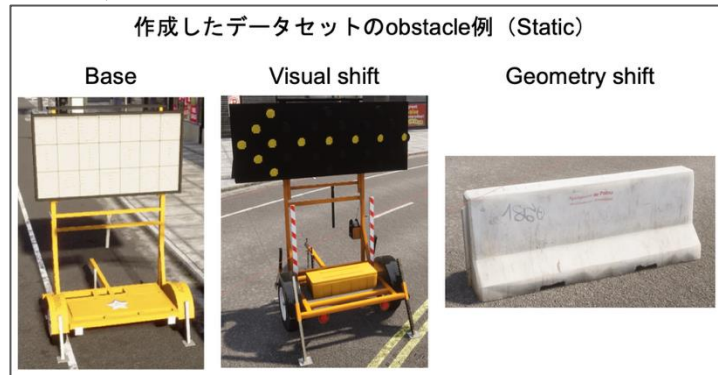


ABCI 3.0開発加速利用 (2025年度) 成果概要 (公開用)

<p>課題名： マルチモーダルAIによる多様なドメインの専門知学習</p>	<p>実施時期：2025年4月15日から2026年3月30日 所属機関名：東北大学 代表者氏名：岡谷貴之</p>
---	--

成果概要：
MLLMのマルチモーダル専門知がどの程度獲得されているかを自動運転タスクで検証するため、既知(ID)とそれを基に生成した未知(OOD)シナリオを含むデータセットを独自に構築し、一般化性能を評価した。OODは外観変化(Visual shift)と形状・サイズ変化(Geometry shift)に分類される。結果、OODで性能低下と衝突増加が確認され、特にGeometry shiftで顕著であった。既存モデルは視覚的特徴には依存できる一方、幾何構造や空間制約に基づく専門知は未知環境に十分汎化できていないことを確認した。

成果のポイント：
本研究は、大規模視覚言語モデル (MLLM) の専門知の獲得の一環として、自動運転分野を対象に実施したものである。既存のMLLMは一般画像には強い一方で、交通シーンにおける運転行動の理解や判断に不可欠な「マルチモーダル専門知」を十分に獲得できていないことが知られている。一方近年では、MLLMを組み込んだモデルを自動運転向けデータセットで学習させることで必要な専門知を付与し、車両の適切な経路推定を可能にする自動運転モデル (例：SimLingo(K.Renz et al., 2025)) が登場し始めている。こうしたモデルは特に学習時に経験したシナリオに対して高い性能を示し、歩行者、車両、信号機といった既知物体が関与する状況では適切な経路を推定できる。しかし実世界では、学習シナリオに含まれない稀な物体 (例：猪、工事車両) に遭遇する可能性がある。こうした未知環境を対象とした評価データセットはまだ整備されていないことから、モデルが専門知を横断的に活用できるレベルで獲得できているか否かは、十分に検証されていない。そこで本研究では、既知シナリオ (ID) とそれをベースに意図的に作成した未知シナリオ (OOD) を含むデータセットを構築し、両者間の性能差を分析することで、専門知の一般化能力を評価する。具体的には、Bench2Drive評価セットを基盤とし、シミュレータ上で既存シナリオ内の物体を新規に導入した未知物体に置き換えることで、OODシナリオを生成する。これらのOODは、(1) 外観のみを変化させるVisual shiftと、(2) 物体の形状やサイズを変化させるGeometry shiftの2種類に分類される。前者は視覚的意味理解の一般化能力を、後者は空間占有や運動制約に基づく行動判断といった、異なる側面から専門知の一般化能力を評価する。実験の結果、OODシナリオにおいて性能低下と衝突率の増加が確認され、特にGeometry shiftで顕著であった。このことは、モデルが視覚的特徴には依存できる一方で、幾何構造や空間的制約に基づく専門知を分布外環境に適用できていないことを示唆する。今後はOODに適用可能なモデルの開発に取り組んでいく。作成したベンチマークおよびデータセットは、今後公開予定である。



Performance (Driving Score ↑ / Number of Collisions ↓)						
Main obstacle	Static			Pedestrian		
OOD type	Base	Visual Shift	Geometry Shift	Base	Visual Shift	Geometry Shift
Rule-based model						
PDM-Lite	100.0 / 0.00	97.4 / 0.71	100.0 / 0.00	100.0 / 0.00	100.0 / 0.00	100.0 / 0.00
End-to-end model (i.e., w/o MLLM)						
DiffAD	42.2 / 9.79	31.1 / 12.02	27.0 / 15.57	60.3 / 4.67	62.7 / 4.12	57.9 / 4.39
TCP	23.7 / 14.65	22.4 / 19.30	16.8 / 36.18	86.1 / 0.71	87.9 / 0.74	83.1 / 1.03
UniAD	24.8 / 16.88	24.5 / 18.27	20.7 / 27.68	47.7 / 5.88	47.9 / 6.79	40.8 / 8.82
VAD	26.2 / 8.43	25.8 / 13.27	19.9 / 31.91	42.4 / 4.76	46.6 / 4.10	48.8 / 5.49
MLLM based model						
ORION	24.5 / 19.02	27.5 / 18.49	16.4 / 34.28	60.9 / 4.70	60.8 / 5.21	51.2 / 5.92
MindDrive	52.0 / 15.77	49.0 / 13.59	51.3 / 15.62	81.5 / 1.65	80.2 / 1.50	65.0 / 4.14
DriveMOE	26.1 / 14.84	28.9 / 18.55	15.3 / 39.67	82.7 / 0.00	72.8 / 0.70	69.3 / 1.68

成果についてより詳細な情報を提供しているWebページ、発表論文などの情報：
 ・本成果は現在論文として投稿準備中であり、公開後に参照可能となる予定。