Semantic Segmentation and Reinforcement Learning in Autonomous Driving

Chuanjin Fan, Wenyao Liu, Jialu Xu and Yiyang Zhao

February 2023

Abstract

Autonomous driving is an integrated system that combines pattern recognition, environment perception, planning, decision-making, and intelligent control. Advances in machine learning have extensively promoted the development of autonomous driving technology. By performing high-quality semantic segmentation of road scenes, the safety of self-driving vehicles can be guaranteed. And the reinforcement learning method is suitable for the intelligent processing of decision-making and control of autonomous driving system in complex traffic scenes, which can help improve the comfort and safety of autonomous driving. In this article, firstly, a brief introduction to autonomous driving technology is given. Secondly, the semantic segmentation techniques are divided into traditional segmentation techniques, traditional segmentation techniques combined with deep learning and deep learning-based segmentation techniques, focusing on the deep learning-based semantic segmentation techniques, and describing them according to three different network training methods: intensely supervised, weakly supervised and unsupervised. Then the performance evaluation indexes of semantic segmentation of road scenes are summarized and compared on this basis to analyze the segmentation results of typical image semantic segmentation methods. Additionally, this paper gives a brief overview of autonomous driving technology, reinforcement learning techniques, and autonomous driving control architecture, as well as an explanation of the fundamental ideas behind and current state of research in reinforcement learning techniques. Then the research history and status of reinforcement learning methods in autonomous driving control are highlighted. Finally, the challenges of semantic segmentation and reinforcement learning applications in autonomous driving and the future development directions are presented.

1 Introduction

Autonomous driving is an integrated system that combines functions such as environment perception, decision planning, and intelligent control. It is an essential part of an intelligent transportation system, and is also a hot spot for research in the field of intelligent vehicles and a new driving force for the growth of automotive industry [111]. The control technology of self-driving vehicles is the critical link in the whole automatic driving system, and is also a key research area for scholars at home and abroad. The autonomous driving system generally adopts a hierarchical structure, in which the function of the control layer is to translate the commands from the decision and planning layer into the actions of each actuator and control each actuator to complete the corresponding actions, so as to follow the path and reasonably control the speed accurately.

The control of autonomous vehicles can be divided into transverse and longitudinal control. Most traditional transverse or longitudinal control methods require accurate mathematical analytical models and precise numerical solutions for the controlled system. However, models with higher accuracy are also generally more complex and have more parameters. Complex models also result in higher computational costs, making the solution difficult and often challenging to ensure real-time performance. With the rapid development of Internet +, big data, and artificial intelligence, researchers have started to develop intelligent vehicle decision-making

and control algorithms based on machine learning methods, opening up a different research line of thought from that of automotive engineering experts.

Most self-driving vehicle systems adopt a layered architecture [72], i.e., four layers of perception layer, decision layer (including motion planning), control layer, and vehicle wire control layer, as shown in Figure 1. The control layer of driverless vehicles, as the intermediate layer between the decision layer and the actuator, is also the most important and has been one of the critical areas of research on driverless technology. The driverless control technology mainly includes lateral and longitudinal control [78]. Lateral control is used to guide the vehicle along a global geometric path by performing appropriate steering motions. The goal of a path-following controller is to minimize the lateral distance between the car and the way and the deviation between the vehicle direction and the path direction, constraining the steering input to smooth the motion to maintain stability [90]. On the other hand, longitudinal control calculates the desired speed and acceleration based on the road shape and controls the throttle and braking system to achieve them while satisfying the vehicle's kinematic constraints, dynamics constraints, and safe distance.

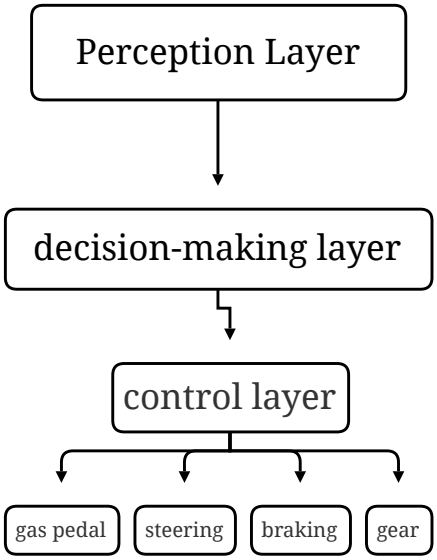

Figure 1: Block diagram of autonomous driving system.

Machine learning, which focuses on how computers acquire knowledge or optimize their skills through experience or exploration of the environment, is one of the fastest growing technology areas today. Image semantic segmentation is one of the most popular approaches based on machine learning that is increasingly being applied to autonomous driving systems. One of them is image semantic segmentation [44, 22], which is one of the core technologies for autonomous driving. Semantic segmentation for road scenes is to classify each pixel in the acquired road scene image into a corresponding category in order to realize the classification of road scene images at the pixel level. Among the technical components of autonomous driving, the processing of environmental information is a key part, which requires high-level road scene semantic segmentation and other related technologies to provide intelligent vehicles with important road condition information to ensure that autonomous vehicles can drive safely. Therefore, in the field of autonomous driving, the semantic segmentation of road scenes plays a critical role and is a hot spot for current research.

In autonomous driving, accuracy and real-time are fundamental indexes. But the accuracy in the actual semantic segmentation will be affected by different driving areas, firstly to overcome the dissimilarity of varying target objects and the similarity of similar target objects, and secondly to pay attention to the complexity of the scene in which the segmented objects are lo-

cated. Finally, external factors such as lighting, shooting conditions, shooting equipment, and shooting distance can also make the target object differ from the picture, which in turn affects the segmentation effect. All these factors greatly enhance the difficulty of image semantic segmentation, which in turn affects the realization of autonomous driving. In summary, semantic segmentation of road scenes is a key and challenging technology in autonomous driving.

Another primary type of machine learning is reinforcement learning (reinforcement learning, RL) [46, 7, 49, 95, 94, 56, 68, 89]. Unlike supervised learning, which is mainly applied to the perception layer of autonomous driving, reinforcement learning is more often applied to the decision and control layers. While traditional controllers generally utilize a priori models consisting of fixed parameters, which cannot anticipate all possible situations that the system must cope with when the robot is used in complex environments (e.g., driving), learning controllers use training information to learn their models step by step [71]. Machine learning can also be combined with traditional control methods, such as learning the cost function of predictive model control (MPC), allowing one to better predict vehicle disturbances and behavior [71]. Due to the high dimensionality, spatial continuity of states and actions, and nonlinearity of the autonomous driving control problem, deep learning is not good at decision making and control, although it has strong perception capabilities. Reinforcement learning, on the other hand, can learn complex control models by continuously exploring the environment. Therefore, deep reinforcement learning (DRL), which combines the two, can provide a new way of thinking for solving perception and decision problems of complex systems by complementing each other. DRL combines the perception capability of deep learning with the decision control capability of reinforcement learning, and can control directly based on the input pixel-level images (or radar data), which is closer to the human way of thinking. Combining deep learning and reinforcement learning allows for more solutions to the autonomous driving control problem.

2 Applications of Reinforcement Learning in Autonomous Driving

2.1 History of Image Semantic Segmentation

From the perspective of evolution, the development process of image semantic segmentation technology is shown in Figure 2.

1. Traditional semantic segmentation stage. Due to the limited computing power at that time, semantic segmentation algorithms in this period mainly relied on image texture, color, and other simple surface features for image segmentation. The segmentation results are relatively rough, with low precision and no relevant annotations.

2. Semantic segmentation stage combining traditional methods with deep learning. This method is similar to the target detection method. In the segmentation process, the traditional method is used to initially process the image to form an image-level effect, and then the feature classifier in the convolutional neural network is used for semantic segmentation, and finally, the image segmentation effect is formed. This method inherits the traditional segmentation method, which has certain deficiencies and defects, and its accuracy is relatively low.

3. Semantic segmentation stage based on deep learning. In the process of autonomous learning and classification of robust features, deep learning technology has shown incomparable advantages and characteristics, and has relatively powerful capabilities. Presently, the semantic segmentation method based on deep learning technology has been popularized and promoted, and has achieved better results than the previous two methods.

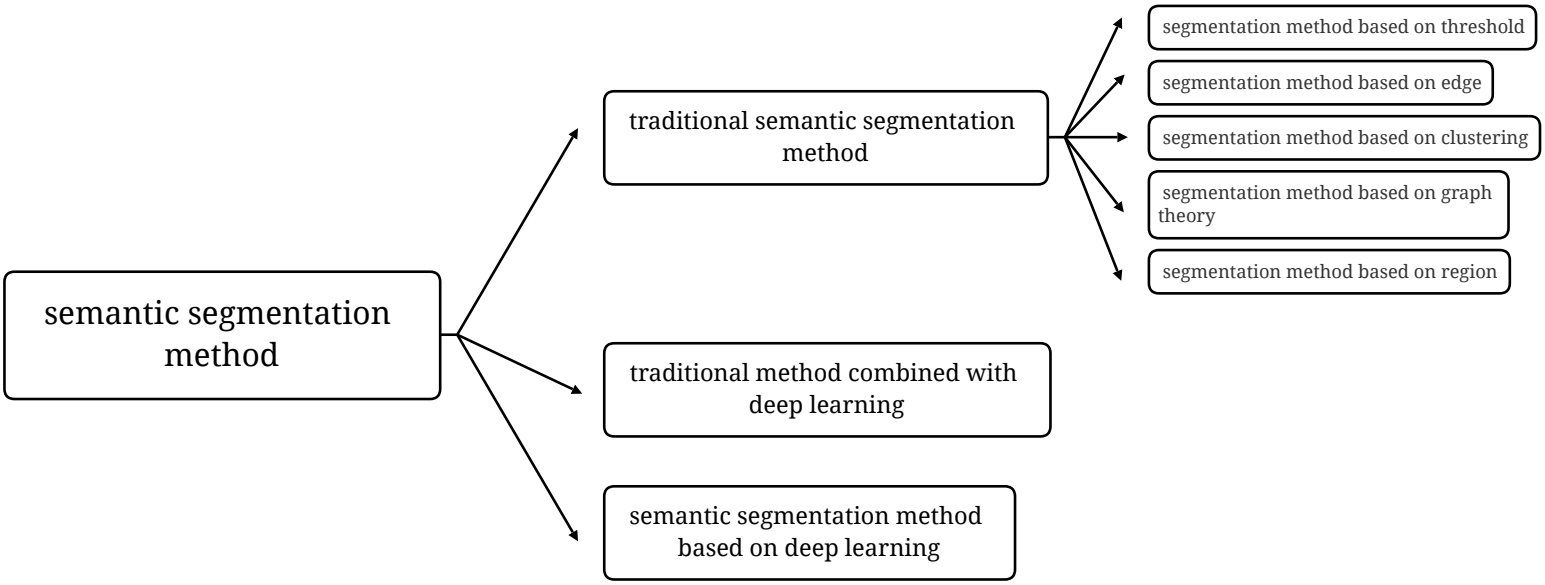

Figure 2: Development history of image semantic segmentation.

2.1.1 Traditional Image Semantic Segmentation Algorithm

Before deep learning was widely used in computer vision, to separate the target from the background, traditional image semantic segmentation methods used multiple features, such as color, grayscale, texture, geometry, etc., to segment the image into multiple independent regions. Traditional image semantic segmentation methods include threshold-based segmentation, edge-based segmentation, clustering-based segmentation, graph-theoretic-based segmentation, and region-based segmentation. Among them, the most commonly used is graph theory-based segmentation, and the "Normalized cut" and "Grab cut" algorithms are the most common techniques based on graph theory-based segmentation [47], which will be explained in the following.

2.1.1.1 Normalized Cut Image Segmentation Algorithm

Researchers proposed a new image segmentation method at the beginning of the 21st century, which takes the picture as a unit and uses it as a basis for segmenting the image. This semantic segmentation method is defined as the Normalized cut algorithm [122]. The idea of its implementation of image segmentation is that it takes the picture as a unit, calculates the weight map, and then segments it into some regions with the same features. The min-cut algorithm, as an essential method, can ideally segment the whole part to be segmented into two parts. However, the min-cut segmentation also has defects such as missing edge corner elements, which makes the final result biased.

2.1.1.2 Grab Cut Image Segmentation Algorithm

The Grab cut [130] algorithm is also based on the theoretical basis of image segmentation, using the hybrid Gaussian model and the Giles energy equation to achieve the modeling effect based on the color space, using an iterative approach to find the optimal solution of the equation, and finally obtaining the optimal parametric solution of the Gaussian model. The proposed algorithm significantly broadens the image segmentation field and realizes the segmentation of color images.

Although Grab cut has improved its segmentation performance, its convenience is poor. Many systems cannot use this technology, and the stability of the operator needs to be con-

sidered. Liu Lei et al. [61] introduced the high-order potential energy term into Grab cut, so that it can better describe the details and associated information of pixels, thus improving the segmentation accuracy of the model.

