16] introduced a traffic sign detection approach built on a CNN, employing a region proposal network (RPN) to identify traffic signs, resulting in an enhancement in the model’s detection accuracy. However, the model size of this method is larger and may be affected by resource limitations. Khan et al. [17] focused on solving the problem of darker areas in complex environments and proposed an intelligent traffic sign recognition system with a lighting pre-processing function to enhance light in the darker areas of the image using brightness enhancement technology. Their approach yielded satisfactory results when tested on the GTSDDB dataset. Jaramillo-Alcazar et al. [18] proposed a traffic optimization system that combines blockchain and computer vision technologies to improve urban transportation efficiency, safety, and sustainability. Simulating urban environments and applying these technologies demonstrated the system’s effectiveness in reducing congestion and enhancing traffic flow.

2.2. Blockchain Technology Application Research in the IoV

In today’s era of data expansion, data privacy protection has become particularly important. In terms of data storage, some researchers have proposed innovative methods. Wang et al. [19] introduced a decentralized storage system and a data-sharing framework based on blockchain and fine-grained access control. However, these methods have not yet fully addressed the issue of user privacy protection. Although blockchain has advantages in data protection, traditional blockchain models may face security issues and risks of private data leakage during public processing. To address the issue of safeguarding data privacy, Zhang et al. [20] proposed an identity-based scheme that aims to enhance data privacy in non-transaction scenarios and implement privacy-preserving encryption methods. In addition, Xu et al. [21] suggested employing homomorphic encryption technology to enhance the data structure of blocks, thereby achieving the encryption of block data. These methods represent continuous exploration and innovation efforts in data privacy protection. Research in this field continues to advance, ranging from on-chain data hashing to off-chain-encrypted on-chain storage.

Ensuring the integrity of data sharing and preventing data leakage are the focuses of current research. In terms of data sharing, researchers combine different blockchain types to achieve data sharing. Liang et al. [22] proposed an industrial IoT transmission technology based on fabric blockchain, using a blockchain-based dynamic secret sharing mechanism and optimizing data storage and transmission, which improves the security and safety of industrial IoT data transmission, as well as reliability, while improving transmission efficiency. In data dissemination, blockchain usually involves multiple parties, so data privacy is of particular concern. Si et al. [23] proposed a blockchain-based lightweight information-sharing security framework for the Internet of Things which is based on blockchain technology. They utilized and enhanced partial blind signature algorithms to secure and control data. Cai et al. [24] proposed an information security guarantee mechanism based on blockchain technology, which can search for trusted private keywords. Using a searchable encryption method, encrypted data can be searched for effectively and safely. For certain wireless frequency interference attacks [25], Shen et al. [26] put forward a support vector machine training model that effectively safeguards the privacy of IoT data through the application of Paillier encryption technology. Magsi et al. [27] proposed a system for detecting and preventing content-poisoning attacks, combining a threshold-based content-caching mechanism with blockchain technology to address content-poisoning.
attacks (CPAs) in Vehicular Named Data Networks. To prevent data leakage, Gao and his team introduced novel technologies aimed at safeguarding sensitive information [28]. They leveraged kleptography algorithms to modify the blockchain signature algorithm and transaction filtering mechanism, thereby establishing a secure communication channel within a public blockchain system and enhancing data transmission concealment. These studies exemplify ongoing innovation in the realm of data privacy protection, ensuring the security and integrity of data.

The Internet of Things (IoT) [29] and the IoV are closely interconnected. The IoV can be viewed as a specific application of the IoT in the transportation sector. It integrates Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication, utilizing technologies such as in-vehicle sensors, V2V communication, and V2I communication to establish an intelligent transportation system. This system enhances traffic management efficiency and strengthens functions such as vehicle safety, intelligent traffic management, remote monitoring, and management. However, the widespread adoption of in-vehicle sensors has led to a decrease in data communication efficiency [30]. To address this issue, Xu et al. [31] proposed the Adaptive Computation Offloading Method (ACOM), designed specifically for 5G-connected vehicle networks. This method aims to optimize task offloading latency and resource utilization in edge computing systems. It generates viable solutions with a multi-objective evolutionary algorithm and obtains the optimal offloading scheme through utility evaluation. In terms of security and privacy protection in vehicular networks, Ren discussed the applications of the IoT and blockchain technology in intelligent transportation systems [32]. These include monitoring traffic, locating emergencies, managing public transportation, and implementing credit token mechanisms to record system changes. Prashar [33] proposed a model combining the IoT and hash graph technology to improve road safety by creating an effective communication network among vehicles. Ghimire [34] introduced a blockchain-empowered federated learning framework, providing a secure, privacy-preserving, and verifiable solution for collaborative learning in vehicular networks to enhance the security of intelligent transportation systems. Chen et al. [35] addressed security, privacy, and efficiency issues in data transactions within vehicular networks, proposing a data transaction framework and an iterative bidirectional auction mechanism based on a consortium blockchain. In terms of data sharing, Joshi introduced a data transfer architecture using blockchain within a cluster-based Vehicular Ad Hoc Network (VANET) [36]. It employs a raindrop optimization algorithm for clustering to protect the privacy and security of vehicular self-organizing networks. Kang et al. [37] proposed the development of a secure point-to-point data-sharing system for vehicle data using consortium blockchain technology to ensure secure, efficient storage and the sharing of data in the edge network of the vehicular IoT.

