

Predição de Geolocalização de Veículo com Alerta de Roubo Usando LSTM, Transformer e TLE

Gustavo V. I. de Macedo¹, Geraldo P. Rocha Filho², João K. M. dos Santos¹, Arthur R. Neves¹, Murilo G. Almeida³, Mariana C. Falqueiro¹, Rodolfo I. Meneguette⁴, André L. M. Serrano⁵, Fábio L. L. de Mendonça¹, Vinícius P. Gonçalves¹

¹Departamento de Engenharia Elétrica – Universidade de Brasília (UnB)
– Brasília – DF – Brazil

²Universidade Estadual do Sudoeste da Bahia (UESB)
– Vitória da Conquista – BA – Brazil

³Universidade Estadual Paulista (UNESP)
– Bauru – SP – Brazil

⁴Universidade de São Paulo (USP)
– São Carlos – SP – Brazil

⁵Departamento de Engenharia de Produção – Universidade de Brasília (UnB)
– Brasília – DF – Brazil

{gustavo.isobe, joao.kleber, rodrigues.neves}@aluno.unb.br,
geraldrocha@uesb.edu.br, murilogo@unesp.br, meneguette@icmc.usp.br,
fabio.mendonca@redes.unb.br, {andrelms, mariana.cabral, vpgvinicius}@unb.br

Abstract. *Vehicles External Sensor Trajectory (EST) Prediction has become a real necessity when combining the OCR capabilities of modern surveillance cameras with vehicle theft alerts [Neto et al. 2021]. In this work, using LSTM, Transformer, and a new proposed Transformer and LSTM Entangled (TLE) architecture, an approach is proposed using street lane data to enhance the prediction of the next external sensor. The models results are remarkable, achieving an MSE of 5.9×10^{-6} in some cases. In this context, it is observed that street lane data are useful information in most cases.*

Resumo. *Predição de Trajetória de Sensor Externo (TSE) veicular se tornou uma necessidade real ao juntar a capacidade de OCR das câmeras modernas de vigilância com os alertas de roubos de veículos [Neto et al. 2021]. Neste trabalho, utilizando as técnicas de deep learning LSTM, Transformer e uma nova arquitetura proposta de Transformer e LSTM Entrelaçados (TLE), propõe-se utilizar os dados de vias das ruas para aprimorar a predição de próximo sensor externo. Os resultados dos modelos surpreendem ao atingir MSE de $5,9 \times 10^{-6}$ em alguns casos. Nesse contexto, observa-se que os dados de vias das ruas são informações úteis na maioria dos casos.*

1. Introdução

Quando um veículo é roubado, o proprietário deve realizar um Boletim de Ocorrência (BO) em uma delegacia física ou virtual. Essa ação gera um alerta que, quando uma

câmera de vigilância identifica por sua placa o veículo roubado, usando OCR - *Optical Character Recognition* - para traduzir a imagem da câmera em texto, é registrado o local e a hora em que foi identificada a placa (possíveis erros de identificação foram desconsiderados). Ao prever a próxima posição discreta de um veículo roubado pode-se auxiliar a recuperação desses veículos.

Para realizar esse trabalho, foram utilizados dados anonimizados dos alertas dos estados do Ceará, de Goiás e do Distrito Federal, que possuem quantidades de registros e complexidades distintas, abordando, assim, 18 cenários diferentes de predições. A Figura 1 ilustra a complexidade e a dispersão dos sensores onde esses alertas foram gerados.

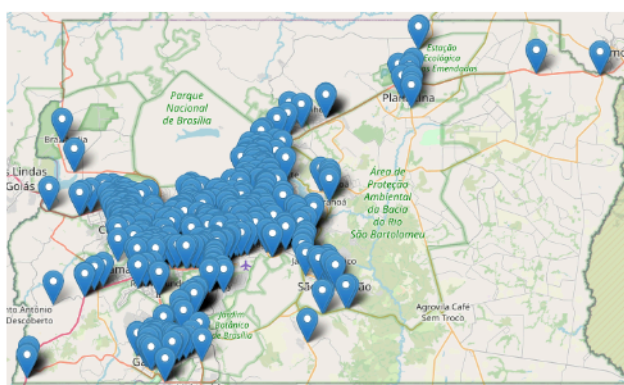


Figura 1. Sensores do DF.

