

Generative Adversarial Network Inverse Perspective Mapping Image Synthesis for Autonomous Vehicle Training

Adizul Ahmad, Ihsan Mohd Yassin*, Mohd Nasir Taib, Megat Syahirul Amin Megat Ali, Fadhlan Hafizhelmi Kamaru Zaman, and Azlee Zabidi

Abstract— Autonomous vehicles (AV) are undergoing extensive research and development due to their disruptive potential and safety advantages. Critical challenges concerning AVs include the precision and accuracy of lane detection. Artificial intelligence (AI) systems for lane detection must be robust, which can be achieved by using many samples to train the system. However, due to the limitations of collection data to account for all road variations is impractical due to the variability of the data involved, especially for highly unique road curvatures. Recent improvements in Generative Adversarial Networks (GANs) make them an attractive tool for generating realistic data to circumvent this problem by generating high-quality images that can represent a variety of road conditions. In this paper, we trained a lightweight GAN architecture on Inverse Perspective Mapping (IPM) images captured by a roof-mounted camera to construct a bird's eye view (BEV) of the road. After training, GAN was able to generate realistic images with a suitable degree of variation. These variations in the generated images have the potential to train a robust navigation computer aboard the AV to perform curvature estimation. Common testing methods for evaluating GANs are presented (such as SSIM, FID, IS, PSNR) demonstrate that the GAN was able to generate statistically proven realistic images with good variability compared to the original training images. Quantitative results show that images generated using the Exponential Moving Average (EMA) technique achieved a PSNR of 15.7947, SSIM of 0.3875, FID of 131.5848, and IS of 10.5015, indicating improved fidelity and realism over standard outputs.

Index Terms— Generative Adversarial Network (GAN), Autonomous Vehicle (AV), Computer Vision, Data Synthesis, Inverse Perspective Mapping (IPM).

I. INTRODUCTION

1.1 Research Background

AVs are self-driving vehicles able to steer and navigate

This manuscript is submitted on 21 February 2025, revised on 24 March 2025, accepted on 28 March 2025. Adizul Ahmad, Ihsan Mohd Yassin, Mohd Nasir Taib, Megat Syahirul Amin Megat Ali and Fadhlan Hafizhelmi Kamaru Zaman, are from the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia.

Azlee Zabidi is from the Faculty of Systems & Software Engineering, College of Computing & Applied Sciences, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia

*Corresponding author
Email address: ihsan.yassin@gmail.com

1985-5389/© 2023 The Authors. Published by UiTM Press. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

without human intervention. They accomplish this by constantly monitoring their environment using an array of sensors and calculating the best action based on the inputs [1]. Autonomous vehicle development has a long history. The first important step towards constructing the first stand-alone AV was taken by Japan's Tsukuba Mechanical Engineering Laboratory in 1977 [2]. Rather than depending on external road equipment, it was guided by machine vision, which uses images from built-in cameras to analyse the surrounding area. The prototype was configured to follow white traffic markings, and its speed could reach over 20 miles per hour. In 2004, the Defense Advanced Development Projects Agency (DARPA) in the United States created the Grand Challenges Program, which accelerated AV research [3], [4]. As a result of the programmes, AVs that could traverse desert terrain were developed, and researchers were also able to put AVs on city streets through DARPA's Urban Challenge Program. AVs are under active research and development by major players in the automotive sector due to its future disruptive potential [5].

Due to the unconstrained nature of the driving environment, Machine Learning (ML) models are often employed to learn the proper controls in response to the information about its surroundings [6]–[9]. Importantly, an autonomous agent should be able to interact with multi-modal dynamic settings in real time while learning previously unseen typical and atypical scenarios.

Recent advances in AI can be attributed in large part to the current acceleration in the development and manufacturing of AVs [10]. However, the precision, stability, and safety of AVs are still in their early stages. Several significant issues concerning AVs include lane detection precision and accuracy, recognition of objects of interest on the road, and the ability to estimate the distance between the AV and those objects [11], [12]. Many ML-based perception, prediction, and planning algorithms are ideally constructed from a large amount of diverse and representative data. Many real-world examples, on the other hand, either do not exist or appear in small numbers in the limited datasets used to train these models [13]. Because such environments are rarely available for training, the agent should be aware of its own capabilities and limitations [14]. The task of gathering training data for this is undoubtedly daunting. In this case, the use of synthetic training data is advantageous.

The use of IPM to convert driver's view samples to BEV images is crucial for AV applications, particularly for accurately observing and analyzing the road's contour and

curvature. IPM provides a top-down view of the road, eliminating perspective distortions inherent in the driver's view. This top view is essential for precise curvature estimation, as it presents a clearer and more consistent representation of the road's geometry. While driver's view samples can be used for training, they often contain significant perspective distortions, occlusions, and variances in appearance due to changes in camera angle and position. These factors complicate the training process and reduce the accuracy of the generated images. By using BEV images obtained through IPM, the training data becomes more uniform and representative of the actual road layout, thereby enhancing the performance and reliability of the trained models for AV steering and navigation tasks.

In this paper, we propose a LGAN trained on IPM images captured by a roof-mounted camera to create a detailed bird's-eye view of road scenes. IPM effectively eliminates distortions, providing a clear view of lane markings and road curvature [15], which are crucial for AV navigation. Our LGAN enhances data diversity by generating realistic and varied IPM images, addressing common challenges like lighting inconsistencies and occlusions. We evaluated the realism of the generated images using established metrics, including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Fréchet Inception Distance (FID), and Inception Score (IS), demonstrating that the LGAN effectively mimics real-world data. These synthetic images improve AV training datasets, especially for robust lane detection and accurate curvature estimation, under diverse driving conditions.

The LGAN [16] was selected for this study due to its ability to achieve high-quality image synthesis while maintaining computational efficiency, making it particularly well-suited for environments with limited resources. Compared to other GAN architectures, LGAN reduces training instability and computational overhead, ensuring faster convergence and stable performance. This is especially beneficial for real-time applications, mobile devices, and personal computing environments where access to high-performance GPUs is limited. By incorporating stabilization techniques and optimizing the generator network, LGAN provides a balance between efficiency and image fidelity, making it a practical choice for applications that require rapid and reliable GAN-based image generation without excessive computational demands.

The remainder of this paper examines the application of LGAN in AV navigation. Section 1.2 explores GANs' role in AV systems, with a focus on road safety and object detection. Section 1.3 explores the applications of IPM in AVs. Section 2.0 covers theoretical foundations, including GANs and Inverse Perspective Mapping (IPM). Section 3.0 outlines the methodology, including camera setup, data collection, IPM transformation to BEV scenes, and GAN training. Section 4.0 presents results, analyzing the quality of GAN-generated IPM images using metrics such as SSIM, PSNR, FID and IS. The paper concludes with key findings and areas for further research.

1.2 GANS in AVs

The diversity and complexity of real-world driving scenarios often exceed the capacity of available datasets, limiting the training effectiveness of AV systems. To address this, GANs have become a valuable tool for generating synthetic data that captures rare or underrepresented events. Applications include pedestrian detection [17], road generation conditioned on steering commands [18], and vehicle trajectory prediction [19], among others. By enriching training datasets and enhancing the realism of simulated environments, GANs contribute significantly to the advancement of perception, planning, and decision-making systems in AVs.