The combination of blockchain and the IoV has also been extensively studied. Kang et al. [37] applied alliance blockchain and smart contracts to Road-Side Units (RSUs) for data storage and sharing and used pseudonyms for anonymous operations to provide user privacy protection. Alfadhli et al. [38] proposed a lightweight multifactor authentication scheme that preserves privacy and combines PUF and multifactor authentication to support V2V interaction without sharing sensitive privacy data with RSUs. In the dynamic vehicle scenario, Chai et al. [39] introduced a hierarchical weighted update federated learning [40] algorithm to facilitate knowledge sharing. Various chains are tasked with recording diverse environmental data to enable the federated learning process, thereby facilitating knowledge sharing. Xiao et al. [41] introduced a combined cluster and blockchain scheme for the secure transmission of real-time information. They devised a distributed consensus mechanism built on the Byzantine fault-tolerance algorithm to guarantee the security of vehicle information communication. Cui et al. [42] presented a secure V2V data-sharing scheme that relies on consortium blockchain technology. In this scheme, vehicles generate ratings for data providers, and trusted agencies periodically broadcast blacklists to enhance data security. These studies reflect an ongoing exploration of the integration of blockchain and the IoV, with the aim of enhancing the efficiency and security of data sharing and
privacy protection. Lu et al. [43] proposed an innovative hybrid blockchain architecture that incorporates an asynchronous federated learning approach. The aim is to enhance the security, reliability, and efficiency of data sharing in the vehicular networking domain by employing deep reinforcement learning for node selection. On the other hand, Tan and their team [44] introduced a vehicle platform based on digital twins. This platform utilizes a consortium blockchain and a proof-of-stake consensus algorithm, along with an inventive incentive mechanism, to facilitate resource sharing among vehicles within urban areas. This approach contributes to the overall utility of vehicular networking. The integration of these two approaches holds the potential to provide a comprehensive, secure, and efficient solution for the field of vehicular networking.

3. Method

This section emphasizes the innovative application of the GGS-YOLO model in the IoV. The model achieves precise detection of traffic signs and ensures the real-time recording of relevant data on the blockchain, introducing unprecedented innovation not only in achieving efficient traffic management but also in enhancing data security and tamper resistance. This unique combination makes traffic sign detection data, secured by the blockchain, highly efficient and trustworthy evidence in court proceedings when traffic accidents occur. Through an in-depth exploration of both blockchain architecture and the GGS-YOLO model, we gain a more comprehensive understanding of this innovative approach, providing new insights and a solid foundation for the reliable development of intelligent urban traffic systems.

3.1. System Architecture

In terms of ensuring the sustainability of a system where the outputs of a target detection model are stored in a distributed blockchain network, we place particular emphasis on several key aspects. Firstly, by adopting smart contracts, we not only ensure the transparency and tamper resistance of the target detection results but also establish a reliable traceability and verification mechanism, promoting a high level of data integrity throughout the long-term operation of the system. Secondly, with the introduction of decentralized storage mechanisms, we not only disperse the storage of large-scale data generated by the model across different nodes in the blockchain network, thereby enhancing the system’s scalability, but also reduce the risk associated with a single node. This enhances the overall stability and reliability of the system, making it more sustainable over time. In terms of data transmission, by leveraging the distributed nature of blockchain, we decompose the transmission process into multiple steps. With smart contract verification, we not only reduce the complexity of data transmission and improve its security but also alleviate the burden on a single node, thereby further enhancing the efficiency of the entire system. The comprehensive improvements and integration of these strategies contribute to ensuring that the system maintains a high level of sustainability throughout its long-term operation.

The system’s overall architecture is depicted in Figure 1, encompassing three key components: data providers, data receivers, and blockchain. Notably, data providers and data receivers share equal roles, serving as both providers and collectors.

1. The term data providers mainly refers to vehicles with data collection functions. Vehicles identify the corresponding traffic sign data using YOLOv5 and then encrypt the collected data and upload them to the blockchain network;
2. Blockchain network refers to a public blockchain platform where data providers or data receivers can conduct data transactions after registering their information. The main part of the blockchain network, the smart contract, consists of four parts: the subject node, the object node, the credit node, and the exchange node. The role of the subject and object nodes is to store the identity information of the provider and receiver and to verify the correctness of the signatures of both parties in the data exchange. The role of the credit node is to evaluate the credit of both parties in the
transaction and determine whether the two parties can conduct data exchange. The exchange node is responsible for access requests of the entire control system;

3. The data receiver is the user of the data, responsible for decrypting the received encrypted data and utilizing the provider’s data to perform real-time processing on traffic signal signs.

Figure 1. System architecture.

In the design and implementation of blockchain systems, employing formal methods is a crucial strategy to ensure their high reliability and security [45]. Given the decentralized nature of blockchain and its applications in critical domains, any security vulnerabilities in the system could lead to severe consequences. Formal methods strengthen the overall security architecture of blockchain systems by providing strict adherence to system specifications and the ability to defend the system against various security threats. Additionally, this approach enhances system transparency and credibility, offering users and participants additional security assurances.

Within this framework, the formal verification of smart contracts becomes particularly essential [46]. Since smart contracts are immutable once deployed and often handle a significant volume of transactions and assets, ensuring their accuracy and intended functionality before deployment is paramount. Through the application of mathematical proof methods, smart contract code can undergo rigorous verification, ensuring the correctness and security of its logic. This formal verification process adds a layer of reliability and security to smart contracts, helping prevent potential vulnerabilities and attacks, thereby maintaining the stable operation of the entire blockchain system.