De um ponto de vista psicossocial, um motorista fora da lei, tende a ter um comportamento anômalo ao volante em relação a um cidadão comum [Haviland and Wiseman 1974]. Ainda neste sentido, esse mesmo motorista, poderá tentar desviar das rotas que tenha ciência de possuir câmeras. Isso dificulta a predição, tanto pelo fator anômalo da trajetória como por reduzir os registros do veículo por câmeras.

Para esse desafio peculiar, neste trabalho traz-se dois novos pontos para análise: identificar se a inclusão de dados de qual via da rua o veículo foi identificado pode melhorar a predição; e se uma nova arquitetura Transformer e LSTM Entrelaçados (TLE) aqui sendo proposta seria capaz de prever melhor em relação às arquiteturas convencionais para esse cenário. Esses dois pontos serão comparados com modelos treinados com e sem os dados de qual via o veículo foi identificado e utilizando, além da nova arquitetura Transformer e LSTM Entrelaçados (TLE), as arquiteturas LSTM em seu estado-da-arte e a Transformer baseada em [Vaswani et al. 2017] como comparativos referenciais.

A seguir, tem-se na seção II os trabalhos relacionados, na seção III os detalhes sobre a predição de geolocalização de veículo, na seção IV os experimentos, na seção V tem-se as conclusões e trabalhos futuros e então os agradecimentos e as referências.

2. Trabalhos Relacionados

A predição de trajetória de próximo sensor externo, devido aos dados esparsos e discretos, tem desafios maiores do que uma predição de trajetória contínua, como a representação da trajetória entre os sensores, uma vez que não é possível definir o trajeto real.

Em [Cruz et al. 2022] é utilizado LSTM e modelo Transformer pré-treinado BERT no contexto de predição de próximo sensor externo em Fortaleza-CE, Brasil, para

veículos roubados. Nesse trabalho o modelo que utiliza, além do LSTM, o BERT e é realizado, ainda, um ajuste fino, obtém o melhor resultado. Nele, não foi utilizado como comparativo o uso da arquitetura Transformer treinada do zero, além do uso híbrido das arquiteturas LSTM e Transformer ser sequencial em vez de entrelaçados e não se utilizou a informação de via da rua para a predição.

Já em [Neto et al. 2021] também é utilizado LSTM para predição de próximo sensor externo em Fortaleza-CE, Brasil, para veículos roubados. No entanto, não é utilizada a arquitetura Transformer e, apesar de trazer o conceito de adicionar semântica às proximidades dos sensores externos, não é utilizada a informação de via da rua nas predições.

[Tsiligkaridis et al. 2020] utiliza dados contínuos de trajetórias parciais de táxis nas arquiteturas LSTM e Transformer para predição de destino, não envolvendo veículos roubados e os dados, apesar de terem um trecho de descontinuidade, não são discretos. Nele a arquitetura LSTM atinge melhores resultados. E em [Capanema et al. 2020] a predição é baseada em dados históricos do próprio motorista que se está prevendo o próximo destino, o que traz uma grande vantagem em relação à maioria das outras situações em que o treinamento é realizado com dados de outros veículos. Já em [Hu et al. 2022] a predição é focada em pequenas distâncias, com até 5 segundos a frente.

Uma linha de trabalhos como [Ladeira et al. 2019b], [Ladeira et al. 2020], [de Souza et al. 2019] e [Brito et al. 2023], incluindo [Almeida et al. 2022] para ônibus, [Ladeira et al. 2019a] que tem seu foco em janela de tempo, além de [de Souza and Villas 2020] e [Karimzadeh et al. 2021] que utilizam arquitetura de *reinforcement learning*, avaliam e fazem predição de rota mais segura, desviando de áreas com maior probabilidade de criminalidade. Isso é uma tendência contrária a este trabalho, pois queremos prever justamente onde um veículo roubado vai estar.

Outros trabalhos de predição de trajetória que utilizam LSTM e Transformer como [Chen et al. 2022], [Xu et al. 2023], [Bae et al. 2022] e [Gaiduchenko et al. 2020] não utilizam dados de veículos e diversificam em relação a qual arquitetura se sobressai.

Considerando, ainda, que, em trabalhos analisados, uns obtiveram melhores resultados com arquitetura LSTM e outros com arquitetura Transformer, traz-se neste trabalho ambas arquiteturas LSTM e Transformer e, adicionalmente, uma nova arquitetura Transformer e LSTM Entrelaçados com o objetivo de verificar uma possível complementação de vantagens de cada arquitetura para atuar neste cenário adverso de predição.

