Behavior Varied Adversarial Pedestrian Generation for Autonomous Vehicle Testing

Anonymous Author(s)
Affiliation
Address
email

Abstract: There exist several datasets for developing self-driving car methodologies. However, manually collected test datasets impose inherent limitations on the variability of test cases, and it is particularly difficult to acquire sufficiently challenging scenarios. In order to overcome these constraints, we propose to automatically generate difficult scenarios in urban environments, by learning the placements of pedestrians such that the induced scenarios are highly challenging for a given autonomous vehicle (AV). In difference to previous methods we separate the pedestrian behavior model from the scenario generation network, which allows for testing the car model with various pedestrian behaviors. We show empirically that it is possible to train a test case scenario generation model with collision avoiding pedestrian model.

Keywords: Autonomous Vehicles, AV Testing, Reinforcement Learning

1 Introduction

Research in autonomous car models has gained momentum in the recent years [1]. With the increase of various autonomous models the need for extensive evaluation of different autonomous models has increased. The lack of varied safety-critical testing can lead to safety issues and can cost lives in applications [2]. There exists several datasets [3, 4, 5, 6, 7, 8, 9, 10] for developing and evaluating self-driving car methodologies. However, manually collected test data has inherent limitations on the variability of test cases. Further it is particularly difficult to manually acquire sufficiently challenging scenarios without undesired safety risks. This has lead to the development of generative test case generation methods [11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29]. All of the above methods either assume that the pedestrian motion can be modelled by a simple constant velocity model or that the pedestrian motion is directly adversarial to the car. In practice however collisions do not occur just when a pedestrian has a perfectly predictable path (i.e. the constant velocity), or when the pedestrian is suicidal and actively seeks to get killed by the car, i.e. adversarial pedestrian. Some collisions occur because pedestrians are distracted, due to occlusions or noise. We wish to alleviate the assumption on pedestrian motion and study if we can also generate test cases where the car and pedestrian collide even though both attempt to avoid one another. To do so we learn a distribution of initial pedestrian locations μ that lead to collisions with the car. To our knowledge we are the first to study the effects of different pedestrian behaviours π on the distribution of pedestrian locations μ.

Separating the problem of pedestrian location distribution μ from modelling pedestrian behavior π allows us to in detail look at the pedestrian location distribution μ. The problem contains some structure that we hope to learn. Pedestrian distribution in scenes is affected by the natural scene semantics. For example pedestrians are more likely to preside near or on sidewalks and on crosswalks. Further pedestrians are less likely to be observed by a car when they are occluded from the car's perspective. Finally clearly the probability of a collision depends on the pedestrian's location in relation to the car and the car's speed. Following these observations we propose a model in Fig. 1 that learns the pedestrian location distribution μ conditioned on the scene semantics S, occlusion map O, temporal mapping of dynamic objects D and a distance dependent prior P to the car.

Submitted to the 5th Conference on Robot Learning (CoRL 2021). Do not distribute.

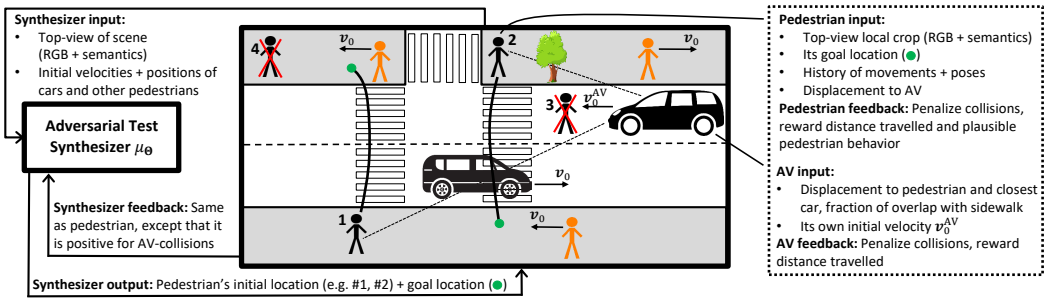

Figure 1: Overview of the proposed safety critical test case generation for autonomous vehicles (AVs). The pedestrian location distribution μ is trained to place out pedestrians with behavior model π such that the car ρ (car to the right) is maximally challenged by the induced scenario. Four example scenarios are shown, where the first two are numbered in ascending order of difficulty. Scenarios #3 and #4 are considered illegal, as in the real world a pedestrian cannot simply appear from nowhere into the line of sight of an AV. Scenario #1: The pedestrian is quite close to the AV and it is not in the line of sight due to the occluding car. Hence this is a challenging scenario. Scenario #2: The pedestrian is close to the AV and is occluded by the tree, so this is a highly challenging scenario.

We propose to treat both the car and the pedestrian as goal driven black-box models to allow for the testing of any pedestrian and car models. The general problem is then to find a pedestrian location distribution μ such that the number of collisions between a black-box car and a black-box pedestrian are maximal in expectation. The problem can be seen as a black box optimization or as a reinforcement learning problem. Bayesian Optimization scales poorly as the number of parameters increases. This in practice means that a parametric distribution needs to be fitted to the pedestrian location distribution μ. We have strong reason to believe that the optimal pedestrian distribution μ is not easy to model since we expect a strong correlation between occluded spaces and high-collision pedestrian locations. To avoid any assumptions about μ's shape we instead use the reinforcement learning approach. The pedestrian and the car are then seen as a part of the unknown model dynamics, and the pedestrian location distribution μ is the learnable stochastic policy. This allows us to model μ as a convolutional neural network μ_Θ.