3.2. Blockchain Network Construction
3.2.1. Node Data Collection

Figure 2 shows the process of node A collecting the “Stop” flag and node D collecting the “No whistle sign” flag. The details of these two data collection processes are updated immediately in the blockchain’s public ledger. The public ledger records all the collected information, regardless of whether a node participates in these information collection processes. This information is visible to all subsequent blockchains. Due to the immutability of the blockchain, the transmission of such information is more secure.
By creating a blockchain containing multiple car nodes, an efficient approach can be established to enable these cars to rapidly perform traffic sign recognition, particularly when one of the cars has been trained using the YOLOv5 algorithm. Assuming 100 cars have not undergone any training, leveraging blockchain technology allows these cars to connect to a shared public ledger, eliminating the cumbersome task of individually training each car. Through blockchain-enabled collective learning, each car can leverage the experience of others, facilitating a form of collaborative learning. This not only enables all cars to learn the positions of traffic signs from the experience of a single car but also avoids redundant efforts, thereby enhancing overall system efficiency. This collective learning approach not only enables cars to adapt to traffic environments more quickly but also facilitates intelligent recognition and adherence to traffic signs within a short time frame. Furthermore, by establishing a distributed learning model on the blockchain, computational resources can be shared across the entire network, achieving sustainability in terms of saving computational resources. This shared computational resource model helps avoid the need for each car to independently perform extensive computations, reducing the overall system’s demand for computational resources and aligning with the principles of sustainable development. Through the application of blockchain technology, an efficient, collaborative, and sustainable system for traffic sign recognition can be realized, providing an innovative solution for the future development of autonomous and self-learning vehicles. Figure 3 illustrates car A updating the “stop” sign in the ledger, exemplifying how shared blockchain data can be used for mutual improvement and knowledge sharing among vehicles in the network.
Figure 4 shows that when any car ledger is updated, the same update is also reflected in other ledgers of the blockchain, which also relatively reflects the shareability of the blockchain. Therefore, as long as one car node receives traffic sign information, any number of car nodes can be updated. This approach can save significant computational resources and distribute acquired data across various nodes to enhance data availability and performance. This is related to sustainability because it ensures that data remain accessible in the face of failures, disasters, or other challenges. With data redundancy, backups, load balancing, and compliance management, this distributed storage system can improve data sustained availability and security, contributing to the principles of sustainability.

3.2.2. Data Sharing

Figure 5 shows the complete process of data sharing. In the access permission request stage, the provider first sends a request to the exchange node, which retrieves the information based on the information of the subject and object nodes, and then returns the result of the request. If the request is successful, the information is passed to the credit node to conduct credit evaluation on both parties, and the result of the request are returned after waiting for the result of the assessment result. Once access permission is approved, the provider can exchange data with the receiver, and all data exchange messages are encrypted using the public keys of both parties to ensure message confidentiality. All interactions are implemented using smart contracts, and nodes interact with contracts and other nodes by sending transactions. All access history and results are stored on the blockchain, and even if some nodes are damaged or compromised, the data-sharing system can still operate reliably.

In this study, the application of smart contracts significantly enhances the efficiency and security of sharing traffic sign detection data in the connected vehicle environment. For instance, when a vehicle detects a newly placed parking sign with its advanced sensor system, this information is instantly recorded on and uploaded to the blockchain, facilitated by the automated execution of smart contracts. Once the traffic sign data are verified and added to the blockchain, they are immediately accessible to all authorized vehicles within the network. This mechanism not only accelerates data transmission but also substantially improves reliability and security due to the immutability of the blockchain, contrasting with traditional centralized data-sharing mechanisms that rely on single servers or databases, posing risks of data transfer delays, security vulnerabilities, and even data loss. Importantly, smart contracts provide a transparent and traceable data-sharing environment. Every transaction, including the upload and download of traffic sign data, is recorded
on the blockchain. This not only promotes data transparency but also facilitates future audits and issue resolution. For example, if an error in traffic sign detection leads to an accident, the relevant data can be traced to determine the root cause. Additionally, compared with traditional data-sharing methods, smart contracts significantly reduce operational and maintenance costs. Automated data validation and transmission reduce the need for manual intervention while lowering the risks associated with data processing errors. Therefore, by leveraging smart contracts to achieve the efficient sharing of traffic sign detection data in connected vehicles, our approach not only enhances the speed and security of data processing but also elevates the overall reliability and transparency of the entire traffic system.

3.2.3. Distributed Storage: Robustness and Security in Connected Vehicles

Distributed storage is a method for dispersing data across multiple nodes in a network. In contrast to traditional centralized storage, its goal is to enhance the security, reliability, and censorship resistance of data. This decentralized storage employs principles of decentralized control, utilizing blockchain technology to ensure data immutability and transparency. Simultaneously, encryption techniques are applied to safeguard data confidentiality. Through data redundancy, replication, and incentive mechanisms, distributed storage systems enhance the overall robustness of the system, effectively mitigating the risks of single points of failure.

In the IoV, distributed storage demonstrates significant advantages, particularly in terms of scalability and risk reduction. Regarding scalability, the system can easily handle scenarios such as urban expansion or an increase in the number of vehicles. The use of distributed storage enables the system to adapt to such growth simply by adding more stor-
age nodes, eliminating the need for costly infrastructure upgrades. This design enhances system flexibility and scalability, allowing it to adapt to evolving data storage requirements.

In terms of risk reduction, distributed storage systems mitigate the risk of system downtime caused by single points of failure through data replication and decentralized storage. Even if a node experiences a technical failure, other nodes can continue to provide services, ensuring the continuous availability of data. Additionally, the dispersed storage of data makes the system more resistant to tampering or centralized network attacks, thereby enhancing overall data security. This security enhancement is particularly crucial in applications such as intelligent transportation systems, which involve sensitive traffic information.

3.3. Traffic Sign Detection Model

The application of real-time traffic sign detection in the IoV and intelligent transportation systems not only enhances the intelligence and efficiency of traffic management but also positively impacts the sustainable development of cities, laying the foundation for further innovation in future intelligent transportation systems. The accuracy and speed of traffic sign detection are crucial components of the entire system.