Este trabalho, portanto, se destaca por utilizar dados de via da rua, com a identificação dos sensores externos, para predição e uso de arquitetura LSTM e Transformer Entrelaçados (TLE) treinados ambos do zero, como observado na Tabela 1:

3. Predição de Geolocalização de Veículo

Para a predição de geolocalização de um veículo, no contexto de próximo sensor externo, tem-se uma margem de imprecisão. Em ruas com mais de uma via, registra-se a mesma latitude e longitude, independente da via em que o veículo foi identificado. Para a predição, considerando um problema de regressão para o modelo de Inteligência Artificial (IA), a predição poderá ser uma geolocalização na qual não existe um sensor externo, necessitando ajuste para uma latitude e longitude de um sensor externo conhecido. Há, ainda, a imprecisão da trajetória real total, tendo apenas posições pontuais e esparsas registradas.

Tabela 1. Resumo comparativo entre trabalhos

#	Artigo	LSTM	Transformer	Transformer + LSTM Entrelaçados	Dados Esparsos/ Discretos	Dados de Veículos Roubados	Uso de Informação de Via da Rua
1	[Chen et al. 2022]	✓	✓	×	×	×	×
2	[Xu et al. 2023]	✓	✓	×	×	×	×
3	[Bae et al. 2022]	✓	✓	×	×	×	×
4	[Gaiduchenko et al. 2020]	✓	✓	×	✓	×	×
5	[Tsiligkaridis et al. 2020]	✓	✓	×	×	×	×
6	[Neto et al. 2021]	✓	×	×	✓	✓	×
7	[Cruz et al. 2022]	✓	×	×	✓	✓	×
8	Este Trabalho	✓	✓	✓	✓	✓	✓

Na Figura 2 apresenta-se uma situação de predição de próximo sensor externo. Nessa imagem verifica-se que o veículo vermelho poderia virar à esquerda, onde seria identificado novamente no sensor externo na parte superior da imagem. Poderia, também, seguir reto e ser identificado no sensor externo mais à direita da imagem. Poderia, ainda, virar à direita e não ser identificado por nenhum outro sensor externo próximo ao ponto inicial. No caso da Figura 2 a predição chegou à conclusão de que o mais provável é o veículo seguir reto. Neste cenário, visualiza-se que utilizar a informação da via da rua na qual o veículo foi identificado pode auxiliar na predição de próximo sensor externo, mesmo não tendo o trajeto exato intermediário entre os pontos.

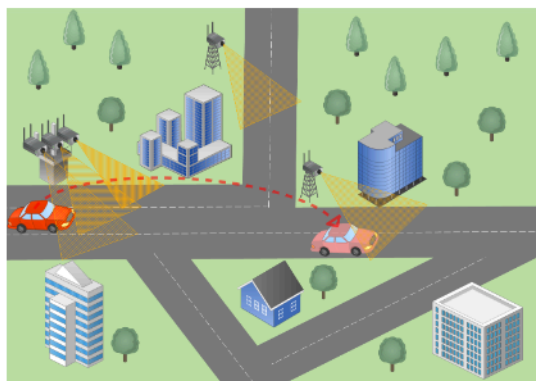


Figura 2. Predição de próximo sensor externo.

Nesta seção, serão definidas as abordagens, as arquiteturas de IA (LSTM, Transformer e TLE) e a metodologia dos experimentos.

3.1. Abordagens e Arquiteturas

Duas abordagens sobre os registros dos sensores externos são utilizadas:

Abordagem 1 *Com identificação da via, além da latitude e da longitude.*

Abordagem 2 *Sem identificação da via, utilizando apenas a latitude e a longitude.*

Para cada abordagem, três arquiteturas foram utilizadas para os treinamentos e predições: LSTM; Transformer; e Transformer e LSTM Entrelaçados (TLE). Essas arquiteturas são detalhadas nas subseções 3.1.1, 3.1.2 e 3.1.3, respectivamente. A função de perda “root mean squared error” (RMSE) calcula a distância euclidiana entre a predição e o ponto real e foi adotada por obter resultados melhores em testes em relação à “mean squared error” (MSE), à “mean absolute error” (MAE) e a outras funções customizadas.

3.1.1. Arquitetura LSTM

A arquitetura LSTM utilizada foi o estado-da-arte descrita em [Gaiduchenko et al. 2020] utilizando algumas características específicas como a função de ativação *Gaussian Error Linear Unit* (GELU) [Hendrycks and Gimpel 2016], *Glorot Initialization* [Sun 2019] e o otimizador AdamW como especificado. Essa arquitetura LSTM, ao ser criada, possui cerca de 3,4 milhões de parâmetros treináveis e é representada na Figura 3.

