Appendix

A Dataset

A.1 Overview

We will publish the DrivingDojo dataset, data format and annotation instructions, AIF benchmark, and code for the baseline method on our project page: https://drivingdojo.github.io.

Terms of use and License. Our dataset is released under the **CC BY-NC 4.0** license, allowing everyone to use it for non-commercial research purposes.

Data maintenance. The data is stored on Google Drive for global accessibility, and we will supply various links (e.g., Hugging Face) for researchers' convenience. We will maintain the data long-term and periodically verify its accessibility.

In the following, we showcase more video examples in our DrivingDojo dataset, the corresponding videos are better illustrated on our project page.

Figure 9: **Examples of rich ego-actions on the DrivingDojo dataset.**

Action completeness. We include more dataset visualizations depicting various ego-actions in Figure 9. From top to bottom, the images show the ego vehicle performing left turns, right turns, going straight, lane-changing, and making emergency brakes during the driving.

Figure 10: **Examples of multi-agent interplay on the DrivingDojo dataset.**

Multi-agent interplay. Interaction plays a crucial role in driving scenarios. It usually means that the ego vehicle has engaged with other road users, leading to changes in the behavior of either the ego vehicle or the other road users. As shown in Figure 10, we present a series of interaction examples in our dataset. In the first scenario, the car suddenly encounters another vehicle crossing its path while moving forward, prompting an abrupt braking maneuver. The second scenario portrays the car encountering an electric scooter unexpectedly crossing its path. Illustrating the third scenario, the car comes across a vehicle in front opening its door, forcing an abrupt brake. In the fourth scenario, the challenge involves encountering a bicycle approaching from the opposite direction, while the fifth scenario involves navigating around a stroller. The sixth scenario showcases encountering road construction ahead, followed by encountering a street sweeper in the seventh scenario. The eighth scenario presents a situation where a car suddenly makes a U-turn from the opposite direction,

prompting an urgent braking response from our vehicle. Subsequent scenarios involve interactions with pedestrians. These diverse interaction scenarios provide a crucial foundation for studying the interaction of real-world simulators.

Figure 11: **Examples of diverse open-world objects on the DrivingDojo dataset.**

Open-world knowledge. In complex driving environments, we often encounter a wide variety of open-world situations. These scenarios can include sudden appearances of unexpected obstacles such as fallen trees, construction barriers, or abandoned vehicles. Typically belonging to the tail end of a long-tail distribution, these scenarios are rare yet crucial for ensuring safe driving. In Figure 11, we showcase a series of examples from the dataset, which fully demonstrate the richness of our dataset in capturing long-tail scenarios. From top to bottom, the examples illustrate encounters with a crane, a towing rope, construction barriers, a fallen roadblock, a vehicle transporting iron pipes, a vehicle transporting tree branches, a herd of sheep, an excavator, a bonfire, and power lines.

A.2 Curation

In this section, we provide the details of the curation procedure of each subset of DrivingDojo dataset and the descriptions of curated actions and interactions. This section supplements the details for Section 3.4 in the main paper.

Action Completeness Ego maneuvers for a car, particularly in the context of autonomous driving, refer to the actions and decisions the vehicle makes to navigate its environment safely and efficiently. Here is an exhaustive list of common ego maneuvers, and some examples in our datasets are shown in Figure 1a in the main paper:

- **Acceleration**: Increasing speed to match traffic flow.
- **Deceleration**: Gradual slowing down for stop signs, traffic lights, or traffic congestion.
- **Lane Keeping**: Maintaining the current lane.
- **Lane Changing**: Changing lanes to overtake slower vehicles or merge into traffic.
- **Turning**: Left/right or U-turns at intersections or roundabouts.
- **Stop and Move on**: Stopping/proceeding at traffic signals or stop signs.
- **Emergency brake**: Abrupt and sudden braking maneuver to avoid a collision or mitigate the impact of a potential hazard.

So, in the DrivingDojo-Action set, the videos follow different action commands, and the actions are mainly from the planning and control (PNC) signals, such as left and right turns, straight crossings, and lane changes. Each curated video clip begins with the PNC issuing a specific command and ends when the command is completed.

