Proceedings of CONF-MCEE 2026 Symposium: Advances in Sustainable Aviation and Aerospace Vehicle Automation
DOI: 10.54254/2755-2721/2026.KA30173

Detection of New Energy Electric Vehicle Charging Ports Based on Image Recognition

Chengyun Li

Jiangsu Tianyi High School, Wuxi, China
Chengyunli318257@outlook.com

Abstract. Charging port recognition and pose estimation are key issues in enabling automatic charging of new energy electric vehicles. Complex lighting, significant vehicle model variations, and varying charging port locations in real-world environments pose challenges to detection. This paper first constructed a charging port dataset encompassing vehicles of varying vehicle models, angles, lighting, and occlusion conditions. Rotation, scaling, brightness adjustment, and local occlusion enhancement are then performed to improve the model's adaptability in complex environments. Secondly, an improved object detection network is designed: a lightweight attention module is introduced in the feature extraction layer, combined with multi-scale feature fusion to enhance the detection of small and occluded targets. Adaptive anchor boxes and an improved loss function are employed in the detection head to improve localization accuracy. Experimental results demonstrated that this method significantly outperforms baseline models in both mAP and recall. The proposed charging port detection scheme based on image recognition has high engineering application value.

Keywords: New Energy, Image Recognition, Electric Vehicle, Charging Port

1. Introduction

In recent years, with the global push for low-carbon mobility and energy transitions, the new energy vehicle industry has entered a period of rapid development [1]. The Chinese government has continuously introduced subsidies, built charging infrastructure, and provided industry support, leading to a rapid increase in electric vehicle ownership and charging demand. Against this backdrop, the intelligent and automated nature of charging equipment and services has become a key focus for improving user experience, reducing operating costs, and increasing charging efficiency [2]. The automatic connection between electric vehicles and charging stations not only reduces manual intervention and improves charging safety, but also provides necessary technical support for the large-scale deployment of new energy vehicles [3].

In the auto-charge process, the detect of charging port and its pose is the crucial visual task. The accuracy of detection of the charging point location directly determines how successfully the machine charging gun inserted, in addition, it’s also significant in the process of safety and stability. However, in the real world, the visual detection is affected by different factors in the world, such as strong light at back, the low intensity of light in the evening, shadow and barriered by car’s structure

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.KA30173

and panels. These complicated situations make manually or traditionally method difficult to keep high quality of precise detect ability [4].

In recent years, the imagine detection technology and technology in deep learning made significant process in target detect and pose estimation [5]. By constructing large scale, variable types database of images, which combined multi-scaled feature extract, attention mechanisms, and single/double depth estimation [6], this model can maintain high detect and locate ability in complicated environment, provide a new solution in the environment of auto-charged background. Recent studies mainly focused on the detections of car license plate, facial detect and sign detect, shows robustness of the deep neural networks model in different environments [7,8]. However, there is still catastrophic lost in detection charging ports on electric car, first, it lacks cover multi models' database of high-quality charging port's images; besides, detections in nowadays network lacks of adaptability in small target and barriers; third, the accuracy of pose estimation and stability need to improve.

Therefore, how to achieve high quality detection and reliable pose estimation has been the research problems in recent dynasty. Research in this passage is to propose a practical solution to solve the problems in charging port visual detection. It has a varied database and can design a improved detection network. This can combined attention light model and multi-scale feature together [9]. This can combine improved detection head and loss function to improve the ability to estimate in the locating ability. This research not only fufills the empty of area in charging port detection, this research also provides a possible way to the system of new energy vehicles in the future.

2. New energy vehiclesn

As the wide spread of the use of new energy vehicles, the multi-types and complexity are being apparent through everyday, this brings a challenge to the auto visual detect system. This research most specific on new energy cars and their charging ports, nor only include charging ports itself but also include cover, car features and some possible signs which may block the visual detection. This research use single or double eye camera system to realize precise detect and estimate of pose. In real world, cars are placed in different environments, such as charging station, gardens or auto charging sets. Every scenario brings specific visual challenge, including angle of car's park, distance between camera and the changing between visual angles; the changing in times of light between day and night; the environment condition in sunny, snowy and foggy days. These factors can significantly decrease the quality of images and decrease the ability to see charging ports. Moreover, extrinsic factors such as dust, stains, scratches and other physical barriers may make the detection task more complicated [10]. As a result, robust and adaptable detection methods are required to maintain high quality in different real-world scenarios.

