

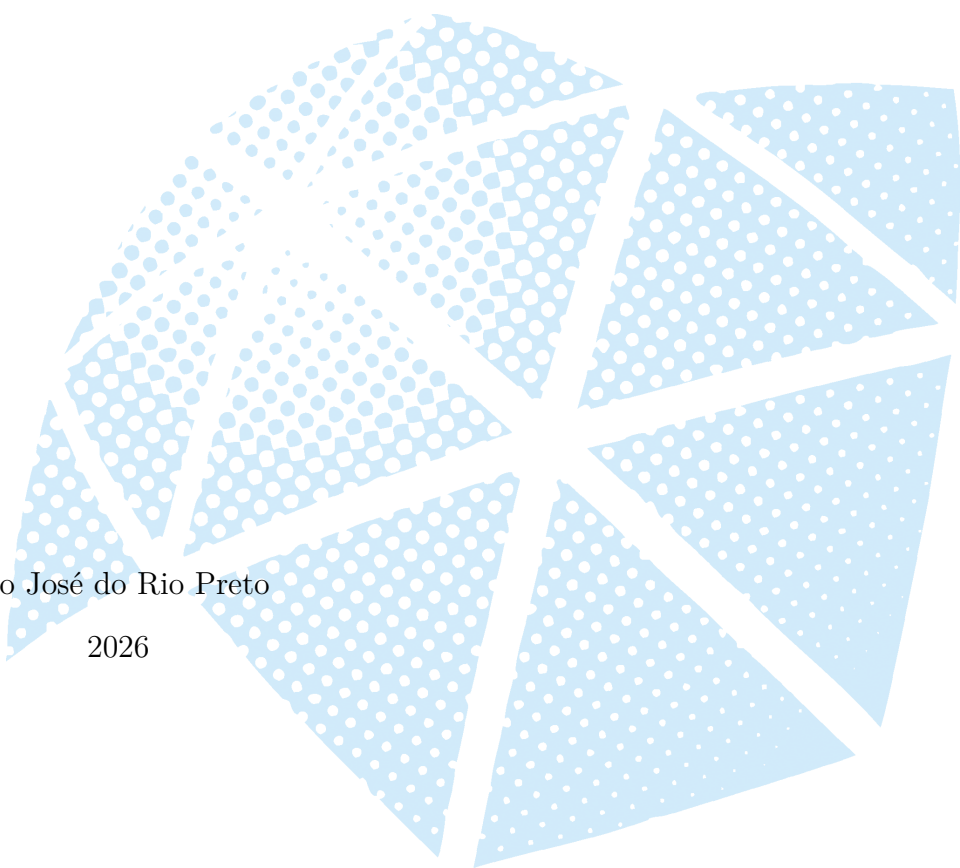
UNIVERSIDADE ESTADUAL PAULISTA - UNESP
Instituto de Biociências, Letras e Ciências Exatas - Câmpus de São José
do Rio Preto

BRUNO BELORTE

Modelagem Espaço-Temporal e Dinâmica do Roubo de Cargas no
Estado de São Paulo

São José do Rio Preto

2026



Bruno Belorte

**Modelagem Espaço-Temporal e Dinâmica do Roubo de Cargas no
Estado de São Paulo**

Tese apresentada à Universidade Estadual Paulista (UNESP), Instituto de Biociências, Letras e Ciências Exatas, São José do Rio Preto, para obtenção do título de Doutor em Matemática.

Área de Concentração: Matemática Computacional

Orientador: Prof. Dr. Wallace Correa de Oliveira Casaca

São José do Rio Preto

2026

B452m Belorte, Bruno
Modelagem Espaço-Temporal e Dinâmica do Roubo de Cargas no Estado de São Paulo / Bruno Belorte. -- São José do Rio Preto, 2026
117 p. : tabs., mapas

Tese (doutorado) - Universidade Estadual Paulista (UNESP), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto
Orientador: Wallace Correa de Oliveira Casaca

1. Redes neurais (Computação). 2. Análise espacial (Estatística). 3. Criminologia. 4. Segurança Pública. 5. Sociologia jurídica. I. Título.

Bruno Belorte

Modelagem Espaço-Temporal e Dinâmica do Roubo de Cargas no Estado de São Paulo

Tese apresentada à Universidade Estadual Paulista (UNESP), Instituto de Biociências, Letras e Ciências Exatas, São José do Rio Preto, para obtenção do título de Doutor em Matemática.

Área de Concentração: Matemática Computacional

Data da defesa: 12/06/2026

Banca Examinadora:

Prof. Dr. Wallace Correa de Oliveira Casaca
UNESP - Instituto de Biociências, Letras e Ciências Exatas - Campus de São José do Rio Preto

Prof. Dr. Lucas Correia Ribas
UNESP - Instituto de Biociências, Letras e Ciências Exatas - Campus de São José do Rio Preto

Prof. Dr. Diego de Castro Rodrigues
Departamento de Matemática Aplicada e Estatística - Universidade de São Paulo

Prof. Dr. Guilherme Freire Roberto
UNESP - Instituto de Biociências, Letras e Ciências Exatas - Campus de São José do Rio Preto

Prof. Dr. Rogério Galante Negri
Departamento de Engenharia Ambiental - UNESP - Campus de São José dos Campos

À minha família, amigos e orientadores.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, à qual agradeço.

Resumo

O roubo de cargas é um dos principais desafios para a segurança pública e a logística no Brasil, gerando altos impactos econômicos. Esta tese investiga a dinâmica desse fenômeno no Estado de São Paulo e propõe uma transição metodológica da estatística espacial clássica para o Aprendizado Profundo em Grafos Espaço-Temporais. Os dados foram estruturados por meio do processo de Descoberta de Conhecimento em Bases de Dados (*Knowledge Discovery in Databases* — KDD) para identificar pontos críticos e corredores rodo-logísticos. Para modelar a propagação do risco pelas rodovias, implementou-se a arquitetura neural A3T-GCN (*Attention Temporal Graph Convolutional Network*). Para corrigir o forte desbalanceamento geográfico dos dados, concentrados na Macrometrópole, desenvolveu-se uma estratégia de Aprendizado Sensível ao Custo (*Cost-Sensitive Learning*) com função de perda ponderada. Os resultados matemáticos demonstram que a rede corrigiu o viés de concentração da Capital e mapeou a difusão do risco criminal em direção ao interior. Sob a ótica da criminologia crítica, a modelagem prova que o roubo de cargas funciona como uma atividade empresarial estável que utiliza a infraestrutura do capital formal. Conclui-se, a partir dos testes do grafo, que intervenções focadas nos nós centralizadores de fluxo (centros logísticos) possuem maior impacto estatístico do que o patrulhamento linear de vias ostensivas, direcionando as estratégias de controle para o estrangulamento financeiro das organizações criminosas.

Palavras-chave: Redes Neurais (Computação). Análise Espacial (Estatística). Criminologia. Segurança Pública. Sociologia Jurídica.

Abstract

Cargo theft is one of the primary challenges to public security and logistics in Brazil, causing severe economic impacts. This thesis investigates the dynamics of this phenomenon in the State of São Paulo, proposing a methodological transition from classical spatial statistics to Spatiotemporal Graph Deep Learning. The data were structured using the Knowledge Discovery in Databases (KDD) process to identify hotspots and routing corridors. To model risk propagation across highways, an A3T-GCN (*Attention Temporal Graph Convolutional Network*) neural architecture was implemented. To address the severe geographical data imbalance concentrated in the Macrometropolis, a Cost-Sensitive Learning strategy with a weighted loss function was developed. The mathematical results demonstrate that the network corrected the Capital's concentration bias and mapped the diffusion of criminal risk toward the state's interior. From the perspective of critical criminology, the model proves that cargo theft operates as a stable corporate activity that leverages formal capital infrastructure. Based on the graph simulation tests, it is concluded that interventions focused on flow-concentrating nodes (logistical hubs) have a greater statistical impact than the linear patrolling of ostensive highways, shifting control strategies toward the financial strangulation of organized crime.

Keywords: Neural Networks (Computer Science). Spatial Analysis (Statistics). Criminology. Public Security. Legal Sociology.

Lista de ilustrações

Figura 1 – Fluxo de trabalho analítico baseado em KDD para a análise espacial do roubo de cargas.	37
Figura 2 – Série temporal de ocorrências de roubo de carga: dados processados <i>versus</i> estatísticas oficiais da SSP-SP.	42
Figura 3 – KDE longitudinal da intensidade de roubo de cargas no estado de São Paulo (2012–2022). <i>Hotspots</i> demonstram notável estabilidade espacial ao longo dos principais corredores logísticos. . .	45
Figura 4 – Distribuição horária dos roubos de carga ao longo dos anos (2012–2022).	47
Figura 5 – Distribuição horária dos roubos de carga por dia da semana (2012–2022).	47
Figura 6 – Distribuição horária das ocorrências agregada por mês (2012–2022). .	48
Figura 7 – Agrupamentos de roubos de carga identificados pelo DBSCAN. . . .	49
Figura 8 – Mapas de agrupamentos territoriais LISA: Alto-Alto (vermelho) indica focos críticos regionais; Baixo-Baixo (azul) representa zonas protegidas.	52
Figura 9 – Aderência da previsão com uma única variável (GRU). O modelo acompanha a direção geral dos dados, mas apresenta um atraso temporal (deslocamento) e tem dificuldade para alcançar os valores dos picos mais altos.	67
Figura 10 – Previsão de longo prazo para 2023 com uma única variável. Os valores previstos se aproximam rapidamente da média histórica devido à falta de outras variáveis.	68
Figura 11 – Predição do modelo RNN Multivariado. A inclusão da matriz espacial permitiu que o modelo acompanhasse com maior fidelidade as flutuações reais da série.	71
Figura 12 – Comparativo visual de aderência: Univariada vs. Multivariada. A variável geográfica destrava a capacidade da rede de modelar a volatilidade aguda.	72

Figura 13 – Processamento residual. A rede tenta rastrear os desvios estocásticos (eixo Y) ao longo do tempo (eixo X), enfrentando alta dispersão informacional.	74
Figura 14 – Desempenho do preditor híbrido (SARIMA + GRU). A linha azul demonstra a instabilidade causada pela tentativa de prever resíduos isolados.	75
Figura 15 – Projeção inicial para a Capital (SP). Sem o ajuste de pesos, a previsão resulta em uma linha quase plana devido à suavização excessiva.	88
Figura 16 – Projeção para a Capital após a transformação logarítmica. A curva acompanha as oscilações, mas permanece em escala comprimida devido à falta de pesos no treinamento.	89
Figura 17 – Desempenho final sob Perda Ponderada. O modelo corrige a trajetória e passa a acompanhar as variações reais da Capital. .	91
Figura 18 – Cartografia Comparativa (2021-2022). Os mapas foram construídos com o auxílio da biblioteca geobr , que fornece os limites territoriais oficiais do IBGE. Na definição das cores de risco, aplicou-se a transformação logarítmica $f(x) = \ln(1 + x)$. Esse ajuste é necessário para evitar que o alto volume de ocorrências da Região Metropolitana oculte os padrões de criminalidade das cidades do interior, permitindo a visualização equilibrada de todo o território. O uso das cores em escala logarítmica revela claramente os corredores espaciais de risco viajar sobre o asfalto.	93
Figura 19 – Evolução Espaço-Temporal e Extrapolação (2021 a 2023). A estagnação territorial do mapa comprova a natureza fixa e enraizada da estrutura corporativa do ilícito.	96

Lista de quadros

Quadro 1 – Resumo estatístico do processo de limpeza e filtragem dos dados.	34
Quadro 2 – Atributos dos registros policiais extraídos e enriquecidos.	35
Quadro 3 – Métricas de desempenho do agrupamento espacial via DBSCAN (2012–2022).	50
Quadro 4 – Estatísticas descritivas da distância geodésica até a base policial rodoviária mais próxima (2012–2022).	55
Quadro 5 – Resultados do Teste Aumentado de Dickey-Fuller (Série Bruta).	61
Quadro 6 – Arquitetura Sequencial das Redes Recorrentes (Univariada e Multivariada).	64
Quadro 7 – Dinâmica das Funções de Ativação e Implicações Algorítmicas.	65
Quadro 8 – Estrutura e parametrização geométrica do tensor de entrada univariado.	66
Quadro 9 – Estrutura e parametrização geométrica do tensor de entrada multivariado.	69
Quadro 10 – Performance preditiva (2012–2022): Univariada vs. Multivariada.	72
Quadro 11 – Diagnóstico de Desempenho por Arquitetura Preditiva.	74
Quadro 12 – Decomposição das Dimensões do Tensor Espaço-Temporal $X \in \mathbb{R}^{N \times F \times T}$	84
Quadro 13 – Dependências Tecnológicas e Bibliotecas de Suporte (STGNN).	86
Quadro 14 – Especificações da Arquitetura Neural Espaço-Temporal (A3T- GCN).	87
Quadro 15 – Comparativo de desempenho entre polos logísticos e cidades do interior.	92
Quadro 16 – Parâmetros de Construção da Cartografia Preditiva.	95

Sumário

1	INTRODUÇÃO	19
1.1	A Dinâmica Estrutural: Desregulamentação Econômica e a Organização do Crime	19
1.2	A Circulação de Armas e o Suprimento Logístico do Crime .	21
1.3	A Geopolítica do Controle Criminal e o Imperialismo Legal .	22
1.4	As Lacunas Metodológicas e a Proposta do Estudo	23
1.5	Hipótese de Trabalho	24
1.6	Objetivos e Contribuições	24
1.7	Estrutura do Trabalho	25
2	REFERENCIAL TEÓRICO	27
2.1	Fundamentação Crítico-Estrutural e Geopolítica do Delito . .	27
2.2	Governança de Dados Criminais e o Desafio Logístico	28
2.3	Criminologia Espacial e Agrupamento Baseado em Densidade	29
2.4	Ruptura Preditiva: De Séries Temporais para Redes em Grafos	30
3	ANÁLISE ESPACIAL E CLUSTERIZAÇÃO	33
3.1	Engenharia e Limpeza de Dados	33
3.1.1	Dados de Infraestrutura Policial e Junção Espacial	34
3.2	Metodologia Analítica	36
3.2.1	Transformação de Dados e Cálculo Geodésico	37
3.2.2	Mineração Espacial: KDE, DBSCAN e LISA	38
3.2.2.1	Operação e Espaço de Fase do Indicador LISA	40
3.3	Resultados e Discussão	41
3.3.1	Curadoria de Dados e Dinâmica Temporal	42
3.3.2	Persistência Espaço-Temporal e Concentração Regional	43
3.3.2.1	Análise das Superfícies de Intensidade Rodoviária	43
3.3.2.2	Concentração Temporal e Testes de Invariância	44
3.3.2.3	Mapeamento de Agrupamentos e Identificação de Zonas Críticas	46

3.3.3	Verificação Estatística da Dependência Espacial	51
3.3.4	Proximidade Geográfica da Infraestrutura Policial	53
3.4	Considerações Finais do Capítulo	54
4	MODELAGEM PREDITIVA DE BASE: ESTATÍSTICA CLÁSSICA VS. REDES NEURAIIS	57
4.1	Preparação dos Dados e das Séries Temporais	57
4.1.1	Agrupamento Mensal e Estruturação Cronológica	58
4.1.2	Padronização e Escalonamento das Variáveis	58
4.1.3	Estruturação dos Dados por Janela Deslizante	59
4.2	Modelagem Estatística Clássica: O Componente Linear (SARIMA)	60
4.2.1	Testes de Estacionariedade e Diferenciação dos Dados	60
4.3	Abordagens Neurais Não Lineares	62
4.3.1	Topologia da Rede e Dinâmica de Ativação	63
4.3.2	Arquitetura Univariada e o Efeito de Suavização	65
4.3.3	Arquitetura Multivariada: O Espaço como Regulador	69
4.4	Avaliação do Modelo Híbrido (SARIMA-GRU)	71
4.5	Consolidação e Diagnóstico Preditivo	73
4.6	Considerações Finais do Capítulo	76
5	PREDIÇÃO AVANÇADA COM REDES NEURAIIS EM GRAFOS ESPAÇO-TEMPORAIS	79
5.1	Modelagem Relacional do Território	79
5.2	Construção da Malha Logística: A Matriz de Adjacência . . .	80
5.3	A Arquitetura A3T-GCN	82
5.3.1	A Estrutura do Tensor Espaço-Temporal	83
5.3.2	Implementação e Dependências Tecnológicas	86
5.3.3	Especificações da Arquitetura e Hiperparâmetros	87
5.4	Aprendizado Sensível ao Custo (<i>Cost-Sensitive Learning</i>) . .	87
5.4.1	Transformação Logarítmica da Variância	88
5.4.2	A Função de Perda Ponderada (<i>Weighted Loss</i>)	89
5.4.3	Deslocamento de Viés e Limitações Geométricas das Arestas	91

5.5	Resultados e Cartografia Preditiva: A Topologia do Risco . . .	92
5.5.1	Análise de Mancha Criminal: O Biênio de Validação (2021–2022) . . .	93
5.5.1.1	Interpretação Topológica e o Impacto da Sensibilidade ao Custo	94
5.5.1.2	Resumo Metodológico da Construção Cartográfica	95
5.5.2	Projeção de Cenários e a Inércia Logística (Extrapolação 2023) . . .	95
5.5.3	Reorientação do Policiamento: Do Modelo Reativo ao Controle Ter- ritorial	97
5.6	Análise Crítica: A Conectividade do Grafo e a Estrutura do Capital	97
5.6.1	Implicações Estratégicas para a Justiça Criminal	98
5.7	Considerações Finais do Capítulo: A Topologia do Risco e a Criminologia em Redes	98
6	CONCLUSÃO GERAL E TRABALHOS FUTUROS	101
6.1	Síntese dos Achados Técnicos e Contribuições do Trabalho .	101
6.2	Implicações Teóricas e a Homologia Algorítmica Estrutural .	102
6.3	O Legado Metodológico da Tese	103
6.4	Limitações e Trabalhos Futuros	104
6.4.1	Expansão Multivariada do Tensor e Variáveis Ambientais	104
6.4.2	Refinamento Topológico com Geoprocessamento Avançado	105
6.4.3	Modelagem por Grafos Heterogêneos Relacionais	105
	REFERÊNCIAS	107

1 Introdução

O roubo de cargas é um fenômeno criminal complexo e de difícil modelagem que afeta diretamente a estabilidade das cadeias globais de suprimento (World Bank, 2023; U.S. Department of State OSAC, 2025). Longe de ser um problema periférico de segurança pública, esse tipo de crime tornou-se uma atividade altamente rentável e enraizada, sobretudo nas economias em desenvolvimento (UNODC, 2023; Overhaul, 2024). Na América Latina, o Brasil apresenta as maiores taxas de incidência do subcontinente (U.S. Department of State OSAC, 2025). A matriz de transportes nacional dita essa vulnerabilidade, uma vez que o país depende do modal rodoviário para escoar mais de 65% de suas mercadorias (International Council on Clean Transportation, 2021).

O Estado de São Paulo está no centro dessa dinâmica. O estado responde por cerca de um terço do Produto Interno Bruto (PIB) brasileiro (GANDHI et al., 2017) e abriga o maior polo industrial e portuário da América do Sul. Essa concentração de riqueza gera um paradoxo: as mesmas rodovias que viabilizam a economia legal servem como eixos para a circulação do capital ilícito. Como resultado, a região Sudeste concentrou 80% dos incidentes nacionais no ano de 2023, com São Paulo liderando as estatísticas isoladamente ao registrar 43% das ocorrências em 2024 (Overhaul, 2023; Overhaul, 2024).

1.1 A Dinâmica Estrutural: Desregulamentação Econômica e a Organização do Crime

Para compreender a dimensão do roubo de cargas paulista, é necessário superar as abordagens tradicionais da criminologia, que costumam tratar o crime organizado como um elemento externo à sociedade. Na realidade, os grandes grupos criminosos expandiram suas atividades em sincronia com a flexibilização dos mercados e a redução do controle estatal sobre fluxos comerciais (PICCELLI, 2013).

Essa configuração reflete o papel ocupado pelo capitalismo periférico na economia global. Conforme a teoria da dependência de Marini (1973), o modelo econômico dependente recorre à exploração intensa das estruturas de circulação para compensar as perdas de valor no mercado internacional. Sob essa ótica, a pilhagem sistemática nas rodovias atua como uma força de expropriação que compõe o próprio metabolismo do mercado local. O lucro gerado pelo roubo de cargas é reintroduzido e absorvido pelas redes do sistema financeiro formal, demonstrando que o crime organizado funciona de forma integrada ao processo de acumulação de riquezas.

O roubo de cargas evoluiu de saques isolados para uma atividade empresarial altamente estruturada (MINGARDI, 2001). A movimentação de recursos nas rodovias passou a exigir mecanismos complexos de engenharia financeira e jurídica. Investigações conduzidas em 2025 revelaram que fundos de investimento sediados nos centros financeiros da capital paulista absorvem e dissimulam os lucros obtidos por essas organizações (InfoMoney, 2025). Esse capital ilícito tem sido utilizado para controlar setores econômicos desregulamentados, incluindo o refino de combustíveis e fraudes em concessões de transporte público na Macrometrópole Paulista (Agência Pública, 2026; JOTA, 2023).

Nesse cenário de integração entre mercados legais e ilegais (SOUZA; MIRANDA, 2023), discursos políticos baseados na flexibilização radical das regras de mercado e na diminuição da fiscalização do Estado terminam por fragilizar os mecanismos de prevenção à lavagem de dinheiro. O caso do Grupo Refit exemplifica essa vulnerabilidade institucional, em que estruturas corporativas foram utilizadas para consolidar uma sonegação fiscal de R\$ 26 bilhões e ocultar ativos por meio de empresas estrangeiras operadas por financiadores de movimentos políticos desregulamentadores (Revista Oeste, 2025; IstoÉ, 2025).

De acordo com a análise de Safatle (2012), pautas moralistas focadas exclusivamente no combate à corrupção estatal costumam operar como ferramentas de despolitização. Ao direcionar a crítica social apenas contra a máquina pública, essa retórica protege as relações estruturais existentes entre grandes corporações sonegadas e os métodos criminosos de controle de mercado.

1.2 A Circulação de Armas e o Suprimento Logístico do Crime

Para garantir a execução dos roubos de carga nas rodovias, as organizações criminosas dependem do acesso contínuo a armamentos e munições. Historicamente, essa demanda era suprida pelo contrabando internacional de armas fabricadas no exterior (DEMORI; OLIVEIRA, 2019). Contudo, a partir de 2019, as redes de suprimento modificaram sua estratégia de abastecimento, aproveitando-se de alterações na legislação nacional de controle de armas.

Entre 2019 e 2022, as regras para a aquisição de armas por Caçadores, Atiradores e Colecionadores (CACs) foram flexibilizadas no Brasil, permitindo a compra de fuzis semiautomáticos e grandes volumes de munição por civis (Fórum Brasileiro de Segurança Pública, 2025). O crime organizado utilizou essa abertura jurídica para recrutar intermediários sem antecedentes criminais, que adquiriam os armamentos legalmente no mercado comercial urbano. Posteriormente, essas armas de calibre militar eram desviadas para as facções por meio do registro de falsos boletins de ocorrência de furto, um mecanismo de abastecimento documentado por institutos de pesquisa em segurança pública (PAZ, 2023a; PAZ, 2023b).

A redução da fiscalização sobre a venda de munições contribuiu para o fortalecimento da capacidade de força dos grupos criminosos (Fórum Brasileiro de Segurança Pública, 2025). Relatórios técnicos apontam que as falhas no rastreamento de projéteis e a falta de transparência institucional sobre o destino de lotes bélicos dificultam a elucidação de crimes cometidos nas rodovias (The Intercept Brasil, 2026; NEVES, 2026a). O afrouxamento dos critérios de controle de produtos controlados facilitou a circulação de armas de fogo de grosso calibre no Estado de São Paulo, afetando diretamente a segurança dos motoristas de transporte de carga (FBSP, 2024). Essa dinâmica impulsionou o faturamento do mercado de armas e ampliou a capacidade operacional e a expansão territorial das facções criminosas (PAZ, 2023b).

1.3 A Geopolítica do Controle Criminal e o Imperialismo Legal

A expansão financeira do crime organizado gerou reações geopolíticas internacionais. Em 2026, o governo dos Estados Unidos classificou as principais facções criminosas brasileiras como Organizações Terroristas Estrangeiras. Esse ato institucional busca estender a jurisdição norte-americana sobre o território nacional, pressionando pela adoção de políticas de segurança pública baseadas em modelos militarizados e em regimes de exceção (BRASIL DE FATO, 2026).

A partir das análises de Fernandes (2024), a criminologia crítica identifica nessa estratégia um exercício de imperialismo legal. O enquadramento do crime organizado como narcoterrorismo distorce a natureza essencialmente econômica do fenômeno. Conforme aponta Neves (2026b), esse rótulo funciona como uma ferramenta discursiva que oculta a lavagem de dinheiro realizada nos centros financeiros globais e justifica intervenções políticas em Estados periféricos.

Na prática, essa classificação permite a aplicação de sanções internacionais contra instituições bancárias e empresas logísticas brasileiras sob a alegação de apoio ao terrorismo (U.S. Department of the Treasury, 2024; Crowell and Moring LLP, 2025). Ao mesmo tempo, a estratégia atua como mecanismo de contenção geopolítica, criando restrições para investimentos de potências concorrentes, como a China, na infraestrutura de transportes da América do Sul (New Eastern Outlook, 2026; Atlantic Council, 2026).

O sistema de justiça criminal reproduz essas contradições ao concentrar a força punitiva na base da pirâmide operacional, penalizando motoristas e pequenos transportadores enquanto preserva as redes de circulação do grande capital. Ao aplicar a economia política da pena ao cenário nacional, Alemany (2019) demonstra que, no capitalismo dependente, o aparato penal atua prioritariamente na gestão e contenção das populações vulneráveis. A repressão policial ostensiva nas estradas atua como controle de território, garantindo a regularidade dos fluxos comerciais sem interferir nas estruturas financeiras que comandam os mercados ilícitos.

1.4 As Lacunas Metodológicas e a Proposta do Estudo

Diante do impacto econômico do roubo de cargas, a segurança pública não pode depender apenas de ações reativas ou de análises estatísticas simplificadas. O fenômeno possui natureza relacional, em que o crime organizado funciona como uma rede de logística e distribuição. O uso de técnicas de Ciência de Dados permite o desenvolvimento de modelos capazes de identificar essas interações espaciais complexas.

Contudo, a literatura especializada apresenta limitações metodológicas. A maioria dos modelos de aprendizado de máquina e de análise de séries temporais assume que as unidades espaciais são estatisticamente independentes, tratando os municípios de forma isolada (ANSELIN, 1988; WANG et al., 2017). Essa abordagem ignora a conectividade territorial imposta pelas rodovias e não consegue prever como o risco de crimes se espalha entre cidades interligadas.

Embora a estatística espacial clássica tenha avançado na identificação de pontos críticos e padrões de autocorrelação (CHAINEDY; TOMPSON; UHLIG, 2008), suas técnicas possuem caráter estático. Os modelos tradicionais falham em analisar o espaço e o tempo de forma conjunta, o que prejudica a identificação de dinâmicas não lineares em redes de transporte de grande escala (XIA et al., 2021; WANG et al., 2024b). Além disso, a literatura dedica pouca atenção ao desbalanceamento geográfico dos dados, concentrados nas regiões metropolitanas, o que reduz a capacidade de previsão dos algoritmos em áreas do interior.

Somado às limitações metodológicas, é necessário destacar que a governança dos dados criminais no Estado de São Paulo envolve fatores políticos e tecnológicos. A infraestrutura de armazenamento e tratamento dos registros policiais está vinculada a contratos de serviços em nuvem com corporações estrangeiras, como a Microsoft, por meio de parcerias coordenadas pela Prodesp (PRODESP, 2024). Esse processo iniciou-se em 2014 com o sistema de inteligência preditiva *Detecta* (MICROSOFT BRASIL, 2014) e evoluiu para o uso de ferramentas automatizadas na análise de Boletins de Ocorrência. Esse arranjo transfere dados sensíveis de segurança para servidores privados externos, gerando dependência tecnológica de plataformas proprietárias terceirizadas.