1.2.1 Road Safety (Object Detection)

Object detection using GANs has emerged as a promising technique for enhancing the safety and perception capabilities of AVs. A GAN comprises two competing networks: a generator that produces synthetic data and a discriminator that distinguishes between real and generated inputs. This adversarial structure allows the system to learn complex visual patterns, enabling AVs to recognize and detect dynamic objects such as pedestrians, cyclists, and other vehicles with greater precision. The integration of GANs into object detection pipelines has contributed to improved navigation and real-time decision-making in complex driving environments.

Numerous GAN-based frameworks have been developed to advance AV navigation. For example, Social GAN (SGAN) incorporates deep learning for trajectory prediction and outperforms traditional Long Short-Term Memory (LSTM) models by integrating object detection (YOLO v3) with real-time tracking (SORT) [18]. Conditional architectures like the Conditional Speed GAN (CSG) simulate realistic pedestrian motion, offering enhanced modeling of multi-agent interactions in traffic scenarios [20]. Other studies focus on addressing specific perceptual tasks, a Conditional GAN was employed for speed bump detection with robust results despite limited data availability [12], while saliency-driven attention mechanisms using conditional GANs improved visual focus in cluttered scenes [17]. Moreover, spatio-temporal GANs (STC-GAN) have demonstrated effectiveness in video frame prediction and pedestrian trajectory estimation using datasets such as Cityscapes and CamVid [21], [22]. These contributions underscore the adaptability and strength of GANs in elevating situational awareness and responsiveness in AV systems.

1.2.2 Navigation Systems

Navigation systems are fundamental to the operation of AVs, requiring high levels of precision, adaptability, and real-time responsiveness. Recent developments have leveraged GANs to enhance these systems, particularly through the prediction of future frames in video sequences, an essential function for anticipating changes in dynamic environments. By generating synthetic data, GANs enable AVs to simulate diverse driving scenarios and foresee potential obstacles, thus improving navigation reliability. This integration of GANs with conventional navigation frameworks allows AVs to more effectively interpret their surroundings and respond proactively

in complex traffic situations.

Path planning, one of the most critical yet challenging components of AV navigation, has also benefited from GAN-based innovations. While traditional algorithms such as Rapidly Exploring Random Trees (RRT) can identify feasible routes, they often fall short in terms of optimality and computational efficiency. To overcome these limitations, the Conditional GAN-based CGAN-RRT* approach integrates the probabilistic modeling capabilities of CGANs with RRT*, enabling the generation of more optimal paths while reducing computational overhead [23]. Further advancements include the Continuous Conditional GAN (CCGAN), which produces varied road images under different steering angles, thereby enriching the training data and improving navigational accuracy [18]. Beyond path planning, GANs also support higher-order functionalities; for example, Video Prediction GANs (VPGANs) have outperformed traditional stochastic models in forecasting vehicle movements [24], while Enhanced Super-Resolution GANs (ESRGANs) significantly improve shared sensory data quality, reducing blind spots and occlusions in real-world environments like KITTI [25]. Collectively, these advancements underscore the transformative role of GANs in making AV navigation systems more intelligent, reliable, and adaptable.

1.2.3 Privacy Protection

The integration of AVs into transportation has highlighted the need for robust privacy protection systems, especially as AVs gather extensive data from multiple sources, such as onboard cameras. This data can inadvertently expose sensitive information, raising privacy concerns. GANs provide a promising solution by generating synthetic datasets that preserve essential features while masking sensitive details, thereby mitigating data leakage risks. GAN-based privacy techniques are effective in securing the data used for navigation and decision-making, ensuring comprehensive coverage of privacy challenges associated with AVs.

Recognizing these concerns, researchers have developed models such as Auto-Driving GAN (ADGAN-I and ADGAN-II), which process visual inputs to produce privacy-preserving outputs. These models were specifically trained to conceal side-channel information and protect the privacy of road users by countering adversarial location inference attacks [26]. ADGAN-I and ADGAN-II leverage GANs to accept photos and generate outputs that obscure sensitive details based on specific item classes, effectively balancing data utility and privacy. Real-world evaluations demonstrate that these models outperform existing methods, offering superior protection against unauthorized data access and enhancing the privacy of images and videos acquired by AVs. This underscores the significant potential of GAN-based frameworks in building secure, privacy-centric AV systems [27].

1.2.4 Dataset Generation

The use of Generative Adversarial Networks (GANs) for dataset generation in autonomous vehicles (AVs) has gained attention due to its potential to reduce data collection time and

enhance model training. By creating synthetic data that mimics real-world scenarios, GANs enable AVs to learn from more diverse datasets than those obtained from real-world driving, including challenging situations like heavy traffic or extreme weather. This automated approach also minimizes the need for manual data labeling in object detection tasks.

Researchers have developed various GAN-based solutions to address limitations in existing datasets. For example, Ped-Cross GAN was customized to generate images of pedestrian behavior, with its loss function adjusted to Wasserstein to avoid exploding gradients [28]. Additionally, a CycleGAN network was employed to synthesize rare traffic conditions, such as nighttime driving, improving the performance of downstream object detection models by incorporating these challenging scenarios into the training process [13], [17]. Another study used a Dynamic Bayesian Network in combination with a bank of GANs to model a vehicle's position and visual input, effectively identifying atypical situations in real-world driving tests [14]. These developments underscore the value of GANs in creating robust, diverse datasets for training AV systems, enhancing their ability to handle complex real-world scenarios.

1.2.5 Lightweight GAN

The Lightweight GAN (LGAN) architecture by [16] was designed to address the challenges of high computational demands and instability in GAN training, particularly for high-fidelity few-shot image synthesis. This architecture is particularly relevant in scenarios where computational resources are limited, such as on mobile or edge devices and personal computers without requiring access to high performance enterprise computational resources. It achieves a balance between model efficiency and image quality, making it a compelling choice for applications requiring fast and stable GAN training.

The Lightweight GAN is specifically optimized for resource-constrained environments, similar to the approaches discussed in the works of several researcher, which focus on reducing model size and computational requirements while maintaining image quality [29], [30]. Additionally, techniques such as knowledge distillation and quantization are employed to achieve significant reductions in model size, as seen in other lightweight GAN implementations [29]. In terms of training stability and speed, LGAN incorporated stabilization methods to address the common issue of training instability in GANs, a problem also tackled by another researcher using self-attention mechanisms and spectral normalization [31].

In terms of speed, the LGAN architecture was designed to accelerate training, akin to the Faster Projected GAN, which achieves a 20% speed increase by optimizing the generator network [32]. Lightweight GAN is capable of producing high-quality images, a critical requirement for few-shot image synthesis tasks, with much faster computational time. This is comparable to the performance of other lightweight models that focus on maintaining image quality despite reduced computational resources [30], [33]. Additionally, the ability to capture long-range dependencies, as highlighted in the work by [33], is crucial for generating high-fidelity images and is a

feature that the Lightweight GAN shares with other advanced models.

In terms of model size and parameters, the Lightweight GAN offers a competitive balance of size and performance compared to its mobile-specific alternatives, making it suitable for real-time applications. Compared to models like MobileFSGAN, which is designed for mobile devices, making it suitable for real-time applications [34][6].

