XXXIII Congresso de Iniciação Científica da Unicamp

Classificação de Comportamento de Motoristas a partir da Análise de Dados Inerciais

Renan Matheus da Silva Florencio - IC, Unicamp
Silvio Fernandes - UFERSA
Profª. Drª. Paula Dornhofer Paro Costa (orientadora) - DCA - FEEC, Unicamp

***Palavras-chave*—Classificação de Comportamento de Motoristas, Machine Learning, Geração de Dados Sintéticos**

I. Introdução

A área de classificação de comportamento de motoristas, ou *Driver Behavior Recognition* (DBR), busca identificar motoristas com comportamentos agressivos ou anormais no trânsito. A detecção de tais comportamentos auxilia na prevenção de acidentes, na emissão de alertas personalizados para motoristas na determinação de padrões de motoristas com maiores chances de causar acidentes. Embora existam diversos estudos na área, eles ainda sofrem de baixa reprodutibilidade, principalmente por usarem majoritariamente dados privados, uma vez que existe uma falta de dados rotulados disponíveis publicamente (a coleta de dados de veículos é custosa e pode ser perigosa). Dessa forma, se torna inviável a comparação entre diferentes técnicas de classificação e de coleta de dados.

Como uma forma de mitigar a falta de dados disponíveis publicamente para classificação de comportamento de motoristas, alguns estudos buscaram a geração de dados sintéticos a partir de simuladores de tráfego [1], [2]. CARLA e SUMO são dois desses simuladores, gratuitos e de código aberto, amplamente utilizados em direção autônoma e sistemas inteligentes de transporte. Embora nenhum deles tenha sido inicialmente desenvolvido para a geração de dados inerciais de alta qualidade, ambos têm sido empregados em contextos nos quais tais dados são de suma importância, como na tomada de decisão autônoma e na classificação de comportamentos de motoristas. No entanto, até onde temos conhecimento, nenhum desses estudos realizou uma verificação da qualidade dos dados gerados em contextos reais, limitando o alcance de suas conclusões.

Este estudo é o primeiro a gerar dados inerciais de comportamentos normais e agressivos a partir de ambos os simuladores e verificar seu realismo, diversidade, coerência e utilidade tomando como base um banco de dados naturalista, o *UAH-DrivesSet* [3]. Além disso, todo o código é disponibilizado, de modo que novos experimentos com diferentes configurações e tratamento de dados possam ser realizados com facilidade. Os resultados mostram que os dados inerciais gerados pelos simuladores não foram capazes de melhorar a performance dos modelos de classificação utilizados (*Random Forests*, *XGBoost* e *Support Vector Machines*) e que podem ser facilmente diferenciados de dados reais. As principais contribuições, portanto, são: (1) a disponibilização dos dados e do repositório utilizados[1], que permitem uma fácil geração e análise dos dados sintéticos dos simuladores para quaisquer configurações; (2) uma análise criteriosa da qualidade dos dados gerados pelos simuladores SUMO e CARLA; (3) uma análise do impacto dos parâmetros de comportamentos de motoristas nos dados gerados.

II. Método

O método se resume na geração de dados sintéticos de comportamentos agressivos e normais de motoristas no trânsito para comparação com o banco de dados reais *UAH-DrivesSet*. Os dados de simuladores foram gerados no mapa padrão *Town01* do simulador CARLA e em sua versão convertida para o simulador SUMO, mantendo as mesmas rotas em ambos os simuladores. Foram testados parâmetros fixos definidos manualmente e parâmetros distintos para cada rota com geração a partir de Grandes Modelos de Linguagem (LLMs). A partir desses dados, foram aplicados testes de realismo, diversidade, coerência e utilidade dos dados por meio de: treinamento em dados reais e validação em dados sintéticos, classificação de dados sintéticos e reais, treinamento em dados aumentados e validação em dados reais, e visualização da projeção dos dados sintéticos e reais [4]. A seção de métodos está dividida em geração de dados sintéticos a partir dos simuladores, métricas de avaliação de qualidade de dados sintéticos e treinamento de modelos de classificação.

