How Generative Adversarial Networks Promote the Development of Intelligent Transportation Systems: A Survey

Hongyi Lin, Yang Liu, Shen Li, and Xiaobo Qu, *Senior Member IEEE*

***Abstract*—In current years, the improvement of deep learning has brought about tremendous changes: As a type of unsupervised deep learning algorithm, generative adversarial networks (GANs) have been widely employed in various fields including transportation. This paper reviews the development of GANs and their applications in the transportation domain. Specifically, many adopted GAN variants for autonomous driving are classified and demonstrated according to data generation, video trajectory prediction, and security of detection. To introduce GANs to traffic research, this review summarizes the related techniques for spatio-temporal, sparse data completion, and time-series data evaluation. GAN-based traffic anomaly inspections such as infrastructure detection and status monitoring are also assessed. Moreover, to promote further development of GANs in intelligent transportation systems (ITSs), challenges and noteworthy research directions on this topic are provided. In general, this survey summarizes 130 GAN-related references and provides comprehensive knowledge for scholars who desire to adopt GANs in their scientific works, especially transportation-related tasks.**

***Index Terms*—Autonomous driving, generative adversarial network (GAN), intelligent transportation system (ITS), traffic anomaly inspection, traffic flow.**

I. Introduction

OVER the past few decades, artificial intelligence has been developing rapidly in the domains of bioinformatics, computer vision, drug design, natural language processing, etc. [1]. As machine learning continues to grow, autonomous vehicles (AVs) and intelligent transportation systems (ITSs) have entered a stage of rapid growth and promoted the innovation and the rise of related technologies and industries such as the automobile industry, 5G new infrastructure construction, and intelligent vehicle infrastructure cooperative systems [2]−[5].

The effectiveness of deep learning in the realm of computers motivates scholars to apply it elsewhere. Scholars have applied neural networks, such as the Transformer, long-short term memory (LSTM), recursive neural network (RNN), feed-forward neural network, etc. to the transportation domain [6], [7]. Specifically, Liu *et al.* [8] established a highly flexible and extendable deep learning architecture to forecast the inbound/outbound passenger flow of the subway and obtained a high level of prediction accuracy due to the simplicity of integrating data from several sources. Xu *et al.* [9] designed a model based on a transformer that combines features of several different vehicles to improve the accuracy of vehicle velocity prediction to quantify the influence of surrounding vehicles on a driving vehicle.

One of the most promising approaches for unsupervised learning on complicated distributions in recent years is the generative adversarial network (GAN), a deep learning model published by Goodfellow *et al.* [10] in 2014. It simultaneously trains the generative model, which depicts the distribution of the data, and the discriminative model, which calculates the likelihood that the sample is drawn from the training set. It aims to generate artificial samples that are indistinguishable from real samples as far as possible. Since they can generate, transform and repair images and data, GANs have high potential research value for transportation research, a field with many data-driven tasks. Currently, some scholars used the GAN to conduct relevant research in the field of transportation, such as autonomous driving, traffic flow data, traffic anomaly inspection, etc.

To the limit of our knowledge, on the one hand, there are a massive number of existing surveys focusing on employing machine learning or deep learning in the transportation domain, but few of them consider the GAN. Deep-learning (DL)-input-based fusion, DL-output-based fusion, and DL-double-stage-based fusion are the three categories into which Liu *et al.* [11] divided deep learning-based urban big data fusion techniques. Elsewhere, others explored deep learning techniques for current hot topics, such as autonomous vehicle (AV) identification, traffic network modeling, traffic signal control, and traffic flow prediction. Wang *et al.* [12] summarized the deep learning-based techniques in transportation as well as other domains from a spatio-temporal perspective. Liu

Manuscript received May 7, 2023; revised June 11, 2023; accepted June 18, 2023. This work was supported by the National Natural Science Foundation of China (52221005, 52220105001, 52272420) and European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie (101025896). Recommended by Associate Editor Xiaohua Ge. *(Corresponding authors: Yang Liu and Shen Li.)*

Citation: H. Y. Lin, Y. Liu, S. Li, and X. B. Qu, "How generative adversarial networks promote the development of intelligent transportation systems: A survey," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 9, pp. 1781–1796, Sept. 2023.

H. Y. Lin and X. B. Qu are with the School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China (e-mail: hy-lin22@mails.tsinghua.edu.cn; xiaobo@tsinghua.edu.cn).

Y. Liu is with the Department of Architecture and Civil Engineering, Chalmers University of Technology, Gothenburg SE-41296, Sweden (e-mail: liuy@chalmers.se).

S. Li is with the Department of Civil Engineering, Tsinghua University, Beijing 100084, China (e-mail: sli299@tsinghua.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2023.123744

et al. [13] gave a comprehensive analysis of machine learning-based approaches for on-demand ride-hailing services, highlighting their potential in assisting with the creation, organization, management, and control of urban ITSs. Ye *et al*. [14] decomposed the existing graph deep learning frameworks, summarized the common deep learning technologies, including GNNs, RNNs, GANs, etc., and analyzed the theories, merits, drawbacks, specific variants, and applications in transportation scenarios. However, none of these surveys consider GAN-related literature, except for a brief portion by Ye *et al*. referencing GANs [14].

On the other hand, there are also numerous overviews of the GAN, whose applications are mainly in image processing, but few involve the application in the transportation domain. In the existing summaries of the GAN, some scholars investigated the background, practical models, advantages and disadvantages, and development trends, pertaining to the GAN and summarized the latest variants of the GAN, and considered its future [15]. A thorough assessment and analysis of current research on GAN models and their uses were provided in [16]. Furthermore, many studies have reviewed the application of GANs in various niches. Kulkarni summarized the application of GANs in music, such as generating sheet music, music matching, etc., and concluded that the GAN offers efficient training for producing music using a set of midi files [17]. With an emphasis on its representative algorithms and their uses in picture editing and restoration, Xia *et al*. [18] provided a comprehensive study of GAN inversions. Wang *et al*. [19] focused on approaches for detecting face images that are generated or synthesized from GAN models and provided a comprehensive review of recent progress in GAN-face detection. An overview of the utilization of GANs in ophthalmic image domains was presented by You *et al*. [20] to discuss significant contributions and to suggest possible future research areas.

In conclusion, a thorough, systematic overview examining recent developments related to GANs in the traffic domain is still lacking. Our study strives to close this gap in knowledge and advance the comprehension of the most recent technological advancement in the transportation community. The goal of this survey is to summarize the literature on the employment of GANs in the transportation domain to see how it promotes the development of ITSs. The key contributions are drawn as follows:

1) The development of GANs is summarized and the existing applications are categorized since their conceptualization;

2) The applications of GANs in the transportation domain are classified and demonstrated, encompassing critical areas such as autonomous driving, traffic flow research, and traffic anomaly inspection;

3) Challenges associated with the integration of GANs into transportation operations are identified, accompanied by suggestions for future research directions to advance the state of the art.

The remainder of this review is structured as follows: Section II provides the fundamental formulation and principle of GANs and summarizes their applications. In Section III, the current development of GANs in autonomous driving is discussed, where the issues of perception, decision-making, prediction, and safety are demonstrated, respectively. Section IV is where the application of GANs is examined in traffic flow research, including the prediction of spatio-temporal sequence information, and the completion and repair of missing traffic data. The applications of GANs in anomaly inspection of traffic facilities and status, such as the crack detection of roads and bridges, inspection of railway flaws, and traffic anomaly monitoring are expounded in Section V. Section VI discusses challenges in deploying GANs for the transportation domain and provides some promising directions of future research. Our work is concluded in Section VII with some final remarks.

II. Principle and Development of GANs

Since it was initially introduced in 2014, the GAN has been extensively studied and has been combined with other machine learning algorithms in some specific applications. For supervised learning, semi-supervised learning, and unsupervised learning problems, numerous novel algorithms have been developed. Currently, there are thousands of variants of the GAN. Since the GAN can generate approximate samples without explicitly modeling any data distribution, it has been frequently employed in various fields. This section primarily demonstrates the principle, evolution, and main uses of the GAN.

A. The Principle of the GAN

A GAN model is made up of two components: a generator G and a discriminator D, akin to the network's traditional form (Fig. 1). The generator tries to represent the distribution of input data for the generated results, called $G(z)$. The distribution of the generated data and the actual data can serve as the discriminator's two inputs. The output of the discriminator $D(x)$ is "1" or "0", which means the output is 100% real or completely generated data, respectively.

It is a binary minimax adversarial challenge to optimize the GAN. The purpose of optimization is that the generator generates results for which the discriminator finds it difficult to identify its source, and for the discriminator, the goal is to discriminate synthetic samples from real samples as accurately as possible. Therefore, G and D form an interactive "game process". At its best, G can generate enough data $G(z)$ to be "fake". It is challenging for D to evaluate whether the data produced by G are accurate or not, so $D(G(z)) = 1$. Then, our goal is to have a generative model that can be utilized to produce synthetic data.

There are no mandatory restrictions on the choice of models of the generator and discriminator in the GAN. Generally, the loss function for directing the training of the discriminator is found as

$$\max V(D,G) = E_{x\sim p_{\text{data}}(x)}[\log D(x)] + E_{z\sim p_z(z)}[\log(1-D(G(z)))] \tag{1}$$

where V means the output value of the loss function, $p_{\text{data}}(x)$ represents the distribution of real data, $p_z(z)$ denotes the distribution of generated data, and E is the mean value. For the generated data, the G is recognized by the D which can be

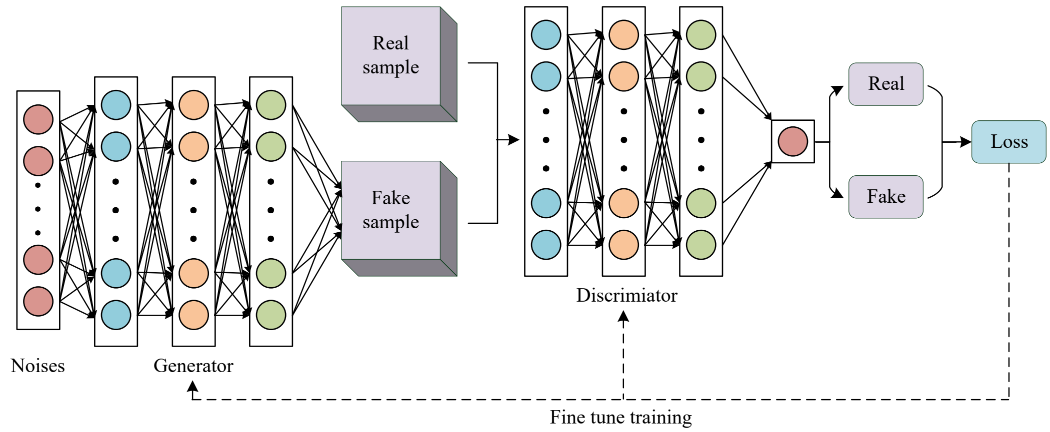

Fig. 1. Structure of the GAN.

