Drive My Way: Preference Alignment of Vision-Language-Action Model for Personalized Driving

Supplementary Material

A. Additional Quantitative Results

To further validate whether DMW produces behaviors consistent with human expectations, we report the detailed user study results on both long-term preference alignment and short-term adaptation to style instructions.

A.1. Long-term Preference Alignment

Building on the results across scenario types in Table 2, the averaged metrics in Table 5 further show clear differences in driver-specific behavior. Drivers with higher average speeds exhibit consistent behavior across scenarios, whereas more cautious drivers tend to maintain larger headways and lower accelerations. These trends are consistent with the cross-scenario patterns observed in Table 2, suggesting that the policy aligns with persistent behavioral traits beyond individual routes.

A.2. Adaptation on Style Instruction

Beyond long-term preferences, users may also express short-term intentions via style instructions depending on situational context. To complement objective metrics in Table 1, and further evaluate whether the policy adapts to these instructions, we conduct a user study in which human evaluators rate trajectories generated under different style prompts for the same scenario. Each evaluator is shown short video clips rendered under *conservative*, *neutral*, and *aggressive* instructions and asked to judge whether the resulting behavior matches the intended style.

Evaluators score each trajectory on a 0-10 scale according to three criteria: (i) how well the behavior follows the instruction, (ii) efficiency, comfort, and smoothness of drive, and (iii) perceived safety. As shown in Table 6, across four representative scenario types, both StyleDrive [11] and DMW outperform SimLingo baseline [42], indicating the benefit of style-aware driving adaptation. StyleDrive [11] demonstrates improved alignment with short-term style instructions compared to SimLingo, but still falls short of DMW. In contrast, DMW consistently achieves the highest ratings across all styles and scenarios. Evaluators consistently observe that under aggressive instructions, DMW produces higher speeds, shorter following distances, and more decisive accelerations, while conservative instructions yield smoother control profiles and larger safety margins. These results confirm that the proposed policy not only responds to explicit style instructions, but does so in a more sensitive manner.

Table 5. Driving metrics across all scenario types.

Driver	DS	Speed	Effic.	Acce.	Headway	AS	Ratings
D1	93.38	8.54	262.68	6.46	40.38	0.92	8.7
D2	95.43	5.57	163.38	5.15	49.82	0.92	8.3
D3	95.30	6.40	185.83	5.56	47.05	0.83	7.8
D4	91.96	8.77	274.19	6.17	39.70	0.83	8.0
D5	92.58	7.48	229.93	5.58	39.08	0.83	7.9
D6	97.71	6.16	190.74	6.08	42.41	0.75	7.4
D7	96.11	7.16	218.25	6.36	40.12	0.83	7.9
D8	95.28	5.82	188.27	5.54	43.95	0.92	8.2
D9	96.16	5.60	178.97	5.84	45.07	0.67	7.0
D10	94.20	8.38	258.28	6.89	40.26	1.00	8.6

A.3. Additional Qualitative Results

Fig. 7 visualizes how DMW responds to aggressive and conservative instructions across safety-critical scenarios. In the lost-of-control scenario, where the ego-vehicle risks losing control due to the bad road conditions, the aggressive instruction prioritizes efficiency: it maintains a higher speed and attempts to pass the unstable area quickly. In contrast, the conservative instruction leads the agent to maintain a smoother trajectory, and preserve vehicle stability. This divergence highlights how the policy adapts its balance between efficiency and comfort based on the given instruction.

A similar pattern emerges in the oncoming-vehicle intrusion scenario. When another vehicle invades the ego lane, the aggressive instruction causes the agent to accelerate and execute a decisive rightward maneuver at relatively high speed. Meanwhile, the conservative instruction prompts early caution: the agent reduces speed, yields space sooner, and performs a safer avoidance.

In another scenario involving a parked vehicle blocking the lane, the agent must decide when to safely overtake. Under an aggressive instruction, corresponding to personal requirements such as “I’m in a hurry” or “I’m running late”. The policy seeks the earliest viable gap and initiates the overtake quickly to reduce waiting time. In contrast, under a conservative instruction, the agent remains patient, yielding until the oncoming lane is fully clear, especially under low-visibility conditions.

In the scenario where the agent needs to make a left turn at an unsignalized junction, under the aggressive instruction, the agent initiates the turn earlier once it identifies a tighter and feasible opening, minimizing delay. In contrast, the conservative instruction causes the agent to wait patiently for a safer gap before turning. This cautious behavior reflects the emphasis on safety and low-risk in complex intersection negotiations. Together, these examples il-

Table 6. User study ratings (0-10) evaluating how well trajectories match intended instructions. Five evaluators (E1-E5) rate trajectories from SimLingo [42], StyleDrive [11], and **DMW**.