1.3 Applications of Inverse Perspective Mapping in AVs

IPM is a fundamental technique in autonomous vehicle navigation, transforming perspective views from vehicle-mounted cameras into bird's-eye views. This transformation is essential for accurate environmental perception, facilitating obstacle detection, path planning, and high-definition mapping. Recent research highlights the significance of IPM in enhancing the accuracy and efficiency of autonomous systems by improving visual data processing. However, challenges such as perspective distortion due to road inclinations and variations in camera properties necessitate ongoing refinement in IPM methodologies to ensure robust performance across diverse driving conditions.

IPM has been successfully applied in various domains of autonomous driving. One notable application is high-definition map construction, where frameworks such as GenMapping leverage IPM to decouple camera parameters from the training process, thereby improving generalization across different visual sensors. This approach enhances mapping accuracy and speed, yielding superior performance in both semantic and vectorized mapping tasks [35]. Additionally, IPM is widely used in camera-based autonomous driving systems, where it aids in transforming camera views into bird's-eye perspectives for improved path planning and obstacle detection. Research suggests that integrating stereo vision and deep learning-based methods can mitigate IPM's inherent distortions and enhance its reliability [36]. Furthermore, IPM plays a crucial role in visual parking and occupancy mapping, offering LiDAR-like performance using visual inputs. By integrating ego-motion data, this method reduces the need for labor-intensive labeling and significantly lowers implementation costs [37].

Several enhancements and techniques have been proposed to optimize IPM's effectiveness. In lane tracking applications, IPM is used to adjust camera perspectives and apply probabilistic Hough transforms, allowing for real-time lane edge detection and steering adjustments with minimal computational overhead. This method eliminates the need for extensive model training, making it an efficient solution for scaled autonomous vehicles [38]. Another key advancement is monocular camera-based bird's-eye view generation, which relies on lane detection and road infrastructure clues to infer waypoints for vehicle motion planning. This technique facilitates navigation in environments where prior location data is unavailable, further expanding the utility of IPM in autonomous systems [39].

Despite its advantages, IPM presents several challenges that require careful consideration. The accuracy of IPM-based

transformations is highly dependent on road conditions and camera calibration, necessitating continuous refinement through correction techniques and alternative methodologies. Moreover, integrating IPM with complementary technologies such as LiDAR and deep learning has shown promise in overcoming its limitations, yet computational demands and susceptibility to adversarial perturbations in image processing remain key concerns [15], [36]. Addressing these issues is critical for ensuring the robustness and efficiency of IPM-based autonomous navigation systems. Moving forward, research should focus on optimizing computational efficiency, enhancing IPM's resilience to environmental variations, and exploring hybrid approaches that combine multiple sensing modalities to achieve more reliable and adaptive autonomous vehicle navigation.

II. THEORETICAL BACKGROUND

The rapid advancement of AV technology has brought forth a critical need for high-quality data and robust perception systems. Among the foundational components of AV perception and navigation are geometric transformations and synthetic image generation techniques that enhance the reliability and diversity of training datasets. This section provides an overview of the theoretical principles underpinning the research, beginning with IPM, a technique essential for transforming perspective views into bird's-eye representations. Following this, the structure and innovations of the proposed LGAN are presented, detailing its mechanisms for efficient image generation. Lastly, a comprehensive discussion of the quantitative metrics used to evaluate the realism and utility of the generated images, namely PSNR, SSIM, FID, and IS ensures a rigorous assessment of image quality for AV navigation tasks.

2.1 Inverse Perspective Mapping (IPM)

IPM is a vital geometric transformation technique for AV, converting driver's view images into BEV images. By using calibrated camera parameters, including position, orientation, and intrinsic properties, IPM reprojects pixels to simulate an overhead view of the road. This transformation eliminates perspective distortions, ensuring that parallel lines appear parallel, and distances are uniformly scaled, providing a clear and accurate depiction of road geometry. Accurate IPM is essential for lane detection and curvature estimation in AV systems and advanced driver-assistance systems (ADAS), as any error in calibration can lead to distortions in the transformed BEV image. The transformation is mathematically represented by Equation 2.1.

$$\begin{pmatrix} U \\ V \\ 1 \end{pmatrix} = K \cdot [R \ v T] \cdot \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (2.1)$$

Where

- $\begin{pmatrix} U \\ V \\ 1 \end{pmatrix}$ are the pixel coordinates in the image.
- K is the camera intrinsic matrix.
- $[R \vee T]$ is the extrinsic matrix combining rotation R and translation T .
- $\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$ are the real-world coordinates.

For the BEV transformation, the Z coordinate (height) is set to zero, simplifying the equation:

$$\begin{pmatrix} U \\ V \\ 1 \end{pmatrix} = K \cdot R \cdot \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} \quad (2.2)$$

IPM is beneficial for AV applications as it removes perspective distortions, ensuring consistent scale and orientation, which is vital for accurate lane detection and curvature estimation. By converting driver's views into standardized BEV, IPM enhances data uniformity and improves training dataset quality. This transformation simplifies the analysis of road features, aiding in more effective machine learning model training. Consequently, IPM-generated BEV images lead to more accurate AV performance in tasks such as lane detection, obstacle recognition, and path planning.

2.2 Generative Adversarial Network

The LGAN, developed for efficient high-resolution image generation, is designed to be computationally light, enabling training on a single GPU within a short period. Key innovations include skip-layer excitation in the generator and autoencoding self-supervised learning in the discriminator. The skip-layer excitation mechanism enhances information flow by adding the output of one layer to the input of a later layer, preserving high-frequency details and stabilizing training. This method not only improves image quality but also accelerates the overall training process, as described mathematically in Equation 2.3.

$$G'(x) = F(x) + x \quad (2.3)$$

Where $G'(x)$ is the enhanced feature map, $F(x)$ is the transformation applied by the generator, and x is the input to the generator.

The LGAN discriminator utilizes self-supervised learning with an auxiliary autoencoder, enhancing its ability to distinguish real from fake images by reconstructing them. The autoencoder is trained alongside the discriminator, with the reconstruction loss integrated into the total loss function, as described in Equation 2.4. Additionally, LGAN improves training efficiency through mixed precision, combining 16-bit and 32-bit floats to speed up training by up to 33% and save up to 40% of memory, detailed in Equation 2.5. Data augmentation

techniques, modeled by Equation 2.6, further boost performance, especially in low-data scenarios.

$$L_D = L_{adv} + \lambda L_{rec} \quad (2.4)$$

where L_{adv} is the adversarial loss, λL_{rec} is the reconstruction loss from the autoencoder, and λ is a weighting factor.

$$Loss = \alpha \cdot Loss_{16-bit} + \beta \cdot Loss_{32-bit} \quad (2.5)$$

Where α and β are scaling factors for the losses computed in different precisions.

$$I' = T(I) + C(I) + A(I) \quad (2.6)$$

Where I' is the augmented image, $T(I)$ is the translation augmentation, $C(I)$ is the color adjustment, and $A(I)$ is the cutout augmentation.

2.3 Quantitative Evaluation of Synthetic Image Quality

To assess the realism of the generated images, several widely used metrics were employed, including PSNR, SSIM, FID, and IS [27], [40]. These metrics provide a comprehensive evaluation of image quality and authenticity, ensuring a robust analysis of the generated images.