Para superar essas barreiras, este estudo propõe uma abordagem baseada em grafos, tratando o território de forma conectada. O roubo de cargas em São Paulo é investigado por meio do processo de Descoberta de Conhecimento em Bases de Dados (*Knowledge Discovery in Databases* — KDD) (ELMASRI; NAVATHE, 2011), utilizando os dados públicos da Secretaria de Segurança Pública (SSP-SP) de 2012 a 2022 para garantir a transparência algorítmica. O modelo implementa Redes Neurais em Grafos (*Graph Neural Networks* — GNNs) para capturar o ritmo temporal e a propagação do risco ao longo da malha viária paulista.

1.5 Hipótese de Trabalho

A hipótese central deste estudo é que o roubo de cargas rodoviárias constitui um fenômeno estrutural de rede cuja dispersão é determinada pela geometria das rotas de escoamento do capital. Consequentemente, o risco criminal não responde de forma eficiente a mecanismos lineares de dissuasão — como a proximidade física de postos policiais —, mas propaga-se de forma previsível pelos eixos da malha viária, sendo passível de rastreamento e antecipação por meio de modelos não lineares espaço-temporais baseados em grafos.

1.6 Objetivos e Contribuições

O objetivo principal desta tese é modelar e analisar a dinâmica espaço-temporal do roubo de cargas em bases matemáticas. O estudo busca fornecer subsídios técnicos ao Ministério Público para o aprimoramento de estratégias de investigação, deslocando o foco dos executores diretos para a desarticulação das estruturas financeiras das organizações criminosas. As contribuições do trabalho são:

- Organização, limpeza e georreferenciamento de uma série histórica (2012–2022) com mais de 36.000 boletins de ocorrência, construindo um processo de processamento auditável.

- Execução de análises geoestatísticas (KDE, DBSCAN) e comprovação da dependência espacial local por meio da técnica LISA (*Local Indicators of Spatial Association*) (ANSELIN, 1995).
- Implementação do modelo estatístico tradicional (SARIMA) e de Redes Neurais Recorrentes (RNN/GRU), demonstrando as limitações de previsões univariadas que excluem os dados geográficos.
- Aplicação da arquitetura neural A3T-GCN sobre uma matriz de adjacência viária, demonstrando matematicamente como o risco se desloca dos polos metropolitanos para o interior do estado devido à inércia logística (BAI et al., 2021).

1.7 Estrutura do Trabalho

O estudo está organizado em seis capítulos. O Capítulo 2 revisa o estado da arte da modelagem criminal e discute os fundamentos das teorias criminológicas. O Capítulo 3 detalha a coleta de dados, o processo de limpeza e a metodologia de mineração via KDD. O Capítulo 4 aborda a modelagem temporal de base, avaliando as limitações do modelo estatístico linear (SARIMA) e das redes recorrentes clássicas. O Capítulo 5 apresenta a arquitetura baseada em grafos espaço-temporais (A3T-GCN), detalhando sua calibração matemática e os resultados cartográficos. O Capítulo 6 apresenta as conclusões institucionais do estudo e aponta diretrizes para investigações futuras no campo da jurimetria preditiva.

2 Referencial Teórico

Este capítulo organiza a fundamentação teórica e o estado da arte da pesquisa em quatro eixos temáticos. Primeiro, discutimos a teoria do crime e sua relação com a estrutura social; segundo, abordamos a governança e os desafios dos dados logísticos; terceiro, examinamos os métodos de análise espacial baseados em densidade e, por fim, detalhamos a transição para modelos preditivos estruturados em Redes Neurais em Grafos (*Graph Neural Networks* — GNNs).

2.1 Fundamentação Crítico-Estrutural e Geopolítica do Delito

Para compreender o roubo de cargas como um fluxo corporativo inercial¹, é necessário transpor os limites abstratos da criminologia liberal e ancorar a análise na economia política da dependência. Embora a matriz epistemológica formulada por Marini (1973) determine que as economias periféricas sustentam-se sobre a superexploração da força de trabalho, as transformações contemporâneas na divisão internacional do trabalho — impulsionadas pela fragmentação das cadeias globais de valor e pela submissão do território à logística de escoamento rápido — aprofundaram esse padrão de subordinação econômica (LUCE, 2018; OSORIO, 2019). Nesse cenário de dependência reconfigurada, o crime corporativo nas rodovias não constitui um resíduo arcaico ou um desvio de rota, mas consolida-se como uma engrenagem que viabiliza a circulação e a realização de mercadorias em mercados secundários e altamente precarizados.

Alemaný (2019) atualiza a premissa clássica de Rusche e Kirchheimer (2004) para o capitalismo brasileiro, demonstrando que o sistema penal atua como um gestor da população excedente. Sob esta ótica, a repressão nas rodovias paulistas não busca apenas erradicar o crime: ela funciona como uma vigilância seletiva que

¹ O termo fluxo corporativo inercial refere-se à natureza do roubo de cargas como uma atividade integrada à dinâmica regular das cadeias de suprimento. Diferente de crimes esporádicos, tal modalidade opera com lógica empresarial e aproveita-se da inércia (estabilidade e recorrência) das rotas logísticas e dos fluxos financeiros, tornando-se uma constante estrutural do capitalismo periférico, conforme discutido por Alemany (2019).

disciplina o território e gerencia os ilegalismos necessários para manter a acumulação de capital na periferia do sistema.

Além desta gestão interna, pressões geopolíticas também moldam o aparato de controle ao crime organizado. Fernandes (2024) descreve o uso do imperialismo legal e do *lawfare* como ferramentas políticas para avançar interesses estrangeiros sob o pretexto de combater o crime transnacional. Nesse contexto, a classificação de facções brasileiras como organizações terroristas pelos Estados Unidos (BRASIL DE FATO, 2026) funciona como instrumento de pressão econômica e jurídica internacional.

A decisão de modelar o roubo de cargas como uma rede topológica baseada estritamente na malha rodoviária nacional responde a esse cenário de forma direta: o modelo não reduz o imperialismo por si só, mas retira dele seu principal vetor de entrada, que é a dependência de tecnologia estrangeira. Desenvolver ferramentas matemáticas independentes impede a necessidade de importação de *softwares* proprietários de monitoramento, que historicamente funcionam como meios de captura de dados sensíveis e de imposição de agendas de segurança militarizadas. A produção de inteligência estatística soberana quebra a assimetria informacional e garante a autonomia decisória do Estado periférico frente às intervenções do centro do capitalismo.

2.2 Governança de Dados Criminais e o Desafio Logístico

A pesquisa em segurança pública depende diretamente da qualidade, do acesso e da granularidade das informações. No roubo de cargas, há uma divisão clara entre estudos que utilizam fontes governamentais abertas e aqueles baseados em telemetria privada corporativa.

No Estado de São Paulo, a Secretaria de Segurança Pública (SSP-SP) constitui a principal fonte para investigações empíricas (JUSTUS et al., 2018; COSTA et al., 2024). Entretanto, a literatura nem sempre detalha os procedimentos de extração e tratamento desses dados, o que dificulta a reprodução exata das pesquisas (MELO et al., 2018; GARCIA-ZANABRIA et al., 2022). No âmbito federal, os pesquisadores recorrem a painéis agregados do Instituto de Pesquisa

Econômica Aplicada (IPEA) ou da Polícia Rodoviária Federal (PRF) (FILHO, 2023), que apresentam restrições de granularidade temporal e espacial.

Por outro lado, o uso de dados privados — como telemetria por GPS e identificação por radiofrequência (RFID) — permite analisar rotas específicas com alta precisão logística (REPOLHO et al., 2019; SONG et al., 2020). Estudos desse tipo são direcionados à otimização de frotas privadas no Brasil e no México (QUEIROZ et al., 2025; CEDILLO-CAMPOS; FLORES-FRANCO; COVARRUBIAS, 2024). Contudo, a disponibilidade dessas bases de dados é severamente restrita por acordos de confidencialidade comercial e regras de segredo industrial. Como essas informações constituem ativos proprietários de seguradoras e gerenciadoras de risco (Overhaul, 2023; Overhaul, 2024), o acesso é vedado à auditoria externa e à pesquisa acadêmica pública.

Esse regime de propriedade cria uma assimetria informacional no território: a previsão exata e a segurança dos fluxos tornam-se mercadorias de circulação restrita, deixando o aparato público dependente de registros estatísticos criminais. Para garantir a transparência e a replicabilidade, esta tese prioriza a soberania informacional ao utilizar exclusivamente os dados públicos abertos da SSP-SP abrangendo o período de 2012 a 2022 (DUARTE et al., 2025).

2.3 Criminologia Espacial e Agrupamento Baseado em Densidade

Identificar o risco logístico exige a superação de métodos geométricos tradicionais que tratam o espaço de forma plana e isotrópica, ignorando o traçado real das rodovias. A literatura tem migrado de centroides tradicionais para métodos de agrupamento baseados em densidade, que permitem identificar formas irregulares e variações locais de intensidade criminal (ESTER et al., 1996; CAMPELLO; MOULAVI; SANDER, 2013).

O algoritmo *K-Means*, por exemplo, mostra-se limitado para a análise de rodovias porque força os agrupamentos de dados a adotarem formatos circulares baseados na distância euclidiana clássica, gerando distorções ao ignorar a conec-

tividade das estradas (GARCIA et al., 2019). Para capturar a forma fluida e linear do crime de cargas, pesquisas recentes têm adotado o algoritmo DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) (SHI; PUN-CHENG, 2019), que rastreia ocorrências ao longo dos eixos viários sem exigir formatos geométricos fixos (AN et al., 2023). Em cenários de alta densidade urbana e eixos de conexão, extensões como o HDBSCAN e técnicas de fatoração de matrizes não negativas (NMF) têm auxiliado na identificação de padrões geográficos ocultos e subjacentes (GARCIA-ZANABRIA et al., 2020).

No entanto, a identificação visual de um aglomerado de ocorrências não é suficiente para confirmar uma estrutura criminal estável. É necessário aplicar testes estatísticos de associação espacial para validar a significância desses padrões. O indicador LISA (*Local Indicators of Spatial Association*) cumpre essa função ao provar matematicamente a existência de autocorrelação espacial local (ANSELIN, 1995; SCHARF, 2022). Por meio do LISA, demonstra-se se o índice de roubos de carga de um determinado município depende diretamente da situação de segurança e da incidência criminal das cidades vizinhas interconectadas pela malha de transporte (PEREZ-FERNANDEZ et al., 2025; SILVA, 2025).

2.4 Ruptura Preditiva: De Séries Temporais para Redes em Grafos

A transição da análise geoestatística exploratória para a modelagem matemática preditiva constitui o núcleo desta pesquisa. Historicamente, os modelos preditivos voltados à segurança pública dividiam-se em modelos puramente cronológicos ou abordagens de aprendizado de máquina supervisionado tradicional, apresentando limitações estruturais para dados dispostos em rede. Modelos como o SARIMA (*Seasonal Autoregressive Integrated Moving Average*) estabeleceram a base para prever volumes agregados a partir de componentes lineares de tendência e sazonalidade histórica (AKINYEMI et al., 2021). Com a expansão computacional, algoritmos baseados em árvores de decisão, como o *Random Forest*, passaram a incorporar variáveis exógenas para estimar o risco em sistemas de transporte (COSTA et al., 2024). Contudo, essas abordagens convencionais operam sob um ponto cego

metodológico: elas assumem a independência estatística das observações espaciais, tratando cada município de forma isolada e falhando em capturar a dinâmica de fluxo que rege o roubo de cargas.

Para superar as limitações do isolamento geográfico tradicional, a fronteira metodológica adota as Redes Neurais em Grafos (*Graph Neural Networks* — GNNs) (KIPF; WELLING, 2017), estrutura que permite formalizar o território como um arranjo relacional composto por nós e arestas (ROSHANKAR; KEYVANPOUR, 2023). A aplicação desses modelos ao tratamento do roubo de cargas exige a formulação de três abordagens matemáticas distintas e complementares ao longo do fluxo de processamento desta tese. O ponto de partida apoia-se na Modelagem Linear de Base (SARIMA), que representa a série temporal Y_t como um processo determinado por defasagens autorregressivas, diferenciações e médias móveis sazonais. Essa abordagem captura a inércia histórica macroeconômica global a partir da estrutura tradicional:

$$\Phi_P(B^s)\phi_p(B)\nabla^d\nabla_s^DY_t = \Theta_Q(B^s)\theta_q(B)Z_t$$

A evolução do fluxo metodológico exige o abandono dos pressupostos lineares rígidos por meio da Modelagem Neural Não Linear, implementada via Redes Neurais Recorrentes com Unidades Recorrentes Fechadas (*Gated Recurrent Unit* — GRU). Essa etapa intermediária processa as sequências cronológicas utilizando vetores de estado ocultos que ativam portas de atualização e de esquecimento. Esse mecanismo permite reter informações e dependências não lineares de longo prazo diretamente na série histórica de roubos, superando as limitações de modelos estatísticos que falham diante da volatilidade aguda dos dados criminais.

O pipeline atinge o seu nível avançado com a Modelagem Espaço-Temporal em Grafo, consolidada na arquitetura A3T-GCN (*Attention Temporal Graph Convolutional Network*). Este modelo integra Convoluções em Grafos (GCN) para capturar a topologia e a conectividade da malha rodoviária por meio de uma matriz de adjacência A , Unidades Recorrentes Fechadas (GRU) para extrair a dinâmica do tempo e mecanismos de atenção adaptativos. A função dessa arquitetura combinada é ponderar matematicamente a importância das defasagens temporais, mapeando de forma conjunta o tempo, o espaço e a infraestrutura viária para rastrear os

padrões de fuga, transbordo e vazamento do risco criminal à medida que este se difunde dos epicentros metropolitanos rumo ao interior do estado (BAI et al., 2021; GU et al., 2023).

3 Análise Espacial e Clusterização

Este capítulo investiga a dinâmica espacial do roubo de cargas no Estado de São Paulo. A partir da consolidação de uma ampla base de registros policiais, são aplicadas técnicas de estatística espacial para identificar padrões de concentração (*hotspots*) e corredores logísticos associados a elevado risco criminal.

3.1 Engenharia e Limpeza de Dados

Esta pesquisa utiliza os registros criminais abertos fornecidos pela Secretaria de Segurança Pública de São Paulo (2025). A base de dados bruta original compreende um total de 66.241 linhas de registros de ocorrências de roubos de carga. O primeiro passo do processamento consistiu na eliminação de duplicatas. Essa duplicidade ocorre porque o sistema policial gera linhas repetidas para um mesmo boletim de ocorrência, primeiro discriminando a ocorrência sem associar os veículos, depois gerando uma linha para cada veículo envolvido. Após a remoção das ocorrências gerais (sem veículos associados), a base foi reduzida para 37.457 ocorrências únicas, garantindo que cada linha represente um fato criminal distinto.

A partir desse conjunto, aplicaram-se filtros sequenciais de consistência cadastral e salvaguarda jurídica. Primeiro, removeram-se 169 registros cujos locais do crime foram identificados como residências, uma vez que a legislação oculta os endereços nesses casos para preservar a privacidade e a segurança das vítimas. Na sequência, identificou-se a ausência completa de dados de localização em 45 logradouros, cujas linhas vazias foram descartadas. Esse refino resultou em 37.243 ocorrências prontas para a etapa de verificação e correção de coordenadas espaciais. O Quadro 1 detalha a volumetria e o impacto de cada etapa de tratamento sobre a base original.

A etapa de georreferenciamento concentrou-se na correção e extração de coordenadas para os registros que apresentavam inconsistências nos campos de latitude e longitude originais. Para esse fim, submeteu-se um lote de 1.415 registros,

Quadro 1 – Resumo estatístico do processo de limpeza e filtragem dos dados.

Etapas do Processamento	Registros	Redução	Motivação Técnica
Base Bruta Cadastral	66.241	0	Volume total de linhas sem tratamento.
Remoção de Duplicatas de BO	37.457	28.784	Consolidação de múltiplos itens por even.
Filtro de Ocorrências Residenciais	37.288	169	Omissão legal de endereço da vítima.
Exclusão de Logradouros Vazios	37.243	45	Inviabilidade de identificação do local.
Geocodificação de Inconsistências	36.403	840	Erro de localização na malha viária.

Fonte: Elaborado pelo autor (2026).

correspondentes a 894 endereços únicos padronizados, à interface de programação de aplicações (API) do OpenStreetMap (HAKLAY; WEBER, 2008; OPENSTREETMAP, 2026). O procedimento localizou com sucesso 376 endereços mapeados, regularizando 575 ocorrências. Por outro lado, 518 endereços não foram localizados pelas bases cartográficas, provocando o descarte de 840 registros inconsistentes.

O banco de dados analítico final consolidou-se com 36.403 ocorrências georreferenciadas, o que representa 54,96% do volume bruto original. Embora os registros oficiais não descrevam os tipos de mercadorias subtraídas ou os calibres de armas empregados pelas quadrilhas, a resolução espacial e o intervalo mensal obtidos nesta série histórica de onze anos são adequados para dar suporte às análises exploratórias e ao treinamento das redes neurais espaço-temporais.

3.1.1 Dados de Infraestrutura Policial e Junção Espacial

Para analisar a relação geométrica entre a atividade criminal e a presença das forças de segurança, mapeou-se a infraestrutura policial do estado por meio das coordenadas de 141 bases operacionais ativas da Polícia Rodoviária Federal

Quadro 2 – Atributos dos registros policiais extraídos e enriquecidos.

Atributo	Domínio	Descrição
ID_DELEGACIA_BASE_POLICIA	Inteiro	Identificador da base policial
NOME	Texto	Nome da base policial
ENDERECO	Texto	Endereço da base policial
TIPO	Inteiro	Classificação (Polícia Estadual ou Federal)
LATITUDE	Real	Latitude
LONGITUDE	Real	Longitude

Fonte: Elaborado pelo autor (2026).

(2025) e da Polícia Militar de São Paulo (2025), conforme detalhado no Quadro 2.

Essas estruturas foram integradas ao banco de dados de crimes por meio de uma operação de junção espacial baseada na proximidade. Para evitar o alto custo computacional de calcular uma matriz exaustiva com todas as combinações possíveis entre os milhares de incidentes e as bases, aplicou-se o algoritmo de busca espacial *Ball Tree* (DOLATSHAH; HADIAN; MINAEI-BIDGOLI, 2015). A rotina foi desenvolvida com o auxílio da biblioteca `scikit-learn` em linguagem `Python`, utilizando uma máscara de validação para isolar as linhas com coordenadas preenchidas. Como o algoritmo exige dados em radianos para aplicar a métrica de Haversine, os valores em graus decimais foram convertidos previamente por funções de transformação angular.

A árvore de busca foi construída tomando como referência as coordenadas das unidades policiais. A consulta foi configurada com o parâmetro de vizinhos mais próximos igual a um ($k = 1$), estabelecendo uma regra de associação exclusiva em que cada um dos 36.403 roubos de carga é vinculado unicamente à base policial que apresentar a menor distância física em relação ao local do crime. Para converter a distância angular resultante em quilômetros, os valores foram multiplicados pelo raio médio da Terra, fixado metodologicamente em 6371 quilômetros. Os índices gerados foram achatados em uma estrutura linear simples para guiar a busca posicional na tabela de infraestrutura por meio do comando de indexação por linha, extraíndo o identificador, o nome e as coordenadas da base correspondente.

A análise descritiva das distâncias calculadas revelou um artefato de registro em cinco ocorrências específicas, nas quais a distância geodésica resultante até a base policial mais próxima foi exatamente igual a zero quilômetro. A verificação detalhada dessas linhas demonstrou que a latitude e a longitude do crime coincidiam perfeitamente com as coordenadas geográficas da própria delegacia (como as unidades de identificação 35 e 8). Esse fenômeno ocorre porque, na ausência de informações precisas sobre o local exato do roubo na rodovia, o agente responsável registrou o endereço da própria repartição policial como o local do fato, configurando um viés administrativo que foi devidamente mapeado e tratado para não distorcer os índices de acessibilidade do modelo.

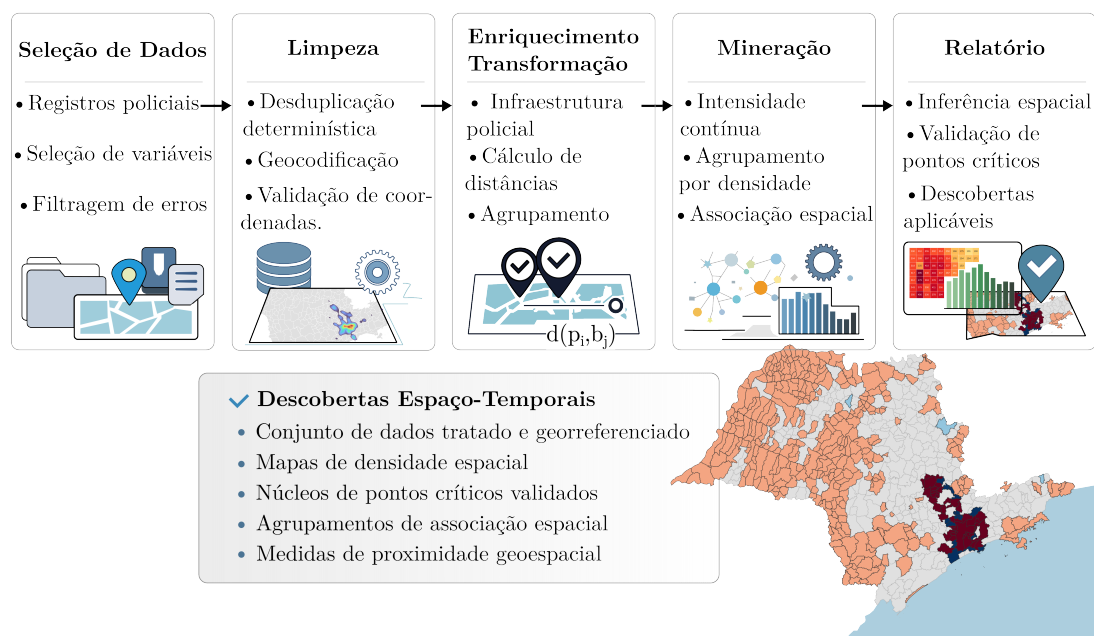
O uso da distância de Haversine calcula a menor extensão entre dois pontos considerando a curvatura terrestre, mas desconsidera as curvas e restrições das estradas, o que tende a subestimar o deslocamento real de uma viatura. Apesar dessa limitação, essa escolha justifica-se nesta tese porque o indicador funciona como uma medida geral de acessibilidade institucional. O cálculo em linha reta poupa o modelo de realizar consultas a sistemas externos de roteamento rodoviário para dezenas de milhares de pontos, além de proteger a base de dados contra inconsistências causadas por reformas asfálticas, interdições de pistas ou mudanças de sentido em trechos viários ocorridas ao longo dos onze anos da série temporal investigada.

3.2 Metodologia Analítica

Este estudo segue o método KDD (SHU; YE, 2023) para garantir um fluxo de trabalho estruturado, transparente e reproduzível. O processo organiza-se a partir das etapas sequenciais de transformação, mineração e avaliação dos dados, cuja representação lógica está detalhada na Figura 1. O objetivo central dessa modelagem é preservar a interpretabilidade visual das ocorrências e, simultaneamente, validar de forma estatística as zonas de risco utilizando ferramentas de análise espacial profunda.

O fluxo foi projetado para aplicar critérios homogêneos de processamento sobre a série histórica de 2012 a 2022. Essa padronização metodológica assegura a

Figura 1 – Fluxo de trabalho analítico baseado em KDD para a análise espacial do roubo de cargas.



Fonte: Elaborada pelo autor (2026).

estabilidade dos resultados e fornece uma base de dados consistente para subsidiar o planejamento e as tomadas de decisão estratégica no setor de segurança pública. Os dados consolidados da SSP-SP e as coordenadas das bases rodoviárias constituem as entradas do sistema, enquanto os mapas de densidade, os agrupamentos de conectividade territorial e os indicadores de associação espacial representam as saídas do pipeline.

3.2.1 Transformação de Dados e Cálculo Geodésico

Após os procedimentos de engenharia e filtragem realizados na base bruta, a etapa de transformação concentrou-se na extração das variáveis estruturais para a modelagem relacional, isolando os vetores de tempo, os identificadores municipais e os pares de coordenadas geográficas. Para incorporar a proximidade das forças de segurança como um fator analítico, foi necessário calcular a distância geodésica

entre os pontos de ocorrência dos roubos e as unidades policiais operacionais mais próximas, considerando a curvatura do planeta.

O cálculo da menor distância sobre a superfície esférica terrestre foi executado por meio da métrica de Haversine (PRASETYA et al., 2020). Tomando como referência a localização espacial de um crime, representada pelo vetor de coordenadas $p_i = (\varphi_i, \lambda_i)$, e a posição da base policial mais próxima, dada por $b_j = (\varphi_j, \lambda_j)$, em que φ e λ correspondem à latitude e à longitude convertidas para radianos, a distância geodésica $d(p_i, b_j)$ é estabelecida pela equação:

$$d(p_i, b_j) = 2R \arcsen \left(\sqrt{\sin^2 \left(\frac{\Delta\varphi}{2} \right) + \cos(\varphi_i) \cos(\varphi_j) \sin^2 \left(\frac{\Delta\lambda}{2} \right)} \right)$$

na qual R representa o raio médio da Terra, fixado em 6371 quilômetros. O parâmetro de consulta de vizinhos mais próximos foi definido como igual a um ($k = 1$), determinando que cada roubo de carga estabeleça um vínculo de associação único e exclusivo com a base policial que apresentar a menor distância física absoluta. Para viabilizar o processamento em larga escala desse volume de dados e mitigar a sobrecarga computacional das consultas iterativas, a busca espacial foi otimizada por meio da estrutura de indexação *Ball Tree*.

3.2.2 Mineração Espacial: KDE, DBSCAN e LISA

A mineração espacial utiliza três métodos complementares para extrair e validar os padrões criminais: a estimativa de densidade por kernel, para modelar o risco contínuo no território; o algoritmo DBSCAN, para agrupar as ocorrências em centros de massa representativos; e o indicador LISA, para validar a significância estatística dos agrupamentos espaciais.

A estimativa de densidade por kernel mapeia a distribuição probabilística das ocorrências, transformando pontos discretos em uma superfície contínua de risco (SCOTT, 2018). A função de densidade acumulada $\hat{f}(x)$ em um vetor de coordenadas de destino x é calculada a partir do somatório ponderado dos eventos vizinhos:

$$\hat{f}(x) = \frac{1}{nh^2} \sum_{i=1}^n K \left(\frac{d(x, x_i)}{h} \right)$$

em que n é o número total de crimes e x_i representa o vetor de coordenadas da i -ésima ocorrência histórica. Adotou-se o kernel Gaussiano bidimensional como função de suavização ponderada, cuja expressão matemática é definida por:

$$K(u) = \frac{1}{2\pi} e^{-\frac{1}{2}\|u\|^2}$$

sendo o argumento determinado pela razão de distância $u = \frac{d(x, x_i)}{h}$. O parâmetro de largura de banda (h) foi calibrado com o valor padrão ajustado unitariamente ($bw_adjust = 1$), escolha que otimiza a janela de busca para equilibrar a sensibilidade visual entre os focos de criminalidade locais e os eixos de difusão regionais ao longo das rodovias.

Como a densidade por kernel gera uma superfície contínua sem delimitar fronteiras de objetos, aplicou-se o algoritmo DBSCAN para extrair os núcleos discretos de risco (CHEN et al., 2023). Esses núcleos representam os centros de gravidade espacial onde os roubos se concentram em um raio delimitado, servindo como base para os nós do grafo espaço-temporal. O agrupamento depende da calibração de dois parâmetros estruturais. O primeiro é o raio de busca (ε), fixado metodologicamente em 1,5 quilômetro para delimitar a zona de influência e conexão direta entre roubos em um mesmo trecho viário ou terminal logístico. O segundo parâmetro é o número mínimo de pontos ($MinPts$), que foi ajustado com valor igual a um ($MinPts = 1$). Essa configuração afasta as heurísticas tradicionais de eliminação de ruídos e funciona como um mecanismo de compressão de dados, garantindo que nenhuma ocorrência isolada seja descartada, mas sim absorvida como um nó representativo ou transformada em um centro de massa individualizado para manter a integridade estatística da série histórica.