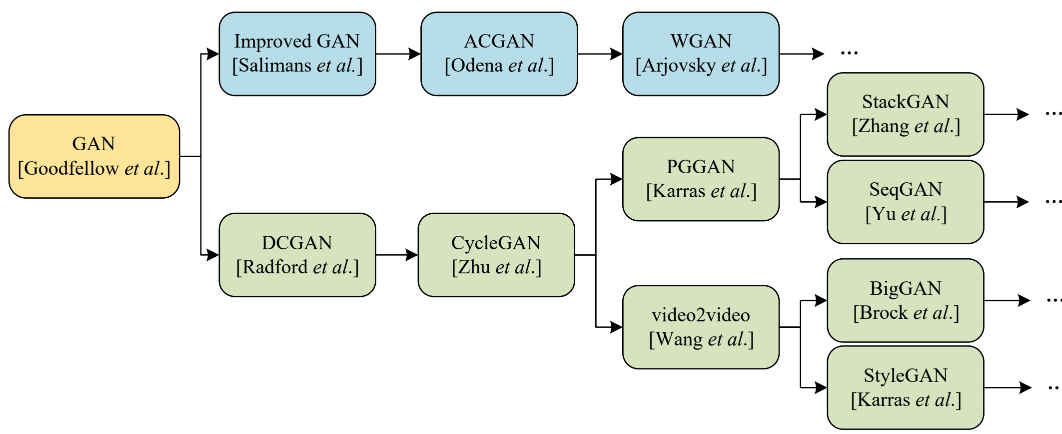

Fig. 2. The evolution of the major variants of the GAN.

inferred that the related loss function of G takes the form depicted as

$$\min V(D,G) = E_{z\sim p_z(z)}[\log(1-D(G(z)))]. \quad (2)$$

Thus, the optimized objective function could be defined as

$$\min_G \max_D V(D,G) = E_{x\sim p_{\text{data}}(x)}[\log D(x)] + E_{z\sim p_z(z)}[\log(1-D(G(z)))]. \quad (3)$$

The $\min_G \max_D V(D,G)$ is understood as follows: Equation (3) is maximized when updating the discriminator, whereas it is required to minimize (3) when updating the generator. To be detailed, on the one hand, we expect that the result $D(x)$, when the parameters of D are updated for the samples from the real distribution x, is as near to 1, that is, $\log D(x)$ should be as large as possible. For the data produced by noise z, we let $D(G(z))$ be as far as possible close to zero, hence the larger the value of $\log(1-D(G(z)))$, the better. Thus, discriminator D is maximized. On the other hand, when parameters of G are updated, the output can match (or at least approximate) the real data, which means the error can be as small as possible. To this end, G is needed to be minimized.

The generated data distribution of G should be compatible with the real data distribution, per the theoretical analysis, to attain the highest generated quality. To fit any sample distribution and identify whether a given new sample is in distribution, a well-trained GAN should be proposed.

B. *The Evolution of the GAN*

Since Goodfellow *et al.* [10] first introduced the GAN model, the GAN has quickly become the most popular generative model. After developing over eight years, the GAN has resulted in a number of valuable architectures, such as the deep convolutional GAN [21], StyleGAN [22], BigGAN [23], StackGAN [24], pix2pix [25], Age-cGAN [26], CycleGAN [27], etc. Fig. 2 shows the evolution of the major variants that generate adversarial networks. The upper branch mainly aims at adopting theoretical cues to improve the stability of the generated adversarial network and solve its training problems or consider different perspectives (such as information theory, model efficiency, etc.) to enrich its structure, whereas the lower branch mainly forms the variant structures and application scenarios of GANs in different application fields, such as actual picture conversion, text to image, image generation, video conversion, and other practical problems.

Specifically, in the upper branch, Salimans *et al.* [28] found techniques to maintain the training process of the GAN and proposed the ImprovedGAN. In the same year, Odena *et al.* [29] proposed a new discriminator to solve the problem of model collapse, called the ACGAN. Later, Arjovsky *et al.* [30] studied the initial problems of the GAN and improved the algorithm to innovate the WGAN. Currently, numerous newly-developed variants appeared to improve the model performance. In the lower branch, Radford *et al.* [21] proposed the DCGAN which adds the Convolutional Neural Networks (CNN) layer to transform an image. Zhu *et al*. [27] further proposed the concept of cycle-consistency loss to deduce similar image-transforming problems, named CycleGAN. Based on the CycleGAN, the PGGAN [31] and video2video [32] were then proposed to resolve the inaccuracy of the transformation of images and videos. For text generation, Zhang *et al*. [24] transformed images based on text-embedded multi-layer architecture and demonstrated the StackGAN, whereas Yu *et al*. [33] developed the SeqGAN to employ reinforcement learning to address the issue of gradient return in text production. In 2019, StyleGAN [22] and BigGAN [23] were developed to synthesize large-scale images and control potential spaces, respectively.

C. Main Applications in the Development of the GAN

During the development of the GAN, several studies have been conducted in the disciplines of medicine, natural language processing, and computer vision. GAN application is divided into four parts, including image generation, image conversion, image editing, and other applications.

1) Image Generation: The earliest research on the GAN was its application in image generation. Goodfellow *et al*. [10] proposed that the GAN can generate new image data sets for the MNIST handwritten digital data set, CIFAR-10 small image data set, portrait data sets, etc. A year later, Radford *et al*. [21] made a similar point by showing a model that generates a new case for the bedroom, showing a way to use the DC-GAN to grow stable GANs on a large scale. It also demonstrated the ability of the GAN (in potential space) to run vector operations, i.e., the result could be obtained by simply inputting the bedroom image. Jin *et al*. [34] trained the GAN in various ways and successfully generated animated heads, such as Japanese anime characters. Inspired by the example of animated characters, Hedge *et al*. [35] used GANs to generate game characters like Pokémon with little success. Later, scientists were not satisfied with generating simple objects or virtual images. In 2017, Karras *et al*. [31] demonstrated the case of generating face photographs, taking celebrities' faces as input. The generated cases featured celebrities' facial features and the photographs were very realistic, making people feel familiar with them. Additionally, Brundage *et al*. [36] also explored the rapid growth of the GAN in the field of portraiture over the past four years.

2) Image Conversion: The GAN is almost omnipresent in image conversion. Isola *et al*. [25] introduced how to use the pix2pix technique for image conversion. They successfully converted semantic images into city and architectural landscape images, satellite images into Google Maps, daytime landscapes into night landscapes, and black and white images into color images. Simultaneously, the famous CycleGAN was proposed to complete a large number of cases of image conversion, such as transforming pictures into artistic painting styles, horses into zebra pictures, summer scenes into winter scenes, and pictures of apples into pictures of oranges [27].

After the realization of image conversion technology, researchers began to study the conversion of words or semantics with images. Zhang *et al*. [24] proposed the StackGAN, which is capable of turning text descriptions of basic items (such as flowers and birds) into lifelike images. Reed *et al*. [37] further used this function and the GAN to complete venture-based transformation, and depict the position of objects (such as a bird), using enclosing boxes and key point prediction. Dash *et al*. [38] presented TAC-GAN which also converted text descriptions of flowers, birds, and other objects into images successfully. The conditional GAN was later used to generate real pictures based on semantic images or sketches, including pictures of city landscapes and home bedrooms, etc. Wang *et al*. [39] also introduced an interactive editor that can manipulate the generated pictures. Yoo *et al*. [40] adopted the GAN to generate images of clothes similar to the clothing atlas or online clothing stores based on photographs of models wearing clothes.

As image conversion techniques develop rapidly, scientists began to convert human features. Huang *et al*. [41] proposed a global and local sensing GAN to create a positive portrait image based on a face from a specified angle, which might be used in a system for face verification or recognition. GANs were also utilized to generate body images, which could structure different but coordinated body shapes. Specifically, cross-domain image conversions, such as transforming block digits into MNIST handwritten digits, or transforming celebrity photographs into emojis or animated expressions were then also realized [42].

3) Image Editing: In addition to generating or converting new images, GANs are often used for image editing. Attempts have been made to edit pictures, such as eliminating rain and snow in pictures by GANs. Perarnau *et al*. [43] used the IcGAN to reconstruct human images based on specific facial features such as hair color, hairstyle, expression, and even gender change. Moreover, Liu and Tuzel [44] proposed the CoGAN by utilizing features such as hair color, expression, and glasses to reconstruct human facial pictures. They also showed the generation of other images, such as scene pictures with color and depth changes. Brock *et al*. [45] proposed a facial photograph editor using a variety of variable autoencoders and GANs, which can quickly modify facial features, including modification of hair color, hairstyle, expression, and posture, and adding facial whiskers. The GAN also has key uses in facial beauty treatment where [26] introduced the method of using GANs to generate facial pictures of different ages and process facial aging.

As GAN image editing technology matures, scientists began to pursue high resolution and clarity of images. Brock *et al*. [23] developed the BigGAN to conduct large-scale GAN training of natural image synthesis and generated composite photographs which were almost identical to real photographs. The super-resolution GAN (SRGAN) model was also proposed to generate images with ultra-high resolution [46]. Bin *et al*. [47] used Conditional GANs to construct different versions of facial images with high quality and ultra-high resolution. Plus, the enhanced perceptual super-resolution network was also introduced to construct high-resolution street view images [48].

4) Other Applications: In addition to computer vision and image processing, as more scholars realized the advantages of the GAN, its range of applications is gradually increasing to video prediction, medical imaging, and three-dimensional reconstruction. As for video prediction, Vondrick *et al*. [49] focused on static factors in the scenarios, achieving continuous prediction of video footage of up to one second duration. In the medical imaging domain, a CycleGAN-based data augmentation model was proposed by Sandfort *et al*. [50] to boost the generalization of CT segmentation. With the potential to accelerate the data collection process, Yu *et al*. [51] used a Conditional GAN-based deep learning framework for de-aliasing and reconstructing MRI pictures from significantly under-sampled data. Furthermore, the GAN can complete or reconstruct the three-dimensional shape of objects, which is the improvement and extension of three-dimensional reconstruction technology. For example, a 3D-VAE-GAN model combines volumetric convolutional networks and GANs to generate 3D objects in probabilistic space [52]. Gadelha *et al*. [53] also used GANs to develop 3D models of objects based on 2D object images from multiple perspectives.

III. GANs in Autonomous Driving

Autonomous driving has witnessed significant advancements, largely due to the rapid development of deep learning techniques in recent years [54]–[58]. AVs have the potential to revolutionize transportation by freeing up manpower, reducing traffic congestion, lowering accident risks, and minimizing energy consumption. As a result, they are poised to capture a substantial market share in the near future.

Generally speaking, as depicted in Fig. 3, a standard autonomous driving system can be broken down into five key components: sensing, localization, scene understanding, decision-making, and control [59]. Deep learning, a powerful solution for autonomous driving, plays a critical role in most of these components. Numerous researchers have explored various deep learning approaches to enhance the capabilities of AVs. Specifically, Bojarski *et al*. [60] proposed an end-to-end deep learning platform for AVs based on CNNs, which demonstrated that CNNs could effectively learn to drive by mapping raw pixels from a single front-facing camera directly to steering commands. In addition, Shalev-Shwartz *et al*. [61] also introduced deep learning-based approaches in multiple aspects, such as perception, high-precision maps, and decision-making processes.

