

Análise de Risco Criminal em Grafos Viários usando Marcações Territoriais de Facções como Features Preditivas

Victor Carvalho Soares de Araújo¹, Raimundo Santos Moura¹

¹PPGCC – Universidade Federal do Piauí (UFPI)
Teresina – PI – Brasil

victor.araujo@ufpi.edu.br, rsm@ufpi.edu.br

Abstract. *This work investigates the association between criminal organization graffiti and spatial patterns of crime in Teresina-PI. Using georeferenced data of graffiti from PCC, B40, and GDE factions, crime records (2022-2023), and urban POIs, a risk classification model was developed on street network graphs. Through LightGBM and SHAP analysis, a statistically robust association between proximity to graffiti and crime occurrence was demonstrated. Territorialization features showed substantial gain in AUPRC over models based solely on POIs and street characteristics, with greater impact on non-violent crimes. This is the first systematic incorporation of criminal territorial markings into spatial risk models, offering a replicable methodology for urban crime analysis in Brazil.*

Resumo. *Este trabalho investiga a associação entre pichações de organizações criminosas e padrões espaciais de criminalidade em Teresina-PI. Utilizando dados georreferenciados de pichações das facções PCC, B40 e GDE, registros criminais (2022-2023) e POIs urbanos, desenvolveu-se modelo de classificação de risco em grafos viários. Mediante LightGBM e análise SHAP, demonstrou-se associação robusta entre proximidade a pichações e ocorrência de crimes. Features de territorialização apresentaram ganho substancial em AUPRC sobre modelos baseados apenas em POIs e características viárias, com maior impacto em crimes não-violentos. Esta é a primeira incorporação sistemática de marcações territoriais criminosas em modelos de risco espacial, oferecendo metodologia replicável para análise criminal urbana no Brasil.*

1. Introdução

A demarcação territorial mediante inscrições urbanas constitui prática recorrente de organizações criminosas para estabelecer domínio espacial [Ley and Cybriwsky 1974]. Tais marcações funcionam como código visual que comunica posse, intimida grupos rivais e demarca fronteiras invisíveis [Hughes et al. 2022], evidenciando como o espaço urbano se torna palco de conflitos territoriais associados à criminalidade organizada.

Este fenômeno manifesta-se de forma expressiva em Teresina-PI, município com população estimada em 860 mil habitantes [IBGE 2023]. A partir de 2022, observou-se intensificação de marcações do Primeiro Comando da Capital (PCC), Bonde dos 40 (B40) e Guardiões do Estado (GDE), coincidente com expansão interestadual destas organizações. A documentação georeferenciada revelou distribuição espacial estruturada,

sugerindo estratégias planejadas de territorialização [Brantingham et al. 2012] e convertendo tais marcações em manifestação cartográfica do domínio criminal.

A análise espacial da criminalidade tem evoluído de mapas de calor para metodologias que reconhecem a concentração de crimes em microlocais específicos [Weisburd 2015] e as associações contextuais entre pontos de interesse urbanos e criminalidade [Silva et al. 2025].

Contudo, no contexto brasileiro, a literatura não registra estudos que utilizem dados geolocalizados de marcações territoriais de facções como *features* em modelos preditivos de risco criminal, lacuna relevante considerando que a territorialização de facções constitui fenômeno urbano crescente em diversas regiões do país. Este artigo pretende responder duas questões de pesquisa:

P1: Existe associação mensurável entre proximidade a pichações de organizações criminosas e ocorrência de crimes?

P2: A magnitude desta associação difere significativamente entre crimes violentos e não-violentos?

Para responder estas questões, desenvolveu-se um modelo de classificação de risco espacial utilizando representação em *line graph* da malha viária de Teresina ao nível de segmentos, incorporando 659 pontos georreferenciados de pichações de facções mapeados através de levantamento de campo, além de 2.481 Pontos de Interesse (POIs) e 35.902 crimes geocodificados ao longo de 24 meses. As *features* utilizadas são majoritariamente estáticas, refletindo condições estruturais do ambiente urbano.