Para comprovar que os agrupamentos identificados possuem relevância estatística e não derivam de flutuações aleatórias, calculou-se o indicador LISA (ANSELIN, 1995; SILVA, 2025). Essa técnica avalia a autocorrelação espacial local, demonstrando se o comportamento do crime em um município é influenciado pela vizinhança. Para cada unidade espacial i , a estatística é formulada por:

$$I_i = \frac{y_i - \bar{y}}{S^2} \sum_j w_{ij} (y_j - \bar{y})$$

em que y_i representa a contagem absoluta de roubos de carga registrados no município i , $\bar{y}$ corresponde à média de ocorrências do estado, S^2 indica a variância global da série e w_{ij} expressa os elementos da matriz de pesos espaciais baseada no critério de contiguidade do tipo Rainha (*Queen*). A significância estatística foi validada em nível de confiança de 95% ($p < 0,05$) por meio de simulações espaciais de Monte Carlo, permitindo mapear os municípios e classificar os eixos rodoviários críticos em agrupamentos estáveis do tipo Alto-Alto.

3.2.2.1 Operação e Espaço de Fase do Indicador LISA

O funcionamento do indicador LISA baseia-se na decomposição do índice de autocorrelação espacial global em componentes locais, permitindo mensurar a contribuição individual de cada município para a formação de agrupamentos territoriais. O núcleo desse cálculo reside na avaliação do desvio do comportamento de uma localidade em relação à média global do estado, ponderado pela soma dos desvios observados em sua vizinhança imediata. Para que essa interação geográfica seja processada matematicamente, faz-se necessária a estruturação de uma matriz de pesos espaciais W , cujos elementos w_{ij} definem a existência e a intensidade dos vínculos de adjacência entre as unidades espaciais. Esta tese adota o critério de contiguidade do tipo Rainha (*Queen*), em que dois municípios são considerados vizinhos caso compartilhem qualquer segmento de fronteira linear ou ponto de vértice geométrico comum. Adicionalmente, aplica-se a normalização por linha sobre a matriz W , transformando os pesos absolutos em frações cuja soma para cada linha é igual à unidade, procedimento que padroniza a escala do cálculo e permite interpretar o somatório vizinho como uma média ponderada local.

A interpretação das dinâmicas territoriais decorrentes da aplicação do LISA fundamenta-se na segmentação do espaço de fase de Moran em quatro quadrantes de associação espacial, os quais discriminam a natureza dos agrupamentos significativos. O primeiro quadrante, denominado Alto-Alto, caracteriza os focos críticos ou *hotspots* regionais, agrupando municípios com elevados índices de roubo de cargas que estão circundados por vizinhanças em idêntica situação de alta criminalidade. Em sentido inverso, o quadrante Baixo-Baixo reúne as localidades de baixa incidência criminal cujos vizinhos imediatos também apresentam registros reduzidos,

delimitando os chamados *coldspots* ou zonas de amortecimento e isolamento do delito. Os dois quadrantes restantes identificam as assimetrias espaciais denominadas *outliers* territoriais, em que o comportamento local diverge da tendência de seu entorno. O quadrante Alto-Baixo mapeia os municípios de alta criminalidade isolados geograficamente por vizinhanças com baixas taxas de ocorrência, enquanto a classe Baixo-Alto aponta as cidades de baixa contagem de roubos imersas em regiões de alta densidade criminal, funcionando frequentemente como franjas de transição e canais por onde o risco se espalha dos polos metropolitanos em direção ao interior.

A validação dessas estruturas espaciais exige comprovar que os agrupamentos identificados possuem relevância estatística, afastando a hipótese nula de que a distribuição observada decorre de flutuações aleatórias no espaço. Como os dados criminais não seguem uma distribuição normal teórica, o cálculo dos níveis de significância é realizado por meio de testes de permutação condicional utilizando simulações estatísticas de Monte Carlo. Nesse procedimento computacional, os valores observados de roubos de carga são redistribuídos de forma aleatória entre os municípios milhares de vezes, mantendo fixa apenas a contagem da localidade em análise para gerar uma distribuição empírica de referência para cada ponto do território. O Índice de Moran Local observado é então comparado com essa distribuição simulada para o cálculo de pseudo- p -valores, estabelecidos sob o patamar crítico de 5% de probabilidade de erro ($p < 0,05$). Esse rigor estatístico garante que apenas as unidades territoriais que apresentarem desvios significativos em relação ao padrão de aleatoriedade sejam coloridas e classificadas no painel multitemporal de mapas da tese, isolando o sinal estrutural do ruído estatístico.

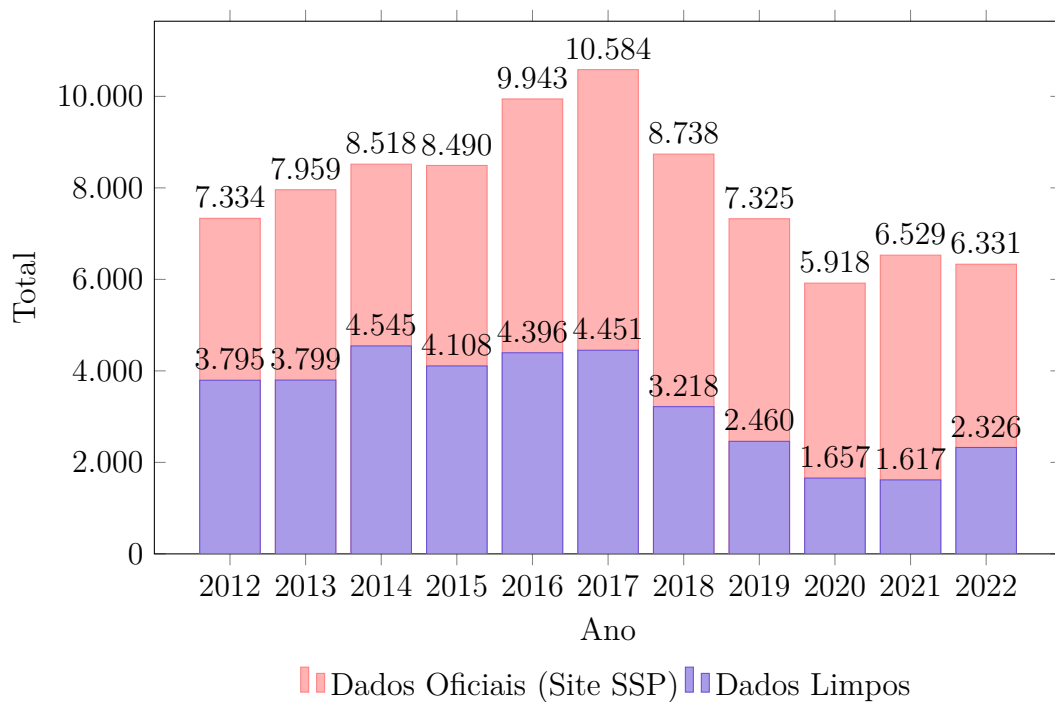
3.3 Resultados e Discussão

Nesta seção, apresentam-se e interpretam-se os padrões espaciais e temporais descobertos por meio do processo de mineração de dados.

3.3.1 Curadoria de Dados e Dinâmica Temporal

O processo de filtragem e engenharia de dados validou 36.372 ocorrências georreferenciadas entre os anos de 2012 e 2022. A Figura 2 apresenta a comparação entre a evolução temporal da base de dados limpa e as estatísticas brutas publicadas originalmente pela Secretaria de Segurança Pública de São Paulo (2025).

Figura 2 – Série temporal de ocorrências de roubo de carga: dados processados *versus* estatísticas oficiais da SSP-SP.



Fonte: Elaborada pelo autor (2026).

Os gráficos revelam uma diferença sistemática de volume entre as duas séries históricas. No ano de 2017, período que registra o topo da série, a base de dados curada consolida 4.451 incidentes espaciais, enquanto os indicadores brutos oficiais apontam 10.584 registros. Essa redução de aproximadamente 58% decorre do critério de deduplicação aplicado, que remove as linhas repetidas geradas quando o sistema policial gera uma ocorrência geral, sem veículos ou partes associadas. Esse

tratamento garante que cada evento criminoso seja contabilizado com materialidade, isto é, tendo um ente associado.

Apesar da divergência na escala dos volumes absolutos, a tendência de comportamento temporal das duas curvas é convergente, apresentando uma correlação de Spearman de $\rho = 0,94$ ($p < 0,001$). O pico de ocorrências identificado em 2017 coincide historicamente com os reflexos da recessão econômica nacional (JUSTUS et al., 2018), cenário em que o desemprego estrutural amplia a vulnerabilidade social e a atratividade dos mercados ilícitos. Por outro lado, a trajetória de queda observada a partir de 2018 reflete o aprimoramento de tecnologias de rastreamento logístico e a alteração de estratégias operacionais das forças de segurança pública e privada (Overhaul, 2025).

3.3.2 Persistência Espaço-Temporal e Concentração Regional

A análise integrada das dimensões geográficas e cronológicas demonstra que o espalhamento do risco não ocorre de maneira aleatória no território paulista, manifestando eixos estáveis de concentração.

3.3.2.1 Análise das Superfícies de Intensidade Rodoviária

A Figura 3 apresenta os mapas de risco anuais gerados por meio da estimativa de densidade por kernel com base Gaussiana, ajustada para equilibrar a visualização dos trechos locais e regionais. Ao avaliar a série histórica de 2012 a 2022, observa-se uma acentuada estabilidade geográfica das manchas de calor. Longe de se comportar de forma errática, a atividade criminal permanece fixada na Região Metropolitana de São Paulo e na Baixada Santista, região que abriga o principal complexo portuário do subcontinente.

Esse comportamento espacial apoia as premissas da Teoria das Atividades Rotineiras, que associa a ocorrência do delito à convergência territorial entre alvos adequados, agentes motivados e a ausência de vigilância eficiente (ARANSIOLA; JUSTUS; CECCATO, 2023). A distribuição das zonas críticas acompanha rigorosamente o traçado dos principais eixos de transporte do estado, com destaque para os complexos viários Anhanguera-Bandeirantes, Washington Luís e Castello

Branco, estruturas que funcionam como os principais corredores de abastecimento e escoamento industrial.

Um aspecto revelado pela análise cartográfica é a independência existente entre o volume total de ocorrências e a estrutura espacial do crime. Mesmo nos períodos em que o número absoluto de roubos sofreu quedas acentuadas, a distribuição geográfica das manchas de calor permaneceu inalterada. Esse cenário indica que as oscilações econômicas e as intervenções policiais de curto prazo conseguem reduzir a quantidade de crimes, mas não alteram a dependência que o crime organizado possui em relação à infraestrutura de transportes. O roubo de cargas opera como uma atividade econômica institucionalizada que se molda aos fluxos da logística formal, mantendo suas bases territoriais fixas independentemente das flutuações do calendário.

3.3.2.2 Concentração Temporal e Testes de Invariância

A Figura 4 detalha a distribuição das ocorrências ao longo das 24 horas do dia para cada ano da série histórica. Os dados revelam que o roubo de cargas é predominantemente diurno, concentrando-se no horário comercial entre 06:00 e 18:00, período que registra a maior movimentação de veículos de transporte de mercadorias. As medianas cronológicas mantêm-se estáveis entre 10:00 e 12:00 ao longo de toda a década, enquanto o intervalo da madrugada apresenta baixa densidade de registros, padrão que converge com as evidências empíricas coletadas em outros cenários internacionais (RAMIREZ, 2024; EMAANI; SAGHAEI, 2025).

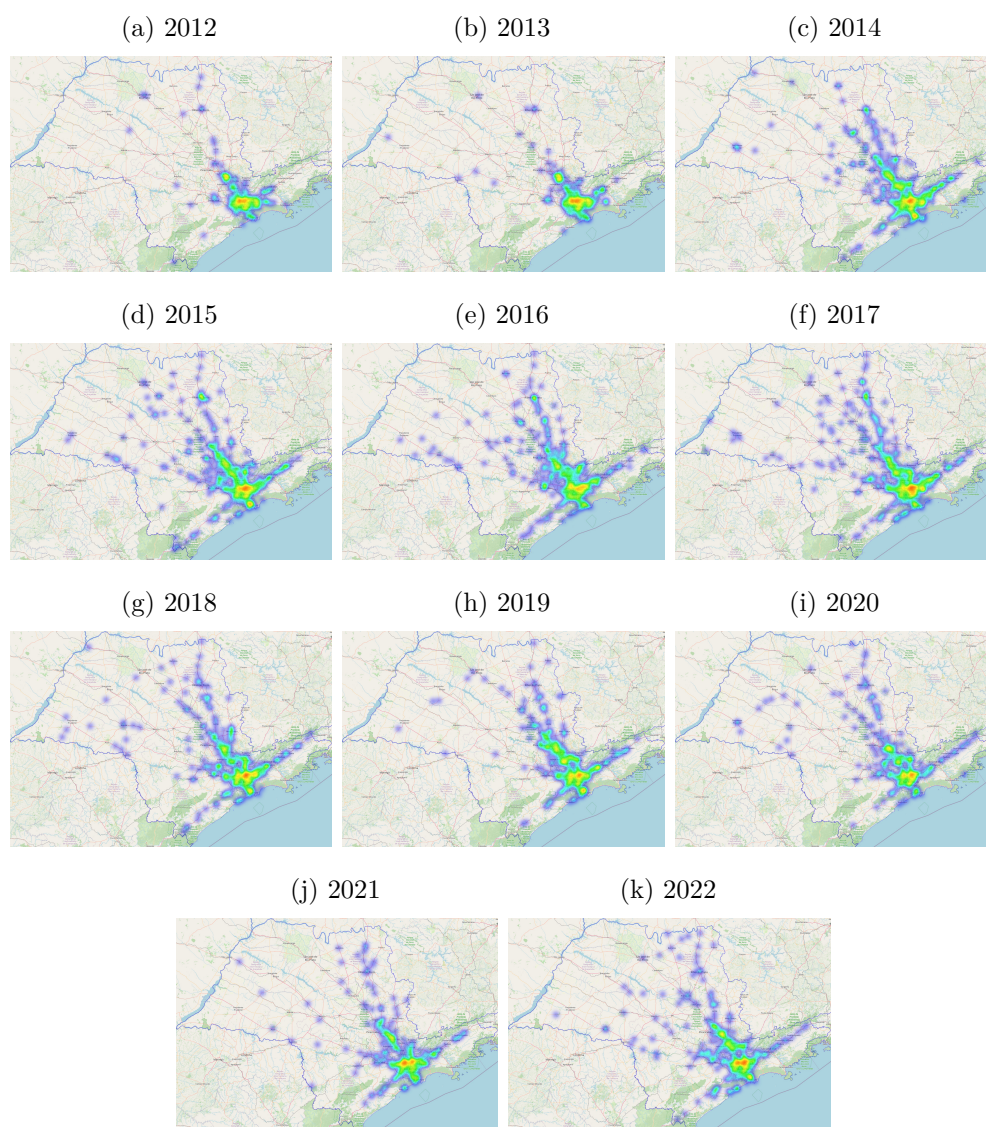
Para validar a estabilidade desse comportamento horário, aplicou-se o teste não paramétrico de Kruskal-Wallis, avaliando a hipótese nula de igualdade entre as medianas temporais dos anos da série histórica:

$$H_0 : \text{Mediana}(H_{2012}) = \text{Mediana}(H_{2013}) = \dots = \text{Mediana}(H_{2022})$$

O teste estatístico resultou em um índice $H = 14,2$ com valor- $p = 0,164$. Como o valor- p é superior ao limite de significância de 0,05, a hipótese nula não é rejeitada, confirmando a invariância temporal da assinatura horária em base anual.

A análise por dias da semana, apresentada na Figura 5, confirma a manutenção do padrão diurno de segunda-feira a sábado. A única alteração ocorre

Figura 3 – KDE longitudinal da intensidade de roubo de cargas no estado de São Paulo (2012–2022). *Hotspots* demonstram notável estabilidade espacial ao longo dos principais corredores logísticos.



Fonte: Elaborada pelo autor (2026).

aos domingos, intervalo que registra uma queda acentuada de casos no período da manhã e elevação progressiva entre 16:00 e 21:00, deslocando a mediana diária para as 17:00. O teste de Kruskal-Wallis confirmou a relevância estatística dessa variação semanal ($H = 846,5$; valor- $p < 0,001$), e as verificações posteriores de Mann-Whitney com correção de Bonferroni ratificaram que o domingo possui uma distribuição horária distinta dos demais dias da semana. Essa dinâmica intradiária repete-se de forma estável ao longo dos meses do ano, conforme indica o gráfico de cristas da Figura 6.

Essa variação identificada aos domingos decorre diretamente das regras de funcionamento do transporte comercial, que enfrenta restrições de tráfego de veículos pesados em rodovias no período matutino e pausa nas atividades dos centros de distribuição. O fluxo logístico é retomado no final da tarde para garantir o abastecimento que deve chegar aos destinos na manhã de segunda-feira.

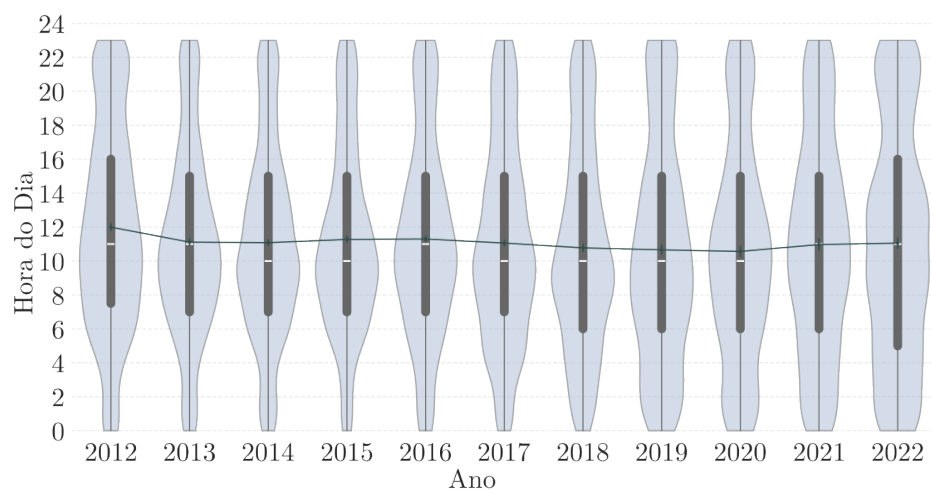
A alta concentração de crimes no final da manhã nos dias úteis reflete a exposição ao risco decorrente do maior volume de caminhões em circulação nas estradas, e não uma eficiência superior das quadrilhas nesse horário. Sob a ótica da criminologia crítica, essa sincronia confirma que o roubo de cargas funciona integrado ao ciclo de reprodução e escoamento do capital comercial. Esse cenário expõe os limites de estratégias policiais puramente reativas, demonstrando que a ação estatal muitas vezes atua na gestão burocrática dos incidentes, cujos limites operacionais permanecem subordinados ao ritmo de circulação das mercadorias.

3.3.2.3 Mapeamento de Agrupamentos e Identificação de Zonas Críticas

Para segmentar as superfícies contínuas do território antes da modelagem em grafos, utilizou-se o algoritmo DBSCAN para identificar agrupamentos discretos de risco espacial. As distâncias foram calculadas com base na métrica esférica de Haversine para garantir o rigor geométrico do mapeamento longitudinal apresentado na Figura 7.

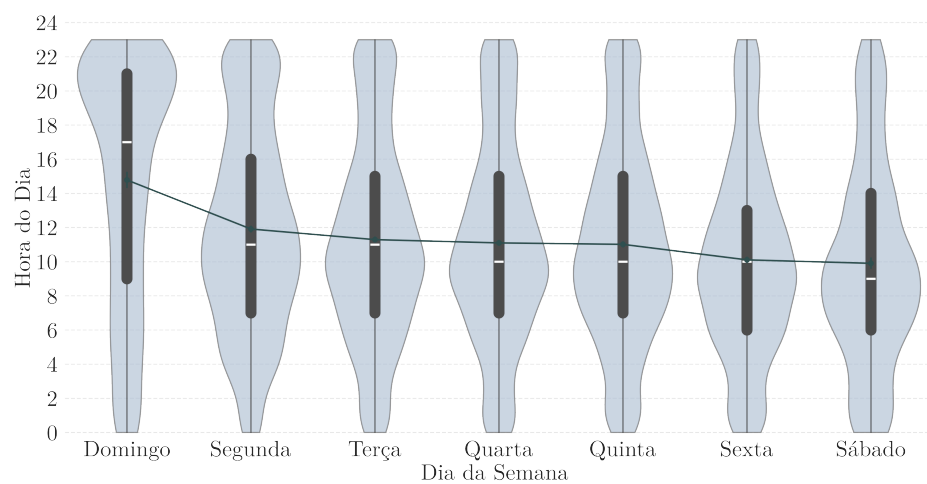
O Quadro 3 reúne as métricas extraídas desse processo de agrupamento. A taxa de compressão espacial expressa o percentual de redução volumétrica alcançado ao concentrar os milhares de coordenadas individuais de crimes em seus respectivos

Figura 4 – Distribuição horária dos roubos de carga ao longo dos anos (2012–2022).



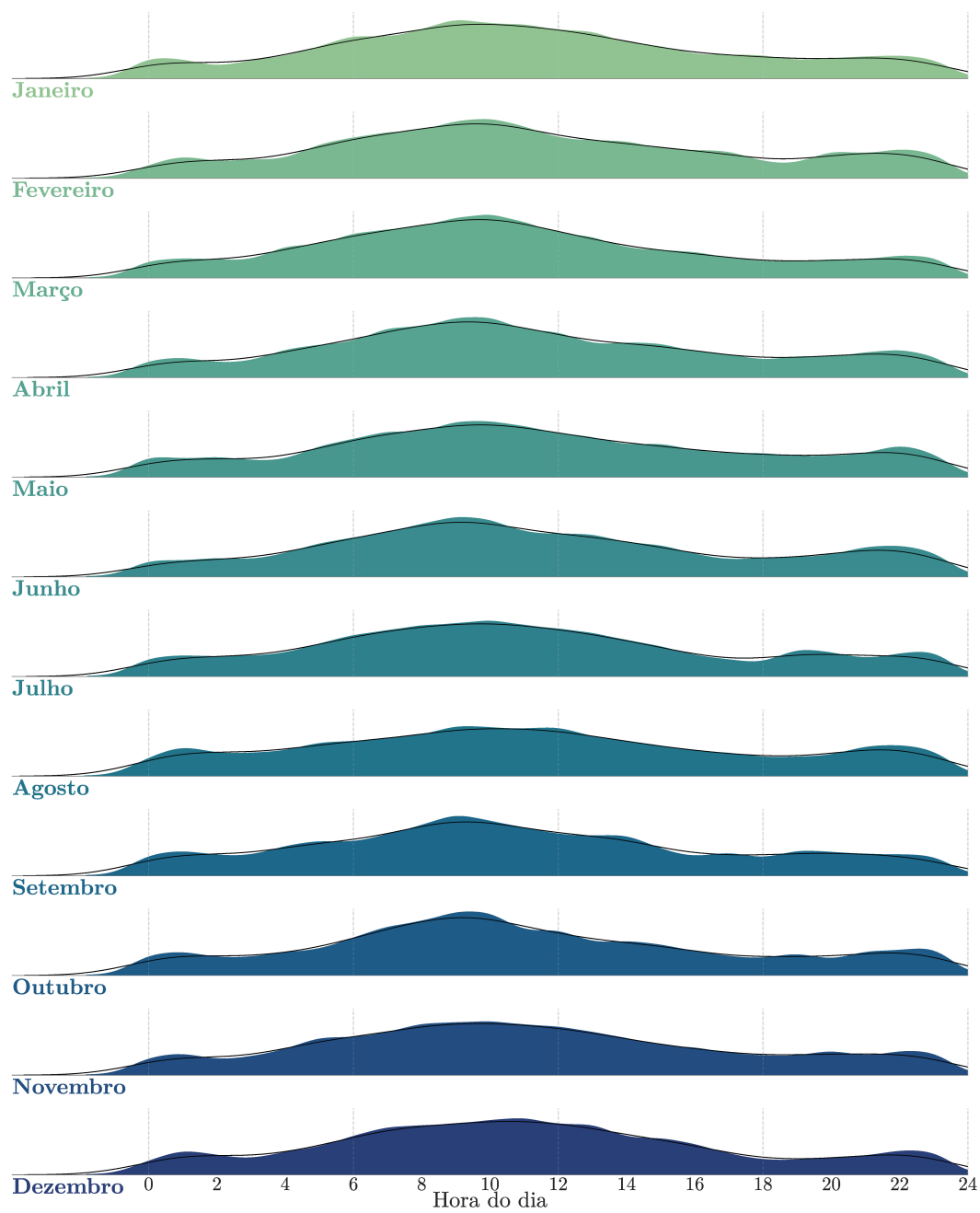
Fonte: Fonte: Elaborada pelo autor (2026).

Figura 5 – Distribuição horária dos roubos de carga por dia da semana (2012–2022).



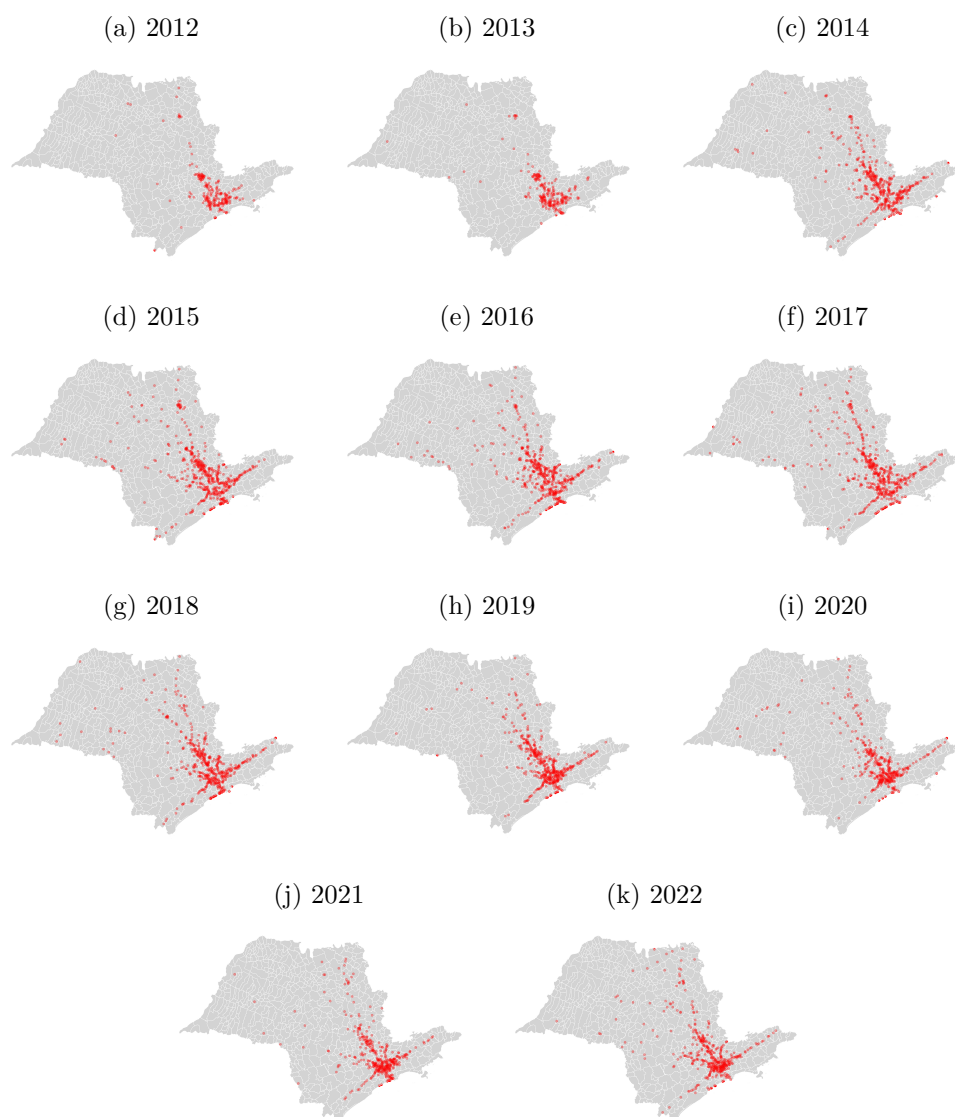
Fonte: Fonte: Elaborada pelo autor (2026).

Figura 6 – Distribuição horária das ocorrências agregada por mês (2012–2022).



Fonte: Fonte: Elaborada pelo autor (2026).

Figura 7 – Agrupamentos de roubos de carga identificados pelo DBSCAN.



Fonte: Elaborada pelo autor (2026).

centros de gravidade ou nós representativos. Os resultados indicam que esse índice superou o patamar de 80% na maioria dos anos analisados. Essa elevada taxa demonstra que os eventos criminosos ocorrem de forma aglutinada dentro de zonas logísticas delimitadas.