After years of iterations of autonomous driving technology, problems with traditional models such as CNN, LSTM, etc. arise. As exemplified, the error tolerance of autonomous driving application scenarios is extremely low, but there is a lack of image data to train autonomous driving models. Also, it is necessary to understand and predict the intention of pedestrians in real-time, but the current method is not accurate enough. Besides, the traditional neural network model is prone to attack and thus produces error output. To overcome these deficiencies, GANs are then introduced to autonomous driving scenarios, as Santana first found that GANs can generate images consistent with the distribution of actual traffic scenes, which can be applied to tasks requiring semi-supervised or unsupervised learning in automatic vehicles, and the constantly updated video frames of actual scenes can optimize the

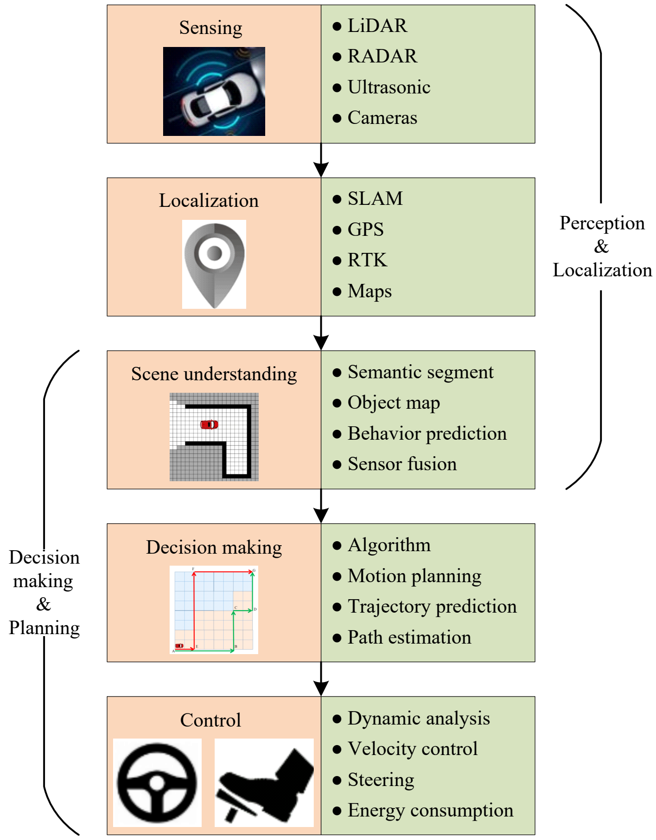

Fig. 3. A common decomposition of an autonomous driving system.

generator of GANs in real time [62]. Currently, GANs applied in autonomous driving are divided into three types: generation and transformation of perceptual data, trajectory prediction based on video scenarios, and security of target detection under jamming as illustrated in Fig. 4.

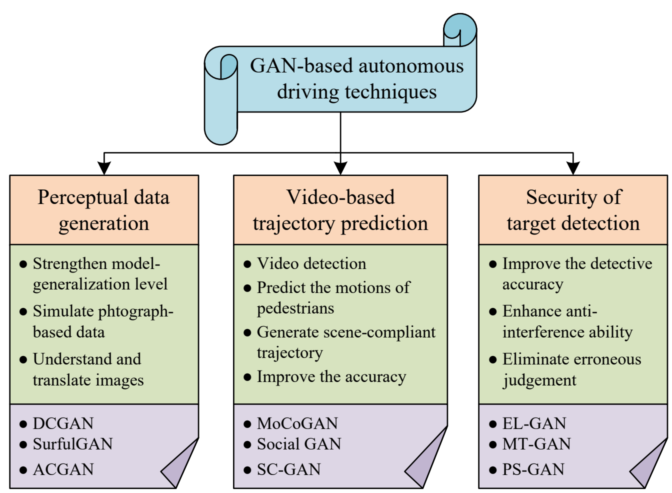

Fig. 4. Major applications and variants of GAN-based autonomous driving techniques.

A. Generation and Transformation of Perceptual Data

To a large extent, the development of autonomous driving hinges on being able to simulate and replay complicated and varied traffic scenarios. Its application scenario has a very low tolerance for errors, which requires the model to have a strong generalization ability. Therefore, a substantial number of pictures is required to train the perception model in the realm of autonomous driving. The image generation models generated for such requirements mainly include the variational autoencoder (VAE) and the GAN. Compared with the VAE, the GAN directly collects samples from real samples and generates new samples whose distribution approaches the real sample's distribution continuously in the training process, rather than pre-set data distribution in the hidden space. Therefore, GAN-generated samples are more like actual pictures and are widely adopted in recent years.

A typical application of the GAN is data generation, which can be used in the simulation of camera data in the autonomous driving domain. For example, SurfelGAN reconstructs realistic camera pictures for innovative self-driving vehicle positions and orientations and objects in the scene based on the limited lidar and camera data [63]. Radford *et al*. [21] proposed the deep convolutional GAN (DCGAN) model, which uses a commutated convolutional layer to generate images and introduces a CNN to distinguish the actual and produced images, laying the structural foundation for advanced GAN models. Odena *et al*. [29] proposed an auxiliary classifier GAN (ACGAN), which added label information to the generator input and reconstructed the label in the discriminator, improving the quality of image generation.

In the perception and decision-making process of autonomous driving, another typical application direction of the GAN is image understanding and translation, which converts the data from one domain to another, such as transforming easily collected daytime scenarios to night-time scenarios which are challenging to obtain, or converting good weather to bad weather. This can be utilized to increase model generalization and the quality of the data. The traditional automatic driving image translation approach changes the images by exchanging attribute codes after encoding the content and attributes of the global image. The exemplary models among them are Disentangled Representation for image-to-image translation (DRIT) and multimodal unsupervised image-to-image translation (MUNIT). Furthermore, Amodio and Krishanswamy [64] replaced the cyclic consistency loss by adding a third network (Siamese) to reduce model complexity and training costs. In more specific tasks of image understanding and translation, a weakly supervised structure-aware image-to-image translation network was developed by Huang *et al*. [65]. It consists of parsing nets for the two domains, respectively, and shows crucial improvement in the detection accuracies. Xu *et al*. [66] empirically analyzed and proposed novel solutions to assess the reliability of GAN-generated data for training and validating vision-based perception modules for AVs, such as object detection and scenario classification.

B. Trajectory Prediction Based on Video Scenarios

Since there are multiple objectives in an autonomous driving environment, comprehending and anticipating the intentions of pedestrians is crucial for the safety of driving. In practical application, AVs will forecast the future course of other targets in the scenario, primarily other cars and pedestrians based on the video collected by the camera, which is also an important direction of related research into autonomous driving [67], [68].

In terms of video detection, other workers proposed a new video-to-video synthesis method based on a generative adversarial learning framework [32] while MoCoGAN suggests breaking down motion and content to generate video for video detection [69]. Moreover, Denton [70] proposed the DRNET, learning to untangle image representation from videos. The GAN is also used for video prediction and video retargeting. Specifically, considering the relatively static background color and dynamic object movement of the composition of a video, the video GAN (VGAN) utilizes a two-stream generator to predict the next frame with the moving foreground generator of the 3D CNN and adopts the static background generator of the 2D CNN to keep the background still [49]. The Pose-GAN mixes the VAE and GAN to estimate the future motions of objects using the present postures and the hidden representations of past postures [71]. In addition, video-based GANs consider both spatial and temporal modeling or the movement between each subsequent frame in a video sequence. To this end, DVD-GAN introduces an extendable, video-specific generator and discriminator architecture while being able to generate longer and higher-resolution videos based on the BigGAN architecture [72].

Predicting the trajectory by the GAN is mainly based on video images detected by vehicles, as Mozaffari *et al*. [73] reviewed the most advanced deep learning-based methods for predicting vehicle behavior, and proved the potential of the behavior prediction function of an AV. Following the popularity of general GAN designs, many studies using adversarial models for trajectory prediction have been promulgated. In chronological order, Gupta *et al*. [74] first introduced the GAN to the task of pedestrian trajectory prediction and proposed the Social GAN to predict socially plausible futures. This newly-developed method promotes a diversity of predictions, encourages diverse predictions and combines tools from sequence prediction to solve the problems of pedestrian path prediction. Recurrent neural networks (RNNs) and mixture density networks are combined to produce probabilistic multimodal predictions in the probabilistic crowd GAN, which Eiffert *et al*. [75] proposed as an extension of recent trajectory prediction work. This showed advancements over the most advanced trajectory prediction techniques. Zhang *et al*. [76] proposed a new adversarial assault that alters typical vehicle trajectories to increase the prediction error and study the adversarial robustness of trajectory prediction algorithms. In the same year, Gómez-Huélamo *et al*. [77] explored how attention affects generative models for motion prediction, considering both physical and social settings to determine the most plausible trajectory.

Furthermore, as the discriminator in most of these GAN-based models does not condition the scene context image, it is unable to discern between a trajectory that follows the actual ground truth and a trajectory that is somewhat similar but does not comply with the scene. To tackle the challenge, the scene context image is added as an input to the discriminator in the Social-BiGAT algorithm, which then concatenates the flattened scene context features with the trajectory embeddings to conduct classification. The discriminator employs a CNN to extract features from the scene context image [78]. To improve prediction accuracy even more, the SC-GAN method was recently carried out by Wang *et al*. [79] which stacks the occupancy grids with the scene context image along the channel dimension after differentially converting the trajectory input into a series of 2D occupancy grids., making the prediction more scene-compliant and accurate. A more recent research proposed by Zhao *et al.* [80] introduced a life-long learning framework based on GANs for efficient and high-precision direct trajectory planning (DTP) tasks, developing a lightweight GDTP model that maps initial/terminal states and control action sequences onto trajectories.

C. Security of Target Detection Under Jamming

As mentioned, autonomous driving is the future of ITSs, but people could not trust this complex technology until its safety has been fully tested. As a field with high safety requirements, the perception of automatic driving environments has numerous target detection and segmentation tasks, but the neural network model is vulnerable to attacks and may result in erroneous outputs, leading to serious traffic accidents. Therefore, the importance of increasing precision and anti-interference ability of target detection in autonomous driving models cannot be overemphasized.

The appearance of the GAN provides a new idea for target detection when driving automatically. In recent years, some studies have applied the GAN to target detection or to deal with countermeasures encountered by automatic driving, some of which have achieved acceptable results. For example, the embedding loss-driven GAN (EL-GAN) was introduced by Ghafoorian *et al*. [81] to address the discussed issue utilizing an embedding loss. This method produced more stable training since it had richer discriminative information and benefited from simultaneously viewing "real" and "fake" predictions.

In the actual scenario, the most common interference is the object being distorted or obscured, making it urgent to innovate a method of detection that is not affected by distortion or occlusion. One approach is to use a data-driven strategy to collect large data sets with examples of objects under different conditions. The classifier could iterate all deformation and occlusion instances in the data set to learn the invariance. Nonetheless, some deformations and occlusions are so rare that they seldom happen in practice. To solve this problem, Wang *et al*. [82] trained an adversarial network to generate samples with occlusion and deformation that were difficult to be classified by the target detector, and the accuracy of target recognition was improved by 2.6%. Li *et al*. [83] established the Perceptual GAN model, which reduces the difference between tiny and big objects to increase small target detection, improves small target detection, and outperforms the most cutting-edge technologies in recognizing small things, including traffic signs and pedestrians. In addition, to retrieve detailed information for more precise identification, Bai *et al*. [84] suggested an end-to-end multi-task generative adversarial network (MT-GAN) that can up-sample small fuzzy images into fine-scale versions thereof.

Research in recent years has further shown that real-world disturbances are aggressive even when captured on camera. The disturbances may lead to erroneous judgment of an AV, resulting in severe traffic accidents. As exemplified, if some stickers or graffiti are attached to a stop sign, it can easily be mistaken for a speed-limit sign by the traffic sign recognition system. Moreover, with pedestrians sometimes hiding in front of the model when distracted by confrontations, and confrontations causing the model to "ignore" roadblocks, researchers found several ways to trick Tesla's autopilot system. A bug in the Tesla model S's automatic wiper and lane recognition system was discovered by researchers who used a noise-generating function to create and patch the required countermeasure image and successfully activated the Tesla's automatic wiper. To cope with this problem, Liu *et al*. [85] presented a perceptual-sensitive GAN (PS-GAN) to create adversarial patches using a GAN that can learn and approximate the distribution of initial instances. Zhang *et al*. [76] proposed a new adversarial attack that increased the prediction error by more than 150%. They also analyzed data augmentation and trajectory smoothing as potential mitigation measures.

IV. GANs in Traffic Flow Research

The objective of traffic flow research is to assess the relationships between traffic flow, velocity, and density by analyzing the interactions between different traffic participants (such as vehicles, pedestrians, and drivers) and infrastructure (such as highways and signal control devices). Traffic flow research is, in essence, based on empirical investigations that depend on accurate measurements of actual data [86].