Multi-agent Interplay The examples of multi-agent interplay are shown in Figure 1b in the main paper. Then we describe the detailed cases of the interactions with dynamic agents.

- **Cutting in/off**: Another vehicle abruptly changes lanes and enters the path of the autonomous vehicle. Ego vehicle changes lanes and enters the path of the other vehicles.
- **Meeting**: Ego vehicle encounters other vehicles traveling in the opposite direction.
- **Blocked**: Ego vehicle is stopped by other agents, such as vehicles, motorcycles, and pedestrians.
- **Overtaking and being overtaken**: Ego vehicle attempting to pass another vehicle and being passed by another vehicle.

In the DrivingDojo-Interplay set, the core data curation strategy is to find the interaction with other agents. The interaction is determined using PNC signals and manually defined rules. The main interaction videos are from PNC dangerous interaction data. PNC conducts a deduction between the ego vehicle and obstacles. When the ego vehicle cannot avoid collision by turning the steering wheel or slowing down slightly, it is a PNC interaction case.

Open-world Knowledge Here, we select some representative and interesting examples from these rare cases and show them in Figure 1c in the main paper. Based on the provided image and the given descriptions, here are the detailed descriptions of each rare case in autonomous driving:

(a) A worker's helmet rolls on the sidewalk next to the road. (b) A soccer ball is seen flying across the road. (c) A water bucket is depicted falling onto the road. (d) Parcel boxes have fallen onto the road. (e) A dog is crossing the road. (f) A rope is floating over the road. (g) The traffic light turns red. (h) A boom barrier blocks the vehicle from moving forward.

As mentioned above, we curated DrivingDojo-Open set in which the videos are more carefully categorized and labeled with text descriptions. The sources are unusual weather, foreign objects on the road surface, floating obstacles, falling objects, taking over cases, and interactions with traffic lights and boom barriers. For curating the foreign objects/obstacles, we manually check and label them by a large number of data annotators.

Dataset Format DrivingDojo dataset provides a file named 'dataset_info.json' that stores information corresponding to each video segment, including the information shown in Table 6. The 'type' represents the major category, 'tag' represents the minor category, and 'remark' provides detailed descriptions of the reasons for hard braking and intervention.

Table 6: The explanation of the information in dataset_info.json.

Information	Detailed explanation
meta_info	weather, location, time, frame number
description	type, tag, remark
videos	the image path for each frame
camera_info	the camera intrinsic parameters and extrinsic matrix for each frame
action_info	the coordinates of the next frame's camera position in the current camera coordinate system

The following is an example directory structure for a dataset:

```
.
dataset_info.json
action_info
    062959_s20-370_1712024694.0_1712024714.0
        0023_next_frame_position_at_current_camera.txt
        0025_next_frame_position_at_current_camera.txt
        0027_next_frame_position_at_current_camera.txt
        ...
    145325_s20-190_1683790938.0_1683790958.0
        0024_next_frame_position_at_current_camera.txt
        0026_next_frame_position_at_current_camera.txt
        0028_next_frame_position_at_current_camera.txt
        ...
    ...
camera_info
    062959_s20-370_1712024694.0_1712024714.0
        0023_camera_parameters.txt
        0025_camera_parameters.txt
        0027_camera_parameters.txt
        ...
    145325_s20-190_1683790938.0_1683790958.0
        0024_camera_parameters.txt
        0026_camera_parameters.txt
        0028_camera_parameters.txt
        ...
    ...
videos
    062959_s20-370_1712024694.0_1712024714.0
        0023_CameraFpgaP0H120.jpg
        0025_CameraFpgaP0H120.jpg
        0027_CameraFpgaP0H120.jpg
        ...
    145325_s20-190_1683790938.0_1683790958.0
        0024_CameraFpgaP0H120.jpg
        0026_CameraFpgaP0H120.jpg
        0028_CameraFpgaP0H120.jpg
        ...
    ...
```

Camera info. The 'camera info' refers to the extrinsic and intrinsic matrices of each frame of a fisheye camera. The world coordinate system is chosen as the East-North-Up (ENU) coordinate system. In the camera coordinate system, the x, y, and z axes respectively point to the right, down,

and forward. We normalize the world coordinate system of the first frame to the origin, which means that the translation variables in the extrinsic matrices of each frame are subtracted by the translation variables of the first frame.