Scenario	Model	Style	E1	E2	E3	E4	E5
Emergency Brake	SimLingo [42]	Conservative	7.4	6.7	6.4	7.3	6.5
		Neutral	7.0	7.3	7.5	7.0	7.1
		Aggressive	6.5	7.6	7.2	6.6	7.4
	StyleDrive [11]	Conservative	8.2	7.5	7.3	8.0	7.2
		Neutral	7.8	8.0	8.2	7.6	7.9
		Aggressive	7.2	8.4	7.9	7.1	8.2
	DMW	Conservative	9.0	8.1	8.0	8.8	7.9
		Neutral	8.4	8.6	9.1	8.1	8.5
		Aggressive	7.8	9.2	8.3	7.6	9.0
Merging	SimLingo [42]	Conservative	7.2	6.4	6.3	7.2	6.3
		Neutral	6.8	7.0	7.3	6.7	6.8
		Aggressive	6.2	7.5	6.9	6.1	7.3
	StyleDrive [11]	Conservative	8.0	7.3	7.2	7.9	7.1
		Neutral	7.6	7.8	8.1	7.4	7.7
		Aggressive	7.0	8.3	7.6	6.9	8.1
	DMW	Conservative	8.9	7.9	7.7	8.7	7.6
		Neutral	8.2	8.4	9.0	8.0	8.3
		Aggressive	7.5	9.1	8.1	7.4	8.9
Overtaking	SimLingo [42]	Conservative	7.3	6.3	6.2	7.2	6.2
		Neutral	6.9	7.1	7.4	6.8	6.9
		Aggressive	6.1	7.5	7.0	6.0	7.3
	StyleDrive [11]	Conservative	8.1	7.2	7.1	8.0	7.0
		Neutral	7.7	7.9	8.2	7.5	7.8
		Aggressive	6.9	8.4	7.7	6.8	8.2
	DMW	Conservative	9.1	7.7	7.5	8.8	7.6
		Neutral	8.3	8.5	9.0	8.0	8.4
		Aggressive	7.4	9.3	8.2	7.2	9.2
Traffic Sign	SimLingo [42]	Conservative	7.4	7.0	6.9	7.3	6.8
		Neutral	7.1	7.3	7.5	7.0	7.1
		Aggressive	6.8	7.6	7.2	6.7	7.4
	StyleDrive [11]	Conservative	8.3	7.8	7.6	8.1	7.7
		Neutral	7.9	8.1	8.3	7.7	8.0
		Aggressive	7.4	8.6	8.0	7.2	8.4
	DMW	Conservative	8.9	8.3	8.1	8.6	8.2
		Neutral	8.5	8.7	8.8	8.2	8.6
		Aggressive	8.0	9.0	8.4	7.8	8.9

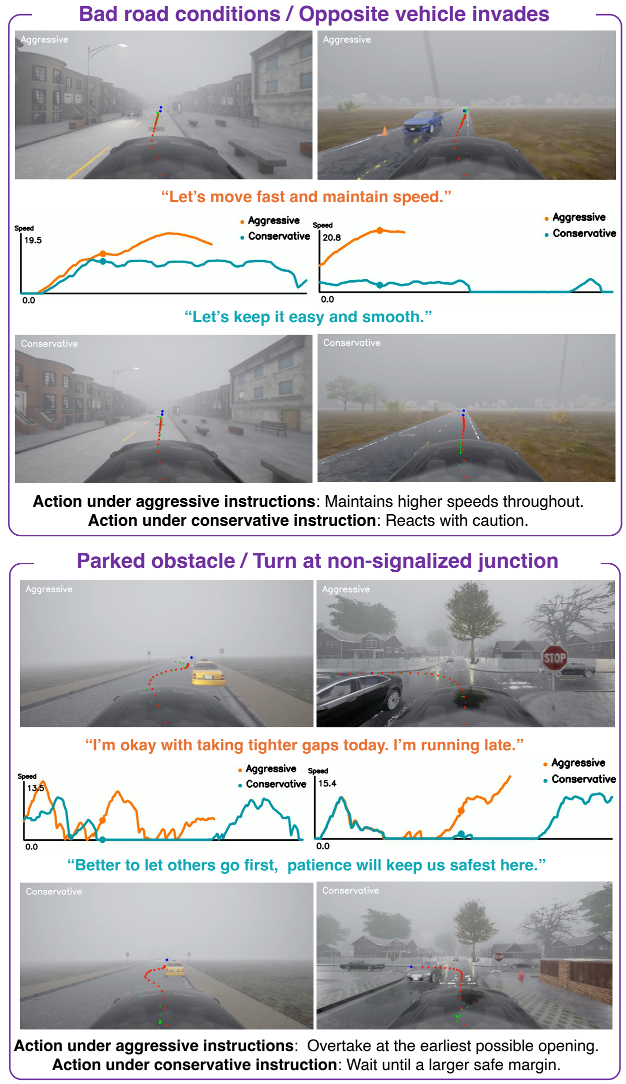

Figure 7. Driving preference under aggressive and conservative instructions. Red waypoints denote distance parametrized (every $1\,\mathrm{m}$) navigation path and green waypoints denote time parametrized (every $0.25\,\mathrm{s}$) trajectory.

lustrate DMW’s ability to adapt in real-time according to short-term style instruction.