[1] https://github.com/H-IAAC/simulated_driving_behavior

A. Geração de Dados Sintéticos

O simulador 2D SUMO (*Simulation of Urban MObility*) surgiu nos anos 2000 com o objetivo de simular redes de tráfego e tem sido muito utilizado por pesquisadores para estudar padrões de mobilidade [5]. Anos depois, o simulador 3D CARLA (*Car Learning to Act*) foi introduzido como uma plataforma focada em direção autônoma, oferecendo suporte para configuração de ambientes e sensores realistas com o apoio do motor físico *Unreal Engine*® [6]. Esses simuladores foram escolhidos por terem sido utilizados em uma ampla gama de tarefas, desde sistemas de assistência a motoristas àqueles que são capazes de tomar decisões autônomas [7].

A geração de dados sintéticos de simuladores foi realizada de modo a explorar o máximo de cada simulador. O mapa escolhido foi um dos mapas padrão 3D do simulador CARLA, *Town01*, que possui uma versão 2D a ser usada no SUMO. Para cada um dos simuladores, foram definidos parâmetros para direção normal e agressiva, divididos em parâmetros fixos, mantidos em todas as rotas, e parâmetros gerados por um LLM, a partir das descrições de cada parâmetro e com diferentes valores para cada rota, aumentando a diversidade das simulações. Foram geradas dez rotas, com o auxílio do LLM, a serem executadas pelos simuladores, passando por diversos pontos de interesse no mapa. Ambas as simulações foram adicionados veículos de fundo, de modo que existam em torno de dez veículos no mapa a todo momento, e todos os semáforos foram congelados para reduzir o tempo e a complexidade das simulações. Todas as coletas foram realizadas a 10 Hz, de acordo com a frequência fornecida pelo *UAH-DriveSet*. O simulador CARLA foi usado no modo síncrono.

Os veículos foram gerados no simulador CARLA com a ferramenta *Traffic Manager* (TM) e somente parâmetros relacionados a semáforos e pedestres foram desconsiderados. O simulador SUMO permite a utilização de diversos modelos de *car-following* e *lane-changing*, que definem o comportamento do veículo na simulação com uma grande quantidade de parâmetros e que podem ter impacto substancial, como mostrado em [8]. Seguindo a recomendação do estudo e do manual do SUMO, foram utilizados os modelos *Extended Intelligent Driver Model* (EIDM) e *SubLane 2015* (SL2015) para *car-following* e *lane-chaning*, respectivamente. Os parâmetros definidos para as simulações estão exibidos nas Tabelas I-IV. Embora o LLM tenha gerado 28 parâmetros para a simulação com SUMO, nem todos foram ajustados manualmente enquanto parâmetros fixos, e somente aqueles que foram ajustados foram exibidos na tabela. Além disso, foram gerados dez conjuntos possíveis de parâmetros pelo LLM para cada simulador, um para cada rotina, e somente um está sendo exibido na tabela. Todos os parâmetros e configurações estão disponíveis no repositório do trabalho.

No total, foram gerados 194 e 201 minutos de dados com o simulador CARLA com parâmetros fixos e de LLM, respectivamente, e 216 e 293 minutos de dados com o simulador SUMO com parâmetros fixos e de LLM, respectivamente.

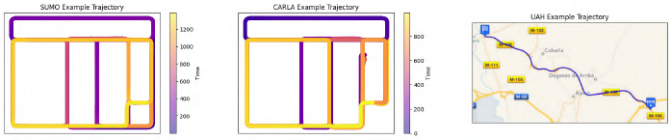

Figura. 1. Exemplos de trajetórias ao longo do tempo do *UAH-DriveSet* e para os simuladores SUMO e CARLA.

A diferença em quantidade de amostras para cada simulador é esperada, pois utilizam diferentes mecanismos de controle de tráfego e de roteamento de veículos, além de momentos em que a simulação no CARLA teve de ser interrompida prematuramente por situações em que o veículo fica preso. Exemplos de trajetórias para os dois simuladores e para o banco de dados real estão mostrados na Figura 1.

Tabela I
PARÂMETROS FIXOS DE COMPORTAMENTO PARA MOTORISTAS COM PERFIL NORMAL E AGRESSIVO NO SIMULADOR SUMO.