Os resultados preliminares demonstram uma associação relevante entre proximidade a territórios de facções e ocorrência de crimes. O modelo de aprendizado de máquina supervisionado *LightGBM* incorporando *features* de pichações de organizações criminosas alcançou Area Under the Precision-Recall Curve (AUPRC) de 0,0334 para crimes gerais, representando capacidade discriminatória 59,9% superior ao modelo *baseline* com *features* viárias e POIs urbanos (AUPRC = 0,0209). A análise de importância de *features* indicou que variáveis relacionadas a facções dominam o sinal discriminatório, com 7 das 10 *features* mais importantes sendo métricas de proximidade a territórios criminosos.

O restante do artigo está organizado da seguinte forma: a Seção 2 apresenta trabalhos relacionados em análise criminal espacial e aprendizado de máquina; a Seção 3 descreve metodologia detalhada incluindo construção do grafo viário, engenharia de *features* e configuração do modelo; a Seção 4 apresenta os resultados experimentais para responder as perguntas de pesquisa; a Seção 5 conclui o artigo apontando direções futuras.

2. Trabalhos Relacionados

2.1. Discretização Espacial e Concentração Criminal

A “lei da concentração criminal” foi demonstrada empiricamente [Weisburd 2015], evidenciando que os crimes se concentram em microlocais específicos ao invés de distribuírem-se homogeneamente em grandes áreas. Este achado tem impulsionado o desenvolvimento de metodologias de discretização espacial refinada, incentivando uma análise do crime com maior nível de detalhe.

O sistema MIRANTE foi proposto para análise visual de criminalidade utilizando rede viária como discretização espacial, argumentando que segmentos de rua constituem unidade mais natural que grades regulares [Garcia-Zanabria and Nonato 2022]. Outra abordagem emprega *line graphs* onde segmentos viários tornam-se nós e adjacências tornam-se arestas, possibilitando representação natural da conectividade urbana e aplicação de algoritmos baseados em grafos [Hassan et al. 2024].

O sistema CriPAV foi desenvolvido para análise de crimes em nível de rua (*street-level*) utilizando interseções de ruas como unidade de análise [García-Zanabria et al. 2022]. Os autores usaram matriz estocástica para calcular probabilidade de crimes e *autoencoders* para agrupamento temporal de *hotspots*, demonstrado que crimes frequentes e não intensos concentram-se em avenidas principais, enquanto crimes esporádicos e intensos distribuem-se em ruas secundárias. A visualização integrada com *Google Street View* permitiu análise temporal de mudanças na infraestrutura urbana correlacionadas a padrões criminais.

Um modelo de grafo integrado foi proposto para previsão diária de crimes em Chicago, USA, usando distritos policiais como unidade espacial [Hou et al. 2022]. Integrando ResNet, GCN e LSTM com atenção, alcançaram MAE de 1,38 incidentes para roubo, embora reconhecessem que atuação policial interfere no risco. Nosso trabalho difere ao modelar territórios de organizações criminosas como *features* em nível de segmento viário, capturando territorialização em microlocais.

No estado do Alagoas, uma metodologia *corner-level* analisou a influência de POIs em criminalidade urbana em Maceió e Arapiraca [Silva et al. 2025], identificando que diferentes tipos de POIs apresentam associações variadas com crime dependendo do contexto socioeconômico. Este trabalho demonstra a importância de considerar heterogeneidade espacial e contextual, princípio incorporado no nosso estudo através de análise estratificada por tipo de crime.

2.2. Integração de *Features* Urbanas Heterogêneas

A integração de dados urbanos heterogêneos (POIs, infraestrutura viária, indicadores socioeconômicos, séries temporais de crimes em discretizações espaciais distintas) constitui desafio metodológico central para análise criminal espacial. A dificuldade reside na incompatibilidade entre diferentes tipos de dados e sistemas de referência espacial: pontos, linhas, polígonos e escalas temporais estáticos vs dinâmicos.