2.1.1.3 The Latest Traditional Semantic Segmentation Algorithm

In 2011, Arbeláez et al. [5] combined both GPB and UCM methods for detection and proposed a new detection algorithm, namely the contour detection method. The algorithm first uses the GPB method to reasonably measure the actual probability of any pixel edge, and then transforms the different closed regions formed by these measurements using the UCM method to form a hierarchical tree-like structure. With continuous research, in 2016, Zhang et al. [124] proposed the random decision forest segmentation method, which is different from the contour method in that the detection method mainly uses different decision trees to combine to form a classifier. In 2017, Pont-Tuest et al. [76] combined the above two detection methods and proposed a new detection method, namely the M CG algorithm. This method first uses the GPB-UCM method to segment the image contours to obtain different block structures, and then uses the classifier formed by the random method for further segmentation. The method achieves an optimal upgrade of the traditional model and traditional methods.

2.1.2 Image Semantic Segmentation by Combining Traditional Methods with Deep Learning Methods

The study found that the prominent feature of the traditional method is to focus on using surface features and external structure features to complete image segmentation, and then manually label [24] it. The advancement of modern technology has promoted the continuous development of deep learning technology and the transformation of semantic segmentation technology. The researchers introduced the deep learning algorithm model into the research of semantic segmentation. Firstly, the traditional method was used for initial segmentation to obtain the target area, and then the convolutional neural network was used to deeply study the characteristics of the target, and a scientific and reasonable classifier was formed, and finally realized segmentation of the target area and complete automatic labeling.

This segmentation method is based on the principle of convolutional neural network, trains the convolutional network, obtains the original contour segmentation area with the help of segmentation tree, super pixel and other technologies, supervises the convolutional network in real time, and performs deep learning, using multiple processing processes such as superpixel segmentation and parameter-free multi-level parsing to obtain the final result. At the same time, researchers used image and depth map technologies in the process of segmenting indoor scenes. The basic process is relatively simple: Firstly, the filter features and convolution features of the image are reasonably extracted, and a scientific classifier is constructed by fusing feature maps of different scales, structures, and levels.

After the RGB image is segmented by superpixels [101], this classifier can be used for further classification. It cannot be ignored that there are many unstable factors in the superpixel segmentation method, and it is easy to produce narrow and wrong classification results. In addition, the superpixel segmentation method has specific difficulties and limitations in the processing of weak boundary image regions.

2.2 Semantic Segmentation Methods Based on Deep Learning

In recent years, with the rapid development of deep learning, the field of semantic segmentation research has also made breakthrough progress. Compared with the traditional semantic segmentation methods, the semantic segmentation methods based on deep learning can obtain more and more advanced semantic information to express the information in images. Since the introduction of deep learning in the field of semantic segmentation, segmentation accuracy has been a hot topic of research as an essential index to measure the effectiveness of semantic segmentation. The model of fully convolutional neural network [86] has initially achieved

semantic segmentation at the pixel level, which has led to a leap forward in segmentation accuracy in this field. Many FCN-based semantic segmentation methods have emerged one after another. This section will introduce the deep learning-based semantic segmentation methods in detail, which are classified into strongly supervised-based semantic segmentation methods, weakly supervised semantic segmentation methods and unsupervised semantic segmentation methods based on the different training methods of the network. Their main advantages and disadvantages are shown in Table 1.

Table 1: Advantage and disadvantage comparison among strongly supervisedweakly supervised and unsupervised semantic segmentation methods.

Type	Advantage	Disadvantage
Strongly supervised	High segmentation accuracy based on densely annotated datasets	Being excessively dependent on dataset marked by dense set, inability to migrate, and poor segmentation accuracy for unknown scenes
Weakly supervised	Only image-level annotated dataset required to complete training	Large number of datasets needed, long time, and lower accuracy than that of strong supervision
Unsupervised	Being independent on manual intensive annotation dataset and strongly adaptable to unknown environment	Being difficult to adapt and no high segmentation accuracy at present

2.2.1 Intensely supervised semantic segmentation-based approach

Manual annotation of samples can reflect a large amount of valid local data and detailed features, which can significantly improve the training effect and segmentation accuracy to a certain extent. The intensely supervised learning model is the most widely used segmentation model, and it is also the algorithmic model with the best effect and the largest influence. The structure of FCN network is shown in Figure 3.

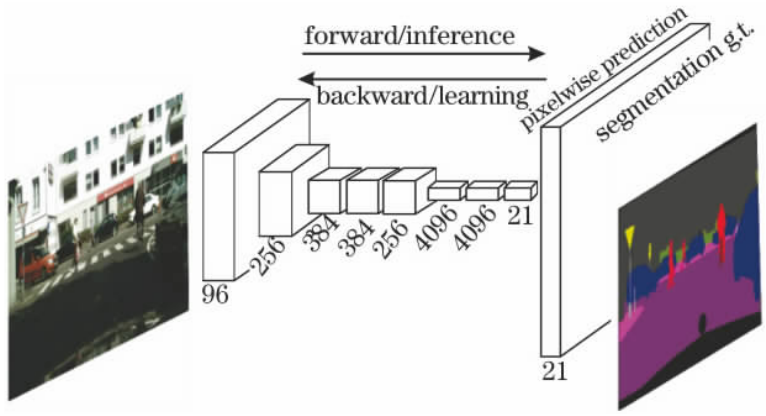

Figure 3: Structural diagram of fully convolutional network [24].

The convolutional and sampling layers in FCN involve various types of up and down, forward and backward, respectively, and these structure types remain unchanged during arbitrary translations in space. Restoring the original resolution size of an image is a common application scenario for full convolutional networks, and the processing often utilizes a form of deconvolution. In the FCN structure, a certain number of fixed-size convolutional layers act as fully-connected layers in a conventional convolutional network, and this structure allows the neural network to slide freely and closely in the image to improve the sliding flexibility of the

convolutional neural network, resulting in a prediction map that contains dense output images. However, FCN still retains the pooling layer in the convolutional neural network (CNN), which increases the perceptual field of the convolutional neural network, but the continuous downsampling causes the loss of details and dramatically affects the segmentation results. Meanwhile, higher sampling rate leads to loss of feature map size and spatial information. To address these problems, based on FCN, researchers have proposed a series of new methods, which we classify into six categories, i.e., segmentation based on expanded field of sensation, segmentation based on probabilistic graph model, segmentation based on feature fusion, segmentation based on encoder-decoder, segmentation based on recurrent neural network (RNN), and segmentation based on generative adversarial network (GAN). network (GAN) based segmentation methods, as shown in Figure 4, where ASPP is a null space convolutional pooling pyramid.

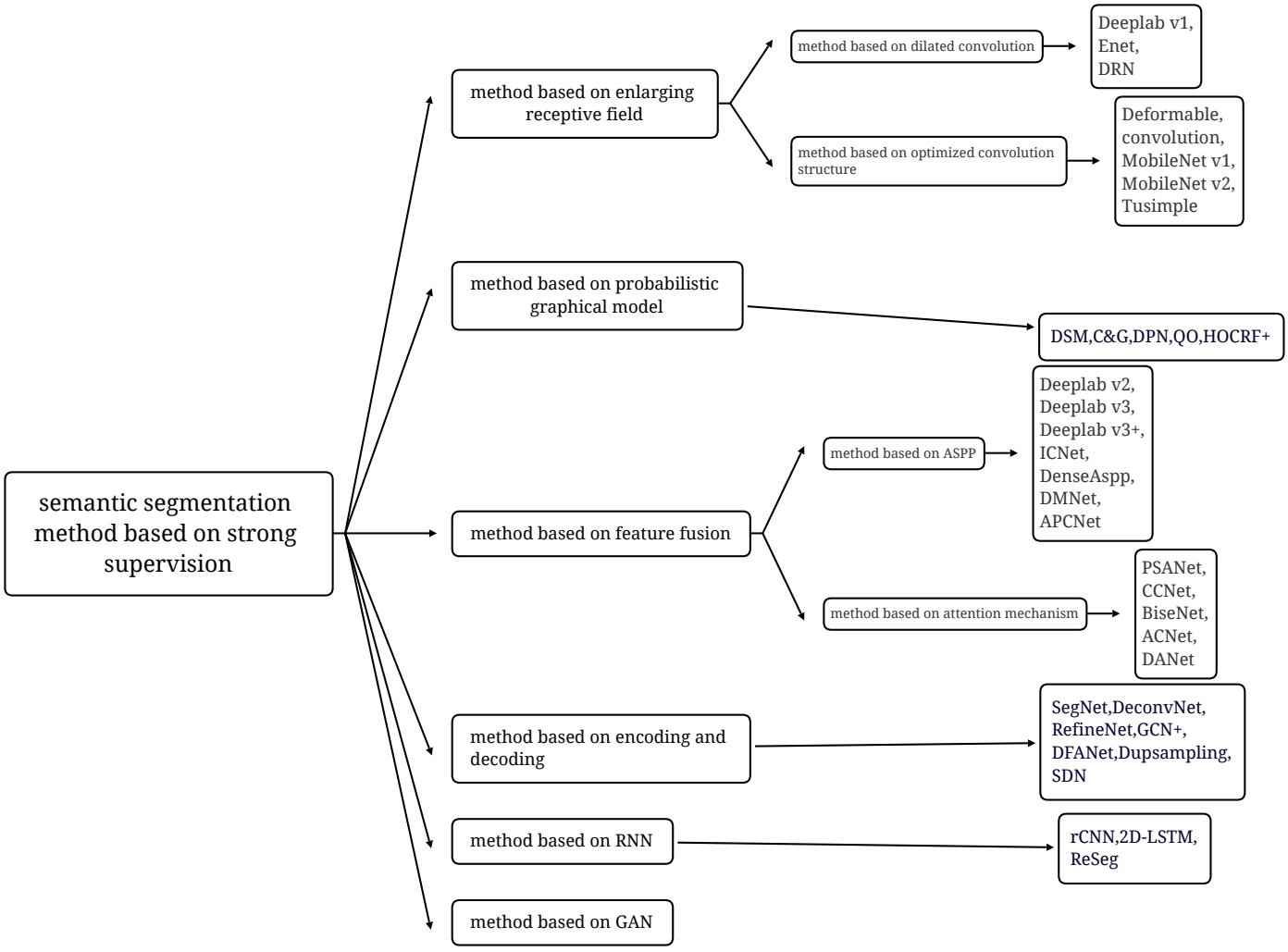

Figure 4: Semantic segmentation method based on strong supervision.

2.2.1.1 Based on an expanded field of sensation approach

Null convolution [120] can be used as a convolution layer for dense prediction, also known as dilated convolution. Null convolution is based on the guaranteed image resolution property to enhance the perceptual field without reducing the coverage. The impact of this method on the resolution of convolutional neural networks is focused on the field of feature response computation. The extended convolution pattern echoes the null convolution. As shown in Figure 5, the 3×3 convolution is chosen, and the perceptual field is compared with the expansion

coefficient of 1, 2 and 4. It is easy to see that the perceptual field is positively correlated with the expansion coefficient, and the size of the perceptual field is 5 times larger when the expansion coefficient is 4 than when the expansion coefficient is 1. Expanded convolution can expand the stacking effect of convolution and increase the size of the receptive field. The null convolution focuses on the improvement of resolution and computational responsiveness, reduces the dependence of the computational process on parameters, and enables the expansion of the convolution kernel field with fewer parameters or factors, and facilitates the acquisition of front and back contents.

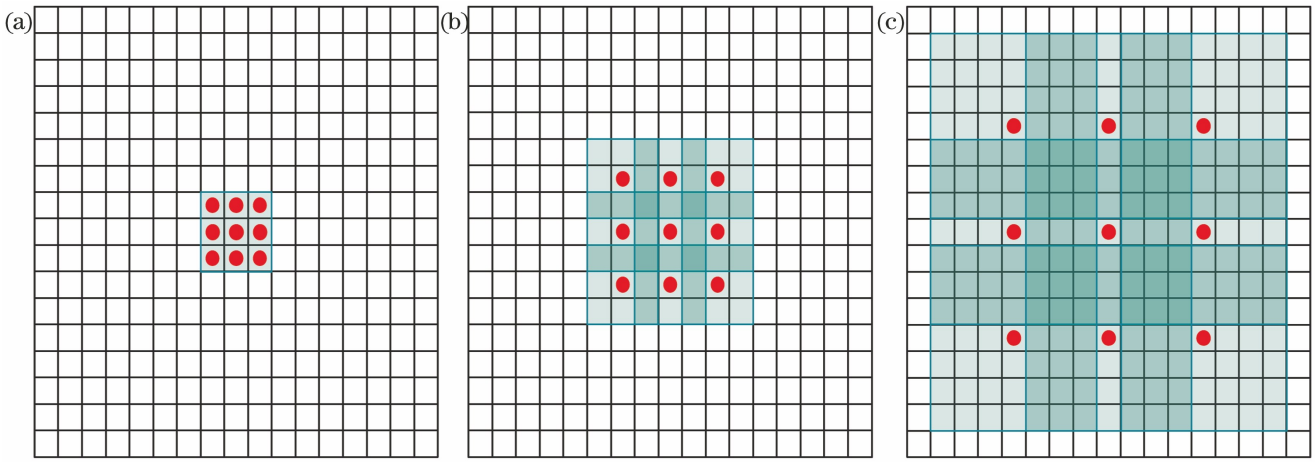

Figure 5: Schematic of expansion convolution [120]. (a)Ordinary convolution;(b) expansion convolution with expansion rate of 2; (c) expansion convolution with expansion rate of 4.

