JOURNAL OF COMPUTING HORIZONS
Vol. 01, No. 01, 2026
DOI: 10.22610/JCH.V1I1.4988

RESEARCH ARTICLE | PEER REVIEWED

A Hybrid Deep Learning and Graph-Based Framework for Crime Hotspot Detection and Prediction

Alina Arif [1], Zuha Rani [1] *, Aqsa Akram [2], Arham Mustafa [1], Umair Muneer Butt [2]

[1] *Department of Artificial Intelligence, University of Management and Technology, Lahore (Sialkot Campus), Pakistan*

[2] *Department of Computer Science, University of Management and Technology, Lahore (Sialkot Campus), Pakistan*

* **Corresponding Author:** zuhamughal0321@gmail.com

ABSTRACT – Prediction and detection of crime hotspots is extremely crucial for public safety and law enforcement. A hybrid deep learning and Graph-Based framework is proposed in this research that integrates GNN (Graph Neural Network), CNN (Convolutional Neural Network), and LSTM (Long Short-Term Memory). The proposed framework is evaluated on the Chicago Crime Dataset (2015-2023) that includes 52917 real records. Our approach captures spatial relationships using beat-level graph construction. The location patterns are learned via CNN-based spatial features, and time-based crime trends are modeled using LSTM-inspired temporal lag features. Class imbalance is handled using the SMOTE technique. The proposed hybrid model is compared against 8 baseline models: Logistic Regression, K-Nearest Neighbors, Random Forest, Gradient Boosting, XGBoost, standalone GNN, CNN, and LSTM, and the results demonstrate that the Hybrid model outperforms all of them in AUC-ROC. This research can assist law enforcement agencies in developing proactive crime prevention strategies.

Keywords: *GNN; LSTM; CNN models; crime hotspot prediction; spatio-temporal; logistic regression.*

I. INTRODUCTION

Crime hotspot prediction is an important research area because crime is often concentrated in distinct areas, a pattern widely acknowledged in spatial criminology [1]. Early approaches leveraged statistical modeling and regression analysis after aggregating crime incidents into geographical regions, so that their relationships with socioeconomic and environmental factors could be examined [2].

Predictive hotspot techniques were later proposed, with most incorporating Kernel Density Estimation (KDE), which defines high-density areas as hotspots by assigning higher weights to nearby events [3]. This approach is implemented as hotspot policing in law enforcement, where police focus their resources on areas with high crime risk [4].

The research subsequently incorporated temporal patterns along with spatial patterns, which in turn led to models such as Spatio-Temporal KDE, which accounted for both space and time [3]. Apart from that, models such as the Self-Exciting Point Process were also employed to model the temporal dependencies among crime events [5, 6].

In recent years, ML and DL approaches have advanced hotspot prediction by enabling models to learn spatial and temporal features automatically. ST-HGnet (Spatio-temporal hierarchical gating network) is a progressive model in contemporary approaches that effectively captures multi-scale spatial and temporal dependencies [7].

Hotspot prediction contributes significantly to policing strategies, as suggested by a Delhi Police study, in which patrolling and resource allocation were enhanced using data-driven hotspot mapping [8]. These Spatial relationships remain unacknowledged because prevailing crime hotspot prediction methods often rely on traditional statistical and machine learning approaches that treat crime events as independent. This limitation prevents models from effectively capturing geographical interactions among neighboring regions. Emerging investigations have proposed graph convolutional networks (GCNs) that model spatial relationships through graph structures, but the full representation of urban crime dynamics remains challenging [9].

Moreover, several approaches are unable to account for temporal aspects, as crime patterns evolve and exhibit long-term dependencies [10]. Deep learning models like CNNs and LSTMs are implemented individually, with CNNs capturing spatial features and LSTMs capturing temporal patterns. Still, due to their isolated use, a complete spatio-temporal comprehension remains missing [11]. Hybrid models (CNN-LSTM or GCN-LSTM) have been proposed to overcome this limitation, but they are unable to effectively model spatial and temporal interdependencies in a fully integrated manner [12].

It is evident from the existing literature that machine learning models ignore spatial dependencies in crime hotspot prediction, resulting in an inadequate capture of geographical interactions [13]. CNN- and LSTM-based deep learning

models, on the other hand, provide partial modeling, where CNN learns spatial patterns and LSTM learns temporal patterns individually. However, a fully-fledged combined understanding is still missing [14].

GNNs model spatial relationships effectively but are limited in capturing temporal dynamics [15]. Therefore, there is a lack of a unified spatio-temporal framework that integrates GNN, CNN, and LSTM in the existing literature to fully address crime hotspot prediction. Subsequent hybrid frameworks like CriPaap have incorporated LSTM and spatio-temporal GNN (ST-GNN) along with clustering techniques to optimize performance, which is indeed a strong step in this direction; however, the scope for fully unified and optimized integration remains.

This research primarily focuses on developing an enhanced hybrid framework that integrates CNN, GNN, and LSTM to effectively model [16]. This framework will leverage GNNs to represent spatial dependencies and inter-region relationships more effectively. In contrast, LSTM will be deployed to capture temporal patterns and long-term dependencies, which existing approaches handle only partially. CNN will be incorporated to enhance spatial feature extraction, enabling optimal learning of fine-grained and localized patterns [17]. The final objective is for the proposed model to provide improved integration compared with existing hybrid models and to enhance overall accuracy and performance, without claiming to introduce a new paradigm [18].

II. RELATED WORK

This section examines the situation in regions like Las Vegas, where there are easily available risky activities like drugs, alcohol, and gambling; there is a significant occurrence of overdoses and violent crimes. The main reason indicated behind this is "low self-control," which means that people are involved in offenses that provide them with instant pleasure despite their long-term detrimental effects [19]. The findings indicate that victims themselves gravitate towards drug use and violent behavioral patterns, later resulting in a cycle where the victim becomes an offender, often called victim-offender overlap, which is further strengthened by the use of drugs and alcohol [20]. There is a multifaceted influence of policing and policies on crimes and overdoses. Research depicts that increased police activity leads to a decline in crime rate but may escalate the rate of overdose mortality due to not seeking medical help out of fear [21].