B. Personalized Driving Dataset

We provide a more detailed description of the collected personalized driving datasets from thirty real drivers.

B.1. Scenarios

In Town 12, we collect twenty routes that cover a diverse set of representative driving scenarios under varying weather and illumination conditions. The scenarios are summarized below (descriptions are taken from https://leaderboard.carla.org/scenarios/):

- **Accident / ParkedObstacle / ConstructionObstacle**: An obstacle (e.g., a construction zone, an accident, or a parked vehicle) is blocking the ego lane. The ego vehicle must change lanes into traffic moving in the same direction to bypass the obstacle.
- **SignalizedJunctionLeftTurn / NonSignalizedJunctionLeftTurn**: The ego vehicle performs an unprotected left turn at an intersection (can occur at both signalized and unsignalized intersections).
- **CrossingBicycleFlow**: The ego vehicle must execute a turn at an intersection while yielding to bicycles crossing perpendicular to its path.
- **StaticCutIn**: Another vehicle cuts into the ego lane from a queue of stationary traffic. It must decelerate, brake, or change lanes to avoid a collision.

- **NonSignalizedJunctionRightTurn / SignalizedJunctionRightTurn / VanillaNonSignalizedTurn**: The ego vehicle makes a right turn at an intersection while yielding to crossing traffic.
- **InterurbanActorFlow**: The ego vehicle leaves the interurban road by turning left, crossing a fast traffic flow.
- **BlockedIntersection**: While performing a maneuver, the ego vehicle encounters a stopped vehicle on the road and must perform an emergency brake or an avoidance maneuver.
- **HazardAtSideLane**: A slow-moving hazard (e.g., bicycle) partially obstructs the ego vehicle's lane. The ego vehicle must either brake or carefully bypass the hazard (bypassing on the lane with traffic in the same direction).
- **ParkingCutIn**: A parked vehicle exits a parallel parking space into the ego vehicle's path. The ego vehicle must slow down to allow the parked vehicle to merge into traffic.
- **VehicleOpensDoorTwoWays**: The ego vehicle needs to avoid a parked vehicle with its door opening into the lane.
- **DynamicObjectCrossing**: A pedestrian suddenly emerges from behind a parked vehicle and enters the lane. The ego vehicle must brake or take evasive action to avoid hitting the pedestrian.
- **EnterActorFlow**: A flow of cars runs a red light in front of the ego when it enters the junction, forcing it to react (interrupting the flow or merging into the flow). These vehicles are 'special' ones, such as police cars, ambulances, or firetrucks.
- **HighwayExit**: The ego vehicle must cross a lane of moving traffic to exit the highway at an off-ramp.
- **ControlLoss**: The ego vehicle loses control due to bad conditions on the road and it must recover, coming back to its original lane.
- **MergerIntoSlowTraffic**: The ego-vehicle merge into a slow traffic on the off-ramp when exiting the highway.

B.2. Auxiliary Information

To enable reliable and interpretable driving preference analysis, we extract environmental information and motion statistics from driving logs. Concretely, we record data at 5 Hz, including:

- **Camera Sensor.** A forward-facing RGB camera with a resolution of 1024×512 and a wide 110° field of view serves as the primary visual sensor.
- **Ego-state and Control Signals.** We store full ego-vehicle kinematics, including linear acceleration, angular velocity, speed, and the world-frame pose (`location`, `rotation`). Human control commands, including throttle, brake, steering, gear, hand-brake status, and reverse flag, together with the speed limit and lane-level attributes such as lane ID, lane type, lane width, and whether the ego vehicle is currently inside a junction.
- **Surrounding Agents.** We store a detailed description of the leading vehicle (`front_vehicle_info`: ID, type, 3D position, velocity, speed, and color), along with all nearby dynamic agents in `other_vehicles` and `walkers`. For each surrounding actor, we log its transform, velocity, bounding-box extent, and other metadata. Nearby traffic lights and stop signs are also recorded, capturing both their spatial relation to the ego and whether they are currently influencing the ego vehicle.
- **Expert Supervision.** Each timestep is paired with privileged control from the PDM-Lite expert, including `throttle`, `brake`, `steer`, and the corresponding `target_speed`.
- **Route Geometry.** We additionally provide route information in the form of two upcoming waypoints along the global route, transformed into the ego frame as `target_point` and `target_point_next`.

All actor-level annotations are saved as JSON files synchronized with the image index.