Currently, commonly used charging ports in electric vehicles can be generally separate into 2 parts, one is DC fast charging ports, the other is AC slow charging ports. First one generally uses 9-pin setting, enabling high current transporting and high-speed charge ability. However, another one usually uses 7-pin structure, this provides less current and more charging time. Except this category, different manufacturers' charging port are quite different especially in location and performance. That are different in interior electric lines and installing methods. This made the detect model in different types of cars and the design of charging port perform difficult in robustness. As the developing of the charging system need, it is not reliable to use traditional one in such complex and changing conditions. In contrast, based on deep learning model like YOLO model [11], it has high precise and strong real-time ability, and this process combined with attention mechanism, which

DOI: 10.54254/2755-2721/2026.KA30173

made the model can focus on related space and neuron networks. It also overcomes the limit of the traditional one. This improvement significantly increases the detect speed and ability, and provide a proper calculation framework for the detection. However, the absence of large-scale, diverse, and specific database limits the stability of this model, especially in complicated and disadvantage environments.

In this background, the detection of charging ports still faces 3 challenges, which must be solve as quick as possible in order to keep reliable and accurate execution. First, complex environment factors significantly affect visual sense. The changing of lights, including strong backlight, reflection or low angle at night, together with rain, snow and other bad weather conditions, which may decrease the quality of images, decrease the accuracy of pose estimate and detection. In addition, the structure barriers and manufacturer difference increase the difficulty of detection. Some charging points are partly of fully covered, decorated or being blocked by surrounding areas, those stains and damages may cover identifying features. Moreover, differences in shape, size and locations of charging ports make it complex, which need an algorithm to identify different targets. Finally, small target detection is still a crucial challenge. The charging port only takes a small part of the whole image in one frame, but their internal specific features such as pin and guide slot, are crucial to achieve precise mechanical splice. Traditional network detect may probably cannot consider both global context and fine-grained features simultaneously, leading to wrong detect, leak or the inferior of the pose estimation.

3. Analysis

The charging port detection technology based on image recognition has significant advantages in the application of automatic charging for new energy vehicles [12]. Firstly, it reduces manual intervention and enables unmanned or semi unmanned charging operations, thereby improving charging safety and operational convenience. By constructing a diverse and representative dataset covering multiple vehicle models, different angles, different lighting conditions, and different degrees of occlusion, and using rich data augmentation techniques such as rotation, scaling, brightness adjustment, and occlusion simulation, the model can maintain robustness in complex real-world scenarios [13]. In addition, by integrating multi-scale feature fusion and attention mechanisms derived from deep learning frameworks, detection networks can more effectively capture features of small, partially occluded, or visually subtle objects, significantly improving localization and detection accuracy. In addition, by applying lightweight network architecture and combining model pruning and quantization methods [14], this algorithm can achieve real-time inference on embedded platforms with limited computing resources, providing a feasible approach for the practical engineering implementation and large-scale deployment of intelligent charging stations.

However, despite there's such advantages, this method still faces many limitations which constrain its properties. The complex lighting conditions and bad weathers is still a big challenge, because strong reflection, backlight, low light intensity and conditions like rain, snow and fog may severely decrease the quality of input of camera, which decrease the accuracy of detection and pose estimation. Furthermore, existing training data may cannot fully capture all possible extreme situations, limits the model adaptability when facing unexpected or rare situations. In addition, the charging port may be blocked by dust, stain, covers or decorations, which increase the difficulty of detection. Moreover, different factories produce different designs of charging ports, including shapes, sizes, pins and locations, which challenges detect model's ability. In pose estimation, monocular vision systems face inherent scale ambiguity [15], while binocular may encounter

DOI: 10.54254/2755-2721/2026.KA30173

unstable feature matching in low texture areas or at long distances [12], which reduces reliability in the future. Last, because the charging port only takes a small part in the whole image, so it is very crucial to introduce some precise features like pins combinatorics. However, traditional detection networks are hard to do this, so that lead to mistakes or missing.