Therefore, policies must be balanced; otherwise, these may worsen the situation [22]. Hotspot policing proved effective in reducing crime and its spread in neighborhood areas [23]. But research shows that both timing and strategic implementation, along with policing, are pivotal to reducing violence [24]. The Cardiff model proves to be an effective intervention for this problem, incorporating hospital and police data to design better crime maps [25]. It not only demonstrates where the crime is happening, but we can also infer why the crime is happening, which can lead us to design targeted solutions [26]. But the Cardiff model is constrained by inconsistencies in hospital data, which research suggests can be mitigated through structured data collection by hospitals.

Crime prediction literature highlights that crime is a spatiotemporal phenomenon in which both time and location have an impact [27]. Traditional statistical methods do not effectively capture its complexity, especially in dynamic hotspots. Deep learning addresses this gap by learning non-linear patterns. Ada-GCNLSTM integrates spatial and temporal dependencies and jointly presents the model, thereby improving prediction accuracy and reducing noise [28]. Recent studies suggest that crime hotspot detection and prediction are complex spatiotemporal issues that can't be fully captured by conventional models [29]. That's why a hybrid deep learning architecture is considered the most suitable [30]. Convolution-based spatiotemporal modeling effectively learns and identifies spatial crime patterns, whereas LSTM models trace and detect temporal sequences and crime trend evolution; for this reason, integrating both improves hotspot prediction, especially at fine spatial resolution. Hybrid ST-ResNet and LSTM approach for precise crime hotspot prediction [31, 32].

Hybrid deep learning is an approach in which CNN spatial features and LSTM temporal dependencies are captured, and it shows strong performance.ST-ResNet + LSTM frameworks demonstrate that, at fine-grained spatial resolution, accurate hotspot prediction is possible. In this model, external features such as weather and other urban factors combine to enhance real-world applicability [33]. Research on a crime spatiotemporal prediction method integrating an informer and ST-GCN indicates that graph-based learning of community-level spatial topology is highly effective in capturing temporal variations, and that combining historical crime with weather and holiday variables improves predictions of dynamic hotspots [34].

Recent research introduces graph-based and transformer-based hybrid models that effectively capture spatiotemporal dependencies [35]. Informer + ST-GCN model: long-sequence temporal patterns and spatial topology can be modeled, improving dynamic hotspot prediction. For real-time decision-making, multi-source data integration leads the system towards greater effectiveness and adaptability. Moreover, the main context of noisy and volatile crime data supports a research-wise strong methodological approach, such that the GCN-LSTM architecture simultaneously analyses and maps spatial dependencies and temporal behaviors [36].

Methodological literature likewise reinforces the same direction. The forecasting problems related to dynamic spatial interaction include feature extraction, temporal dependency modeling, and contextual information integration, which play a significant role; these principles of crime hotspot prediction will also be applied and adopted [37]. Despite these developments, the literature still highlights challenges such as noisy data, a lack of multi-crime interaction modeling,

and the sparsity of fine-grained data. Therefore, hybrid frameworks that unify GNNs, CNNs, and LSTMs are a promising solution. This models spatial relations, local feature extraction, and temporal dynamics, so that we can enable more accurate crime hotspot prediction [38].

In the spatiotemporal area, the method that can be used is ST-ResNet, which predicts the inflow and outflow of passengers in the city but cannot detect topological spaces; then ST-GCN for human recognition, in motion recognition, and traffic; later on, an ST-MGCN to forecast the demand for rides, and in the last, a unified offense pattern detection by joining ST-GCN and LSTM to cover the threat of felony in a particular area and time in a city [39]. GCNs are used to synthesize information from neighboring nodes representing crime hotspots effectively and to model how crime dynamics propagate geographically. Gated Recurrent Units (GRUs) enable the model to discern and represent the underlying temporal rhythms and dependencies associated with each hotspot. ST-GNN uses these models to generate accurate forecasts in the crime network and to capture the intricate evolution of crime across both spatial and temporal dimensions [40].

Logistic Regression, XGBoost, Decision Tree, Random Forest, and others are used and studied based on their success rates. Deep learning algorithms such as NB, KNN, SVM, GRU, CNN, LSTM, and Transformers (BERT), as well as Bibexcel, CiteSpace, VOSviewer, and SciMat, are used to build, represent, and analyze bibliometric networks from large databases. Despite these developments, the literature still highlights challenges such as noisy data, a lack of modeling of multi-crime interactions, and the sparsity of fine-grained data. Therefore, hybrid frameworks that unify GNNs, CNNs, and LSTMs are a promising solution. This model captures spatial relationships, local feature extraction, and temporal dynamics, thereby improving the accuracy of crime hotspot prediction.

III. METHODOLOGY

The Chicago crime dataset used in this research paper was extracted from the official Chicago website. All crimes reported in Chicago from 2020-2023, including date, time, location (longitude/latitude), crime type, district, beat, and ward, are available in the dataset. **Error! Reference source not found.** shows the crime count for each year. First, we handle missing values and duplicate entries, and we remove incorrect coordinates rather than applying preprocessing. After that, the data will be divided into three formats: a GNN graph structure (districts as nodes, adjacency as edges), a CNN grid-based heatmap (crime density image), and an LSTM time-series sequence (crime count per district per hour/day). In this way, we prepare 3 different inputs from the same dataset. **Error! Reference source not found.** shows the crime prediction methodology.