Diferente de infrações urbanas pulverizadas como o furto de pedestres, que se difundem de maneira dispersa acompanhando a densidade residencial, o roubo de cargas vincula-se diretamente à infraestrutura comercial fixa do estado. Essa concentração indica que intervenções táticas focadas nas rodovias e nos terminais de transbordo apresentam maior potencial de eficiência do que o patrulhamento genérico.

Quadro 3 – Métricas de desempenho do agrupamento espacial via DBSCAN (2012–2022).

Ano	Total de Eventos	Núcleos (Aglomera-dos)	Taxa de Compressão	Silhouette
2012	3.768	234	93,8%	-0,272
2013	3.818	244	93,6%	-0,196
2014	4.539	449	90,1%	-0,172
2015	4.110	507	87,7%	-0,104
2016	4.400	488	88,9%	-0,120
2017	4.451	513	88,5%	0,003
2018	3.217	487	84,9%	0,203
2019	2.456	476	80,6%	0,439
2020	1.670	411	75,4%	0,441
2021	1.607	400	75,1%	0,395
2022	2.330	438	81,2%	0,589

Fonte: Elaborado pelo autor (2026).

A rigidez das redes de transporte impede o deslocamento simples das ocorrências para caminhos alternativos. Como o crime de carga depende do escoamento de grandes volumes de riqueza, as quadrilhas enfrentam barreiras para migrar para estradas vicinais ou zonas rurais de baixa densidade, locais onde o tráfego de carretas pesadas é inexistente e a geometria das vias inviabiliza as manobras de abordagem. Essa configuração força a criminalidade a permanecer nos corredores

logísticos principais.

A evolução histórica do coeficiente Silhouette confirma uma alteração estrutural na organização espacial do delito. Entre 2012 e 2016, os valores negativos demonstram que as ocorrências se organizavam na forma de faixas contínuas ao longo das rodovias, gerando sobreposição nas fronteiras dos agrupamentos. A partir de 2018, em sincronia com a redução do volume bruto de casos, o índice tornou-se positivo e atingiu 0,589 no ano de 2022. Esse aumento do indicador comprova que o fenômeno tornou-se espacialmente mais isolado e focado, concentrando-se em pontos estratégicos do interior do estado, como entroncamentos rodoviários e pátios de descanso de postos de combustíveis (DIAZ; CERNA; LIZA, 2025). Observa-se também a formação de aglomerados lineares que se estendem até a divisa do Vale do Paraíba com o Estado do Rio de Janeiro, evidenciando rotas estruturadas para facilitar a fuga interestadual e justificando o uso de abordagens baseadas em grafos para modelar essas conexões.

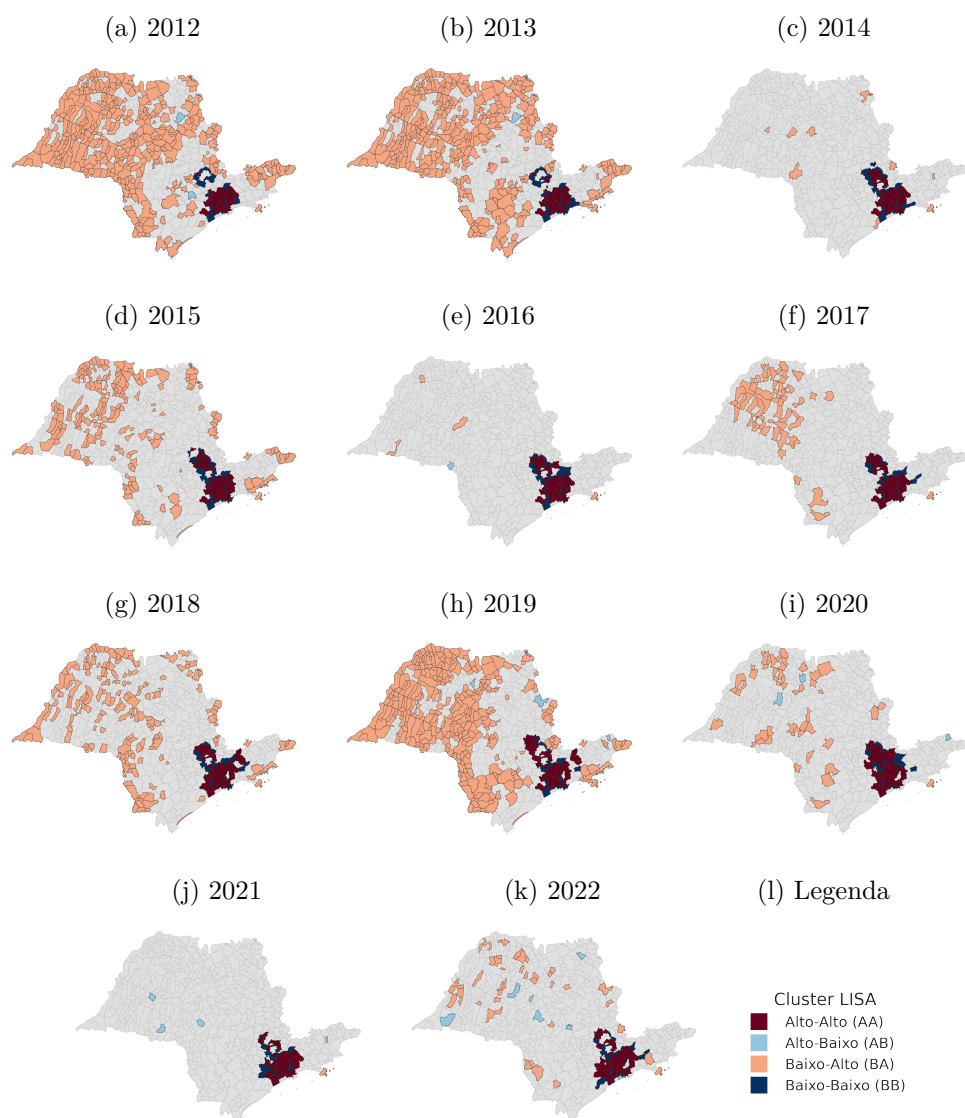
3.3.3 Verificação Estatística da Dependência Espacial

Para comprovar que a concentração observada possui relevância estatística, aplicaram-se os Indicadores Locais de Associação Espacial (LISA) (ANSELIN, 1995). A Figura 8 apresenta o Índice de Moran Local calculado anualmente para os 645 municípios paulistas.

A análise estatística local valida que a distribuição do crime é governada por processos de dependência espacial estáveis. Os agrupamentos classificados como Alto-Alto identificam municípios de alta incidência criminal cercados por vizinhos em idêntica situação crítica, delimitando os polos regionais de risco acumulado ($p < 0,05$) ancorados na Região Metropolitana de São Paulo e na Baixada Santista. Por sua vez, as áreas marcadas como Baixo-Baixo representam municípios de baixa atividade criminosa circundados por cidades seguras, mapeando extensas zonas de proteção nas regiões oeste e sudoeste do estado.

Esse descolamento entre a oscilação do volume anual de roubos e a estabilidade geográfica das manchas críticas decorre da fixidez das redes de transporte. Enquanto os fatores econômicos alteram a quantidade total de ocorrências ao longo

Figura 8 – Mapas de agrupamentos territoriais LISA: Alto-Alto (vermelho) indica focos críticos regionais; Baixo-Baixo (azul) representa zonas protegidas.



Fonte: Fonte: Elaborada pelo autor (2026).

do tempo, a localização dos crimes não muda porque a infraestrutura física que comanda o fluxo de capitais permanece estática.

As categorias periféricas auxiliam na compreensão desse processo de difusão. As zonas do tipo Baixo-Alto, que reúnem cidades de baixo índice cercadas por municípios de alta criminalidade, funcionam como franjas de transição geográfica por onde o risco se espalha a partir do centro metropolitano. Esses resultados demonstram que as ações de segurança pública não devem focar em municípios de forma isolada, mas sim em corredores regionais integrados. Essa dependência espacial e a estabilidade da topografia de vizinhança fornecem o embasamento empírico necessário para a aplicação de Redes Neurais em Grafos, estruturadas especificamente para processar e prever dinâmicas de rede no território.

3.3.4 Proximidade Geográfica da Infraestrutura Policial

O Quadro 4 apresenta as métricas descritivas da distância em linha reta entre os locais das ocorrências e as bases de policiamento rodoviário mais próximas. Os indicadores indicam que o roubo de cargas não se restringe a territórios isolados ou desprovidos de presença estatal.

A metade dos crimes catalogados ocorre a menos de 7,12 quilômetros de distância de uma unidade policial ativa, e 25% dos casos concentram-se em um raio de até 4,08 quilômetros de proximidade. Sob as condições de deslocamento em rodovias, em que veículos de emergência mantêm velocidade média próxima a 80 km/h, essas extensões espaciais correspondem a tempos estimados de resposta operacional de três a cinco minutos. Essa proximidade geográfica decorre do planejamento urbano da infraestrutura de segurança, cuja distribuição acompanha os trechos de maior tráfego da malha viária. As organizações criminosas não fogem da presença física da polícia; elas optam por atuar nos eixos de maior densidade logística, assumindo o risco da proximidade institucional devido à alta rentabilidade das mercadorias em circulação.

A distância mínima computada de 0,00 km mapeia cinco casos específicos registrados nas coordenadas exatas da base policial, viés administrativo gerado quando a ausência de dados precisos sobre o local do roubo leva o agente a registrar

o endereço da própria delegacia como o local do fato. No extremo oposto, a distância máxima de 85,75 km reflete o espaçamento da infraestrutura no interior do estado, ocorrendo em trechos viários secundários afastados dos eixos de escoamento industrial principal.

A análise da proximidade institucional indica que a estratégia das quadrilhas divide-se em dois comportamentos principais. O primeiro caracteriza-se pela atuação concentrada nas proximidades imediatas das bases e dos centros de transbordo, onde o grande volume de mercadorias de alto valor compensa o risco de uma resposta rápida da polícia. O segundo comportamento baseia-se na exploração de rotas secundárias e eixos de menor fiscalização, nos quais os criminosos aproveitam a distância das redes de socorro para interceptar alvos específicos.

Essa dinâmica reflete o conceito de efeito de borda discutido na criminologia ambiental, em que os agentes calculam o valor do alvo contra o risco real de captura, atuando em áreas onde julgam ser possível neutralizar a resposta policial (KEZIOU; DOULKERIDIS, 2025). A baixa taxa de letalidade identificada na série histórica (0,06%) confirma o caráter empresarial dos grupos criminosos, cujo objetivo central concentra-se na expropriação de mercadorias e na regularidade financeira dos receptadores, evitando confrontos armados que possam atrair operações policiais e paralisar a circulação das redes de distribuição.

A ocorrência de roubos em raios próximos às bases policiais indica que a simples instalação de postos físicos fixos é insuficiente para inibir o delito. Existe uma coexistência entre o controle estatal e a atividade ilícita sobre a mesma infraestrutura viária. Sob a ótica materialista, os postos de policiamento são posicionados para tentar assegurar a regularidade das artérias de circulação do capital. Contudo, o crime organizado opera de forma integrada a esse sistema econômico, tratando a presença física do Estado não como uma barreira intransponível, mas como um custo operacional calculável dentro de seu modelo de negócios.

3.4 Considerações Finais do Capítulo

Este capítulo analisou como funcionam os roubos de carga em São Paulo, cruzando dados estatísticos com a rede de transporte do estado. O uso das ferramen-

Quadro 4 – Estatísticas descritivas da distância geodésica até a base policial rodoviária mais próxima (2012–2022).

Métrica Analítica	Valor Computado
Número Total de Incidentes Espaciais	36.372
Distância Média Global	8,07 km
Desvio Padrão da Amostra	5,94 km
Distância Mínima Registrada	0,00 km
Percentil 25	4,08 km
Mediana (Percentil 50)	7,12 km
Percentil 75	10,44 km
Distância Máxima Identificada	85,75 km
Indicadores de Gravidade e Violência	
Vítimas Fatais Catalogadas (Óbitos)	23
Taxa de Letalidade Absoluta	0,06%

Fonte: Elaborado pelo autor (2026).

tas KDE, DBSCAN e LISA mostrou que os crimes não acontecem por acaso, mas são organizados ao longo das principais rodovias paulistas (CEDILLO-CAMPOS; FLORES-FRANCO; COVARRUBIAS, 2024).

Para além dos mapas, a permanência dessas concentrações exige uma explicação econômica e social. Os resultados apontam três conclusões que contrariam as ideias tradicionais de segurança pública:

Primeiro, a concentração dos agrupamentos em rodovias com grande volume de transporte prova que o crime organizado funciona como uma atividade econômica estruturada. Trata-se de uma operação de alto valor que depende da mesma infraestrutura construída para a circulação legal de mercadorias. Como explicam Poulantzas (2019) e Mascaro (2013), as estradas são a base material de um Estado organizado para a reprodução do capital. O asfalto que atende ao mercado formal serve também como o caminho para o roubo das cargas.

Segundo, o pico de roubos às 11:00 em dias úteis reflete a sincronia do crime com o horário comercial. Sob a ótica de Pachukanis (2017) e de Rusche e Kirchheimer (2004), o roubo de carga funciona como uma atividade profissional. Ele opera no momento de maior movimentação econômica, aproveitando a lógica

de mercado discutida por Lukács (2003).

Por fim, a curta distância entre os locais dos roubos e as bases policiais (mediana de 7,12 km) contesta a eficácia da Teoria da Dissuasão¹ (NAGIN, 2013). Com base nos estudos de Foucault (2014) e Zaffaroni (1991), nota-se que a presença da polícia nas rodovias cumpre um papel de controle do território, mas não consegue deter esse modelo de crime empresarial. A base policial é uma estrutura fixa, enquanto o transporte do capital ilícito é dinâmico. O sistema econômico absorve essa convivência porque o prejuízo financeiro é repassado ao mercado: o boletim de ocorrência serve principalmente para acionar as seguradoras, garantindo que o ciclo de circulação continue sem interrupções.

Embora os métodos baseados em densidade localizem com precisão os pontos críticos, eles são estáticos. Eles mostram a concentração no espaço, mas não conseguem capturar como os municípios dependem uns dos outros ao longo do tempo nem prever cenários futuros.

Por essa razão, compreendida a dimensão espacial e econômica do problema, é necessário avançar para os modelos de previsão.

O Capítulo 4 iniciará essa etapa analisando as limitações dos modelos estatísticos tradicionais (SARIMA) e testando a capacidade das Redes Neurais Recorrentes (RNNs) para prever o risco nas rodovias. Há uma contradição no uso dessas tecnologias: embora algoritmos semelhantes sejam utilizados para o policiamento preditivo convencional (*Predictive Policing*), o modelo proposto nesta tese busca fornecer ferramentas para rastrear e conter a circulação das redes de capital ilícito.

¹ A Teoria da Dissuasão (*Deterrence Theory*), comum na criminologia tradicional, assume que a presença da polícia e a certeza da punição desestimulam o crime. Nesta tese, o conceito é questionado: em atividades criminosas com estrutura empresarial e logística, a fiscalização policial é tratada pelos criminosos como um custo operacional previsível, e não como uma barreira que impede a ação.

4 Modelagem Preditiva de Base: Estatística Clássica vs. Redes Neurais

Historicamente, a segurança pública projeta índices criminais usando métodos estatísticos clássicos. Esses métodos assumem que os crimes seguem padrões lineares e estáveis ao longo do tempo (BOX et al., 2015). Contudo, o objetivo central desta etapa é avaliar se precisamos adotar modelos matemáticos mais complexos (não lineares). Para isso, comparamos os métodos estatísticos tradicionais com técnicas de aprendizado profundo (*Deep Learning*) (GOODFELLOW; BENGIO; COURVILLE, 2016; WANG et al., 2024a).

Começamos a nossa análise testando o modelo SARIMA, que faz previsões com base apenas no histórico de registros da própria série temporal. Em seguida, aplicamos Redes Neurais Recorrentes (RNNs) utilizando cenários com uma única variável e com múltiplas variáveis (CHO et al., 2014). Por fim, avaliamos um modelo híbrido (SARIMA-GRU) focado em corrigir os erros de previsão do método estatístico (XIAO et al., 2024). Os resultados mostram que incluir dados geográficos — como a distância até as bases policiais — melhora a capacidade de aprendizado da rede (LIU et al., 2022). Isso confirma que a previsão de roubos de carga exige uma análise que considere o território.

4.1 Preparação dos Dados e das Séries Temporais

Os registros criminais brutos exigem estruturação metodológica para que possam ser processados de forma eficiente por arquiteturas de aprendizado profundo. A transformação dos dados originais em tensores organizou-se em três etapas sequenciais executadas em linguagem *Python*: a agregação cronológica mensal, o escalonamento de amplitude das variáveis e a estruturação das sequências por meio de janelas deslizantes (*sliding windows*).

4.1.1 Agrupamento Mensal e Estruturação Cronológica

Os roubos de cargas registrados entre janeiro de 2012 e dezembro de 2022 foram consolidados em uma série temporal de periodicidade mensal. A escolha da escala mensal atenuou a volatilidade excessiva e a inflação de zeros características dos registros diários quando agregados territorialmente, conferindo maior estabilidade estatística para a identificação de componentes de tendência e sazonalidade de longo prazo (CATLETT et al., 2019). Embora essa agregação oculte as variações de curtíssimo prazo, como a dinâmica intradiária e semanal analisada no capítulo anterior, ela atende aos requisitos de planejamento estratégico e de modelagem macro temporal de base. A variável descritiva de distância até as bases policiais foi convertida para a média mensal das distâncias das ocorrências registradas em cada intervalo, integrando a dimensão espacial ao ritmo cronológico exigido pelos modelos.

A estruturação dos dados utilizou os recursos da biblioteca `pandas`. O alinhamento cronológico contínuo e sem interrupções — exigência metodológica para o funcionamento de modelos autorregressivos — foi assegurado pela aplicação do método `asfreq('MS')`, tratando os raros intervalos sem registros por meio de técnicas de preenchimento sequencial (`ffill` e `bfill`). Esse procedimento consolidou uma série histórica com $T = 132$ observações mensais ao longo dos onze anos de escopo da pesquisa. Formalmente, a série temporal é representada pelo conjunto $Y = \{y_1, y_2, \dots, y_T\}$, em que y_t expressa o volume agregado de crimes no mês $t \in \{1, \dots, 132\}$, servindo como base para a projeção do comportamento futuro do fenômeno.

4.1.2 Padronização e Escalonamento das Variáveis

Diferente das abordagens estatísticas lineares clássicas, as redes neurais recorrentes, como as arquiteturas baseadas em Unidades Recorrentes Fechadas (GRU) (CHO et al., 2014), apresentam alta sensibilidade à escala das variáveis de entrada. O ajuste paramétrico por meio do algoritmo de retropropagação através do tempo (*backpropagation through time*) (RUMELHART; HINTON; WILLIAMS, 1986) pode sofrer instabilidades numéricas, como a explosão ou o desaparecimento do

gradiente, caso as magnitudes dos dados sejam muito discrepantes (GIAMPAOLO, 2023).

Como o volume mensal de crimes se posiciona na ordem das centenas e a distância média até os postos policiais atinge a escala dos milhares de metros, aplicou-se o escalonamento linear por amplitude por meio da função `MinMaxScaler` da biblioteca `scikit-learn` sobre a variável espacial. Essa transformação mapeia os valores originais estritamente para o intervalo fechado $[0, 1]$, equalizando o peso matemático inicial atribuído aos diferentes atributos sem alterar as propriedades distributivas subjacentes (SHIH, 2022). O volume de roubos (y_t) foi mantido em sua escala inteira original para simplificar a interpretação direta e a auditoria das projeções numéricas geradas pelo modelo.

4.1.3 Estruturação dos Dados por Janela Deslizante

O treinamento supervisionado das arquiteturas de aprendizado profundo requer a conversão da série univariada contínua em pares de matrizes de características preditoras e vetores de alvos determinados. Essa reformulação foi implementada via método de janela deslizante com suporte computacional da biblioteca `NumPy` (TANG, 2017). Fixou-se uma janela temporal histórica de $w = 3$ meses para capturar o comportamento sequencial recente. Para cada período t , os valores observados nas três defasagens cronológicas anteriores são combinados em um vetor de entrada $X_t = [y_{t-3} \ y_{t-2} \ y_{t-1}]$, associado ao valor futuro imediato y_t que atua como a resposta esperada no modelo.

Essa organização estrutura os dados no formato de tensores multidimensionais compatíveis com os requisitos de entrada das bibliotecas `TensorFlow` e `Keras` (CHOLLET et al., 2015; ABADI et al., 2016). A base temporal foi cindida de forma estritamente cronológica na proporção de 80% para o conjunto de treinamento e 20% para o conjunto de teste. Sob esse critério, as primeiras 105 observações mensais, correspondentes ao intervalo entre janeiro de 2012 e setembro de 2020, foram destinadas à otimização dos pesos da rede neural. Os 27 meses restantes, abrangendo o período de outubro de 2020 a dezembro de 2022, foram isolados para a validação externa fora do período de treino (*out-of-sample*), configuração que impede o vazamento de dados (*data leakage*) e garante uma avaliação imparcial da

capacidade de generalização do modelo.

4.2 Modelagem Estatística Clássica: O Componente Linear (SARIMA)

A primeira abordagem preditiva desenvolvida nesta tese concentra-se no isolamento dos componentes lineares e sazonais da série histórica de roubos de carga. Utilizou-se a metodologia tradicional autoregressiva integrada de médias móveis sazonais com o objetivo de identificar a tendência estrutural dos incidentes e mapear os resíduos estocásticos que não podem ser explicados por formulações lineares de dependência passada. Para garantir a validade estatística das inferências e evitar o comportamento espúrio dos coeficientes, é necessário assegurar que a série temporal apresente a propriedade de estacionariedade, mantendo sua média e sua variância constantes ao longo do tempo (BOX et al., 2015).

4.2.1 Testes de Estacionariedade e Diferenciação dos Dados

A verificação formal sobre a presença de tendências estocásticas e o diagnóstico de estabilidade da série foram realizados por meio do Teste Aumentado de Dickey-Fuller (ADF) (DICKY; FULLER, 1979), utilizando rotinas computacionais do pacote `statsmodels`. A formulação estrutural desse teste baseia-se em uma equação de regressão linear por mínimos quadrados expressa por:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{j=1}^p \delta_j \Delta Y_{t-j} + \varepsilon_t \quad (4.1)$$

em que $\Delta Y_t = Y_t - Y_{t-1}$ representa a primeira diferença da série no mês t , α corresponde ao termo de intercepto ou constante de deriva (*drift*), β expressa o coeficiente associado à tendência temporal determinística linear t , e γ identifica o parâmetro crítico testado sobre o nível defasado Y_{t-1} , cujo valor determina a existência de uma raiz unitária. Os termos δ_j representam os coeficientes das primeiras diferenças defasadas até o limite do lag p , introduzidos para corrigir problemas de correlação serial nos resíduos, enquanto ε_t constitui o termo de erro estocástico modelado como ruído branco.

A mecânica de extração do valor- p nesse procedimento não segue as distribuições simétricas tradicionais da estatística paramétrica. A hipótese nula (H_0) estabelece que a série possui uma raiz unitária, o que é representado matematicamente por $\gamma = 0$, indicando um comportamento não estacionário. A partir da estimação dos parâmetros, calcula-se a estatística de teste τ , correspondente à razão entre o coeficiente estimado $\hat{\gamma}$ e o seu respectivo erro padrão. Sob a hipótese nula, essa razão não segue a distribuição t de Student devido à natureza não estacionária do regressor defasado, convergindo para funcionais do movimento Browniano. Consequentemente, o valor- p é computado a partir do confronto do índice τ obtido com as tabelas de distribuição assintótica e superfícies de resposta de aproximação numérica desenvolvidas por MacKinnon, rejeitando-se a presença de raiz unitária caso o valor- p se posicione abaixo do nível de significância de 5% ($\alpha = 0,05$).

O Quadro 5 reúne as métricas extraídas da aplicação do teste sobre a série original de roubos de carga. O valor- p apurado de 0,42 indicou a impossibilidade de rejeitar a hipótese nula de raiz unitária, confirmando que a sequência bruta de crimes manifesta um comportamento não estacionário e não converge espontaneamente para uma média histórica estável.

Quadro 5 – Resultados do Teste Aumentado de Dickey-Fuller (Série Bruta).

Métrica Analítica	Valor Observado	Critério de Decisão ($\alpha = 0,05$)	Conclusão do Teste
Estatística ADF (τ)	-1,73	Rejeita-se H_0 se $\tau < \text{Valor Crítico}$ (5%)	A estatística observada é superior ao valor crítico.
Valor- p de MacKinnon	0,42	Rejeita-se H_0 se $\text{Valor-}p < 0,05$	A hipótese nula não foi rejeitada (Série não estacionária).
Valor Crítico (1%)	-3,48		
Valor Crítico (5%)	-2,88	$-1,73 \not< -2,88$	

Fonte: Elaborado pelo autor (2026).

Para estabilizar a série e atender às premissas de modelagem do componente

linear, aplicou-se a transformação de diferenciação de primeira ordem, que corresponde à derivada discreta da série temporal. Se a trajetória original manifesta uma tendência de crescimento linear, a execução de sua derivada discreta resulta em uma série estável cujo valor médio é deslocado para uma constante que representa a taxa de crescimento por período, também denominada componente de deriva. Essa transformação não anula o sentido físico ou a interpretação do fenômeno econômico, uma vez que a taxa constante de variação é preservada no intercepto da equação diferencial, permitindo que os polinômios autoregressivos e de médias móveis modelem as flutuações residuais estacionárias de curto prazo de forma rigorosa.

Após a aplicação da derivada discreta com o parâmetro de integração igual a um ($d = 1$), submeteu-se a nova série diferenciada ao teste de Dickey-Fuller. A estatística apurada deslocou-se para o valor de $-12,71$, resultando em um valor- p de MacKinnon igual a $0,00$. Como o indicador de probabilidade posicionou-se abaixo do limite crítico estabelecido, a hipótese nula de raiz unitária foi formalmente rejeitada, atestando que a derivada discreta eliminou a tendência estocástica e conferiu a estacionariedade necessária para prosseguir com a identificação dos parâmetros autorregressivos do modelo. Na etapa final de projeção, o operador do modelo executa a operação inversa à diferenciação, realizando a reintegração dos valores previstos para reconstituir a tendência de crescimento e a escala nominal original dos boletins de ocorrência.

4.3 Abordagens Neurais Não Lineares

Como a modelagem estatística linear univariada mostrou-se limitada para capturar a volatilidade e as quebras estruturais presentes na série histórica de roubos de carga, tornou-se necessária a transição para modelos não lineares baseados em aprendizado profundo. Optou-se pela implementação da arquitetura baseada em Unidades Recorrentes Fechadas (GRU) (CHO et al., 2014).

A escolha da GRU em detrimento das redes de Longa Memória de Curto Prazo (LSTM) justifica-se pelo seu menor custo computacional e maior velocidade de convergência, uma vez que a GRU unifica as portas de esquecimento e entrada em

uma única porta de atualização, operando com um número reduzido de parâmetros que atenua o risco de sobreajuste (*overfitting*) em séries temporais de dimensão macro. Da mesma forma, afastou-se o uso de arquiteturas baseadas em mecanismos de atenção total (*Transformers*), visto que esses modelos carecem de viés indutivo temporal e exigem volumes massivos de dados para otimizar suas matrizes de projeção, tornando-se inadequados para o processamento de sequências contínuas curtas de escala mensal.

A dinâmica matemática interna da célula GRU processa o vetor de entrada do mês atual x_t e o vetor de estado oculto anterior h_{t-1} por meio de um sistema de equações que determina o fluxo da informação ao longo do tempo. O estágio inicial calcula a porta de reinicialização r_t e a porta de atualização z_t pelas seguintes formulações:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (4.2)$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (4.3)$$

nas quais W_r e W_z representam as matrizes de pesos de cada porta e σ corresponde à função de ativação sigmoide, responsável por mapear os valores estritamente para o intervalo aberto entre zero e um para calibrar a intensidade do fluxo. Na sequência, gera-se o estado oculto candidato $\tilde{h}_t$, cuja ativação utiliza a tangente hiperbólica ($\tanh$) para comprimir a amplitude dos dados entre os limites de menos um e mais um:

$$\tilde{h}_t = \tanh(W \cdot [r_t \odot h_{t-1}, x_t]) \quad (4.4)$$

em que o operador $\odot$ expressa o produto de Hadamard para multiplicação elemento a elemento, permitindo que a porta de reinicialização decida quanta informação do passado será descartada. Por fim, o vetor de estado oculto final h_t é consolidado pela interpolação linear controlada pela porta de atualização:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (4.5)$$

4.3.1 Topologia da Rede e Dinâmica de Ativação

Para assegurar uma comparação metodologicamente justa, configurou-se uma topologia sequencial idêntica para os experimentos univariado e multivariado,

modificando-se exclusivamente a dimensão do tensor de entrada. Essa estrutura atua como um funil de processamento em que a camada recorrente inicial extrai os componentes temporais e as camadas totalmente conectadas subsequentes projetam esses recursos em subespaços de maior dimensionalidade para mapear as interações não lineares. O Quadro 6 detalha a distribuição das camadas e a função computacional de cada nível da rede.