A biblioteca *CityHub* foi desenvolvida para fusão de *datasets* urbanos em grafos viários [Salinas et al. 2022], suportando camadas para POIs, áreas administrativas, zonas censitárias e eventos espaço-temporais. Esta biblioteca foi utilizada para integrar POIs e crimes geocodificados ao grafo viário, permitindo agregação sistemática de *features* heterogêneas em representação unificada.

Uma metodologia baseada em *Graph Autoencoders (GAE)* foi proposta para fusão de *features* estáticas e dinâmicas em análise criminal urbana [Pocco et al. 2025], demonstrando que componentes estáticos (infraestrutura, POIs) e dinâmicos (séries temporais) contribuem com informação complementar. Embora nosso trabalho utilize *features* majoritariamente estáticas, esta abordagem de fusão multimodal inspirou a incorporação simultânea de *features* viárias, POIs e *features* de facções em arquitetura unificada.

2.3. Machine Learning para Análise Criminal Espacial

A aplicação de aprendizado de máquina em análise criminal evoluiu de modelos simples de regressão para arquiteturas sofisticadas incluindo *gradient boosting* e redes neurais em grafos. A escolha do algoritmo requer consideração de múltiplos fatores: i) interpretabilidade, ii) eficiência computacional, iii) robustez ao desbalanceamento de classes e iv) capacidade de processar *features* heterogêneas.

O algoritmo *Gradient Boosting Decision Trees (GBDT)*, especialmente o *LightGBM* [Ke et al. 2017], tem demonstrado desempenho competitivo em classificação tabular. Suas inovações (*Gradient-based One-Side Sampling* e *Exclusive Feature Bundling*) reduzem custo computacional preservando precisão, tornando-o particularmente adequado para análise criminal com desbalanceamento extremo de classes, esparsidade de *features* e outras particularidades desse domínio.

Um estudo comparou *Graph Neural Networks* dinâmicas com *Random Forest* para predição criminal na cidade de São Paulo [Hassan et al. 2024], demonstrando que métodos *ensemble* tradicionais podem superar GNNs em problemas com *features* estáticas e desbalanceamento extremo. Este achado motivou a escolha de *LightGBM*, complementada por análise *SHapley Additive exPlanations (SHAP)* [Lundberg and Lee 2017] para interpretabilidade.

Explicações contrafactuais para análise criminal [Raimundo et al. 2025] demonstram que interpretabilidade constitui requisito crítico para aceitação em sistemas de apoio à decisão policial, evidenciando que modelos devem fornecer *insights* acionáveis além de predição acurada. Nosso estudo incorpora múltiplas técnicas de interpretabilidade: *feature importance* via ganho de informação, *SHAP values* para decomposição aditiva, e análise de correlação entre *features* e contribuições SHAP.

Embora trabalhos anteriores tenham incorporado variáveis socioeconômicas [Bogomolov et al. 2014], infraestrutura urbana [Silva et al. 2025, Garcia-Zanabria and Nonato 2022], estratégias policiais [Hou et al. 2022], dados de redes sociais [Gerber 2014] e padrões de mobilidade [Wang et al. 2016], nenhum estudo usou dados geolocalizados de territórios de organizações criminosas como *features* em modelos de análise de risco criminal.

A ausência de estudos sobre o tema deve-se à dificuldade de obtenção de dados sobre domínio territorial, frequentemente confidenciais ou incompletos. Pichações constituem fonte observacional que permite mapeamento sistemático. Este trabalho contribui nesse aspecto com dados georreferenciados de três organizações criminosas em Teresina (PCC, B40, GDE), obtidos por levantamento de campo sistemático.

3. Metodologia

3.1. Dados e Área de Estudo

Os dados utilizados são provenientes do Sistema de Despacho de Ocorrências Policiais (SysPM/PMPI), disponibilizados pela Diretoria de Inteligência da Polícia Militar do Piauí para fins de pesquisa acadêmica. Após processo de unificação, filtragem, normalização e delimitação espacial e temporal, restaram 35.653 registros geocodificados em Teresina ao longo de 24 meses (jan/2022 a dez/2023).