Action info. The 'action info' represents the coordinates of the next frame's camera position in the current camera coordinate system. Let the transformation matrix from the camera to the world coordinate system be $\begin{pmatrix} R & T \\ 0^3 & 1 \end{pmatrix}$. The calculation method for the action info A_n of the n-th frame is shown in formula 3. The orientation of xyz axes in matrix A_n is consistent with the camera coordinate system, where the x, y, and z axes respectively point to the right, down, and forward.

$$\begin{pmatrix} A_n \\ 1 \end{pmatrix} = \begin{pmatrix} R_n & T_n \\ 0^3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} T_{n+1} \\ 1 \end{pmatrix} \tag{3}$$

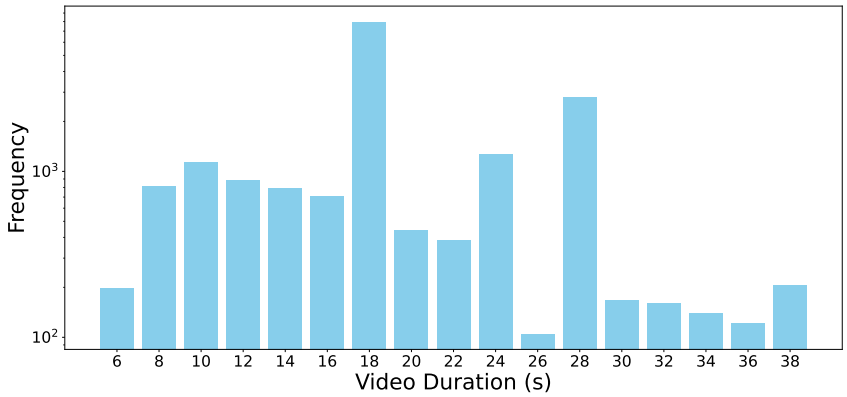

Figure 12: **Distribution of DrivingDojo video duration.**

Video info. The video is stored as a sequence of individual image frames. The distribution of video duration is shown in Figure 12, with the majority of videos lasting around 20 seconds.

B Implementation Details

B.1 Experiment Setup

During the experiment, we employed two settings for training the model. In the first setting, the focus is on visual prediction: the model predicts subsequent video content based solely on the initial frame image. In the second setting, we employ action-controlled video generation. Here, alongside the initial frame image, action information for the subsequent frames is provided to the model, enabling it to predict the ensuing video content.

Visual prediction. In this setup, we trained a high-resolution version of the model, 1024×576 resolution for 14 frames, aimed at better capturing the generation of long-tail objects. Additionally, we developed a low-resolution version of the model, 576×320 resolution for 30 frames, to simulate various vehicle behaviors and interaction events. We fine-tune all parameters of the U-Net model.

Action instruction following. In this setup, we trained model using 576×320 resolution for 30 frames. We fine-tune all U-Net parameters together with a new action encoder.

B.2 Training

We initialize the model using the SVD-XT checkpoint. Following SVD, our model is trained with the EDM framework [32]. During training, we set the fps to 5 and the motion_bucket_id to 127. We

utilize the AdamW optimizer [35] with a learning rate of 1×10^{-5}. The training process is conducted on 16 NVIDIA A100 (80G) GPUs with 32 batch size for 50K iterations. To allow classifier-free guidance [25], we drop out action feature with a ratio of 20%.

B.3 Evaluation

During inference, we generate videos using the DDIM sampler for 25 steps.

Visual Quality. To evaluate the quality of the generated video, we utilize FID (Frechet Inception Distance) [23] and FVD (Frechet Video Distance) [46] as the main metrics. For FID calculation on videos, we randomly select 5,000 frames for evaluation. Additionally, for FVD calculation, we generate 256 videos for evaluation. The results are the average of 10 calculations. We use the official UCF FVD evaluation code[2].

