Application and Challenges of Automotive Visual Perception Technology Based on Deep Learning

Yicheng Fu

{ yicheng.22@intl.zju.edu.cn }

Department of Civil and Environmental Engineering, International Campus, Zhejiang University, Jiaxing, Zhejiang, China

Abstract. Under the trend of the deep integration of artificial intelligence and the automotive industry, the development frontier in the transportation field is intelligent connected vehicles. The core prerequisite for achieving autonomous driving is to endow vehicles with precise environmental perception capabilities. This article systematically reviews the current application status and challenges of automotive visual perception technology based on deep learning. The article first provides an overview of the basic paradigms of deep learning and its core role in key visual tasks of autonomous driving. Next, a detailed analysis will be conducted on the typical applications of these technologies in advanced driver assistance systems and high-level autonomous driving, as well as their implementation principles. However, when this technology is widely applied, it still faces many severe challenges such as environmental adaptability, safety and reliability. Therefore, at the end of this article, the future development trends such as sensor fusion, model lightweighting and algorithm optimization are prospected. The analysis results show that deep learning visual perception technology is a key driving force for the development of intelligent vehicles. However, to achieve full implementation, continuous breakthroughs are still needed in terms of technological robustness and engineering.

Keywords: Deep Learning, Autonomous Driving, Automotive Visual Perception

1 Introduction

Intelligent connected vehicles have become a new focus of development. As the core driving force, autonomous driving technology can build a safer and more convenient transportation system. However, the prerequisite for achieving this goal is to endow vehicles with environmental perception capabilities similar to those of human drivers, ensuring that vehicles can accurately and in real time identify and understand the dynamic environment around them, such as road structures, signal signs, and various unexpected situations. Visual sensors like cameras, which are the closest to human visual experience and relatively low-cost, have become an indispensable part of autonomous driving systems [1]. However, traditional computer vision methods that rely heavily on manually designed features are difficult to meet the strict requirements of autonomous driving for perception accuracy and reliability in complex real

ICMEEA 2026, April 24-26, Singapore, Singapore

DOI 10.4108/eai.24-4-2026.2364826

traffic scenarios. But deep learning has brought about a huge change in the field of computer vision. It can automatically learn hierarchical feature representations from massive data through deep neural networks and has demonstrated unprecedented performance in various tasks, greatly promoting the development of automotive visual perception technology. Models based on deep learning not only perform well in standard scenarios on clear and sunny days but also show great potential in extreme and harsh environments.

Integrating deep learning technology into automotive visual perception faces many severe challenges. For example, the performance of automotive intelligent driving models highly depends on large-scale high-quality labeled data. However, obtaining and labeling data for extreme cases is extremely costly. In addition, under the limited computing resources of on-board platforms, achieving efficient inference of complex models is also an engineering challenge to be solved.

Therefore, this paper aims to systematically review the development status, application achievements, and challenges of automotive visual perception technology based on deep learning, and look forward to future research directions and development trends, hoping to provide researchers in related fields with a clear technical panorama and valuable references.

2 Core Technologies of Visual Perception Based on Deep Learning

2.1 Environmental Perception and Image Acquisition

Configuration of In-Vehicle Vision Sensors. As an irreplaceable core component of the intelligent vehicle's environmental perception system, the visual sensor is a front-end perception device that collects raw image information. Unlike sensors such as lidar and radar that directly output distance or speed, the visual sensor can obtain various images around the vehicle that contain rich texture and semantic information [2]. Theoretically, it can acquire all the necessary information required for autonomous driving, such as lane lines, drivable areas, and traffic signs. Although the video signal itself contains little and rather indirect information, and it needs to go through multiple layers of algorithms to obtain effective information, it is also regarded as one of the key factors for achieving major breakthroughs in functions such as ADAS and autonomous driving due to its huge potential. From the earliest monocular night vision, reversing camera, and 360-degree parking assistance technology, to the current integrated lateral and longitudinal control of autonomous driving, the visual sensor has been officially applied to mass-produced models and has achieved good application effects in some real scenarios.