Parâmetro	**Normal**	**Agressivo**
Distância mínima ao veículo da frente (m)	2.5	1.0
Velocidade máxima teórica (m/s)	200	250
Capacidade de aceleração (m/s²)	2.6	3.5
Capacidade de frenagem (m/s²)	4.5	5.5
Multiplicador da velocidade da via	1.0	1.2
Frenagem de emergência (m/s²)	9.0	10.0

Tabela II
EXEMPLO DE PARÂMETROS GERADOS POR LLM DE COMPORTAMENTO PARA MOTORISTAS COM PERFIL NORMAL E AGRESSIVO NO SIMULADOR SUMO. UM CONJUNTO DE PARÂMETROS É GERADO PARA CADA ROTINA.

Parâmetro	**Normal**	**Agressivo**
Distância mínima ao veículo da frente (m)	1.9	1.0
Velocidade máxima teórica (m/s)	32.8	48.4
Capacidade de aceleração (m/s²)	2.2	4.5
Capacidade de frenagem (m/s²)	4.0	6.3
Frenagem de emergência (m/s²)	7.1	11.9
Multiplicador da velocidade da via	0.7	0.7

Tabela III
PARÂMETROS FIXOS DE COMPORTAMENTO PARA MOTORISTAS COM PERFIL NORMAL E AGRESSIVO NO SIMULADOR CARLA.

Parâmetro	**Normal**	**Agressivo**
Distância mínima ao veículo da frente (m)	3	1
Chance de ignorar semáforo (%)	100	100
Chance de ignorar placas de pare (%)	0	20
Chance de ignorar veículos (%)	0	10
Chance de manter na faixa lenta (%)	30	0
Prob. de trocar para faixa esquerda (%)	5	20
Prob. de trocar para faixa direita (%)	5	20
Diferença da velocidade alvo vs. limite (%)	20	-15

B. UAH-DrivesSet

UAH-DrivesSet é um banco de dados coletado a partir do aplicativo de celular de monitoramento *DriveSafe*. Ele é

Tabela IV
EXEMPLO DE PARÂMETROS GERADOS POR LLM DE COMPORTAMENTO PARA MOTORISTAS COM PERFIL NORMAL E AGRESSIVO NO SIMULADOR CARLA. UM CONJUNTO DE PARÂMETROS É GERADO PARA CADA ROTINA.

Parâmetro	**Normal**	**Agressivo**
Distância mínima ao veículo da frente (m)	2.7	1.24
Chance de ignorar semáforo (%)	3.7	9.0
Chance de ignorar placas de pare (%)	0.9	5.2
Chance de ignorar veículos (%)	0.9	3.6
Chance de manter na faixa lenta (%)	82.8	37.3
Prob. de trocar para faixa esquerda (%)	10.8	25.2
Prob. de trocar para faixa direita (%)	8.6	14.9
Diferença da velocidade alvo vs. limite (%)	3.3	-2.7

composto de dados de seis veículos e motoristas diferentes realizando três comportamentos (normal, agressivo e sonolento) em dois tipos de estradas (rodovias e estradas secundárias), resultando em mais de 500 minutos de direção com leituras de diversos sensores, como acelerômetro, giroscópio, GPS e câmera [3]. Ele foi escolhido por ser o único banco de dados público encontrado a rotular comportamentos de motoristas e por ser um dos bancos de dados mais citados na área de DBR [7]. Os sensores relevantes para este estudo e fornecidos pelo banco de dados são acelerômetro, giroscópio e GPS (contendo velocidade e ângulo). Dados de acelerômetro e giroscópio são fornecidos para os três eixos em 10 Hz e dados de GPS são fornecidos em 1 Hz.

Embora o banco de dados tenha rótulos para comportamento sonolento, nos voltaremos apenas para os comportamentos normais e agressivos neste estudo, por ser particularmente complexo de modelar o comportamento "sonolento"nos simuladores. Os dados de cinco motoristas performando tais comportamentos foram unidos em único conjunto de dados que foi dividido em treino e validação. Para isso, os dados de GPS foram preenchidos em decorrência da diferença de frequência de coleta, o que foi feito utilizando o valor mais próximo de GPS nos instantes de tempo em que não foram coletados. Por fim, foram obtidos aproximadamente 250 minutos de dados reais para treinamento e 60 minutos de dados para validação.

C. Métricas de Qualidade de Dados Sintéticos

Seguindo métodos encontrados na literatura, como o utilizado em [4], foram avaliados o realismo, a diversidade, a coerência e a utilidade dos dados.