Action instruction following (AIF). For each generated video with action instructions, we estimate the camera poses for each frame in the video, align the scale of the estimated trajectory with the instruction trajectory, and compare the vehicle motion in each frame with the respective action instructions. We estimate the ego trajectories in generated videos using the offline visual structure-from-motion (SfM) implementation COLMAP [41, 42]. We found that moving objects significantly impact the quality of the reconstruction, so we used instance masks to occlude foreground moving objects during the reconstruction process. For videos generated based on initial images from DrivingDojo, we fix the camera intrinsic parameters as the ground truth values for videos from DrivingDojo. For videos generated from initial images with unknown camera intrinsics (e.g. images from OpenDV-2K), we estimate the camera intrinsics together with the camera extrinsics of images. We perform feature point extraction, feature point matching, and sparse scene reconstruction with the official implementation of COLMAP[3] to estimate the poses of cameras. In our experiments, we generate videos in 30 frames and align the scale of estimated trajectories with the instruction trajectories based on the motions in the first $N = 10$ frames. We report the mean value of the absolute error between estimated motions and instruction motions in all video frames.

C Visualizations

In this section, we show the model generation demos trained on the DrivingDojo dataset. As shown in Figure 13, our model can generate high-resolution, complex driving scenarios.

Figure 13: **Examples of high-resolution and complex scenarios generation.** For illustration purposes, we represent each video example with a single frame.

[2] https://github.com/SongweiGe/TATS/

[3] https://github.com/colmap/colmap

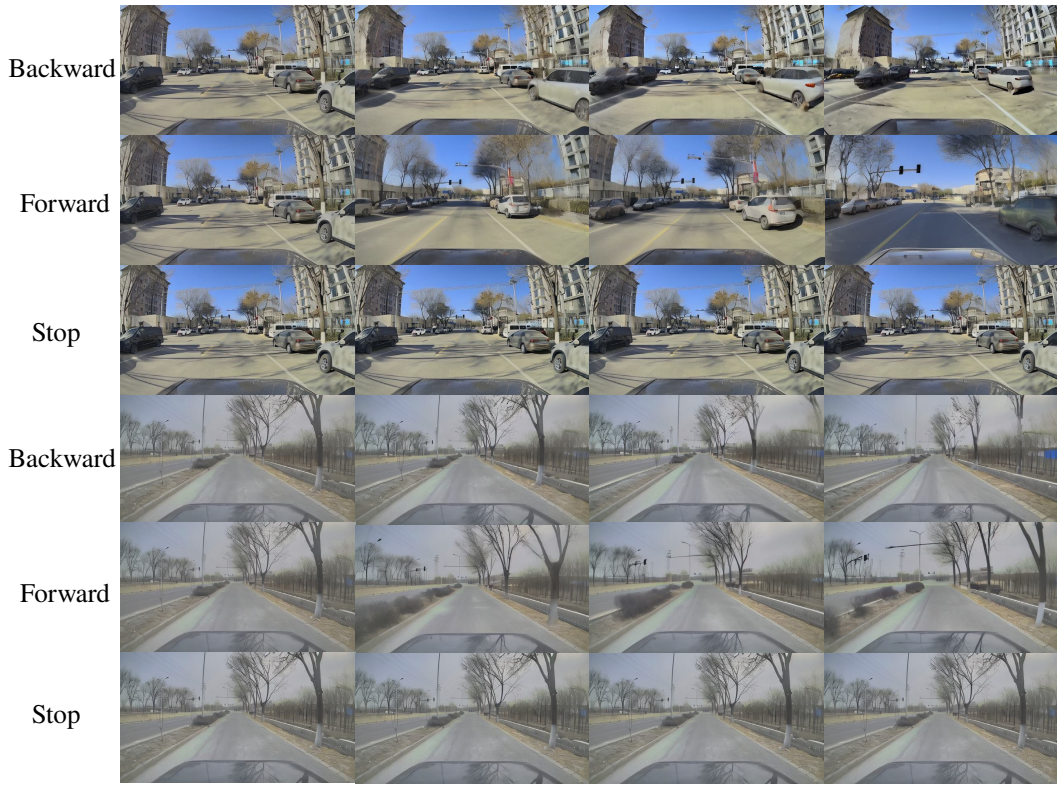

Figure 14: **Examples of diverse action-based video generation.**

C.1 Diverse Actions

As shown in Figure 14, we demonstrate how actions control the generation of different futures, such as moving forward, backward, and stopping.