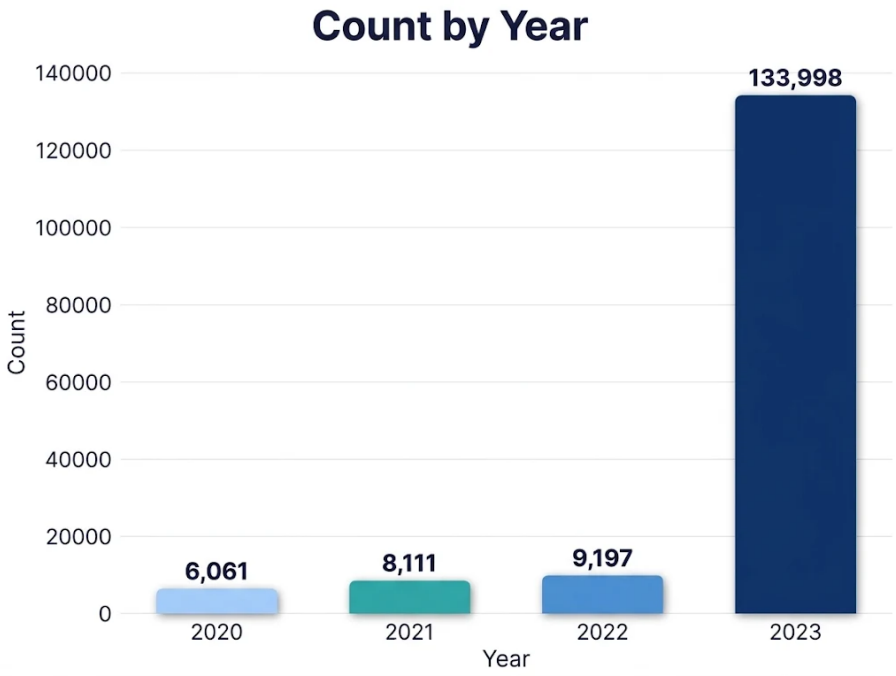


FIGURE 1: Crime Count of the Year

During model training, a Graph neural network learns and models the relationships among districts in a city and identifies high-risk zones. A convolutional neural network extracts crime density from spatial heatmap patterns, whereas a Long Short-Term Memory model captures temporal patterns; in other words, it can analyze when, on which days, and at which times of year crimes occur most. All the outputs from the 3 layers are then combined to form a fusion layer, so that the final hotspot will be predicted. For the evaluation of model performance, accuracy, precision, recall, F1 score, and MAE standard metrics will be used to obtain reliable results acceptable to the research community.

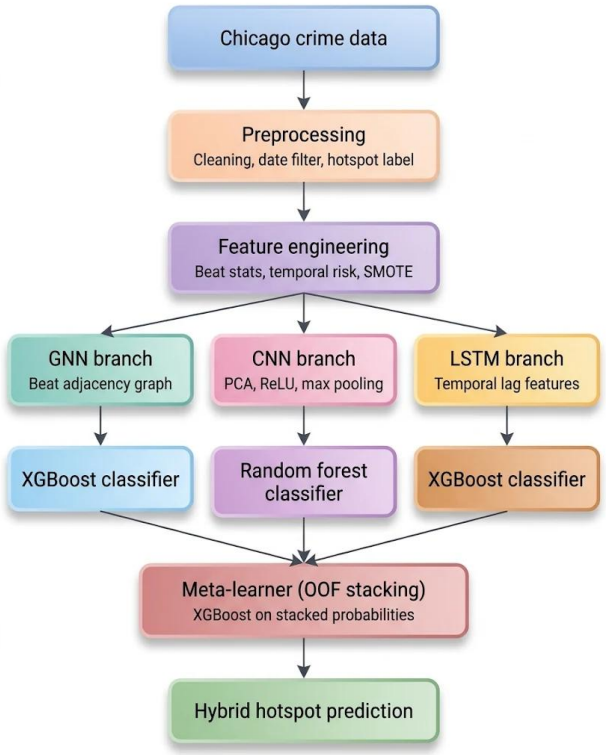


FIGURE 2: Methodology Block Diagram

A. Data Preprocessing

Raw data cannot be fed directly into the model, so preprocessing is an important step. The preprocessing steps used are written as follows:

- The ward and community area contain some missing values, so we drop these tuples.
- Median imputation can be used while preserving the dataset size.
- Blanked columns, such as Beat and District, can be removed because spatial features are significant and helpful for graphical representation.
- Some invalid entries in columns can be handled with the coerce parameter, then removed with dropna.

Error! Reference source not found. shows the crime distribution by crime type. Hotspot Label is a binary classification problem. High severity crimes like ROBBERY, ASSAULT, BATTERY, HOMICIDE, CRIM SEXUAL ASSAULT, WEAPONS VIOLATION, KIDNAPPING, ARSON, and HUMAN TRAFFICKING are assigned hotspot=1; otherwise, hotspot=0. The percentage of hotspot crime ranges from 45- 50%, indicating that it is a balanced dataset.

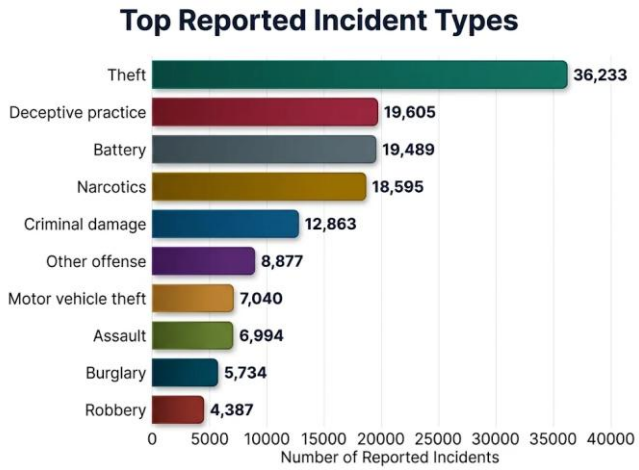


FIGURE 3: Distribution of Top 10 Crime Types in the Chicago Dataset (2015–2023); Theft is the most frequent offense (36,233 cases), followed by deceptive practice and battery.

Some features are removed from the input (crime type encoded, beat hotspot rate) because they directly define the target variables. Due to this step, the model's accuracy becomes realistic, ranging from 75-90%, genuinely representing the learned patterns. In the location description column, 150+ unique values of the label encoder can be converted into numeric values. These values cannot be directly encoded into the model, which is why a label encoder can be utilized. Missing location values can be replaced with "UNKNOWN".