If the on-board visual system is to achieve comprehensive environmental perception, its processing flow is usually a step-by-step refinement. The initial step is "early vision", which involves processing the intermediate products (dense or sparse) generated after processing a single frame, two-frame difference, and multi-frame fusion, such as depth and motion vectors. Then it enters the second step "mid-level vision", where it detects and recognizes corresponding traffic participants (vehicles/pedestrians) and traffic elements (traffic signs) by leveraging appearance cues and geometric motion cues, and finally abstracts all the information into forms such as occupancy grids (DRMs), dynamic Stixels, or XYZC target lists, providing direct input for subsequent trajectory planning and decision control[3].

Image Data Preprocessing and Augmentation. Image data preprocessing and augmentation involve optimizing and expanding the raw images captured by on-board cameras. This technology can address the challenges posed by the complex real-world environment. The focus is on restoring the quality and normalizing single-frame images to provide clear and standardized input for subsequent perception algorithms. This includes advanced image restoration algorithms based on deep learning that can effectively suppress image degradation caused by adverse weather and vehicle movement. The raw data usually contains noise and inconsistencies. If used directly for training, it can significantly affect the accuracy and reliability of the model. Through a series of steps such as data cleaning, integration, transformation, and data reduction, the raw data is transformed into high-quality data suitable for algorithm analysis. This process can even take up 50% to 80% of the entire classification project's time [4].

Based on preprocessing, the process of data augmentation involves applying a series of random yet reasonable transformations to images during the training phase, thereby expanding the training dataset in a cost-effective and large-scale manner through algorithmic means. Its core value lies in introducing artificial diversity to force the model to learn invariance to irrelevant transformations, thereby significantly enhancing its generalization ability and effectively preventing overfitting. Specific techniques include basic geometric transformations and color space perturbations. More importantly, the generation of synthetic data through generative adversarial networks or specialized simulation engines provides the model with extremely important "negative samples" and corner case training data, systematically filling the distribution blind spots of real datasets.

2.2 Image Processing and Object Recognition

Object Detection Technology. Object detection is a fundamental computer vision technology that enables environmental perception in autonomous driving systems. It uses computer vision techniques to identify and classify objects from predefined categories within complex scenes.

Researchers have enhanced object detection performance by employing the Faster R-CNN algorithm. Applying a dual-channel feature extraction network and a region proposal optimization strategy helps to improve vehicle detection accuracy. Experimental results reveal that under adverse weather conditions, this approach achieves higher detection accuracy compared to single-sensor systems and LiDAR-based methods. This highlights the effectiveness of heterogeneous sensor fusion in enhancing the reliability and environmental robustness of object detection systems [5].

Semantic Segmentation and Instance Segmentation Techniques. As the key image segmentation technology for achieving pixel-level environmental understanding in autonomous driving, it consists of two parts: semantic segmentation and instance segmentation [6]. Among them, the classification of each pixel point in the image is carried out, and it is classified into predefined semantic categories such as roads, vehicles, sky, and buildings to obtain a dense semantic annotation map, which enables vehicles to understand the overall arrangement and combination of the current scene. This is semantic segmentation. Technologies such as U-Net based on the encoder-decoder structure and the DeepLab series, all belong to further processing techniques such as the division of drivable areas and lane line recognition based on semantic segmentation. Instance segmentation is a further enhancement on the basis of semantic segmentation, in addition to achieving pixel-level classification, it also needs to distinguish different individuals within the same semantic category. Generally, it combines target detection to achieve the technology of clearly outlining the contour of a single instance when locating a single independent target.

To achieve these goals, researchers have developed numerous representative architectures. For example, the fully convolutional network (FCN) and U-Net, by using the encoder-decoder structure and combining skip connections to restore spatial details, effectively solve the problem of semantic segmentation; while models such as Mask R-CNN, on the basis of target detection, add a parallel mask prediction branch, thereby being able to simultaneously complete the positioning, classification, and pixel-level segmentation of the target, becoming the classic paradigm of instance segmentation. Additionally, the DeepLab series models, by introducing dilated convolutions and spatial pyramid pooling to capture multi-scale contextual information, further improve the accuracy of segmentation boundaries [7]. These technologies have shown great application potential in multiple fields such as autonomous driving and medical image analysis.

In autonomous driving, semantic segmentation is used to understand the overall topological structure of the road, while instance segmentation is used to precisely perceive the geometric shape and occupied space of obstacles. The two complement each other, jointly constructing a comprehensive scene understanding from the macro layout to the micro individual.