Com base no Código Penal Brasileiro [Brasil 1940], foram classificados como “violentos” sete tipos de ocorrências: *Homicídio doloso*, *Lesão corporal grave/gravíssima dolosa*, *Latrocínio*, *Feminicídio*, *Estupro*, *Lesão corporal seguida de morte* e *Roubo*. Incorporou-se também 2.481 POIs extraídos do *OpenStreetMap*, distribuídos em 17 *features* relacionadas a transporte, educação, saúde, lazer, comércio e segurança pública.

Para as pichações, foram mapeados 659 pontos georreferenciados de organizações criminosas nas vias de Teresina entre os meses de nov/2022 e fev/2023, coletados por meio de levantamento de campo in loco com registro fotográfico e georreferenciamento. A distribuição de pichações por organização foi: PCC (321 pontos), B40 (326 pontos) e GDE (12 pontos). A Figura 1 mostra de forma gráfica o panorama geral dos dados de crimes na cidade de Teresina e a Figura 2 mostra exemplos de pichações.

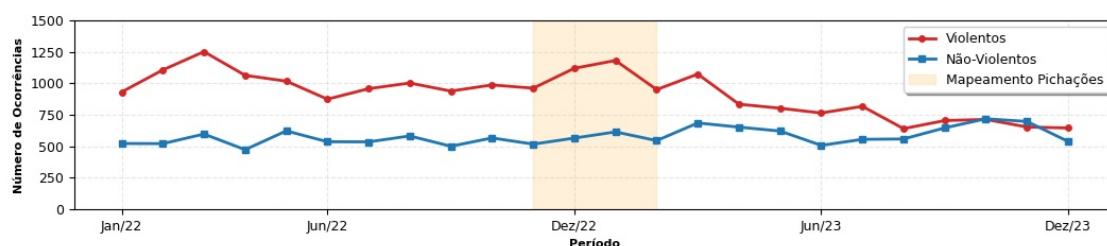


Figura 1. Evolução temporal de crimes e mapeamento de pichações.

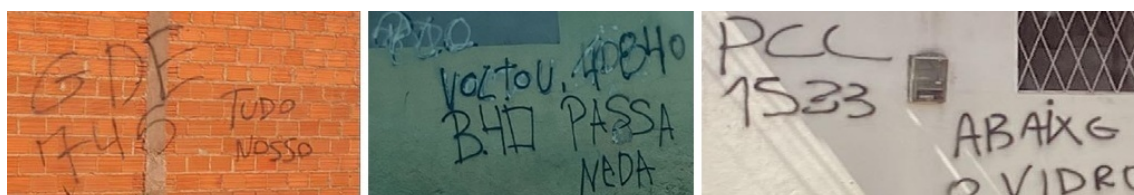


Figura 2. Exemplos de pichações georreferenciadas: GDE, B40 e PCC.

3.2. Construção do Grafo Viário

Extraíu-se a rede viária do *OpenStreetMap* [Hassan et al. 2024] e converteu-se em *line graph* onde nós representam segmentos viários e arestas representam adjacências (244.053 conexões). Os crimes foram vinculados aos segmentos via *spatial join* (*GeoPandas*) em raio de 200 m, associando cada crime a um único segmento.

A estrutura do grafo foi utilizada para engenharia de *features*, convertendo a conectividade espacial em variáveis numéricas. Para cada segmento, foram calculados seu grau de conectividade e estatísticas agregadas da vizinhança, como a soma e a média de crimes ocorridos nos segmentos adjacentes em períodos anteriores.

Cada segmento viário recebeu 49 *features*, agrupadas em três categorias:

- **Features Viárias (14 features):** Representação *one-hot encoding* dos tipos de via, segundo classificação do *OpenStreetMap*.
- **POIs Urbanos (17 features):** Para cada categoria de POI, computou-se contagem de pontos em *buffer* de 200 m ao redor do segmento. Estas *features* capturam infraestrutura urbana e possibilidades de crimes.

- **Features de Facções (18 features):** Para cada organização $g \in \{PCC, B40, GDE\}$, computaram-se 6 *features* baseadas em proximidade espacial e *smoothing* através de vizinhança no grafo.