3.3.1. YOLOv5 Model

YOLOv5 is indeed a target detection model that was open-sourced by Ultralytics in May 2020. It is the first YOLO series algorithm written using the PyTorch framework. This model is better suited for deployment in production environments, offers ease of deployment, and boasts fewer parameters. The YOLOv5 model is categorized into various versions, denoted as YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x based on the network’s depth and width, with each variant being tailored to specific computational requirements and performance levels. Indeed, in most cases, as the computational complexity of a model increases, its detection accuracy tends to gradually improve. This paper is a modified model based on the latest 7.0 version of YOLOv5. YOLOv5s consists of an input terminal (input), backbone network (backbone), neck network (neck), and prediction network (head). Conv is the most basic part of the network, consisting of Conv2d, BatchNorm2d, and SiLU activation functions. The data are initially augmented using Mosaic data in the input stage, which serves to enhance data diversity and enhance training efficiency. The backbone network uses a simplified version of BottleneckCSP, the C3 module, which greatly reduces the number of parameters. The neck network uses the PANet network and adopts the FPN+PAN structure [47]. In contrast, the FPN structure transfers deep semantic features to shallower layers, thereby augmenting semantic information at multiple scales. The PAN structure transfers shallow positioning information to deeper layers, thereby improving positioning capabilities at multiple scales. The neck network also incorporates a fast spatial pyramid pooling (SPPF) module to facilitate the fusion of local and global features. Finally, the resulting adaptive size output is passed to the neck network. In YOLOv5, there is a challenge in balancing accuracy and inference speed, as improving model accuracy may increase the computational burden, thus negatively impacting performance. Particularly, ensuring precise detection and localization of small objects in images, especially in scenarios with occlusions or dense target presence, poses a challenging problem. Ensuring real-time performance is crucial for key applications such as autonomous driving and video surveillance, making it an urgent challenge to ensure that the model can perform rapid and efficient inference in real-time scenarios. To address these issues, researchers often refine the model architecture to strike a balance between accuracy and inference speed, involving the optimization of network structures and the introduction of attention mechanisms, among other improvements [48]. Additionally, the adoption of advanced data augmentation techniques and sample generation methods, along with optimized post-processing techniques, holds the promise of enhancing the model’s capability in detecting small targets and improving robustness, especially when facing challenges like sample imbalance and annotation issues. Strengthening research in domain adaptation
and transfer learning is also crucial to ensuring model versatility, enabling adaptation to diverse datasets and environments, and ultimately enhancing performance across various application scenarios [49]. The network structure of YOLOv5 is shown in Figure 6.

Figure 6. YOLOv5 network structure.

3.3.2. Coordinate Attention Mechanism

The primary purpose of introducing attention mechanisms is to enable a model to focus more selectively on specific parts of the input data, enhancing its ability to capture and process important information. This introduction helps the model to more effectively learn the relationships within the input data, mitigating the impact of irrelevant or redundant information and thereby improving overall model performance. By selectively concentrating on specific regions relevant to the task, attention mechanisms contribute to boosting the model’s accuracy, reducing interference from irrelevant information, and simultaneously enhancing the model’s robustness and adaptability. Additionally, this mechanism aids in addressing scale-related challenges, allowing the model to handle targets of different scales more effectively. Moreover, it alleviates computational burdens to some extent, improving the efficiency of the model’s inference process.

To enhance the YOLOv5 feature extraction network’s ability to accurately capture position signals and concentrate on critical feature regions, an integrated coordinate attention module (CA) is introduced into the existing network architecture [50]. The introduction of CA significantly enhances the model’s performance in multiple aspects. Firstly, this mechanism markedly improves spatial localization accuracy, reducing the risks of false positives and false negatives. This results in more precise and reliable performance, particularly when dealing with challenges such as occlusion, deformation, and complex scenes in tasks like object detection, enhancing the model’s robustness. Secondly, the coordinate attention mechanism contributes to the fine-grained learning of features within input data by allowing deep learning models to discern subtle variations, thereby increasing the model’s sensitivity to minor changes. Lastly, by optimizing the utilization of computational resources, this mechanism improves the overall computational efficiency of the model, providing higher efficacy for practical applications. Overall, the coordinate attention mechanism proves instrumental in advancing spatial localization, bolstering robustness, and enhancing computational efficiency, thereby positively impacting various performance metrics of the model. The CA module uses one-dimensional feature encoding, modeling both channel and spatial directions to retain original spatial information and rich position information. By combining parallel information from both channel and spatial dimensions, the model enhances its focus on important features while reducing background noise interference, generating better quality feature maps and a better positioning of sensitive
areas. The coordinate attention mechanism can be broken down into two fundamental components: coordinate information embedding and coordinate attention generation. These components work together to establish a lasting connection with the target. Figure 7 is a structural diagram of the coordinate attention mechanism.

Figure 7. Coordinate attention mechanism.

The first step is coordinate information embedding. To ensure the accuracy of capturing spatial position information, we decompose the global pooling operations, as shown in Equation (1). Specifically, we split two-dimensional global pooling into two separate one-dimensional encoding processes.

\[ Z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_c(i, j) \]  

Specifically, we extract information separately from both the horizontal and vertical dimensions. We employ a pair of pooling kernels with sizes \((H, 1)\) and \((1, W)\) to perform feature encoding in the horizontal and vertical directions for each channel of the input feature, thus obtaining perceptual features in both directions, as shown in Equations (2) and (3).

\[ Z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} X_c(h, i) \]  

\[ Z_c^w(w) = \frac{1}{H} \sum_{0 \leq j \leq H} X_c(j, w) \]
where the subscript \( c \) represents the channel of the feature vector, \( Z^h_c \) represents the information we obtain in the vertical direction, and \( Z^w_c \) represents the perceptual information we obtain in the horizontal direction. Here, \( X_c \) represents the input of the previous layer network.

The second step involves generating coordinated attention. After gathering information from both the vertical and horizontal directions, the next step is to combine it. Then, a convolutional transformation function, denoted by \( F_1 \), is applied to compress the channel. This process produces an attention graph \( f \) that captures spatial information in both the horizontal and vertical directions, and the symbol \( \delta \) signifies the activation function used in the process, as shown in Equation (4).

\[
f = \delta \left( F_1 \left( \left[ Z^h, Z^w \right] \right) \right)
\]

(4)

Once the attention map is obtained, tensor \( f \) is divided into two separate tensors, \( f^h \) and \( f^w \), along the spatial dimension. After a \( 1 \times 1 \) convolutional transformation and the sigmoid activation function, the final region weights are obtained. The weight in the vertical direction is denoted by \( g^h \), while the weight in the horizontal direction is represented by \( g^w \), as shown in Equations (5) and (6). The adjusted attention map is expressed as \( y_c \), with \( \sigma \) signifying the sigmoid function and \( F^h \) and \( F^w \) indicating \( 1 \times 1 \) convolution operations, as shown in Equation (7).

\[
g^h = \sigma \left( F^h \left( f^h \right) \right)
\]

(5)

\[
g^w = \sigma \left( F^w \left( f^w \right) \right)
\]

(6)

\[
y_c(i, j) = x_c(i, j) \times g^h(i) \times g^w(j)
\]

(7)

In this paper, the coordinate attention mechanism is incorporated into both C3 modules. This enhancement not only reduces the network’s depth, saving a significant amount of computational resources, but also significantly improves efficiency. The new C3-CA module is capable of capturing global feature information comprehensively, thereby utilizing the attention mechanism more effectively.