DeepLabv1 network model was proposed by Chen et al. [16], DeepLabv1 innovatively applied the null convolution to the VGG16 network, by converting the fully connected layers of VGG16 into convolutional layers and adjusting all convolutional layers after the fourth and fifth pooling layers of the VGG model to null convolution with different expansion rates respectively, the sensory field was restored to the original image size, In 2016, Paszke et al. [74] proposed a real-time segmentation model, namely ENet. this segmentation model mainly uses bottleneck module thinking to perform serial operations on multiple null convolutions to adjust the actual area size of the perceptual field, which effectively solves the problem of continuously decreasing feature resolution. In 2017, Yu et al. [121] proposed a DRN network model. It was shown that this model is based on ResNet network and uses the null convolution to replace the normal convolution to maintain the actual resolution of the original image and the effective perceptual field area of the original network. The model uses two different expansion rates of the null convolution to replace the end convolution layer of ResNet in order to continuously enhance the spatially effective information. In order to avoid the tessellation effect caused by the recycling of the null convolution, the residuals and the maximum pooling layer are removed, and finally the pixel output operation is realized by the entire convolution method.

When using CNN method for semantic segmentation of images, the pooling process will continuously increase the effective range of the perceptual field and fuse the background information. However, it should not be neglected that this process also causes a continuous decrease in image resolution, resulting in the loss of some spatial information. A reasonable solution to this problem is to optimize the convolution structure and use the optimized convolution structure for convolution and pooling operations.

Dilated convolution can quickly and effectively acquire depth features of images, expand the field of perception, and preserve the location information of specific pixels. However, it is easy to form certain spatial loopholes when performing convolutional operations, and problems such

as data loss and message loss can occur. In the paper [25], the researchers used hybrid dilated convolution (HDC) to replace the BI algorithm with a dense upsampling algorithm, and the HDC approach consists of a series of dilated convolution modules that can further expand the receptive field while maintaining the local information related features. The above methods can effectively expand the receptive field, but due to the relatively fixed shape of the convolution kernel, these methods are relatively weak in simulating geometric transformations, poor in adapting to graphical changes, and poor in extracting features of irregularly shaped objects. In the paper [21], the researchers introduced a learnable offset by using a sampling operation with a certain offset to finally adjust the shape of the convolution kernel to make it variable and proposed the basic concept of deformable convolution in convolution processing. This convolution mode can effectively expand the perceptual field, increase the image area, improve the self-adaptability of semantic segmentation to graphic transformation, and improve the precision and accuracy of segmentation. In the process of depthwise separable convolution, less computation reduces the performance consumption. Segmentation models for mobile devices usually have two modes: point-by-point convolution and depth convolution. Among them, point-by-point convolution mainly uses 1×1 convolution, while deep convolution uses different convolution kernels in each channel. The actual segmentation effect of deep convolution is not good, and often only the basic features of the low-dimensional space can be extracted. To solve this problem, the literature [29] continuously increases the convolutional dimension before the start of deep convolution, so that deep convolution can operate in high-dimensional space.

2.2.1.2 Segmentation method based on probabilistic graphical model

The probabilistic graph model is used for post-processing of CNN to efficiently optimize object boundaries in a structured predictive manner, capturing image contextual information and enabling a balance between the utilization of local and global features.

By combining two models, namely, random field of conditions and CNN, we can reasonably predict the relevant information in the process of information transmission and effectively reduce the amount of redundant calculations, thus achieving an increase in computing efficiency. This method can obtain relatively rich data information and improve the operation efficiency. However, in the process of structure prediction, this method can only input images into monadic or paired items, and it is difficult to achieve structure prediction in medium and high items, and the segmentation accuracy is relatively low. Therefore, Arnab et al. [6] embedded two different forms of higher-order potential terms into CNN to train the depth, and the segmentation quality was improved. In addition, to optimize the segmentation model and improve the segmentation quality, Vemulapalli et al. [98] used Gaussian condition random field to optimize the segmentation results. Some scholars have fused the two models, FCN and CRF, and proposed two different segmentation models, namely SegModel [87] network model and DFCNDCRF [43] network model.

2.2.1.3 Feature fusion-based segmentation method

The methods in Section 2.2.1.1 all utilize serial operations with different expansion rates of hole convolution to continuously increase the perceptual field and to extract semantic features in depth. However, cycling repeatedly using hole convolution will inevitably produce a tessellation effect and also cause some features to be lost, and occupy a large amount of runtime space and consume a lot of memory. the probabilistic model graph-based methods in Section 2.2.1.2 also have problems in terms of excessive computation, long training time, and large memory consumption. Feature fusion is the summation or splicing fusion of the extracted feature maps. In the feature extraction stage, the semantic information of the feature map is enriched by fusing the multi-scale feature information. In the feature utilization stage, the global effective information is utilized to improve the segmentation accuracy by fusing the features at different levels. The feature fusion-based method captures the implicit contextual information in the image by fusing features of different layers and regions, which can effectively improve the segmentation rate and segmentation effectiveness, and also significantly reduce the

operation consumption.

Lin et al. [59] proposed a feature pyramid network. In the structure setting process, this network enriches the semantic information of features at each scale by adjusting the connection form of higher-level features and lower-level features. DeepLabv2 [17] introduced the band-hole convolution and pyramid pooling based on DeepLabv1, and replaced the VGG-16 network with the ResNet network. Context capture was achieved by sampling using different scales, and the process was based on the convolution of voids with multiple sampling rates of the input. The classification effect is enhanced based on the mining of convolutional image features and the extraction of content image features, and the above processing is based on the premise that the resolution of the feature map is not affected. DeepLabv3 [103, 18] improved the ASPP structure by introducing the Resnetblock module and aiming at extracting salient features, and proposed the null convolution model which realized the pooling effect of modules and spatial elements. Combining the above two methods, Yang et al. [116] conducted an in-depth study and proposed the DenseASPP network. He et al. [34] found that although ASPP can handle the change of graphic scale, it is difficult to achieve a new balance in the change of scale and expansion rate. Therefore, this group proposed a dynamic multiscale network with the help of which dynamic convolutional semantic perception and estimation were achieved. To improve the ability of the network to aggregate contextual information, Zhao et al. [128] proposed a pyramidal scene resolution network, and then Zhao et al. [127] proposed an image cascade network with real-time segmentation features from the perspective of compressed PSPNet. Wu et al. [109] proposed a new joint pyramid sampling model in 2019 for the alternative network of expanded convolution. This method can effectively acquire samples with high-resolution mapping features, which can significantly reduce the accuracy loss and memory consumption.

The traditional fixed convolutional structure mainly obtains information with the help of segmentation framework of FCN, but only short distance information can be obtained. In order to acquire long-distance contextual information, researchers have proposed methods such as dilation convolution. However, this class of methods does not form dense information during the process of acquiring information. Therefore, in the process of semantic segmentation, Zhao et al. [129] reasonably introduced the attention mechanism and proposed the PSANet network model to aggregate information at different locations by pre-drawing the attention graph [123]. This method realizes the calculation of the relationship between each pixel with the help of a huge attention map, which is relatively complicated to calculate during the operation and has a relatively high memory usage. A series of innovative network modules have been proposed to improve the quality and efficiency of the partitioning. The following are three of the more popular modules: the CCNet [26] algorithm module can perform high-end segmentation by inserting a fully convolved arbitrary neural network; the BiSeNet [118] module can perform the global information integration operation without any sampling process, which effectively reduces the operation cost and increases the computation speed; the ACNet [64] module combines the self-attentive auxiliary algorithm mode and the parallel algorithm mode to improve the computation speed. The ACNet [64] module combines the self-attentive auxiliary algorithm model and parallel branching architecture to perform the balancing operation on the depth image features. In recent years, the self-attentive mechanism has been increasingly effective in the practice of semantic segmentation. Researchers have incorporated this mechanism into the basic process of semantic segmentation. Innovative approaches such as dual attention networks [112] have been proposed to effectively reduce the spatio-temporal complexity. In the process of performing semantic segmentation, the attention mechanism is reasonably introduced to learn the relevant information, and by adjusting and optimizing the attention mechanism, a new decimal cross module and a self-attention module are formed as a way to obtain global information and perceive the information at each level, which makes the capture of information and internal features easier.

2.2.1.4 Code-decoder based methods

The basic idea of the method is that the encoder extracts the main feature information of

the image through a series of convolution-pooling operations, and then gradually restores the spatial dimension of the image through the upsampling-transposition convolution structure of the decoder. Relying on the basic method of encoder-decoder, feature processing and upsampling operations can be performed on low-resolution graphics, which can effectively solve the problem of decreasing resolution and highly restore the spatio-temporal information of pixels and dimensional data of graphics.

SegNet [50] and U-net [80] are two typical coder-decoder structures for image semantic segmentation, and the structure of SegNet network is shown in Figure 6. SegNet uses VGG-16 network, and the dense feature map is output using this network, and the recovery of the dense map is achieved by The recovery of the dense graph is achieved by the convolution calculation of the sparse image [110]. Subsequently, researchers also proposed a BayesianSegNet network based on the SegNet network, which solved the problem that the prior probability could not give the confidence of the classification result [114] by introducing Bayesian network and Gaussian process, and improved the learning ability of the network.Noh et al. [70] proposed a fully symmetric DeconvNet network based on FCN, which based on the complementarity of FCN and inverse convolution network, using FCN to extract the overall shape and inverse convolution network to extract the fine boundaries, which can cope with objects of different scale sizes, but also can better identify the details of objects and improve the efficiency of segmentation.

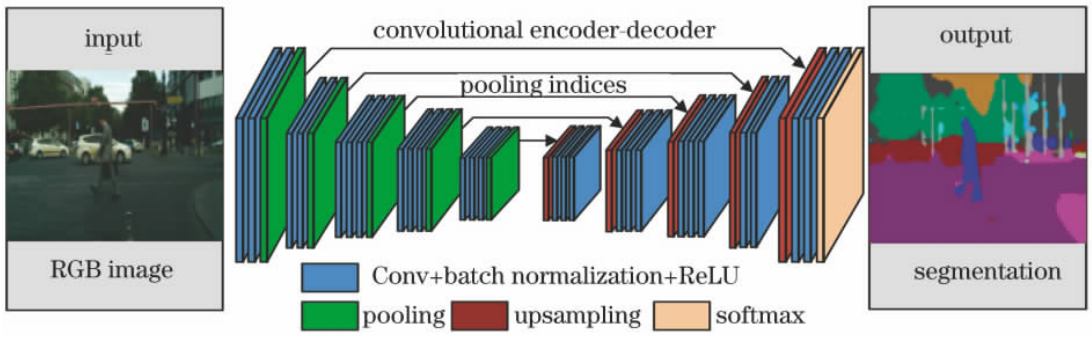

Figure 6: Structural diagram of SegNet network [50].

The researchers proposed a relatively homogeneous architecture of U-Net network, which has different and complementary encoding and decoding structures to refine the details and restore the effect. Although the U-Net semantic segmentation model achieves good segmentation results, it can only handle 2D images. Researchers proposed the V-Net network [77], which is a 3D symmetric semantic segmentation model combining 3D volume, full convolution and neural networks, solving the problem of insufficient training labeled datasets and providing computational advantages compared with other segmentation models. To solve the problem of data information loss during semantic segmentation, Lin et al. [58] proposed the RefineNet network in 2017. This network uses a chained residual connectivity model, which can recover and effectively fuse the missing information in the segmentation process, and then form a relatively clear predicted image. Taken together, the method can effectively fuse high-level features and low-level features, and use advanced thinking such as constant mapping and residual connectivity to obtain good training results. The model can obtain relatively superior segmentation results in different scene environments. Based on the coder and decoder structures, the researchers have studied the segmentation method in depth and presented a series of fruitful research results. For example, the method of early sampling is applied to reduce the application of decoders, thus forming a simplified version of the ENet structure and achieving the basic goals of reducing network redundancy and decreasing the number of parameters.

In order to reasonably improve the structure of encoding-decoding models, researchers have made several improvements: 1) continuously improve the actual speed of semantic seg-

mentation, for example, using models such as ENet and LEDNet, which strongly promote the realization of real-time segmentation targets; 2) effectively fuse multiple resolution features, for example, the DUpsamling [96] module, which can be used to learn sampling; 3) continuously extending the effective range of perceptual fields to improve the segmentation accuracy, e.g., GCN module; 4) capturing multi-scale and multi-level information to ensure effective recovery of target information, e.g., SDN module.