2.3.1 Peak Signal-to-Noise Ratio (PSNR)

The PSNR is a fundamental metric in image processing, widely used to evaluate the quality of images generated by GANs. It quantifies the ratio between maximum pixel intensity and noise that degrades image quality, based on the mean squared error (MSE) between the original and generated images. Expressed in decibels (dB), a higher PSNR indicates closer resemblance to the original image, signifying minimal distortion. In IPM-based applications, high PSNR values confirm that the GAN has effectively replicated crucial structural and geometric features, essential for accurate autonomous vehicle navigation. The formula for PSNR is presented in equation 2.7:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX^2}{MSE} \right) \quad (2.7)$$

- Where
- **MAX** is the maximum possible pixel value of the image. For an 8-bit image, this value is 255.
 - **MSE** is the mean squared error between the original and the generated images.

2.3.2 Structural Similarity Index (SSIM)

SSIM is a crucial metric for evaluating the quality of images generated by GANs, particularly in assessing perceptual realism. Unlike traditional metrics like PSNR that focus on pixel-level differences, SSIM considers luminance, contrast, and structural information, making it more aligned with human visual perception[40]. In this research, SSIM is used to measure how closely generated IPM images resemble ground truth data,

ensuring accurate representation of patterns and textures critical for autonomous vehicle navigation. The SSIM index is mathematically expressed in Equation 2.8.

$$SSIM(x, y) = [L(x, y)]^\alpha \cdot [C(x, y)]^\beta \cdot [S(x, y)]^\gamma \quad (2.8)$$

Where

- x and y are the two images being compared.
- $L(x, y)$ are the luminance, contrast, and structure components, respectively.
- α, β and γ are parameters that determine the relative importance of each component (typically set to 1).

The SSIM value ranges from -1 to 1, where 1 indicates perfect structural similarity between the two images, and values closer to 1 suggest higher image quality.

2.3.3 Fréchet Inception Distance (FID)

The FID is a metric that evaluates the quality and diversity of generated images by comparing their feature distributions to those of real images. Unlike metrics such as PSNR and SSIM, which assess individual images, FID examines the entire image set as a distribution. It is calculated by passing both real and generated images through a pre-trained Inception v3 model to extract features, assuming a multivariate Gaussian distribution. FID then computes the Fréchet distance between these distributions, as mathematically represented in Equation 2.9.

$$FID = \|\mu_r - \mu_g\|^2 + Tr\left(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{\frac{1}{2}}\right) \quad (2.9)$$

Where:

- μ_r and μ_g are the mean vectors of the real and generated image features, respectively.
- Σ_r and Σ_g are the covariance matrices of the real and generated image features, respectively.
- $\|\mu_r - \mu_g\|^2$ is the squared distance between the means of the two distributions.
- $Tr\left(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{\frac{1}{2}}\right)$ represents the trace of the covariance matrices.

2.3.4 Inception Score (IS)

The IS is derived from the predictions of a pre-trained Inception v3 model, which was initially designed for image classification tasks on the ImageNet dataset. The Inception v3 model outputs a probability distribution over different classes for each input image. The Inception Score is calculated based on two key properties: the image's predicted class distribution and the distribution of predictions across the entire dataset.

The IS is calculated as follows equation 2.10:

$$IS = exp\left(E_x\left[KL(p(y \vee x) \vee p(y))\right]\right) \quad (2.10)$$

Where:

- $p(y \vee x)$ is the conditional label distribution given an image x .
- $p(y)$ is the marginal label distribution across the generated dataset.
- $KL(p(y \vee x) \vee p(y))$ is the Kullback-Leibler divergence between the conditional and marginal distributions.

In practice, a higher Inception Score suggests that the GAN has successfully generated a set of images that are both sharp and diverse, characteristics that are essential for effective training datasets in autonomous vehicle navigation.

III. METHODOLOGY

The research involved five key steps as shown in Fig. 1. Initially, a camera was mounted and calibrated on the roof of a vehicle to capture images while the vehicle was in motion. Using the calibration parameters, the acquired images were transformed into BEV images through an IPM transformation. These BEV images served as the dataset for training LGAN. Following the training phase, the output was rigorously analyzed and validated to assess the effectiveness of the LGAN in generating realistic IPM images.

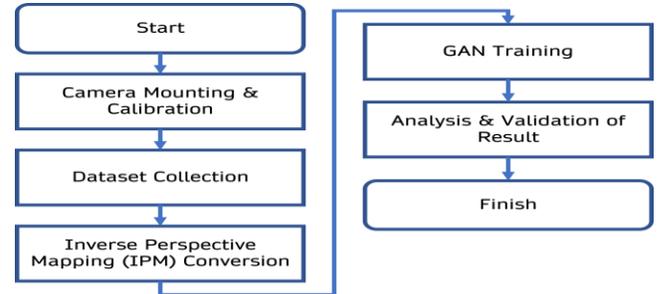


Fig. 1. Research flowchart

3.1 Hardware Description

The specifications of the computer used in the experiments are shown in Table I.

TABLE I: EXPERIMENT HARDWARE SPECIFICATION

Item	Specification
Central Processing Unit (CPU)	AMD Threadripper 3990x
Graphics Processing Unit (GPU)	3 × GTX 1080 Ti
Random Access Memory (RAM)	64 GB
Operating System	Linux Ubuntu 20.04.3 LTS

3.2 Camera Mounting and Calibration

The Point Grey Blackfly camera (BFLY-U3-23S6C-C) featuring a Sony IMX249 CMOS sensor with a resolution of 1920×1200 pixels and a frame rate of 41 FPS was mounted on the roof of a Proton Persona 1.6L at an 8-degree angle and 1.4 meters height. Prior to data collection, the camera was

calibrated using a 6x8-inch checkerboard, which was repositioned to capture various angles. MATLAB calibration software detected checkerboard edges to determine accurate camera settings, and reprojection errors were used to refine these parameters, ensuring precise image acquisition.

3.3 IPM For BEV Conversion

IPM was used to transform the driver’s view samples into BEV. Samples of the conversion are shown in Fig. 2. The IPM used the camera parameters to synthesize the BEV images. The images were then resized into size 1024 x 1024 as inputs to LGAN using bicubic interpolation.



Fig. 2. Samples of BEV conversion

3.4 Data Collection

The road images were taken in an around UiTM Dengkil Branch, Dengkil, Malaysia, during daylight hours in sunny weather. The photographs were taken at a frame rate of approximately 30 frames per second, yielding 12,207 JPEG still images. 1,795 images were removed (Fig. 3) from the dataset due to several factors such as camera initialization, pedestrian, and road occlusion, resulting in the final 10,412 images. Several training samples are shown on Fig. 4.

In developing a vision system for an autonomous vehicle (AV) using Inverse Perspective Mapping (IPM), it was essential to begin with controlled conditions to establish a reliable baseline. Sunny weather provides optimal lighting, reducing potential distortions from shadows, rain, or fog. This ensures that errors in the perspective transformation are attributed to the IPM method itself rather than external environmental factors. By first validating the system under ideal conditions, we ensure that fundamental geometric transformations and feature extractions, such as lane markings and object distances, are accurately processed before introducing additional complexities.