3.3.3. Loss Function Improvement

IoU loss is a loss function that directly measures the IoU value between the predicted bounding box and the true bounding box and optimizes it as the goal. When the IoU value is low, the gradient of the loss function becomes very small, resulting in a long training time. In YOLOv5, CIou Loss is introduced as the loss function of bounding box regression, and binary cross-entropy loss (BEC Loss) is used for handling classification tasks. In this paper, the Wise-IoU v3 loss function is employed for bounding box regression in the model, replacing CIou [51].

The method adopted by Wise-IoU v3 offsets the negative impact of geometric factors, such as distance and aspect ratio, on the model [52]. In the case of a high overlap between the prediction box and the target box, the loss function is implemented by weakening the penalty of the geometric factors. This allows the model to improve its generalization capability with minimal training intervention. The schematic diagram of the Wise-IoU v3 loss function used in this study is shown in Figure 8.
Figure 8. Wise-IoU v3 schematic diagram.

In the diagram, $W$ and $H$ correspond to the predicted box’s width and height, while $W_{gt}$ and $H_{gt}$ represent the actual box’s width and height. $W_i$ and $H_i$ stand for the width and height of the overlapping area between the predicted box and the actual box. $S_u$ represents the combined area of overlap between the predicted box and the actual box, while $L_{iou}$ is a measure of the loss incurred when gauging the level of overlap between the predicted box and the actual box, as shown in Equations (8) and (9).

$$S_u = WH + W_{gt}H_{gt} - W_iH_i$$  \hspace{1cm} (8)

$$L_{iou} = 1 - \frac{W_iH_i}{S_u}$$  \hspace{1cm} (9)

The $\beta$ outlier and $\lambda$ gradient gain are mainly determined by the hyperparameters $\alpha$ and $\delta$. $\beta$ is utilized to gauge the extent of irregularity in the predicted box. A reduced irregularity corresponds to an anchor box of superior quality. By suitably tuning these hyperparameters, it becomes possible to establish a non-monotonic focus function. This function, in turn, allocates smaller gradient gains to prediction boxes with substantial outliers. This adjustment effectively reduces the harmful impact of gradient on prediction boxes with lower-quality training samples. Combining the above formulas yields the formula for Wise-IoU v3, as shown in Equation (11).

$$\lambda = \frac{\beta}{\delta \alpha^{\beta-\delta}}$$  \hspace{1cm} (10)

$$L_{W_iou-v3} = L_{iou} \exp \left[ \frac{(x_p - x_{gt})^2 + (y_p - y_{gt})^2}{\left(W_{gt}^2 + H_{gt}^2\right)^{\alpha}} \right]^\lambda$$  \hspace{1cm} (11)

3.3.4. BiFPN Feature Fusion Network

Feature pyramid networks (FPNs) use a top-down approach to merge features, as shown in Figure 9a [53]. The objective is to combine low-level positional information...
with high-level semantic information in order to improve the network’s feature extraction capability. Nonetheless, as features are transmitted through the network, there is potential for the loss of more detailed feature information. To address the limitation of feature pyramid networks that can only transmit information in a top-down manner, Liu et al. [47] proposed the Path Aggregation Network (PAN), as shown in Figure 9b. The PAN introduces an information channel that flows in a bottom-up manner within the feature pyramid structure. This approach helps mitigate the loss of feature information and ultimately enhances detection performance. Nevertheless, when employing this structure for small object detection, the rate of feature reuse is limited, and there is a risk of losing some information due to extensive upsampling and downsampling operations applied to the original features. The Google Brain team proposed an improved version of PANet called Bidirectional Feature Pyramid Network (BiFPN) [54], which is shown in Figure 9c. BiFPN improves the structure of PANet by removing the part with only one input node.

![Figure 9. FPN, PANet, and BiFPN structures.](image)

This article is inspired by the BiFPN idea and improves the neck of YOLOv5. Based on the original PAN structure, it introduces skip connections, which helps to effectively fuse features and improve detection accuracy.

3.3.5. GGS Module

Depth-wise separable convolution (DSC) is a lightweight convolutional structure that can effectively reduce the computational burden but may lead to a significant decrease in detection accuracy. DSC splits the information in each channel during processing, which may result in the loss of a large number of hidden connections, thereby reducing the feature extraction ability and making it less powerful than ordinary convolution operations. To overcome this shortcoming, this study adopts an improved DSC structure, namely, GSConv [55]. Compared with standard convolution, GSConv has fewer parameters and is more suitable for building lightweight detection models.