2.2.1.5 Recurrent neural network based approach

Another widely used and effective operational model in deep learning is the cyclic nerve network [117] model. The main advantage of this model is that in addition to learning the current information, it can also integrate the sequence information and construct a global modeling algorithm to improve the combined utilization of graphical information. Based on this idea, Visin et al. [99] proposed the ReSeg network by combining the local information obtained by CNN approach, the global features obtained by RNN approach, and also by drawing on the image segmentation model. Influenced by the image segmentation network ReNet, Li et al. [53] proposed the LSTM-CF network model, which can effectively use the two basic features of depth image and luminosity. However, since this model only uses the LSTM approach, the flexibility and diversity of image processing are relatively poor. Therefore, Liang et al. [54] proposed a Graph-LSTM network, which sets arbitrary superpixels as the reference valid nodes and constructs an adaptive image for the image. After that, the team adjusted the operation mode of the network and optimized the structure from the perspective of coding hierarchy. Considering the lack of effective interconnection between the FCN and the fully connected CRF model, Zheng et al. [131] proposed the CRFasRNN network model, which incorporates the relevant learning and reasoning process of CRF into the RNN operation.

The RNN is capable of retaining relevant information, recursively processing historical data and historical memories, extracting sequential information from images, and reasonably modeling the semantic relationships of images to obtain relevant data information. At the same time, the network model can be deeply integrated with the convolutional layer, and the spatial features of the convolutional layer can be effectively extracted by integrating into the nerve network structure, and the deep extraction of pixel features can also be realized.

2.2.1.6 Generative adversarial network-based approach

Similar to the pyramid network structure, the generative adversarial network [102] can replace CRF to a certain extent to complete the acquisition of image information features, which can achieve continuous spatial expansion and ensure the consistency of spatial features without additional training time and training difficulty.

In order to continuously reduce the inconsistency between labels and images, Luc et al. [63] introduced the GAN technique for the first time in 2016, and used discriminators to perform recognition operations on labels and segmentation domains in the process of semantic segmentation. In the medical field, since the U-Net network cannot well solve the realistic problem of inconsistent and unbalanced pixel categories, Xue et al. [113] proposed an adversarial network model based on multi-scale and multi-level functions, and used discriminators to deeply learn the local attributes and global structural features of segmented objects during graph segmentation, so as to obtain the different pixel-to-pixel In addition, the GAN model also has the ability to recognize the spatial relationships between pixels. In addition, the GAN model has the ability to identify true and false data and continuously generate new data. The GAN model has some instability in the process of application, especially for large data images, and the interpretability and extensibility of the method are not sufficient.

2.2.2 A weakly supervised semantic segmentation based approach

Strongly supervised semantic segmentation-based methods require a large number of pixel-level annotated training samples. Since it takes a lot of time and effort to obtain pixel-level

semantic annotation samples, and there are limitations in training by pixel-level annotation samples, weakly supervised semantic segmentation-based methods have started to emerge. In this paper, based on different types of supervised information, weakly supervised image semantic segmentation-based methods are classified into six categories: methods based on bounding box-level annotation, methods based on scribble-level annotation, methods based on point-level annotation, methods based on image-level annotation, methods based on hybrid annotation, and methods based on additional data sources.

2.2.2.1 Bounding box level annotation based approach

The method based on bounding box annotation uses a rectangular region including the whole object as a training sample and provides annotation information. Although this annotation method is one of the more complex among many annotation methods, it contains more semantic information, is less costly, and has better segmentation performance.

Dai et al. [20] proposed a BoxSup network model by using candidate regions with the help of FCN network. The model uses the bounding box labeled image as the training sample, and the MCG algorithm is chosen to calculate the original candidate region, which is then entered into the FCN network as "supervised information" for further optimization and upgrading. Then, the valid range of the candidate region is predicted, and the region is repeatedly optimized and upgraded until the result converges to a reasonable range. In the face of classification problems, DeepCut [79] mainly performs image segmentation by iterative iterative operations to continuously improve the segmentation accuracy and image precision.

In the traditional weakly supervised learning process, a simple iterative approach is commonly used for model training, and the final results often differ significantly from the actual labels.Song et al. [91], in their study of image segmentation methods, reasonably used a bounding box-driven classification region masking model to perform deletion operations on irrelevant regions, thereby obtaining pixel-level segmentation regions and fill rates, and subsequently applied a fill-rate-guided adaptive The fill-rate-induced adaptive loss algorithm model is then applied to correct and delete the erroneous pixels in the proposal that have completed labeling. The model algorithm mainly relies on the bounding box supervision algorithm to annotate and segment the image data, which can minimize the undesirable effects caused by erroneous annotation.

2.2.2.2 Graffiti-level annotation-based approach

The method sets the training samples reasonably in the annotation process, and the segmentation method taken is relatively simple. The samples are mainly graffiti-level images, which are relatively less difficult to obtain and effectively reduce the task of manual annotation.

In the literature [10], an annotation method using randomly scribbled points as supervised information was proposed. The method uses pixel points to annotate the image and sets the scribble points as supervised information in the practical operation, which effectively combines the advantages of supervised information, CNN network model as a function, and obtains good segmentation results. The literature [57] proposed the ScribbleSup model algorithm. The model takes some images containing scribble lines or scribble points as samples and unfolds the annotation in a scribble manner. The algorithm can be roughly divided into two stages: the first stage is the automatic tagging stage, which mainly forms pixel blocks of different shapes based on the graffiti lines, and then uses the pixel blocks as the basic nodes for automatic modeling, and finally performs the tagging process for all images; the second stage is the image training stage, which mainly trains the model for the images already formed in the first stage, and finally obtains reasonable segmentation results.

2.2.2.3 Point-level annotation-based approach

In essence, the instance point labeling method is a weak labeling method, which is mainly implemented by providing location information and identifying the center location. Compared

with other algorithms, the point-level supervision is better and the final result is more superior with the same budget premise.

In order to obtain good segmentation effect, researchers effectively integrate the advantageous features of point-level supervision and loss function to further enhance the effect supervision of semantic segmentation. Considering that the segmentation object contains four extreme points, Maninis et al. [66] proposed a CNN framework that can achieve semi-automatic segmentation based on this, namely DeepExtreme Cut

2.2.2.4 Image level annotation based approach

In contrast, the image-level annotation has multiple advantages and features, the annotation process is relatively simple, does not require the use of pixel annotation, sample acquisition is relatively easy, and the overall workload is relatively small. Therefore, this method also gradually becomes the mainstream method in the process of weakly supervised learning. In the research process, the researcher scientifically builds the association structure between graphic labels and pixels by reasonably introducing a multi-instance learning model, and uses algorithms such as superpixels to perform smoothing operations on various types of labels. The researcher uses the expectation maximization approach to reasonably predict and evaluate the pixel-level labels, and uses these labels as training samples to update the data model to maximize the expectation-maximization process. The researchers performed decomposition operations on the relevant feature maps to form the initial multi-channel features. Different channels have different local features, and the basic feature map of the multichannel is formed after the pooling operation, and the feature label information is subsequently learned for this map.

Compared with pixel-level annotation, the method of image-level annotation is somewhat simple and crude, and it is difficult to achieve good and expected segmentation results. Kolesnikov et al. [48] proposed the SEC algorithm. Influenced and inspired by this algorithm, Huang et al. [41] used the SRG region growth method to supervise the seed region to obtain relevant information and finally form a scientifically reasonable pixel label. Influenced and inspired by the null convolution, Wang et al. [104] applied the MDC algorithm model to the field of image denoising. In the case of missing external data or incomplete supervisory information loss, Ahn et al. [3] obtained accurate segmentation labels by effectively using AffinityNet network as a way to compensate for the missing relevant information. During the study, Zhou et al. [132] performed an image-level labeling of the supervised information and continuously improved the efficiency of instance segmentation using response peaks. In contrast, the segmentation method has a simple operation process, low sample acquisition cost, and the segmentation can be achieved by categorical labeling, which gradually improves the semantic segmentation quality and the actual effect of point-by-point localization.

In addition, Wei et al. [108] used saliency as the basic feature for effective extraction of additional knowledge information and proposed an SCT model algorithm. This method detected the regions with saliency features from bottom up, obtained the relationship between the region map and label information, and then gradually inferred the segmentation mask of the image, and used this as supervised information to start the learning training.

2.2.2.5 Hybrid annotation based approach

In summary, the above methods have significant advantages in terms of cost and time reduction, and can significantly reduce the actual demand for data training. However, it cannot be ignored that there are certain limitations of weak annotation methods, and one type of annotated data alone does not achieve good segmentation results. In the process of labeling, the segmentation effect can be improved if other types of data can be fused to achieve complementary advantages.

Semi-supervised learning-based segmentation methods typically use two types of labeled images, relatively few of which are pixel-level and relatively many of which are weakly labeled.

Researchers proposed the randomized gradient descent [82] algorithm model, and by operating on a combination of the two types of images, the superior performance of a single type of image is obtained that is not comparable. Hong et al. [37] proposed a semi-supervised segmentation framework model for DecoupledNet. The model operates differently on segmentation and classification items, where the classification network mainly uses image-level data in the model learning process, and then optimizes and upgrades the segmentation network using training examples. The method is relatively scalable because there is no repetitive cycle of operations.

2.2.2.6 Methods based on additional data sources

The graffiti and point-level annotations described above are generally more difficult to obtain directly and require human interaction. Compared to pixel annotations, this type of annotation information is less difficult to obtain, but the main purpose of implementing weakly supervised learning is to minimize human interactions. Therefore, researchers usually introduce some additional data and use stronger supervisory information to avoid the use of manual annotation.

Compared with single images, video information is relatively less difficult to obtain, and videos are now more commonly distributed. The researcher uses class tags as keywords and web libraries as search sources to obtain relevant video information by applying fully automated search. At the same time, the rational use of classifiers to optimize the relevant video intervals can obtain better retrieval results. In addition, the researcher introduces an attention mechanism in the new codec structure, which can migrate the irrelevant knowledge to the weakly supervised segmentation operation process.

2.2.3 Unsupervised semantic segmentation based approach

Numerous studies have shown that if a nerve network has a large amount of training data, the network tends to have relatively good operational properties. In practice, well-trained networks usually do not perform well if a certain size of data set is set. An effective solution is to adopt a relatively intensive manual labeling approach to repeatedly train the network. Another solution is to synthesize the data using a combination of computers with automatic semantic annotation and then iterative data training. In this process, the cyclic data synthesis training can degrade the performance of data usage and reduce the operational effect to some extent. On the whole, the best solution is to introduce unsupervised application methods, construct reasonable labeling regions, and continuously reduce the error of the labeled data.

In general, the basic process of adaptive training in an unsupervised manner is to construct reasonable spanning domains with the help of DomainShift minimization. Researchers proposed the DC [23] method, which effectively uses a binary domain classifier to achieve a uniform layout of labels. Since then, after intensive research, Tzeng et al. [97] further proposed a new adversarial discriminative domain adaptation method to continuously optimize the relevant model with the help of adversarial training model. To solve the problem of cross-domain segmentation, Hoffman et al. [36] proposed the FCNWild method. Zhang et al. [125] proposed the FCAN adaptive network model, which effectively combines the dual adaptive networks of image domain and feature domain, and improves the quality of semantic segmentation by using the merged images.

Figure 7 provides a summary of the weakly supervised and unsupervised semantic segmentation methods based on classification are summarized.

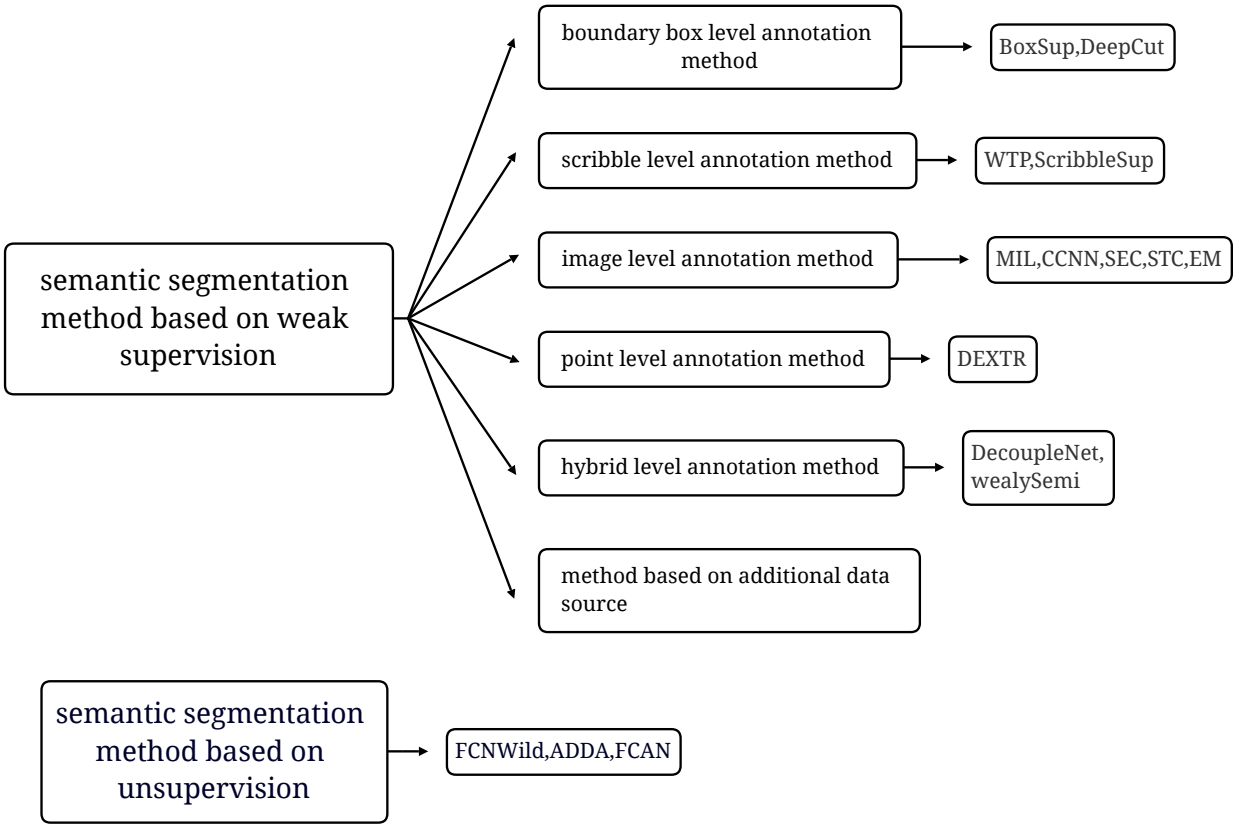

Figure 7: Semantic segmentation method based on weak supervision and unsupervision.