In our formulation the pedestrian location distribution μ and the car motion ρ are playing an indirect constrained min-max game. The pedestrian behavior π can be adversarial to the car or the pedestrian location distribution μ, or even not directly cooperative to either side. The problem then becomes a three agent mini-max game with up to two agents per team. Viewing traffic as a multi-agent game is a well studied topic within multi agent reinforcement learning [30]. We extend the concept by viewing the pedestrian location distribution μ as an additional agent. The study of the equilibrium is beyond this paper however it is clear that if the car ρ, pedestrian behavior π and pedestrian location distribution μ are unconstrained then the minimax problem has a trivial solution. If the pedestrian is always initialized arbitrary close to the front of the car then this will always lead to a collision. If the pedestrian and the car always stand still or always move in opposite directions, then there are never any collisions. To avoid trivial solutions sufficient amount of additional constraints need to modelled.

To our knowledge we are the first to provide a proof of concept study of the effects of the pedestrian behavior policies π on the pedestrian location distribution μ in the domain of test case generation for autonomous vehicles. To do so we utilize the semantic pedestrian motion generation model of [31]. In experiments we articulate the studied pedestrian policies by [32] to ensure that plausible human dynamics are proposed by all studied pedestrian behaviours π. We note that our results provide some proof of concept insights to the qualities of the pedestrian behavior model π on test case generation. In the future work we hope to extend the experiments with realistic pedestrian detection noise modelling for different sensors, with state of the art car models ρ, and by extending the model to all of the observable traffic entities.

2 Related Work

Previous work [14, 16, 17] have studied the generation of pedestrian trajectories η such that the number of collisions is maximized. Where η includes implicitly the pedestrian position distribution μ and the pedestrian behavior π. When the objective J is maximized with respect to η then both π and μ

become adversarial to the car's objective J_ρ. This leads to the pedestrian only behaving suicidal. This is unnecessarily limiting as the pedestrian behavior could also be helpful in avoiding collisions. In testing we wish to ultimately ensure that the car can avoid collisions both with adversarial pedestrians but also with collision avoiding pedestrians. To allow for both we separate η explicitly into the initial pedestrian location μ distribution and the pedestrian behavior π. In [11] pedestrians are modelled with constant velocity, and are initialized from set positions. The AV is retrained in a loop with the test case generator. [14] existing trajectories are adapted to become adversarial. The suggested method is interesting but requires a varied ground truth dataset and the generated data is dependent on variability of the existing dataset.

Further the study of visual relations in data from the autonomous vehicle's perspective [33, 34, 35] has been raised recently. In [33] a visual prior is learnt for where pedestrians and other objects can appear in a camera mounted on a AV. This illustrates that the appearance of pedestrians in scenes should be structured to be visually plausible. A similar problem of realistic object placement in lidar scenes is studied in [34]. Finally [35] show that visual cues from an on-board camera can be used to learn walkeable areas in a scene.

3 Methodology

To illustrate the relation between the different models' loss functions, let the number of collisions between the car model ρ and the pedestrian model be measured by an indicator function I that is 1 if a collision occurs between the car and the pedestrian. In the proceeding, we will assume that all models' loss functions can be composed as an expectation over the number of collisions and other loss components. The car model ρ is learnt by optimizing a loss function J^ρ, that can be assumed to minimize the $\mathbb{E}[I]$ under some desired behavioral constraints B_ρ. The pedestrian spatial distribution model's loss μ has a loss function J^μ that maximizes $\mathbb{E}[I]$ under some behavioural constrains B_ρ. Collectively the problem can then be expressed as

$$\min_\rho \mathbb{E}_{\mu,\rho,\pi}[I] + f_\rho(\rho) \text{ s.t. } \rho \in B_\rho \tag{1}$$

$$\max_\mu \mathbb{E}_{\mu,\rho,\pi}[I] + f_\mu(\mu) \text{ s.t. } \mu \in B_\mu \tag{2}$$

where f_μ, f_ρ contain any other terms of J^μ and J^ρ respectively. Equations 1 express the general optimization problem of ρ and μ when the pedestrian behaviour π is constant or independent of $\mathbb{E}[I]$. If the pedestrian behaviour model π is collision avoiding then the problem gains a third equation that minimizes $\mathbb{E}[I]$. If the pedestrian behaviour model π is adversarial then the third equation will maximize $\mathbb{E}[I]$. From this it it is clear that the behaviour policy π affects the solution of both μ and ρ. Further depending on the choice of π the set of applied constraints B_ρ, B_μ may need to be adjusted to ensure that none of the models converge to a trivial solution. Previous work has considered the case of adversarial or constant velocity pedestrian behaviour policy π. In our experiments we illustrate that with sufficient constraints on μ, π, ρ a non-trivial solution exists even for collision avoiding π.

3.1 Special Case: Three Policy Gradient Agents

We adopt a reinforcement learning (RL) perspective and study the problem as a sequential problem, where at each timestep the different agents make decisions conditioned on their current observation s_t of the current state of the world. The actors take actions a_t, which affect the state of the world in the next timestep – moreover, in training a reward r may be given as feedback from the environment. The proposed solution consist of the following components:

- A goal driven car model $\rho(s_t^\rho)$ with a model specific state s_t^ρ at timestep t;
- A pedestrian behavior model $\pi(s_t^\pi, g^\pi)$ with a state s_t^π at timestep t, and a goal location g^π;
- A learnable distribution $\mu(S, D, O, P)$ over pedestrian initial locations;
- External world dynamics and noise $p(s_{t+1}|s_t)$, where s_t is the current state of the world, including static objects, the autonomous vehicle, the pedestrian, as well as any external traffic participants (e.g. other cars and pedestrians). Any additional external dynamics or noise could be included here, such as the car's dynamics.