The decision to record data exclusively in sunny weather follows a standard incremental validation approach commonly used in AV development. Initial testing under clear conditions helps isolate and fine-tune system parameters without interference from adverse weather artifacts. Once the core system is validated, the next phases will involve testing under diverse conditions, including rain, fog, and low-light environments. This step-by-step methodology prevents compounding errors and allows for systematic debugging, making it easier to identify whether performance issues arise from the IPM framework itself or from environmental variations.

3.5 GAN Training

GAN training involves two competing networks: the generator and the discriminator. Both are optimized simultaneously using the Adaptive Moment Estimation (ADAM) algorithm, chosen for its computational efficiency and modest memory requirements, particularly in handling networks with many parameters. Unlike traditional backpropagation, ADAM requires minimal parameter tuning, making it suitable for complex training processes.

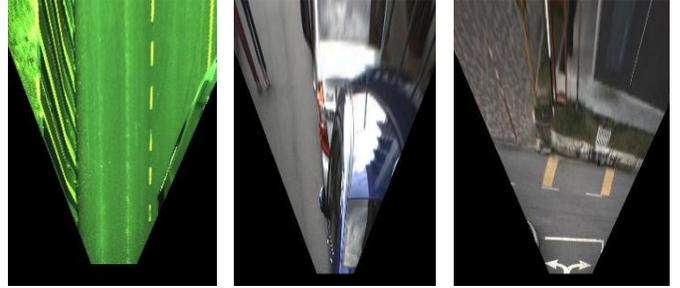


Fig. 3. Removed samples (from left: initialization, containing vehicles, camera tilted during data collection).



Fig. 4. Accepted training samples

The GAN structure in this study follows the LGAN implementation, which integrates skip-layer excitation in the generator and autoencoding self-supervised learning in the discriminator [16]. These modifications enhance training speed, even on modest hardware, by improving information flow and stability. During training, the generator attempts to create synthetic images that deceive the discriminator, while the discriminator strives to accurately differentiate between real and fake images. This adversarial process refines both networks until the discriminator can no longer distinguish between authentic and generated images, leading to high-quality outputs. For details on the LGAN architecture, refer to [16]. The training parameters for LGAN are shown in Table II.

TABLE II: TRAINING PARAMETER VALUES

Parameter	Value
Training Algorithm	Adaptive Moment Estimation (ADAM)
Epochs	200,000
Output Size	512 × 512 pixels

IV. RESULTS AND DISCUSSION

The application of GANs for generating realistic inverse perspective mapping IPM images represents a significant advancement in autonomous vehicle navigation systems. As shown in Table III, this study produced 1,653 images from an

original dataset of 1,307, achieving a 75% acceptance rate, with 1,238 images deemed suitable for further use. This outcome highlights both the potential of GANs to augment training datasets and their limitations, underscoring the need for continued refinement.

Table IV demonstrates that the LGAN effectively expanded the dataset, enhancing robustness by increasing data diversity. However, the rejection of 415 images points to challenges such as mode collapse, insufficient diversity, or inaccuracies in perspective rendering. These limitations indicate areas for improvement in the GAN architecture and training process to further enhance image quality. Rejected images, while inevitable, provide critical feedback during the GAN training, as the generator iteratively learns to produce higher-quality outputs. This competitive dynamic between the generator and discriminator ensures continuous refinement, leading to more realistic and reliable results.

TABLE III: THE DISTRIBUTION OF ORIGINAL IMAGES

Type	Acceptable
Straight	503
Curve Left	400
Curve Right	404
Total	1307

TABLE IV: THE DISTRIBUTION OF THE ACCEPTED GENERATED IMAGES

Type	Acceptable
Straight	686
Curve Left	434
Curve Right	118
Reject	415
Total	1653

4.1 Comparison Between Accepted Images and Original Images

A comparative analysis between the accepted IPM images (Fig. 5(a)) and the original samples (Fig. 5(b)) reveals a high degree of visual similarity, underscoring the suitability of the generated images for training autonomous vehicle navigation systems. The accepted images retain critical visual features such as well-defined edges, consistent lane markings, and clearly delineated road boundaries. These attributes are essential for accurate lane detection and localization, enabling machine learning models to interpret and navigate complex road environments with greater reliability.

In addition to geometric consistency, the accepted images exhibit high fidelity in texture and color, closely mirroring the visual characteristics of the original dataset. As shown in Fig. 5(a), the realistic rendering of road surface textures and color variations enhances the authenticity of the synthetic data, thereby improving model performance under real-world driving conditions. The perspective transformations in the accepted images preserve accurate spatial proportions, which are crucial for tasks such as distance estimation and path planning in autonomous systems.

Another significant quality of the accepted IPM images is their minimal presence of visual noise and artifacts. In contrast to rejected samples, these images are free from distortions or

speckles that could compromise learning accuracy. Lighting and shadow consistency further reinforces their realism, contributing to more robust depth perception and spatial orientation. The clarity of lane markings, realistic shading, and faithful replication of surface textures collectively demonstrate the LGAN's effectiveness in generating high-quality, application-ready training data. Overall, the accepted images reflect the LGAN's potential to enrich datasets with realistic, diverse samples that enhance the generalization capabilities of autonomous vehicle navigation models.

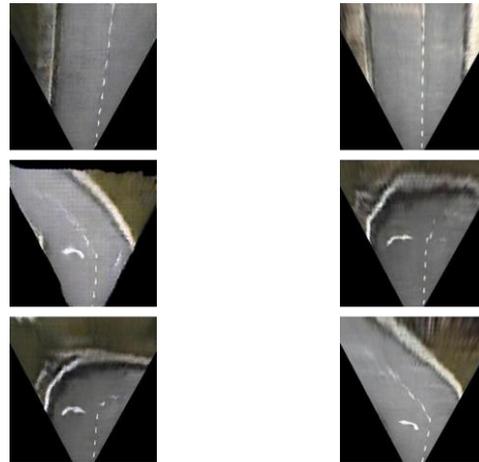


Fig. 5(a). Accepted sample

Fig. 5(b). Sample from original

Fig. 5. Comparison between accepted images and original images

4.2 Comparison Between Rejected Images and Original Images

A comparison between the rejected inverse perspective mapping (IPM) images (Fig. 6(a)) and the original images (Fig. 6(b)) highlights several limitations that make the rejected samples unsuitable for training autonomous vehicle navigation systems. The most prominent issue is the lack of visual clarity and structural consistency. Blurred and misaligned lane markings particularly around curves are frequently observed, undermining the accuracy of lane detection. These distortions compromise the ability of machine learning models to extract reliable spatial cues, which are essential for precise lane positioning and safe navigation in dynamic driving environments.

Moreover, the rejected images exhibit unnatural texture patterns, color inconsistencies, and significant perspective distortions. These artifacts introduce discrepancies that deviate from real-world conditions, potentially confusing learning algorithms during training. Perspective inaccuracies result in misrepresented road geometries, further degrading spatial awareness. Additionally, the presence of visual noise such as random speckles, streaks, and geometric warping reduces image quality and can adversely affect model generalization. In contrast, the original images offer clean, geometrically accurate data suitable for learning robust navigation behaviors. Overall, the rejected samples in Fig. 6(a) demonstrate critical deficiencies that diminish their utility, emphasizing the need for improvements in GAN architecture, loss functions, and image

quality control to ensure the generation of high-fidelity IPM images for autonomous driving applications.