As shown in Figure 10, GSConv consists of standard convolution, DSC, and random sorting. First, downsampling is performed with a convolution operation; then, the results of the two convolution operations are merged together using deep separable convolution. Finally, a random sorting operation is performed to ensure that the channel numbers of the previous two convolution operations are arranged adjacent to each other, maintaining implicit connections among channels and increasing the diversity of features. This approach preserves the quality of the feature map while minimizing its impact on the model’s inference time.
In YOLOv5, we adopt an improved convolutional module called GSs, which consists of three key components, i.e., GSConv, Group Norm (GN) [56], and the SMU activation function [57], replacing the traditional CBS convolutional module consisting of a standard convolutional layer, Batch Norm (BN), and the SiLU activation function. In most cases, although BN can improve the speed of training and convergence, when the batch size is too small, performance may decrease. This occurs because the preceding batch of data may not accurately represent the distribution of the entire input data, leading to a gradual rise in the error rate of BN. GN is a new normalization method that can replace BN. Its working mechanism is slightly different from that of BN, as shown in Figure 11. The key idea of GN is to normalize the information within each channel independently, rather than relying on the statistical information of the entire batch as in BN. This makes GN a better choice for certain specific application scenarios, particularly when dealing with small batch sizes or inconsistent input data sizes, as it tends to deliver more stable performance. In addition, GN can reduce computational complexity, because the normalization of each group is performed independently, as shown in Figure 12. As illustrated in Figure 13, when comparing the error rates of the two, it becomes evident that GN outperforms the other.
Figure 12. Group normalization principle diagram.

Figure 13. BN and GN error rates.

The SMU activation function represents an enhanced iteration of Leaky ReLU, offering improved smoothness and non-linearity. It aids in the effective training of deep neural networks. SMU achieves smoothness by dividing the input into positive and negative parts and using a logarithmic function to process the negative part. This design equips SMU to excel in various tasks, including classification, object detection, and semantic segmentation. It is worth noting that \( \text{erf} \) refers to the Gaussian error function, as shown in Equations (12) and (13).

\[
f_{\text{SMU}}(x, \alpha) = \left(1 + \alpha\right)x + \left(1 - \alpha\right) \cdot \text{erf}
\left(\mu \cdot (1 - \alpha)\right)
\]

(12)

\[
\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-t^2} dt
\]

(13)

The GGS convolutional module effectively reduces the computational burden with GSConv while incorporating the GN layer to significantly reduce the impact of the batch size on training performance. This combination not only improves the trainability of the model but also exhibits strong performance in feature extraction. Using the SMU
activation function, the smoothness and non-linearity of this activation function enable it to better extract and represent features. It helps the model better capture complex patterns and information in the input data. Experimental results have demonstrated that the utilization of the GGS convolutional module can substantially enhance the training performance of neural networks. This enhancement is evident in higher detection accuracy and improved overall performance. This holds particular significance for applications in computer vision, such as object detection. It can enhance model accuracy while optimizing resource utilization and training time.

3.3.6. GGS-YOLO Detection Model

This paper proposes a new target detection model named GGS-YOLO. The structure of the model is shown in Figure 14. This model is an improvement and innovation based on the original YOLOv5. The major enhancements include replacing the CBS module with GGS, incorporating the CA mechanism ahead of both C3 modules to enhance the network’s feature extraction capabilities. This aims to make the model more intelligent in selecting and enhancing key features. Furthermore, we also replace the original CIoU with Wise-IoU to enhance the precision of target detection, leading to improved accuracy. The introduction of the BiFPN feature fusion network could help to better integrate multilevel feature information, thus improving performance in target detection. The introduction of BiFPN makes GGS-YOLO more reliable in detecting targets in various complex scenarios.

Figure 14. GGS-YOLO structure.

4. Experimentation and Evaluation

In this section, firstly, we report the determination of the appropriate values of the Wise-IoU v3 hyperparameters through experiments. Next, we report on the ablation experiments conducted to evaluate the improvement in the overall performance of the improved parts. Finally, we report on the comparative experiments conducted to detail the advantages and improvements of the proposed method compared with other methods.

4.1. Dataset

The dataset utilized in this paper was Chinese Traffic Signs Dataset (CCTSDB), which was created by Changsha University of Science and Technology in 2021 [38]. The dataset comprises a total of 12,000 images, and each image is accompanied by corresponding annotation information. These images cover a variety of road conditions, including highways, urban roads, and rural roads, as well as various complex weather conditions, such as rainy
days, snowy days, foggy days, weak light at night, and strong light during the day and night. These conditions significantly amplify the complexity of traffic sign detection. The dataset categorizes traffic signs into three main groups: mandatory signs, prohibitory signs, and warning signs. For experimental purposes, the dataset was split into two subsets, with 9000 images being used for training the model and 3000 images being employed for evaluating the model’s performance. Some images from the dataset are shown in Figure 15.

![Figure 15. CCTSDB.](Image)

4.2. Laboratory Environment

This experiment was carried out on the Ubuntu 20.04 operating system, using a system with an Intel(R) Xeon(R) Gold 5222 CPU @ 3.80 GHz, an NVIDIA RTX 3090 24 GB GPU, and total system memory of 128 GB. The experiment used Python version 3.8 and CUDA 11.8 acceleration. In terms of deep learning framework, PyTorch 2.0.1 was used for model training. During the training stage, the following parameters were set: the total training period (epochs) was 300; the batch size per batch was 16; and eight processes (workers) were used to accelerate the data processing and the training process.

4.3. Evaluation

In this paper, the performance of the model was evaluated from various angles, and the following metrics were employed: precision, recall, average precision (AP), mean average precision (mAP), and parameter count (Params). Precision assessed the fraction of traffic signs that the model correctly predicted among all the model’s predictions. Recall quantifies the ratio of all correctly predicted traffic signs to all actual signs. Here, TP stands for the count of true-positive samples, which are actual positive signs correctly predicted as positive by the model. FP represents the count of true-negative samples incorrectly predicted as positive by the model, and FN represents the count of true-positive samples erroneously predicted as negative by the model.

\[
P = \frac{TP}{TP + FP} \tag{14}
\]

\[
R = \frac{TP}{TP + FN} \tag{15}
\]

AP quantifies the area under the precision–recall curve and serves as a comprehensive metric that considers both precision and recall in the evaluation of model performance. mAP is employed to assess the accuracy of multi-class object detection. It is calculated as
the average of the AP values for all classes. In this context, \( N \) represents the total number of classes, and \( AP_n \) represents the AP value of the \( n \)th class.

\[
AP = \int_0^1 P(t) dt
\]  

(16)

\[
mAP = \frac{\sum_{n=1}^{N} AP_n}{N}
\]

(17)

Furthermore, to assess the model’s size and speed, we incorporated parameter count as an indicator. A smaller number of parameters indicates a more lightweight model.