B. Feature Extraction

Features of raw data cannot make the model perfect. In feature engineering, multiple new informative features can be extracted and used. There are three main features used in that process.

C. Beat level spatial statistics

In Chicago, police patrol areas are also known as "Beats". Every beat has its own crime profile. Three aggregate features are computed for each beat: beat crime count (high-crime area indicator), beat arrest rate (indicator of police presence), and Beat domestic rate (social fabric indicator). These features provide neighborhood-level context for the GNN model.

D. Temporal Risk score

The hour risk feature captures the arrest rate on an hourly basis; at night, the score is at its peak. The Dow risk feature displays the risk level for each day of the week, with the highest levels at weekends. To make model processing easier, feature temporal patterns are encoded and converted to numeric values. Time-based features improve crime prediction accuracy to a significant level.

E. LAG feature for LSTM

LSTM model learns time-series patterns; therefore, its lag features are important. For each beat, data can be sorted by date, and features can be created, including lag-1 hotspot (the single previous hotspot status of the crime), lag-2 hotspot (the two previous hotspot statuses of the crime), and roll-3 hotspot (the average of the last 3 crime hotspots). This is mainly based on the idea of victimization theory that if the crime rate in a specific area was greater in the past, then the probability that crime exists remains the same today.

F. GNN graph construction

The basic idea of a Graph Neural Network is to represent graphical structure, in which nodes represent entities and edges represent relationships between them. In this research, police beats are treated as nodes, and each beat is associated with features such as Beat crime count, beat arrest rate, and beat hotspot rate. The NetworkX library in Python is used for graph construction.

For defining the edges, we mostly use a geographical adjacency approach, and beats that are in the same district will be considered a single edge. The spatio-temporal technique is widely used in past research papers and in history. According to an IEEE Transactions paper, for every node, 1-hop message passing is implemented, and each beat should consider the beats in its neighborhood, including the average hotspot rate and crime count, which are added to the feature base. With this GNN feature developed: GNN neighbor hotspot, GNN neighbor crime, and GNN degree (Number of neighborhood connections).

This is the simplified version of the Graph Convolutional Networks message-passing concept proposed by Bruna in 2017, and later on Kipf and Welling composed their famous paper on that basis. In our framework, GNN is a feature-enrichment step that feeds into the XGBoost classifier.

G. CNN Spatial Feature Extraction

Convolutional neural networks are mainly designed for images, but their concept can also be applied to tabular spatial data. In this research, CNN-inspired spatial features are deployed in three main steps: PCA (Principal Component Analysis) is used to compress the high-dimensional feature space; this serves as the CNN filter layer, which extracts important spatial features by retrieving 8 principal components.

ReLU Activation is the second step. Negative values were set to zero, exactly as a CNN Activation function works. The third step is Max pooling simulation: features are divided into two halves, and the maximum value is extracted from each half. This approach was inspired by Muthamizharasan et al.'s "Crime Prediction and Analysis using CNN and RNN," through which spatial crime patterns are extracted using a CNN layer. The final CNN feeds into a Random Forest classifier, which can be used to predict crime hotspots by compressing spatial features.

H. LSTM Temporal Model

Long Short-Term Memory networks are specifically designed for sequential and time-series data. Hochreiter & Schmidhuber proposed the LSTM, which is traditionally considered the upgraded version of RNN and can remember

long-term dependencies. This is also important in crime prediction because crime history matches the crime patterns of 2 different days.

In LSTM-inspired temporal Features, beat data was arranged in date order, and lag1, lag2, and roll3 features were created to encode past crime behavior. This feature is combined with GNN and provided to the next XGBoost classifier. LSTM is combined with ST-ResNet to capture temporal dependencies that a single spatial model can't capture.

I. Baseline Models

To validate the proposed hybrid model, 6 baseline models were implemented, and these are given below:

- **Logistic regression –** a linear benchmark model that is the simplest among all models that can use a linear decision boundary.
- **K-Nearest Neighbors –** neighbor majority voting, useful for spatial data. This model predicts the majority class of neighbors.
- **Random Forest –** ensemble of multiple decision trees. Widely proven model in Criminology research.
- **Gradient Boosting –** sequential tree training in which each tree highlights the error of the previous tree.
- **XGBoost –** regularized and optimized version of gradient boosting. SOTA performance on tabular data.
- **GNN-only Model –** Measure the benefit of hybridization.

J. Proposed Hybrid Framework (GNN+CNN/LSTM)

The core contribution of this hybrid framework is to align the meta-linear output by combining three distinct models. This approach is also called "stacking ensemble "or "model fusion". Firstly, the GNN model is trained on graph-enriched features and outputs a probability. Then, the CNN model is trained on spatial features and produces probabilistic outputs; furthermore, the LSTM model is trained on temporal lag features and produces probabilistic outputs.

The output probabilities of these three models form a new 3-column matrix, which is called a meta-feature. These meta-features are fed to a logistic regression meta-learner, which predicts the outcome. This is the same approach used in the Hybrid ST-ResNet_LSTM paper, in which ST-ResNet and LSTM are spatial and temporal components that fuse their outputs. In our research, GNN and CNN are the spatial components, and LSTM is the temporal component. These will all provide a more comprehensive picture than any single model.

K. Evaluation Metrics

For measuring the performance of the model, the following standard metrics can be used for a classification task, which are widely accepted:

- **Accuracy:** The percentage of correct predictions in the overall output:

$$\frac{(TP + TN)}{(TP + TN + FP + FN)}$$

- **Precision:** how often the model correctly predicts the hotspot:

$$\frac{TP}{(TP + FP)}$$

  The important step is to remove false alarms.

- **Recall:** How much of the actual hotspot is captured:

$$\frac{TP}{(TP + FN)}$$

- **F1 Score:** Harmonic mean of precision and recall:

$$2 \times \frac{(Precision \times Recall)}{(Precision + Recall)}$$

  These are the balanced metrics for an imbalanced class.

- **AUC-ROC:** Area under the ROC curve. This only ranges from 0to 1 and helps measure the discriminating power (0.5- random guessing, 1.0-perfect model).