Although these challenges still exist, but there is still potential that the new energy car has. In order to has reliable and large-scale set, it is needed to overcome current bottleneck. Although these challenges still exist, detections based on charging port in smart charging on new energy cars still perform extremely potentials. To achieve truly reliable and large-scale deployment, further advancements are required to overcome current technological bottlenecks. Future research can be pursued along three main directions. First, the integration of multimodal perception methods—such as combining image data with LiDAR or other sensor information—can enhance the model's robustness and resilience in complex and dynamic environments. Second, the construction of a larger, more diverse, and high-quality dataset that covers a wider range of vehicle models, extreme weather conditions, and challenging usage scenarios will improve the generalization and reliability of detection models. Third, the development of pose estimation strategies that combine deep learning with geometric or physical constraints can improve both the accuracy and stability of docking predictions. Such approaches may involve hybrid optimization of image-based predictions and spatial constraints derived from vehicle geometry, further enhancing robustness in practical deployment. Collectively, these improvements are expected to drive the maturation of automatic charging technologies, providing strong technical support for the widespread adoption of intelligent charging solutions and facilitating the large-scale popularization of new energy vehicles, ultimately contributing to a safer, more efficient, and user-friendly charging ecosystem.

4. Suggestions for improving the charging port detection system

The current research highlights significant limitations in dataset diversity and coverage, which directly constrain the generalization ability of charging port detection systems. To address these challenges, future work should systematically expand dataset coverage through partnerships with major electric vehicle manufacturers such as Tesla, BYD, NIO, and Xiaopeng. Collecting charging port images from a wide range of vehicle models, including DC fast charging (9-pin) and AC slow charging (7-pin) interfaces, will ensure coverage of different design and manufacturing standards. In addition, the dataset should be enriched with extreme environmental conditions such as heavy rain, snow, fog, strong sunlight, and low-light parking lot settings. Incorporating occlusion scenarios—such as partial coverings, dust accumulation, and decorative elements—will enhance robustness in real-world applications. To supplement physical data collection, Generative Adversarial Networks (GANs) can be employed to generate realistic synthetic images of charging ports in rare but critical scenarios, improving data diversity while maintaining statistical consistency with real-world distributions. Ensuring high-resolution imaging standards (minimum 1080p) will allow fine details of charging port pins and guide slots to be captured for precise docking operations.

While YOLO-based methods have demonstrated satisfactory performance, there remains considerable potential for improvement in handling complex cases and meeting the precision demands of autonomous charging. Future research should consider integrating hierarchical attention mechanisms operating at multiple scales, enabling global attention for port localization and local attention for detailed pin detection. A hybrid architecture combining Vision Transformers (ViTs) with CNN backbones could enhance feature representation and capture long-range dependencies critical for pose estimation. By using an adaptive feature aggregation module to enhance the Feature Pyramid Network (FPN), feature weights can be dynamically adjusted based on the size of the

DOI: 10.54254/2755-2721/2026.KA30173

charging port and input characteristics, thereby improving consistency between different models and camera distances. Domain specific convolution kernels designed for circular and rectangular geometric shapes can further improve detection accuracy. In addition, the dual head architecture combining object detection and keypoint regression can achieve accurate pose estimation required for docking. The standardized definition of key points will ensure consistency across manufacturers, while integrating uncertainty prediction mechanisms will allow the system to request manual intervention when confidence is below a safety threshold.

Finally, it is crucial to address the limitations of unimodal vision. Multimodal perception should combine two-dimensional visual data with three-dimensional LiDAR point clouds to improve depth and pose estimation through early feature level fusion and later decision level fusion. For double-eye camera system, it is needed to develop a cubic matching algorithm to geometric shape of the charging port so that it can overcome the challenge of textured surface and keep real-time effiency. In extreme weathers when the perform of traditional camera is poor, the thermal imaging can provide additional advantages. At the same time, safety and reliability is still crucial. Fault safety mechanisms with multi-stage protocols should detect and confirm before connection in order to decrease the risk of damage and broken. Real time anomaly detection and learning framework makes the system can suit for new design and environment, together protect the privacy. Standard correction program, communication protocols and base test will confirm the operation ability between manufacturer and infrastructure providers. Separated-stage strategy, which means starting from private parking lots and expanded to public networks, which will achieve gradual, reliable applications and continuous performance improvements.