- *Realismo*: Fidelidade com a qual dados individuais representam dados do mundo real, como comportamento e propriedades estatísticas.
- *Diversidade*: Variedade e cobertura dos dados sintéticos, indicando sua capacidade de cobrir uma grande variedade de diferentes variações dos dados reais, como a capacidade de representar comportamentos agressivos de uma ampla gama de motoristas.
- *Coerência*: Consistência e sentido lógico dos dados sintéticos como um todo, indicando, por exemplo, que motoristas agressivos tenham distribuições diferentes de motoristas normais, de maneira consistente.
- *Utilidade*: Capacidade dos dados sintéticos em cumprirem seu propósito, como aumentar a performance de modelos de classificação de comportamentos. É a métrica mais utilizada em estudos de geração de dados sintéticos.

D. Teste de Qualidade de Dados Sintéticos

Para verificar cada métrica, os seguintes testes foram aplicados: treinamento em dados reais e teste em dados sintéticos (TRTS), pontuação discriminativa, ou *Discriminative Score* (DS), pontuação preditiva, ou *Predictive Score* (PS), e visualização da projeção 2D *t-distributed Stochastic Neighbor Embedding (t-SNE)*. Para cada uma das métricas de avaliação foram treinados modelos de aprendizado de máquina e aplicadas varreduras de hiperparâmetros, quando necessário. Os modelos utilizados foram *RandomForest Classifier* (RF), *Support Vector Classifier* (SVC) e *XGBoost Classifier* (XGB), além do modelo de projeção t-SNE.

- Treinamento em dados Reais, Teste em dados Sintéticos (TRTS): Utiliza-se os dados do UAH-DrivesSet como treinamento e dados sintéticos como validação. Todos os modelos foram treinados com varreduras de hiperparâmetros. Bons desempenhos dos modelos nesse teste indicam realismo e utilidade dos dados.
- Pontuação Discriminativa ou *Discriminative Score* (DS): Utiliza-se uma mistura de dados sintéticos e reais, rotulados como sintéticos ou reais, para treinamento e validação. Somente modelo RF foi treinado para esse teste e sua performance foi a máxima possível. Bons desempenhos nessa tarefa indicam que os dados reais são facilmente distinguíveis dos sintéticos, mensurando o realismo.
- Pontuação Preditiva ou *Predictive Score* (PS): Utiliza-se os dados reais aumentados com dados sintéticos para treinamento e dados reais para validação. Foram treinados modelos RF, SVC, e XGB com varredura de hiperparâmetros e aumentações em 20%, 60% e 100% de dados sintéticos. Caso modelos treinados com dados aumentados apresentem um desempenho melhor que modelos treinados somente em dados reais, os dados sintéticos têm utilidade e coerência com a realidade.
- Visualização t-SNE: A projeção t-SNE permite a visualização de dados de alta dimensão em somente duas dimensões, se tornando uma métrica qualitativa de avaliação do realismo e da diversidade dos dados. Uma grande intersecção dos dados sintéticos e reais indica que os dados sintéticos são realistas e possuem uma boa diversidade.

Os modelos foram treinados somente com os sensores inerciais fornecidos por cada um dos simuladores e que também estão presentes no *UAH-DrivesSet*. Para o simulador SUMO, foram usados: magnitude da aceleração e da velocidade e ângulo. Para o simulador CARLA, foram usados: magnitude da aceleração, acelerômetro e giroscópio nos eixos x, y e z, e ângulo. Para realizar a avaliação usando *UAH-DrivesSet* como dados de treino e validação, foram utilizados: acelerômetro e giroscópio nos eixos x, y e z, ângulo e velocidade, tendo sido

excluídos sensores que não são contemplados por nenhum dos simuladores para que a comparação seja justa. Os sensores de posição foram descartados por não terem definição no mapa virtual, que não possui latitude e longitude. Todos os resultados de treinamentos que incluem dados dos simuladores utilizam este mesmo conjunto de sensores, dentre eles as aumentações de dados.