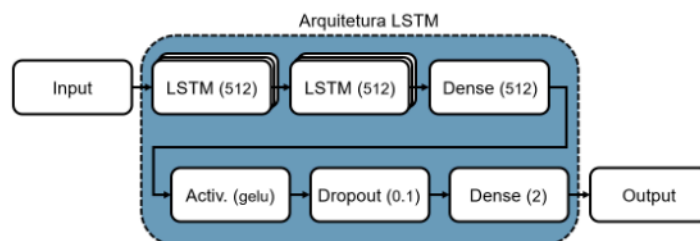


Figura 3. Arquitetura LSTM.

3.1.2. Arquitetura Transformer

A arquitetura Transformer utilizada foi baseada em [Vaswani et al. 2017] e possui cerca de 100 mil parâmetros treináveis. A arquitetura Transformer por realizar execuções em paralelo internamente, diferentemente da LSTM que tem execução sequencial, juntando com a menor quantidade de parâmetros treináveis, o tempo de treinamento dessa arquitetura se torna significativamente menor em relação aos demais modelos deste trabalho.

Em adaptação ao modelo proposto em [Vaswani et al. 2017], considerando que os tipos de dados de entrada não são palavras, testes foram realizados e a arquitetura padrão foi ajustada para sua melhor performance e está representada na Figura 4.

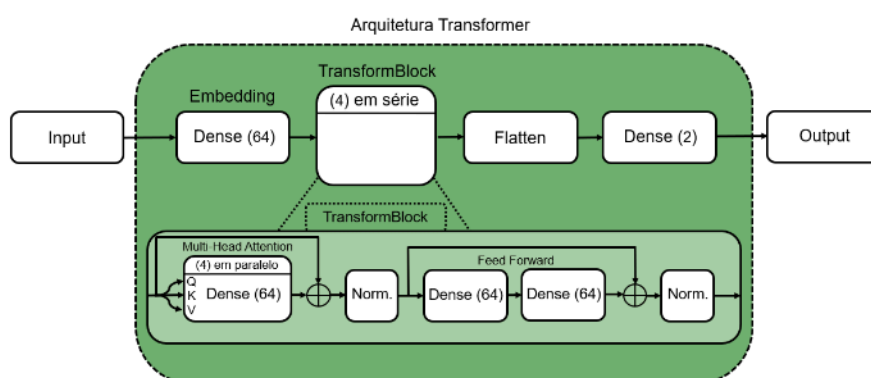


Figura 4. Arquitetura Transformer.

3.1.3. Arquitetura TLE - Transformer e LSTM Entrelaçados

Quando o GPT-3 desbancou o LSTM na área de Processamento de Linguagem Natural (PLN), gerou-se a expectativa de que a arquitetura Transformer seria melhor do

que a LSTM em predição de sequências, mesmo em área distinta de PLN. No entanto, existem artigos nos quais um modelo LSTM tem melhor performance do que modelos Transformer, como em [Bae et al. 2022] e [Gaiduchenko et al. 2020], e em outros artigos o modelo Transformer obtém melhores resultados como em [Chen et al. 2022], [Xu et al. 2023] e [Tsiligkaridis et al. 2020]. Logo, ainda não se tem uma arquitetura com melhor desempenho independente do contexto.

Ainda em [Cruz et al. 2022] há o uso de uma arquitetura híbrida de LSTM e Transformer, utilizando o modelo pré-treinado BERT [Devlin et al. 2018]. Naquele contexto, a arquitetura híbrida obteve os melhores resultados, deixando um viés de superação para uma arquitetura híbrida em relação às arquiteturas LSTM e Transformer individualizadas.

Com essas observações, é proposta uma nova arquitetura em que se adota uma arquitetura Transformer como base e se insere uma camada LSTM em 3 pontos distintos da arquitetura, por isso chamada de Transformer e LSTM Entrelaçados (TLE). Esse modelo conta com cerca de 1 milhão de parâmetros treináveis e é representado na Figura 5.