4.4. Experimental Results

We introduced the Wise-IoU v3 loss function to improve model performance. These hyperparameters may have different effects on different models and datasets. To attain the optimal performance of our model, we carried out a series of controlled experiments encompassing five different combinations of hyperparameters, as shown in Table 1. The bolded values in Table 1 represent the best results obtained from these experiments. These experiments aimed to determine which parameter configuration could achieve the best detection performance. After careful adjustment and comparison, we found that the model achieved the highest level of average accuracy when \( \alpha = 1.2 \) and \( \delta = 6 \), which means that the overall detection effect reached a state close to optimal performance.

<table>
<thead>
<tr>
<th>((\alpha, \delta))</th>
<th>(P)%</th>
<th>(R)%</th>
<th>(mAP@0.5)%</th>
<th>Params/10^9</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1.0, 7)</td>
<td>89.2</td>
<td>79.9</td>
<td>86.5</td>
<td>5.67</td>
</tr>
<tr>
<td>(1.2, 6)</td>
<td><strong>89.8</strong></td>
<td>81.3</td>
<td><strong>86.9</strong></td>
<td>5.67</td>
</tr>
<tr>
<td>(1.4, 5)</td>
<td>88.5</td>
<td><strong>81.7</strong></td>
<td>86.5</td>
<td>5.67</td>
</tr>
<tr>
<td>(1.6, 4)</td>
<td>87.8</td>
<td>81.5</td>
<td>86.4</td>
<td>5.67</td>
</tr>
<tr>
<td>(2.5, 2)</td>
<td>88.2</td>
<td>79.6</td>
<td>86.3</td>
<td>5.67</td>
</tr>
</tbody>
</table>

Based on the results of the ablation experiments presented in Table 2, it is evident that each enhancement has a positive influence on the performance of the model. The optimal outcomes from these ablation experiments are highlighted in bold in Table 2. First, model A introduces the CA mechanism based on YOLOv5, which improves accuracy by 8.7% while reducing the number of parameters by \( 0.6 \times 10^7 \). This improvement not only increases mAP by 0.9% but also helps better capture the target in the feature extraction process, while preserving shallow details, thus obtaining more effective feature details. Model B adopts Wise-IoU as the loss function, which increases the mAP value by 0.5%. This suggests that the loss function associated with the dynamic non-monotonic focusing mechanism contributes to the optimization of the model. Model C introduces BiFPN, which improves all indicators, especially the recall rate (by 5.1%). This indicates that the improved model has significantly improved target detection ability. Model D adopts the GGS module, which, although accuracy fluctuates, increases the mAP value by 0.9% while reducing the model parameters by \( 0.48 \times 10^7 \). This indicates that the GGS module not only helps to build a lightweight model but also effectively assists in network feature extraction. Model H is the improved GGS-YOLO, which reduces the number of parameters by \( 0.34 \times 10^7 \) compared with YOLOv5 while increasing the mAP value by 4.9%. This achieves a detection effect with fewer parameters and higher average accuracy.
Table 2. Ablation experiment.

<table>
<thead>
<tr>
<th>Method</th>
<th>CA</th>
<th>Wise-IoU v3</th>
<th>BiFPN</th>
<th>GGS</th>
<th>P/%</th>
<th>R/%</th>
<th>mAP@0.5/%</th>
<th>Params/10⁶</th>
</tr>
</thead>
<tbody>
<tr>
<td>YOLOv5s</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>75.5</td>
<td>72.5</td>
<td>81.8</td>
<td>7.1</td>
</tr>
<tr>
<td>A</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>84.2</td>
<td>73.6</td>
<td>82.7</td>
<td>7.04</td>
</tr>
<tr>
<td>B</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>79.1</td>
<td>76.2</td>
<td>82.3</td>
<td>7.1</td>
</tr>
<tr>
<td>C</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>80.8</td>
<td>77.6</td>
<td>82.4</td>
<td>7.1</td>
</tr>
<tr>
<td>D</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>74.6</td>
<td>78.7</td>
<td>82.7</td>
<td>6.62</td>
</tr>
<tr>
<td>E</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>77.7</td>
<td>80.5</td>
<td>85.6</td>
<td>6.7</td>
</tr>
<tr>
<td>F</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>85.2</td>
<td>81.7</td>
<td>84.9</td>
<td>6.72</td>
</tr>
<tr>
<td>H</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>85.6</td>
<td>82.3</td>
<td>86.7</td>
<td>6.76</td>
</tr>
</tbody>
</table>

4.5. Comparative Test

To delve deeper into the exceptional performance of GGS-YOLO in traffic sign detection, we conducted an extensive comparison on the CCTSDB dataset. This comparison involved evaluating GGS-YOLO against several other detection algorithms, including Faster R-CNN, YOLOv3, YOLOv5s, YOLOv7-tiny, YOLOv8n, and YOLOv8s. The bolded values in Table 3 represent the best results obtained from these comparison experiments. Due to multiple downsampling operations, the classic algorithm Faster R-CNN suffers from the problem of losing small target information, resulting in an mAP value of only 73.4%. However, it is worth noting that although YOLOv7-tiny performs well in terms of accuracy, with accuracy of 87.5%, GGS-YOLO achieves a higher mAP value. In summary, the overall results demonstrate that GGS-YOLO outperforms other models on the CCTSDB dataset, while maintaining a relatively low parameter count. This superiority is particularly evident in the context of small target detection, underscoring its exceptional performance. This holds significant importance in the field of traffic sign detection. As depicted in Figure 16, the enhanced model exhibits a faster convergence rate, characterized by lower loss values and improved model convergence capabilities.

Figure 16. Loss curves of different models.
Table 3. Comparative experiment.