Monocular Depth Estimation. Monocular depth estimation is a highly challenging and crucial visual perception technology for autonomous driving. Its core task is to infer the three-dimensional geometric structure of a scene from a single camera's two-dimensional image, assigning precise distance (depth) values to each pixel and generating a depth map. Unlike solutions relying on stereo vision or active sensors, this technology attempts to solve the typical "ill-posed problem" - there are infinitely many solutions to inferring the three-dimensional world from two-dimensional information. Its success heavily depends on the scene geometry prior and semantic understanding capabilities learned by deep learning models from massive amounts of data. In recent years, significant progress has been made in this field through two paradigms: supervised learning (relying on ground truth depth labels collected by LiDAR) and more scalable self-supervised learning (utilizing the motion parallax between consecutive frames in video sequences as self-supervisory signals).

Take the Monodepth2 model proposed by Godard et al. as an example. This framework uses a depth network to predict depth maps and a pose network to estimate camera motion simultaneously. Its key contribution lies in introducing a per-pixel minimum reprojection loss to effectively handle occlusion problems; designing an automatic mask to filter out the interference of static pixels; and employing full-resolution multi-scale sampling to reduce artifacts. Experiments on the KITTI dataset show that this model significantly outperforms other self-supervised methods of the same period in multiple metrics [8], demonstrating its effectiveness in generating high-quality, sharp depth maps in unseen scenes.

Through monocular depth estimation, low-cost three-dimensional environmental perception can be achieved. This cannot only be directly used for obstacle distance perception and forward collision warning but also the output depth maps can be converted into pseudo-LiDAR point clouds, thereby enabling a wide range of existing three-dimensional object detection and segmentation algorithms based on pseudo-LiDAR point clouds to be implemented with pure visual solutions. This plays a crucial role in enhancing the performance and redundancy of pure visual systems.

3 Typical Applications of Deep Learning Visual Perception in Intelligent Vehicles

3.1 Application of Advanced Driver Assistance Systems

Adaptive Cruise Control. As a core function of advanced driver assistance systems, the Adaptive Cruise Control (ACC) system has achieved a technological leap from traditional cruise control to intelligent vehicle following by relying on deep learning visual perception technology. This system continuously collects road information in front of the vehicle through on-board cameras and uses deep learning algorithms to analyze and predict key parameters such as the position, distance, and relative speed of the vehicle ahead in real time, providing precise speed adjustment and safe distance control for the vehicle.

In the ACC system, the visual sensor generates a continuous stream of visual information by scanning the environment, providing a feasible solution for the car to detect its surroundings. The application of deep learning algorithms enables the vehicle to process historical data and real-time information comprehensively. Therefore, it possesses the ability to analyze and adapt to various traffic conditions in real time. It can accurately identify different types of vehicles ahead and adjust the following strategy according to their driving characteristics [9].

High-performance chips with the characteristics of advanced artificial intelligence technologies provide drivers with more intelligent driving assistance and safety guarantees. They achieve functions like driving environment perception, behavior prediction and intelligent decision-making [10]. This system not only maintains the set cruise speed but also automatically adjusts the speed when it detects the front vehicle decelerating or changing lanes. When the vehicle ahead leaves the lane or accelerates, the system gradually returns to the set cruise speed. The process is smooth and natural, significantly enhancing driving comfort and safety. With continuous learning and optimization, the ACC system is moving towards a more intelligent and humanized direction.

Automatic Emergency Braking. In terms of complex traffic environments, the automatic emergency braking system based on computer vision has significant advantages. The deep learning model trained with a large amount of real scene data can effectively handle the target detection tasks under challenging conditions, involving light changes and shadow occlusions. When the system detects a potential collision threat, a multi-level warning mechanism will be immediately activated. A slight braking will be adopted to remind the driver to pay attention. If the driver does not respond in time and the collision risk continues to increase, the system will perform full braking to minimize the collision damage [11]. During this process, the visual sensor can provide a clearer perception effect than the human eye, especially in night or obstructed vision environments.

Modern automatic emergency braking systems adopt a multi-sensor fusion strategy. They integrate visual perception with other technologies like millimeter-wave radar, ultrasonic sensors to build a more reliable safety protection system. The visual system, which is specialized in target classification and recognition, can distinguish different types of obstacles and predict their behavior patterns. The radar system, which can provide precise distance and speed measurements, complements it. This fusion solution significantly enhances the robustness and reliability of the system and ensures that the emergency braking function can still work under various adverse weather conditions [12].