Traffic flow research is a task dealing with various types of time-series data, so several scholars have contributed to this subject by introducing deep learning-based techniques like the RNN and its variant, long-short term memory. Nonetheless, classic models can hardly reflect fluctuations in real data, resulting in a reduction in accuracy. Besides, there are roads with few vehicles and pedestrians or not enough sensors, and sometimes parts of the traffic flow data are missing for technical reasons, making it impossible for the usual models to predict. To this end, the GAN is a feasible way to handle data generation more effectively by addressing the issue of traffic data sparsity as well as improving the predicting accuracy.

When the GAN is applied in traffic prediction tasks, Generator G is typically adopted to generate future traffic circumstances based on historical observations. After that, Discriminator D is trained using both the created data and the real data to come. Upon training, G discovers the distribution of real traffic flow data that may be used to forecast future traffic conditions using a lot of historical data. Additionally, depending on the specific traffic tasks, the generator or discriminator of the GAN can be any model, such as RNNs and Seq2Seq.

Generally, the applications of the GAN in traffic flow research mainly consist of two parts, traffic spatio-temporal (ST) data forecasting and traffic sparse data completion, where approaches of model evaluation are also important. The major applications and variants of GAN-based traffic flow research are illustrated in Fig. 5.

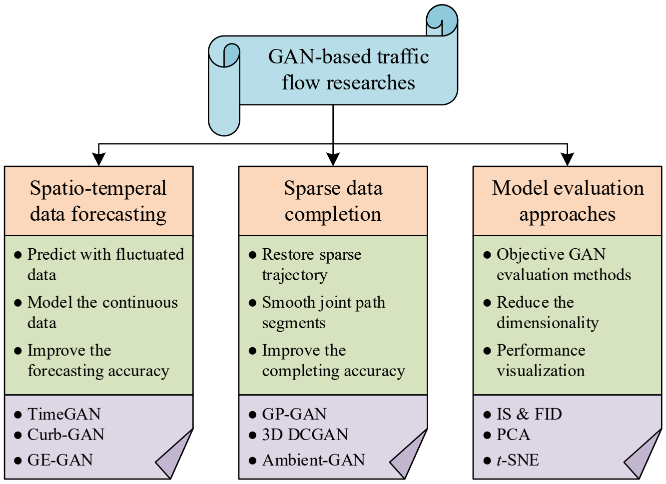

Fig. 5. Major applications and variants of GAN-based traffic flow research.

A. Traffic Spatio-Temporal Data Forecasting

Currently, growing interest has been shown in creating effective urban applications employing spatio-temporal data, such as air and water quality forecasting, crowd flows prediction, vehicle velocity prediction, and demand prediction have drawn increasing attention. Real-time analysis and precise prediction abilities are warranted for these applications. For instance, a reliable model for predicting taxi demand can assist drivers in increasing their income by making it simple for them to find passengers. It can also reduce overall traffic and carbon emissions in a city.

Spatio-temporal data related to traffic flow is generally divided into discrete time series and continuous time series. Data points in discrete time series are separated by time intervals, such as the subway passenger flow data, which can be collected once a day, whereas each moment in a continuous time series has a corresponding data value, such as the real-time traffic flow at a given crossroad. Due to the large number of zero gradients, the GAN is difficult to generate discrete data, that is, the distribution on the discrete object is not differentiable concerning its parameters. In the actual training process, the generator is guided by the loss gradient of the discriminator output, which changes the output of the generator slightly, bringing it closer to the desired output.

The common models for data-driven time-series prediction are mainly neural network models with various structures, including linear regression, random forest, etc., but the character of such models is that they do not reflect fluctuations in real data. Fortunately, the GAN is essentially designed for continuous data modeling, and several GAN-based approaches to ameliorate this problem: A one-step probabilistic prediction model ForGAN whose core is the Conditional GAN was developed [87], and Saxena and Cao [88] presented a new model based on the deep generative adversarial networks (D-GAN) to predict the spatio-temporal characteristics in unsupervised learning more accurately. To improve the accuracy of prediction, some scholars combined the GAN with self-attention mechanisms or reinforcement learning (RL). As exemplified, the prevailing TimeGAN, which blends the control offered by supervised training with the adaptability of an unsupervised paradigm, was proposed, achieving excellent performance in generating realistic time-series data [89]. In 2021, Zhang *et al*. [90] developed the SATP-GAN, which combines self-attention, the GAN, and RL mechanisms to predict traffic, where the GAN is used to generate new prediction data and RL is used to adjust parameters, showing an improvement in accuracy by 6.5%−9.1%.

In the application of traffic spatio-temporal data prediction, the Curb-GAN proposed by Zhang *et al*. [91] allows flow estimation of continuous periods based on dynamic traffic demands and enables planners to estimate the planning schemes more accurately before implementing them. The TrafficGAN model was recently offered as a solution to the problem of traffic estimation, but it ignores the temporal autocorrelations of traffic status and only provides snapshot estimates [92]. In addition, Xu *et al*. [93] introduced a road estimation framework called GE-GAN where data from neighboring links were introduced to forecast the traffic states using graphs to illustrate the road network. Besides, attempts have also been made by Jin *et al.* [94] who introduced PL-WGAN, a short-term traffic speed prediction approach using WGANs for urban road networks, which outperforms state-of-the-art deep learning models in predicting link traffic speed. Lv *et al.* [95] utilized GANs in parallel transportation systems for traffic data generation, modeling, prediction, and control, highlighting their potential as a specific algorithmic support.

B. Traffic Sparse Data Completion

With the progress of sensor technology and the popularity of traffic software, countries all over the world have laid much traffic-data acquisition equipment in the road networks, such as underground induction coil detectors, speed measurement radars, etc., while people also widely rely on map navigation and other functions to travel. After long-term accumulation, we have come to possess massive historical traffic data, which provides data support for ITSs.

In traffic flow research, a crucial topic is to estimate accurate travel time and energy consumption. Nonetheless, many links of roads may not be traversed by any GPS-equipped vehicles at this time, and neither can we select a trajectory precisely spanning a query path. Besides, for a path segment with trajectories, combining a number of trajectories to calculate the time of the whole journey causes information losses at each node [96]. To fill the aforementioned gaps, researchers have proposed different approaches, such as directly ignoring missing values which is commonly used in discrete time-series processing [97]. Yu *et al*. [98] not only paid attention to its own temporal evolution but also focused on the adjacent temporal evolution, especially the effects between temporal indices with influences. For example, a certain section's traffic volume is directly related to the traffic volume of upstream and downstream sections. In [99], considering the external attributes of multi-source time-series data, the explainable correlation relationship is constructed, and the data are completed by combining multi-source timing sequences and their relationships.

Nevertheless, since the existing methods of data completion are very rough, a better method is developed to complete the data more accurately. Considering the image hybrid restoration function of GANs as Wu *et al*. [100] proposed the GP-GAN, which blended real photographs and successfully integrates photographs of fields, mountains and large objects, and Pathak *et al*. [101] adopted a text encoder to repair and complete pictures, that is, to fill a missing part of an image, we reasonably infer that the GAN can also perform sparse trajectory completion. To the limit of our knowledge, few such studies have utilized the GAN for traffic sparse data completion. Among them, a GAN model based on the Ambient-GAN and 3D-DCGAN was proposed to infill the missing data pertaining to the traffic flow. By using the Beijing taxi traffic data for simulative experiments, the network can recover the missing data to a fairly high level compared with the DeepST, Bayesian CP, and other algorithms [102]. Plus, Wang *et al*. [102] demonstrated a road network traffic data supplement method, which uses road network data based on spatio-temporal information compensation as the input for training, to realize the reconstruction of missing data of two-dimensional infographics in road networks [103]. A more recent approach proposed a semantic-guiding adversarial network (TrajSGAN) for generating human trajectories by combining sequence-based and image-based approaches, achieving superior performance in simulating human mobility and studying epidemic diffusion [104].

C. Approaches of Model Evaluation

Any model needs evaluative indices to judge whether the model is good or bad. For the GAN, if the generated sample is similar to the real sample, subjective evaluation has the following problems: when a vast number of images are produced, observing a tiny part of images may not represent the quality of all images. Meanwhile, when the generated pictures are very real, it is subjectively considered to be a good GAN, but there may be overfitting phenomena that are difficult to detect with the naked eye.

Since problems occur in subjective evaluation, scholars have proposed objective evaluation methods of the GAN, such as the inception score (IS), frechet inception distance (FID), etc. [105]. Specifically, IS mainly evaluates the following two aspects: definition and diversity, which can be calculated, thus

$$IS\,(G) = \exp(E_{x\sim p_g} D_{KL}(p(y|x)\|\hat{p}\,(y))) \tag{4}$$

where $x \sim p_g$ represents a picture produced by a generator, and $(y|x)$ means the probability distribution of each category for image x. Then, it takes the average of the probability distribution vectors of each image to find the edge distribution of the generated image on all categories, and the specific formula is given by

$$\hat{p}\,(y) = \frac{1}{N}\sum_{i=1}^{N} p\left(y|x^{(i)}\right). \tag{5}$$

D_{KL} means to find the divergence of KL, which is formulated thus

$$D_{KL}(P\|Q) = \sum_{i=1}^{N} P\,(i)\log\frac{P\,(i)}{Q\,(i)}. \tag{6}$$

IS cannot reflect overfitting or train classification models on one data set to evaluate generating models trained on another data set. On the other hand, the FID can determine the separation between the generated feature vector and the actual image, which can be formulated as

$$FID\left(P_r, P_g\right) = \mu_r - \mu_g + T_r\left(C_r + C_g - 2\left(C_r C_g\right)^{1/2}\right) \tag{7}$$

where μ_g and C_g represent the mean and variance of the feature of the generated image, respectively, while μ_r and C_r are the mean and variance of the feature of the real image, respectively. The FID method is more robust and computationally efficient.

Compared with image data, much more dimensions of data exist in traffic flow research, but it is challenging to accurately estimate high dimensional time-series data at present. However, dimensionality reduction is required initially. In terms of qualitative evaluation of the GAN based on time series, the principal component analysis (PCA) and *t*-distributed stochastic neighbor embedding (*t*-SNE) are usually used for dimensionality reduction visualization analysis to compare the similarity between the generated data distribution and the original data distribution [106].

PCA is a statistical analysis technique that uses a linear transformation to reduce the original many variables to a small number of new comprehensive variables. These new variables sometimes referred to as principal components, reflect the information of the original variable without losing or greatly reducing the loss of that information because they are independent. After dimensionality reduction, the PCA aims to maximize the internal information of the data and analyze the significance of the direction by calculating the variance of the data in the projection direction. The basic mathematical principle is as follows:

$$\max_{\omega} \frac{1}{m-1}\sum_{i=1}^{m} \omega^T (x_i - \bar{x})^2 \tag{8}$$

where m is the number of data involved in dimensionality reduction and x_i denotes the concrete vector representation of the random data i and $\bar{x}$ is the average vector of all the data involved in dimensionality reduction. A column vector matrix defined as W contains all feature mapping vectors. An algebraic linear transformation of this matrix yields an optimized objective function

$$\min_{\omega} tr\left(W^T A W\right), \ \text{s.t. } W^T W = I \tag{9}$$

where tr is the trace of the matrix, and A denotes the covariance matrix, expressed as

$$A = \frac{1}{m-1}\sum_{i=1}^{m} (x_i - \bar{x})(x_i - \bar{x})^T. \tag{10}$$

Therefore, the output of the PCA is $Y = W^T X$, and the top k greatest eigenvalues of the data covariance matrix are represented as column vectors in the optimal W, which can reduce the original dimension of x to k dimensions.