Quadro 6 – Arquitetura Sequencial das Redes Recorrentes (Univariada e Multivariada).

#	Tipo de Camada	Neurônios	Função e Observação Computacional
1	Input (Entrada)	-	Geometria (3, 1) na Univariada e (3, 2) na Multivariada.
2	GRU Recorrente	64	Extração de dependências e memória sequencial de longo prazo.
3	Dense 1 (Densa)	32	Primeira aproximação não linear e combinação de recursos.
4	Dense 2 (Densa)	32	Aprofundamento do mapeamento e ajuste de alta dimensão.
5	Saída (Dense)	1	Camada linear para projeção do valor escalar contínuo.

Fonte: Elaborado pelo autor (2026).

A integração das funções de ativação ao longo da topologia obedece aos critérios de estabilização dos gradientes e modelagem do domínio numérico do problema. Conforme sintetizado no Quadro 7, as ativações não lineares internas da célula GRU gerenciam a retenção de dados, enquanto as camadas densas intermediárias utilizam a Unidade Linear Retificada (ReLU), expressa por $f(x) = \max(0, x)$, para introduzir esparsidade e afastar o fenômeno do desaparecimento

do gradiente durante o ajuste dos pesos. Na camada final de regressão, aplica-se a ativação linear, escolha que permite à rede estimar valores absolutos contínuos sem impor limites máximos artificiais que prejudicariam a capacidade do modelo de mensurar picos severos de criminalidade.

Quadro 7 – Dinâmica das Funções de Ativação e Implicações Algorítmicas.

Camada Alvo	Função de Ativação	Mecanismo de Atuação no Treinamento
Célula GRU	Sigmoide e Tangente Hiperbólica	Atuam nas portas internas limitando os valores nos intervalos $[0, 1]$ e $[-1, 1]$ para suavizar o fluxo do erro.
Camadas Densas	Unidade Linear Retificada (ReLU)	Zera respostas negativas para acelerar a convergência e evitar o travamento dos neurônios.
Camada de Saída	Linear ($f(x) = x$)	Projeta a saída diretamente no espaço dos números reais contínuos para o cálculo das ocorrências.

Fonte: Elaborado pelo autor (2026).

4.3.2 Arquitetura Univariada e o Efeito de Suavização

O primeiro experimento consistiu no treinamento da rede utilizando unicamente o vetor cronológico histórico de roubos de carga. Os dados de entrada foram estruturados na geometria tridimensional padronizada de tensores no formato (amostras, 3, 1), arranjo matricial que organiza as sequências temporais para o processamento paralelo em redes recorrentes. O Quadro 8 detalha a função estrutural, as nomenclaturas técnicas e os valores exatos assumidos por cada um dos três eixos coordenados desse tensor ao longo da modelagem univariada da série bruta.

Para compreender a aplicação prática dessa geometria tridimensional sobre o banco de dados manipulado, considera-se o exemplo do mapeamento executado

Quadro 8 – Estrutura e parametrização geométrica do tensor de entrada univariado.

Eixo	Componente do Tensor	Dimensão	Descrição e Papel Operacional
0	Amostras (<i>Samples</i>)	129	Quantidade de fatias temporais ou janelas deslizantes completas extraídas da série histórica.
1	Passos de Tempo (<i>Time Steps</i>)	3	Extensão da janela retroativa de observação (<i>lookback window</i>) adotada como memória.
2	Atributos (<i>Features</i>)	1	Quantidade de variáveis preditoras simultâneas em uso (contagem mensal de crimes).

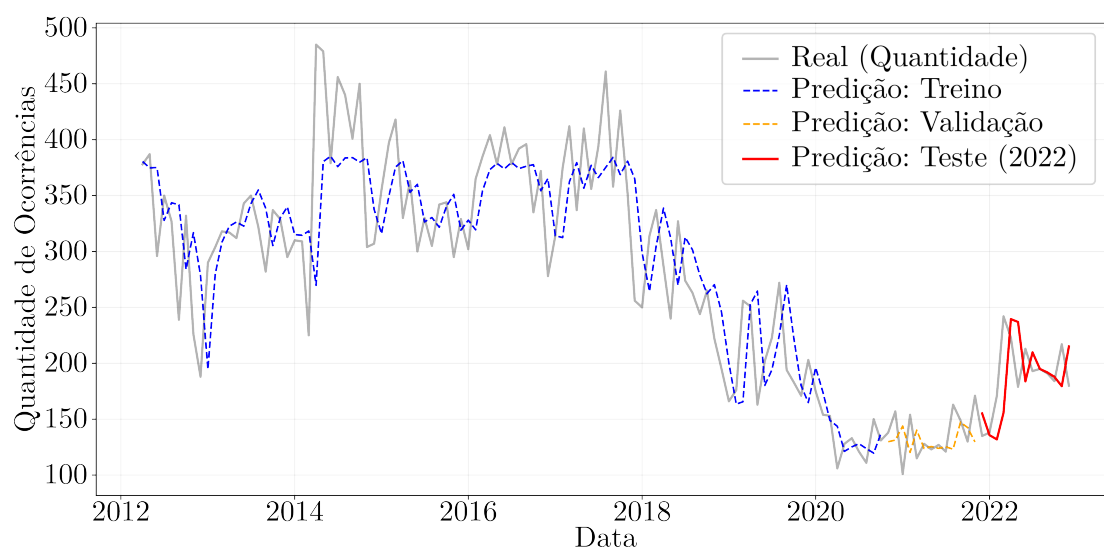
Fonte: Elaborado pelo autor (2026).

para o quarto mês da série histórica, correspondente a abril de 2012. Como a extensão dos passos de tempo foi fixada em três meses ($w = 3$), o algoritmo realiza uma varredura retrógrada e monta a estrutura preditora combinando ordenadamente o volume de roubos registrados em janeiro, fevereiro e março de 2012. Esse bloco de dados sequenciais constitui a primeira amostra do conjunto, estruturada computacionalmente como uma pequena matriz isolada de três linhas e uma coluna. O alvo preditivo associado a essa linha específica é a contagem absoluta de crimes ocorridos no próprio mês de abril de 2012.

O algoritmo executa esse deslocamento iterativo avançando um passo de cada vez ao longo dos 132 meses totais que compõem os onze anos da série. Como os três primeiros meses da sequência cronológica funcionam obrigatoriamente como a janela inicial de memória histórica, a extração gera de forma exata 129 amostras tridimensionais independentes conectadas aos seus respectivos alvos futuros. Esse tensor consolidado alimenta a camada recorrente primária, e o processo de otimização utiliza o Erro Quadrático Médio (MSE) como função de perda para guiar o cálculo dos gradientes no algoritmo de retropropagação, ajustando as matrizes de pesos internos a cada iteração de treino.

Os resultados obtidos no modelo univariado, ilustrados na Figura 9, evidenciam o comportamento clássico de redes recorrentes expostas a séries temporais de alta volatilidade. Na ausência de variáveis exógenas reguladoras, o modelo tende a adotar uma postura conservadora, gerando uma curva suavizada que converge para a média histórica da série. Essa limitação estrutural impede a rede de antecipar picos repentinos de criminalidade ou de se ajustar com precisão às cristas da série histórica.

Figura 9 – Aderência da previsão com uma única variável (GRU). O modelo acompanha a direção geral dos dados, mas apresenta um atraso temporal (deslocamento) e tem dificuldade para alcançar os valores dos picos mais altos.

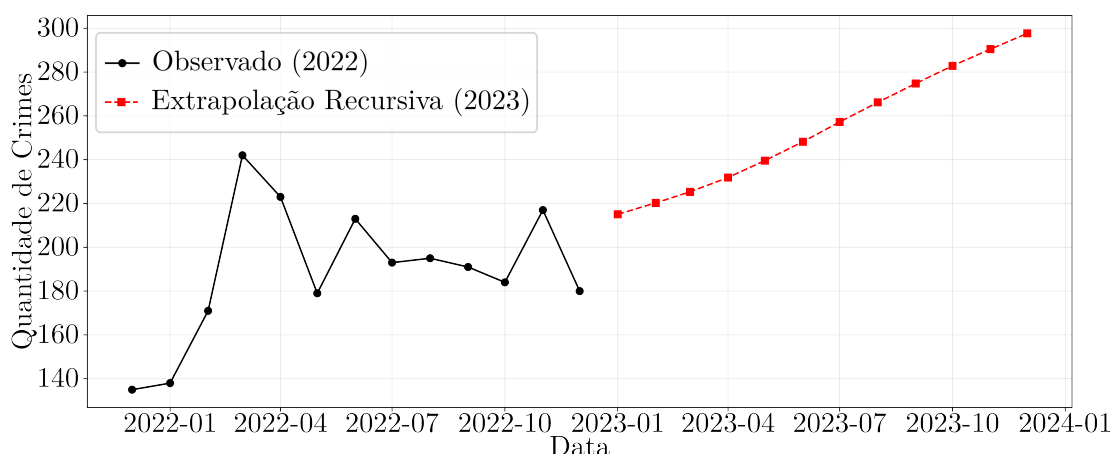


Fonte: Elaborada pelo autor (2026).

Adicionalmente, a análise visual revela um acentuado atraso de fase na curva estimada. Embora o perfil da predição pareça mimetizar o comportamento dos dados reais, a rede opera sob um efeito de deslocamento temporal em que a previsão para o mês seguinte assume aproximadamente o valor verificado no mês anterior. Esse fenômeno demonstra que o modelo univariado não realiza uma antecipação real da tendência, mas comporta-se de forma puramente reativa.

A distribuição cronológica estrita adotada na divisão dos dados destinou os últimos 27 meses (outubro de 2020 a dezembro de 2022) para o conjunto de teste, intervalo marcado por uma fase estável e de baixos índices criminais decorrente da retração pós-pandêmica e de melhorias na segurança privada. Embora o teste em um período de calmaria possa sugerir um desbalanceamento, essa configuração metodológica constitui um teste rigoroso de capacidade de generalização. Expor o algoritmo treinado em fases de alta criminalidade a um cenário de estabilização por baixa comprova se a rede neural extraiu as propriedades matemáticas intrínsecas do fenômeno ou se apenas memorizou as oscilações passadas.

Figura 10 – Previsão de longo prazo para 2023 com uma única variável. Os valores previstos se aproximam rapidamente da média histórica devido à falta de outras variáveis.



Fonte: Elaborada pelo autor (2026).

Para avaliar o desempenho do modelo fora da amostra de treino, calculou-se o Erro Percentual Médio Absoluto (MAPE), que resultou no patamar de 17,04% (Figura 10). A introdução do MAPE ao lado do MSE justifica-se por critérios metodológicos distintos: enquanto o MSE atua estritamente como a função de perda diferenciável para o ajuste dos gradientes, o MAPE é aplicado como métrica de avaliação por sua propriedade de invariância de escala, traduzindo o desvio em termos percentuais relativos ao volume real e facilitando a interpretação gerencial do

erro. Os resultados ratificam que a ausência de dados geográficos limita o horizonte preditivo da rede univariada, confinando suas estimativas à reprodução linear do tempo (WANG et al., 2024a).

4.3.3 Arquitetura Multivariada: O Espaço como Regulador

Com o objetivo de anular o atraso de fase característico do modelo univariado e capturar a dinâmica territorial do delito, o tensor de entrada foi expandido para o formato (amostras, 3, 2). Essa alteração metodológica incorpora a distância média mensal até as bases policiais rodoviárias como a segunda variável preditora da rede neural. A intenção dessa configuração é introduzir a proximidade da infraestrutura estatal como um elemento de regularização espacial, forçando as portas de controle da rede GRU a ponderarem conjuntamente a cronologia e a geografia do risco (LIU et al., 2022). O Quadro 9 descreve a organização interna, as especificações geométricas e o papel analítico de cada um dos três eixos que constituem essa nova estrutura tridimensional.

Quadro 9 – Estrutura e parametrização geométrica do tensor de entrada multivariado.

Eixo	Componente do Tensor	Dimensão	Descrição e Papel Operacional
0	Amostras (<i>Samples</i>)	129	Quantidade de blocos temporais deslizantes calculados sobre a série histórica consolidada.
1	Passos de Tempo (<i>Time Steps</i>)	3	Tamanho da janela de retroobservação (<i>lookback window</i>) mantida como memória histórica.
2	Atributos (<i>Features</i>)	2	Variáveis preditoras simultâneas em uso (volume mensal de crimes e distância policial escalonada).

Fonte: Elaborado pelo autor (2026).

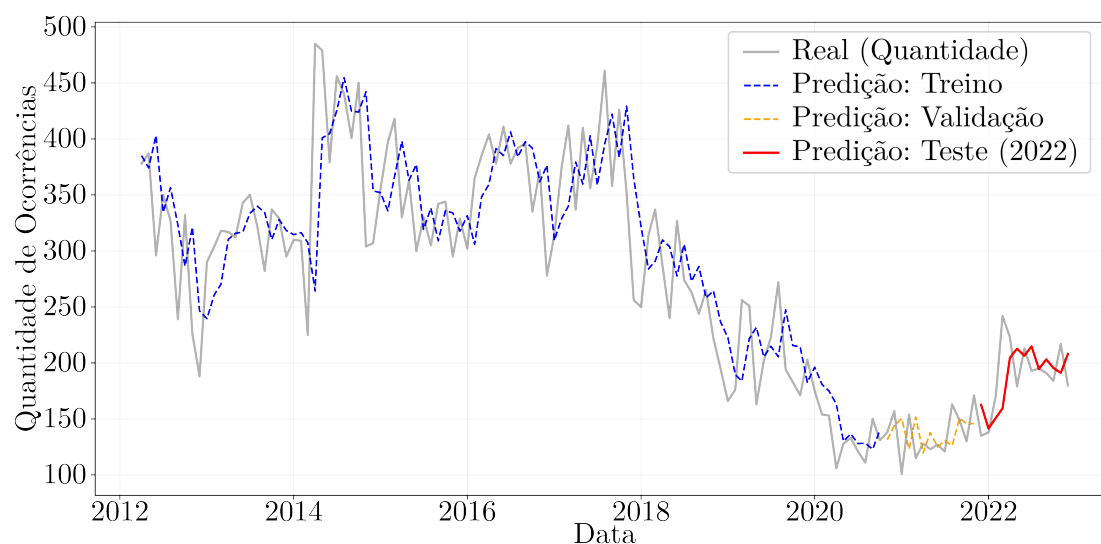
Para demonstrar a mecânica de alimentação dessa estrutura multivariada sobre a base de dados real, toma-se novamente o exemplo da previsão formulada para o mês de abril de 2012. Mantendo-se a janela de memória de três meses ($w = 3$), o algoritmo executa a varredura retrospectiva nos meses de janeiro, fevereiro e março de 2012. Contudo, diferentemente do arranjo univariado, cada um desses meses passados agora fornece um vetor composto por dois elementos numéricos ordenados: a contagem absoluta de roubos daquele período e o valor correspondente da distância geodésica média escalonada no intervalo $[0, 1]$. Essa combinação gera uma matriz tridimensional compacta de três linhas por duas colunas para cada passo da série, vinculada estritamente ao alvo contínuo do mês de destino, que permanece representado pela quantidade isolada de crimes ocorridos em abril de 2012.

A execução iterativa dessa rotina computacional processa sequencialmente as 129 amostras estruturadas, preservando o ordenamento cronológico e geográfico dos dados. A inclusão dessa variável espacial alterou significativamente a dinâmica de aprendizado do modelo, conforme ilustrado no diagnóstico preditivo da Figura 11. A presença do atributo geográfico atuou como um fator de quebra da colinearidade temporal pura, eliminando o atraso de fase (*lag shift*) que reduzia a utilidade prática da abordagem univariada. No período de validação por teste, a curva multivariada pode manifestar um alinhamento visual menos colado a cristas e picos isolados se comparada ao experimento univariado anterior; no entanto, essa mudança reflete a correção do erro de atraso. Enquanto o modelo univariado simulava aderência por simplesmente replicar o comportamento passado com um passo de atraso, a formulação multivariada alcança uma antecipação real da tendência ao processar as restrições físicas e institucionais do território (ARANSIOLA; JUSTUS; CECCATO, 2023).

A superação das limitações temporais torna-se evidente no confronto direto entre as duas abordagens apresentado na Figura 12. A inserção da variável de proximidade institucional destravou a capacidade da rede de rastrear as mudanças de comportamento da série histórica que os modelos focados estritamente na cronologia pura atrasavam ou suavizavam.

O Quadro 10 consolida os índices estatísticos de validação das duas ar-

Figura 11 – Predição do modelo RNN Multivariado. A inclusão da matriz espacial permitiu que o modelo acompanhasse com maior fidelidade as flutuações reais da série.



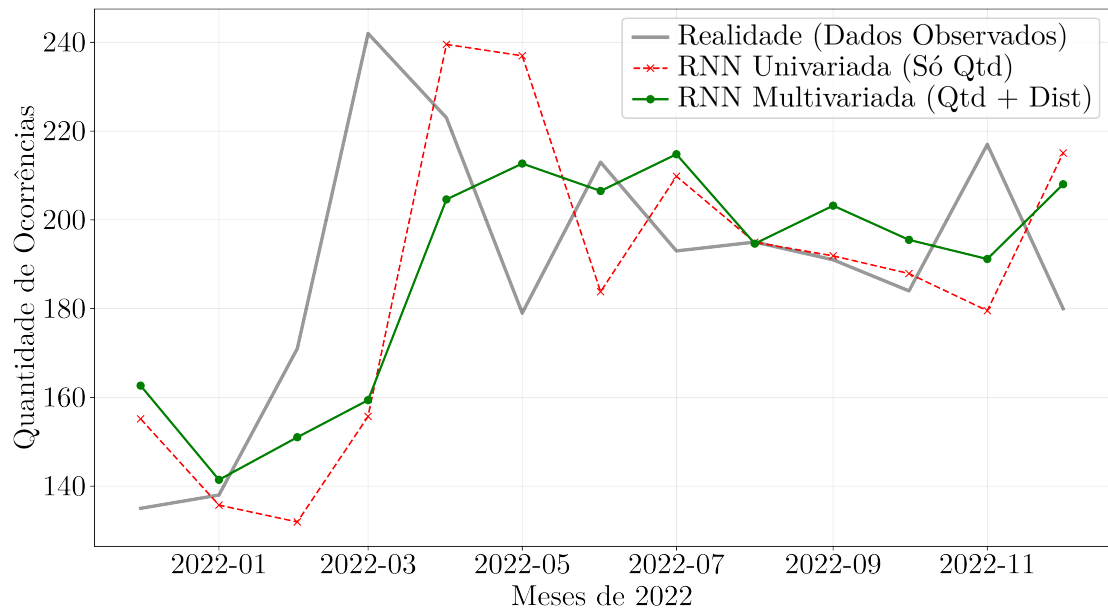
Fonte: Elaborada pelo autor (2026).

quiteturas neurais. Adotou-se o Erro Absoluto Médio (MAE) como critério de corte principal por sua robustez analítica na presença de valores extremos, visto que o MAE não eleva o erro ao quadrado e evita a distorção das métricas por desvios residuais isolados (CHAI; DRAXLER, 2014). A injeção do componente de infraestrutura policial reduziu o MAPE para o patamar de 10,91%, representando um ganho de precisão de 35,97% em relação à base temporal pura e comprovando matematicamente que a distribuição física das forças de segurança regula a variação do roubo de cargas.

4.4 Avaliação do Modelo Híbrido (SARIMA-GRU)

Estudos sobre previsão de consumo de energia e mercado financeiro indicam bons resultados com o uso de modelos híbridos de decomposição (XIAO et al., 2024). Testamos essa estratégia aplicando o modelo SARIMA para capturar a tendência

Figura 12 – Comparativo visual de aderência: Univariada vs. Multivariada. A variável geográfica destrava a capacidade da rede de modelar a volatilidade aguda.



Fonte: Elaborada pelo autor (2026).

Quadro 10 – Performance preditiva (2012–2022): Univariada vs. Multivariada.

Métrica de Avaliação	GRU Univariada	GRU Multivariada	Variação Relativa
MAE	31,01	21,41	–30,96%
MAPE	17,04%	10,91%	–35,97%

Fonte: Elaborado pelo autor (2026).

de longo prazo do tempo, enquanto a rede GRU ficou responsável por prever os resíduos (os erros que o método estatístico linear não consegue explicar) (ZHENG et al., 2021).

Definição 4.1 (Predição Híbrida Agregada). A previsão final ($\hat{Y}_t$) é obtida pela soma entre a tendência estatística e a correção feita pela rede neural:

$$\hat{Y}_t = \max(0, \hat{L}_t + \hat{N}_t) \quad (4.6)$$

em que $\hat{L}_t$ corresponde à projeção gerada pelo SARIMA e $\hat{N}_t$ é a projeção gerada pela GRU. Adicionamos a função limite $\max$ porque não existem quantidades negativas de roubos na realidade física.

Para realizar o experimento, calculamos a diferença entre os dados reais e as previsões do SARIMA. Esses erros foram usados como entrada para a rede neural. Para evitar que o modelo apenas memorizasse ruídos aleatórios, utilizamos o mecanismo de parada antecipada (*Early Stopping*) no **TensorFlow**, que interrompeu o treinamento na 19ª época ao notar que o desempenho no teste parou de melhorar. Isso evita o superajuste dos dados (*overfitting*) (PRECHELT, 2012).

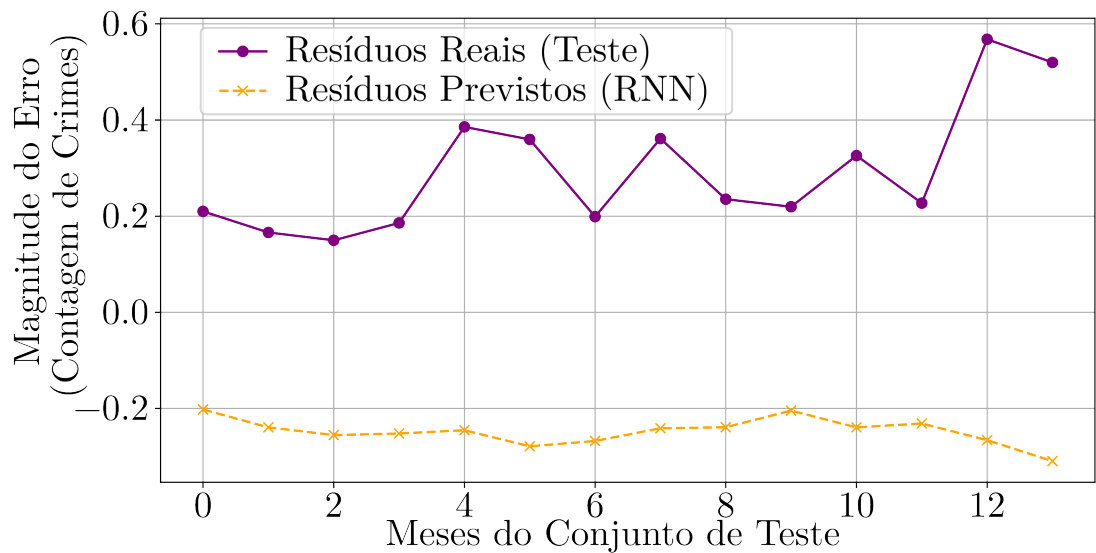
A Figura 13 mostra o comportamento da rede recorrente ao analisar os erros gerados pelo SARIMA. O modelo buscou mapear apenas os desvios deixados pelo método estatístico, o que torna clara a dificuldade de prever variações aleatórias (ruídos) quando elas estão separadas do volume total de crimes que as originou.

4.5 Consolidação e Diagnóstico Preditivo

Os experimentos realizados fornecem um diagnóstico claro sobre o comportamento dos modelos aplicados à previsão do roubo de cargas. O Quadro 11 resume o desempenho, as taxas de erro e a viabilidade de cada método testado.

A comparação traz um resultado inesperado: o modelo híbrido (elogiado em estudos de engenharia elétrica e finanças (ZHENG et al., 2021)) apresentou o pior desempenho em nossa base de dados, com um erro médio próximo a 66%. Isso confirma que remover a tendência dos dados por meio de métodos estatísticos tradicionais retira o contexto necessário para o aprendizado da rede neural. Ao

Figura 13 – Processamento residual. A rede tenta rastrear os desvios estocásticos (eixo Y) ao longo do tempo (eixo X), enfrentando alta dispersão informacional.



Fonte: Elaborada pelo autor (2026).

Quadro 11 – Diagnóstico de Desempenho por Arquitetura Preditiva.

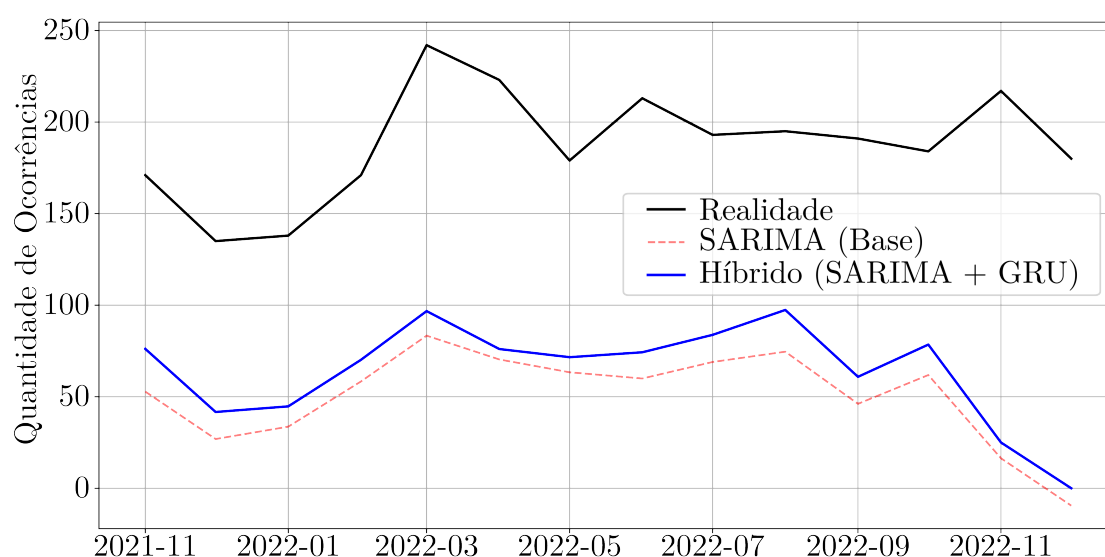
Arquitetura	MAPE	Observação Analítica
SARIMA	> 21%	Adequado para descrever a tendência de longo prazo, mas não se adapta a mudanças rápidas de curto prazo.
Híbrido (SARIMA+GRU)	65,91%	Maior índice de erro. A separação dos resíduos fez o modelo perder a visão do conjunto dos dados (XIAO et al., 2024).
GRU Univariada	17,04%	Identificou a direção geral dos dados, mas suavizou as previsões aproximando-as da média histórica e ignorou os picos.
GRU Multivariada (ST-Crime)	10,91%	Maior precisão do teste. A inclusão da variável de distância corrigiu o atraso temporal da rede.

Fonte: Elaborado pelo autor (2026).

analisar apenas os resíduos isolados, o modelo processa uma sequência de variações sem padrão claro (ARANSIOLA; JUSTUS; CECCATO, 2023).

A Figura 14 mostra a falha de previsão do modelo híbrido. A soma da tendência estatística com a estimativa dos resíduos gera uma curva instável, que não acompanha o comportamento real do roubo de cargas ao longo do tempo.

Figura 14 – Desempenho do preditor híbrido (SARIMA + GRU). A linha azul demonstra a instabilidade causada pela tentativa de prever resíduos isolados.



Fonte: Elaborada pelo autor (2026).