C.2 Dynamic Interaction

As shown in Figure 15, we observe that choosing different actions can influence the behavior of other vehicles, resulting in different responses from the world model. For instance, in the first example, if we choose to proceed slowly, the vehicle on the left decides to stop and yield. Conversely, if our vehicle stops, the left vehicle perceives an obstruction and slightly reverses to make way. In the second example, when we choose to brake, the right vehicle quickly cuts in front of us, while if we choose to proceed straight, the right vehicle waits in place.

C.3 Open-world Knowledge

As illustrated in Figure 16, we demonstrate the model's ability to simulate various open-world objects, such as encountering construction zones, rare objects like ladders or balloons on the road, and simulating a puddle of water on the ground.

D License of Assets

We report licenses of all artifacts used in this work in this section.

Figure 15: **Simulation of interaction with other agents.**

Figure 16: **Simulation of various open-world objects on the road.**

Model We use the pre-trained stable video diffusion [2] checkpoints from the huggingface platform. These checkpoints are released under the stable video diffusion non-commercial community license agreement[4] for research purpose.

Our Dataset Our dataset is collected and curated by the autonomous driving team of Meituan Inc. The road test and data collection procedures conform to privacy and security requirements of local authorities. The authors have obtained the permission for publicly releasing this dataset from both the management team and the company legal team. All personal identifiable information has been removed by both algorithm and subsequent manual inspection. We release the dataset under the CC BY-NC 4.0 license.

Other Datasets We use other public datasets in this work including nuScenes [6], ONCE [37] and OpenDV-2k [54]. The nuScenes [6] dataset is released under the CC BY-NC-SA 4.0 license with Dataset Terms [5]. The ONCE dataset is also released under the CC BY-NC-SA 4.0 license with

[4] https://huggingface.co/stabilityai/stable-video-diffusion-img2vid-xt/blob/main/LICENSE

[5] https://www.nuscenes.org/terms-of-use

Dataset Terms [6]. The OpenDV-2K dataset is constructed from publicly licensed datasets and youtube videos that the authors claimed to support academic usage licenses.

E Datasheet

E.1 Motivation

- **For what purpose was the dataset created?** Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description.

  We introduce DrivingDojo, the first dataset tailor-made for training interactive world models with complex driving dynamics. Our dataset features video clips with a complete set of driving maneuvers, diverse multi-agent interplay, and rich open-world driving knowledge, laying a stepping stone for future world model development.

- **Who created the dataset (e.g., which team, research group) and on behalf of which entity (e.g., company, institution, organization)?**

  Institute of Automation, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Meituan Inc., and Centre for Artificial Intelligence and Robotics, HKISI_CAS.

- **Who funded the creation of the dataset?**If there is an associated grant, please provide the name of the grantor and the grant name and number.

  This work was supported in part by the National Key R&D Program of China (No. 2022ZD0116500), the National Natural Science Foundation of China (No. U21B2042, No. 62320106010), and in part by the 2035 Innovation Program of CAS, and the InnoHK program.

- **Any other comments?**

  No.

E.2 Composition

- **What do the instances that comprise the dataset represent (e.g., documents, photos, people, countries)?** Are there multiple types of instances (e.g., movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description.

  The instances of our DrivingDojo dataset are videos with ego actions and DrivingDojo-Open subset is also with text descriptions for each scene.

- **How many instances are there in total (of each type, if appropriate)?**

  There are 17.8k videos for the whole DrivingDojo dataset, in which the DrivingDojo-Action subset has 7.9k videos, DrivingDojo-Interplay subset has 6.2k videos, and DrivingDojo-Open has 3.7k videos.

- **Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set?** If the dataset is a sample, then what is the larger set? Is the sample representative of the larger set (e.g., geographic coverage)? If so, please describe how this representativeness was validated/verified. If it is not representative of the larger set, please describe why not (e.g., to cover a more diverse range of instances, because instances were withheld or unavailable).