2.3 Comparison of Algorithm Performance

2.3.1 Datasets and Performance Evaluation Metrics

This section takes road scenes as the core background, describes the commonly used datasets in road scene semantic segmentation and the commonly used performance metrics to evaluate the effect of road scene semantic segmentation, and then compares the performance of different semantic segmentation methods under different datasets to analyze and summarize the methods applicable to road scene semantic segmentation.

2.3.1.1 Urban Road Scenes Dataset

A large number of researchers have focused on studying urban road scenes dataset, capturing multidimensional information using multiple sensors in real-world scenarios, and constructing large urban roads dataset through a large number of fine annotations, which have greatly contributed to the development of visual understanding in complex urban street scenes. Commonly used autonomous driving datasets are shown in Table 2, and common traffic sign datasets are shown in Table 3.

2.3.1.2 Runtime

In practice of segmentation, the structural properties and the actual effectiveness of the segmentation framework need to be reasonably measured in order to achieve good results. The performance metrics of segmentation networks are described below in terms of both execution time and accuracy.

Runtime or processing speed is a very valuable metric. In many application areas, real-time performance is a very important property, so the runtime is needed to measure the

Table 2: Common automatic driving datasets.

Dataset	Year	Number of categories	Total amount of data	Area	Environment
CamVid [11]	2009	32	700	Europe	Day
KITTI [28]	2013	10		Germany and America	Day
Oxford Robotcar [65]	2014		2×10^7	Oxford	All weather conditions
Cityscapes [19]	2016	34	20000	Germany, Switzerland and France	Spring, summer, and autumn
SYNTHIA [81]	2016	11	13407		Various scenes
Comma.ai	2016			America	
Mapillary Vistas [69]	2017	66	25000	America, Europe, Africa, Asia, and Oceania	Complex weather
Apollo Scape [40]	2018	28	143906	China	Complex weather
BDD100K [119]	2018	10	10000	Multiple cities around the world	Various scenes
Udacity,s Driving [13]	2018	3, 8	9420, 15000		
NuScenes	2019	23	14×10^5	Boston and Singapore	Day
$D^2 - City$	2019	12		China	Complex weather
Waymo [30]	2019		3000	America	Complex weather

real-time performance of the segmentation method. However, runtime is difficult to compare due to the different levels of hardware back-end implementation. Therefore, the segmentation efficiency of segmentation methods should be judged by comparing the runtimes under the same conditions. For fields with high real-time requirements such as unmanned driving, runtime is a very important evaluation criterion.

2.3.1.3 Accuracy

There are several metrics to evaluate the performance of pixel semantic segmentation. The pixel accuracy (PA) reflects the accuracy of pixel-class correspondence, and the mean pixel accuracy (mPA) can be obtained by averaging it, and there are also intersection ratio metrics, such as Mean Intersection over Union (mIoU) and Frequency-Weighted intersection over Union (FWIoU). The mIoU is often used to measure the performance of semantic segmentation models.

The pixel accuracy is calculated by taking the number of accurate pixels in the prediction category as the object and obtaining the accuracy rate by ratio calculation. The formula is:

$$P = \frac{\sum_{i=0}^{n} p_{ii}}{\sum_{i=0}^{n} \sum_{j=0}^{n} p_{ij}} \tag{1}$$

where: P is the pixel accuracy; p_{ii} is the number of accurately classified pixels. p_{ij} is the number of the misclassified pixels; i and j are the attributed class numbers; n is the total number of categories.

The correspondence between pixels and division classes is not necessarily accurate. So therefore, the accuracy is characterized by the mean pixel accuracy (mPA) metric, which is

Table 3: Common traffic sign datasets.

Dataset	Summary
KUL Belgium Traffic Sign [31]	Dataset of traffic signs in Belgium
German Traffic Sign [38]	German traffic annotated dataset
STSD [133]	More than 20,000images containing 3488traffic signs
LISA [52]	7855 annotations with more than 6610 frames
Tsinghua-Tencent 100K [92]	Dataset with 100000 pictures, including 30000traffic sign examples

calculated as

$$M = \frac{1}{n+1} \sum_{i=0}^{n} \frac{p_{ii}}{\sum_{j=0}^{n} p_{ij}} \tag{2}$$

where M is the mean pixel accuracy.

The Mean Intersection over Union is calculated as:

$$m = \frac{1}{n+1} \sum_{i=0}^{n} \frac{p_{ii}}{\sum_{j=0}^{n} p_{ij} + \sum_{j=0}^{n} p_{ji} - p_{ii}} \tag{3}$$

where m is the mIoU.

The frequency-weighted cross-merge ratio is a new evaluation index after the improvement of mloU index, which is calculated as:

$$F = \frac{1}{\sum_{i=0}^{n}} \sum_{j=0}^{n} p_{ij} \sum_{i=0}^{n} \frac{\sum_{j=0}^{n} p_{ij} p_{ii}}{\sum_{j=0}^{n} p_{ij} + \sum_{j=0}^{n} p_{ji} - p_{ii}} \tag{4}$$

where F is the Frequency-Weighted intersection over Union. The representativeness and simplicity of the mIoU metric are outstanding, and it is currently the most frequently used and common accuracy evaluation metric in the field of image semantic segmentation.

2.3.2 Comparison of Algorithm Performance

The autonomous driving domain studied in this paper requires real-time and efficient segmentation networks. In addition to comparing the segmentation accuracy of different segmentation networks, their real-time performance is investigated in terms of both the number of parameters and the operation rate for networks applicable to semantic segmentation of road scenes.

2.3.2.1 Experimental comparison of traditional semantic segmentation methods

Among the traditional image semantic segmentation methods, classical algorithms such as N-Cut and Grab cut have been widely used, but the efficiency of these algorithms is low. On this basis, the improved algorithms such as GPB-UCM, Random DecisionForest and MCG absorb the advantages of the classical algorithms and have better performance in terms of the quality of the generated image segmentation blocks and the time complexity of the algorithm, but the traditional segmentation methods cannot meet the requirements of the semantic segmentation of road scenes in terms of the number of classification and segmentation accuracy. The analysis of traditional image semantic segmentation methods is summarized in Table 4.

Table 4: Analysis and summary of traditional image semantic segmentation methods [105].

Method	Year	Contribution
Normalized cut	2000	Dividing graph into k subgraphs and then minimizing them
Grab cut	2004	Using image texture and boundary information dependent on small amount of manual intervention to obtain better foreground and background segmentation
GPB-UCM	2011	Using probability of each pixel as an edge, detecting target contour, generating contour map, and completing segmentation with complex steps and high complexity
Random Decision Forest	2016	Combining multiple decision trees into classifier
MCG	2017	On basis of GPS-UCM, using generated multiple contour segmentation blocks when combined with random forest classifier to get prediction object

2.3.2.2 Experimental comparison of strongly supervised semantic segmentation-based methods

For the number of parameters and operation rate of the algorithms, several algorithms with high representativeness and real-time performance were selected from the strongly supervised image semantic segmentation based methods and analyzed and compared on the Cityscapes test dataset in this paper, and the speed analysis comparison is shown in Table 5.

Table 5: Speed analysis of algorithms [45].

Model	Parameter	Time/ms	mIoU/%
FCN-8		500	63.1
DeepLab	250.8	4000	63.1
SegNet	29.5	89.2	57
CRF-RNN		700	74.7
ENet	0.4	135.4	57
DeepLab v2	44	4000	70.4
PSPNet	250.8	1288	81.2
DUC+HDC		900	80.1
DenseASPP	28.6	500	80.6
ESPNet	0.4		60.3
BiSeNet1	5.8	13	68.4
BiSeNet2	49	21	74.7

From 5, we can see that there are still big differences between various algorithms in terms of segmentation speed, among which BiSeNet, ICNet and DFANet algorithms are faster, real-time and suitable for real-time image semantic segmentation. biSeNet proposes a shallow network for high-resolution images and a deep network for fast downsampling, which strikes a balance between classification ability and perceptual field. In contrast, FCN and FCN-based DeepLab v1 and DeepLabv2 have long running time and cannot meet the demand of real-time image segmentation. Among the DeepLab series, DeepLabv3+ has the best segmentation effect, mainly because it absorbs the advantages of the DeepLab series methods and combines the

depth-separable convolution to simplify the model and improve the segmentation efficiency, thus achieving a balance between the accuracy and speed of image semantic segmentation. The segmentation speed of other algorithms is lower than that of FCN, and they cannot meet the demand of real-time image segmentation and are not applicable to dynamic scene segmentation. Therefore, balancing segmentation accuracy and segmentation speed is still the most important task in the unmanned field.

2.3.2.3 Experimental comparison of weakly supervised semantic segmentation-based methods

The weakly supervised image semantic segmentation based methods were compared on the most representative datasets, as shown in 6 [45], with the main factors compared being supervision information, key techniques, the use of PGM methods or not, experimental datasets and evaluation metrics.

Table 6: Analysis and summary of image semantic segmentation method based on weak supervision [45].

Supervision information	Model	Year	Key technology	PGM	Dataset	mIoU/%
Frame level	BoxSup	2015	MCG		PASCAL VOC 2012	75.2
					PASCAL-CONTEXT	40.5
	DeepCut	2016	CRF	CRF		
Scribble level	WTP	2016	Objectness		PASCAL VOC 2012	49.1
	ScribbleSup	2015	Hyperpixel	CRF	PASCAL VOC 2012	71.3
Image level	MIL	2015	MCG		ImageNet	42.0
	CCNN	2015	Class Size		PASCAL VOC 2012	42.4
	SEC	2016	Saliency detection algorithm	CRF	PASCAL VOC 2012	50.7
	STC	2015	Saliency detection algorithm	CRF	PASCAL VOC 2012	49.8
	AugFeed	2016	MCG	CRF	PASCAL VOC 2012	54.34
	EM	2017	Saliency detection algorithm	CRF	PASCAL VOC 2012	58.71
Image level and pixel level	Decoupled	2015		CRF	PASCAL VOC 2012	66.6
Image level, frame level and pixel level	WeaklySemi	2015		CRF	PASCAL VOC 2012	73.9

As can be seen from Table 6, in the weakly supervised semantic segmentation-based methods, although the image-level label is relatively easy to obtain, it contains too little useful information to obtain accurate segmentation results. Although the form of bounding box labeling is more complicated, it can provide supervised information within the target location, so it has better segmentation results compared with other weakly supervised semantic segmentation-based methods. Overall, although the weakly supervised image segmentation technique greatly reduces the requirement of labeling the dataset and lowers the research cost, it contains too little useful information, and the gap between the segmentation effect and segmentation performance is large compared with the strongly supervised semantic segmentation based algorithm, which cannot meet the segmentation requirements in the unmanned field, but it will be a hot spot for future research in this field.

2.3.2.4 Experimental comparison of unsupervised semantic segmentation-based methods

The unsupervised image semantic segmentation based methods were compared on the most representative datasets, as shown in 7 [105], with the main factors compared being the key techniques, the use of PGM methods or not, the experimental dataset and the evaluation metrics.

Table 7: Analysis and summary of image semantic segmentation method based on unsupervision [105].

Model	Year	Key technology	Dataset	mIoU/%
FCNWild	2016	Domain adaptive full convolution adversarial training	Cityscapes	27.1
ADDA	2017	Adversarial training	NYU Depth v2	
FCAN	2018	Image domain adaptive network and feature adaptive network	Cityscapes	47.75

The unsupervised image semantic segmentation based method mainly uses virtual scenes, doing data annotation of real scenes, and then completes semantic segmentation. This method reduces the annotation cost and simplifies the segmentation process, but it requires objective knowledge and understanding of the differences between virtual scenes and real scenes, and the differences in texture and lighting can often reduce the accuracy and precision of image segmentation in real scenes and produce certain segmentation bias. A large amount of research data shows that the effective accuracy of the current unsupervised image semantic segmentation-based methods is not high, and further improving its segmentation accuracy and segmentation quality will be the focus and hot spot of future research.