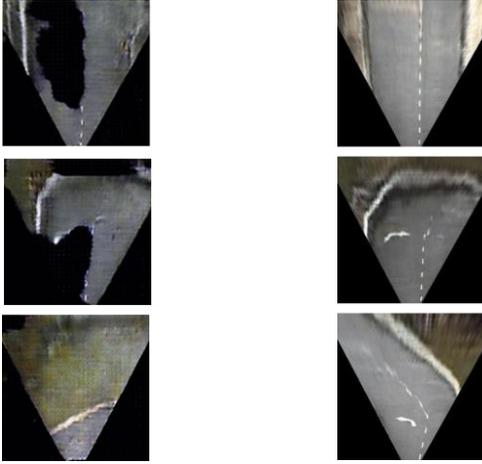


Fig 6(a). Rejected sample

Fig. 6(b). Sample from original

Fig. 6. Comparison between accepted images and original images

4.3 Images Training Generated

In this study, LGANs were utilized to generate IPM images for autonomous vehicle training. To improve output stability and quality, the Exponential Moving Average (EMA) technique was applied alongside standard (normal) image generation. EMA operated by averaging the generator's weights across multiple iterations, thereby smoothing fluctuations during training. This approach helped reduce noise and enhanced the consistency of the generated images.

A visual comparison between normal and EMA-generated images revealed clear distinctions in quality and structure. As illustrated in Fig. 7(a), normal outputs frequently exhibited artifacts, irregular textures, and inconsistencies resulting from unstable training dynamics. Conversely, EMA outputs, shown in Fig. 7(b), displayed improved clarity, uniform textures, and well-defined features. This improvement demonstrated the effectiveness of EMA in mitigating training noise, preserving key structural elements, and enhancing the realism of synthetic IPM images. Overall, EMA proved to be a valuable method for producing higher-fidelity data suitable for training robust autonomous vehicle navigation models.

4.4 Iteration Analysis

The iterative development of the LGAN in generating inverse perspective mapping (IPM) images is illustrated in Figure 8. At 0 iterations (Fig. 8(a)), the LGAN is entirely untrained, producing a blank black image that reflects the absence of learned features or structure. By 5,000 iterations (Fig. 8(b)), the model begins to generate rudimentary geometric patterns; however, these outputs remain noisy and incoherent, indicating that the network is only beginning to capture the underlying data distribution.

Improvements become more evident at 10,000 iterations (Fig. 8(c)), where triangular shapes start to form with greater clarity and reduced noise, although challenges related to texture and contrast remain. By 15,000 iterations (Fig. 8(d)), the

network exhibits more stable and consistent outputs, with clearer geometric patterns and fewer artifacts, suggesting that the LGAN has begun to internalize the structural characteristics of IPM images.

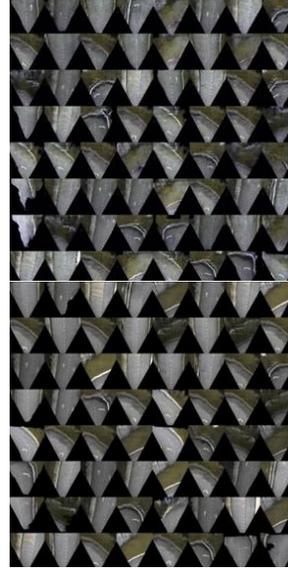


Fig. 7(a). Normal image

Fig. 7(b). EMA image

Fig. 7. Comparison between Normal and EMA

From 20,000 to 30,000 iterations (Fig. 8(e)–8(g)), the LGAN produces high-fidelity, visually coherent images that exhibit accurate textures, consistent shading, and realistic spatial distribution. This stage marks the model's convergence toward the target data distribution, demonstrating its ability to synthesize structured and high-quality IPM images. The progression from untrained randomness to refined outputs highlights the effectiveness of the LGAN's learning process, reinforcing its suitability for generating synthetic data in autonomous vehicle navigation tasks.

4.5 Quantitative Performance Assessment

Four primary metrics, PSNR, SSIM, FID, and IS are used to evaluate the quality of LGAN-generated images, each providing unique insights into image fidelity and realism. PSNR measures pixel-level accuracy by comparing generated images to reference images, focusing on image fidelity. SSIM extends this by assessing structural similarity, capturing luminance and contrast aspects that align more closely with human perception. FID compares the feature distribution between real and generated images, offering a comprehensive measure of realism and diversity. IS evaluates both the recognizability and variety of generated images using a pre-trained classifier, thus addressing quality and diversity across classes. Collectively, these metrics provide a robust framework for assessing the fidelity, perceptual quality, and distributional alignment of GAN outputs.

Table V presents a comparative analysis of normal images and those generated using EMA technique, with each image set evaluated through FID, IS, PSNR, and SSIM. The data suggests that the EMA technique significantly enhances image quality and realism. For example, FID, which quantifies the similarity between generated and real image distributions, is substantially

lower for EMA images (131.5848) than for normal images (174.2456), indicating that EMA images are more closely aligned with real data. Similarly, PSNR, which evaluates pixel fidelity, improves from 15.3717 for normal images to 15.7947 for EMA images, suggesting lower distortion levels. SSIM also increases from 0.3491 to 0.3875, indicating improved structural similarity and a closer match to reference images.

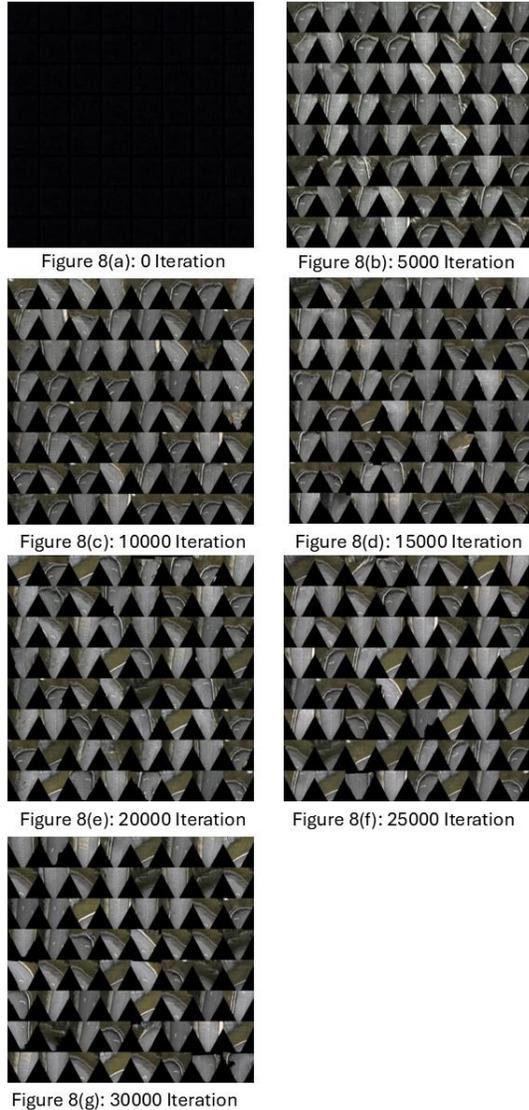


Fig. 8. Iteration Between 0 Iteration and 30,000 Iteration

TABLE V: COMPARATIVE ANALYSIS OF TWO SETS OF GENERATED IMAGES

Test	FID	IS	PSNR	SSIM
Normal Images	174.2456	10.8954 ± 1.2970	15.3717	0.3491
EMA Images	131.5848	10.5015 ± 1.2386	15.7947	0.3875