In the *t*-SNE, the heavy-tailed distribution in the low-dimensional map is introduced as a one-degree-of-freedom *t*-distribution, which is identical to a Cauchy distribution [107]. In this case, the joint probabilities q_{ij} can be defined as

$$q_{ij} = \frac{\left(1 + y_i - y_j{}^2\right)^{-1}}{\sum_{k \neq l}\left(1 + y_k - y_l{}^2\right)^{-1}} \tag{11}$$

where y_i, y_j, y_k and y_l are map points. Then, when P and the t-based joint probability distribution Q are compared, the gradient of the Kullback-Leibler divergence is given by

$$\frac{\partial C}{\partial y_i} = 4 \sum_j (p_{ij} - q_{ij})\left(y_i - y_j\right)\left(1 + y_i - y_j{}^2\right)^{-1} \tag{12}$$

where C is the cost function. In summary, the t-SNE is better than linear reduction and other nonlinear dimensionality reduction methods when processing complex data. However, it is extremely difficult to establish the network when there are too many samples of these evaluation approaches, and the gradient descent is too slow. Therefore, it is necessary to improve on existing methods and propose even better new methods.

V. GANs in Traffic Anomaly Inspection

Anomaly inspection of traffic infrastructures and status exerts a significant influence on transportation safety. Abnormal traffic conditions will greatly decrease traffic efficiency and even cause accidents, so it is necessary to detect and monitor them. If anomalies occur, alarm and rescue will be conducted to remove the inconvenience caused by the incident as soon as possible and restore normal traffic.

The anomaly inspection in the transportation domain usually contains two parts, infrastructure detection, and traffic status monitoring. On the one hand, the traditional detecting techniques of infrastructures include a rangefinder, theodolite, level, total station, and other measuring methods, but with the rapid development of electronic devices and space technology, many advanced techniques have appeared [108]. On the other hand, conventional traffic status monitoring methods mainly consist of electromagnetic induction loop type and wave type, which detect the vehicle information based on the frequency change of reflected waves when the vehicle passes by. However, these approaches cannot provide comprehensive traffic information, so it is restricted by significant inherent limitations [109].

To overcome the above problems, some scholars have adopted deep learning-based approaches to traffic anomaly inspection. Nonetheless, it is difficult to define abnormal traffic behavior or acquire anomaly information in the actual scenario resulting from numerous factors, such as weather conditions, illumination changes, etc. [110]. Fortunately, GANs are capable of modeling high-dimensional data distributions and are useful for anomaly-detection jobs involving complicated data sets, which have displayed cutting-edge performance. As a result, most existing models and theories built on GANs are appropriate for detecting traffic anomalies. Fig. 6 shows the major applications and variants of GAN-based traffic anomaly inspection.

A. Damage Detection of Roads and Bridges

Following structural deterioration or exterior damage, transportation infrastructure, such as roads and bridges, will

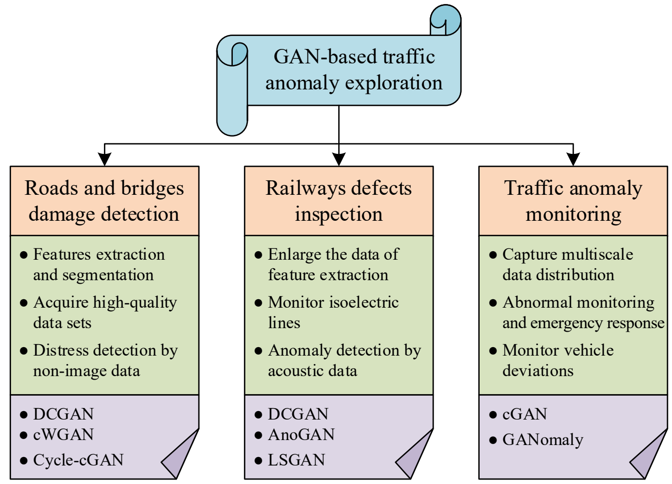

Fig. 6. Major applications and variants of GAN-based traffic anomaly inspection.

unavoidably experience surface damage. Therefore, regular inspection is required to ensure the safety of transportation infrastructures. There are various forms of road defects, among which cracks are the key parts of road and bridge detection. The classic approach relies on the segmentation achieved by the gray difference between the crack and the normal part, but it is easily influenced by noise, leading to weak generalization capacity [111]. As artificial intelligence progresses, deep learning exhibits significant generalization capability and robustness in extracting crack picture features. Nevertheless, deep learning-based crack detection still faces some challenges; for example, cracks come in many different forms and shapes, and their characteristics are complicated. Moreover, low-contrast crack images are prevalent, making it challenging to segment and extract cracks. Besides, the issue of data imbalances prevents the quick adoption of innovative methods due to the lack of excellent data sets.

To extract and segment features accurately, Zhang *et al*. [112] suggested an end-to-end training crack-patch-only supervised GAN to solve the problem using a huge crack image fed into an asymmetric U-shaped generator. Another unsupervised visual inspection framework proposed by Zhai *et al*. [113] utilized the DCGAN to learn the feature representation of typical surface samples. The first three convolution layers were employed as feature extractors to enable multi-scale feature fusion and segmentation based on the trained discriminator's sensitivity to the anomalous regions, demonstrating the efficiency and resilience of the road crack database [114]. Additionally, Gao *et al*. [115] introduced a GAN-based approach to road crack segmentation. The segmentation performance was examined on the CFD and AigleRN data sets using the U-Net and cross-layer concatenate networks as generators to successfully prevent loss and gradient evaporation.

Moreover, aiming at obtaining high-quality data sets, the GAN can provide an improved crack detection strategy, which can extract feature distribution from a tiny data set of actual crack pictures and generate many similar pictures to enhance crack-detecting efficiency. Generally, most of the existing approaches have been verified in public data sets, but detection can only be implemented in simple scenarios due to the relatively simple environment of image acquisition. To this end, Mei and Giil [116] gathered pictures of road cracks in various natural scenes and produced EdmCrack600 data set with pixel-level labeling. Huang *et al*. [117] combined cWGAN and a linked domain graph and provided high-quality data sets for road crack detection. For fracture detection, a self-supervised Cycle-cGAN was presented, which transforms crack images into ground truth-like images with comparable structural patterns [118].

Furthermore, while most contemporary GAN-based fracture detection methods concentrate on pictures, the GAN is also used to detect distress in non-image data. To be specific, Mao *et al*. [119] applied the GAN to identify distress in bridge health by transforming time-series data into Gramian Angular Field (GAF) images to enable network training. Besides, laser scanning and ultrasonic detection remain the standard technologies for locating bridge faults currently, whereas attempts have been made to improve the technique based on the GAN. For example, Lee and Shin [120] proposed a data-generating strategy depending on the data imbalance observed before and after the occurrence of damage, which can analyze complicated data independently.

B. Railway Defect Inspection

Once the defects such as track and infrastructure damage occur in railway transportation systems, it may result in serious accidents. To this end, it is of great necessity to regularly inspect tracks and allied systems for flaws and problems, while rail tracks, catenaries, fasteners, bogies, and other such objects are among the main detection objects in routine inspection tasks. Since the railway high-voltage circuit is dangerous and may cause widespread safety problems, manual detection is not feasible, and it is necessary to conduct reliable real-time visual monitoring. A viable substitute is to feed real-time video pictures into a trained deep neural network for real-time error detection. Nevertheless, this method still has certain inescapable challenges. As exemplified, given the size of the railway system that needs to be maintained and the amount of time and personnel needs to keep tracks in good condition, visual inspection equipment that is highly mobile and reliable is needed. Moreover, the external environment severely influences inspection tasks, and the captured photographs are frequently tainted with grease and dust. Plus, images of railway defects are limited and difficult to obtain.

Since the flaws and data of anomaly are limited, it is impossible to apply supervised deep learning techniques based on tagged training data. Fortunately, the unsupervised GAN learns the distribution of features from extensive unlabeled data sets, which have a superior preponderance in railway defect inspection. Yang *et al*. [121] who generated data features from unlabeled data using common catenary photographs as training data, proposed a DCGAN-based anomaly-detection approach to detect bird nests. The Catenary Supporting Components developed by Lyu *et al.* [122] combine a CNN with a GAN to detect flaws and respond to them and show great performance in judging insulator anomalies.

Moreover, the isoelectric line is a significant section of the high-speed railway, which is prone to failure for automobiles running for a long time. To monitor the isoelectric lines, the Faster R-CNN was adopted to extract features from input images and locate isoelectric lines accurately and the DCGAN was introduced to obtain the antagonistic illustration of features, whereas isoelectric line flaws were found by the anomaly classification criterion [123]. A method that integrates a CNN and an AnoGAN for unsupervised anomaly detection was created by Xue and Gao [124], dealing with the real-time monitoring of turnout problems.

Plus, researchers employed the GAN to detect abnormalities through audio data in addition to image-based railway defect inspection. Wang *et al*. [125] developed an enhanced regularized least-squares GAN (LSGAN) for acoustic emission signal denoising to combat the noise interference brought on by the mechanical contact between wheels and rails. Compared with traditional denoising methods, the LSGAN retains more details of crack signals and eliminates most statistical and mechanical noise.

C. Traffic Status Monitoring

Traffic status monitoring, another perspective topic in traffic anomaly inspection, plays a crucial role in ensuring traffic safety. The traffic anomaly is affected by such factors as the high-density traffic flow, weather conditions, changes in camera perspective and illumination, target occlusion, the low resolution of data acquisition, and lack of real scene data, making it more challenging for accurate and timeous detections [126].

The cGAN which jointly learns the creation of latent space and high-dimensional image space is a feasible way for anomaly inspection. Akcay *et al*. [127] employed the encoder-decoder-encoder model to learn the data distribution of normal samples to better measure the anomalies of the new data distribution by the greater distance between the generated picture and the latent vector. To solve the problem of limited samples, attempts have also been undertaken to obtain the multiscale distribution of the normal data in the image space [128].

In addition, once mechanical failure breaks down on a running vehicle, traffic congestion or even serious accidents may happen. Therefore, abnormal monitoring and emergency response are very necessary. To replicate the anomaly inspection process, Sun *et al*. [129] created a defect prediction approach and tested it using data from Isuzu vehicles. Moreover, Qiu *et al*. [130] established a multi-pattern recognition system based on the GAN to assess the score difference between the projected signal and the actual signal to accomplish unsupervised driving anomaly detection. The newly-developed approach is capable of monitoring vehicle departures from the typical driving track more precisely.

VI. Challenges and Promising Directions

The original goal of the GAN was to generate deceitful pictures, which has achieved spectacular success. Although the GAN has been extensively adopted in transportation-related research, numerous challenges still occur during the experiment and provide broad prospects for future studies. In this respect, this research outlines the difficulties encountered when implementing GAN-based solutions for transportation issues and highlights the following possible prospective research topics:

1) Enhancing the Model Performance With Reinforcement Learning in Autonomous Driving: Due to the inherent instability of the GAN, it is difficult to train, and the model suffers from some potential issues, such as non-convergence, gradient diminution or disappearance, poor interpretability, and mode collapse. This instability can lead to continuous oscillations, causing divergence, a situation that could seriously impact autonomous driving's safety. Goodfellow *et al*. [10] proposed that GANs can be smoothly integrated into the reinforcement learning (RL) framework. As exemplified, while addressing programming problems with RL, GANs can learn the conditional probability distribution of actions, thereby enabling the agent to choose logical outcomes based on the responses generated from different actions. Current research has successfully incorporated GANs into various RL methods, such as imitation learning, policy gradient algorithms, and actor-critic algorithms. This combination is expected to enhance the efficiency and stability of GAN-based models, consequently improving the response speed and accuracy of autonomous driving systems.