Por outro lado, o desempenho superior e estável da Rede Recorrente Multivariada comprova a importância de usar dados do espaço nas previsões. Os resultados mostram que a localização e a proximidade das bases policiais são fundamentais para prever os roubos. Isso confirma que o crime depende da estrutura geográfica do território, e não apenas do histórico do tempo.

4.6 Considerações Finais do Capítulo

Neste capítulo, analisamos o comportamento do roubo de cargas ao longo do tempo. Os testes provaram que tentar prever crimes usando apenas modelos temporais, sem considerar o espaço geográfico, prejudica o aprendizado da rede neural e aumenta os erros de previsão (LIU et al., 2022; WANG et al., 2024a).

O alto índice de erro do modelo híbrido (SARIMA-GRU) traz uma importante lição técnica e social. Computacionalmente, ele mostra que os dados criminais são complexos e não podem ser divididos em blocos isolados. Sob a ótica da criminologia crítica, esse resultado confirma as análises de Baratta (1982): o crime organizado nas redes de transporte não é um problema isolado ou um desvio passageiro que possa ser separado da sociedade. Ele funciona de forma integrada e aproveita a própria estrutura da economia e da logística legal.

O melhor desempenho do modelo multivariado também expõe as limitações das estratégias tradicionais de segurança pública. A superioridade dessa abordagem não se resume a um decréscimo marginal nas métricas absolutas de erro, mas reside em uma correção estrutural qualitativa: a injeção da distância rodoviária atuou como um operador de regularização espacial. Esse componente quebrou a colineariedade temporal pura e eliminou o atraso de fase (*lag shift*) verificado no experimento univariado. Enquanto a rede univariada simulava aderência por simplesmente replicar o passado de forma reativa, o modelo multivariado utiliza a âncora geográfica para estabilizar a variância do sistema, impedindo que o algoritmo gere tendências explosivas ou linearmente suavizadas durante as extrapolações recursivas de longo prazo.

Esse comportamento contraria a Teoria da Dissuasão (NAGIN, 2013). Como explicam Rusche e Kirchheimer (2004), a presença física da polícia não impede a ação de organizações criminosas estruturadas. Pelo contrário, as quadrilhas calculam a distância até o policiamento como mais um fator de risco ou custo operacional em seu modelo de negócios (ARANSIOLA; JUSTUS; CECCATO, 2023).

Para a atuação jurídica do Ministério Público, como titular da ação penal, a margem de erro estável do modelo (em torno de 10%) funciona como uma prova importante. A regularidade com que esses crimes acontecem demonstra que

as quadrilhas não agem por desespero individual ou impulsos passageiros, mas operam com um planejamento logístico contínuo. Diante disso, o sistema de justiça criminal não deve enquadrar o roubo de cargas rodoviárias como um crime de roubo comum praticado em concurso de pessoas (Artigo 157, §2º, do Código Penal). O correto é aplicar a tipificação de Organização Criminosa (Lei 12.850/2013). Esse enquadramento jurídico fornece os instrumentos necessários para investigar e bloquear os fluxos financeiros do mercado clandestino de mercadorias.

Embora o modelo multivariado tenha trazido avanços, ele possui uma limitação técnica. Incluir a distância como um número simples na entrada da rede impede o modelo de compreender o traçado real e a conectividade da malha rodoviária, além de ignorar como o risco de crimes se espalha entre cidades vizinhas. Para superar essa barreira, o próximo capítulo investigará o uso de Redes Neurais Convolucionais em Grafos Espaço-Temporais (A3T-GCN / STGNN) (BAI et al., 2021). O objetivo é fazer com que o modelo processe diretamente as conexões e a estrutura geométrica das rodovias paulistas.

5 Predição Avançada com Redes Neurais em Grafos Espaço-Temporais

O capítulo anterior demonstrou que analisar o tempo e o espaço de forma isolada limita a precisão das previsões. Este capítulo integra ambas as dimensões por meio de Redes Neurais em Grafos Espaço-Temporais (*Spatio-Temporal Graph Neural Networks* — STGNNs). Para facilitar a compreensão da tecnologia utilizada, o modelo pode ser dividido entre gênero e espécie. O gênero é a STGNN, que representa uma grande família de arquiteturas matemáticas desenvolvidas para mapear sistemas que mudam no tempo e no espaço simultaneamente. A espécie é a Rede Neural Convolutacional em Grafo com Atenção Temporal (*Attention Temporal Graph Convolutional Network* — A3T-GCN) (BAI et al., 2021), que identifica o modelo matemático exato e as equações aplicadas neste trabalho.

O objetivo desta etapa é demonstrar que o roubo de cargas não ocorre de maneira isolada no território, mas se espalha por uma rede rodoviária conectada. Ao representar a malha de transportes do Estado de São Paulo na forma de um grafo, propõe-se um modelo preditivo capaz de rastrear a circulação do risco criminal. Essa resposta estruturada serve para subsidiar o planejamento tático da segurança pública e auxiliar no combate às bases financeiras e logísticas do crime organizado.

5.1 Modelagem Relacional do Território

Até este estágio da pesquisa, a atividade criminal do roubo de cargas foi tratada sob a perspectiva de séries temporais independentes. No Capítulo 4, evidenciou-se a capacidade das unidades recorrentes fechadas em extrair a ciclicidade e as quebras de tendência do delito. Contudo, essa formulação univariada e multivariada esbarrou em um limite geométrico, visto que os algoritmos calculam o risco de cada município de forma isolada, omitindo a conectividade e a influência mútua existentes entre os nós da rede rodoviária.

Na realidade da logística criminal, o deslocamento dos grupos organizados não respeita as divisões político-administrativas das fronteiras municipais. O delito manifesta-se como um fluxo que se propaga ao longo da infraestrutura física de transportes. Um incremento na densidade de ocorrências no município de Campinas, por exemplo, altera a distribuição de probabilidade do risco em localidades vizinhas e economicamente integradas, como Jundiaí e Sumaré, utilizando como canais de difusão os complexos viários Anhanguera e Bandeirantes.

Para modelar o risco criminal com base nessas restrições de conectividade viária, adota-se a metodologia das Redes Neurais em Grafos (*Graph Neural Networks* — GNNs) (KIPF; WELLING, 2017; ROSHANKAR; KEYVANPOUR, 2023). Sob essa abordagem, os municípios deixam de ser interpretados como unidades isoladas e passam a constituir os elementos de uma estrutura topológica interconectada. A inclusão da malha de transportes no treinamento do algoritmo permite que o cálculo do risco de cada nó incorpore o comportamento dos vizinhos adjacentes na rede, superando as limitações do espaço euclidiano tradicional (SHAO et al., 2024).

Formalmente, o território do estado é modelado como um grafo não direcionado $G = (V, E, A)$. O conjunto V representa os vértices ou nós do sistema, sendo definido por $V = \{v_1, v_2, \dots, v_N\}$, em que $N = 645$ corresponde à totalidade dos municípios paulistas inseridos na matriz analítica. O conjunto E descreve as arestas ou arcos que interconectam os nós, estabelecidos pela presença de rodovias estaduais ou federais que cruzam as divisas municipais de forma direta. A estrutura de conectividade do grafo é codificada pela matriz de adjacência espacial $A \in \mathbb{R}^{N \times N}$, cujos elementos a_{ij} determinam a existência de relação de vizinhança entre as localidades i e j . Se dois municípios compartilham uma rota logística rodoviária direta, atribui-se o valor unitário ao elemento correspondente ($a_{ij} = 1$), enquanto a ausência de conexão viária direta é mapeada pelo valor nulo ($a_{ij} = 0$), fixando-se a diagonal principal da matriz em zero para anular a autoadjacência espontânea.

5.2 Construção da Malha Logística: A Matriz de Adjacência

A modelagem espacial da atividade criminal exige a conversão da geografia física do Estado de São Paulo em uma estrutura matemática relacional de suporte.

A malha logística é formalizada como um grafo não direcionado $G = (V, E, A)$, no qual V constitui o conjunto de vértices ou nós, representado por $N = 289$ municípios selecionados por manifestarem recorrência estatística do delito, critério adotado para isolar os polos de risco e afastar o fenômeno da esparsidade extrema de dados na rede neural. O conjunto E define as arestas ou arcos que indicam a existência de conexão viária direta entre as unidades espaciais, enquanto $A \in \mathbb{R}^{N \times N}$ expressa a matriz de adjacência ponderada responsável por mensurar a intensidade da ligação física e a acessibilidade entre cada par de cidades.

O estabelecimento dos pesos contidos na matriz de adjacência baseia-se em uma função de decaimento capaz de modelar o princípio de que a influência de uma ocorrência criminal diminui com o aumento da distância geográfica. Para isso, computou-se inicialmente a distância euclidiana em linha reta entre as coordenadas de latitude e longitude correspondentes aos centroides geométricos dos polígonos municipais, extraídos diretamente da malha digital oficial de municípios do Instituto Brasileiro de Geografia e Estatística (IBGE) referente ao ano de 2024. A opção pela distância linear entre os centroides do *shapefile* oficial, em detrimento do traçado viário real extraído de servidores de mapas rodoviários públicos, fundamenta-se em restrições de custo computacional e limitações de requisições a APIs externas. O cálculo de rotas viárias exatas para matrizes completas de origem e destino exige custos financeiros elevados em plataformas comerciais ou demanda tempos de processamento inviáveis devido aos atrasos obrigatórios introduzidos para evitar o bloqueio de servidores de código aberto. A aplicação da distância em linha reta a partir dos centroides oficiais, quando submetida a processos de normalização global por amplitude, fornece uma aproximação estatisticamente estável, simétrica e escalável da proximidade regional.

Para converter a matriz de distâncias normalizadas em coeficientes relacionais de proximidade, aplicou-se uma função de decaimento baseada no Kernel Gaussiano (SHUMAN et al., 2013; GU et al., 2023). O preenchimento de cada elemento $W_{i,j}$ da matriz de adjacência ponderada obedece à formulação matemática definida por:

$$W_{i,j} = \begin{cases} \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right), & \text{se } \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \geq \tau \\ 0, & \text{caso contrário} \end{cases} \quad (5.1)$$

em que d_{ij} representa a distância euclidiana normalizada entre os centroides das cidades i e j . O parâmetro de largura de banda do kernel foi fixado no valor de $\sigma = 0,1$, atuando como o fator de escala que calibra a sensibilidade e o alcance da zona de influência geográfica das manchas térmicas. O limiar de corte para esparsificação foi estabelecido em $\tau = 0,1$, índice encarregado de isolar e zerar as conexões fracas ou excessivamente distantes, forçando o modelo a concentrar o aprendizado nas interações de vizinhança de maior impacto estrutural.

Os elementos posicionados na diagonal principal da matriz foram preenchidos com valores nulos ($W_{i,i} = 0$), procedimento que elimina a autoajacência espontânea e impede que as operações convolucionais espaciais gerem auto-loops redundantes durante a agregação de mensagens nos nós. Esse processo de calibração gerou uma topologia de rede com densidade espacial estável de aproximadamente 24%, configuração em que cada município possui uma média de 68 nós vizinhos conectados diretamente. No contexto do roubo de cargas, essa ampla conectividade em rede é condizente com a realidade operante das organizações criminosas, cujas ações de transbordo e rotas de fuga utilizam eixos rodoviários integrados que interconectam múltiplos municípios de maneira célere.

5.3 A Arquitetura A3T-GCN

Para analisar a estrutura do território e a variação do tempo de forma conjunta, esta tese aplica a Rede Neural Convolucional em Grafo com Atenção Temporal (*Attention Temporal Graph Convolutional Network* — A3T-GCN) (BAI et al., 2021). Apresentar essa arquitetura neste estágio da pesquisa faz parte de uma estratégia de desenvolvimento por etapas. Para avaliar o desempenho de um modelo espaço-temporal complexo, é necessário construir primeiro modelos de base mais simples para servir de comparação. Sem o teste isolado do tempo com o SARIMA e da proximidade geográfica com a GRU multivariada, não seria possível medir o ganho real de precisão trazido pela modelagem em grafos. Por ser um método especializado, detalha-se o funcionamento de suas equações para demonstrar como ele se ajusta ao problema do roubo de cargas. O modelo opera combinando três componentes em uma única sequência de aprendizado contínuo. O

primeiro componente realiza convoluções em grafos para extrair as características da malha rodoviária. O segundo utiliza Unidades Recorrentes Fechadas (GRU) para acompanhar a tendência do tempo. O terceiro componente aplica um mecanismo de atenção adaptativo, que calcula a importância de diferentes meses do passado para prever a distribuição do risco no futuro.

5.3.1 A Estrutura do Tensor Espaço-Temporal

Para processar as dimensões de espaço e tempo de forma integrada, transformamos os registros de ocorrências em um bloco de dados tridimensional, chamado de Tensor Espaço-Temporal e representado por $X \in \mathbb{R}^{N \times F \times T}$. Nessa estrutura, o primeiro eixo corresponde ao número de nós $N = 289$. Esse valor representa os municípios paulistas que registraram roubos de carga no período analisado. As cidades que não tiveram nenhuma ocorrência ao longo da série histórica foram removidas da análise para evitar dados vazios (*sparsity*) e impedir que o ruído estatístico prejudicasse os cálculos do modelo. O segundo eixo mostra a quantidade de atributos $F = 1$, que indica o uso do volume mensal de crimes normalizados. O terceiro eixo define o alcance da memória temporal da rede, fixado em $T = 12$ meses (Quadro 12).

A escolha de usar uma janela histórica de doze meses baseia-se em motivos computacionais e no próprio comportamento do crime organizado. No planejamento de segurança pública, o ciclo de doze meses capta o movimento completo de um ano do roubo de cargas. Esse período é muito influenciado por fatores da economia e da logística legal, como as safras agrícolas, as férias coletivas das indústrias e o aumento no transporte de mercadorias para as festividades de fim de ano (JUSTUS et al., 2018). Do ponto de vista computacional, fixar o tempo em doze meses evita falhas comuns no treinamento de redes recorrentes, como o desaparecimento ou o crescimento exagerado dos valores dos gradientes (PASCANU; MIKOLOV; BENGIO, 2013). Essa configuração garante que o mecanismo de atenção guarde uma memória recente útil, sem diluir a importância dos meses passados e mantendo os cálculos estáveis, padrão adotado no desenvolvimento original da arquitetura (BAI et al., 2021). Com essa organização em três dimensões, o algoritmo processa os dados de todas as cidades ao mesmo tempo e consegue mapear como o risco criminal

se espalha pelas rodovias.

Quadro 12 – Decomposição das Dimensões do Tensor Espaço-Temporal $X \in \mathbb{R}^{N \times F \times T}$.

Eixo / Matriz	Valor	Significado no Fenômeno	Justificativa Arquitetural (GNN)
Nós (N)	289	Municípios paulistas ativos na série histórica filtrada.	Define a dimensão espacial estável do grafo para a agregação e propagação de mensagens.
Atributos (F)	1	Volume de roubos mensais estatisticamente normalizados.	Eixo obrigatório de canais de entrada (<i>in_channels</i>). É necessário para a multiplicação algébrica pelas matrizes de pesos da convolução, mesmo em cenários univariados, permitindo também futuras expansões de variáveis.
Tempo (T)	12	Passos retrospectivos (janela deslizante de contexto).	Extensão da sequência cronológica alimentada de forma encadeada nas células GRU e ponderada pelo bloco de atenção.

Fonte: Elaborado pelo autor (2026).

O funcionamento da arquitetura A3T-GCN começa pelo processamento do espaço geográfico. Para realizar essa operação, o modelo utiliza a Convolução em Grafos baseada na formulação clássica de Kipf e Welling (2017). Em cada mês t , o algoritmo calcula como o crime se espalha pelos municípios vizinhos conectados pelas rodovias por meio da equação matricial:

$$f(A, X_t) = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} X_t W \quad (5.2)$$

na qual $\tilde{A} = A + I_N$ representa a matriz de adjacência modificada pela adição da matriz identidade I_N . Essa operação funciona como um ajuste que garante que o município preserve o peso do seu próprio histórico de crimes durante os cálculos. A

matriz $\tilde{D}$ é uma matriz de grau diagonal usada para normalizar os dados de forma simétrica, o que evita que cidades com muitas conexões rodoviárias desequilibrem o modelo. O termo X_t representa a quantidade de roubos registrada no mês t e W indica a matriz de pesos que o algoritmo ajusta ao longo do treinamento.

Depois de calcular as relações do espaço para cada mês, o resultado entra diretamente nas portas de controle da Unidade Recorrente Fechada (GRU) para identificar a tendência do tempo, seguindo a estrutura matemática de Cho et al. (2014). O modelo processa os estados internos da rede de forma encadeada por meio das seguintes equações lineares e não lineares:

$$z_t = \sigma(W_z \cdot [f(A, X_t), H_{t-1}] + b_z) \quad (5.3)$$

$$r_t = \sigma(W_r \cdot [f(A, X_t), H_{t-1}] + b_r) \quad (5.4)$$

$$\tilde{H}_t = \tanh(W_h \cdot [f(A, X_t), r_t \odot H_{t-1}] + b_h) \quad (5.5)$$

$$H_t = (1 - z_t) \odot H_{t-1} + z_t \odot \tilde{H}_t \quad (5.6)$$

nas quais z_t representa a porta de atualização, encarregada de medir quanta informação do passado será levada para o futuro, e r_t identifica a porta de reinicialização, responsável por decidir quais dados antigos serão descartados. Os termos W_z, W_r, W_h indicam as matrizes de pesos e b_z, b_r, b_h são os vetores de ajuste de viés de cada porta. A letra σ representa a função de ativação sigmoide. O estado oculto candidato $\tilde{H}_t$ utiliza a função tangente hiperbólica ($\tanh$) para limitar os valores numéricos, e a combinação final gera o vetor de estado oculto H_t .

O grande diferencial da arquitetura A3T-GCN é a inclusão do mecanismo de atenção temporal sobre a sequência de meses gerada $\{H_1, H_2, \dots, H_T\}$, conforme proposto no modelo original de Bai et al. (2021). Esse módulo calcula a importância de cada mês do ano por meio de uma operação de projeção definida pela equação:

$$e_t = v_{att}^\top \tanh(W_{att}H_t + b_{att}) \quad (5.7)$$

em que W_{att} e b_{att} são os pesos e os vieses do bloco de atenção, e $v_{att}^\top$ é o vetor que extrai a nota de importância bruta e_t do período analisado. Para transformar essas notas em probabilidades comparáveis, o modelo aplica a função de normalização

exponencial *softmax*, gerando os pesos de atenção finais α_t :

$$\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^T \exp(e_k)} \quad (5.8)$$

O vetor de contexto final C é obtido pela soma ponderada de todos os estados ocultos multiplicados pelos seus respectivos pesos de atenção, expresso por $C = \sum_{t=1}^T \alpha_t H_t$. Esse mecanismo faz com que o algoritmo ignore variações aleatórias da série temporal e foque nos meses que realmente ajudam a entender o crime. Por fim, o vetor C passa por uma camada totalmente conectada com ativação linear, que transforma os dados em um único valor real. Esse número representa a quantidade de roubos prevista para o município no mês seguinte ($T + 1$).

5.3.2 Implementação e Dependências Tecnológicas

O código foi desenvolvido em linguagem **Python**, valendo-se do ecossistema de aprendizado profundo do **PyTorch**. A manipulação de dados estruturados em topologias não Euclidianas exigiu a integração de extensões de processamento geométrico acelerado por hardware em unidades de processamento gráfico (GPU T4). O Quadro 13 discrimina as dependências tecnológicas, as versões de compilação de nível estável empregadas e o papel analítico desempenhado por cada biblioteca no experimento espacial.

Quadro 13 – Dependências Tecnológicas e Bibliotecas de Suporte (STGNN).

Biblioteca	Versão	Função no Experimento Computacional
PyTorch	2.5.1+cu121	Processamento de tensores e execução de grafos de computação em GPU.
PyG (Geometric)	2.6.1	Construção da estrutura do grafo e gerenciamento das conexões entre os municípios.
PyG Temporal	0.1.14	Implementação nativa e otimização dos operadores da célula A3T-GCN.
Geobr	1.7.0	Carregamento e vetorização da malha municipal oficial de 2024 do IBGE.

Fonte: Elaborado pelo autor (2026).

5.3.3 Especificações da Arquitetura e Hiperparâmetros

A rede neural foi calibrada para resolver um problema de regressão contínua multipasso com saída única (*single-step forecasting*), estimando o nível de risco de forma pontual para o mês subsequente à janela de observação. O Quadro 14 demonstra a arquitetura sequencial do modelo, detalhando o fluxo de dados desde a leitura do grafo de vizinhança até a camada densa terminal. O processo de convergência utilizou o otimizador Adam com uma taxa de aprendizado de 0,01, aplicando o Erro Quadrático Médio (MSE) como função de perda para orientar o ajuste dos parâmetros matemáticos.

Quadro 14 – Especificações da Arquitetura Neural Espaço-Temporal (A3T-GCN).

#	Tipo de Camada	Unidades	Papel Operacional no Fluxo Preditivo
1	Input (Graph)	-	Captura da matriz de adjacência de 289 nós e tensores de atributos.
2	A3T-GCN Layer	32	Processamento convolucional espacial e filtragem por atenção temporal.
3	Activation (ReLU)	-	Introdução de não linearidade e imposição do piso de não negatividade.
4	Linear (Dense)	1	Colapso dos canais ocultos latentes em uma projeção escalar contínua.
5	Output Vector	-	Consolidação do vetor contendo as estimativas de crimes por município.

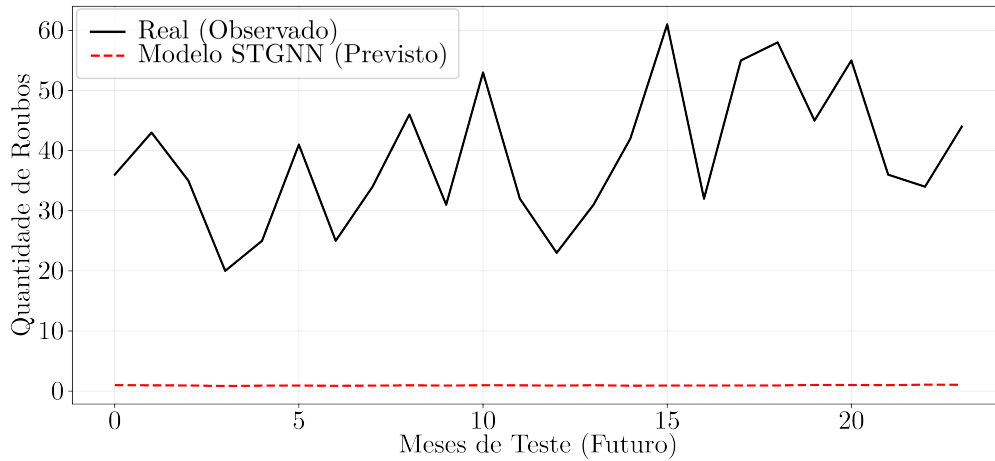
Fonte: Elaborado pelo autor (2026).

5.4 Aprendizado Sensível ao Custo (*Cost-Sensitive Learning*)

Ao executar o modelo pela primeira vez, identificamos um comportamento típico em redes neurais baseados em grafos, a suavização excessiva (*oversmo-*

othing) (LI; HAN; WU, 2018; RUSCH; BRONSTEIN; MISHRA, 2023). Como mostrado na Figura 15, o modelo inicial apresentou uma forte subestimação das ocorrências na Capital, gerando uma linha de previsão quase plana próxima ao valor de 0,77 roubos, muito abaixo da realidade histórica. Isso ocorre porque o algoritmo busca minimizar o erro médio global das 289 cidades juntas. Para manter a taxa de acerto alta no estado inteiro, o sistema prefere achatar os picos da capital a gerar falsos alarmes em municípios pequenos do interior.

Figura 15 – Projeção inicial para a Capital (SP). Sem o ajuste de pesos, a previsão resulta em uma linha quase plana devido à suavização excessiva.



Fonte: Elaborada pelo autor (2026).

Para forçar o modelo a reconhecer a alta concentração de crimes na região metropolitana, aplicamos duas correções, a primeira a transformação logarítmica dos dados e a segunda a punição assimétrica dos erros de previsão.

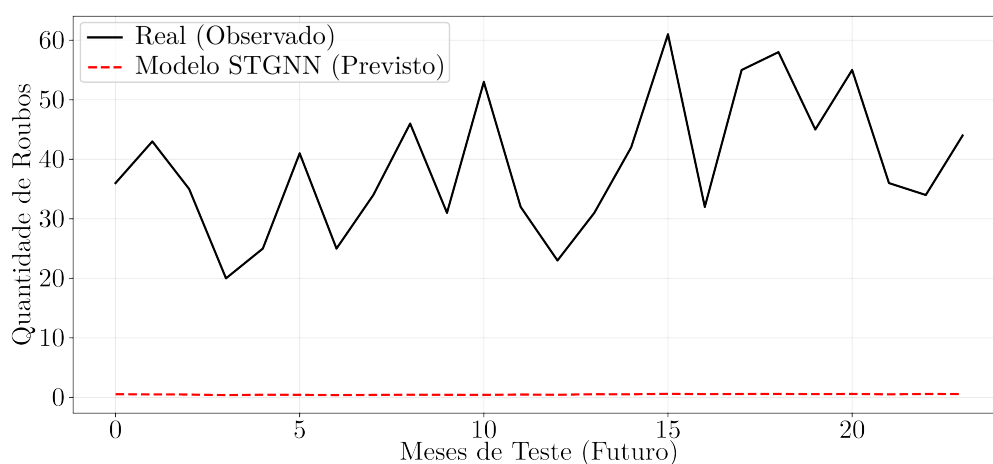
5.4.1 Transformação Logarítmica da Variância

Comprimos os dados de criminalidade aplicando a equação $f(x) = \ln(1+x)$ antes de iniciar o treinamento, adicionando o valor 1 para evitar o cálculo de logaritmo de zero. Essa transformação reduz o impacto dos valores extremos e faz com que a rede foque na proporção de variação do crime, e não apenas nos

números absolutos (GOODFELLOW; BENGIO; COURVILLE, 2016; ZHA et al., 2023). Com esse ajuste, pequenos aumentos de casos no interior ganham relevância matemática comparável às variações da capital, ajudando o modelo a identificar como o risco se espalha pelas rodovias.

A Figura 16 mostra o resultado dessa primeira correção. O modelo passa a reagir às oscilações da capital, mas a linha de previsão ainda aparece achatada e visualmente semelhante à anterior. Essa semelhança ocorre porque a transformação logarítmica altera a escala dos gráficos (comprimindo os dados no eixo vertical), mas ainda não altera o peso do erro no treinamento da rede. Sem a punição direcionada, o algoritmo continua suavizando os picos para reduzir o erro global.

Figura 16 – Projeção para a Capital após a transformação logarítmica. A curva acompanha as oscilações, mas permanece em escala comprimida devido à falta de pesos no treinamento.



Fonte: Elaborada pelo autor (2026).

5.4.2 A Função de Perda Ponderada (*Weighted Loss*)

Para resolver o problema da suavização de forma definitiva, adotamos o Aprendizado Sensível ao Custo (*Cost-Sensitive Learning*) (HU et al., 2024; JOHNSON; KHOSHGOFTAAR, 2019). Configuramos a função de perda do modelo

para que os erros de previsão nos municípios com maior volume de crimes recebessem uma punição matemática mais severa.

Definição 5.1 (Erro Quadrático Médio Ponderado). O cálculo do erro com aplicação de pesos (L_w) é definido pela fórmula:

$$L_w = \frac{1}{N} \sum_{i=1}^N \omega_i \cdot (y_i - \hat{y}_i)^2 \quad (5.9)$$

em que:

- ω_i : é o peso aplicado ao município i , calculado pela razão entre o volume histórico de crimes da cidade e a média de roubos de todo o estado;
- $(y_i - \hat{y}_i)^2$: é o erro quadrático da previsão para aquele município.