To simplify notation we will use s_t to denote the current time step's global state (therefore including s_t^ρ, s_t^π). Similarly we denote by a_t a vector containing the action taken by the car a_t^ρ and the action taken by the pedestrian a_t^π the at timestep t.

In our experiments all three of the learnable models ρ, π, μ are modelled by a policy gradient model and share the loss function

$$J = \mathbb{E}_{s_0 \sim \mu, a_t^p \sim \pi, a_t^c \sim \rho, , s_t \sim p} \left[\sum_{t=0}^{T} \gamma r(s_t, a_t^\pi, a_t^\rho, s_{t+1}) \right], \tag{3}$$

where $r = [r_\mu, r_\pi, r_\rho]$ is a reward vector containing the reward r_μ of the pedestrian distribution policy μ, the reward r_π of the pedestrian behaviour policy π and the reward r_ρ of the car policy ρ. The loss functions' dependence on I is expressed in the different reward functions. To simplify notations let the state, action history $(s_0, a_0, ..., s_T, a_T, s_{T+1})$ of an episode be denoted τ, and the discounted cumulative reward $R = \sum_{t=0}^{T} \gamma^t r(s_t, a_t, s_{t+1})$. We can express (3) as $\mathbb{E}[R] = \int_\tau p_\tau(\tau) R(\tau) d\tau$, where p_τ is the probability density function of τ. Assuming finite sequences allows the model to be rewritten showing the relation between μ and the rewards attained during an the episode,

$$\mathbb{E}[R] = \int_\tau \sum_{t=0}^{T} \gamma^t r(s_t, a_t, s_{t+1}) \mu_\Theta(s_0) \prod_{t=0}^{T} \pi(a_t^p | s_t) \rho(a_t^c | s_t) p(s_{t+1} | s_t, a_t) \mathrm{d}\tau. \tag{4}$$

From (4) it is clear that μ can be seen as a policy gradient model with the one step reward $\sum_{t=0}^{T} \gamma^t r(s_t, a_t, s_{t+1})$. Now if all of the environment dynamics are known, we could try to find the closed form solution of the integral (4). Using Bellman equations would then allow for a white-box treatment of the car and pedestrian dynamics. Otherwise Monte Carlo estimates can be used to solve the problem. Following REINFORCE [36] the car ρ and the pedestrian distribution μ can be learnt simultaneously as shown in the supplement. Further even the pedestrian behaviour policy can be learnt simultaneously with ρ and μ if π can be learnt only by sampling from μ, ρ and π.

3.2 Pedestrian Initial Spatial Distribution

In our experiments the pedestrian initial distribution μ is modelled as a policy gradient agent. Its objective is to provide an initial position x_0 to the pedestrian model such that the pedestrian collides with the AV ρ. This is achieved by performing roll outs sampling initial states x_0 for the pedestrian from μ and by sampling pedestrian and car moves from π and ρ respectively. The roll-outs are evaluated by a reward function r_μ that rewards collisions between the AV with position y_t and the pedestrian x_t, where $t \in [0, T]$ is the timestep. The model μ observes the current scene (S, D, O, P) of size $w \times h$ containing the AV ρ at position y_0 and outputs a distribution of size $w \times h$ over the input scene. The pedestrian's initial position $x_0 \sim P\mu(S, D, O, P)$ is considered to be the action taken by the policy gradient agent μ. Since the initializer μ cannot control the actions of the behaviour policy π further the model μ does not receive a new state in response to the chosen action, only a reward r_μ.

To avoid sampling from locations that could not possibly lead to a collision we introduce a prior for the μ. Given that the pedestrian has a maximal speed of $3m^{-1}$ there exists a cone of points h from which the pedestrian can reach the car's constant velocity trajectory. The prior for the points in h is $\|x - y_0\|^{-1}$ where y_0 is the initial position of the car. The prior $P(x)$ is 0 within the braking distance $v_\mu^2/(2g * 0.8)$ of the car assuming dry road conditions (0.8 as friction coefficient). This is to avoid sampling from the trivial initializations within the car's braking distance, thus leading to an inevitable collision. The points x that are on the constant velocity estimate of the car's trajectory receive a 0 prior. Finally for all other points the prior is $\frac{1}{||x - y_0||}^2$. The prior encourages the pedestrian agent to be initialized close to the car, to encourage collisions. Without a prior a large number of samples of x_0 may be needed to encounter a collision. For a sample prior see Fig. 2.

Note that in order for a collisions to occur when the pedestrian behaviour model π is collision avoiding, the pedestrian should be initialized close enough to the AV such that collision avoidance will not be trivial. Alternatively the pedestrian's observation of the car should be noisy enough for the pedestrian to miss the exact location of the car or to incorrectly internally predict the car's future motion. Lastly additional spatial constrains such as street width, external obstacles and traffic agents can be used to enforce a smaller movement space for the pedestrian, leaving the pedestrian fewer options to avoid a collision with the car. Finally to allow the initializer to control the pedestrian enough to ensure collisions we utilize a goal g^π driven pedestrian behaviour model. In the presented experiments the pedestrian's goal is set to be the reflection of the pedestrian's position in the car's trajectory. This goal location enforces the pedestrian and car trajectories to cross. In the supplementary material we present also a model for learning the goal positions g^π of the pedestrian. Further temporal or

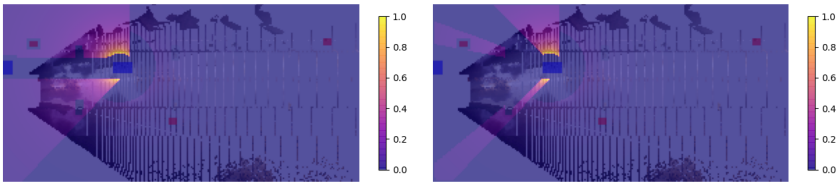

Figure 2: *Left:* A top view image of a sample prior P of μ. In red are other pedestrians, and in blue are cars. Note how the prior implies a higher likelihood of pedestrian initial placement which are close to the AV (as these correspond to challenging scenarios). *Right:* The same prior after a multiplication with the occlusion map O.