Por se tratarem de dados de séries temporais, que não podem ser fornecidos diretamente para modelos de aprendizado de máquina, foi utilizada a técnica de janelas deslizantes, *sliding windows* ou *rolling windows*, que realiza a quebra da série temporal em janelas de tempo, cada uma com o rótulo do comportamento do motorista que gerou os dados. A série temporal é dividida em diversas janelas, cada uma composta por dez instantes de tempo e com diferença de cinco instantes de tempo entre elas (sobreposição de 50%), que são linearizadas e tratadas como amostras independentes. Embora esse processamento cause a perda de dependências temporais explícitas entre os instantes de tempo, modelos de aprendizado de máquina são capazes de obter bons desempenhos em séries temporais tratadas dessa maneira [9]. Além disso, os dados foram escalados e os comportamentos foram codificados em 0 para normal e 1 para agressivo.

III. Resultados

Um dos resultados principais do trabalho é o repositório disponibilizado, que permite a recriação de todos os experimentos realizados de maneira simples com quaisquer parâmetros, mapas e configurações de simulação de ambos os simuladores. Torna-se, assim, mais simples a comparação entre métodos de aperfeiçoamento dos simuladores (como ajuste de parâmetros de comportamento), propostas de arquiteturas de aprendizado de máquina e geração de dados sintéticos. Além do código, todos os dados gerados foram disponibilizados de maneira organizada e todos os parâmetros de simulação e dos modelos de aprendizado de máquina estão detalhados no repositório.

A Tabela V apresenta os valores de *Train Real, Test Syntehtic* (TRTS). A partir dela é visível que nenhum dos modelos foi capaz de ter uma boa performance, sendo o melhor resultado um F1-score de 54%, muito próximo da aleatoriedade. O resultado sugere que os simuladores não foram capazes de gerar padrões similares aos padrões reais, uma vez que o treinamento em dados reais não foi permitiu que os modelos obtivessem bons resultados em distinguir motoristas normais e agressivos gerados nos simuladores. Embora os dados gerados com LLM tenham obtidos os melhores desempenhos em ambos os simuladores, indicando que a variabilidade dos parâmetros não é irrelevante, o conhecimento de comportamentos reais não se aplicou aos comportamentos gerados sinteticamente, tendo pouco realismo e baixa utilidade.

A Tabela VI apresenta os valores de *Discriminative Score* (DS). A partir dela, vê-se que os modelos foram capazes de distinguir perfeitamente entre os dados sintéticos e reais ao serem treinados em bancos de dados mesclados de ambos os simuladores, mesmo sem uma varredura de hiperparâmetros. Este resultado sugere que amostras individuais de dados sintéticos são facilmente distinguíveis de amostras individuais de dados reais, ratificando seu baixo realismo tanto com parâmetros fixos quanto com parâmetros gerados pelo LLM.

A tabela VII apresenta os valores de *Predictive Score* (PS), o teste mais utilizado para medição de utilidade de dados sintéticos. Nela, vê-se que treinamento e validação somente com o *UAH-DriveSet*, sem nenhuma aumentação de dados, levou ao melhor F1-score de 79%, seguido por diversos bancos de dados com aumentação a partir do simulador SUMO e as piores performances para aumentações usando o simulador CARLA. Embora a aumentação de dados tenha falhado para ambos os simuladores, dado que a performance dos modelos se tornou pior, os resultados mostram que dados gerados pelo simulador SUMO tendem a perturbar menos o desempenho. Dessa forma, estudos que buscam realizar um ajuste mais fino de parâmetros de simulação do SUMO, como [10], podem auxiliar na geração de dados mais úteis e coerentes, embora os outros testes também não tenham sido favoráveis para o simulador. Não houve diferença substancial na geração de parâmetros com o LLM ou mantendo-os fixos.

Por fim, as projeções t-SNE 2D são mostradas na Figura III, com rotulamento de dados reais e sintéticos. Embora os agrupamentos não estejam distantes, a divisão dos dados reais e sintéticos é muito clara e quase não existem regiões de interseção de dados sintéticos e reais, principalmente para dados do simulador CARLA, corroborando com os resultados apresentados nos três testes anteriores e explicitando a também falta de diversidade nos dados, mesmo nos dados com parâmetros variados gerados pelo LLM.

Tabela V
Métricas de performance para TRTS ordenadas por F1-score. Nenhum modelo treinado em dados reais teve bom desempenho em dados sintéticos. XGB foi o melhor modelo em todos os testes.