From the above analysis, it can be seen that in the field of autonomous driving, the strongly supervised semantic segmentation method is still the mainstream road scene segmentation method, and the segmentation accuracy should be considered along with the segmentation efficiency. The weakly supervised and unsupervised image semantic segmentation methods reduce the labeling cost, but the segmentation effect is not obvious at present, the segmentation boundary is rough and discontinuous, and improving its segmentation accuracy is a hot spot for future research.

3 Reinforcement Learning in Autonomous Driving

3.1 Reinforcement Learning

Reinforcement learning is an algorithm based on Markov decision process (MDP, [1]). An agent interacts with the environment through actions while completing a task and generates a new state, while receiving a reward from the environment. In this cycle, training data is continuously generated during the interaction between the agent and the environment. The reinforcement learning algorithm uses the generated data to iteratively optimize its behavioral strategy as it continues to explore the environment. After several iterations, the agent eventually learns the optimal strategy to accomplish the corresponding task. This is illustrated in Figure 8.

Reinforcement learning has been studied in psychology for almost a century, and this work has strongly impacted artificial intelligence. Reinforcement learning can be viewed as the reverse engineering of specific cognitive learning processes. Samuel [84] was the first to use a method similar to today's temporal differential optimization for delayed rewards, and this is considered the earliest research related to reinforcement learning. Reinforcement learning was first proposed by Minsky [67] and was independently introduced into control theory by Waltz and Fu [100]. By the 1980s, Barto and Sutton's [8] work led to reinforcement learning becoming a popular area of machine learning research [9, 8].

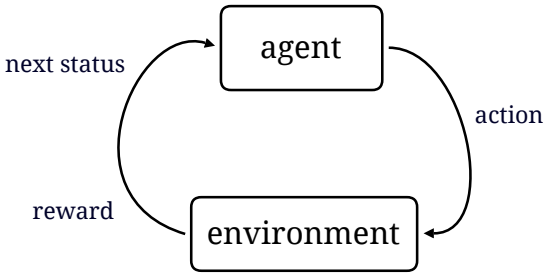

Figure 8: Reinforcement learning model.

A Markov decision process can be represented by a five-tuple $(\boldsymbol{S}, \boldsymbol{A}, P, R, \gamma)$. Where $\boldsymbol{S}$ is the state space, $\boldsymbol{A}$ is the action space, P is the transition probability, R is the reward, and γ is the decay coefficient of the reward. The goal of reinforcement learning is to maximize the desired cumulative reward, so reinforcement learning is ultimately an optimization problem with the objective function

$$J^*(\theta) = \max_{\theta} \mathrm{E}\left\{\sum_{k=0}^{H} \gamma^k r_k\right\} \tag{5}$$

where θ is the parameter to be optimized.

Q-learning [107] is currently the most widely used reinforcement learning algorithm. Watkins and Dayan [106] gave the first complete proof of Q-learning convergence. However, Q-learning is generally based on table settings and can only solve reinforcement learning problems in lower dimensional, discrete state/action spaces. Minh et al. [68] introduced deep neural networks to Q-learning, giving birth to the deep Q network (DQN) algorithm. Deep learning has a powerful feature extraction capability, which provides reinforcement learning the ability to extract features directly from the state space at the high-dimensional pixel level and train them. DQN algorithms have achieved human expert level performance in many video games.

Although DQN solves the problem of high-dimensional state space, it can only handle discrete, low-dimensional action space. However, practical control tasks, especially autonomous driving tasks, often have continuous, high-dimensional action spaces. One solution is to discretize the action space. However, the number of actions grows exponentially with increasing degrees of freedom, and DQNs are challenging to explore and train effectively under huge action spaces. Lillicrap et al. [56] proposed the deep deterministic policy gradient algorithm (DDPG), which successfully implemented the reinforcement learning problem in continuous state and action spaces. The algorithm combines DQN and Actor-Critic algorithms, using deep neural networks to approximate the value function and the policy function, where the value function is updated by the Bellman equation and the policy function is updated by gradient descent. However, the DDPG algorithm has a high degree of parameter fragility, and its hyperparameters often have to be carefully set for different problems to obtain good training results. The most crucial event in the field of deep reinforcement learning in 2016 was the victory of the AlphaGo program developed by Google's DeepMind team over the Korean professional Go player Lee Sedol with a score of 4:1, which achieved world attention [88]. AlphaGo combines supervised and reinforcement learning, first using supervised learning to initialize the strategy network, then iteratively updating the value and strategy networks through reinforcement learning, and then searching based on Monte Carlo search trees. However, AlphaGo still relies to some extent on a priori knowledge of human games. To improve this, Silver et al. developed AlphaZero in 2017, beating all previous Go algorithms and also achieving optimal results in chess and Japanese shogi [89]. AlphaZero relies entirely on Monte Carlo search trees and deep residual networks [35], and does not require any Haarnoja et al. [32] proposed the soft actor-critic method (SAC) in order to address the high sample complexity and hyperparameter crispness of reinforcement learning and to improve the stability of training. This

method introduces the concept of maximum entropy into reinforcement learning and exceeds the efficiency and final performance of the DDPG method.

3.2 Application of Reinforcement Learning in Autonomous Driving Control

The control of autonomous vehicles is based chiefly on rule design, and the number of rules in complex scenarios increases exponentially, and the rules may conflict with each other. It can be challenging to test and validate autonomous vehicles in complicated conditions for safety concerns. In order to cope with complex traffic scenarios, control algorithms need to learn autonomously through data-driven or interaction with the environment, and be able to cope with sophisticated conditions autonomously after sufficient training. Reinforcement learning is one such data-driven autonomous learning method. Unlike supervised learning, reinforcement learning is more suitable for decision making and control of autonomous driving. The autonomous driving task can be considered as a partially observable Markov decision process (POMDP). In POMDP, an agent (i.e., an autonomous vehicle) observes the environment through sensors to obtain an observation I_t, then takes an action a_t in state s_t, and subsequently receives a reward R_{t+1} in the environment, and transitions to the next state s_{t+1}.

The first international application of deep reinforcement learning to vehicle control was by Lange et al. [51], who used a DQN approach to train under a micro-racing simulator and achieved good results with a level of control that even exceeded that of human players. In 2016, Sallab et al. [83] implemented lane keeping control on the open racing car simulator (TORCS) using deep reinforcement learning methods and compared the DQN method in discrete space with the DDAC method in continuous action space, demonstrating that the DDAC method could achieve excellent control and smooth trajectories [83]. Due to the introduction of deep learning methods, Sallab et al. [83] proposed the idea of end-to-end deep reinforcement learning. Thanks to the powerful feature extraction capability of deep neural networks, combined with reinforcement learning methods to train the agent, the original image can be directly mapped to the output of the actuator, and it surpasses simple supervised learning type of end-to-end control in terms of robustness. Deep reinforcement learning is also gradually replacing traditional reinforcement learning methods. Chae et al. [15] used the DQN algorithm to train an agent to learn to handle a pedestrian crossing scene and achieve autonomous vehicle braking control. Zong et al. [134] used the DDPG algorithm to train an agent's acceleration and steering control to achieve autonomous obstacle avoidance and tested it in a TORCS environment. Shalev-Shwartz et al. [85] used reinforcement learning combined with long short term memory networks (LSTM) algorithms to solve the longitudinal control of autonomous driving and the management of merging into traffic circles in a game environment. Yang Shun of Jilin University used deep learning combined with DDPG to propose a deep reinforcement learning control method based on visual scene understanding.

As the research on the application of reinforcement learning in autonomous driving is heating up, to improve the training efficiency of reinforcement learning, Microsoft proposed a framework for distributed cloud-based deep reinforcement learning in 2018, which dramatically reduces the training time [93]. Referring to the process of human learning, Liang et al. [55] combined imitation learning and reinforcement learning and proposed a controlled imitation reinforcement learning approach with good control results in an open-source simulator. This method first initializes the weights of the control network by imitation learning, then performs reinforcement training by the DDPG method. This not only solves the problem that DDPG is sensitive to hyperparameters, but also can better adapt to complex environments than imitation learning alone. Xiangmin Han et al. [33] used the DDPG algorithm to achieve automatic longitudinal control of autonomous driving. It enables intelligent vehicles to complete adaptive cruising and continuously improve during the self-learning process, which results in the control level of human drivers.

3.3 Challenges of Reinforcement Learning in Autonomous Driving

Autonomous driving experiments are hazardous, so most current reinforcement learning models use video game simulation engines for training and simulation, such as TORCS, grand theft auto V (GAT5), and so on. However, there are considerable differences between real and virtual environments, and the stability and robustness of the models can often only be verified using data set validation or offline data playback, while simulator-based training often leads to poor reliability when migrating the trained models to natural environments due to modeling errors. The emergence of generative adversarial networks (GAN) provides an idea to solve this problem. In order to cope with disturbances to improve the robustness of the model, Pinto et al. [75] proposed anti-reinforcement learning by combining adversarial learning and reinforcement learning with training an agent. Yang et al. [115] proposed DU-drive, an end-to-end control framework from virtual to reality, using a framework similar to conditional generative adversarial network (conditional GAN) for image parsing and controlling vehicle motion. Ferdowsi et al. [27] of the Department of Electrical and Computer Engineering at Virginia Tech propose an deep adversarial reinforcement learning framework to address the "safety" problem of self-driving systems to address the safety of self-driving cars.

Multi-intelligent reinforcement learning [12] is also a current direction of reinforcement learning development. In a natural traffic environment, there is not only one participant in the traffic, and the driver's decision and control are often the result of multiple traffic participants playing against each other. Reinforcement learning is based on the theory of Markovian decision processes, while many reinforcement learning algorithms are only approximations of Markovian processes. In autonomous driving applications, for example, the transition of states does not necessarily depend only on the actions taken by the agent, but also on the actions taken by other participants in the environment. Multi-intelligent reinforcement learning is designed to address this problem, with methods such as minimax-Q learning [60], Nash-Q learning [39], etc. It is undeniable that multi-intelligent training is more complex than single-intelligent.

The role of the reward function in reinforcement learning is to guide the agent to continuously optimize its strategy to obtain the expected future cumulative maximum reward. The reward function in most reinforcement learning paradigms is usually coded manually by the system designer. Some obvious reward functions can usually be found for specific reinforcement learning problems, such as the score in a game, profit in a financial situation, etc. However, for some real-world applications of reinforcement learning problems, the reward functions are not only unknown, but also require weighing the needs of many different aspects. If the reward function is not set correctly, the agent may converge in the wrong direction or learn a suboptimal strategy. In autonomous driving applications, the reward function needs to be charged not only for safety and comfort, but also for making the agent more compatible with the driving habits of human drivers. However, the control behavior of human driving is more complex and requires weighing multiple demands and constraints during the driving process, so it is changing to manually specify a reasonable reward function to guide the training of an agent. And an unreasonable reward function can cause a trained model to converge to a local minimum or even poor performance. The smart driving team of Beijing Union University analyzed the driving data to get the characteristics of human drivers and designed the reward function of reinforcement learning to achieve the longitudinal control of unmanned driving, which makes the agent more consistent with human driving habits in longitudinal control [73]. Imitation learning and reverse reinforcement learning provide an effective solution for the acquisition of real reward functions and how to make the performance of an agent closer to that of a human, and have become another research hotspot in unmanned driving [4, 2].

4 Conclusion and Outlook

Autonomous driving technology is a hot spot for research in the world of vehicle engineering and a new growth driver for the automotive industry, and is also an essential part of the intelligent transportation systems currently being developed in various countries. The control

system of self-driving vehicles as the critical link of vehicle behavior layer is crucial for the safety and comfort of vehicle driving. Most traditional control methods are based on precise mathematical analytical models or rule-based designs. The complex and changing traffic scenarios in real traffic environments make it challenging to design accurate mathematical models, and the number of rules grows exponentially with the complexity of the traffic scenarios. With the continuous development of applications such as autonomous driving, there are also higher demands on model size, computational cost, and segmentation accuracy in recent years. The emergence of semantic segmentation and reinforcement learning has made it possible to design data-driven control systems or interact with the environment for autonomous learning. Fully trained learning controllers can also better cope with complex operating conditions.

This paper first introduces the development status and challenges of semantic segmentation of road scenes. The semantic segmentation techniques are classified into traditional models, traditional methods combined with deep learning, and deep learning-based models. The deep learning-based model is highlighted, which is further subdivided into strongly supervised image semantic segmentation-based methods, weakly supervised image semantic segmentation-based methods, and unsupervised image semantic segmentation-based methods. The representative algorithms of each type of method are analyzed and compared for road scenarios, and the technical characteristics, advantages, and disadvantages of each kind of method are summarized. In addition, the application of reinforcement learning in autonomous driving species is also introduced. In general, the application of semantic segmentation and reinforcement learning techniques in autonomous driving is still evolving, but there are some areas for improvement.