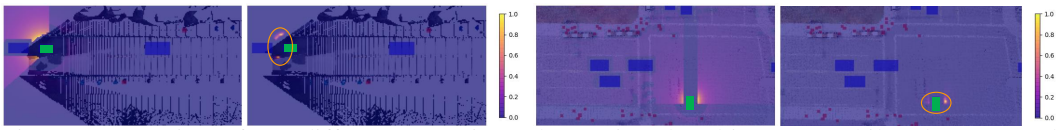

Figure 3: Top-view of two different scenarios. The AV is colored in green, while other cars and pedestrians are colored in blue and dark red, respectively. In each of the two examples, the left part shows the prior distribution and the right part shows the final initial distribution. *Left example:* The prior induces a high likelihood for initializing a pedestrian close to the AV (as is expected, since that corresponds to a challenging scenario), but the probability map is very smeared out. Comparing with the final initial distribution, the probability map is much less scattered and more peaked close to the AV (indicated also with an external orange ellipsoid, to more clearly show where the probability mass is). *Right example:* We see a similar phenomenon as in the left example, in the dense dataset.

behavioural constraints could be added to the pedestrian behaviour model to allow μ more control of the pedestrian behaviour after initialization.

Pedestrians can be initialized in occluded spaces to ensure that the generated pedestrians are visually plausible when added to existing data. The prior P can be used to enforce pedestrian initialization in occluded spaces. This can be done by replacing the prior P with OP, where O is the occlusion map from the AV's perspective. In the experiments it is assumed that the AV observes front-camera images with a 90-degree view. If the AV instead has no direct blind spots then the occlusion map needs to be blurred before adding to the prior to encure non-zero gradients when training μ. The pedestrian distribution model μ observes the scene as $s^\mu = (S, D, PO)$. Here S contains the top view RGB image and semantic labels of the scene (possibly constructed from a reconstruction). The same semantic labels are used as in the pedestrian model §3.4. The dynamic mapping D contains the constant velocity predictions of the external cars and pedestrians. It is the reciprocal of the dynamic map used in the pedestrian behaviour model §3.4, and contains a separate channel for cars and pedestrians. Further details are given in the supplementary. Finally μ observes the product of the prior and the occlusion map PO. The prior indicates to μ which car should be challenged by the pedestrian initialization.

We use a small architecture for μ, see the supplement for details. It consists of two convolutional layers, where the output of each convolutional layer is interpolated back to the original image size and added to each-other. The sum of the up-scaled convolutional layers is passed through softmax giving rise to the output. An entropy term is subtracted from the loss J^μ to encourage enough exploration of the scene space before convergence.

The reward function of μ is defined as the cumulative sum of rewards $r^\mu(x_0) = \sum_t \gamma^t r_\mu(x_t, a_t^\pi, y_t, a_t^\rho)$ attained at each timestep t in the episode. The reward r_μ encourages initial positions x_0 that lead the pedestrian to pedestrian-like goal reaching motion while avoiding collisions with all external objects, cars and pedestrians but encouraging collisions with the AV. The reward $r_\mu(x_0)$ is a slightly adapted multiplicative reward consisting of the reward components used in the pedestrian behaviour model §3.4 and an additional term R_ρ that is an indicator function that is 1 when π and ρ collide.

3.3 Autonomous Car Model

The car model ρ is intentionally simple. The focus of this work is to show that a collision avoiding pedestrian behaviour can be successfully used in autonomous car test case generation given enough constraints on the problem. The car is a policy gradient model with three inputs; $|x_t - y_t|$ distance to the pedestrian agent in the current timestep t, distance to the closest car d_t, and car's intersection with the sidewalk δ_t. The car's speed c_t is sampled from $\mathcal{N}(\text{sigmoid}(w^T s^\rho + b), \sigma_\rho)$, where w, b are learnt weights, and $\sigma_\rho = 0.1$. The sampled speed is then scaled by the maximal speed of 70km/h. The car's initial position y_0 and direction are chosen randomly among the external car's constant velocity future trajectories.

The car ρ is assumed to have a constant direction and the policy gradient model learns to select the speed of the car. In a number of cases the speed control can be enough to avoid collisions. Allowing direction changes complicates the learning objective of the AV somewhat as the car receives two conflicting objectives: to move to a goal location further ahead and to avoid collisions. There is vast literature on AV's and on balancing such conflicting objectives, and in the future we aim to replace the minimal car model with a state of the art AV model. This change in the car model would also require additional constraints to the learning problem in the form of realistic noise modelling on pedestrian detection. Without introducing additional constraints a state of the art car model may outperform the pedestrian initial distribution model μ in early training leading to 0 gradients.

The reward r_ρ function provides a negative reward for collisions with cars, people and static objects. The car also receives a negative reward proportional to fraction of overlap with the sidewalk. To motivate the car to move, a positive reward is given at the end of the episode for travelled distance $\|y_0 - y_T\|$. The exact reward structure is provided in the supplementary.