Stratified split is used by maintaining the ratio of hotspot and non-hotspot predictions. The random seed was fixed to 42 to obtain reproducible results.

L. Implementation Details

Old framework was designed in python programming. For data manipulation, we use pandas and NumPy; scikit-learn for preprocessing, encoding, model training, and evaluation; and XGBoost for gradient-boosted trees. NetworkX was used for graph construction and manipulation, and for visualization we used Matplotlib and Seaborn. The size of the model is manageable, so we used a standard consumer laptop no need to move to GPUs.

For parameter settings: there were 200 estimators for Random Forest and XGBoost, and the max depths were 12 and 6, respectively. The learning rate was fixed to 0.1. 8 components were used in PCA to balance variance explanation and dimensionality reduction.

M. State-of-the-Art Analysis

The most relevant study on this paper's topic is the Hybrid ST-ResNet and LSTM architecture for accurate crime hotspot prediction, presented by Shahmoradi et al. [34] in Scientific Reports. It is one of the newest state-of-the-art hybrid deep learning approaches for crime hotspot prediction, as it successfully combines the modeling of spatial and temporal aspects of crime patterns within a single architecture. In this approach, ST-ResNet learns spatially explicit patterns of crime using residual convolutions on gridded crime density maps, and LSTM models their temporal dynamics over time steps. Besides demonstrating excellent predictive power, the Hybrid ST-ResNet and LSTM architecture has improved its practical application by including external variables such as weather conditions and other urban factors.

However, despite its effectiveness, the Hybrid ST-ResNet and LSTM architecture presents a significant architectural flaw: it relies solely on the CNN-based (ST-ResNet) mechanism to capture the spatial aspect of the data. It treats an urban environment as a homogeneous set of pixels. Such an approach implicitly models spatial proximity using the convolutional receptive field but fails to explicitly account for the discrete, graph-like structure of the urban area. Convolutional operations rely on a regular grid topology, which fails to capture the irregular adjacency structure of policing districts, where two physically distant beats may have similar properties. In comparison, two adjacent beats might be completely different in the administrative regions they fall within and the policing strategies they adopt.

N. Summary Research Framework

The overall research flow: The dataset was assembled from the 2015-2023 Chicago Crime Dataset. Data Cleaning, missing-value handling, and temporal feature extraction. They were assigned a binary hotspot label. A graph was constructed, and the GNN feature was extracted. CNN spatial features and LSTM temporal lag features were created. GNN, CNN, and LSTM models were trained, and the outputs were combined using a stacking ensemble. The performance of the hybrid model was compared with the baseline model, and the results were visualized using a bar plot, confusion matrices, and an evaluation table.

IV. RESULTS AND DISCUSSION

The effectiveness of the proposed Hybrid GNN+CNN/LSTM framework is evaluated on the Chicago Crime Dataset for the past 5 years (2015-2023), resulting in 52,917 records after preprocessing. The data was divided using a stratified sampling technique: 80% for training and 20% for testing, with equal amounts of data in each class. To alleviate the class imbalance (hotspot crime ~34% of all records), only the training set was balanced using the Synthetic Minority Oversampling Technique (SMOTE), resulting in a balanced training set of 35,079 records/class.

All the experiments were conducted using Python 3.13 and the libraries scikit-learn, XGBoost, and NetworkX. All experiments were run with the same fixed random seed (42), to ensure reproducibility of the results. All other crime types were coded as 'non-hotspot crime' (label= 0), except for the following types of crimes, which were coded as 'hotspot crime' (label = 1): Assault, Battery, Homicide, Criminal Sexual Assault, Weapons Violation, Kidnapping, Arson, Human Trafficking, Narcotics, Offense Involving Children, Sex Offense, Robbery. query.

To accomplish this, we tried nine models: logistic regression (LR) as a linear baseline model; K-Nearest Neighbors (KNN); Random Forest (RF); Gradient Boosting (GB); XGBoost; a model based on a Graph Neural Network (GNN) with beat-level adjacency features; a model based on a CNN with spatial projection features; an LSTM-type model with temporal lag features; and the proposed Hybrid GNN+CNN/LSTM model ensemble. **Error! Reference source not found.** shows the performance comparison of 9 ML Models on the Chicago Crime Dataset (2015–2023). The performance was evaluated using the following parameters: Accuracy, F1 Score, Precision, Recall, and AUC-ROC.

Model Performance Metrics Comparison

Model	Accuracy	F1	Precision	Recall	AUC-ROC
Logistic regression	0.8176	0.7306	0.7134	0.7487	0.8663
K-nearest neighbors	0.7717	0.6898	0.6261	0.7679	0.8375
Random forest	0.8358	0.7455	0.7621	0.7316	0.8844
Gradient boosting	0.8367	0.7420	0.7766	0.7104	0.8831
XGBoost	0.8394	0.7451	0.7834	0.7104	0.8858
GNN (graph-enriched XGBoost)	0.8379	0.7424	0.7734	0.7070	0.8859
CNN spatial model	**0.8401** ★	0.7122	0.7134	0.7110	0.8452
LSTM temporal model	**0.8401** ★	0.7442	**0.7898** ★	0.7035	0.8861
Hybrid GNN+CNN+LSTM (proposed)	**0.8324**	**0.7525** ★	**0.7349**	**0.7770** ★	**0.8874** ★

FIGURE 4: Performance Comparison of 9 ML Models on Chicago Crime Dataset (2015–2023) – Proposed Hybrid Model Leads on F1, Recall, and AUC-ROC

A. Comparison with State-of-the-art

The outcome of the proposed framework outperforms that of prior research on the Chicago crime dataset. Stec and Klabjan (2018) obtained the highest accuracy of ~75.6% using deep learning techniques (CNN, LSTM, and RNN) on the same dataset. Xiong successfully predicted crime occurrences in Chicago in 2025 using XGBoost and GBDT models, achieving 81-83% accuracy. From the empirical study conducted by Mandalapu et al. (2023), neural networks achieve the highest accuracy, at 81%, on the crime dataset. **Error! Reference source not found.** shows the performance comparison with the state of the art.