Traffic Sign Recognition. In the field of traffic sign recognition, the widely used convolutional neural network improves recognition accuracy and significantly enhances the system's ability to recognize intact or damaged traffic signs. Compared with traditional computer vision methods, the deep learning methods have stronger representational capabilities in extracting feature information. Even when traffic signs are integrated into complex road backgrounds, they can maintain good detection performance [13].

Modern traffic sign detection algorithms mainly adopt two types of methods based on deep learning: one is the two-stage detection algorithms represented by RCNN, Fast-RCNN, and Faster-RCNN; the other is the single-stage detection algorithms represented by the YOLO series. Researchers continuously improve and optimize these algorithms, such as reducing computational load by introducing depthwise separable convolution or improving detection accuracy by replacing loss functions. The improved algorithms can maintain high accuracy while effectively solving the problems of large algorithm computation and large model size in traditional methods.

3.2 Advanced Autonomous Driving Applications

Behavior Prediction and Trajectory Planning. Based on the deep learning model, through the analysis of the real-time video streams captured by the vehicle-mounted cameras, key information such as lane lines, traffic signals, pedestrians, and vehicles can be accurately identified, thereby endowing the autonomous driving system with comprehensive, accurate, and real-time road perception capabilities [14]. This technology not only achieves static recognition of environmental elements, but more importantly, it can also realize dynamic prediction of the behavioral intentions of other traffic participants. The traditional intention recognition methods mainly rely on statistical learning techniques such as support vector machines and naive Bayes, conduct analysis based on manually designed features. With the continuous progress of deep learning technology, methods based on convolutional neural networks have gradually become mainstream. By recognizing the movement patterns and speed changes of pedestrians, modern visual perception systems can predict the possible movement paths of pedestrians. Based on this, the system can adjust the vehicle speed or change the driving route in advance to avoid potential collisions.

End-to-end Autonomous Driving. Deep learning visual perception is represented by the cutting-edge application in the field of intelligent vehicles - end-to-end autonomous driving technology, which directly maps the original visual input to driving control instructions. This technical solution breaks the limitations of the traditional modular architecture [15]. Autonomous vehicles that can achieve more intelligent driving behaviors have emerged by delegating scene understanding and driving decisions to neural networks. This method avoids the problem of information loss and error accumulation among the sub-modules of the traditional architecture and provides a more unified and efficient solution for intelligent vehicles.

In terms of the technical implementation path, end-to-end autonomous driving mainly presents two different paradigms, namely imitation learning and reinforcement learning. Imitation learning, also known as behavior cloning, uses the method of supervised learning to train the driving decision model to imitate the driving behavior of experts. With this method, basic driving skills can be quickly acquired, but when dealing with scenarios not present in the training data, the performance may be poor [16]. Waymo's fifth-generation system adopts a multi-sensor fusion architecture (5 laser radars + 29 cameras), achieving object detection up to

300 meters away with 360-degree laser radars, and complementing it with 6 millimeter-wave radars for complementary perception. While Tesla's pure visual solution has a hardware cost as low as $300, it uses its self-developed FSD Chip 3.0 to achieve 450 TOPS computing power, reducing the industry average cost of $2,450 by 72%. The reinforcement learning method learns the optimal driving strategy by interacting with the environment. Although the training process is more complex, it has stronger generalization ability and adaptability. The integration of deep learning technology provides strong technical support for both methods. It enables neural networks to process complex visual information and make reasonable driving decisions. Comparative case studies show that XNGP of Xiao Peng which used a BEV + laser radar fusion model, achieved a 98% lane change success rate in the NOA scenario of the central area of Shanghai [2].

4 Challenges and Development Trends

4.1 Technical Performance Challenges

The environmental adaptability of visual sensors is a core challenge. The main reason is that their recognition accuracy under adverse weather conditions varies. The performance of image-based 3D object detection methods significantly declines when the target is covered or is at a long distance. For example, the accuracy of monocular depth estimation in rainy and snowy weather can decrease by up to 30% [17]. Therefore, enhancing the system's robustness requires the use of self-supervised learning frameworks and multi-sensor fusion technologies.