Criou-se *snapshots* mensais gerando 1.905.240 instâncias (79.385 segmentos $\times$ 24 meses), com *labels* binárias por presença de crime. A divisão temporal foi: treino (meses 1-18, 1.428.930 amostras), validação (meses 19-20) e teste (meses 21-24, 317.540 amostras), analisando estabilidade temporal das associações identificadas.

Os hiperparâmetros do modelo *LightGBM* [Ke et al. 2017] foram otimizados via *grid search* com *5-fold cross-validation* temporal (`learning_rate=0.05`, `num_leaves=31`, `class_weight='balanced'`). A etapa de pré-processamento incluiu: i) transformação `log1p` para *features* de distância; ii) padronização preservando esparsidade; e iii) *One-hot encoding* para tipos de via OSM.

Por fim, foi implementado *early stopping* para monitorar a AUPRC no conjunto de validação com janela de 50 iterações, convergindo tipicamente em 250-350 árvores.

4. Resultados

4.1. Estudo de Ablação: Evidência quantitativa para P1

Para responder a primeira pergunta “P1: Existe associação mensurável entre proximidade a territórios de organizações criminosas e ocorrência de crimes?”, conduziu-se estudo de ablação com duas configurações incrementais de *features*. A Figura 3 apresenta os resultados para crimes gerais no conjunto de teste, considerando os modelos *Baseline* e *Completo*.

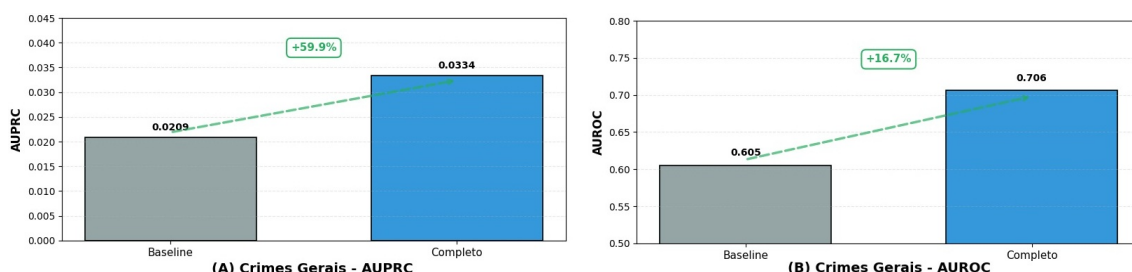


Figura 3. Desempenho comparativo do estudo de ablação para crimes gerais.

O modelo *Baseline* usando POIs urbanos e *features* viárias alcançou AUPRC de 0,0209, representando desempenho $1,95\times$ superior ao *baseline* aleatório (0,0107). O modelo *Completo* usando incorporação de *features* de facções proporcionou ganho de +59,9% (AUPRC: 0,0334), demonstrando que territórios de organizações criminosas contribuem com informação relevante além da infraestrutura urbana convencional. O ganho se mantém consistente considerando a métrica *Area Under the ROC Curve* (AUROC) (*Baseline*: 0,605 e *Completo*: 0,706).

Este resultado constitui evidência quantitativa direta para a pergunta P1, ou seja, existe associação mensurável e substantiva entre proximidade de territórios de facções e ocorrência de crimes. A representatividade do ganho indica que esta associação não é marginal, mas representa componente fundamental para compreensão da dinâmica criminal de Teresina.

4.2. Análise por Tipo de Crime: Resposta a P2

Para responder a segunda pergunta “P2: A magnitude desta associação difere significativamente entre crimes violentos e não-violentos?”, replicou-se o estudo de ablação estratificado por tipo de crime: violentos e não violentos (ver a Figura 4).