Although the IS shows a slight decrease for EMA images from 10.8954 ± 1.2970 to 10.5015 ± 1.2386 this reduction is minimal. IS measures both image diversity and quality, and the slight drop may suggest a small compromise in diversity while overall realism and fidelity improve. This trade-off appears favorable, as the reduced FID and increased PSNR and SSIM

demonstrate enhanced realism and structural accuracy in EMA images compared to normal outputs. In summary, the EMA technique proves effective in producing high-quality and realistic IPM images, with minimal impact on diversity, making it a valuable enhancement for GAN performance in applications requiring high fidelity and realism.

V. CONCLUSION

This research has demonstrated the effectiveness of using LGANs to generate realistic IPM images for enhancing the training of AV navigation systems. By increasing the dataset from 1,307 to 1,653 images with a 75% acceptance rate the GAN significantly improved data diversity, a critical factor in training robust models for lane detection and curvature estimation. The contrast between accepted and rejected images highlights the importance of visual fidelity in terms of texture, color, and structural consistency, which directly affects the utility of synthetic images for AV training.

Moreover, the integration of the EMA technique during LGAN training further enhanced image quality. EMA generated images exhibited greater clarity, reduced noise, and superior structural stability compared to standard outputs. Quantitative analysis confirmed these improvements, with EMA images achieving lower FID, higher PSNR, and improved SSIM. These findings validate EMA as an effective technique for stabilizing GAN training and producing high-quality IPM images suitable for real-world AV applications.

Beyond technical validation, the findings have promising implications for real-world AV development. The ability to generate diverse and realistic IPM images offers a scalable solution for augmenting datasets in data-scarce environments, particularly under rare or hazardous driving conditions. These synthetic datasets can accelerate the training of perception models for commercial AVs, support the testing of navigation algorithms in simulated environments, and reduce the dependence on expensive and time-consuming data collection processes. Practical implementation could involve deploying lightweight GAN architectures in edge-computing platforms on AVs or within cloud-based training pipelines.

Future research should aim to further generalize the model across varying environmental conditions, including rain, fog, and nighttime scenarios. This may involve combining real-world datasets with synthetic augmentation strategies or integrating adaptive IPM methods driven by deep learning. By addressing these practical considerations, this work takes a significant step toward scalable, efficient, and reliable data augmentation for real-world AV training.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support of Universiti Teknologi MARA. This work was funded by the Ministry of Higher Education (MOHE) under the Fundamental Research Grant Scheme (600-RMC/FRGS 5/3 (008/2024)).

REFERENCES

[1] K. Dietmayer, H. Winner, M. Maurer, K. Bengler, C. Stiller, and B. Farber, "Three Decades of Driver Assistance Systems: Review and Future