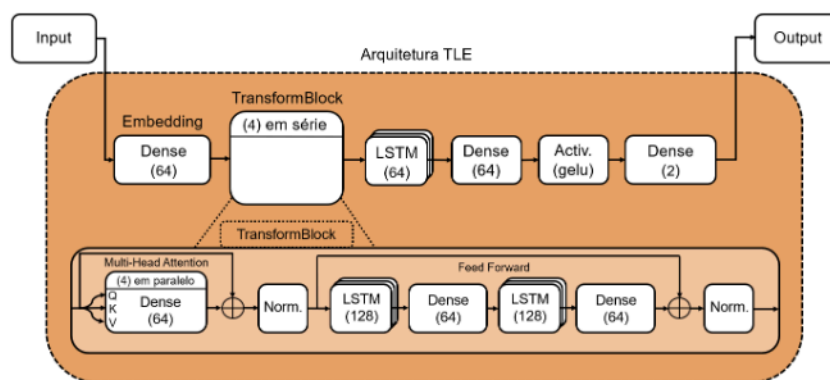


Figura 5. Arquitetura TLE.

3.2. Metodologia

Neste trabalho são comparadas três arquiteturas diferentes: o estado-da-arte da LSTM (fig. 3); uma arquitetura Transformer padrão (fig. 4); e uma nova arquitetura TLE (fig. 5) com potencial de superar as anteriores por conter elementos de ambas. Cada arquitetura foi treinada com os registros de sensores externos que geraram alertas de veículos roubados ou furtados, durante o período de junho de 2021 a setembro de 2023, em três estados brasileiros: Ceará (CE), Goiás (GO) e Distrito Federal (DF).

O problema é abordado como um problema de regressão, devido à natureza contínua das coordenadas geográficas, à grande quantidade de possíveis pontos geográficos e para uma melhor capacidade de generalização do modelo, como prever pontos que não apareceram nos dados de treinamento.

Em cada combinação de arquitetura e estado brasileiro foi treinado o modelo considerando três valores distintos de sequências para, com base nessas sequências, o modelo prever a próxima geolocalização que o veículo seria identificado. Assim, foram realizados treinamentos com sequências de 2 localizações para prever a 3ª, abordando um cenário minimalista, sequências de 10 localizações para prever a 11ª, abordando um cenário mediano e que, com base em testes, tende a ser o ponto ótimo de predição e

sequências de 30 localizações para prever a 31ª, abordando um cenário de maior volume de histórico, para analisar a premissa de que quanto mais histórico, melhor é a previsão.

Com base em [Joseph 2022] onde se propõe uma divisão ótima de dados de treinamento e teste na proporção de $\sqrt{p} : 1$, onde p é o número de parâmetros do modelo, faz-se uma ponderação do resultado e divide-se em base de treinamento com 95% dos registros e base de teste com os 5% dos registros restantes. Essa divisão condiz com o contexto de alta variabilidade dos dados e modelos complexos de *deep learning*. Ao todo, foram 3.213 previsões para o CE: 2.226 para a 3ª localização, 966 para a 11ª e 21 para a 31ª. Para o GO, foram 16.506 previsões, com 7.413 para a 3ª localização, 5.460 para a 11ª e 3.633 para a 31ª. Já para o DF, foram 22.428 previsões, sendo 14.952 para a 3ª localização, 5.460 para a 11ª e 2.016 para a 31ª

Como boa prática em treinamentos de modelos de *deep learning*, os dados de entrada foram normalizados entre 0 e 1 e os dados de teste também normalizados considerando apenas os dados de treinamento. Em seguida, foi aplicado o conceito de janela deslizante subdividindo sequências grandes em subsequências de tamanho fixo e, então, essas subsequências de entrada foram embaralhadas randomicamente.

Para analisarmos se a hipótese de que a informação de via da rua auxilia na previsão, adota-se as Abordagens 1 e 2 para cada um desses cenários. Portanto, cada modelo foi treinado considerando os dados de ID do sensor externo e desconsiderando esse dado, para avaliar se essa informação melhora a previsão ou se, ao contrário, gera ruído no treinamento dificultando a previsão. Ao todo foram 18 treinamentos para cada arquitetura, totalizando 54 treinamentos, para gerar, então, as previsões desses modelos treinados. Para plotar em mapas os pontos e trajetos foram utilizados a linguagem *python* e o *OpenStreetMap*.

4. Experimentos

Como métricas de avaliação para os resultados dos testes, foram utilizadas as métricas RMSE (*Root Mean Squared Error*), MSE (*Mean Squared Error*) e MAE (*Mean Absolute Error*) que são métricas convencionais para problemas de regressão [Brownlee 2021] [Botchkarev 2018].

A tabela 2 traz os dados de avaliação das previsões para a base de teste das métricas RMSE, MSE e MAE para cada arquitetura e seus respectivos 18 cenários.