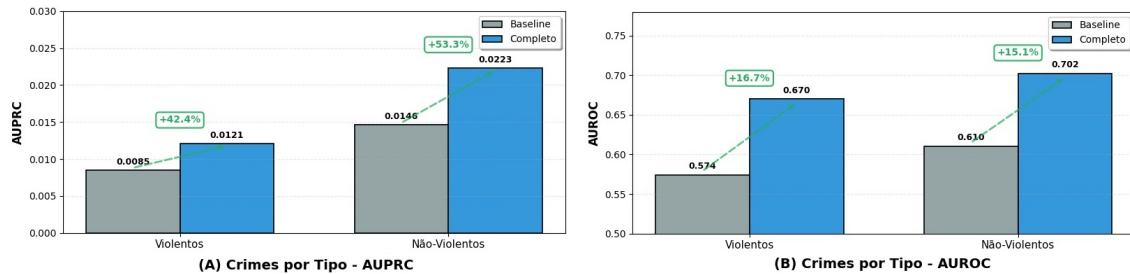


Figura 4. Desempenho do modelo para crimes violentos e não-violentos.

Crimes violentos apresentam ganho de +42,4% com incorporação de *features* de facções, enquanto não-violentos apresentam +53,3%. A diferença de 10,9 pontos percentuais evidencia que a associação entre territórios de facções e criminalidade apresenta magnitude diferenciada por tipo de crime, respondendo afirmativamente a pergunta P2.

A diferença absoluta em AUPRC (violentos: 0,0121 vs não-violentos: 0,0223) reflete não apenas prevalências distintas, mas também padrões espaciais diferenciados. Este achado sugere que crimes não-violentos podem estar mais fortemente estruturados pela economia ilegal territorializada, enquanto crimes violentos podem envolver componentes situacionais adicionais não capturados pelas *features* estáticas utilizadas.

4.3. Feature Importance: Dominância de Facções

A análise de importância via ganho de informação acumulado, métrica que quantifica a redução de impureza proporcionada por cada *feature* nas divisões das árvores [Ke et al. 2017], revela dominância das *features* de facções no modelo para crimes gerais (Tabela 1).

Tabela 1. Importância das 10 Features Mais Relevantes

Rank	Feature	Categoria	Importância (%)
1º	dist_pcc_m	Facções	33,8
2º	dist_b40_m	Facções	29,9
3º	dist_gde_m	Facções	4,4
4º	dist_pcc_m_nbr	Facções	4,2
5º	cnt_pcc_200m_nbr	Facções	4,1
6º	poi_transporte	POIs	4,1
7º	poi_escola	POIs	3,0
8º	cnt_pcc_200m	Facções	1,5
9º	poi_restaurante	POIs	1,5
10º	cnt_b40_200m	Facções	1,5

A análise da importância agregada das 49 *features* por categoria revela clara dominância das facções, que representam 80,4% (≈ 49.890) da importância total. Em contrapartida, as contribuições de POIs e *features* viárias foram 14,2% (≈ 8.835) e 5,3%

(≈ 3.275), respectivamente. Reforçando essa tendência, sete das dez *features* mais relevantes do modelo, quando se considera o *smoothing* espacial, estão relacionadas a facções.

É possível observar um padrão: *features* de distância (*dist_pcc_m*, *dist_b40_m*) têm importância aproximadamente $20\times$ maior que *features* de contagem (*cnt_pcc_200m*, *cnt_b40_200m*), sugerindo que efeito de facções sobre risco criminal é função contínua da distância ao invés de efeito binário de limiar. As *features* de vizinhança (*_nbr_*) aparecem consistentemente no top-10 (4 e 5 do rank), indicando que os segmentos cercados por vizinhos próximos a facções apresentam risco elevado.

4.4. Análise SHAP: Relações Direcionais e Heterogeneidade

Para investigar como as *features* individuais influenciam as previsões, calcularam-se *SHAP values* [Lundberg and Lee 2017] via *TreeExplainer*. O SHAP decompõe cada previsão em contribuições aditivas baseadas em teoria de jogos cooperativos, permitindo análise granular da influência de cada variável no risco predito. Devido ao custo computacional em *datasets* extensos (complexidade $O(n \times m \times t)$), utilizou-se uma amostra estratificada de 2.000 instâncias do conjunto de teste, preservando a proporção das classes.