The contradiction between algorithm complexity and real-time requirements is also significant. Although deep learning models can improve the understanding of complex scenes, their high demand for computing resources makes it difficult to meet the millisecond response requirements of vehicle-mounted systems. Data quality and annotation accuracy also directly affect system performance. Obtaining high-quality annotated data requires a huge investment [18], and the significant differences in road environments in different regions limit the model's generalization ability. Edge cases such as abnormal driving behaviors, sudden situations, and rare scenarios lack sufficient training samples, and synthetic data generation techniques need to be used to fill the distribution blind spots.

4.2 Safety and Cost Challenges

Algorithm reliability is a core challenge in security. Uncovered edge scenarios in the training set are prone to leading to misjudgments. Harsh weather conditions, complex lighting, and adversarial attacks all pose security risks [19]. By processing multi-sensor conflicting data using Dempster-Shafer evidence theory, the confidence level of target detection can be increased by 15% in rainy and foggy weather; Bayesian neural networks achieve a 99.9% recall rate in complex road conditions by providing confidence interval outputs. Waymo's fifth-generation system adopts a five-redundancy sensor architecture to achieve a failure rate of less than 0.1%; XNGP of Xiao Peng reduces the false detection rate of construction sections from 8% to 2% through the BEV perception framework and dynamic map matching [2].

Cost challenges restrict the industrialization process of the technology. In terms of hardware, the costs of GPUs, dedicated AI chips, and high-resolution camera solutions are high; in the algorithm development process, data annotation and model optimization require continuous investment from professional teams, and the specific adaptation to different application scenarios further increases the R&D costs [20]. Industrial deployment also requires a large amount of testing verification and system integration investment.

4.3 Future Development Trend

Future technologies will develop towards intelligence and efficiency. The generalization ability and robustness of models will be continuously enhanced through data augmentation, transfer learning and model fusion. Sensor fusion technology is becoming increasingly mature. Deep learning algorithms integrate information from multiple sensors to provide a unified output for decision-making systems, improving perception accuracy and stability. Lightweight models such as the YOLO series can achieve real-time processing while maintaining accuracy. Combined with dedicated chips and edge computing, higher-performance local processing capabilities will be realized, enhancing system response speed and security.

5 Conclusion

Currently, the deep learning visual perception technology, which moves from theoretical research to practical application, plays a crucial role in the process of automotive intelligence. For the vehicle execution system, discussions at present mainly focus on the application status of advanced auxiliary driving systems and high-level autonomous driving. For the current intelligent connected vehicles, visual sensing technology is an extremely important component. With this visual sensor technology, intelligent connected vehicles can better understand the surrounding scenes. From the human perspective, the human observation range is relatively small. With this technology, the detection range can be expanded, and it can conduct good detection of pedestrians, vehicles and other potential hazards regardless of day or night or in various environmental conditions.

In the application of advanced auxiliary driving systems, deep learning technology demonstrates proficiency in scenarios such as adaptive cruise control, automatic emergency braking, and lane keeping assistance. Using image perception as the data input method, after introducing deep learning and neural network models, it gradually transforms from a black box to an intelligent tool that can learn and reason on its own. By integrating information from various sensors and aggregating this information based on time factors, it generates output result values and enables the decision-making system to achieve more accurate output. In the application of autonomous driving, deep learning has great potential for achieving complex scene understanding and behavior prediction.

In the face of many difficulties and technical bottlenecks, deep learning visual perception still needs to be continuously improved in terms of technical performance, safety performance, and cost reduction. In the direction of intelligent driving, computer vision integrates multiple sensors through more refined environmental perception, scene adaptability improvement, and decision-making intelligence enhancement. This helps to achieve higher-level autonomous driving. With the continuous improvement of algorithm deep learning capabilities and sensor accuracy, technological progress will inevitably lead to the reduction of costs for large-scale commercialization. One of the key points of deep learning model optimization is the refinement of modeling work in time and space dimensions. Such improvement is beneficial to enhancing the system's perception ability in complex environments. In the future, unmanned driving will achieve better integration with other systems like intelligent transportation systems and intelligent cities, thereby promoting a fundamental change in the overall transportation ecosystem. Predictably, it can bring people a safer, more efficient and sustainable transportation experience.