1. The accuracy of the semantic segmentation algorithm needs to be further improved. The core of driverless lies in the refined perception and judgment of the surrounding environment, such as changes in the surrounding weather, changes in traffic lights, and oncoming vehicles and pedestrians during driving, which requires accurate segmentation of the input segmentation objects.

2. Real-time semantic segmentation techniques [62, 14]. At this stage, the accuracy rate is still the key metric for evaluating semantic segmentation network models. Still, as autonomous technology continues to mature, the segmentation efficiency has an increasing impact, which requires minimizing the response time while maintaining a high accuracy rate.

3. Weakly-supervised or unsupervised semantic segmentation-based techniques. The current segmentation effect of weakly supervised and unsupervised semantic segmentation-based methods needs to be more satisfactory. Using as little annotation information as possible to improve the accuracy of the network model is the trend of future development.

4. The application of 3D data [42, 126]. Three-dimensional data is crucial to real scenes, and the segmentation objects of most current semantic segmentation methods are two-dimensional scenes, so the application of three-dimensional data will be a hot spot for future research.

References

[1] A Abakuks. Reviewed work: Markov processes: characterization and convergence. by sn ethier, tg kurtz. *Biometrics*, 43(2):113–122, 1987.

[2] P. Abbeel and A. Ng. Apprenticeship learning via inverse reinforcement learning. *Proceedings of the twenty-first international conference on Machine learning*, 2004.

[3] Jiwoon Ahn and Suha Kwak. Learning pixel-level semantic affinity with image-level supervision for weakly supervised semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4981–4990, 2018.

[4] R. Amit and Maja J. Matari. Learning movement sequences from demonstration. *Proceedings 2nd International Conference on Development and Learning. ICDL 2002*, pages 203–208, 2002.

[5] Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik. Contour detection and hierarchical image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 33(5):898–916, 2010.

[6] Anurag Arnab, Sadeep Jayasumana, Shuai Zheng, and Philip HS Torr. Higher order conditional random fields in deep neural networks. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*, pages 524–540. Springer, 2016.

[7] Peter L Bartlett. An introduction to reinforcement learning theory: Value function methods. In *Advanced Lectures on Machine Learning: Machine Learning Summer School 2002 Canberra, Australia, February 11–22, 2002 Revised Lectures*, pages 184–202. Springer, 2003.

[8] Andrew G Barto and Richard S Sutton. Landmark learning: An illustration of associative search. *Biological cybernetics*, 42(1):1–8, 1981.

[9] Andrew G Barto, Richard S Sutton, and Charles W Anderson. Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE transactions on systems, man, and cybernetics*, (5):834–846, 1983.

[10] Amy Bearman, Olga Russakovsky, Vittorio Ferrari, and Li Fei-Fei. What's the point: Semantic segmentation with point supervision. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VII 14*, pages 549–565. Springer, 2016.

[11] Gabriel J Brostow, Julien Fauqueur, and Roberto Cipolla. Semantic object classes in video: A high-definition ground truth database. *Pattern Recognition Letters*, 30(2):88–97, 2009.

[12] Lucian Busoniu, Robert Babuska, and Bart De Schutter. Multi-agent reinforcement learning: A survey. In *2006 9th International Conference on Control, Automation, Robotics and Vision*, pages 1–6, 2006.

[13] Alexander Buyval, Aidar Gabdullin, Ruslan Mustafin, and Ilya Shimchik. Realtime vehicle and pedestrian tracking for didi udacity self-driving car challenge. In *2018 IEEE international conference on robotics and automation (ICRA)*, pages 2064–2069. IEEE, 2018.

[14] Huang X G Cai Y and Zhang Z A. Real-time semantic segmentation algorithm based on featurefusion technology. *Laser & Optoelectronics Progress*, 57:021011, 2020.

[15] Hyunmin Chae, Chang Mook Kang, ByeoungDo Kim, Jaekyum Kim, Chung Choo Chung, and Jun Won Choi. Autonomous braking system via deep reinforcement learning. In *2017 IEEE 20th International conference on intelligent transportation systems (ITSC)*, pages 1–6. IEEE, 2017.

[16] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Semantic image segmentation with deep convolutional nets and fully connected crfs. *arXiv preprint arXiv:1412.7062*, 2014.

[17] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2017.

[18] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 801–818, 2018.

[19] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3213–3223, 2016.

[20] Jifeng Dai, Kaiming He, and Jian Sun. Boxsup: Exploiting bounding boxes to supervise convolutional networks for semantic segmentation. In *Proceedings of the IEEE international conference on computer vision*, pages 1635–1643, 2015.

[21] Jifeng Dai, Haozhi Qi, Yuwen Xiong, Yi Li, Guodong Zhang, Han Hu, and Yichen Wei. Deformable convolutional networks. In *Proceedings of the IEEE international conference on computer vision*, pages 764–773, 2017.

[22] Liuyuan Deng, Ming Yang, Zhidong Liang, Yuesheng He, and Chunxiang Wang. Fusing geometrical and visual information via superpoints for the semantic segmentation of 3d road scenes. *Tsinghua science and technology*, 25(4):498–507, 2020.

[23] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. In *International conference on machine learning*, pages 647–655. PMLR, 2014.

[24] Abdel Hamid Elhofi and Hany Ahmed Helaly. Comparison between digital and manual marking for toric intraocular lenses: a randomized trial. *Medicine*, 94(38), 2015.

[25] Yuchun Fang, Yifan Li, Xiaokang Tu, Taifeng Tan, and Xin Wang. Face completion with hybrid dilated convolution. *Signal Processing: Image Communication*, 80:115664, 2020.

[26] Shouting Feng, Zhongshuo Zhuo, Daru Pan, and Qi Tian. Ccnet: A cross-connected convolutional network for segmenting retinal vessels using multi-scale features. *Neurocomputing*, 392:268–276, 2020.

[27] Aidin Ferdowsi, Ursula Challita, Walid Saad, and Narayan B Mandayam. Robust deep reinforcement learning for security and safety in autonomous vehicle systems. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 307–312. IEEE, 2018.

[28] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research*, 32(11):1231–1237, 2013.

[29] Golnaz Ghiasi and Charless C Fowlkes. Laplacian pyramid reconstruction and refinement for semantic segmentation. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14*, pages 519–534. Springer, 2016.

[30] Zhicheng Gu, Zhihao Li, Xuan Di, and Rongye Shi. An lstm-based autonomous driving model using a waymo open dataset. *Applied Sciences*, 10(6):2046, 2020.

[31] Anjan Gudigar, Shreesha Chokkadi, U Raghavendra, and U Rajendra Acharya. An efficient traffic sign recognition based on graph embedding features. *Neural Computing and Applications*, 31:395–407, 2019.

[32] Tuomas Haarnoja, Aurick Zhou, Kristian Hartikainen, George Tucker, Sehoon Ha, Jie Tan, Vikash Kumar, Henry Zhu, Abhishek Gupta, Pieter Abbeel, et al. Soft actor-critic algorithms and applications. *arXiv preprint arXiv:1812.05905*, 2018.

[33] Liang J Pan F Han X M, Bao H and Xuan Z X. An adaptive cruise control algorithm based on deep reinforcement learning. *Computer Engineering*, pages 32–35, 2018.

[34] Junjun He, Zhongying Deng, and Yu Qiao. Dynamic multi-scale filters for semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3562–3572, 2019.

[35] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[36] Judy Hoffman, Dequan Wang, Fisher Yu, and Trevor Darrell. Fcns in the wild: Pixel-level adversarial and constraint-based adaptation. *arXiv preprint arXiv:1612.02649*, 2016.

[37] Seunghoon Hong, Hyeonwoo Noh, and Bohyung Han. Decoupled deep neural network for semi-supervised semantic segmentation. *Advances in neural information processing systems*, 28, 2015.

[38] Sebastian Houben, Johannes Stallkamp, Jan Salmen, Marc Schlipsing, and Christian Igel. Detection of traffic signs in real-world images: The german traffic sign detection benchmark. In *The 2013 international joint conference on neural networks (IJCNN)*, pages 1–8. Ieee, 2013.

[39] Junling Hu and Michael P. Wellman. Nash q-learning for general-sum stochastic games. *J. Mach. Learn. Res.*, 4:1039–1069, 2003.

[40] Xinyu Huang, Peng Wang, Xinjing Cheng, Dingfu Zhou, Qichuan Geng, and Ruigang Yang. The apolloscape open dataset for autonomous driving and its application. *IEEE transactions on pattern analysis and machine intelligence*, 42(10):2702–2719, 2019.

[41] Zilong Huang, Xinggang Wang, Jiasi Wang, Wenyu Liu, and Jingdong Wang. Weakly-supervised semantic segmentation network with deep seeded region growing. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7014–7023, 2018.

[42] Yang J and Dang J S. Recognition and segmentation of three-dimensional point cloud based on deep cascade convolutional neural network. *Optics and Precision Engineering*, 28:1187–1199, 2020.

[43] Jindong Jiang, Zhijun Zhang, Yongqian Huang, and Lunan Zheng. Incorporating depth into both cnn and crf for indoor semantic segmentation. In *2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, pages 525–530. IEEE, 2017.

[44] ZHOU Jimiao, LI Bijun, and CHEN Shizeng. A real time semantic segmentation method based on multi-level feature fusion. *Bulletin of Surveying and Mapping*, (1):10, 2020.

[45] ZW Jing, HY Guan, DF Peng, and YT Yu. Survey of research in image semantic segmentation based on deep neural network. *Computer Engineering*, 46(10):1–17, 2020.

[46] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of artificial intelligence research*, 4:237–285, 1996.

[47] Muhammad Waseem Khan. A survey: Image segmentation techniques. *International Journal of Future Computer and Communication*, 3(2):89, 2014.

[48] Alexander Kolesnikov and Christoph H Lampert. Seed, expand and constrain: Three principles for weakly-supervised image segmentation. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14*, pages 695–711. Springer, 2016.

[49] Vijay R Konda and John N Tsitsiklis. Onactor-critic algorithms. *SIAM journal on Control and Optimization*, 42(4):1143–1166, 2003.

[50] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.

[51] Sascha Lange, Martin Riedmiller, and Arne Voigtländer. Autonomous reinforcement learning on raw visual input data in a real world application. In *The 2012 international joint conference on neural networks (IJCNN)*, pages 1–8. IEEE, 2012.

[52] Eunseop Lee and Daijin Kim. Accurate traffic light detection using deep neural network with focal regression loss. *Image and Vision Computing*, 87:24–36, 2019.

[53] Zhen Li, Yukang Gan, Xiaodan Liang, Yizhou Yu, Hui Cheng, and Liang Lin. Lstm-cf: Unifying context modeling and fusion with lstms for rgb-d scene labeling. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*, pages 541–557. Springer, 2016.

[54] Xiaodan Liang, Xiaohui Shen, Jiashi Feng, Liang Lin, and Shuicheng Yan. Semantic object parsing with graph lstm. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, pages 125–143. Springer, 2016.

[55] Xiaodan Liang, Tairui Wang, Luona Yang, and Eric Xing. Cirl: Controllable imitative reinforcement learning for vision-based self-driving. In *Proceedings of the European conference on computer vision (ECCV)*, pages 584–599, 2018.

[56] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

[57] Di Lin, Jifeng Dai, Jiaya Jia, Kaiming He, and Jian Sun. Scribblesup: Scribble-supervised convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3159–3167, 2016.

[58] Guosheng Lin, Anton Milan, Chunhua Shen, and Ian Reid. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1925–1934, 2017.

[59] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125, 2017.

[60] Michael L. Littman. Markov games as a framework for multi-agent reinforcement learning. In *International Conference on Machine Learning*, 1994.

[61] Lei Liu, Zhiguo Shi, Haoru Su, Hong Li, et al. Image segmentation based on higher order markov random field. 2013.

[62] He Yuqing Lu Wenchao, Pang Yanwei and Wang Jian. Real-time and accurate semantic segmentation based on separable residual modules. *Laser & Optoelectronics Progress*, 56:051005, 2019.

[63] Pauline Luc, Camille Couprie, Soumith Chintala, and Jakob Verbeek. Semantic segmentation using adversarial networks. *arXiv preprint arXiv:1611.08408*, 2016.

[64] Chao Luo, Xin Wang, Xiaojie Li, Yucheng Chen, Jiliu Zhou, Kunlin Cao, Qi Song, Xi Wu, and Youbing Yin. Acnet: Attention-based convolution network with additional discriminative features for dcm classification (s). In *SEKE*, pages 535–700, 2019.

[65] Will Maddern, Geoffrey Pascoe, Chris Linegar, and Paul Newman. 1 year, 1000 km: The oxford robotcar dataset. *The International Journal of Robotics Research*, 36(1):3–15, 2017.