Dados comparados	F1-score	Acurácia	Recall	Precisão
sumo_llm	0.54	0.55	0.55	0.54
sumo_fixed	0.50	0.51	0.51	0.50
carla_llm	0.50	0.50	0.50	0.50
carla_fixed	0.43	0.59	0.59	0.34

Tabela VI
Métricas de performance para DS ordenadas por F1-score. Todos os modelos são RF e obtiveram performance perfeita, indicando dados sintéticos facilmente separáveis.

Dados comparados	F1-score	Acurácia	Recall	Precisão
sumo_fixed	1.00	1.00	1.00	1.00
sumo_llm	1.00	1.00	1.00	1.00
carla_fixed	1.00	1.00	1.00	1.00
carla_llm	1.00	1.00	1.00	1.00

IV. Conclusão

Os resultados do experimento validam o funcionamento do repositório criado e demonstram que os simuladores SUMO e CARLA não foram capazes de apresentar bons resultados

Tabela VII
MÉTRICAS DE PERFORMANCE PARA PS ORDENADAS POR F1-SCORE PARA DIVERSOS DATASETS. OS SUFIXOS *fixed* E *llm* INDICAM A ORIGEM DOS PARÂMETROS, OS SUFIXOS *20, 60, 100* INDICAM A PORCENTAGEM DE DADOS SINTÉTICOS. A MELHOR PERFORMANCE FOI DO MODELO QUE NÃO RECEBEU NENHUM DADO SINTÉTICO.
MELHOR MODELO: ⋆ RF; † XGB; ⋄ SVC.

Dados comparados	**F1-score**	**Acurácia**	**Recall**	**Precisão**
⋄ uah	0.79	0.79	0.79	0.79
† sumo_uah_fixed_20	0.74	0.74	0.74	0.74
† sumo_uah_llm_60	0.73	0.73	0.73	0.73
† sumo_uah_fixed_100	0.73	0.73	0.73	0.73
⋆ sumo_uah_llm_20	0.73	0.73	0.73	0.73
⋆ sumo_uah_llm_100	0.73	0.73	0.73	0.73
⋆ sumo_uah_fixed_60	0.73	0.73	0.73	0.73
† carla_uah_fixed_20	0.66	0.67	0.67	0.68
† carla_uah_llm_20	0.66	0.67	0.67	0.69
† carla_uah_llm_60	0.65	0.66	0.66	0.69
⋆ carla_uah_fixed_100	0.64	0.64	0.64	0.64
⋆ carla_uah_llm_100	0.63	0.63	0.63	0.63
⋆ carla_uah_fixed_60	0.61	0.61	0.61	0.61
⋆ carla_llm	0.44	0.48	0.48	0.46
† sumo_llm	0.42	0.49	0.49	0.48
⋆ carla_fixed	0.38	0.49	0.49	0.48
⋆ sumo_fixed	0.34	0.50	0.50	0.75

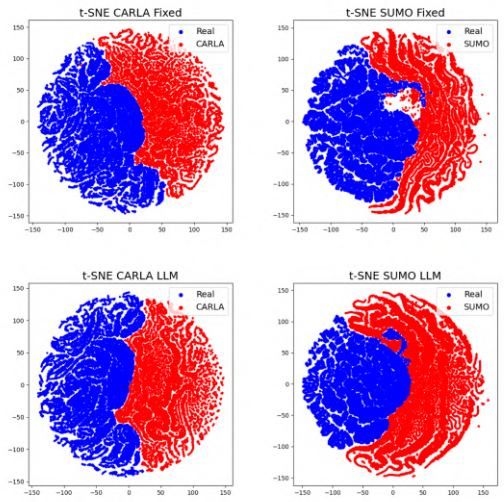

Figura. 2. Projeções t-SNE 2D comparando dados reais com dados sintéticos para simuladores CARLA e SUMO com parâmetros fixos e gerados por LLM. Nota-se claramente a divisão entre dados sintéticos e reais.

em nenhum dos testes (realismo, diversidade, coerência e utilidade) com os parâmetros utilizados, se mostrando insuficientes para aplicações reais com o banco de dados *UAH-DriveSet*. Dentre as possíveis causas da falha estão a qualidade da simulação e dos modelos de tráfego, a qualidade dos sensores gerados (desde aplicação de ruído até modelos físicos) e a dissimilaridade entre o cenário real e o cenário simulado.