A Tabela 2 apresenta a correlação de Spearman entre *features* de distância a facções e seus *SHAP values*. Para crimes gerais e não-violentos, o PCC apresenta forte correlação negativa ($r = -0,696$ e $-0,662$), indicando que proximidade aumenta o risco criminal. Para o B40, a correlação é mais fraca ($r = -0,183$ a $-0,251$), sugerindo relação não-linear. A correlação positiva do GDE ($r = +0.375$) pode refletir concentração regional, supressão territorial ou sensibilidade a *outliers* (12 pontos mapeados).

Tabela 2. Correlações SHAP Distância $\times$ SHAP Value

Facção	Crimes Gerais	Crimes Violentos	Crimes Não-Violentos
PCC	$r = -0,696$ ($p < 0,001$)	$r = -0,612$ ($p < 0,001$)	$r = -0,662$ ($p < 0,001$)
B40	$r = -0,183$ ($p < 0,001$)	$r = -0,118$ ($p < 0,001$)	$r = -0,251$ ($p < 0,001$)
GDE	$r = +0,375$ ($p < 0,001$)	$r = +0,353$ ($p < 0,001$)	$r = +0,380$ ($p < 0,001$)

Como o SHAP assume independência entre *features*, a alta correlação espacial entre distâncias a facções e variáveis de vizinhança fragmenta a atribuição de importância, subestimando contribuições individuais. Contudo, a consistência dos sinais na Tabela 2 confirma a robustez das associações. Essa heterogeneidade entre organizações justifica a escolha do *LightGBM* e a análise estratificada.

4.5. Validação Operacional: Análise Top-K

Para validar que as associações estatísticas identificadas têm magnitude e persistência temporal, conduziu-se análise Top-K no conjunto de teste. Esta métrica mensura a proporção de crimes capturados pelos k segmentos de maior risco predito, uma abordagem de validação recorrente em estudos recentes de modelagem de risco espacial [Kronkvist et al. 2025].

Os 5% dos segmentos classificados como maior risco concentram 19,3% dos crimes gerais ao longo dos quatro meses de teste. A diferença entre violentos e não-violentos reflete maior previsibilidade espacial de crimes não-violentos, consistente com análise de ablação.

Tabela 3. Cobertura de Crimes por Percentil (Conjunto de Teste)

Top-K	N Segmentos	Crimes Gerais	Crimes Violentos	Crimes Não-Violentos
1%	794	7,0%	5,6%	8,5%
3%	2.382	13,8%	9,8%	16,4%
5%	3.969	19,3%	15,9%	21,7%
10%	7.939	29,6%	26,1%	32,6%
20%	15.877	45,1%	40,9%	46,4%

A cobertura Top-5% manteve estabilidade temporal (Set: 20,4%, Out: 21,0%, Nov: 18,4%, Dez: 17,3%), consistente com a média agregada de 19,3%. Esta variabilidade moderada, comparável a flutuações sazonais típicas em cidades de médio porte, valida que o modelo identifica padrões estruturais ao invés de artefatos temporais.

5. Conclusão

Este trabalho investigou a associação entre territórios de organizações criminosas e padrões espaciais de criminalidade através de modelagem em grafos viários. Utilizando um *dataset* georreferenciados de pichações em Teresina, demonstrou-se evidência quantitativa robusta que responde às perguntas de pesquisa formuladas. Em resposta a P1, *features* de facções contribuíram com ganho de +59,9% em AUPRC, representando 80,4% da importância do modelo. Quanto a P2, crimes não-violentos apresentaram maior ganho preditivo (+53,3%) que violentos (+42,4%), sugerindo forte associação com territorialização criminal. Os 5% dos segmentos de maior risco concentram 19,3% das ocorrências, demonstrando utilidade prática para priorização de recursos policiais.

Trabalhos futuros incluem: i) comparação com modelos espaciais clássicos e GNN para explorar mais profundamente a estrutura relacional do grafo; e ii) *transfer learning* para testar generalização em outras cidades.