2) Improving the Accuracy of Spatio-Temporal Prediction in Traffic Flow Research: Applying GANs of image and video prediction is not suitable for predicting the spatio-temporal data of traffic flow, such as bus demand, metro passenger flows, etc. The reason for this is that while the image production process can account for differences in appearance between input and output images, it is unable to handle spatial alterations. Also, the video prediction algorithm considers spatial changes, although it heavily depends on the prior image in the image sequence to anticipate the subsequent image, that is, the time dynamics are not considered. At present, the application of the initial GAN to traffic spatio-temporal data prediction is ineffective. The potential reasons are demonstrated as follows.

a) The GAN with input noise hardly covers possible future space adequately;

b) The GAN cannot make good use of additional information not considered by the existing model, such as the weather;

c) The GAN does not support inference networks, but understanding inference networks is a necessary step in improving the model.

Therefore, there is much potential to improve the accuracy of spatio-temporal prediction in traffic flow research.

3) Predicting Activity-Travel Patterns in the Field of Travel Demand Analysis by GANs: In addition to the prediction of traffic flows, another promising research direction where GANs could be applied is in predicting activity-travel patterns in the field of travel demand analysis. Traditional methods in this field often rely on complex models, which may not fully capture the complexity and variability of human activity-travel behavior. For example, Liao *et al.* [131] incorporated space-time constraints and activity-travel time profiles in a multi-state supernetwork for individual activity-travel scheduling, demonstrating the challenges of calibrating such models.

Applying GANs to travel demand analysis could provide a data-driven, flexible approach for modeling and predicting human activity-travel patterns. Their ability to learn intricate, high-dimensional, and non-linear relationships can improve the accuracy of demand forecasts and reduce calibration needs. Furthermore, GANs could also help in the generation of synthetic populations and realistic activity-travel schedules, aiding the development of more robust and efficient transportation planning strategies. As such, the application of GANs in travel demand analysis represents an exciting and valuable area of future research within the context of intelligent transportation systems.

4) Completing Sparse or Missing Traffic Flow Data With GANs: As mentioned in Section IV, there may not be any GPS-equipped automobiles on many route segments at this time, and some traffic flow data have been lost for some reason, blocking some potential research. Additionally, due to the trade-off between the number of trajectories traveling the path and its length, losses appear during combining the trajectory segments. The completion of sparse or missing data is a crucial topic in current traffic flow research, however, compared with image completion, the sample length of traffic flow data generally varies and changes significantly with time. Although the GAN is expected to augment the sparse data of some sections, which is likely to produce excellent results, data preprocessing is necessary and study-worthy at the machine learning level.

5) Designing Evaluation Indices for GANs Generating Time-Series Data: Researchers have reached a certain consensus on how to evaluate the produced distribution inferred from the distribution of training data using image-based GANs. However, for time-series GANs, there is no consensus on the assessment indicators for generating data due to the small number of publications that have been published. Researchers have developed different approaches, such as the PCA and t-SNE mentioned in Section IV, but none are yet universally viable. Worse still, GANs can also generate fake information, which may have a significant influence on security. Therefore, future research should design evaluation indices for the GAN-generating time-series data.

6) Solving the Problems of Data Imbalance or Interference by GANs: In traffic anomaly inspection tasks, GANs predominantly rely on public road crack image datasets. Nonetheless, certain scenarios are challenging to train with these existing data sets. The imminent need is to address and rectify the data imbalance issue by generating more image data covering diverse scenarios and categories using GANs, and capturing additional risk types yet to be encountered by the network. Aiming at alleviating the background interference problems in images, methods such as image style transfer or road damage detection can be further studied to create useful data sets by simulating images in various settings, which should enhance the present detection capability of GAN models.

VII. CONCLUSIONS

This survey investigates and summarizes the recent progress on the GAN and its applications in the transportation domain. To enhance the performance of autonomous driving, the generation and transformation of perceptual data, trajectory prediction based on video scenarios, and security of target detection under jamming receive a lot of attention. Traffic flow research emphasizes spatio-temporal data forecasting and sparse data completion, both of which can be improved by GAN-based models. Commonly-used model evaluation indices are additionally mentioned in this part. Aiming at improving the efficiency of traffic anomaly inspection, we exert effort towards defect detection on roads and bridges, railway defect inspection, and traffic anomaly monitoring. Based on the applications of GAN-based methods, five challenges that remain to be addressed are further provided along with promising future perspectives. This survey will help the reader better understand how the GAN is employed in the transportation domain and offer useful suggestions for the future development of the framework.

REFERENCES

[1] I. Ahmed, S. Din, G. Jeon, F. Piccialli, and G. Fortino, "Towards collaborative robotics in top view surveillance: A framework for multiple object tracking by detection using deep learning," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 7, pp. 1253–1270, 2021.

[2] B. Peng, M. F. Keskin, B. Kulcsár, and H. Wymeersch, "Connected autonomous vehicles for improving mixed traffic efficiency in unsignalized intersections with deep reinforcement learning," *Comm. Transportation Research*, vol. 1, p. 100017, 2021.

[3] J. Dong, S. Chen, M. Miralinaghi, T. Chen, and S. Labi, "Development and testing of an image transformer for explainable autonomous driving systems," *J. Intelligent and Connected Vehicles*, vol. 5, no. 3, pp. 235–249, 2022.

[4] H. Ding, W. Li, N. Xu, and J. Zhang, "An enhanced eco-driving strategy based on reinforcement learning for connected electric vehicles: Cooperative velocity and lane-changing control," *J. Intelligent and Connected Vehicles*, vol. 5, no. 3, pp. 316–332, 2022.

[5] F. Liang, G. Zhiwei, W. Tao, G. Jinfeng, and D. Feng, "Collision avoidance model and its validation for intelligent vehicles based on deep learning LSTM," *J. Autom. Safety and Energy*, vol. 13, no. 1, p. 104, 2022.

[6] S. Li, Y. Liu, and X. Qu, "Model controlled prediction: A reciprocal alternative of model predictive control," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 6, pp. 1107–1110, 2022.

[7] Y. Liu, C. Lyu, Y. Zhang, Z. Liu, W. Yu, and X. Qu, "DeepTSP: Deep traffic state prediction model based on large-scale empirical data," *Comm. Transportation Research*, vol. 1, p. 100012, 2021.

[8] Y. Liu, Z. Liu, and R. Jia, "DeepPF: A deep learning based architecture for metro passenger flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 18–34, 2019.

[9] M. Xu, H. Lin, and Y. Liu, "A deep learning approach for vehicle velocity prediction considering the influence factors of multiple lanes," *Electronic Research Archive*, vol. 31, no. 1, pp. 401–420, 2023.

[10] I. Goodfellow, J. Pouget-Abadie, M. Mirza, *et al*., "Generative adversarial networks," *Communi. ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[11] J. Liu, T. Li, P. Xie, S. Du, F. Teng, and X. Yang, "Urban big data fusion based on deep learning: An overview," *Infor. Fusion*, vol. 53, pp. 123–133, 2020.

[12] S. Wang, J. Cao, and P. Yu, "Deep learning for spatio-temporal data mining: A survey," *IEEE Trans. Knowledge and Data Engineering*, vol. 34, no. 8, pp. 3681–3700, 2020.

[13] Y. Liu, R. Jia, J. Ye, and X. Qu, "How machine learning informs ride-hailing services: A survey," *Comm. Transportation Research*, vol. 2, p. 100075, 2022.

[14] J. Ye, J. Zhao, K. Ye, and C. Xu, "How to build a graph-based deep learning architecture in traffic domain: A survey," *IEEE Trans. Intelligent Transportation Systems*, vol. 23, no. 5, pp. 3904–3924, 2020.

[15] K.-F. Wang, C. Gou, Y.-J. Duan, Y.-L. Lin, X.-H. Zheng, and F. Wang, "Generative adversarial networks: The state of the art and beyond," *Acta Autom. Sinica*, vol. 43, no. 3, pp. 321–332, 2017.

[16] A. Aggarwal, M. Mittal, and G. Battineni, "Generative adversarial network: An overview of theory and applications," *Inter. J. Infor. Management Data Insights*, vol. 1, no. 1, p. 100004, 2021.

[17] R. Kulkarni, R. Gaikwad, R. Sugandhi, P. Kulkarni, and S. Kone, "Survey on deep learning in music using GAN," *Int. J. Eng. Res. Technol*, vol. 8, no. 9, pp. 646–648, 2019.

[18] W. Xia, Y. Zhang, Y. Yang, J.-H. Xue, B. Zhou, and M.-H. Yang, "Gan inversion: A survey," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 45, no. 3, pp. 3121–3138, 2022.

[19] X. Wang, H. Guo, S. Hu, M.-C. Chang, and S. Lyu, "GAN-generated faces detection: A survey and new perspectives," arXiv preprint arXiv: 2202.07145, 2022.

[20] A. You, J. K. Kim, I. H. Ryu, and T. K. Yoo, "Application of generative adversarial networks (GAN) for ophthalmology image domains: A survey," *Eye and Vision*, vol. 9, no. 1, pp. 1–19, 2022.

[21] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," arXiv preprint arXiv: 1511.06434, 2015.

[22] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of stylegan," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, 2020, pp. 8110–8119.

[23] A. Brock, J. Donahue, and K. Simonyan, "Large scale GAN training for high fidelity natural image synthesis,"arXiv preprint arXiv: 1809.11096, 2018.

[24] H. Zhang, T. Xu, H. Li, *et al.*, "StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks," in *Proc. IEEE Int. Conf. Computer Vision*, 2017, pp. 5907–5915.

[25] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2017, pp. 1125–1134.

[26] G. Antipov, M. Baccouche, and J.-L. Dugelay, "Face aging with conditional generative adversarial networks," in *Proc. IEEE Inte. Conf. Image Processing*, 2017, pp. 2089–2093.

[27] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Inte. Conf. Computer Vision*, 2017, pp. 2223–2232.

[28] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training GANs," in *Proc. 30th Int. Conf. Neural Information Proc. Syst.,* 2016, pp. 2234–2242.

[29] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier GANs," in *Proc. Int. Conf. Machine Learning*, 2017, pp. 2642–2651.

[30] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *Proc. Int. Conf. Machine Learning*, 2017, pp. 214–223.

[31] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation,"arXiv preprint arXiv: 1710.10196, 2017.

[32] T. Wang, M. Liu, J. Zhu, *et al.*, "Video-to-video synthesis," arXiv preprint arXiv: 1808.06601, 2018.

[33] L. Yu, W. Zhang, J. Wang, and Y. Yu, "SeqGAN: Sequence generative adversarial nets with policy gradient," in *Proc. AAAI Conf Artificial Intelligence*, 2017, vol. 31, p. 1.

[34] Y. Jin, J. Zhang, M. Li, Y. Tian, H. Zhu, and Z. Fang, "Towards the automatic anime characters creation with generative adversarial networks," arXiv preprint arXiv: 1708.05509, 2017.

[35] M. B. Hedge, M. Nelson, T. Pengilly, and M. Weatherford, "PokéGAN: P2P (Pet to Pokémon) Stylizer," *SMU Data Science*

Review, vol. 5, no. 2, p. 10, 2021.

[36] M. Brundage, S. Avin, J. Clark, *et al.*, "The malicious use of artificial intelligence: Forecasting, prevention, and mitigation," arXiv preprint arXiv: 1802.07228, 2018.

[37] S. E. Reed, Z. Akata, S. Mohan, S. Tenka, B. Schiele, and H. Lee, "Learning what and where to draw," in *Prpc. 30th Annual Conf. Neural Information Proc. Syst.,* 2016, pp. 217-225.