- Perspectives,” *IEEE Intell. Transp. Syst. Mag.*, vol. 6, no. 4, pp. 6–22, 2014, doi: 10.1109/mits.2014.2336271.
- [2] Nguyen and Tuan C, “History of Self-Driving Cars,” *ThoughtCo*, 26-Jan-2021. [Online]. Available: <https://www.thoughtco.com/history-of-self-driving-cars-4117191>. [Accessed: 05-May-2021].
 - [3] S. D. Pendleton et al., “Perception, planning, control, and coordination for autonomous vehicles,” *Machines*, vol. 5, no. 1, 2017, doi: 10.3390/machines5010006.
 - [4] S. E. Shladover, “Connected and automated vehicle systems: Introduction and overview,” *J. Intell. Transp. Syst. Technol. Planning, Oper.*, vol. 22, no. 3, 2018, doi: 10.1080/15472450.2017.1336053.
 - [5] N. Zulkifli et al., “Fast region convolutional neural network lane and road defect detection for autonomous vehicle application,” *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 1.3 S1, 2019, doi: 10.30534/ijatcse/2019/6681.32019.
 - [6] Y. Tian et al., “Lane marking detection via deep convolutional neural network,” *Neurocomputing*, vol. 280, pp. 46–55, Mar. 2018, doi: 10.1016/j.neucom.2017.09.098.
 - [7] J. Li, X. Mei, D. Prokhorov, and D. Tao, “Deep Neural Network for Structural Prediction and Lane Detection in Traffic Scene,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 28, no. 3, pp. 690–703, Mar. 2017, doi: 10.1109/TNNLS.2016.2522428.
 - [8] R. Karthika and L. Parameswaran, “An automated vision-based algorithm for out of context detection in images,” *Int. J. Signal Imaging Syst. Eng.*, vol. 11, no. 1, pp. 1–8, 2018, doi: 10.1504/IJSISE.2018.090601.
 - [9] S. Kadam, S. Bhandari, P. Sontakke, and S. Sawant, “GPU-based video-processing traffic signals for high-density vehicle areas,” *Int. J. Signal Imaging Syst. Eng.*, vol. 13, no. 1, pp. 1–9, 2024, doi: 10.1504/IJSISE.2024.139977.
 - [10] B. Ashwini and B. N. Yuvaraju, “Application of machine learning approach in detection and classification of cars of an image,” *Int. J. Signal Imaging Syst. Eng.*, vol. 10, no. 1/2, p. 8, 2017, doi: 10.1504/IJSISE.2017.084564.
 - [11] Cao, Song, Song, Xiao, and Peng, “Lane Detection Algorithm for Intelligent Vehicles in Complex Road Conditions and Dynamic Environments,” *Sensors*, vol. 19, no. 14, p. 3166, Jul. 2019, doi: 10.3390/s19143166.
 - [12] S. O. Patil, V. V. Sajith Variyar, and K. P. Soman, “Speed Bump Segmentation an Application of Conditional Generative Adversarial Network for Self-driving Vehicles,” in *Proceedings of the 4th International Conference on Computing Methodologies and Communication, ICCMC 2020*, 2020, doi: 10.1109/ICCMC48092.2020.ICCMC-000173.
 - [13] W. Xu, N. Souly, and P. P. Brahma, “Reliability of GAN Generated Data to Train and Validate Perception Systems for Autonomous Vehicles,” in *Proceedings - 2021 IEEE Winter Conference on Applications of Computer Vision Workshops, WACVW 2021*, 2021, doi: 10.1109/WACVW52041.2021.00023.
 - [14] M. Ravanbakhsh et al., “Learning Self-Awareness for Autonomous Vehicles: Exploring Multisensory Incremental Models,” *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, 2021, doi: 10.1109/TITS.2020.2984735.
 - [15] H.-J. Yoon, R. Holmes, H. Jafarnejadsani, and P. Voulgaris, “Real-time Adversarial Image Perturbations for Autonomous Vehicles using Reinforcement Learning,” *ACM Trans. Cyber-Physical Syst.*, Jan. 2025, doi: 10.1145/3712303.
 - [16] B. Liu, Y. Zhu, K. Song, and A. Elgammal, “Towards Faster and Stabilized GAN Training for High-fidelity Few-shot Image Synthesis,” pp. 1–22, Jan. 2021.
 - [17] F. Lateef, M. Kas, and Y. Ruichek, “Saliency Heat-Map as Visual Attention for Autonomous Driving Using Generative Adversarial Network (GAN),” *IEEE Trans. Intell. Transp. Syst.*, 2021, doi: 10.1109/TITS.2021.3053178.
 - [18] X. Ding, Y. Wang, Z. Xu, W. J. Welch, and Z. J. Wang, “Continuous Conditional Generative Adversarial Networks: Novel Empirical Losses and Label Input Mechanisms,” pp. 1–49, 2020.
 - [19] L. W. Kang, C. C. Hsu, I. S. Wang, T. L. Liu, S. Y. Chen, and C. Y. Chang, “Vehicle trajectory prediction based on social generative adversarial network for self-driving car applications,” in *Proceedings - 2020 International Symposium on Computer, Consumer and Control, IS3C 2020*, 2020, doi: 10.1109/IS3C50286.2020.00133.
 - [20] S. Julka, V. Sowrirajan, J. Schloetterer, and M. Granitzer, “Conditional Generative Adversarial Networks for Speed Control in Trajectory Simulation,” pp. 436–450, 2022, doi: 10.1007/978-3-030-95470-3_33.
 - [21] M. Qi, Y. Wang, A. Li, and J. Luo, “STC-GAN: Spatio-Temporally Coupled Generative Adversarial Networks for Predictive Scene Parsing,” *IEEE Trans. Image Process.*, vol. 29, pp. 5420–5430, 2020, doi: 10.1109/TIP.2020.2983567.
 - [22] L. Huang, J. Zhuang, X. Cheng, R. Xu, and H. Ma, “STI-GAN: Multimodal Pedestrian Trajectory Prediction Using Spatiotemporal Interactions and a Generative Adversarial Network,” *IEEE Access*, vol. 9, pp. 50846–50856, 2021, doi: 10.1109/ACCESS.2021.3069134.
 - [23] N. Ma, J. Wang, J. Liu, and M. Q. H. Meng, “Conditional Generative Adversarial Networks for Optimal Path Planning,” *IEEE Trans. Cogn. Dev. Syst.*, 2021, doi: 10.1109/TCDS.2021.3063273.
 - [24] Z. Hu, T. Turki, and J. T. L. Wang, “Generative Adversarial Networks for Stochastic Video Prediction with Action Control,” *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2982750.
 - [25] J. Guo, Q. Yang, S. Fu, R. Boyles, S. Turner, and K. Clarke, “Towards Trustworthy Perception Information Sharing on Connected and Autonomous Vehicles,” in *Proceedings - 2020 International Conference on Connected and Autonomous Driving, MetroCAD 2020*, 2020, doi: 10.1109/MetroCAD48866.2020.00021.
 - [26] Z. Xiong, Z. Cai, Q. Han, A. Alrawais, and W. Li, “ADGAN: Protect Your Location Privacy in Camera Data of Auto-Driving Vehicles,” *IEEE Trans. Ind. Informatics*, vol. 17, no. 9, 2021, doi: 10.1109/TII.2020.3032352.
 - [27] A. Obukhov and M. Krasnyanskiy, “Quality Assessment Method for GAN Based on Modified Metrics Inception Score and Fréchet Inception Distance,” 2020, pp. 102–114.
 - [28] J. Spooner, V. Palade, M. Cheah, S. Kanarachos, and A. Daneshkhan, “Generation of pedestrian crossing scenarios using ped-cross generative adversarial network,” *Appl. Sci.*, vol. 11, no. 2, 2021, doi: 10.3390/app11020471.
 - [29] A. Yadav, B. S. Avinash, and B. Kumar, “Lightweight Generative Adversarial Network (LGAN) Architectures for Resource-Constrained Hardware Applications,” in *2024 IEEE 5th India Council International Subsections Conference (INDISCON)*, 2024, pp. 1–5, doi: 10.1109/INDISCON62179.2024.10744337.
 - [30] Y. Gong, G. Duan, Z. Ma, and M. Xie, “Lightweight semantic image synthesis with mutable framework,” *J. Electron. Imaging*, vol. 32, no. 03, Jun. 2023, doi: 10.1117/1.JEI.32.3.033027.
 - [31] B. Cao and H. Lei, “LGEGAN: A Lightweight Evolutionary Generative Adversarial Network with Statistic Global Information,” in *2023 42nd Chinese Control Conference (CCC)*, 2023, pp. 8282–8287, doi: 10.23919/CCC58697.2023.10241151.
 - [32] C. Wang, Z. Li, Y. Hao, L. Wang, and X. Li, “Faster Projected GAN: Towards Faster Few-Shot Image Generation,” Jan. 2024.
 - [33] B. Li and T. Lukasiewicz, “Lightweight Long-Range Generative Adversarial Networks,” Sep. 2022.
 - [34] H. Yu, H. Zhu, X. Lu, and J. Liu, “Migrating Face Swap to Mobile Devices: A Lightweight Framework and a Supervised Training Solution,” in *2022 IEEE International Conference on Multimedia and Expo (ICME)*, 2022, pp. 1–6, doi: 10.1109/ICME52920.2022.9859806.
 - [35] S. Li, K. Yang, H. Shi, S. Wang, Y. Yao, and Z. Li, “GenMapping: Unleashing the Potential of Inverse Perspective Mapping for Robust Online HD Map Construction,” Sep. 2024.
 - [36] N. Markó, P. Kőrös, and M. Unger, “Inverse Perspective Mapping Correction for Aiding Camera-Based Autonomous Driving Tasks,” in *SMTS 2024*, 2024, p. 67, doi: 10.3390/engproc2024079067.
 - [37] X. Mu, H. Ye, D. Zhu, T. Chen, and T. Qin, “Inverse Perspective Mapping-Based Neural Occupancy Grid Map for Visual Parking,” in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 8400–8406, doi: 10.1109/ICRA48891.2023.10160849.
 - [38] M. Rao, L. Paulino, V. Robila, I. Li, M. Zhu, and W. Wang, “Training-Free Lane Tracking for 1/10th Scale Autonomous Vehicle Using Inverse Perspective Mapping and Probabilistic Hough Transforms,” in *2022 IEEE Integrated STEM Education Conference (ISEC)*, 2022, pp. 406–411, doi: 10.1109/ISEC54952.2022.10025098.
 - [39] M. A. Rafique, M. I. Hussain, S. Dubey, K. Sayfullokh, and M. Jeon, “A Monocular Camera Bird-Eye-View Generation using Lane Markers Prior,” in *2022 22nd International Conference on Control, Automation and Systems (ICCAS)*, 2022, pp. 1137–1142, doi: 10.23919/ICCAS55662.2022.10003957.
 - [40] I. H. El-Shal, O. M. Fahmy, and M. A. Elattar, “License Plate Image Analysis Empowered by Generative Adversarial Neural Networks (GANs),” *IEEE Access*, vol. 10, pp. 30846–30857, 2022, doi: 10.1109/ACCESS.2022.3157714.