3.4 Pedestrian Model

The collision avoiding pedestrian behaviour policy π is the CARLA SPL-goal model [31]. The π is a policy gradient agent that is trained by alternatively optimizing π for the maximum likelihood objective for pedestrian trajectory forecasting and for the policy gradient objective for collision avoidance. The reward function r_π of π encourages motion in pedestrian dense areas R_{ped} and penalizes collisions with cars (including ρ), other pedestrians and static objects in the reward terms R_{coll}. The reward component R_g encourages movement towards the goal location g^π, and R_ϕ penalizes unnaturally large motions. The reward function is further detailed in the supplementary.

The model observes a local crop $S_t(x_t)$ of size $5m \times 5m$ of the semantic labels and RGB top view image of scene S and a local crop $D_t(x_t)$ of the dynamic occupancy map D_t. Further the state s^π contains a history of past movements and poses taken by the pedestrian in the past $N = 12$ timesteps, the displacement to the closest car and the displacement to the goal g^π. The policy gradient model takes a step a_t^π that is described by a discretized direction and a speed, that is then articulated by the Human Locomotion Network. In the majority of the experiments the pedestrian models weights are kept constant to not deviate from the learnt pedestrian motion.

It should be noted that the CARLA SPL-goal model is trained to avoid collisions on a dataset where the cars have much lower average speed than what can be exhibited by ρ. This provides sufficient constraints on the collision avoidance abilities of π to ensure that the initial pedestrian distribution can learn in early stages. The pedestrian model's incorrect internal estimate of the car's motion acts as a noise on the pedestrian's prediction of the car's motion, and could be instead be replaced by a noise model. Alternatively with a more detailed pedestrian model the test case generator could be given more control to induce collisions by controlling π's latent behaviour and thus affecting the pedestrian's distractedness, urgency or possibly mood.

4 Experiments

We experiment on a dataset gathered from CARLA[37]. Training data is collected from Town 1 and consists of 100 training and 50 validation scenes. The test set consists of 37 scenes from Town 2. For each scene a 3d reconstruction of RGB and semantic segmentation is created from a car's perspective. We also introduce a smaller more object-dense dataset consisting of 4 different simulations of 5 scenes, these scenes are gathered from a drone's perspective. The latter dataset is only used for training and will referred to as the denser CARLA dataset. In all experiments the scenes $51m \times 25.6m$

Table 1: *Left:* The proposed pedestrian initial distribution model π generates more collisions with the base car model than the baselines; the priors P, PO and random initialization from occluded spaces. *Right:* An ablation studying the effect of the prior and dataset density during the training of μ shows that the model appears to be robust to changes in prior during training. The dataset leads to fewer collisions but increases the variability of the generated data.

	Random O	Prior P	Prior PO	$PO\mu$
#. collisions	0.06	0.10	0.09	**0.21**
Avg distance	8.1	8.0	8.0	**6.9**
π-entropy	0.83	0.62	0.64	0.22

Training prior:	P	OP	P-D
#. collisions	**0.21**	**0.21**	0.17
Avg distance	**6.9**	7.1	7.4
π-entropy	0.22	0.24	0.28

are voxelized into $20cm \times 20cm \times 20cm$ voxels. The Adam optimizer with a learning rate of 0.005 for μ and a learning rate of 0.03 for ρ is chosen. The weights of μ are initialized randomly, and the weights of ρ are initialized to 1. All of the pedestrian location distribution models μ showed highest validation performance after 2 epochs. During initial experimentation it was noted that the car model ρ had trouble learning without the presence of an initializer. Therefore all pedestrian initial distribution models μ are evaluated with a *base car model* that is trained on two scenes of the dataset for 200 epochs with an initializer that is trained on the training dataset for 10 epochs. A discount rate $\gamma = 0.99$ and a trajectory length of $T = 30$ are used to train the car model, and a trajectory length of $T = 100$ is used to train the pedestrian initial distribution models μ. Each scene is evaluated for 10 epochs with a sequence length $T = 100$. During testing the pedestrian π and car models ρ are evaluated by performing the mode and mean actions respectively. The following metrics are reported,

- #. collisions - number of collisions the car model ρ has with pedestrians on average.
- Avg distance - average distance travelled by the car during an episode.
- π-entropy - entropy of the pedestrian policy during the length of an episode.

In Table 1 *left* the proposed pedestrian initial distribution model $PO\mu$ from the §3.2 is shown to outperform the prior distribution P, the occlusion-enforced prior distribution PO, and *Random O* the random initialization of pedestrians from occluded spaces assuming a 360 field of view. The proposed $PO\mu$ has the lowest π-entropy illustrating that the pedestrian initializing distribution $PO\mu$ has learnt to initialize the pedestrian such that the pedestrian's direction of movement is as predictable as possible to ensure collisions. To the *left* in Fig. 4 the prior P and the corresponding scene's $P\mu$ distribution are visually compared. The $P\mu$ has learnt decisively to initialize the pedestrian near the car, and with a higher probability towards the sidewalk than the road.

In Table 1 *right* the effect of the prior during training is studied. A model trained with a prior P without an occlusion map $P\mu$ is then tested with the prior PO, and it performs similarly to the model $PO\mu$ for which the learning and testing priors are PO. This suggests that the model is relatively robust to the choice of prior. For a visual comparison of P and PO see Fig. 2. The last column in Table 1 *right* shows the number of collisions produced by a model P-D that is trained on the smaller but more obstacle dense dataset, but tested on the larger CARLA dataset's test set. The π-entropy suggests that P-D appears to lead to a slightly more varied pedestrian motion, and therefore a lower collision rate as the initialization has less control over the π roll out. The distribution of $P\mu$-D is visualized in Fig. 4 to the right on a scene from the denser dataset.