[38] A. Dash, J. C. B. Gamboa, S. Ahmed, M. Liwicki, and M. Z. Afzal, "TAC-GAN-text conditioned auxiliary classifier generative adversarial network," arXiv preprint arXiv: 1703.06412, 2017.

[39] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2018, pp. 8798–8807.

[40] D. Yoo, N. Kim, S. Park, A. S. Paek, and I. S. Kweon, "Pixel-level domain transfer," in *Proc. Computer Vision–ECCV: 14th European Conference, Amsterdam, The Netherlands, Proceedings, Part VIII 14*, 2016, pp. 517–532.

[41] R. Huang, S. Zhang, T. Li, and R. He, "Beyond face rotation: Global and local perception GAN for photorealistic and identity preserving frontal view synthesis," in *Proc. IEEE Int. Conf. Computer Vision*, 2017, pp. 2439–2448.

[42] Y. Taigman, A. Polyak, and L. Wolf, "Unsupervised cross-domain image generation," arXiv preprint arXiv: 1611.02200, 2016.

[43] G. Perarnau, J. Van De Weijer, B. Raducanu, and J. M. Álvarez, "Invertible conditional gans for image editing," arXiv preprint arXiv: 1611.06355, 2016.

[44] M.-Y. Liu and O. Tuzel, "Coupled generative adversarial networks," in *Proc. 30th Int. Conf. Neural Information Proc. Syst.,* 2016, pp. 469–477.

[45] A. Brock, T. Lim, J. M. Ritchie, and N. Weston, "Neural photo editing with introspective adversarial networks," arXiv preprint arXiv: 1609.07093, 2016.

[46] C. Ledig, L. Theis, F. Huszár, *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf Computer Vision and Pattern Recognition*, 2017, pp. 4681–4690.

[47] H. Bin, C. Weihai, W. Xingming, and L. Chun-Liang, "High-quality face image SR using conditional generative adversarial networks," arXiv preprint arXiv: 1707.00737, 2017.

[48] S. Vasu, N. Thekke Madam, and A. Rajagopalan, "Analyzing perception-distortion tradeoff using enhanced perceptual super-resolution network," in *Proc. European Conf. Computer Vision Workshops*, 2018, pp. 114–131.

[49] C. Vondrick, H. Pirsiavash, and A. Torralba, "Generating videos with scene dynamics," *Advances Neural Information Processing Systems,* vol. 29, 2016.

[50] V. Sandfort, K. Yan, P. J. Pickhardt, and R. M. Summers, "Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks," *Scientific Reports*, vol. 9, no. 1, p. 16884, 2019.

[51] S. Yu, H. Dong, G. Yang, *et al*., "Deep DE-aliasing for fast compressive sensing MRI," arXiv preprint arXiv: 1705.07137, 2017.

[52] J. Wu, C. Zhang, T. Xue, B. Freeman, and J. Tenenbaum, "Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling," in *Proc. 30th Int. Conf. Neural Information Proc. Syst.,* 2016, pp. 82–90.

[53] M. Gadelha, S. Maji, and R. Wang, "3D shape induction from 2D views of multiple objects," in *Proc. Int. Conf. 3D Vision*, 2017, pp. 402–411.

[54] A. Li, P. Zhao, X. Liu, A. Mansourian, K. W. Axhausen, and X. Qu, "Comprehensive comparison of E-scooter sharing mobility: Evidence from 30 European cities," *Transportation Research Part D:Transport and Environment*, vol. 105, p. 103229, 2022.

[55] J. Ke, H. Zheng, H. Yang, and X. M. Chen, "Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 591–608, 2017.

[56] J. Bao, X. Shi, and H. Zhang, "Spatial analysis of bikeshare ridership with smart card and POI data using geographically weighted regression method," *IEEE Access*, vol. 6, pp. 76049–76059, 2018.

[57] J. Wu and X. Qu, "Intersection control with connected and automated vehicles: A review," *J. Intelligent and Connected Vehicles*, vol. 5, no. 3, pp. 260–269, 2022.

[58] L. Maoyue, L. Hongyu, H. Xiangmei, X. Guangqi, and Y. Wei, "Surrounding vehicle recognition and information map construction technology in automatic driving," *J. Automotive Safety and Energy*, vol. 13, no. 1, p. 131, 2022.

[59] B. Kiran, I. Sobh, V. Talpaert, *et al*., "Deep reinforcement learning for autonomous driving: A survey," *IEEE Trans. Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4909–4926, 2021.

[60] M. Bojarski, D. Del Testa, D. Dworakowski, *et al.*, "End to end learning for self-driving cars," arXiv preprint arXiv: 1604.07316, 2016.

[61] S. Shalev-Shwartz, N. Ben-Zrihem, A. Cohen, and A. Shashua, "Long-term planning by short-term prediction," arXiv preprint arXiv: 1602.01580, 2016.

[62] E. Santana and G. Hotz, "Learning a driving simulator," arXiv preprint arXiv: 1608.01230, 2016.

[63] Z. Yang, Y. Chai, D. Anguelov, *et al.*, "SurfelGAN: Synthesizing realistic sensor data for autonomous driving," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, 2020, pp. 11118–11127.

[64] M. Amodio and S. Krishnaswamy, "TraVeLGAN: Image-to-image translation by transformation vector learning," in *Proc. IEEE/CVF Conf. computer Vision and Pattern Recognition*, 2019, pp. 8983–8992.

[65] S.-W. Huang, C.-T. Lin, S.-P. Chen, Y.-Y. Wu, P.-H. Hsu, and S.-H. Lai, "AugGAN: Cross domain adaptation with GAN-based data augmentation," in *Proc. European Conf. Computer Vision*, 2018, pp. 718–731.

[66] W. Xu, N. Souly, and P. P. Brahma, "Reliability of GAN generated data to train and validate perception systems for autonomous vehicles," in *Proc IEEE/CVF Winter Conf. Applications Computer Vision*, 2021, pp. 171–180.

[67] W. Wang, Y. Zhang, J. Gao, *et al*., "GOPS: A general optimal control problem solver for autonomous driving and industrial control applications," *Communi. Transportation Research*, vol. 3, p. 100096, 2023.

[68] D. Fei, M. Guanyu, T. En, Z. Nan, B. Jianmin, and Z. Dengyin, "Multi-channel high-resolution network and attention mechanism fusion for vehicle detection model," *J. Automotive Safety and Energy*, vol. 13, no. 1, p. 122, 2022.

[69] S. Tulyakov, M.-Y. Liu, X. Yang, and J. Kautz, "MoCoGAN: Decomposing motion and content for video generation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2018, pp. 1526–1535.

[70] E. L. Denton, "Unsupervised learning of disentangled representations from video," in *Proc. 31st Int. Conf. Neural Information Proc. Syst.,* 2017, pp. 4417–4426.

[71] J. Walker, K. Marino, A. Gupta, and M. Hebert, "The pose knows: Video forecasting by generating pose futures," in *Proc. IEEE Int. Conf. Computer Vision*, 2017, pp. 3332–3341.

[72] A. Clark, J. Donahue, and K. Simonyan, "Adversarial video generation on complex datasets," arXiv preprint arXiv: 1907.06571,2019.

[73] S. Mozaffari, O. Y. Al-Jarrah, M. Dianati, P. Jennings, and A. Mouzakitis, "Deep learning-based vehicle behavior prediction for autonomous driving applications: A review," *IEEE Trans. Intelligent Transportation Systems*, vol. 23, no. 1, pp. 33–47, 2020.

[74] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, and A. Alahi, "Social GAN: Socially acceptable trajectories with generative adversarial networks," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2018, pp. 2255–2264.

[75] S. Eiffert, K. Li, M. Shan, S. Worrall, S. Sukkarieh, and E. Nebot, "Probabilistic crowd GAN: Multimodal pedestrian trajectory

prediction using a graph vehicle-pedestrian attention network," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5026–5033, 2020.

[76] Q. Zhang, S. Hu, J. Sun, Q. A. Chen, and Z. M. Mao, "On adversarial robustness of trajectory prediction for autonomous vehicles," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, 2022, pp. 15159–15168.

[77] C. Gómez-Huélamo, M. V. Conde, M. Ortiz, S. Montiel, R. Barea, and L. M. Bergasa, "Exploring attention GAN for vehicle motion prediction," in *Proc. IEEE 25th Int. Conf. Intelligent Transportation Systems*, 2022, pp. 4011–4016.

[78] V. Kosaraju, A. Sadeghian, R. Martín-Martín, I. Reid, H. Rezatofighi, and S. Savarese, "Social-BIGAT: Multimodal trajectory forecasting using bicycle-GAN and graph attention networks," *Advances in Neural Information Processing Systems,* vol. 32, 2019.

[79] E. Wang, H. Cui, S. Yalamanchi, M. Moorthy, and N. Djuric, "Improving movement predictions of traffic actors in bird's-eye view models using GANs and differentiable trajectory rasterization," in *Proc. 26th ACM SIGKDD Int. Conf. Knowledge Discovery & Data Mining*, 2020, pp. 2340–2348.

[80] C. Zhao, Y. Zhu, Y. Du, F. Liao, and C.-Y. Chan, "A novel direct trajectory planning approach based on generative adversarial networks and rapidly-exploring random tree," *IEEE Trans. Intelligent Transportation Systems*, vol. 23, no. 10, pp. 17910–17921, 2022.

[81] M. Ghafoorian, C. Nugteren, N. Baka, O. Booij, and M. Hofmann, "El-GAN: Embedding loss driven generative adversarial networks for lane detection," in *Proc. European Conf. Computer Vision Workshops*, 2018, pp. 256–272.

[82] X. Wang, A. Shrivastava, and A. Gupta, "A-fast-RCNN: Hard positive generation via adversary for object detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2017, pp. 2606–2615.

[83] J. Li, X. Liang, Y. Wei, T. Xu, J. Feng, and S. Yan, "Perceptual generative adversarial networks for small object detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2017, pp. 1222–1230.

[84] Y. Bai, Y. Zhang, M. Ding, and B. Ghanem, "SOD-MTGAN: Small object detection via multi-task generative adversarial network," in *Proc. European Conf. Computer Vision*, 2018, pp. 206–221.

[85] A. Liu, X. Liu, J. Fan, *et al.*, "Perceptual-sensitive GAN for generating adversarial patches," in *Proc. AAAI Conf. Artificial Intelligence*, 2019, vol. 33, no. 1, pp. 1028–1035.

[86] W. Zhu, J. Wu, T. Fu, J. Wang, J. Zhang, and Q. Shangguan, "Dynamic prediction of traffic incident duration on urban expressways: A deep learning approach based on LSTM and MLP," *J. Intelligent and Connected Vehicles*, vol. 4, no. 2, pp. 80–91, 2021.

[87] A. Koochali, P. Schichtel, A. Dengel, and S. Ahmed, "Probabilistic forecasting of sensory data with generative adversarial networks–forGAN," *IEEE Access*, vol. 7, pp. 63868–63880, 2019.

[88] D. Saxena and J. Cao, "D-GAN: Deep generative adversarial nets for spatio-temporal prediction," arXiv preprint arXiv: 1907.08556, 2019.

[89] J. Yoon, D. Jarrett, and M. Van der Schaar, "Time-series generative adversarial networks," *Advances in Neural Information Processing Systems,* vol. 32, 2019.

[90] L. Zhang, J. Wu, J. Shen, *et al.*, "SATP-GAN: Self-attention based generative adversarial network for traffic flow prediction," *Transportmetrica B:Transport Dynamics*, vol. 9, no. 1, pp. 552–568, 2021.