  The DrivingDojo dataset is sampled from a data pool of around 7500 hours. About representativeness, please refer to the Data Curation section (Sec. 3.4 and Sec. A.2).

- **What data does each instance consist of?** "Raw" data (e.g., unprocessed text or images) or features? In either case, please provide a description.

  DrivingDojo-Action and DrivingDojo-Interplay subsets consist of videos and ego actions, and DrivingDojo-Open subset consists of videos, ego actions, and text descriptions.

- **Is there a label or target associated with each instance?** If so, please provide a description.

  Yes. There is a text description label for each instance in DrivingDojo-Open subset, which describes the open-world knowledge in the scene.

[6] https://once-for-auto-driving.github.io/terms_of_use.html

- **Is any information missing from individual instances?** If so, please provide a description, explaining why this information is missing (e.g., because it was unavailable). This does not include intentionally removed information, but might include, e.g., redacted text.

  No.

- **Are relationships between individual instances made explicit (e.g., users' movie ratings, social network links)?** If so, please describe how these relationships are made explicit.

  No.

- **Are there recommended data splits (e.g., training, development/validation, testing)?** If so, please provide a description of these splits, explaining the rationale behind them.

  No. There is no need for the validation/testing split. We care about zero-shot generation.

- **Are there any errors, sources of noise, or redundancies in the dataset?** If so, please provide a description.

  Yes. The sources of noise may be inaccurate poses, camera noises, and human-sourced text noises.

- **Is the dataset self-contained, or does it link to or otherwise rely on external resources (e.g., websites, tweets, other datasets)?** If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time; b) are there official archival versions of the complete dataset (i.e., including the external resources as they existed at the time the dataset was created); c) are there any restrictions (e.g., licenses, fees) associated with any of the external resources that might apply to a dataset consumer? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate.

  Yes. the DrivingDojo dataset is self-contained.

- **Does the dataset contain data that might be considered confidential (e.g., data that is protected by legal privilege or by doctor–patient confidentiality, data that includes the content of individuals' non-public communications)?** If so, please provide a description.

  No.

- **Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety?** If so, please describe why.

  No.

E.3 Collection Process

- **How was the data associated with each instance acquired?** Was the data directly observable (e.g., raw text, movie ratings), reported by subjects (e.g., survey responses), or indirectly inferred/derived from other data (e.g., part-of-speech tags, model-based guesses for age or language)? If the data was reported by subjects or indirectly inferred/derived from other data, was the data validated/verified? If so, please describe how.

  DrivingDojo dataset is collected using the platform of Meituan's autonomous delivery vehicles.

- **What mechanisms or procedures were used to collect the data (e.g., hardware apparatuses or sensors, manual human curation, software programs, software APIs)?** How were these mechanisms or procedures validated?

  DrivingDojo dataset is collected using the platform of Meituan's autonomous delivery vehicles with fish-eye RGB cameras. The cameras are calibrated. The text labels are manually validated.

- **If the dataset is a sample from a larger set, what was the sampling strategy (e.g., deterministic, probabilistic with specific sampling probabilities)?**

  Please refer to the Data Curation section (Sec. 3.4 and Sec. A.2).

- **Who was involved in the data collection process (e.g., students, crowdworkers, contractors) and how were they compensated (e.g., how much were crowdworkers paid)?**

  The data collectors are employed by Meituan Inc. and are paid by Meituan Inc.

- **Over what timeframe was the data collected?** Does this timeframe match the creation timeframe of the data associated with the instances (e.g., recent crawl of old news articles)? If not, please describe the timeframe in which the data associated with the instances was created.

  The data are collected from May 2022 to May 2024. This timeframe matches the creation timeframe of the data associated with the instances.

- **Were any ethical review processes conducted (e.g., by an institutional review board)?** If so, please provide a description of these review processes, including the outcomes, as well as a link or other access point to any supporting documentation.

  Yes. The ethical review is conducted before the release by Meituan Inc.

E.4 Preprocessing/cleaning/labeling

- **Was any preprocessing/cleaning/labeling of the data done (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)?** If so, please provide a description. If not, you may skip the remaining questions in this section.

  No.