Analisando os resultados e os gráficos das Figuras 6 e 7 percebe-se que a Abordagem 1 favorece uma melhor performance média para as 3 quantidades de sequências de entrada ($n_{seq} = (2, 10, 30)$) para previsão de próximo sensor externo. Para a média geral, a Abordagem 1 promove uma melhoria média de 6,5% no MSE das previsões.

Em relação à análise entre as 3 arquiteturas, os gráficos das Figuras 8 e 9 mostram que a LSTM levou vantagem no MSE na maior parte das avaliações.

Na média geral, a LSTM obteve uma performance 12,4% melhor do que a TLE e 33,3% melhor do que a Transformer. No entanto, existe um cenário, que é para o $n_{seq} = 2$, em que a nova proposta de arquitetura TLE (Transformer e LSTM Entrelaçados) obtém o melhor MSE médio, sendo 0,74% melhor em comparação à LSTM e 2,3% melhor em comparação à Transformer.

Tabela 2. Resultados das avaliações das predições da base de teste.

Parâmetros \ Arquiteturas			LSTM			Transformer			TLE		
Estado	n_seq	Abord.	RMSE	MSE	MAE	RMSE	MSE	MAE	RMSE	MSE	MAE
CE	2	1	0.0072	5.36×10^{-5}	0.0052	0.0171	2.93×10^{-4}	0.0110	0.0112	1.26×10^{-4}	0.0084
	2	2	0.0074	5.59×10^{-5}	0.0053	0.0159	2.52×10^{-4}	0.0102	0.0120	1.45×10^{-4}	0.0086
	10	1	0.0070	4.86×10^{-5}	0.0053	0.0134	1.80×10^{-4}	0.0115	0.0096	9.15×10^{-5}	0.0080
	10	2	0.0076	5.72×10^{-5}	0.0057	0.0103	1.07×10^{-4}	0.0087	0.0102	1.06×10^{-4}	0.0089
	30	1	0.0102	1.03×10^{-4}	0.0102	0.0281	7.87×10^{-4}	0.0259	0.0154	2.37×10^{-4}	0.0153
	30	2	0.0120	1.45×10^{-4}	0.0107	0.0121	1.46×10^{-4}	0.0095	0.0138	1.91×10^{-4}	0.0119
GO	2	1	0.0076	1.20×10^{-4}	0.0026	0.0091	1.92×10^{-4}	0.0030	0.0058	6.85×10^{-5}	0.0024
	2	2	0.0075	1.19×10^{-4}	0.0026	0.0092	1.83×10^{-4}	0.0031	0.0064	8.68×10^{-5}	0.0024
	10	1	0.0024	5.90×10^{-6}	0.0017	0.0027	7.53×10^{-6}	0.0020	0.0024	5.97×10^{-6}	0.0017
	10	2	0.0029	8.74×10^{-6}	0.0023	0.0030	9.18×10^{-6}	0.0022	0.0028	8.05×10^{-6}	0.0020
	30	1	0.0025	6.52×10^{-6}	0.0018	0.0036	1.35×10^{-5}	0.0025	0.0036	1.31×10^{-5}	0.0025
	30	2	0.0026	6.67×10^{-6}	0.0018	0.0036	1.35×10^{-5}	0.0025	0.0037	1.37×10^{-5}	0.0025
DF	2	1	0.0899	0.0083	0.0654	0.0888	0.0081	0.0642	0.0890	0.0082	0.0640
	2	2	0.0900	0.0083	0.0654	0.0892	0.0082	0.0644	0.0894	0.0082	0.0645
	10	1	0.0798	0.0065	0.0582	0.0956	0.0095	0.0702	0.0780	0.0062	0.0568
	10	2	0.0801	0.0066	0.0584	0.1000	0.0104	0.0744	0.0867	0.0077	0.0619
	30	1	0.0694	0.0051	0.0463	0.0989	0.0098	0.0692	0.0855	0.0074	0.0633
	30	2	0.0697	0.0051	0.0463	0.1124	0.0127	0.0802	0.0859	0.0075	0.0570

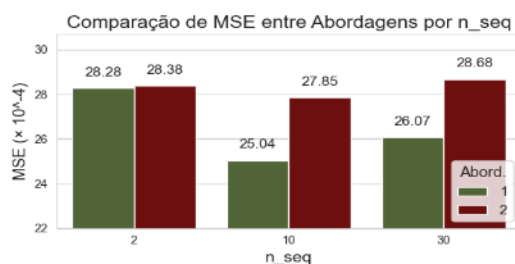


Figura 6. Comparação entre abordagens 1 (com ID sensor) e 2 (sem ID sensor).