The effects of the pedestrian behaviour model π on the pedestrian initial distribution model μ can be seen in Table 2. The tested pedestrian behaviour policies are the following,

- *SPL-goal* - the goal reaching collision avoiding pedestrian model described in §3.4
- *SPL+ϵ* - a distracted SPL-goal pedestrian. With a 0.3 probability at each timestep the pedestrian will not notice the car for $m \sim Po(2)$ timesteps.
- *Adv. SPL* - an adversarial SPL-goal agent. A *SPL-goal* that is trained further simultaneously with μ with the R^{μ} reward.
- *CV* - constant velocity motion articulated by [32]. The movement direction is towards the goal and the speed is drawn from a Gaussian with $\mu = 1.23$ and $\sigma = 0.3$ (same as [38]).

The number of collisions of the pedestrian initial distribution models are similar for the collision avoiding *SPL-goal*, the distracted *SPL+ϵ* and the adversarial pedestrian policy *Adv. SPL*. The pedestrian initializer μ does not learn to utilize the distractedness of *SPL+ϵ*, as the added noise is

Table 2: *Left:* Changing the collision avoiding pedestrian behaviour model *SPL-goal* for a distracted pedestrian *SPL*+ϵ or an adversarial pedestrian *Adv. SPL* does not change the number of collisions. A constant velocity pedestrian leads to a lower number of collisions and entropy.*Right:* Training the π and μ simultaneously *Sim* results in metrics similar to those of separately trained models. This is confirmed by testing the *Alt-*μ,*Sim-*μ against the *baseline car*, and the *Alt-*ρ,*Sim-*ρ against $PO\mu$

	SPL	SPL+ϵ	A. SPL	CV	Alt-μ, ρ	Sim-μ, ρ	Alt-μ	Sim-μ	Alt-ρ	Sim-ρ
#. collisions	**0.21**	**0.21**	0.20	0.12	0.20	0.19	**0.20**	0.18	0.40	**0.22**
Avg distance	**6.9**	7.1	7.1	8.7	7.1	7.1	**7.2**	7.3	5.1	**6.9**
π-entropy	**0.22**	0.21	0.11	0	0.11	**0.24**	**0.24**	0.24	0.18	0.22

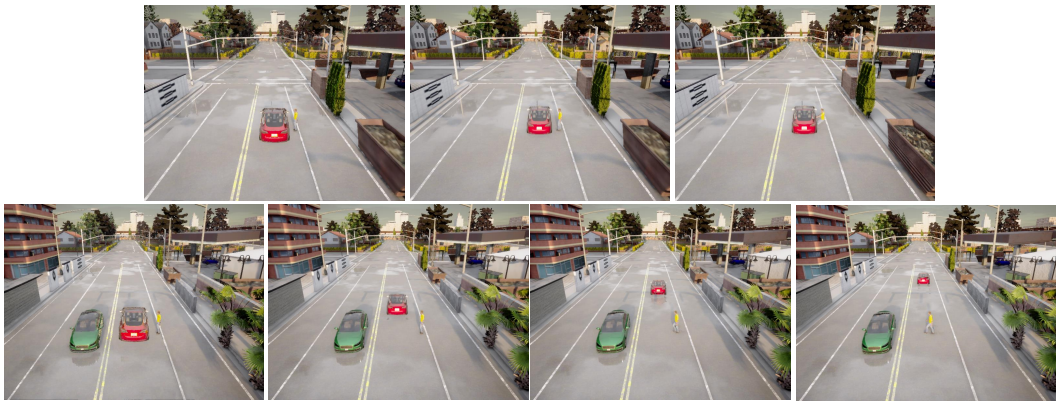

Figure 4: Sample trajectories of the *Sim-*μ, ρ model, sub-sampled at 5 frames from frame 0. *First row:* The car changes speed thus causing the pedestrian to incorrectly estimate the car's motion and walk into the car. *Second row:*The pedestrian successfully waits for the car to pass before crossing the road

unstructured, and thus unpredictable. Natural noise in the pedestrian's observation of the car could be expected to be more structured, for example high noise is expected near occluded spaces. The entropy of the adversarial *Adv. SPL* pedestrian halves compared to the *SPL-goal* suggesting that the produced trajectories are less varied, but this does not affect the collision rate. The *CV* model results in fewer collisions than the *SPL-goal*, it is possible that the car model has learnt to avoid linear motion pedestrians during its training. Since the car and the pedestrian initial distribution model are interdependent it is of interest to train the models in a cyclic manner. The car and the initial pedestrian model are trained simultaneously, *Sim*, for 14 epochs. The resulting μ and ρ and are evaluated against one another in Table 2 *Right* in column *Sim-*μ, ρ, giving rise to collision rates similar to the base model $PO\mu$ in Table 1. The pedestrian initial distribution model of *Sim-*μ, ρ is evaluated against the *base car model* in column *Sim-*μ, and the car model of *Sim-*μ, ρ is evaluated against the base model $PO\mu$ in column *Sim-*ρ. This is compared to alternative training *Alt-*μ, ρ where the car is trained for 2 epochs followed by training μ for 1 epoch for a total of 14 epochs. It can be noted that when testing the *Alt-*μ, ρ car against the *Alt-*μ, ρ pedestrian initial distribution results in low π-entropy even compared to *Alt-*μ suggesting that the pedestrian initial distribution model of *Alt-*μ, ρ optimizes the the initialization for the *Alt-*ρ car model. Unfortunately the car model of *Alt-*ρ collides often when evaluated with the $PO\mu$ from Table 1. The simultaneously trained μ and ρ are comparable in collisions and entropy to the $PO\mu$ and the *baseline car* model. Sample trajectories if the *Sim-*μ, ρ can be seen in Fig. 4. Further visualizations are provided in the supplement.