O peso de cada cidade foi definido de forma automatizada a partir do histórico de dados, obedecendo à seguinte equação de normalização:

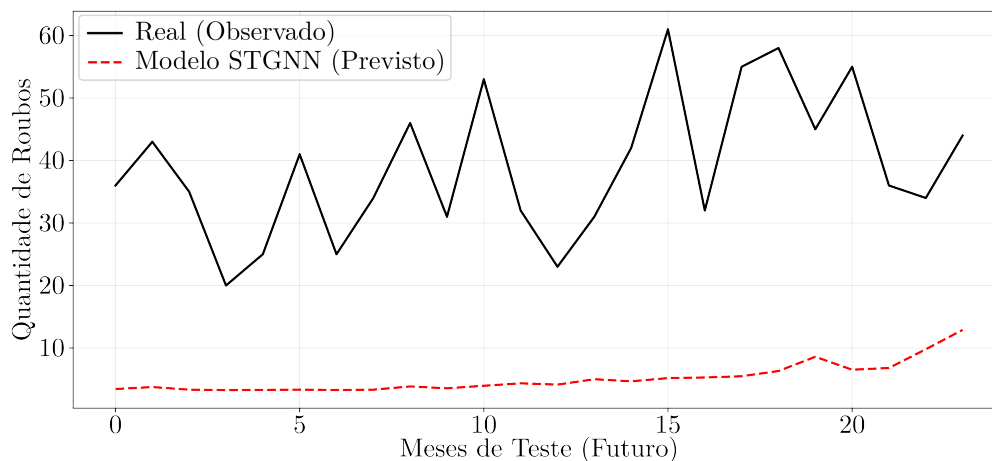
$$\omega_i = \frac{V_i}{\frac{1}{N} \sum_{j=1}^N V_j} \quad (5.10)$$

em que V_i representa o número total de roubos registrados no município i ao longo de todo o período analisado. Para o município de São Paulo (código IBGE 3550308), esse cálculo resultou em um peso punitivo elevado de $\omega \approx 107,04$. Esse fator força o otimizador a priorizar o ajuste da capital durante o treinamento.

A Figura 17 confirma a eficácia dessa estratégia. Sob a influência do peso punitivo, o modelo passou a acompanhar os picos e vales da criminalidade na capital. Esse ajuste reduziu o Erro Absoluto Médio (MAE) local de São Paulo de 38,63 para 33,89 roubos por mês.

Apesar de o modelo multivariado acompanhar a tendência dos dados (Figura 17), os valores previstos (linha vermelha) ainda se mantêm numericamente abaixo dos dados reais, apresentando uma média prevista de 5,14 roubos frente a uma média real de 39,04. Como os resultados já foram revertidos da escala logarítmica para valores absolutos, essa diferença reflete a persistência da suavização excessiva causada pela estrutura do grafo. Como a capital possui 68 conexões com municípios menores cujos índices de roubo são próximos a zero, a previsão de São Paulo é puxada para baixo pela influência de sua vizinhança.

Figura 17 – Desempenho final sob Perda Ponderada. O modelo corrige a trajetória e passa a acompanhar as variações reais da Capital.



Fonte: Elaborada pelo autor (2026).

5.4.3 Deslocamento de Viés e Limitações Geométricas das Arestas

O aumento da sensibilidade nos grandes polos logísticos gerou um efeito colateral estatístico nos municípios menores, conhecido como Deslocamento de Viés (*Bias Shift*) (CLEVERT; UNTERTHINER; HOCHREITER, 2016). Cidades com pouca ou nenhuma ocorrência histórica, como Araçatuba, passaram a apresentar previsões médias próximas a 5 roubos mensais, gerando falsos positivos. O Quadro 15 mostra esse comportamento nas estimativas do modelo após a aplicação da função de perda ponderada.

Esse aumento geral nas previsões ocorre porque o algoritmo de otimização (Adam), pressionado pela punição de 107 vezes aplicada aos erros de São Paulo, eleva o patamar mínimo das previsões em toda a rede. Para o modelo, o custo computacional de gerar um falso positivo em uma cidade pequena é menor do que o risco de prever valores próximos a zero na Capital e sofrer a punição multiplicada.

Essa distorção está diretamente ligada à forma como as arestas do grafo foram construídas. Os pesos das ligações consideraram apenas a distância em linha reta entre os centros geográficos dos municípios, criando uma conectividade

Quadro 15 – Comparativo de desempenho entre polos logísticos e cidades do interior.

Município	Código IBGE	Média Real Mensal	Média Prevista Mensal	MAE Local
São Paulo	3550308	39,04	5,14*	33,90
Guarulhos	3518800	12,92	4,22	8,78
Cubatão	3513504	8,21	3,59	6,41
São Caetano do Sul	3548807	0,37	5,44	5,06
Araçatuba	3502804	0,00	4,98	4,97

Fonte: Elaborado pelo autor (2026). (*) Valores em escala logarítmica reconstruída.

puramente geométrica. Na realidade do transporte de cargas, o uso de outros fatores para calcular o peso das arestas — como o traçado real das rodovias, o volume diário de caminhões e o tipo de mercadoria movimentada — mitigaria esse problema. Esses fatores adicionais mudariam a estrutura do grafo, separando áreas que estão próximas em linha reta, mas distantes pelas rotas logísticas reais. Isso impediria que o alto risco da Capital elevasse artificialmente as previsões de cidades do interior sem histórico de crimes.

Apesar de ser um viés estatístico, esse comportamento revela a capacidade do modelo de mapear o risco de contágio geográfico pelas estradas paulistas. Como as rodovias conectam fortemente o estado, a rede identifica que cidades menores, mesmo sem registros recentes, estão estruturalmente expostas ao risco por estarem ligadas aos mesmos corredores utilizados para a fuga ou transbordo de mercadorias. Sob a ótica do planejamento de segurança pública, o modelo atua como um indicador de vulnerabilidade territorial, auxiliando na antecipação de riscos e no posicionamento de ações preventivas ao longo das rotas de circulação do capital ilícito.

5.5 Resultados e Cartografia Preditiva: A Topologia do Risco

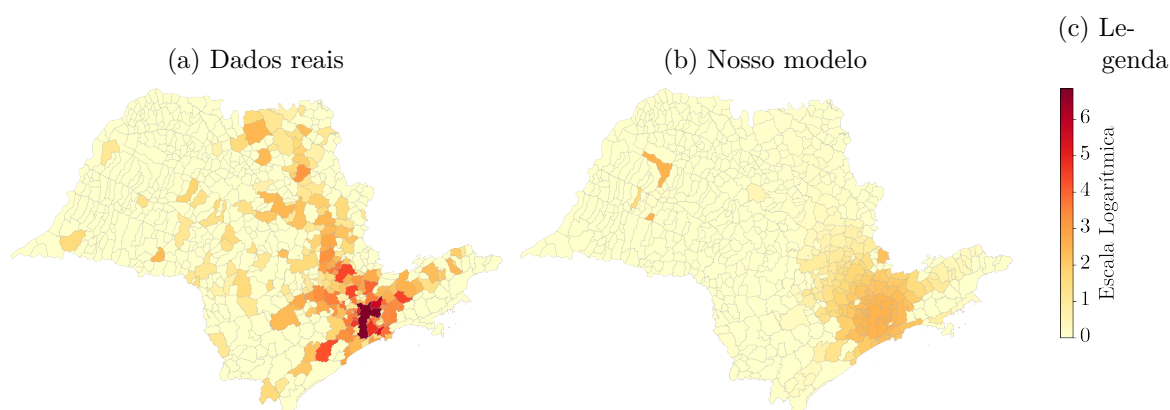
A análise isolada de tabelas e taxas de erro é insuficiente para o planejamento estratégico da segurança pública. Para tornar os resultados aplicáveis,

os valores numéricos gerados pela rede neural foram convertidos em mapas de risco, transformando dados matemáticos em áreas coloridas que representam a concentração geográfica dos roubos de carga em São Paulo.

5.5.1 Análise de Mancha Criminal: O Biênio de Validação (2021–2022)

A Figura 18 apresenta a comparação entre os boletins de ocorrência reais e o volume de risco estimado pelo modelo A3T-GCN para o biênio de 2021 e 2022.

Figura 18 – Cartografia Comparativa (2021–2022). Os mapas foram construídos com o auxílio da biblioteca `geobr`, que fornece os limites territoriais oficiais do IBGE. Na definição das cores de risco, aplicou-se a transformação logarítmica $f(x) = \ln(1 + x)$. Esse ajuste é necessário para evitar que o alto volume de ocorrências da Região Metropolitana oculte os padrões de criminalidade das cidades do interior, permitindo a visualização equilibrada de todo o território. O uso das cores em escala logarítmica revela claramente os corredores espaciais de risco viajar sobre o asfalto.



Fonte: Elaborada pelo autor (2026).

O mapeamento comparativo revela duas características importantes sobre o comportamento do modelo em grafos e o funcionamento da rede neural. A primeira delas é a suavização das cores no mapa gerado pelo algoritmo. O mapa do modelo apresenta tons mais claros se comparado ao cenário real. Enquanto a realidade exhibe colorações escuras e concentradas na Região Metropolitana, o modelo projeta uma mancha expandida em tons de laranja. Essa perda de intensidade local é o reflexo da

suavização excessiva (*oversmoothing*), um fenômeno característico das convoluções em grafos. Como o algoritmo calcula o risco compartilhando informações entre os 68 vizinhos conectados de cada cidade, o valor extremo da Capital acaba diluído pela média dos municípios próximos com criminalidade baixa ou nula. Essa variação de tom não indica um erro de localização, mas uma limitação matemática do modelo ao tentar equilibrar as grandes concentrações urbanas com o restante do estado.

A segunda característica marcante é o surgimento de uma mancha alaranjada isolada na região de Araçatuba, no interior paulista, onde os dados reais apontam ausência de ocorrências. Esse comportamento ilustra o deslocamento de viés (*bias shift*) provocado pela aplicação da função de perda ponderada durante o treinamento. Para evitar punições matemáticas severas devido ao peso de 107 vezes aplicado aos erros da Capital, o otimizador eleva o nível básico de risco em toda a estrutura do grafo. Como o modelo foi construído utilizando conexões baseadas apenas na distância geométrica em linha reta entre os centroides oficiais do IBGE, o risco concentrado na área macrometropolitana acabou sendo transmitido para eixos distantes do interior, gerando esse falso positivo na porção oeste do estado.

Apesar dessas diferenças de intensidade e dos falsos positivos locais, o modelo reproduz com fidelidade o traçado dos principais eixos logísticos do estado. A projeção destaca de forma clara os corredores formados pelas rodovias Anhanguera, Bandeirantes e Castello Branco. O mapa preditivo confirma a eficácia do algoritmo para identificar a direção do movimento do crime e a tendência de espalhamento do risco pelas estradas, cumprindo seu papel de ferramenta de apoio ao planejamento estratégico da segurança pública.

5.5.1.1 Interpretação Topológica e o Impacto da Sensibilidade ao Custo

As áreas de risco realçadas nos mapas demonstram o efeito prático da função de perda ponderada. Ao penalizar severamente os erros na Capital, o modelo elevou o nível de alerta para as cidades conectadas a ela pela malha rodoviária. Os padrões projetados indicam que o risco de roubo não respeita as fronteiras administrativas dos municípios, mas acompanha a conectividade das estradas. Esse comportamento confirma a importância de utilizar grafos para modelar o território e compreender como a criminalidade se propaga espacialmente.

5.5.1.2 Resumo Metodológico da Construção Cartográfica

Para garantir a transparência e a reprodutibilidade do método, o Quadro 16 resume os componentes, dados e ferramentas computacionais utilizados na elaboração dos mapas de risco.

Quadro 16 – Parâmetros de Construção da Cartografia Preditiva.

Componente	Especificação Técnico-Científica
Base de Dados	Secretaria de Segurança Pública de SP (SSP-SP).
Malha Geográfica	Biblioteca <code>geobr</code> (IBGE 2024) manipulada pelo <code>Geopandas</code> .
Período de Teste	Biênio acumulado (Janeiro/2021 a Dezembro/2022).
Transformação Visual	Aplicação da fórmula $f(x) = \ln(1 + x)$ para corrigir a saturação visual da cor.
Lógica Preditiva	Aprendizado Sensível ao Custo (<i>Cost-Sensitive Learning</i>).

Fonte: Elaborado pelo autor (2026).

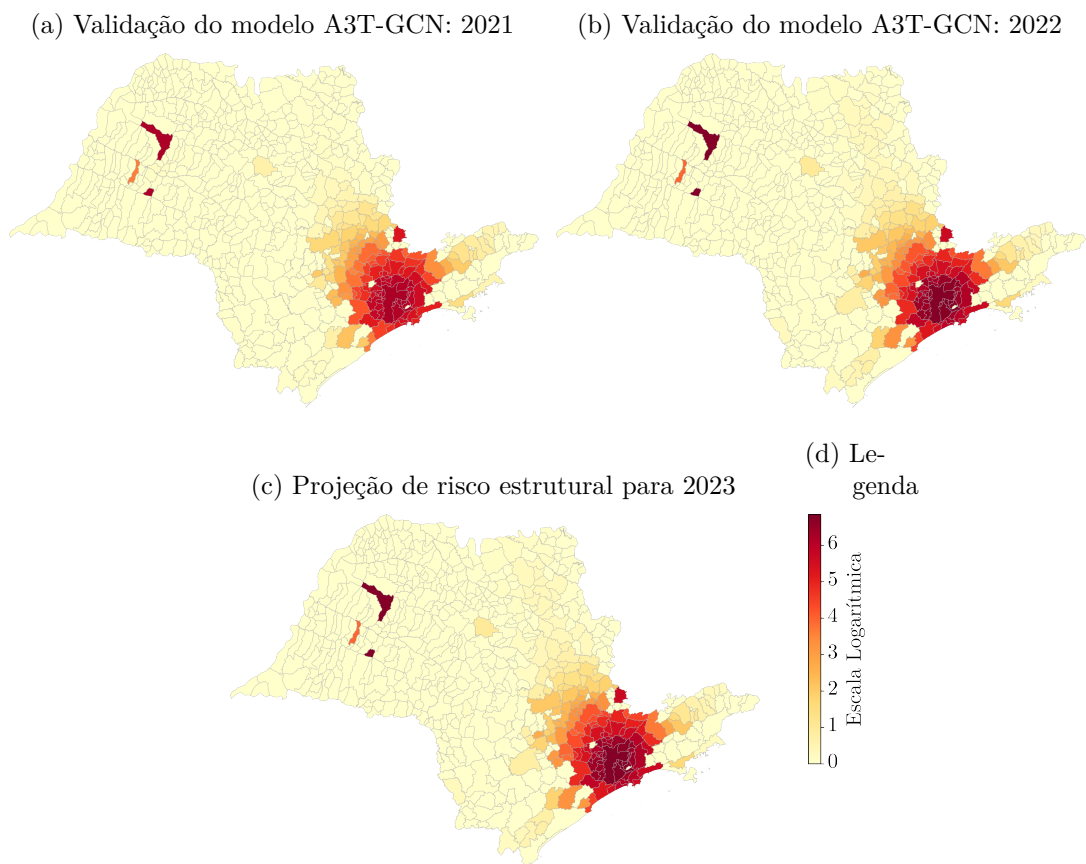
5.5.2 Projeção de Cenários e a Inércia Logística (Extrapolação 2023)

O modelo também foi utilizado para realizar a extrapolção de dados para o ano de 2023. A estabilidade visual quase idêntica das manchas de risco ao longo dos anos de 2021, 2022 e na projeção de 2023 reflete diretamente o comportamento de previsão constante observado nos gráficos de linha do horizonte temporal. Por causa da forte suavização e do viés da média, o algoritmo A3T-GCN tende a convergir para um estado de equilíbrio estático de longo prazo nas saídas mensais.

Em vez de acompanhar flutuações e variações abruptas mês a mês, a rede neural especializou-se em projetar o risco estrutural estável do território. Esse fenômeno demonstra a chamada inércia logística criminal. Os resultados indicam que variações de curto prazo têm pouca influência na tendência do crime quando comparadas a fatores geográficos fixos, como as redes de transporte e as zonas industriais. As quadrilhas dependem diretamente do acesso aos grandes

centros urbanos tanto para o armazenamento de produtos roubados quanto para a movimentação financeira. Portanto, enquanto o traçado das rodovias e os polos logísticos permanecerem fixos, a distribuição das manchas de calor preditivas apresentará pouca mudança no tempo.

Figura 19 – Evolução Espaço-Temporal e Extrapolação (2021 a 2023). A estagnação territorial do mapa comprova a natureza fixa e enraizada da estrutura corporativa do ilícito.



Fonte: Elaborada pelo autor (2026).

5.5.3 Reorientação do Policiamento: Do Modelo Reativo ao Controle Territorial

A dispersão do risco calculada pelo modelo indica as rotas mais prováveis de deslocamento do crime. Atualmente, o policiamento costuma atuar de forma reativa, posicionando viaturas com base nas ocorrências do mês anterior. A modelagem em grafos propõe uma estratégia diferente, baseada na presença fixa em pontos de conexão da malha viária. Ao demonstrar matematicamente que o perigo se desloca da Região Metropolitana para o interior pelas rodovias, o modelo confirma que a segurança das rotas do interior depende diretamente do controle e do desmantelamento das bases financeiras e receptoras localizadas no grande centro logístico do estado.

5.6 Análise Crítica: A Conectividade do Grafo e a Estrutura do Capital

Além do desempenho técnico, os resultados podem ser analisados sob a ótica da criminologia crítica. Os dados estruturados no formato de tensor pelo modelo A3T-GCN confirmam a premissa de que o roubo de cargas depende da mesma infraestrutura de transportes construída para a circulação regular de mercadorias (PACHUKANIS, 2017). O comportamento do crime não ocorre de forma isolada, mas se apoia na malha rodoviária, nos corredores industriais e nos fluxos econômicos do mercado formal.

A matriz de conectividade descreve a organização territorial do estado. O funcionamento do modelo baseado em grafos espaciais está alinhado à teoria do Estado de Mascaro (2013), demonstrando que os grupos criminosos operam com uma estrutura semelhante à de empresas comerciais. O roubo de cargas segue uma lógica de mercado em que os lucros obtidos nas rodovias são privatizados pelas quadrilhas, enquanto os custos com segurança e os prejuízos são transferidos para o setor público e para a sociedade por meio da demanda por policiamento.

5.6.1 Implicações Estratégicas para a Justiça Criminal

No âmbito do processo penal, os mapas de risco indicam que focar a política criminal apenas na prisão dos executores diretos do roubo nas rodovias — como motoristas de apoio e receptadores locais — é uma estratégia insuficiente para desarticular o problema.

A precisão do modelo em apontar o risco concentrado nas grandes vias apoia a atuação do Ministério Público. Como o modelo demonstra que a Região Metropolitana funciona como o centro de distribuição do risco, os Grupos de Atuação Especial de Combate ao Crime Organizado (GAECO) passam a ter fundamentos técnicos para direcionar investigações contra financiadores e lavadores de dinheiro em indústrias e distribuidoras locais.

Esses resultados também estão alinhados às análises de Zaffaroni (1991) e Zaffaroni (2012) sobre a responsabilidade no controle de riscos. Grandes empresas de transporte, quando informadas tecnicamente sobre os riscos geográficos das rotas, compartilham a responsabilidade caso deixem de adotar medidas preventivas. A escolha de não investir em sistemas de rastreamento e segurança para maximizar lucros pode caracterizar uma conduta negligente. Com esses dados preditivos, o Ministério Público ganha fundamentação técnica para propor Termos de Ajustamento de Conduta (TAC) e exigir investimentos em segurança por parte das empresas de logística.

5.7 Considerações Finais do Capítulo: A Topologia do Risco e a Criminologia em Redes

A aplicação do modelo A3T-GCN ao roubo de cargas integrou de forma direta as dimensões de espaço e tempo. A substituição de tabelas isoladas por um tensor tridimensional ($X \in \mathbb{R}^{289 \times 1 \times 12}$) permitiu identificar padrões persistentes e quantificar o comportamento do crime no território paulista.

O problema do achatamento das previsões em torno da média, causado pela suavização excessiva, foi corrigido com ajustes no treinamento da rede neural. A aplicação de transformações logarítmicas e do aprendizado sensível ao custo

(*Cost-Sensitive Learning*) aumentou o peso dos erros na Região Metropolitana de São Paulo. Essa abordagem reduziu a suavização estatística e permitiu que o modelo acompanhasse os picos de ocorrências e as variações das grandes operações criminosas nas rodovias.

A estabilidade das áreas de risco nas projeções para períodos futuros confirma o caráter estrutural do roubo de cargas. O espalhamento do crime está diretamente ligado à infraestrutura logística e industrial de São Paulo, dependendo da conexão das rodovias e dos grandes centros urbanos.

Esse cenário mostra que operações policiais isoladas têm efeito limitado contra quadrilhas organizadas em rede ao longo de eixos de transporte consolidados. O combate eficiente exige ações contínuas para desarticular as conexões financeiras e operacionais desses grupos, concentrando os esforços nos centros metropolitanos que comandam a circulação do capital ilícito.

6 Conclusão Geral e Trabalhos Futuros

Nesta tese, investigou-se o fenômeno do roubo de cargas no Estado de São Paulo por meio do processamento de dados e do uso de arquiteturas de Redes Neurais em Grafos Espaço-Temporais (*Spatio-Temporal Graph Neural Networks* — STGNN). O objetivo central foi demonstrar, por meio de métricas numéricas e modelagem computacional, que esse tipo de crime segue um padrão geográfico e dinâmico diretamente ligado à estrutura da malha rodoviária paulista.

Os resultados obtidos confirmam que o risco de roubo de cargas não se distribui de maneira aleatória pelo território. Em vez disso, o crime se comporta como um fluxo que utiliza os eixos de transporte formais para a circulação e o escoamento de mercadorias, validando a necessidade de abordagens que analisem o espaço e o tempo de forma integrada.

6.1 Síntese dos Achados Técnicos e Contribuições do Trabalho

A trajetória de testes realizada nesta pesquisa evidenciou as limitações das abordagens tradicionais e justificou o aumento gradual da complexidade dos modelos. Os experimentos iniciais baseados em séries temporais puras, como o modelo estatístico SARIMA e a rede recorrente GRU univariada, mostraram que tentar prever a criminalidade ignorando a geografia gera uma suavização excessiva dos dados. Esses modelos de base foram incapazes de capturar os picos de ocorrência e a volatilidade do crime, pois calculavam o risco de cada município de maneira isolada.

A principal contribuição computacional deste trabalho foi a aplicação da arquitetura A3T-GCN sobre um grafo representativo do território. Ao conectar os municípios por meio de uma matriz de adjacência baseada na distância entre os centroides oficiais do IBGE, o algoritmo conseguiu capturar como o risco criminal se propagará pelas estradas. O modelo demonstrou que um aumento de ocorrências na Região Metropolitana gera um efeito de contágio que eleva o nível de perigo nas

idades do interior conectadas pelos mesmos eixos viários.

Os fenômenos identificados nos mapas como suavização espacial (*oversmoothing*) e deslocamento de viés (*bias shift*) revelaram-se características importantes do ajuste do algoritmo. A aplicação do aprendizado sensível ao custo, com um peso maior para os erros na Capital, forçou o otimizador a buscar um estado de equilíbrio estável. Esse comportamento gerou uma linha de tendência constante nas projeções de longo prazo, confirmando que o modelo funciona como um mapeamento estável de vulnerabilidade territorial, indicando com precisão os principais corredores rodoviários de risco.

6.2 Implicações Teóricas e a Homologia Algorítmica Estrutural

As implicações teóricas dos resultados obtidos exigem uma leitura realista do papel desempenhado pelas estratégias tradicionais de segurança pública. O patrulhamento ostensivo rodoviário possui aplicação prática, mas sua eficácia é limitada por restrições estruturais de efetivo que reduzem a capilaridade da fiscalização nas estradas. Diante de uma presença policial rarefeita, a atuação fardada opera primordialmente na esfera psicológica, gerando uma sensação de segurança e exercendo uma dissuasão local contra a criminalidade amadora. No entanto, para as redes estruturadas do roubo de cargas, essa presença física pontual torna-se uma variável de baixa probabilidade de interferência. A imobilidade territorial da mancha criminal ao longo do horizonte histórico analisado demonstra que os grupos organizados operam sob uma lógica empresarial, na qual a escassez de fiscalização é calculada e internalizada como um risco operacional previsível em suas planilhas de custos.

Diante dessa limitação do policiamento tradicional, o ineditismo desta tese consolida-se na construção de uma intersecção entre o desenvolvimento de ferramentas de alto desempenho computacional e a interpretação teórica estrutural do território. Enquanto a literatura de engenharia frequentemente adota uma postura gerencial que trata o crime como uma patologia isolada, e a literatura jurídica e sociológica foca em análises textuais sem validação empírica, este trabalho utiliza o modelo A3T-GCN para demonstrar o alinhamento existente entre o

processamento matemático e a dinâmica material do espaço geográfico.

Essa convergência permite propor o conceito de homologia algorítmica estrutural, expandindo a formulação clássica desenvolvida por Evgeni Pachukanis. Na teoria social estrutural, a homologia identifica estruturas internas semelhantes em fenômenos de naturezas distintas. Pachukanis demonstrou que o mercado necessita reduzir mercadorias diferentes a uma medida abstrata comum, que é o valor de troca, da mesma forma que o sistema jurídico reduz indivíduos diferentes a uma ficção abstrata e igualitária, que é o sujeito de direito, tornando as estruturas do mercado e do direito espelhos mútuos.

Esta tese estende essa relação para o ambiente digital ao demonstrar que a circulação do comércio legal e o deslocamento das redes criminosas compartilham a mesma infraestrutura física. Ao atingir a estabilidade nas métricas de erro, o modelo A3T-GCN não cria uma nova realidade geográfica. O algoritmo apenas mapeia as redes logísticas que sustentam o fluxo de riquezas. O código da rede neural atua, portanto, como um reflexo exato das redes de distribuição, provando que o comportamento do modelo é homólogo ao metabolismo de circulação do capital no território.

6.3 O Legado Metodológico da Tese

Para consolidar esta abordagem como uma pesquisa replicável, a tese define um protocolo estruturado em quatro etapas. Esse fluxo de trabalho independe do modelo matemático utilizado. Isso significa que o método continua válido mesmo se as redes neurais em grafos forem substituídas por outras arquiteturas de aprendizado profundo ou por modelos estatísticos tradicionais.

A primeira etapa foca na engenharia de dados e na criação dos tensores. Nela, aplicamos técnicas de Ciência de Dados para extrair, limpar e formatar os registros criminais. A segunda etapa introduz a lente crítica: analisamos a topologia da rede e os mapas preditivos não como números neutros, mas como o reflexo exato das contradições econômicas e logísticas do território paulista no capitalismo periférico.

A terceira etapa foca no refinamento dialético, momento em que aplicamos o conceito de superação (*Aufhebung*). O motor do método é o conflito entre a previsão inicial do algoritmo e a sua falha em registrar os extremos de criminalidade nos grandes polos logísticos. O retorno ao código para aplicar a Função de Perda Ponderada (*Weighted Loss*) não é apenas uma correção de software, mas a verdadeira prática de pesquisa. O modelo evolui ao transformar a desigualdade estrutural do roubo de cargas em um parâmetro matemático.

O ciclo termina na quarta etapa, com a validação espacial e a redação final. Ela une a precisão exigida pela Matemática Aplicada à visão estrutural da Criminologia Crítica. Esse passo a passo garante que o protocolo possa ser reproduzido para investigar outros fenômenos de segurança pública que parasitam a infraestrutura viária.

6.4 Limitações e Trabalhos Futuros

Apesar dos avanços alcançados na modelagem espaço-temporal, a arquitetura atual possui pontos de simplificação que limitam sua capacidade de previsão de curto prazo. Essas barreiras definem caminhos claros para a continuidade da pesquisa, focados no aperfeiçoamento algorítmico e na engenharia de dados.

6.4.1 Expansão Multivariada do Tensor e Variáveis Ambientais

A versão atual do modelo opera com uma estrutura de atributos simplificada, utilizando um tensor univariado no qual a dimensão de características de entrada está fixada em $F = 1$, representando apenas o volume histórico de roubos. Como proposta de trabalho futuro, projeta-se a expansão desse bloco de dados para um formato multivariado ($F > 1$), o que resolve a limitação apontada sobre a tridimensionalidade do tensor.

Essa ampliação será feita por meio da inserção de condições climáticas específicas coletadas no momento de cada ocorrência, tais como o índice de precipitação pluvial, a presença de névoa e o horário do pôr do sol. A inclusão dessas variáveis meteorológicas associadas a indicadores do fluxo econômico permitirá que

a rede neural correlacione a visibilidade e as condições de aderência da pista com a oportunidade de atuação das quadrilhas, elevando a capacidade de previsão do modelo.

6.4.2 Refinamento Topológico com Geoprocessamento Avançado

A matriz de adjacência do grafo foi construída utilizando aproximações de distância em linha reta entre as coordenadas centrais das cidades. Embora essa técnica funcione como uma aproximação, ela desconsidera o traçado real das rodovias e as barreiras físicas do terreno.