- **Was the "raw" data saved in addition to the preprocessed/cleaned/labeled data (e.g., to support unanticipated future uses)?** If so, please provide a link or other access point to the "raw" data.

  N/A.

- **Is the software that was used to preprocess/clean/label the data available?** If so, please provide a link or other access point.

  N/A.

- **Any other comments?**

  No.

E.5 Uses

- **Has the dataset been used for any tasks already?** If so, please provide a description.

  The DrivingDojo dataset has been used for driving world models. The experiments are in Sec. 5 in the main paper and Sec. C in the appendix.

- **Is there a repository that links to any or all papers or systems that use the dataset?** If so, please provide a link or other access point.

  Yes. Please refer to the webset:https://drivingdojo.github.io.

- **What (other) tasks could the dataset be used for?**

  The DrivingDojo dataset could be used for training end-to-end autonomous driving models.

- **Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labeled that might impact future uses?** For example, is there anything that a dataset consumer might need to know to avoid uses that could result in unfair treatment of individuals or groups (e.g., stereotyping, quality of service issues) or other risks or harms (e.g., legal risks, financial harms)? If so, please provide a description. Is there anything a dataset consumer could do to mitigate these risks or harms?

  No.

- **Are there tasks for which the dataset should not be used?** If so, please provide a description.

  Due to the known biases of the dataset, under no circumstance should any models be put into production using the dataset as is. It is neither safe nor responsible. As it stands, the dataset should be solely used for research purposes in its uncurated state.

- **Any other comments?**

  No.

E.6 Distribution

- **Will the dataset be distributed to third parties outside of the entity (e.g., company, institution, organization) on behalf of which the dataset was created?** If so, please provide a description.
  Yes, the dataset will be open-source.
- **How will the dataset will be distributed (e.g., tarball on website, API, GitHub)?** Does the dataset have a digital object identifier (DOI)?
  On our website: https://drivingdojo.github.io.
- **When will the dataset be distributed?**
  We have released some demos on the project page. The whole DrivingDojo dataset will be public in the camera-ready version.
- **Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)?** If so, please describe this license and/or ToU, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms or ToU, as well as any fees associated with these restrictions.
  DrivingDojo dataset will be distributed under the CC BY-NC 4.0 license.
- **Have any third parties imposed IP-based or other restrictions on the data associated with the instances?** If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms, as well as any fees associated with these restrictions.
  No.
- **Do any export controls or other regulatory restrictions apply to the dataset or to individual instances?** If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any supporting documentation.
  No.
- **Any other comments?**
  No.

E.7 Maintenance

- **Who will be supporting/hosting/maintaining the dataset?**
  Institute of Automation, Chinese Academy of Sciences and Meituan Inc. will maintain DrivingDojo dataset.
- **How can the owner/curator/manager of the dataset be contacted (e.g., email address)?**
  The main maintainer Yuqi Wang's e-mail: wangyuqi2020@ia.ac.cn.
- **Is there an erratum?** If so, please provide a link or other access point.
  There is no erratum for our initial release. Errata will be documented as future releases on the dataset website.
- **Will the dataset be updated (e.g., to correct labeling errors, add new instances, delete instances)?** If so, please describe how often, by whom, and how updates will be communicated to dataset consumers (e.g., mailing list, GitHub)?
  Yes. We will update the DrivingDojo dataset. Especially, we will adapt to end-to-end autonomous driving tasks in the future. The update will be released on the website and GitHub.
- **If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (e.g., were the individuals in question told that their data would be retained for a fixed period of time and then deleted)?** If so, please describe these limits and explain how they will be enforced.
  N/A.
- **Will older versions of the dataset continue to be supported/hosted/maintained?** If so, please describe how. If not, please describe how its obsolescence will be communicated to dataset consumers.
  Yes. We will maintain the older versions of the dataset on the website and GitHub.

- **If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so?** If so, please provide a description. Will these contributions be validated/verified? If so, please describe how. If not, why not? Is there a process for communicating/distributing these contributions to dataset consumers? If so, please provide a description.

  Yes. The dataset is open source under the CC BY-NC 4.0 license. So it is open to other contributors.

- **Any other comments?**

  No.