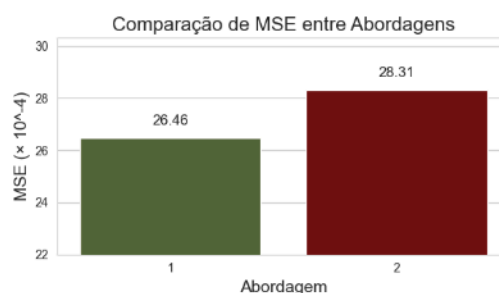


Figura 7. Comparação geral entre abordagens 1 e 2.

5. Conclusões e Trabalhos Futuros

Os resultados mostram que, em média, a informação de qual via o veículo foi identificado traz um efeito positivo na predição de próximo sensor externo, com um ganho médio geral do MSE de 6,5% em relação aos mesmos registros desconsiderando a informação de via da rua. Com isso conclui-se que, mesmo em um contexto de comportamento anômalo do motorista, a informação da via da rua, aqui estudada, traz pistas de para onde o veículo está indo, favorecendo a predição de próximo sensor externo.

Em relação à proposta da nova arquitetura TLE, não teve destaque amplo, mas atingiu resultado melhor na média geral do MSE para $n_{seq} = 2$, ou seja, 2 localizações de entrada para predizer a 3ª geolocalização. Mesmo com uma vantagem pequena (0,74% em relação à LSTM e 2,3% em relação à Transformer), mostra um potencial importante, valendo novos empregos e análises dessa arquitetura, pois, no contexto da predição de próximo sensor externo para veículos roubados, quanto menor a quantidade de localizações históricas para predizer a próxima, melhor para recuperar o veículo, minimizando as chances de uso do veículo roubado para cometer outros crimes subsequentes.

Estudos futuros incluindo dados de horários da identificação do veículo podem auxiliar ainda mais a predição. Assim como utilizar a nova proposta de arquitetura TLE em contextos diferentes de predição de sequências podem gerar resultados interessantes.

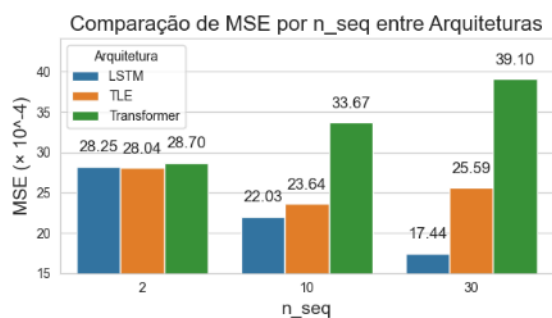


Figura 8. Comparação por sequências entre arquiteturas LSTM, TLE e Transformer.

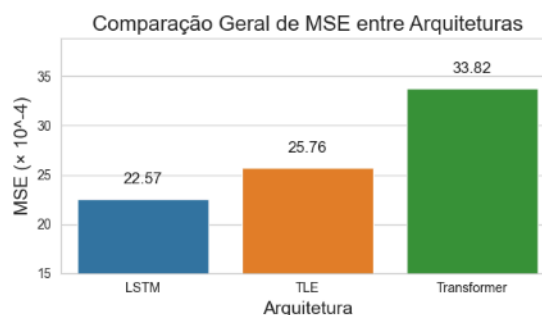


Figura 9. Comparação geral entre arquiteturas LSTM, TLE e Transformer.

Agradecimentos

Os autores agradecem à FAPESP, ao CNPq, ao Laboratório LATITUDE, ao CNI e à UnB.