[66] Kevis-Kokitsi Maninis, Sergi Caelles, Jordi Pont-Tuset, and Luc Van Gool. Deep extreme cut: From extreme points to object segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 616–625, 2018.

[67] Marvin Minsky. Steps toward artificial intelligence. *Proceedings of the IRE*, 49(1):8–30, 1961.

[68] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *nature*, 518(7540):529–533, 2015.

[69] Gerhard Neuhold, Tobias Ollmann, Samuel Rota Bulo, and Peter Kontschieder. The mapillary vistas dataset for semantic understanding of street scenes. In *Proceedings of the IEEE international conference on computer vision*, pages 4990–4999, 2017.

[70] Hyeonwoo Noh, Seunghoon Hong, and Bohyung Han. Learning deconvolution network for semantic segmentation. In *Proceedings of the IEEE international conference on computer vision*, pages 1520–1528, 2015.

[71] Chris J Ostafew, Angela P Schoellig, Timothy D Barfoot, and Jack Collier. Learning-based nonlinear model predictive control to improve vision-based mobile robot path tracking. *Journal of Field Robotics*, 33(1):133–152, 2016.

[72] Brian Paden, Michal Čáp, Sze Zheng Yong, Dmitry Yershov, and Emilio Frazzoli. A survey of motion planning and control techniques for self-driving urban vehicles. *IEEE Transactions on intelligent vehicles*, 1(1):33–55, 2016.

[73] Feng Pan and Hong Bao. Reinforcement learning model with a reward function based on human driving characteristics. *2019 15th International Conference on Computational Intelligence and Security (CIS)*, pages 225–229, 2019.

[74] Adam Paszke, Abhishek Chaurasia, Sangpil Kim, and Eugenio Culurciello. Enet: A deep neural network architecture for real-time semantic segmentation. *arXiv preprint arXiv:1606.02147*, 2016.

[75] Lerrel Pinto, James Davidson, Rahul Sukthankar, and Abhinav Gupta. Robust adversarial reinforcement learning. In *International Conference on Machine Learning*, pages 2817–2826. PMLR, 2017.

[76] Jordi Pont-Tuset, Pablo Arbelaez, Jonathan T Barron, Ferran Marques, and Jitendra Malik. Multiscale combinatorial grouping for image segmentation and object proposal generation. *IEEE transactions on pattern analysis and machine intelligence*, 39(1):128–140, 2016.

[77] Li QB and Su D. Multi-organ abdominal imagesegmentation based on v-net. pages 89–91, 2019.

[78] Rajesh Rajamani. *Vehicle dynamics and control*. Springer Science & Business Media, 2011.

[79] Martin Rajchl, Matthew CH Lee, Ozan Oktay, Konstantinos Kamnitsas, Jonathan Passerat-Palmbach, Wenjia Bai, Mellisa Damodaram, Mary A Rutherford, Joseph V Hajnal, Bernhard Kainz, et al. Deepcut: Object segmentation from bounding box annotations using convolutional neural networks. *IEEE transactions on medical imaging*, 36(2):674–683, 2016.

[80] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.

[81] German Ros, Laura Sellart, Joanna Materzynska, David Vazquez, and Antonio M Lopez. The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3234–3243, 2016.

[82] Mukhopadhyay s. Stochastic gradient descent forlinear systems with sequential matrix entry accumulation. *Signal Processing*, 171:107494, 2020.

[83] Ahmad El Sallab, Mohammed Abdou, Etienne Perot, and Senthil Yogamani. End-to-end deep reinforcement learning for lane keeping assist. *arXiv preprint arXiv:1612.04340*, 2016.

[84] Arthur L. Samuel. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.*, 44:206–227, 1967.

[85] Shai Shalev-Shwartz, Nir Ben-Zrihem, Aviad Cohen, and Amnon Shashua. Long-term planning by short-term prediction. *arXiv preprint arXiv:1602.01580*, 2016.

[86] Evan Shelhamer, Jonathan Long, and Trevor Darrell. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell. 2017 Apr*, pages 640–651, 2017.

[87] Falong Shen, Rui Gan, Shuicheng Yan, and Gang Zeng. Semantic segmentation via structured patch prediction, context crf and guidance crf. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1953–1961, 2017.

[88] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[89] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv:1712.01815*, 2017.

[90] Jarrod M Snider et al. Automatic steering methods for autonomous automobile path tracking. *Robotics Institute, Pittsburgh, PA, Tech. Rep. CMU-RITR-09-08*, 2009.

[91] Chunfeng Song, Yan Huang, Wanli Ouyang, and Liang Wang. Box-driven class-wise region masking and filling rate guided loss for weakly supervised semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3136–3145, 2019.

[92] Shijin Song, Zhiqiang Que, Junjie Hou, Sen Du, and Yuefeng Song. An efficient convolutional neural network for small traffic sign detection. *Journal of Systems Architecture*, 97:269–277, 2019.

[93] Mitchell Spryn, Aditya Sharma, Dhawal Parkar, and Madhur Shrimal. Distributed deep reinforcement learning on the cloud for autonomous driving. In *Proceedings of the 1st International Workshop on Software Engineering for AI in Autonomous Systems*, pages 16–22, 2018.

[94] Richard S. Sutton. Introduction: The challenge of reinforcement learning. *Machine Learning*, 8:225–227, 1992.

[95] Richard S Sutton, Andrew G Barto, et al. *Introduction to reinforcement learning*, volume 135. MIT press Cambridge, 1998.

[96] Zhi Tian, Tong He, Chunhua Shen, and Youliang Yan. Decoders matter for semantic segmentation: Data-dependent decoding enables flexible feature aggregation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3126–3135, 2019.

[97] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7167–7176, 2017.

[98] Raviteja Vemulapalli, Oncel Tuzel, Ming-Yu Liu, and Rama Chellapa. Gaussian conditional random field network for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3224–3233, 2016.

[99] Francesco Visin, Marco Ciccone, Adriana Romero, Kyle Kastner, Kyunghyun Cho, Yoshua Bengio, Matteo Matteucci, and Aaron Courville. Reseg: A recurrent neural network-based model for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 41–48, 2016.

[100] M Waltz and K Fu. A heuristic approach to reinforcement learning control systems. *IEEE Transactions on Automatic Control*, 10(4):390–398, 1965.

[101] CY Wang, JZ Chen, and W Li. Review on superpixel segmentation algorithms. *Appl. Res. Comput*, 31(1):6–12, 2014.

[102] Kun-Feng Wang, Chao Gou, Yan-Jie Duan, Yi-Lun Lin, Xin-Hu Zheng, and FY Wang. Generative adversarial networks: the state of the art and beyond. *Acta Automatica Sinica*, 43(3):321–332, 2017.

[103] Panqu Wang, Pengfei Chen, Ye Yuan, Ding Liu, Zehua Huang, Xiaodi Hou, and Garrison Cottrell. Understanding convolution for semantic segmentation. In *2018 IEEE winter conference on applications of computer vision (WACV)*, pages 1451–1460. Ieee, 2018.

[104] Yanjie Wang, Guodong Wang, Chenglizhao Chen, and Zhenkuan Pan. Multi-scale dilated convolution of convolutional neural network for image denoising. *Multimedia Tools and Applications*, 78:19945–19960, 2019.

[105] YR Wang, QL Chen, and JJ Wu. Research on image semantic segmentation for complex environments. *Computer Science*, 46(9):36–46, 2019.

[106] Christopher JC Watkins. H.; dayan, p. q-learning. *Machine learning*, 8(3):279–292, 1992.

[107] Christopher John Cornish Hellaby Watkins. Learning from delayed rewards. 1989.

[108] Yunchao Wei, Xiaodan Liang, Yunpeng Chen, Xiaohui Shen, Ming-Ming Cheng, Jiashi Feng, Yao Zhao, and Shuicheng Yan. Stc: A simple to complex framework for weakly-supervised semantic segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(11):2314–2320, 2016.

[109] Huikai Wu, Junge Zhang, Kaiqi Huang, Kongming Liang, and Yizhou Yu. Fastfcn: Rethinking dilated convolution in the backbone for semantic segmentation. *arXiv preprint arXiv:1903.11816*, 2019.

[110] Zongsheng Wu, Weiping Fu, and Gaining Han. Road scene understanding based on deep convolutional neural network. *Computer engineering and applications*, 53(22):8–15, 2017.

[111] You-chun Xu, R Wang, BJAE Li, and B Li. A summary of worldwide intelligent vehicle. *Automotive engineering*, 23(5):289–295, 2001.

[112] Haolan Xue, Chang Liu, Fang Wan, Jianbin Jiao, Xiangyang Ji, and Qixiang Ye. Danet: Divergent activation for weakly supervised object localization. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6589–6598, 2019.

[113] Yuan Xue, Tao Xu, Han Zhang, L Rodney Long, and Xiaolei Huang. Segan: Adversarial network with multi-scale l 1 loss for medical image segmentation. *Neuroinformatics*, 16:383–392, 2018.

[114] Yunyang YAN, Xuexin QU, Quanyin ZHU, et al. Confidence measure method of classification results based on outlier detection. *Journal of Nanjing University: Natural Science*, 55(1):102–109, 2019.

[115] Luona Yang, Xiaodan Liang, Tairui Wang, and Eric Xing. Real-to-virtual domain unification for end-to-end autonomous driving. In *Proceedings of the European conference on computer vision (ECCV)*, pages 530–545, 2018.

[116] Maoke Yang, Kun Yu, Chi Zhang, Zhiwei Li, and Kuiyuan Yang. Denseaspp for semantic segmentation in street scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3684–3692, 2018.

[117] Wu Y X Yang L and Wang J L. Research on recurrent neural network. pages 1–6, 2016.

[118] Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, and Nong Sang. Bisenet: Bilateral segmentation network for real-time semantic segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 325–341, 2018.

[119] Fisher Yu, Haofeng Chen, Xin Wang, Wenqi Xian, Yingying Chen, Fangchen Liu, Vashisht Madhavan, and Trevor Darrell. Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2636–2645, 2020.

[120] Fisher Yu and Vladlen Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*, 2015.

[121] Fisher Yu, Vladlen Koltun, and Thomas Funkhouser. Dilated residual networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 472–480, 2017.

[122] YANG Yu-peng, ZHAO Wei-dong, WANG Zhi-cheng, and CHEN Gang. Research on graph-based normalized cut image segmentation method. *Computer and Modernization*, 1(01):113, 2010.

[123] Zhang L Yuan J J and Chen Y H. Deep neuralnetwork based on attention convolution module forimage recognition. *Computer Engineering and Applications*, 55:9–16, 2019.

[124] Chunjie Zhang, Zhe Xue, Xiaobin Zhu, Huanian Wang, Qingming Huang, and Qi Tian. Boosted random contextual semantic space based representation for visual recognition. *Information Sciences*, 369:160–170, 2016.

[125] Yiheng Zhang, Zhaofan Qiu, Ting Yao, Dong Liu, and Tao Mei. Fully convolutional adaptation networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 6810–6818, 2018.

[126] Liu L L Zhang A W and Zhang X Z. Multi-feature 3d road point cloud semantic segmentation method based on convolutional neural network. *Optics and Precision Engineering*, 47:0410001, 2020.

[127] Hengshuang Zhao, Xiaojuan Qi, Xiaoyong Shen, Jianping Shi, and Jiaya Jia. Icnet for real-time semantic segmentation on high-resolution images. In *Proceedings of the European conference on computer vision (ECCV)*, pages 405–420, 2018.

[128] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2881–2890, 2017.

[129] Hengshuang Zhao, Yi Zhang, Shu Liu, Jianping Shi, Chen Change Loy, Dahua Lin, and Jiaya Jia. Psanet: Point-wise spatial attention network for scene parsing. In *Proceedings of the European conference on computer vision (ECCV)*, pages 267–283, 2018.

[130] Qiuhua Zheng, Wenqing Li, Weihua Hu, and Guohua Wu. An interactive image segmentation algorithm based on graph cut. *Procedia Engineering*, 29:1420–1424, 2012.

[131] Shuai Zheng, Sadeep Jayasumana, Bernardino Romera-Paredes, Vibhav Vineet, Zhizhong Su, Dalong Du, Chang Huang, and Philip HS Torr. Conditional random fields as recurrent neural networks. In *Proceedings of the IEEE international conference on computer vision*, pages 1529–1537, 2015.

[132] Yanzhao Zhou, Yi Zhu, Qixiang Ye, Qiang Qiu, and Jianbin Jiao. Weakly supervised instance segmentation using class peak response. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3791–3800, 2018.

[133] Yingying Zhu, Chengquan Zhang, Duoyou Zhou, Xinggang Wang, Xiang Bai, and Wenyu Liu. Traffic sign detection and recognition using fully convolutional network guided proposals. *Neurocomputing*, 214:758–766, 2016.

[134] Xiaopeng Zong, Guoyan Xu, Guizhen Yu, Hongjie Su, and Chaowei Hu. Obstacle avoidance for self-driving vehicle with reinforcement learning. *SAE International Journal of Passenger Cars-Electronic and Electrical Systems*, 11(07-11-01-0003):30–39, 2017.