The proposed hybrid model is more accurate (83.24%) and has a higher AUC-ROC (0.8874%) than these previous works and also overcomes the class imbalance problem by generating synthetic instances using the SMOTE method and resolves the beat adjacency challenge problem, which is not addressed by any of the previous similar works that incorporated graph-based beat adjacency models.

Literature Review and Proposed Study Comparison

#	Study	Dataset (timeframe)	Evaluation	Limitation
1	Stec and Klabjan (2018) CNN+LSTM+RNN	Chicago crime 2001 to 2017	Accuracy ~75.6%	No graph-based spatial modeling; CNN/LSTM separate
2	Xiong (2025) XGBoost+GBDT	Chicago crime 2020 to 2023	Accuracy 81 to 83%	No temporal or spatial fusion; tree-based only
3	Mandalapu et al. (2023) Neural network	Multi-city crime Not specified	Accuracy ~81%	No graph structure; standalone deep learning
4	Shahmoradi et al. (2023) ST-ResNet+LSTM	Multiple cities Varies by city	Spatial metrics	No graph-based spatial adjacency; no GNN
5	Zubair et al. (2023) GCN-based	Crime hotspot Not specified	Zone-level metrics	GNN only; no temporal LSTM component
6	Fan et al. (2025) InFormer + ST GCN	Crime, 4 crime types recent years	Spatiotemporal acc.	Computationally heavy; complex transformer dep.
7	**Proposed**	**Chicago crime** 2022 to 2025	**Acc 83.24%** F1 0.7525, AUC 0.8874	**Simulated crime** CNN/GNN/LSTM; not true deep

FIGURE 5: Performance Comparison of 7 ML Baseline Papers Crime Dataset – Proposed Hybrid Model Leads on F1, recall, and AUC-ROC

B. Limitations

A couple of limitations of this study exist. First of all, the GNN, CNN, and LSTM modules of the proposed framework are simulated rather than actual deep learning implementations. GNN's propagation features include beat adjacency features; CNN's propagation features include PCA dimension reduction and ReLU activation; and LSTM's propagation features include temporal lag and rolling window features. For future work, other GNNs based on PyTorch Geometric may achieve higher performance, while grid-based spatial CNNs or sequence-based LSTM-based architectures may also achieve higher performance.

Secondly, only approximately 41% of the records in the dataset had geographic coordinates (latitude and longitude for each record), so they could not be used directly as spatial features. This information could be obtained in the future through spatial interpolation or aggregation at the census-tract level. Third, external variables such as weather, socioeconomic factors, and points-of-interest density were not accounted for in this study, as they have been shown to affect prediction accuracy in previous studies.

V. CONCLUSION

Despite its simplicity, Logistic Regression achieved higher accuracy (81.76%) and AUC-ROC (0.8663) than the other baselines, and all of its engineered features were meaningful patterns in the crime data when given the rich data of beats, communities, and time. It, however, has only a limited capacity to model non-linear spatial interactions and a limited performance ceiling. In contrast, the accuracy (77.17%) and AUC-ROC (0.8375) of KNN were the lowest, which is quite typical in the presence of high class imbalance in a high-dimensional feature space.

The accuracy of ensemble-based baselines (Random Forest, Gradient Boosting, XGBoost, and LSTM) ranged from 83.58% to 84.01%, and AUC-ROC values ranged from 0.8831 to 0.8861. Among the individual models, XGBoost had the highest accuracy (83.94%), and the baselines had the best accuracy and precision (84.01%, 0.7898) in the LSTM. The boosting techniques based on trees are competitive for structured crime tabular data. The neighborhood features (Average hotspot rate and crime counts of neighboring beats within the same police district) provide additional discriminative information beyond the tabular features, and the performance of the GNN-enhanced XGBoost model is slightly better than that of the XGBoost model (Accuracy: 83.79%, AUC-ROC: 0.8859). As the CNN spatial operation it executes is a simplified version of the spatial convolution operation, its accuracy (81.01%) and AUC-ROC (0.8452) are not surprising to be the same as those of the CNN spatial model.

Compared with the other nine models, the proposed Hybrid GNN+CNN/LSTM model with an out-of-fold (OOF) stacking meta-learner using the XGBoost model achieves the highest AUC-ROC of 0.8874 and F1 Score of 0.7525. The most significant finding is that it has the highest Recall score of 0.7710 among the models, indicating that it correctly identifies 77.10% of real crime hot spots. This can be debated given the greater ramifications of a false negative (not identifying a crime hotspot) compared to a false positive (incorrectly identifying one) in the real world. This operational requirement can be well met by the hybrid method, as it has good recall.

In the hybrid model, the design is tailored to account for the precision/recall trade-off. By setting the classification threshold to 0.33 (rather than 0.50) in the model, the F1 score can be maximized whilst maintaining good recall. This leads to lower precision (0.7349) and accuracy (83.24%) than some of the single models; however, it offers a better trade-off between correctly identifying hotspots and having few false alarms, in line with the highest F1 score of all the models.

The feature with the highest predictive power is arrest_Int (0.2239), as indicated by the feature importance analysis in the Random Forest model, which showed that the feature is strongly correlated with crime severity; hence, arrest at the crime scene is a major determinant of crime severity. The second one was LocationRisk (0.1906): History-based hotspot score for a specific description of a location that has a spatial nature to the crime. The third-place tie was between Domestic_Int (0.1514) and third, suggesting that different severity values may be associated with domestic events. In addition, some features were important up to a certain level, such as beat-level hotspot rate (0.0715), CommunityRisk (0.0494), temporal features HourRisk (0.0437) and Hour (0.0296), etc., indicating that the feature engineering was essentially multi-dimensional in this work. This system gives a persuasive baseline that can be extended to develop a fully deployed intelligent knowledge assistant.