5 Conclusion

The problem of adversarial pedestrian test case generation can advantageously be decomposed into the pedestrian initial distribution mdoel learning and the pedestrian behaviour model learning. This problem decomposition allows for collision avoiding pedestrian models to be used to evaluate AVs. Collision avoiding pedestrian models mimic cooperative pedestrians and the experiments have shown that a collision inducing pedestrian initial distribution is learnable even for coo-operative pedestrian behaviours. Further the problem decomposition also allows for a number of pedestrian behaviour models to be used in the testing of an AV.

References

[1] Z. Zhu and H. Zhao. A survey of deep rl and il for autonomous driving policy learning. *arXiv preprint arXiv:2101.01993*, 2021.

[2] P. Kohli and A. Chadha. Enabling pedestrian safety using computer vision techniques: A case study of the 2018 uber inc. self-driving car crash. *CoRR*, abs/1805.11815, 2018. URL http://arxiv.org/abs/1805.11815.

[3] M.-F. Chang, J. Lambert, P. Sangkloy, J. Singh, S. Bak, A. Hartnett, D. Wang, P. Carr, S. Lucey, D. Ramanan, et al. Argoverse: 3d tracking and forecasting with rich maps. In *CVPR*, 2019.

[4] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun. Vision meets robotics: The KITTI dataset. *I. J. Robotics Res.*, 32(11):1231–1237, 2013. doi:10.1177/0278364913491297. URL https://doi.org/10.1177/0278364913491297.

[5] P. Sun, H. Kretzschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. Caine, V. Vasudevan, W. Han, J. Ngiam, H. Zhao, A. Timofeev, S. Ettinger, M. Krivokon, A. Gao, A. Joshi, Y. Zhang, J. Shlens, Z. Chen, and D. Anguelov. Scalability in perception for autonomous driving: Waymo open dataset, 2019.

[6] B. Zhou, H. Zhao, X. Puig, T. Xiao, S. Fidler, A. Barriuso, and A. Torralba. Semantic understanding of scenes through the ade20k dataset. *IJCV*, 2018.

[7] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom. nuscenes: A multimodal dataset for autonomous driving. In *CVPR*, 2020.

[8] J. Behley, M. Garbade, A. Milioto, J. Quenzel, S. Behnke, C. Stachniss, and J. Gall. Semantickitti: A dataset for semantic scene understanding of lidar sequences. In *ICCV*, 2019.

[9] X. Huang, X. Cheng, Q. Geng, B. Cao, D. Zhou, P. Wang, Y. Lin, and R. Yang. The apolloscape dataset for autonomous driving. In *CVPR Workshops*, 2018.

[10] R. Kesten, M. Usman, J. Houston, T. Pandya, K. Nadhamuni, A. Ferreira, M. Yuan, B. Low, A. Jain, P. Ondruska, et al. Lyft level 5 av dataset 2019. urlhttps. *level5. lyft. com/dataset*, 2:5, 2019.

[11] Y. Abeysirigoonawardena, F. Shkurti, and G. Dudek. Generating adversarial driving scenarios in high-fidelity simulators. In *International Conference on Robotics and Automation, ICRA 2019, Montreal, QC, Canada, May 20-24, 2019*, pages 8271–8277. IEEE, 2019. doi:10.1109/ICRA.2019.8793740. URL https://doi.org/10.1109/ICRA.2019.8793740.

[12] A. Hamdi, M. Mueller, and B. Ghanem. SADA: semantic adversarial diagnostic attacks for autonomous applications. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 10901–10908. AAAI Press, 2020. URL https://aaai.org/ojs/index.php/AAAI/article/view/6722.

[13] P. Gupta, D. Coleman, and J. Siegel. Towards safer self-driving through great pain (physically adversarial intelligent networks). *ArXiv*, abs/2003.10662, 2020.

[14] J. Wang, A. Pun, J. Tu, S. Manivasagam, A. Sadat, S. Casas, M. Ren, and R. Urtasun. Advsim: Generating safety-critical scenarios for self-driving vehicles. *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[15] D. Karunakaran, S. Worrall, and E. Nebot. Efficient falsification approach for autonomous vehicle validation using a parameter optimisation technique based on reinforcement learning, 2020.

[16] D. Karunakaran, S. Worrall, and E. M. Nebot. Efficient statistical validation with edge cases to evaluate highly automated vehicles. In *23rd IEEE International Conference on Intelligent Transportation Systems, ITSC 2020, Rhodes, Greece, September 20-23, 2020*, pages 1–8. IEEE, 2020. doi:10.1109/ITSC45102.2020.9294590. URL https://doi.org/10.1109/ITSC45102.2020.9294590.

[17] W. Ding, B. Chen, M. Xu, and D. Zhao. Learning to collide: An adaptive safety-critical scenarios generating method. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2020, Las Vegas, NV, USA, October 24, 2020 - January 24, 2021*, pages 2243–2250. IEEE, 2020. doi:10.1109/IROS45743.2020.9340696. URL https://doi.org/10.1109/IROS45743.2020.9340696.

[18] N. Tsoi, M. Hussein, J. Espinoza, X. Ruiz, and M. Vázquez. SEAN: social environment for autonomous navigation. In M. Obaid, O. Mubin, Y. Nagai, H. Osawa, Y. Abdelrahman, and M. Fjeld, editors, *HAI '20: 8th International Conference on Human-Agent Interaction, Virtual Event, Australia, November 10-13, 2020*, pages 281–283. ACM, 2020. doi:10.1145/3406499.3418760. URL https://doi.org/10.1145/3406499.3418760.