Para os próximos trabalhos, propõe-se a migração dos dados geográficos para um sistema de banco de dados espacial baseado na extensão *PostGIS* do *PostgreSQL*. O uso de ferramentas de geoprocessamento e APIs de roteamento rodoviário permitirá calcular a distância viária real e o tempo de viagem entre os municípios. Ao considerar fatores práticos como a localização de postos de pesagem, pedágios e a velocidade média das vias, o grafo passará a espelhar os caminhos logísticos reais utilizados no transporte de cargas.

6.4.3 Modelagem por Grafos Heterogêneos Relacionais

Como desdobramento de longo prazo, projeta-se a expansão da modelagem topológica por meio do uso de Grafos Heterogêneos Relacionais, definidos matematicamente pela estrutura $G = (V, E, \mathcal{T}_v, \mathcal{T}_e)$, na qual $\mathcal{T}_v$ e $\mathcal{T}_e$ representam os conjuntos de tipos de vértices e arestas, respectivamente. Essa formulação permite superar as restrições de grafos homogêneos ao incorporar múltiplos tipos de entidades e interações em uma mesma rede de aprendizado.

O objetivo dessa expansão é analisar o princípio constitucional da eficiência administrativa sob uma perspectiva quantitativa. O grafo heterogêneo mapeará não apenas a malha física das rodovias, mas a rede de relações institucionais e financeiras composta por nós que representam os órgãos públicos, as concessionárias de transporte, os provedores de infraestrutura e os trabalhadores operacionais na ponta do sistema logístico.

Por meio da mensuração dos fluxos de dados e ativos financeiros entre esses nós distintos, o algoritmo fornecerá uma base empírica para avaliar como as dinâmicas de delegação e a fragmentação de contratos influenciam a vulnerabilidade do território. Essa modelagem servirá para subsidiar investigações jurimétricas avançadas e fornecer parâmetros numéricos estruturados para o desenho de políticas públicas focadas na preservação e no fortalecimento de infraestruturas estratégicas estatais.

Referências

ABADI, M. et al. Tensorflow: A system for large-scale machine learning. In: *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*. [S.l.: s.n.], 2016. p. 265–283. 59

Agência Pública. *Master: como CPMI do INSS e CPI do Crime Organizado se conectam*. 2026. Disponível em: <<https://apublica.org/2026/03/master-como-cpmi-do-inss-e-cpi-do-crime-organizado-se-conectam/>>. Acesso em: 1 abr. 2026. 20

AKINYEMI, E. K. et al. Modelling theft criminal offence in kwara state using arima. *International Journal of Research and Scientific Innovation (IJRSI)*, v. 8, n. 4, 2021. 30

ALEMANY, F. R. *Punição e estrutura social brasileira*. Dissertação (Dissertação de Mestrado) — Faculdade de Direito, Universidade de São Paulo, São Paulo, 2019. 22, 27

AN, X. et al. Strp-dbscan: A parallel dbscan algorithm based on spatial-temporal random partitioning for clustering trajectory data. *Applied Sciences*, v. 13, n. 20, 2023. 30

ANSELIN, L. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers, 1988. 23

ANSELIN, L. Local indicators of spatial association—lisa. *Geographical analysis*, Wiley Online Library, v. 27, n. 2, p. 93–115, 1995. 25, 30, 39, 51

ARANSIOLA, T. J.; JUSTUS, M.; CECCATO, V. Space-time dynamics of cargo theft: evidence from são paulo, brazil. *Journal of Transportation Security*, Springer, v. 16, n. 1, p. 1–28, 2023. 43, 70, 75, 76

Atlantic Council. *After Maduro: Latin America’s policy community reassesses the US-China balance*. 2026. Disponível em: <<https://www.atlanticcouncil.org/in-dept-research-reports/issue-brief/after-maduro-latin-americas-policy-community-reassesses-the-us-china-balance/>>. Acesso em: 1 abr. 2026. 22

BAI, J. et al. A3t-gcn: Attention temporal graph convolutional network for traffic forecasting. *ISPRS International Journal of Geo-Information*, MDPI, v. 10, n. 7, p. 485, 2021. 25, 32, 77, 79, 82, 83, 85

- BARATTA, A. *Criminologia crítica e crítica do direito penal: introdução à sociologia do direito penal*. Rio de Janeiro: Revan, 1982. 76
- BOX, G. E. P. et al. *Time Series Analysis: Forecasting and Control*. 5th. ed. [S.l.]: John Wiley & Sons, 2015. 57, 60
- BRASIL DE FATO. *Após reunião com Flávio, EUA classificam PCC e Comando Vermelho como organizações terroristas*. 2026. Disponível em: <<https://www.brasildefato.com.br/2026/05/28/apos-reuniao-com-flavio-eua-classificam-pcc-e-comando-vermelho-como-organizacoes-terroristas/>>. Acesso em: 20 de junho de 2026. 22, 28
- CAMPELLO, R. J.; MOULAVI, D.; SANDER, J. Density-based clustering based on hierarchical density estimates. In: SPRINGER. *Pacific-Asia conference on knowledge discovery and data mining*. [S.l.], 2013. p. 160–172. 29
- CATLETT, C. et al. Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments. In: *Pervasive and Mobile Computing*. [S.l.: s.n.], 2019. v. 53, p. 62–74. 58
- CEDILLO-CAMPOS, M. G.; FLORES-FRANCO, J. E.; COVARRUBIAS, D. A physical internet-based analytic model for reducing the risk of cargo theft in road transportation. *Computers & Industrial Engineering*, Elsevier, v. 190, p. 110016, 2024. 29, 55
- CHAI, T.; DRAXLER, R. R. Root mean square error (rmse) or mean absolute error (mae)?—arguments against avoiding rmse in the literature. *Geoscientific Model Development*, Copernicus Publications, v. 7, n. 3, p. 1247–1250, 2014. Disponível em: <<https://gmd.copernicus.org/articles/7/1247/2014/>>. 71
- CHANEY, S.; TOMPSON, L.; UHLIG, S. The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal*, v. 21, p. 4–28, 02 2008. 23
- CHEN, Y. et al. Survey on dbscan acceleration algorithms for large scale data. *Journal of Computer Research and Development*, v. 60, n. 9, p. 2028–2047, 2023. 39
- CHO, K. et al. Learning phrase representations using rnn encoder-decoder for statistical machine translation. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. [S.l.]: Association for Computational Linguistics, 2014. p. 1724–1734. 57, 58, 62, 85
- CHOLLET, F. et al. *Keras*. GitHub, 2015. Disponível em: <<https://github.com/fchollet/keras>>. 59

CLEVERT, D.-A.; UNTERTHINER, T.; HOCHREITER, S. Fast and accurate deep network learning by exponential linear units (elus). In: *Proceedings of the 4th International Conference on Learning Representations (ICLR)*. [S.l.: s.n.], 2016. 91

COSTA, A. et al. Previsão de roubo de cargas no estado de são paulo: um estudo de caso. *Revista de Gestão e Secretariado*, v. 15, p. e3786, 06 2024. 28, 30

Crowell and Moring LLP. *Cartels, Foreign Terrorist Organizations, and the High Stakes for Businesses*. 2025. Disponível em: <<https://www.crowell.com/en/insights/client-alerts/cartels-foreign-terrorist-organizations-and-the-high-stakes-for-businesses>>. Acesso em: 1 abr. 2026. 22

DEMORI, L.; OLIVEIRA, C. *Arsenal global: rodamos as ruas do Rio depois de tiroteios e achamos balas da Guerra Fria*. The Intercept Brasil, 2019. Disponível em: <<https://www.intercept.com.br/2019/12/16/rio-municoes-balas-eua-bosnia-russia-guerra-fria/>>. Acesso em: 4 abr. 2026. 21

DIAZ, F.; CERNA, N.; LIZA, R. Artificial intelligence and crime in latin america: A multilingual bibliometric review (2010–2025). *Information*, v. 16, n. 11, 2025. ISSN 2078-2489. Disponível em: <<https://www.mdpi.com/2078-2489/16/11/1001>>. 51

DICKEY, D. A.; FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, Taylor & Francis, v. 74, n. 366a, p. 427–431, 1979. 60

DOLATSHAH, M.; HADIAN, A.; MINAEI-BIDGOLI, B. Ball-tree: Efficient spatial indexing for constrained nearest-neighbor search in metric spaces. *arXiv preprint arXiv:1511.00628*, abs/1511.00628, 2015. Disponível em: <<https://arxiv.org/abs/1511.00628>>. 35

DUARTE, A. L. S. et al. Data protection in brazil: Applying text mining in court documents. *Engineering Proceedings*, v. 87, n. 1, 2025. ISSN 2673-4591. Disponível em: <<https://www.mdpi.com/2673-4591/87/1/57>>. 29

ELMASRI, R.; NAVATHE, S. B. *Sistema de Banco de Dados*. 6. ed. [S.l.]: Pearson Education do Brasil, 2011. 24

EMAANI, S.; SAGHAEI, A. Driver anomaly detection in cargo terminal. *Heliyon*, v. 11, n. 2, p. e41567, 2025. ISSN 2405-8440. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2405844024175986>>. 44

ESTER, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: *KDD Proceedings*. [S.l.: s.n.], 1996. v. 96, p. 226–231. 29

FBSP. *18º Anuário Brasileiro de Segurança Pública*. São Paulo, 2024. Disponível em: <<https://forumseguranca.org.br/>>. 21

FERNANDES, L. E. *A Internacional da Lava Jato: imperialismo, nova direita e o combate à corrupção como farsa*. São Paulo: Autonomia Literária, 2024. ISBN 97865549701. 22, 28

FILHO, L. S. *Estimativa das equações de oferta e demanda de segurança pública nas rodovias federais do Brasil 2020 - 2022*. Dissertação (Mestrado) — Universidade de Brasília, Brasília, 2023. Disponível em: <<http://www.rlbea.unb.br/jspui/handle/10482/47992>>. 29

Fórum Brasileiro de Segurança Pública. O rastro da pólvora e a cegueira deliberada: a quem interessa o descontrole das munições no brasil? *Fonte Segura*, 2025. Disponível em: <<https://fontesegura.forumseguranca.org.br/o-rastro-da-polvora-e-a-cegueira-deliberada-a-quem-interessa-o-descontrole-das-municoes-no-brasil/>>. Acesso em: 20 de junho de 2026. 21

FOUCAULT, M. *Vigiar e punir: nascimento da prisão*. 42. ed. Petrópolis: Editora Vozes, 2014. 56

GANDHI, O. et al. Energy intensity trend explained for sao paulo state. *Renewable and Sustainable Energy Reviews*, Elsevier, v. 77, p. 1046–1054, 2017. 19

GARCIA, G. et al. Crimanalyzer: Understanding crime patterns in são paulo. *IEEE transactions on visualization and computer graphics*, IEEE, v. 27, n. 4, p. 2313–2328, 2019. 30

GARCIA-ZANABRIA, G. et al. Mirante: A visualization tool for analyzing urban crimes. In: *2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*. [S.l.: s.n.], 2020. p. 148–155. 30

GARCIA-ZANABRIA, G. et al. Cripav: Street-level crime patterns analysis and visualization. *IEEE Transactions on Visualization and Computer Graphics*, v. 28, n. 12, p. 4000–4015, 2022. 28

GIAMPAOLO, F. *Foundations of the Learning Paradigm in Neural Networks*. [S.l.]: Springer, 2023. 59

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. Cambridge, MA: MIT Press, 2016. Disponível em: <<http://www.deeplearningbook.org>>. 57, 89

- GU, J. et al. Dynamic correlation adjacency matrix based graph neural networks for traffic flow prediction. *Sensors*, v. 23, n. 6, p. 2897, 2023. Disponível em: <<https://www.mdpi.com/1424-8220/23/6/2897>>. 32, 81
- HAKLAY, M.; WEBER, P. Openstreetmap: User-generated street maps. *IEEE Pervasive Computing*, v. 7, n. 4, p. 12–18, 2008. 34
- HU, Z. et al. Cost-sensitive gnn-based imbalanced learning for mobile social network fraud detection. *IEEE Transactions on Computational Social Systems*, IEEE, v. 11, n. 2, p. 2675–2690, 2024. 89
- InfoMoney. *Operação Carbono Oculto: entenda o esquema bilionário do PCC no setor de combustíveis*. 2025. Disponível em: <<https://www.infomoney.com.br/brasil/operacao-carbono-oculto-entenda-o-esquema-bilionario-do-pcc-no-setor-de-combustiveis/>>. Acesso em: 1 abr. 2026. 20
- International Council on Clean Transportation. *Brazil Freight Assessment*. 2021. Disponível em: <<https://theicct.org/wp-content/uploads/2021/12/brazil-freight-assessed-sept21.pdf>>. Acesso em: 10 out. 2025. 19
- IstoÉ. *A ligação entre a operação contra a Refit e o MBL*. 2025. Disponível em: <<https://istoe.com.br/escandalo-fiscal-refit-conecta-com-mbl>>. Acesso em: 1 abr. 2026. 20
- JOHNSON, J. M.; KHOSHGOFTAAR, T. M. Survey on deep learning with class imbalance. *Journal of Big Data*, Springer, v. 6, n. 1, p. 27, 2019. 89
- JOTA. *A privatização da Sabesp e os limites da eficiência privada*. 2023. Disponível em: <<https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/a-privatizacao-da-sabesp-2>>. Acesso em: 1 abr. 2026. 20
- JUSTUS, M. et al. Crime against trading: The case of cargo theft in são paulo. In: _____. *Retail Crime: International Evidence and Prevention*. Cham: Springer International Publishing, 2018. p. 297–323. ISBN 978-3-319-73065-3. 28, 43, 83
- KEZIOU, P.; DOULKERIDIS, C. Hot spot analysis for big trajectory data in road networks. *Proceedings of the Workshops of the EDBT/ICDT 2025 Joint Conference*, 2025. 54
- KIPF, T. N.; WELLING, M. Semi-supervised classification with graph convolutional networks. In: *Proceedings of the 5th International Conference on Learning Representations (ICLR)*. OpenReview.net, 2017. Disponível em: <<https://openreview.net/forum?id=SJU4ayYgl>>. 31, 80, 84

- LI, Q.; HAN, Z.; WU, X.-M. Deeper insights into graph convolutional networks for semi-supervised learning. In: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*. [S.l.]: AAAI Press, 2018. p. 3538–3545. 88
- LIU, H. et al. Spatial impact of environmental and socio-economic factors on property crime. *ISPRS International Journal of Geo-Information*, MDPI, v. 11, n. 8, p. 428, 2022. 57, 69, 76
- LUCE, M. S. *Teoria marxista da dependência – problemas e categorias, uma visão histórica*. São Paulo: Expressão Popular, 2018. 27
- LUKÁCS, G. *História e consciência de classe: estudos sobre a dialética marxista*. São Paulo: Martins Fontes, 2003. 56
- MARINI, R. M. *Dialética da dependência*. Instituto de Estudos Socioeconômicos, 1973. Disponível em: <<https://marxists.architexturez.net/portugues/marini/1973/mes/dialetica.htm>>. Acesso em: 1 abr. 2026. 20, 27
- MASCARO, A. L. *Estado e Forma Política*. São Paulo: Boitempo, 2013. 55, 97
- MELO, S. N. D. et al. Spatial/temporal variations of crime: A routine activity theory perspective. *International journal of offender therapy and comparative criminology*, SAGE Publications Sage CA: Los Angeles, CA, v. 62, n. 7, p. 1967–1991, 2018. 28
- MICROSOFT BRASIL. *Parceria entre Governo de SP e Microsoft amplia ações de inteligência policial*. 2014. Disponível em: <<https://news.microsoft.com/pt-br/parceria-entre-governo-de-sp-e-microsoft-amplia-acoes-de-inteligencia-policial/>>. Acesso em: 20 de junho de 2026. 23
- MINGARDI, G. O estado e o crime organizado. *São Paulo em Perspectiva*, Fundação Seade, v. 15, n. 3, p. 53–63, 2001. 20
- NAGIN, D. S. Deterrence in the twenty-first century. *Crime and Justice*, University of Chicago Press, v. 42, n. 1, p. 199–263, 2013. 56, 76
- NEVES, I. *COMO AS FORÇAS ARMADAS AJUDAM O CRIME ORGANIZADO*. Canal História Pública, 2026. Vídeo no YouTube. Disponível em: <<https://www.youtube.com/watch?v=213nLdTskJ4>>. Acesso em: 4 abr. 2026. 21
- NEVES, I. *TRUMP QUER CLASSIFICAR PCC E CV COMO TERRORISTAS*. Canal História Pública, 2026. Vídeo no YouTube. Disponível em: <https://www.youtube.com/watch?v=a2DL8c86_ZY>. Acesso em: 4 abr. 2026. 22

New Eastern Outlook. *Shield of the Americas: The Geopolitics Behind US-Led Counter-Cartel Coalition*. 2026. Disponível em: <<https://journal-neo.su/2026/03/15/shield-of-the-americas-the-geopolitics-behind-us-led-counter-cartel-coalition/>>. Acesso em: 1 abr. 2026. 22

OPENSTREETMAP. *OpenStreetMap API Documentation*. 2026. Disponível em: <<https://www.openstreetmap.org/>>. Acesso em: 29 set. 2025. 34

OSORIO, J. *O Estado no centro da mundialização: padrões de acumulação e forças políticas na América Latina*. São Paulo: Outras Expressões, 2019. 27

Overhaul. *Brazil Annual Report 2023*. 2023. Disponível em: <<https://www.over-haul.com/reports/brazil-annual-report-2023>>. Acesso em: 8 dez. 2025. 19, 29

Overhaul. *Brazil Annual Cargo Theft Report 2024*. 2024. Disponível em: <<https://www.over-haul.com/reports/brazil-annual-report-2024>>. Acesso em: 8 dez. 2025. 19, 29

Overhaul. *Brazil Cargo Theft Report Q3: 2025*. 2025. Disponível em: <<https://www.over-haul.com/reports/brazil-q3-2025-cargo-theft-report>>. Acesso em: 8 dez. 2025. 43

PACHUKANIS, E. B. *Teoria Geral do Direito e Marxismo*. São Paulo: Boitempo, 2017. 55, 97

PASCANU, R.; MIKOLOV, T.; BENGIO, Y. On the difficulty of training recurrent neural networks. In: *International Conference on Machine Learning (ICML)*. [S.l.: s.n.], 2013. p. 1310–1318. 83

PAZ, I. S. da. *Cresce a apreensão de armas de estilo militar no Sudeste, revela pesquisa do Instituto Sou da Paz*. 2023. Acesso em: 23 maio 2026. Disponível em: <<https://soudapaz.org/cresce-a-apreensao-de-armas-de-estilo-militar-no-sudeste-revela-pesquisa-do-instituto-sou-da-paz/>>. 21

PAZ, I. S. da. *Mercado ilegal de armas no Sudeste se moderniza impulsionado por decretos de Bolsonaro, revela estudo inédito*. 2023. Acesso em: 23 maio 2026. Disponível em: <<https://soudapaz.org/mercado-ilegal-de-armas-no-sudeste-se-moderniza-impulsionado-por-decretos-de-bolsonaro-revela-estudo-inedito/>>. 21

PEREZ-FERNANDEZ, O. et al. Spatio-temporal mapping of violence against women: An urban geographic analysis based on 911 emergency reports in monterrey. *ISPRS International Journal of Geo-Information*, v. 14, n. 10, 2025. ISSN 2220-9964. Disponível em: <<https://www.mdpi.com/2220-9964/14/10/367>>. 30

- PICCELLI, A. M. Neoliberalismo, crime organizado e milícia nos morros cariocas nos anos 1990 e 2000. *Universidade Estadual de Londrina (UEL)*, 2013. Disponível em: <<https://repositorio.uel.br/bitstreams/558ff10d-82ff-44f7-8dba-6fc0689a3f6c/download>>. 19
- Polícia Militar de São Paulo. *Bases e Quartéis da Polícia Militar Rodoviária*. 2025. Disponível em: <https://www4.policiamilitar.sp.gov.br/unidades/cprv/?page_id=249>. Acesso em: 4 set. 2025. 35
- Polícia Rodoviária Federal. *Unidades da Polícia Rodoviária Federal em São Paulo: endereço e telefone*. 2025. Disponível em: <<https://www.gov.br/prf/pt-br/canais-de-atendimento/unidades-prf/sao-paulo>>. Acesso em: 4 set. 2025. 35
- POULANTZAS, N. *Poder político e classes sociais*. Campinas: Unicamp, 2019. 55
- PRASETYA, D. A. et al. Resolving the shortest path problem using the haversine algorithm. *Journal of Critical Reviews*, v. 7, n. 1, p. 62–64, 2020. 38
- PRECHELT, L. Early stopping—but when? In: *Neural Networks: Tricks of the Trade*. [S.l.]: Springer, 2012. p. 53–67. 73
- PRODESP. *Microsoft - Tabela de Preços de Produtos e Serviços: Situação do Acordo*. 2024. Acordo vigente a partir de 22 de agosto de 2024, pelo prazo de 60 meses. Disponível em: <<https://www.prodesp.sp.gov.br/clientes/acordos-de-produtos-e-servicos/microsoft>>. Acesso em: 20 de junho de 2026. 23
- QUEIROZ, K. V. et al. Analysis of strategies to combat cargo theft and robbery in peripheral communities of são paulo, brazil, using a paraconsistent expert system. *Logistics*, v. 9, n. 1, 2025. ISSN 2305-6290. Disponível em: <<https://www.mdpi.com/2305-6290/9/1/37>>. 29
- RAMIREZ, J. L. H. *A Geographic Intelligence Model of Criminal Groups: A study of cargo theft on Mexican Highways*. Tese (Doutorado) — UCL (University College London), 2024. 44
- REPOLHO, H. M. et al. Cargo theft weighted vehicle routing problem: modeling and application to the pharmaceutical distribution sector. *Soft Computing*, Springer, v. 23, n. 14, p. 5865–5882, 2019. 29
- Revista Oeste. *Comentarista político e membro do MBL é investigado em operação contra a Refit*. 2025. Disponível em: <<https://revistaoste.com/politica/comentarios-politico-e-investigado-em-operacao-contra-a-refit/>>. Acesso em: 1 abr. 2026. 20

- ROSHANKAR, R.; KEYVANPOUR, M. R. Spatio-temporal graph neural networks for accurate crime prediction. In: *13th International Conference on Computer and Knowledge Engineering (ICCKE)*. IEEE, 2023. p. 168–173. Disponível em: <<https://ieeexplore.ieee.org/document/10326223>>. 31, 80
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, Nature Publishing Group, v. 323, n. 6088, p. 533–536, 1986. 58
- RUSCH, T. K.; BRONSTEIN, M. M.; MISHRA, S. A survey on oversmoothing in graph neural networks. *arXiv preprint arXiv:2303.10993*, 2023. Disponível em: <<https://doi.org/10.48550/arXiv.2303.10993>>. 88
- RUSCHE, G.; KIRCHHEIMER, O. *Punição e estrutura social*. Rio de Janeiro: Editora Revan, 2004. 27, 55, 76
- SAFATLE, V. *Quando a indignação não leva a nada*. 2012. Acesso em: 23 maio 2026. Disponível em: <<https://sp.cut.org.br/noticias/safatle-quando-a-indignacao-nao-leva-a-nada-3525>>. 20
- SCHARF, H. Local indicators of spatial association (lisa). In: _____. *Wiley StatsRef: Statistics Reference Online*. Hoboken, NJ: John Wiley & Sons, 2022. p. 1–9. ISBN 9781118445112. 30
- SCOTT, D. W. Kernel density estimation. In: _____. *Wiley StatsRef: Statistics Reference Online*. Hoboken, NJ: John Wiley & Sons, 2018. ISBN 9781118445112. 38
- Secretaria de Segurança Pública de São Paulo. *Portal da Transparência da Secretaria de Segurança Pública de São Paulo*. 2025. Disponível em: <<http://www.ssp.sp.gov.br/transparenciassp/Consulta.aspx>>. Acesso em: 5 jul. 2025. 33, 42
- SHAO, Z. et al. Pre-training enhanced spatial-temporal graph neural network for multivariate time series forecasting. In: *Proceedings of the 28th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. [S.l.]: ACM, 2024. 80
- SHI, Z.; PUN-CHENG, L. S. Spatiotemporal data clustering: A survey of methods. *ISPRS International Journal of Geo-Information*, v. 8, n. 3, 2019. ISSN 2220-9964. Disponível em: <<https://www.mdpi.com/2220-9964/8/3/112>>. 30
- SHIH, T. Impact of data normalization methods on time series prediction. *IEEE Access*, 2022. 59

- SHU, X.; YE, Y. Knowledge discovery: Methods from data mining and machine learning. *Social Science Research*, Elsevier, v. 110, p. 102817, 2023. 36
- SHUMAN, D. I. et al. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine*, v. 30, n. 3, p. 83–98, 2013. 81
- SILVA, R. A. d. *Análise exploratória de dados de indicadores criminais nos municípios do estado de São Paulo: tendências, correlações e desigualdades regionais*. Dissertação (Dissertação de Mestrado) — Universidade Estadual Paulista (Unesp), São José do Rio Preto - SP, 08 2025. Disponível em: <<https://repositorio.unesp.br/entities/publication/30130f97-51f0-4f18-9528-b2e1219a8a47>>. 30, 39
- SONG, R. et al. Fraud detection of bulk cargo theft in port using bayesian network models. *Applied Sciences*, v. 10, n. 3, 2020. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/10/3/1056>>. 29
- SOUZA, K. C. d. A. e.; MIRANDA, P. F. M. O lado brilhante da criminalidade: neoliberalismo e crime organizado. *Revista Emancipação*, v. 23, 2023. 20
- TANG, J. Modeling temporal-spatial correlations for crime prediction. In: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. [S.l.: s.n.], 2017. p. 497–506. 59
- The Intercept Brasil. *Exército esconde origem de balas usadas em homicídios no Brasil*. The Intercept Brasil, 2026. Disponível em: <<https://www.intercept.com.br/2026/01/07/exercito-origem-balas-homicidios-brasil/>>. Acesso em: 4 abr. 2026. 21
- UNODC. *Transnational Organized Crime in the Americas: Threats and Logistics*. Vienna, 2023. 19
- U.S. Department of State OSAC. *Brazil Country Security Report*. 2025. Disponível em: <<https://www.osac.gov/Content/Report/7ce0bb41-df62-4e92-b4b0-1d02a39be392>>. Acesso em: 10 out. 2025. 19
- U.S. Department of the Treasury. *2024 National Money Laundering Risk Assessment (NMLRA)*. 2024. Disponível em: <<https://home.treasury.gov/news/press-releases/jy2080>>. Acesso em: 1 abr. 2026. 22
- WANG, B. et al. Deep learning for real-time crime forecasting and its ternarization. In: *arXiv preprint arXiv:1711.08833*. [s.n.], 2017. Disponível em: <<https://arxiv.org/abs/1711.08833>>. 23

- WANG, X. et al. Spatiotemporal crime prediction considering structural sparsity and environmental features. *Frontiers in Big Data*, 2024. 57, 69, 76
- WANG, Z. et al. *Uncertainty-Aware Crime Prediction With Spatial Temporal Multivariate Graph Neural Networks*. 2024. Disponível em: <<https://arxiv.org/abs/2408.04193>>. 23
- World Bank. *Logistics Performance Index 2023: Connecting to Compete*. Washington, DC, 2023. 19
- XIA, L. et al. Spatial-temporal sequential hypergraph network for crime prediction with dynamic multiplex relation learning. In: *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*. International Joint Conferences on Artificial Intelligence Organization, 2021. p. 1631–1637. Disponível em: <<http://dx.doi.org/10.24963/ijcai.2021/225>>. 23
- XIAO, H. et al. A method for filling traffic data based on feature-based combination prediction model. *Scientific Reports*, Nature Publishing Group, v. 14, 2024. 57, 71, 74
- ZAFFARONI, E. R. *Em busca das penas perdidas: a perda de legitimidade do sistema penal*. Rio de Janeiro: Revan, 1991. 56, 98
- ZAFFARONI, E. R. *A palavra dos mortos: conferências de criminologia cautelar*. São Paulo: Saraiva, 2012. 98
- ZHA, K. et al. Rank-n-contrast: Learning continuous representations for regression. In: *Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS)*. [S.l.: s.n.], 2023. 89
- ZHENG, C. et al. A load forecasting method of power grid host based on sarima-gru model. In: *National Conference of Theoretical Computer Science*. [S.l.]: Springer, 2021. p. 135–153. 73