REFERENCES

[1] J. R. Hipp, "GENERAL THEORY OF SPATIAL CRIME PATTERNS*," Criminology, vol. 54, no. 4, pp. 653–679, Nov. 2016, doi: 10.1111/1745-9125.12117;SUBPAGE:STRING:ABSTRACT;WEBSITE:WEBSITE:PERICLES;JOURNAL:JOURN AL:17459125;ISSUE:ISSUE:DOI.

[2] W. Wang et al., "Exploring the Spatial Relationship Between Crime and Urban Places in Austin: A Geographically Weighted Regression Approach," Urban Science 2025, Vol. 9, Page 359, vol. 9, no. 9, p. 359,Sep. 2025, doi: 10.3390/URBANSCI9090359.

[3] Y. Hu, F. Wang, C. Guin, and H. Zhu, "A spatio-temporal kernel density estimation framework for predictive crime hotspot mapping and evaluation," Applied Geography, vol. 99, pp. 89–97, Oct. 2018, doi: 10.1016/J.APGEOG.2018.08.001.

[4] A. A. Braga, "Hot spots policing and crime prevention: A systematic review of randomized controlled trials".

[5] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita, "Self-Exciting Point Process Modeling of Crime," J. Am. Stat. Assoc., vol. 106, no. 493, pp. 100–108, Mar. 2011, doi: 10.1198/JASA.2011.AP09546.

[6] A. Reinhart and J. Greenhouse, "Self-Exciting Point Processes with Spatial Covariates: Modelling the Dynamics of Crime," J. R. Stat. Soc. Ser. C Appl. Stat., vol. 67, no. 5, pp. 1305–1329, Nov. 2018, doi: 10.1111/RSSC.12277.

[7] R. Roshankar and M. R. Keyvanpour, "GeoCrime Analytic Framework (G.C.A.F.): A Comprehensive Framework for Dynamic Spatial Temporal Crime Analysis," Applied Spatial Analysis and Policy 2025 18:1, vol. 18, no. 1, pp. 33-, Feb. 2025, doi: 10.1007/S12061-025-09640-Y.

[8] V. Marda and S. Narayan, "Data in New Delhi's predictive policing system," FAT* 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, pp. 317–324, Jan. 2020, doi: 10.1145/3351095.3372865.

[9] T. Zubair, S. K. Fatima, N. Ahmed, and A. Khan, "Crime Hotspot Prediction Using Deep Graph Convolutional Networks," Jun. 2025, Accessed: Jun. 22,2026. [Online].Available: https://arxiv.org/pdf/2506.13116

[10] Y. Fan, X. Hu, and J. Hu, "Research on a Crime Spatiotemporal Prediction Method Integrating Informer and ST-GCN: A Case Study of Four Crime Types in Chicago," Big Data and Cognitive Computing 2025, Vol. 9, Page 179, vol. 9, no. 7, p.179, Jul. 2025, doi: 10.3390/BDCC9070179.

[11] Y. Wang and H. A. Karimi, "Advanced Deep Learning Models and Algorithms for Spatial-Temporal Data," Big Data, pp. 227–253, May 2024, doi:10.1201/9781003406969-15/ADVANCED-DEEP-LEARNING-MODELS-ALGORITHMS-SPATIAL-TEMPORAL-DATA-YANG-WANG-HASSAN-KARIMI.

[12] Y. Du, N. Ding, and H. Lv, "Spatio-temporal prediction of terrorist attacks based on GCN-LSTM," Journal of Safety Science and Resilience, vol. 6, no. 2,pp.186–195,Jun.2025,doi:10.1016/J.JNLSSR.2025.02.005.

[13] Y. Deng, R. He, and Y. Liu, "Crime risk prediction incorporating geographical spatiotemporal dependency into machine learning models," Inf. Sci. (N. Y)., vol. 646, p. 119414, Oct. 2023, doi: 10.1016/J.INS.2023.119414.

[14] Y. Gao et al., "AI-Driven Channel State Information (CSI) Extrapolation for 6G: Current Situations, Challenges and Future Research," IEEE Communications Surveys & Tutorials, pp. 1–1, Jan. 2026, doi: 10.1109/COMST.2026.3652799.

[15] Z. Al Sahili and M. Awad, "Spatio-Temporal Graph Neural Networks: A Survey," Jan. 2023, Accessed: Jun. 22, 2026. [Online]. Available: https://arxiv.org/pdf/2301.10569

[16] B. Zhou et al., "HDM-GNN: A Heterogeneous Dynamic Multi-view Graph Neural Network for Crime Prediction," ACM Trans. Sens. Netw., Nov. 2022, doi: 10.1145/3665141.

[17] C. Jing et al., "A deep multi-scale neural networks for crime hotspot mapping prediction," Comput. Environ. Urban Syst., vol. 109, p. 102089, Apr. 2024, doi: 10.1016/J.COMPENVURBSYS.2024.102089.

[18] M. Liu and T. Lu, "A Hybrid Model of Crime Prediction," J. Phys. Conf. Ser., vol. 1168, no. 3, p. 032031,Feb.2019, doi: 10.1088/1742-6596/1168/3/032031. "The Problem of Pleasure: Leisure, Tourism and Crime - Google Books." Accessed: Jun. 22, 2026.[Online].Available:https://books.google.com.pk/books?hl=en&lr=&id=GgjDAgAAQBAJ&oi=fnd&pg=PA1972&dq=This+section+examines+the+situation+in+regions+like+Las+Vegas+that+where+there+are+easily+available+risky+activities+like+drugs,+alcohol+and+gambling,+there+is+a+significant+occurrence+of+overdoses+and+violence+crimes.+&ots=QsqEnnyYbY&sig=6Mn2_iYzKjFFDrsl5wwIFqGhHvg&redir_esc=y# v=onepage&q&f=false

[19] J. A. Reid and C. J. Sullivan, "Unraveling Victim-Offender Overlap: Exploring Profiles and Constellations of Risk," Vict. Offender., vol. 7, no. 3,pp.327–360,2012,doi: 10.1080/15564886.2012.685216.