References

[1]. Dongliang G., Meng W., Zhiqiang Z., Juan S.: Research on visual sensor occlusion test tools for autonomous vehicles. Journal of Physics: Conference Series,2902(1),012008-012008 (2024)

[2]. Reshma K & Sangeetha Gopan G. S.: Implementation of Vision and Lidar Sensor Fusion Using Kalman Filter Algorithm. International Journal of Sciences: Basic and Applied Research (IJSBAR),55(2),183-198. (2021)

[3]. Maharana, K., Surajit M., Bhushankumar N.: A review: Data pre-processing and data augmentation techniques. Global Transitions Proceedings 3.1: 91-99. (2022)

[4]. Liu, Z., et al.: Robust target recognition and tracking of self-driving cars with radar and camera information fusion under severe weather conditions. IEEE Transactions on Intelligent Transportation Systems 23.7: 6640-6653. (2021)

[5]. Papadeas I., Tsochatzidis L., Amanatiadis A, Pratikakis I.: Real-Time Semantic Image Segmentation with Deep Learning for Autonomous Driving: A Survey. Applied Sciences,11(19),8802-8802. (2021).

[6]. Ghosh, S., et al.: Understanding deep learning techniques for image segmentation. ACM computing surveys (CSUR) 52.4: 1-35. (2019)

[7]. Godard, C., et al.: Digging into self-supervised monocular depth estimation. Proceedings of the IEEE/CVF international conference on computer vision. (2019)

[8]. Meinhardt, T., et al.: Trackformer: Multi-object tracking with transformers. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. (2022)

[9]. Iman T., Zadeh M., Moosa A., Amir T.: Cooperative adaptive cruise control system for electric vehicles through a predictive deep reinforcement learning approach. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering,238(7),1751-1773. (2024)

[10]. Coelingh, E., Andreas E., and Mattias B.: Collision warning with full auto brake and pedestrian detection-a practical example of automatic emergency braking. 13th International IEEE Conference on Intelligent Transportation Systems. IEEE (2010)

[11]. Alsuwian T., Saeed B., Amin A. Autonomous Vehicle with Emergency Braking Algorithm Based on Multi-Sensor Fusion and Super Twisting Speed Controller. Applied Sciences,12(17),8458-8458. (2022)

[12]. Li, Y., Norman H., Sharath P.: Intelligent headlight control using learning-based approaches. 2011 IEEE intelligent vehicles symposium (IV). IEEE, (2011)

[13]. Wei W., Shiwei C., Zhongjiao S., Yuchen W.: Neuroadaptive high-order fully-actuated system approach for roll autopilot with unknown uncertainties. Aerospace Science and Technology,155(P1),109567-109567. (2024)

[14]. Hengyang S., Meng L., Zhihao C., Yanjun H., Hong C.: Semantic Shapley-based counterfactual explanations for end-to-end autonomous driving. Engineering Applications of Artificial Intelligence,159(PB),111638-111638. (2025)

[15]. Rui Z., Ziguo C., Yuze F., Fei G., Yuzhuo M.: BETAV: A Unified BEV-Transformer and Bézier Optimization Framework for Jointly Optimized End-to-End Autonomous Driving.Sensors,25(11),3336-3336. (2025)
[16]. Qiang G., et, al.: Progress, challenges and trends on vision sensing technologies in automatic/intelligent robotic welding: State-of-the-art review. Robotics and Computer-Integrated Manufacturing,89,102767. (2024)
[17]. Parul P., Amit R., Pawan K.: Enhanced zebra optimization algorithm for reliability redundancy allocation and engineering optimization problems. Cluster Computing, 28(4), 267-267. (2025)
[18]. Bugaj M., et, al.: Deep Learning-based Sensing and Extended Reality Technologies, Visual Recognition and Geospatial Mapping Tools, and Virtual Simulation and Spatial Cognition Algorithms in Digital Twin Cities and Immersive 3D Environments.Geopolitics, History, and International Relations,15(1),31-45. (2023)
[19]. Juan L., Nadine M.: Opacity in car augmented reality head-up displays: users' preferences, visual attention, and situation awareness..Applied ergonomics,129,104610. (2025)
[20]. Dalong L., Lijuan X.: Robot path planning and obstacle avoidance algorithm based on visual perception. Neural Computing and Applications, (prepublish),1-18. (2025)