O simulador CARLA, embora 3D e munido de um motor físico, apresenta variações muito bruscas de aceleração e desaceleração dos veículos, levando a um padrão pouco realista de acelerômetro durante as simulações e, consequentemente, a resultados ruins quando aplicados a cenários reais. O simulador SUMO, embora apresente melhores modelos de tráfego e direção, fornecendo dados mais próximos do esperado, não é voltado para a avaliação de sensores e não fornece diversos eixos de acelerômetro e giroscópio, limitando seu alcance para aplicações em cenários reais e dificultando seu uso para aumentação de dados. Além disso, a diferença entre as vias em que foi coletado o *UAH-DriveSet* e aquelas do mapa virtual *Town01* podem ter afetado os resultados obtidos.

Por fim, os resultados salientam a necessidade de validação em ambientes reais de modelos treinados em ambientes virtuais com os simuladores CARLA e SUMO, reiterando que a validação em dados sintéticos pode ser criticamente enviesada devido à dificuldade de geração de dados sintéticos de veículos no tráfego. É fundamental que estudos utilizando os simuladores tenham bases de comparação sólidas para que seus resultados sejam conclusivos. Os próximos passos do trabalho envolvem a definição de parâmetros de forma mais rigorosa a partir da comparação com dados reais para verificar a melhora em performance, o treinamento dos modelos com bancos de dados sintéticos pré-processados de modo a tornar os sensores dos simuladores mais realistas, e a aplicação de extratores de características antes do treinamento dos modelos.

REFERÊNCIAS

[1] S. B. Brahim, H. Ghazzai, H. Besbes, and Y. Massoud, "A Machine Learning Smartphone-based Sensing for Driver Behavior Classification," in *2022 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2022, pp. 610–614.

[2] B. Chah, A. Lombard, Y. Mualla, A. Bkakria, A. Abbas-Turki, and R. Yaich, "Building a Database of Simulated Driver Behaviors Using the SUMO Simulator," in *Intelligent Systems and Applications*, K. Arai, Ed. Cham: Springer Nature Switzerland, 2024, pp. 536–555.

[3] E. Romera, L. M. Bergasa, and R. Arroyo, "Need data for driver behaviour analysis? Presenting the public UAH-DriveSet," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE Press, 2016, p. 387–392. [Online]. Available: https://doi.org/10.1109/ITSC.2016.7795584

[4] N. Alzahrani, J. Cała, and P. Missier, "Experience: A comparative analysis of multivariate time-series generative models: A case study on human activity data," *J. Data and Information Quality*, vol. 16, no. 3, Oct. 2024. [Online]. Available: https://doi.org/10.1145/3688393

[5] P. Alvarez Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, and E. Wießner, "Microscopic Traffic Simulation using SUMO," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, November 2018, pp. 2575–2582. [Online]. Available: https://elib.dlr.de/127994/

[6] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An Open Urban Driving Simulator," in *Proceedings of the 1st Annual Conference on Robot Learning*, 2017, pp. 1–16.

[7] S. Sarker, B. Maples, I. Islam, M. Fan, C. Papadopoulos, and W. Li, "A Comprehensive Review on Traffic Datasets and Simulators for Autonomous Vehicles," 2025.

[8] M. Schrader, M. Al Abdraboh, and J. Bittle, "Comparing Measured Driver Behavior Distributions to Results from Car-Following Models using SUMO and Real-World Vehicle Trajectories from Radar: SUMO Default vs. Radar-Measured CF model Parameters," *SUMO Conference Proceedings*, vol. 4, p. 41–54, Jun. 2023. [Online]. Available: https://www.tib-op.org/ojs/index.php/scp/article/view/214

[9] S. Bouhsissin, N. Sael, and F. Benabbou, "Driver behavior classification: A systematic literature review," *IEEE Access*, vol. 11, pp. 14 128–14 153, 2023.

[10] A. Kusari, P. Li, H. Yang, N. Punshi, M. Rasulis, S. Bogard, and D. J. LeBlanc, "Enhancing SUMO simulator for simulation based testing and validation of autonomous vehicles," in *2022 IEEE Intelligent Vehicles Symposium (IV)*, 2022, pp. 829–835.