5. Conclusion

Technology based on image detections has advantage in applications in new energy cars. Firstly, in reduce manually interception, successfully achieve non-people or semi non-people charging operators, increasing charging safety and operating convenience. By constructing datasets from different sizes, angles, lights and blocks of multi-vehicles, and use spin, magnify, light and barriers to simulate large amount of advanced technology, this model can keep robustness in the complex real world. Moreover, by deep learning framework multi-scale combination feature and attention mechanisms, it is able to capture small, part blocked or subtle features, so that improve locating ability and causing a more precise detection. Furthermore, using light framework application, combined model pruning and quantization methods. This algorithm can achieve real time logical thinking on a limited computing resource, which provide an available method for actual engineering and large-scale smart charging stations.

Another important achievement is the integration of multimodal perception capabilities, it combines visual detections and deep sense, pose estimate together. This development represents the significant progress in the car sensor application. This framework successfully solves the problems of the long time technology in vehicle charging problems, that is, lack of standard algorithm to solve the charging terminal in cases of facing different manufacturer. Through the test procedure, the framework performs same performance in fast and slow speed of charge, this satisfies current market demand. Except the charging application, these contribution also provide the wide spread of method note, such as the system of auto-driving car which need high-precise locating ability and robot control task.

Although such achievements are made, but there is still some limitation need to research and improve. In extreme weathers such as heavy snow and foggy days, system performers need to improve, or even need extra clean mechanisms to keep accuracy. Same, the real-time processing

DOI: 10.54254/2755-2721/2026.KA30173

calculation is still a keeping challenge, emphasize the need of continuously improve of hardware and technology. However, this research has significant meaning to the basic traffic and auto driving technology. Through reliable auto charge detect system, this research can help solve the availability of charge bottleneck, promote electric cars wider use. Last, this technology improvement will promote the changing from traditional fossil funerals to clan and recycle funerals, so that support sustainable goals

References

[1] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 17, 779-788.
[2] Tadic, V., Odry, A., Vizvari, Z., Kiraly, Z., Felde, I. and Odry, P. (2024) Electric Vehicle Charging Socket Detection Using YOLOv8s Model. Acta Polytechnica Hungarica, 21, 150-165.
[3] Wang, C.Y., Bochkovskiy, A. and Liao, H.Y.M. (2023) YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 23, 7464-7475.
[4] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. (2017) Attention Is All You Need. Advances in Neural Information Processing Systems, 30, 61-73.
[5] Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B. and Belongie, S. (2017) Feature Pyramid Networks for Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 10, 2117-2126.
[6] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 13, 7132-7141.
[7] Sun, S., Mo, B., Xu, J., Li, D., Zhao, J. and Han, S. (2024) Multi-YOLOv8: An Infrared Moving Small Object Detection Model Based on YOLOv8 for Air Vehicle. Neurocomputing, 588, 127685.
[8] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 18, 770-778.
[9] Eigen, D., Puhrsch, C. and Fergus, R. (2014) Depth Map Prediction from a Single Image Using a Multi-Scale Deep Network. Advances in Neural Information Processing Systems, 27, 170-179.
[10] Godard, C., Mac Aodha, O. and Brostow, G.J. (2017) Unsupervised Monocular Depth Estimation with Left-Right Consistency. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 17, 270-279.
[11] Kershenbaum, A., Akçay, Ç., Babu-Saheer, L., Barnhill, A., Best, P., Cauzinille, J. and Dunn, J.C. (2025) Automatic Detection for Bioacoustic Research: A Practical Guide from and for Biologists and Computer Scientists. Biological Reviews, 100, 620-646.
[12] Scharstein, D. and Szeliski, R. (2002) A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. International Journal of Computer Vision, 47, 7-42.
[13] Li, X., Wang, W., Wu, L., Chen, S., Hu, X., Li, J., Tang, J. and Yang, J. (2020) Generalized Focal Loss: Learning Qualified and Distributed Bounding Boxes for Dense Object Detection. Advances in Neural Information Processing Systems, 33, 21002-21012.
[14] Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R. and Ren, D. (2020) Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. Proceedings of the AAAI Conference on Artificial Intelligence, 34, 12993-13000.
[15] Chen, L., Zhang, H., Xiao, J., Nie, L., Shao, J., Liu, W. and Chua, T.S. (2017) SCA-CNN: Spatial and Channel-Wise Attention in Convolutional Networks for Image Captioning. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 13, 5659-5667.