Agradecimentos

O presente trabalho foi financiado parcialmente pela SSP/PI, com apoio da Fundação Cultural e de Fomento à Pesquisa, Ensino, Extensão e Inovação (FADEX). Agradecemos também a equipe técnica da Diretoria de Inteligência da Polícia Militar do Piauí (PMPI) pelo apoio e disponibilidade dos dados.

Referências

- Bogomolov, A., Lepri, B., Staiano, J., Oliver, N., Pianesi, F., and Pentland, A. (2014). Once upon a crime: Towards crime prediction from demographics and mobile data. In *Proc. of the ACM Int. Joint Conference on Pervasive and Ubiquitous Computing*.
- Brantingham, P. J., Tita, G. E., Short, M. B., and Reid, S. (2012). The ecology of gang territorial boundaries. *Criminology*, 50(3):851–885.
- Brasil (1940). Decreto-lei nº 2.848, de 7 de dezembro de 1940. código penal. http://www.planalto.gov.br/ccivil_03/decreto-lei/del2848.htm.
- Garcia-Zanabria, G. and Nonato, L. (2022). Visual crime pattern analysis. In *Anais Estendidos da SIBGRAPI*, Porto Alegre, RS, Brasil. SBC.

- García-Zanabria, G., Raimundo, M., Poco, J., Nery, M., Silva, C., Adorno, S., and Nonato, L. (2022). Cripav: Street-level crime patterns analysis and visualization. *IEEE Transactions on Visualization and Computer Graphics*, 28(12):4000–4015.
- Gerber, M. (2014). Predicting crime using twitter and kernel density estimation. *Decision Support Systems*, 61.
- Hassan, W., Cabral, M., Ramos, T., Filho, A. C., and Nonato, L. (2024). Modeling and predicting crimes in the city of são paulo using graph neural networks. In *Proc. XXXIV Brazilian Conference on Intelligent Systems (BRACIS)*, pages 372–386.
- Hou, M., Hu, X., Cai, J., Han, X., and Yuan, S. (2022). An integrated graph model for spatial-temporal urban crime prediction based on attention mechanism. *ISPRS International Journal of Geo-Information*, 11(5):294.
- Hughes, L. A., Schaible, L. M., and Kephart, T. (2022). Gang graffiti, group process, and gang violence. *Journal of Quantitative Criminology*, 38:365–384.
- IBGE (2023). Censo demográfico 2022: Primeiros resultados. <https://www.ibge.gov.br/estatisticas/sociais/saude/22827-censo-demografico-2022.html>.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, pages 3146–3154.
- Kronkvist, K., Borg, A., Boldt, M., and Gerell, M. (2025). Predicting public violent crime using register and openstreetmap data: A risk terrain modeling approach across three cities of varying size. *Applied Spatial Analysis and Policy*, 18(9).
- Ley, D. and Cybriwsky, R. (1974). Urban graffiti as territorial markers. *Annals of the Association of American Geographers*, 64(4):491–505.
- Lundberg, S. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems 30*, pages 4765–4774.
- Pocco, X., Hassan, W., Salinas, K., Molchanov, V., and Nonato, L. G. (2025). Exploring urban factors with autoencoders: Relationship between static and dynamic features.
- Raimundo, M., Garcia-Zanabria, G., Nonato, L., and Poco, J. (2025). Countercrime - using counterfactual explanations to explore crime reduction scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 31(10):9008–9023.
- Salinas, K., Gonçalves, T., Barella, V., Vieira, T., and Nonato, L. (2022). Cityhub: A library for urban data integration. In *Anais da Conf. on Graphics, Patterns and Images*.
- Silva, D., Vieira, T., Costa, E., Paiva, A., and Nonato, L. (2025). A street corner-level methodology to analyze the influence of points of interest on urban crime. *Socio-Economic Planning Sciences*, 102:102297.
- Wang, H., Kifer, D., Graif, C., and Li, Z. (2016). Crime rate inference with big data. In *Proc. of the ACM SIGKDD Int. Conference on Knowledge Discovery and Data Mining*.
- Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology*, 53.