[19] C. B. Schwab Benedikt and T. Kolbe. Spatio-semantic road space modeling for vehicle–pedestrian simulation to test automated driving systems. *Sustainability*, 3799, 2020.

[20] W. Ding, B. Chen, B. Li, K. J. Eun, and D. Zhao. Multimodal safety-critical scenarios generation for decision-making algorithms evaluation. *IEEE Robotics and Automation Letters*, 6(2):1551–1558, 2021. doi:10.1109/LRA.2021.3058873.

[21] P. Parashar, A. Cosgun, A. Nakhaei, and K. Fujimura. Modeling preemptive behaviors for uncommon hazardous situations from demonstrations. *CoRR*, abs/1806.00143, 2018. URL http://arxiv.org/abs/1806.00143.

[22] M. Wen, J. Park, and K. Cho. A scenario generation pipeline for autonomous vehicle simulators. *Human-centric Computing and Information Sciences*, 10, 12 2020. doi:10.1186/s13673-020-00231-z.

[23] A. Demetriou, H. Alfsvåg, S. Rahrovani, and M. Chehreghani. A deep learning framework for generation and analysis of driving scenario trajectories. *ArXiv*, abs/2007.14524, 2020.

[24] D. Nishiyama, M. Y. Castro, S. Maruyama, S. Shiroshita, K. Hamzaoui, Y. Ouyang, G. Rosman, J. A. DeCastro, K. Lee, and A. Gaidon. Discovering avoidable planner failures of autonomous vehicles using counterfactual analysis in behaviorally diverse simulation. In *23rd IEEE International Conference on Intelligent Transportation Systems, ITSC 2020, Rhodes, Greece, September 20-23, 2020*, pages 1–8. IEEE, 2020. doi:10.1109/ITSC45102.2020.9294715. URL https://doi.org/10.1109/ITSC45102.2020.9294715.

[25] B. Schütt, T. Braun, S. Otten, and E. Sax. Sceml: A graphical modeling framework for scenario-based testing of autonomous vehicles. In *Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems*, MODELS '20, page 114–120, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450370196. doi:10.1145/3365438.3410933. URL https://doi.org/10.1145/3365438.3410933.

[26] W. Ding, M. Xu, and D. Zhao. CMTS: A conditional multiple trajectory synthesizer for generating safety-critical driving scenarios. In *2020 IEEE International Conference on Robotics and Automation, ICRA 2020, Paris, France, May 31 - August 31, 2020*, pages 4314–4321. IEEE, 2020. doi:10.1109/ICRA40945.2020.9197145. URL https://doi.org/10.1109/ICRA40945.2020.9197145.

[27] Z. W. Sun X., Zhang Y. Building narrative scenarios for human-autonomous vehicle interaction research in simulators. *Advances in Simulation and Digital Human Modeling*, 1206, 2020.

[28] M. Koren and M. J. Kochenderfer. Efficient autonomy validation in simulation with adaptive stress testing. In *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019, Auckland, New Zealand, October 27-30, 2019*, pages 4178–4183. IEEE, 2019. doi:10.1109/ITSC.2019.8917403. URL https://doi.org/10.1109/ITSC.2019.8917403.

[29] D. Lee, S. Liu, J. Gu, M. Liu, M. Yang, and J. Kautz. Context-aware synthesis and placement of object instances. In S. Bengio, H. M. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, pages 10414–10424, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/c6969ae30d99f73951cb976b88a457af-Abstract.html.

[30] Q. Dai, X. Xu, W. Guo, S. Huang, and D. Filev. Towards a systematic computational framework for modeling multi-agent decision-making at micro level for smart vehicles in a smart world, 2020.

[31] M. Priisalu, C. Paduraru, A. Pirinen, and C. Sminchisescu. Semantic synthesis of pedestrian locomotion. In *Proceedings of the Asian Conference on Computer Vision (ACCV)*, November 2020.

[32] D. Holden, T. Komura, and J. Saito. Phase-functioned neural networks for character control. *ACM Trans. Graph.*, 36(4):42:1–42:13, 2017. doi:10.1145/3072959.3073663. URL https://doi.org/10.1145/3072959.3073663.

[33] O. Makansi, Ö. Çiçek, K. Buchicchio, and T. Brox. Multimodal future localization and emergence prediction for objects in egocentric view with a reachability prior. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, pages 4353–4362. IEEE, 2020. doi:10.1109/CVPR42600.2020.00441. URL https://doi.org/10.1109/CVPR42600.2020.00441.

[34] S. Manivasagam, S. Wang, K. Wong, W. Zeng, M. Sazanovich, S. Tan, B. Yang, W. Ma, and R. Urtasun. Lidarsim: Realistic lidar simulation by leveraging the real world. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, pages 11164–11173. IEEE, 2020. doi:10.1109/CVPR42600.2020.01118. URL https://doi.org/10.1109/CVPR42600.2020.01118.

[35] J. Sun, H. Averbuch-Elor, Q. Wang, and N. Snavely. Hidden footprints: Learning contextual walkability from 3d human trails. In A. Vedaldi, H. Bischof, T. Brox, and J. Frahm, editors, *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part XVIII*, volume 12363 of *Lecture Notes in Computer Science*, pages 192–207. Springer, 2020. doi:10.1007/978-3-030-58523-5_12. URL https://doi.org/10.1007/978-3-030-58523-5_12.

[36] R. J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 1992.

[37] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun. Carla: An open urban driving simulator. In *CoRL*, 2017.

[38] S. Chandra and A. K. Bharti. Speed distribution curves for pedestrians during walking and crossing. *Procedia-Social and Behavioral Sciences*, 104:660–667, 2013.