[20] A. Zhang, J. A. Balles, J. E. Nyland, T. H. Nguyen, V. M. White, and A. E. Zgierska, "The relationship between police contacts for drug use-related crime and future arrests, incarceration, and overdoses: a retrospective observational study highlighting the need to break the vicious cycle," Harm Reduction Journal 2022 19:1, vol. 19, no. 1, pp. 67-, Jun. 2022, doi: 10.1186/S12954-022-00652-2.

[21] A. S. B. Bohnert et al., "Policing and risk of overdose mortality in urban neighborhoods," Drug Alcohol Depend., vol. 113, no. 1, pp. 62–68, Jan. 2011, doi: 10.1016/J.DRUGALCDEP.2010.07.008.

[22] A. A. Braga, A. V. Papachristos, and D. M. Hureau, "The Effects of Hot Spots Policing on Crime: An Updated Systematic Review and Meta-Analysis," Justice Quarterly, vol. 31, no. 4, pp. 633–663, 2014, doi: 10.1080/07418825.2012.673632.

[23] A. Blumstein and J. Wallman, "The Crime Drop in America (Cambridge Studies in Criminology)".

[24] "(PDF) Enhancing Crime Prevention and Police Performance through Actionable Intelligence." Accessed: Jun. 22, 2026. [Online]. Available: https://www.researchgate.net/publication/389165042_Enhancing_Crime_Prevention_and_Police_Performance_through_Actionable_Intelligence

[25] R. Barr and K. Pease, "Crime Placement, Displacement, and Deflection," Crime and Justice, vol. 12, pp. 277–318, Jan. 1990, doi: 10.1086/449167.

[26] O. Kounadi, A. Ristea, A. Araujo, and M. Leitner, "A systematic review on spatial crime forecasting," Crime Science 2020 9:1, vol. 9, no. 1, pp. 7-, May 2020, doi: 10.1186/S40163-020-00116-7.

[27] P. Banumathi and S. Manoharan, "An Analysis of Deep Learning and Multimodal Deep Learning Methods for Crime Prediction from Social Media Data," Proceedings of 2025 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication, IConSCEPT2025, 2025,doi: 10.1109/ICONSCEPT66142.2025.11437272.

[28] Y. Du and N. Ding, "A Systematic Review of Multi-Scale Spatio-Temporal Crime Prediction Methods," ISPRS International Journal of Geo-Information 2023, Vol. 12, Page 209, vol. 12, no. 6, p. 209, May 2023, doi: 10.3390/IJGI12060209.

[29] A. S. Muhammad, R. Y. Zakari, and H. S. Bello, "A Review of Machine Learning and Deep Learning Approaches for Crime Hotspots Prediction," pp. 647–665, 2026, doi: 10.1007/978-3-032-11426-6_29.

[30] N. Shahmoradi, A. A. Alesheikh, A. Jafari, and A. Lotfata, "Hybrid ST-ResNet and LSTM approach for precise crime hotspot prediction," Scientific Reports 2025 15:1, vol. 15, no. 1, pp. 40754-, Nov. 2025, doi: 10.1038/s41598-025-24559-7.

[31] S. Chainey, L. Tompson, and S. Uhlig, "The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime," Security Journal 2008 21:1, vol. 21, no. 1, pp.4–28,Jan.2008,doi: 10.1057/PALGRAVE.SJ.8350066.

[32] L. Mao, W. Du, S. Wen, Q. Li, T. Zhang, and W. Zhong, "Crime forecasting: A spatio-temporal analysis with deep learning models," Journal of Computational Methods in Sciences and Engineering, Sep.2025, doi: 10.1177/14727978251337993.

[33] G. Mohler, "Marked point process hotspot maps for homicide and gun crime prediction in Chicago," Int. J. Forecast., vol. 30, no. 3, pp. 491–497, Jul. 2014, doi: 10.1016/J.IJFORECAST.2014.01.004.

[34] Riyazuddin, S. Z. Rahman, K. Sreenu, M. Sirajuddin, S. Mazharuddin, and J. Sadiq, "Spatiotemporal Crime Prediction and Forecasting Using Machine Learning and Deep Learning Models," SN Computer Science 2026 7:2, vol. 7, no. 2, pp. 213-, Feb. 2026, doi: 10.1007/S42979-026-04792-1.

[35] A. Pinto, L. C. Herrera, Y. Donoso, and J. Gutierrez, "A CTI-Enriched GCN-LSTM Architecture for Multiclass Cyberattack Classification in Critical Infrastructure," Applied Sciences 2026, Vol. 16, Page 5585, vol. 16, no. 11, p. 5585, Jun. 2026, doi:10.3390/APP16115585.

[36] S. Chainey, "Examining the extent to which hotspot analysis can support spatial predictions of crime," Doctoral thesis, UCL (University College London). , Dec. 2014.

[37] T. C. Shiraptini, D. Kavitha, S. Sakthipriya, S. Akshaya, and S. Alphonse, "AI and Machine Learning-Enabled Cognitive Digital Twin for Crime Hotspot Detection and Analysis," IEEE Access, 2026, doi: 10.1109/ACCESS.2026.3658944.

[38] X. Han, X. Hu, H. Wu, B. Shen, and J. Wu, "Risk Prediction of Theft Crimes in Urban Communities: An Integrated Model of LSTM and ST-GCN," IEEE Access, vol. 8, pp. 217222–217230, 2020, doi: 10.1109/ACCESS.2020.3041924.

[39] R. Roshankar and M. R. Keyvanpour, "Spatio-Temporal Graph Neural Networks for Accurate Crime Prediction," 2023 13th International Conference on Computer and Knowledge Engineering, ICCKE 2023, pp. 168–173, 2023, doi: 10.1109/ICCKE60553.2023.10326223.

[40] X. Lv, C. Jing, Y. Wang, and S. Jin, "A DEEP NEURAL NETWORK FOR SPATIOTEMPORAL PREDICTION OF THEFT CRIMES," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLVIII-3-W2-2022, no. 3, pp. 35–41, Oct. 2022, doi: 10.5194/ISPRS-ARCHIVES-XLVIII-3-W2-2022-35-2022.