[91] Y. Zhang, Y. Li, X. Zhou, X. Kong, and J. Luo, "Curb-GAN: Conditional urban traffic estimation through spatio-temporal generative adversarial networks," in *Proc. 26th ACM SIGKDD Int. Conf. Knowledge Discovery & Data Mining*, 2020, pp. 842–852.

[92] Y. Zhang, Y. Li, X. Zhou, X. Kong, and J. Luo, "TrafficGAN: Off-deployment traffic estimation with traffic generative adversarial networks," in *Proc. IEEE Int. Conf. Data Mining*, 2019, pp. 1474–1479.

[93] D. Xu, C. Wei, P. Peng, Q. Xuan, and H. Guo, "GE-GAN: A novel deep learning framework for road traffic state estimation," *Transportation Research Part C: Emerging Technologies*, vol. 117, p. 102635, 2020.

[94] J. Jin J, D. Rong, T. Zhang, *et al*., "A GAN-based short-term link traffic prediction approach for urban road networks under a parallel learning framework," *IEEE Trans. Intelligent Transportation Syst.*, vol. 23, no. 9, pp. 16185–16196, 2022.

[95] Y. Lv, Y. Chen, L. Li, and F.-Y. Wang, "Generative adversarial networks for parallel transportation systems," *IEEE Intelligent Transportation Systems Magazine*, vol. 10, no. 3, pp. 4–10, 2018.

[96] Y. Wang, Y. Zheng, and Y. Xue, "Travel time estimation of a path using sparse trajectories," in *Proc. 20th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2014, pp. 25–34.

[97] I. M. Baytas, C. Xiao, X. Zhang, F. Wang, A. K. Jain, and J. Zhou, "Patient subtyping via time-aware LSTM networks," in *Proc. 23rd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2017, pp. 65–74.

[98] H.-F. Yu, N. Rao, and I. S. Dhillon, "Temporal regularized matrix factorization for high-dimensional time series prediction," in *Proc. 30th Int. Conf. Neural Information Proc. Syst.,* 2016, pp. 847–855.

[99] Z. Liu, Y. Yang, W. Huang, Z. Tang, N. Li, and F. Wu, "How do your neighbors disclose your information: Social-aware time series imputation," in *Proc. World Wide Web Conf.*, 2019, pp. 1164–1174.

[100] H. Wu, S. Zheng, J. Zhang, and K. Huang, "GP-GAN: Towards realistic high-resolution image blending," in *Proc. 27th ACM Int. Conf. Multimedia*, 2019, pp. 2487–2495.

[101] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2016, pp. 2536–2544.

[102] X. Wang, D. Xu, and F. Gu, "3D smodel inpainting based on 3D deep convolutional generative adversarial network," *IEEE Access*, vol. 8, pp. 170355–170363, 2020.

[103] W. Li, L. Min, Y. Jia-qing, Z. Ling-yu, P. Ke, and L. Zheng-xi, "Urban traffic flow data recovery method based on generative adversarial network," *J. Transportation Systems Engineering and Information Technology*, vol. 18, no. 6, p. 63, 2018.

[104] G. Xiong, Z. Li, M. Zhao, *et al.*, "TrajSGAN: A Semantic-guiding adversarial network for Urban trajectory generation," *IEEE Trans. Computational Social Systems,* 2023.

[105] A. Borji, "Pros and cons of GAN evaluation measures," *Computer Vision and Image Understanding*, vol. 179, pp. 41–65, 2019.

[106] Z. Chen, S. Nie, T. Wu, and C. G. Healey, "High resolution face completion with multiple controllable attributes via fully end-to-end progressive generative adversarial networks," arXiv preprint arXiv: 1801.07632, 2018.

[107] T. Chen, X. Zhai, M. Ritter, M. Lucic, and N. Houlsby, "Self-supervised gans via auxiliary rotation loss," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, 2019, pp. 12154–12163.

[108] F. Ni, J. Zhang, and M. N. Noori, "Deep learning for data anomaly detection and data compression of a long: Span suspension bridge," *Computer—Aided Civil and Infrastructure Engineering*, vol. 35, no. 7, pp. 685–700, 2020.

[109] Y. Xu, X. Ouyang, Y. Cheng, *et al.*, "Dual-mode vehicle motion pattern learning for high performance road traffic anomaly detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops*, 2018, pp. 145–152.

[110] V. Reddy, C. Sanderson, and B. C. Lovell, "Improved anomaly detection in crowded scenes via cell-based analysis of foreground speed, size and texture," in *Proc. CVPR Workshops*, 2011, pp. 55–61.

[111] A. Mohan and S. Poobal, "Crack detection using image processing: A critical review and analysis," *Alexandria Engineering J.*, vol. 57, no. 2, pp. 787–798, 2018.

[112] K. Zhang, Y. Zhang, and H.-D. Cheng, "CrackGAN: Pavement crack detection using partially accurate ground truths based on generative adversarial learning," *IEEE Trans. Intelligent Transportation Systems*, vol. 22, no. 2, pp. 1306–1319, 2020.

[113] W. Zhai, J. Zhu, Y. Cao, and Z. Wang, "A generative adversarial network based framework for unsupervised visual surface inspection," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2018, pp. 1283–1287.

[114] H. Oliveira and P. L. Correia, "CrackIT—An image processing toolbox for crack detection and characterization," in *Proc. IEEE Int. Conf. Image Processing*, 2014, pp. 798–802.

[115] Z. Gao, B. Peng, T. Li, and C. Gou, "Generative adversarial networks for road crack image segmentation," in *Proc. Int. Joint Conf. Neural Networks,* 2019, pp. 1–8.

[116] Q. Mei and M. Gül, "A cost effective solution for pavement crack inspection using cameras and deep neural networks," *Construction and Building Materials*, vol. 256, p. 119397, 2020.

[117] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2017, pp. 4700–4708.

[118] K. Zhang, Y. Zhang, and H. Cheng, "Self-supervised structure learning for crack detection based on cycle-consistent generative adversarial networks," *J. Computing in Civil Engineering*, vol. 34, no. 3, p. 04020004, 2020.

[119] J. Mao, H. Wang, and B. F. Spencer Jr, "Toward data anomaly detection for automated structural health monitoring: Exploiting generative adversarial nets and autoencoders," *Structural Health Monitoring*, vol. 20, no. 4, pp. 1609–1626, 2021.

[120] K. Lee and D. H. Shin, "Generative model of acceleration data for deep learning-based damage detection for bridges using generative adversarial network," *Journal of KIBIM*, vol. 9, no. 1, pp. 42–51, 2019.

[121] P. Yang, W. Jin, and P. Tang, "Anomaly detection of railway catenary based on deep convolutional generative adversarial networks," in *Proc. 3rd IEEE Advanced Information Technology, Electronic and Automation Control Conf.*, 2018, pp. 1366–1370.

[122] Y. Lyu, Z. Han, J. Zhong, C. Li, and Z. Liu, "A generic anomaly detection of catenary support components based on generative adversarial networks," *IEEE Trans. Instrumentation and Measurement*, vol. 69, no. 5, pp. 2439–2448, 2019.

[123] Y. Lyu, Z. Han, J. Zhong, C. Li, and Z. Liu, "A GAN-based anomaly detection method for isoelectric line in high-speed railway," in *Proc. IEEE Int. Instrumentation and Measurement Technology Conf.*, 2019, pp. 1–6.

[124] L. Xue and S. Gao, "Unsupervised anomaly detection system for railway turnout based on GAN," in *Proc. Journal of Physics: Conf. Series*, 2019, vol. 1345, no. 3. p. 032069.

[125] K. Wang, X. Zhang, Q. Hao, Y. Wang, and Y. Shen, "Application of improved least-square generative adversarial networks for rail crack detection by AE technique," *Neurocomputing*, vol. 332, pp. 236–248, 2019.

[126] V. Saligrama and Z. Chen, "Video anomaly detection based on local statistical aggregates," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2012, pp. 2112–2119.

[127] S. Akcay, A. Atapour-Abarghouei, and T. P. Breckon, "Ganomaly: Semi-supervised anomaly detection via adversarial training," in *Proc.14th Asian Conf. Computer Proc. Vision, Perth, Australia, Revised Selected Papers, Part III 14*, 2019, pp. 622–637.

[128] Y. Sun, W. Yu, Y. Chen, and A. Kadam, "Time series anomaly detection based on GAN," in *Proc. 6th Int. Conf. Social Networks Analysis, Management and Security*, 2019, pp. 375–382.

[129] S. Akçay, A. Atapour-Abarghouei, and T. P. Breckon, "Skip-ganomaly: Skip connected and adversarially trained encoder-decoder anomaly detection," in *Proc. Int. Joint Conf. Neural Networks*, 2019, pp. 1–8.

[130] Y. Qiu, T. Misu, and C. Busso, "Driving anomaly detection with conditional generative adversarial network using physiological and can-bus data," in *Proc. Int. Conf. Multimodal Interaction*, 2019, pp. 164–173.

[131] F. Liao, T. Arentze, and H. Timmermans, "Incorporating space–Time constraints and activity-travel time profiles in a multi-state supernetwork approach to individual activity-travel scheduling," *Transportation Research Part B: Methodological*, vol. 55, pp. 41–58, 2013.

Hongyi Lin received the B.S. degree in automotive engineering from the School of Mechanical and Automotive Engineering, South China University of Technology in 2022. Currently, he is a Ph.D. candidate in mechanical engineering at the School of Vehicle and Mobility, Tsinghua University, under the guidance of Professor Xiaobo Qu. His research interests include intelligent vehicles and the applications of machine learning in intelligent transportation systems.

Yang Liu received the Ph.D. degree in transportation engineering from Southeast University in 2021. Currently, He is a Marie Curie Fellow at the Department of Architecture and Civil Engineering, Chalmers University of Technology, Sweden. His research interests include intelligent transportation systems, AI techniques, and data mining, with applications in complex real-world problems such as autonomous vehicles and urban mobility.

He serves as the Associate Editor of *IEEE Transactions on Intelligent Vehicles* and *Journal of Intelligent* and *Connected Vehicles*. In addition, he is a Young Editor for two top-tier academic journals, including *The Innovation* and *IEEE/CAA Journal of Automatica Sinica*.

Shen Li received the Ph.D. degree in transportation engineering from the University of Wisconsin- Madison, USA in 2018. He is a Research Associate at Tsinghua University. His research interests include intelligent transportation systems (ITS), architecture design of CAVH system, vehicle-infrastructure cooperative planning and decision method, traffic data mining based on cellular data, and traffic operations and management.

Xiaobo Qu received the Ph.D. degree in transport engineering from National University of Singapore. Currently, he is a Chair Professor at the School of Vehicle and Mobility, Tsinghua University. He is also an elected Member of Academia Europaea – the Academy of Europe since August 2020, and a Fellow of the Institution of Engineering and Technology (IET). His research is focused on enhancing urban mobility systems by incorporating emerging technologies, particularly for emergency services and connected automated vehicles.

He has Authored or Co-Authored over 120 journal articles published in top-tier journals, including 14 ESI highly cited papers. He serves as an Editor for 10 journals, including the Editor in Chief of *Communications in Transportation Research*, and Executive Editor in Chief of *Journal of Intelligent and Connected Vehicles*. To date, he has secured research funding well above 10 million Euros from the Australian Research Council, Swedish Innovation Agency Vinnova, STINT, and the European Union. He has been invited to serve as a panel/assessor for prestigious funding scheme such as the Australian Research Council Centre of Excellence (35 million AUD each), Future Fellowship (career grant, 1 million AUD each), Netherlands NWO VICI (career grant, 2.5 million Euros each), Hong Kong Research Council theme-based scheme (30 million HKD each), Singapore Ministry of Education Thematic Research Program (5 million SD each), and the European Research Council, etc.