<table>
<thead>
<tr>
<th>Method</th>
<th>P/%</th>
<th>R/%</th>
<th>mAP@ 0.5/%</th>
<th>Params/10⁹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faster R-CNN</td>
<td>59.8</td>
<td>76.7</td>
<td>73.4</td>
<td>144.5</td>
</tr>
<tr>
<td>YOLOv3</td>
<td>61.6</td>
<td>72.8</td>
<td>63.2</td>
<td>61.5</td>
</tr>
<tr>
<td>YOLOv4</td>
<td>74.9</td>
<td>73.2</td>
<td>82.6</td>
<td>96.9</td>
</tr>
<tr>
<td>YOLOv5s</td>
<td>75.5</td>
<td>72.5</td>
<td>81.8</td>
<td>7.1</td>
</tr>
<tr>
<td>YOLOv7-tiny</td>
<td>87.5</td>
<td>81.2</td>
<td>82.4</td>
<td>6.2</td>
</tr>
<tr>
<td>YOLOv8n</td>
<td>78.9</td>
<td>70.3</td>
<td>76.9</td>
<td>3</td>
</tr>
<tr>
<td>YOLOv8s</td>
<td>85.4</td>
<td>76.4</td>
<td>85.4</td>
<td>11.2</td>
</tr>
<tr>
<td>Ours</td>
<td>85.6</td>
<td><strong>82.3</strong></td>
<td><strong>86.7</strong></td>
<td><strong>6.76</strong></td>
</tr>
</tbody>
</table>

Figure 17 shows the visual detection results of YOLOv5s and GGS-YOLO in the CCTSDB dataset. In the upper section, you can observe YOLOv5s, while the lower part displays GGS-YOLO. In the initial column of images, it’s evident that both YOLOv5s and our proposed approach accurately detect traffic signs. However, the traffic signs detected by our method exhibit higher confidence, indicating that our proposed CA and BiFPN method is more precise in localizing traffic signs. In the second column of images, the original YOLOv5s mistakenly detected a sign on the billboard next to it, while our method was able to accurately distinguish between traffic signs and billboards. In the third column of images, we assessed the detection performance in challenging environments. The results indicate that our method, when compared to YOLOv5s, exhibited superior confidence in detecting traffic signs under such complex conditions.

![YOLOv5s and GGS-YOLO Comparison](image)

**Figure 17.** (a–c) Comparison of detection results between YOLOv5s and GGS-YOLO.

The experimental results underscore the significance and effectiveness of integrating innovative technologies in the field of traffic sign detection. This research indicates that by integrating and optimizing innovative technologies, the accuracy of traffic sign detection markedly improves. These integration and optimization not only enhance the model’s comprehensive ability to recognize traffic signs in complex road environments but are also crucial to ensuring the safety and efficiency of autonomous driving and connected vehicle technologies. It is noteworthy that by reducing model parameters, the efficiency and simplicity of the model are also improved, which is particularly important for resource-constrained in-vehicle systems.
5. Conclusions

This paper introduces a traffic sign detection algorithm called GGS-YOLO, which is an improved version of YOLOv5. To enhance the sustainability of the traffic sign detection framework, the algorithm incorporates blockchain technology. By combining blockchain technology with traffic sign detection, the goal is to improve the robustness and security of the IoV, while significantly saving computational resources. The GGS-YOLO algorithm employs a series of improvement measures, including the introduction of the C3-CA module, which enhances the ability to capture global feature information using improved CA. To enhance the multi-scale feature fusion capability, improvements were made to the neck section by introducing the BiFPN feature fusion network. Additionally, the Wise-IoU v3 bounding box loss function was introduced to better guide the model in learning accurate boundaries of targets. To improve the lightweight performance of the model, the algorithm introduces the GGS module, incorporating GSconv technology. Experimental results demonstrate that the model performs well in terms of accuracy. By integrating blockchain technology, not only can the sustainability of storage and computational resources be ensured, but the security of data transmission is also enhanced. However, this solution still faces some issues that require further research, including how to compress the number of data transmitted in blockchain and addressing potential false positives and false negatives in the detection model. These challenges will be the focus of future work.

Author Contributions: Conceptualization, Y.L. and Q.Q.; Methodology, Y.L. and H.Z.; Software, Y.L., H.Z. and J.L.; Validation, H.Z., J.L. and N.N.X.; Formal analysis, Y.L., Y.Z. and N.N.X.; Investigation, Y.L., Q.Q. and Y.Z.; Writing—original draft preparation, Y.L. and Q.Q.; Writing—review and editing, Y.L., H.Z. and N.N.X.; Project administration, H.Z.; Funding acquisition, Y.L. and H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by National Natural Science Foundation of China under grant 61963017; in part by Shanghai Science and Technology Program, China, under grant 23010501000; in part by Shanghai Educational Science Research Project, China, under grant C2022056; in part by Humanities and Social Sciences of Ministry of Education Planning Fund, China, under grant 22YJAZHA145.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

References


18. Jaramillo-Alcazar, A.; Govea, J.; Villegas-Ch, W. Advances in the Optimization of Vehicular Traffic in Smart Cities: Integration of Blockchain and Computer Vision for Sustainable Mobility. Sustainability 2023, 15, 15736. [CrossRef]


32. Ren, Q.; Man, K.L.; Li, M.; Gao, B.; Ma, J. Intelligent design and implementation of blockchain and Internet of things–based traffic system. Int. J. Distr. Sens. Netw. 2019, 15, 1500147719870653. [CrossRef]


Sustainability 2024, 16, 171


41. Xiao, H.; Zhang, W.; Li, W.; Chronopoulos, A.T.; Zhang, Z. Joint clustering and blockchain for real-time information security transmission at the crossroads in C-V2X networks. IEEE Internet Things J. 2021, 8, 13926–13938. [CrossRef]


52. Li, Y.; Fan, Q.; Huang, H.; Han, Z.; Gu, Q. A Modified YOLOv8 Detection Network for UAV Aerial Image Recognition. Drones 2023, 7, 304. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.