Referências

- Almeida, V. G., Silva, T. R., and Silva, F. A. (2022). Se for, vá na paz: Construindo rotas seguras para veículos coletivos urbanos. In *Anais do XL Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 140–153. SBC.
- Bae, K., Lee, S., and Lee, W. (2022). Transformer networks for trajectory classification. In *2022 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pages 331–333. IEEE.
- Botchkarev, A. (2018). Evaluating performance of regression machine learning models using multiple error metrics in azure machine learning studio. Available at SSRN 3177507.
- Brito, M., Martins, B., Santos, C., Medeiros, I., Araújo, F., Seruffo, M., Oliveira, H., Cerqueira, E., and Rosário, D. (2023). Personalized experience-aware multi-criteria route selection for smart mobility. In *Anais do XLI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 504–517. SBC.
- Brownlee, J. (2021). Regression metrics for machine learning. <https://machinelearningmastery.com/regression-metrics-for-machine-learning/>. [On-line; Acesso em: 12-01-2024].
- Capanema, C. G. S., Silva, F. A., and Silva, T. R. d. M. B. (2020). Mfa-rnn: Uma rede neural recorrente para predição de próximo local de visita com base em dados esparsos. In *Anais do XXXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 127–140. SBC.
- Chen, X., Zhang, H., Zhao, F., Cai, Y., Wang, H., and Ye, Q. (2022). Vehicle trajectory prediction based on intention-aware non-autoregressive transformer with multi-attention learning for internet of vehicles. *IEEE Transactions on Instrumentation and Measurement*, 71:1–12.
- Cruz, L. A., Coelho da Silva, T. L., Magalhães, R. P., Melo, W. C. D., Cordeiro, M., de Macedo, J. A. F., and Zeitouni, K. (2022). Modeling trajectories obtained from external sensors for location prediction via nlp approaches. *Sensors*, 22(19):7475.
- de Souza, A. M., Braun, T., Botega, L. C., Villas, L. A., and Loureiro, A. A. (2019). Safe and sound: Driver safety-aware vehicle re-routing based on spatiotemporal information. *IEEE Transactions on Intelligent Transportation Systems*, 21(9):3973–3989.

- de Souza, A. M. and Villas, L. A. (2020). Vem tranquilo: Rotas eficientes baseado na dinâmica urbana futura com deep learning e computação de borda. In *Anais do XXXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 351–364. SBC.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Gaiduchenko, N. E., Gritsyk, P. A., and Malashko, Y. I. (2020). Multi-step ballistic vehicle trajectory forecasting using deep learning models. In *2020 International Conference Engineering and Telecommunication (En&T)*, pages 1–6. IEEE.
- Haviland, C. V. and Wiseman, H. (1974). Criminals who drive. In *Proceedings: American Association for Automotive Medicine Annual Conference*, volume 18, pages 432–439. Association for the Advancement of Automotive Medicine.
- Hendrycks, D. and Gimpel, K. (2016). Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*.
- Hu, H., Wang, Q., Du, L., Lu, Z., and Gao, Z. (2022). Vehicle trajectory prediction considering aleatoric uncertainty. *Knowledge-Based Systems*, 255:109617.
- Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 15(4):531–538.
- Karimzadeh, M., Aebi, R., de Souza, A. M., Zhao, Z., Braun, T., Sargento, S., and Villas, L. (2021). Reinforcement learning-designed lstm for trajectory and traffic flow prediction. In *2021 IEEE wireless communications and networking conference (WCNC)*, pages 1–6. IEEE.
- Ladeira, L. Z., de Souza, A. M., Rocha Filho, G. P., Silva, T. H., Sanches, M. F., and Villas, L. A. (2019a). Martini: Towards a mobile and variable time window identification for spatio-temporal data. In *2019 IEEE Latin-American Conference on Communications (LATINCOM)*, pages 1–6. IEEE.
- Ladeira, L. Z., de Souza, A. M., Rocha Filho, G. P., Silva, T. H., and Villas, L. A. (2019b). Serviço de sugestão de rotas seguras para veículos. In *Anais do XXXVII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 608–621. SBC.
- Ladeira, L. Z., de Souza, A. M., Silva, T. H., Rocha Filho, G. P., Peixoto, M. L. M., and Villas, L. A. (2020). Cerva: Roteamento contextual para veículos com risco espaço-temporal. In *Anais do XXXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 379–392. SBC.
- Neto, J. S. D. S., Da Silva, T. L. C., Cruz, L. A., de Lira, V. M., de Macêdo, J. A. F., Magalhaes, R. P., and Peres, L. G. (2021). Predicting the next location for trajectories from stolen vehicles. In *2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 452–456. IEEE.
- Sun, R. (2019). Optimization for deep learning: theory and algorithms. *arXiv preprint arXiv:1912.08957*.
- Tsiligkaridis, A., Zhang, J., Taguchi, H., and Nikovski, D. (2020). Personalized destination prediction using transformers in a contextless data setting. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7. IEEE.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Xu, Y., Wang, Y., and Peeta, S. (2023). Leveraging transformer model to predict vehicle trajectories in congested urban traffic. *Transportation research record*, 2677(2):898–909.