



# Predicting Crime and Other Uses of Neural Networks in Police Decision Making

Steven Walczak\*

School of Information and Florida Center for Cybersecurity, University of South Florida, Tampa, FL, United States

Neural networks are a machine learning method that excel in solving classification and forecasting problems. They have also been shown to be a useful tool for working with big data oriented environments such as law enforcement. This article reviews and examines existing research on the utilization of neural networks for forecasting crime and other police decision making problem solving. Neural network models to predict specific types of crime using location and time information and to predict a crime's location when given the crime and time of day are developed to demonstrate the application of neural networks to police decision making. The neural network crime prediction models utilize geo-spatiality to provide immediate information on crimes to enhance law enforcement decision making. The neural network models are able to predict the type of crime being committed 16.4% of the time for 27 different types of crime or 27.1% of the time when similar crimes are grouped into seven categories of crime. The location prediction neural networks are able to predict the zip code location or adjacent location 31.2% of the time.

**Keywords:** crime, literature review, location, neural network, police, temporal reasoning, geo-spatial reasoning

## INTRODUCTION

Crime is a global concern that impacts individuals and society on a daily basis and negatively affects society (Costa, 2010). The ever increasing population along with the rise in urbanization has led to dramatic increases in criminal activities (Fajnzylber et al., 2002; Zhang, 2016), particularly in urban settings (Tumulak and Espinosa, 2017; Stebbins, 2019; Zhu et al., 2019). While increased police presence has been documented to reduce increases in crime (Chalfin and McCrary, 2018), police need to be able to plan for and respond effectively to criminal events, especially those that affect personal or public safety. Police and other forces of civil order would benefit directly from intelligence to help combat all forms of crime and intelligence is a requirement for managing and fighting “intelligent” forms of crime (Tastle, 2013).

Intelligence requires information or data and a methodology for analyzing the data. Law enforcement agencies are an excellent source of large amounts of data. As an example, the number of arrest records available from 2017 to 2019 for the city of Detroit are 81,440, 82,197, and 83,893, respectively, (City of Detroit, 2020) or almost a quarter million records for this 3 year period.

Artificial intelligence and more precisely machine learning, provide mechanisms for improving police knowledge of current and potential crimes and facilitating complex anti-criminal decision making. Machine learning techniques scale well to very large amounts of data (Mena, 2011).

## OPEN ACCESS

### Edited by:

Simon Garnier,  
New Jersey Institute of Technology,  
United States

### Reviewed by:

Colleen M. Berryessa,  
Rutgers University,  
Newark, United States  
Neil David Shortland,  
University of Massachusetts Lowell,  
United States

### \*Correspondence:

Steven Walczak  
swalczak@usf.edu

### Specialty section:

This article was submitted to  
Forensic and Legal Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 27 July 2020

**Accepted:** 24 August 2021

**Published:** 07 October 2021

### Citation:

Walczak S (2021) Predicting Crime  
and Other Uses of Neural Networks  
in Police Decision Making.  
Front. Psychol. 12:587943.  
doi: 10.3389/fpsyg.2021.587943

Neural networks (NNs) are a form of machine learning based on the functionality of the human brain and its neuro-cognitive processing. NNs are the most powerful and accurate clustering technology available (Mena, 2011), enabling high quality information to support police decision making.

This article serves two purposes. First is to assist law enforcement agencies understanding of the capabilities of NNs and the extent of past and current research. A literature review is used to facilitate this knowledge sharing for law enforcement automated decision making and planning researchers and developers. Second is to demonstrate through original research the potential for NNs to be used for analysis of crimes that may facilitate crime preparedness and response by police officers. Existing research as shown through the accompanying literature review focuses on prediction of large areas of generalized increases in criminal activity or the increase in activity for a specific crime. The research presented will be able to predict a specific type of crime as soon as the location and time are known, improving on the larger time horizons of existing research. The other NN model will show that when a type of crime and the time of day is known, a specific area of the city may be immediately identified where that crime is most likely to have occurred, which improves on existing research that uses much larger time horizons. While the research demonstrates that NNs are capable of performing a wide range of crime analyses, they should be viewed as advisory tools or for secondary confirmatory analysis (Tastle, 2013).

## BACKGROUND ON NEURAL NETWORKS

NNs are a machine learning methodology that excels at classification and prediction problems. Because NNs learn, they are capable of solving arbitrarily complex problems (Hornik, 1991). NNs have been widely applied to solve problems in various domains including business (Vellido et al., 1999), engineering (Bose, 1994), medicine (Baxt, 1991), and the sciences (Sundararaj et al., 1999), and more recently to crime sciences (Oatley et al., 2006), which will be reviewed in the next section.

NNs learn through one of two mechanisms: supervised learning which requires historic data with known results, and unsupervised learning which learns patterns directly from the data. Both techniques require that the NN be trained in order to learn and various training algorithms exist for use in each type of NN learning, with backpropagation being the most popular for supervised learning (Jayalakshmi and Santhakumaran, 2011) and self-organizing maps (SOM being the most prevalent for unsupervised learning; Oatley et al., 2006). Hybrid methods that combine supervised and unsupervised learning also exist, such as the popular convolutional NN (CNN).

All NNs have an input layer that defines the variables provided to the NN for learning and an output layer that defines the desired classification or prediction solution. Supervised learning and hybrid models also contain one or more hidden layers with each layer being fully connected to the next layer with a weighted connection. Learning occurs in this type of NN by determining the error of a training prediction from

the actual value and propagating this error backwards through the network to modify the weights of the connections to better align the prediction with the correct output value. A sample supervised learning NN architecture is shown in **Figure 1**.

Validation or testing of NN models is done with data that has never been used in the training of the model being evaluated. This is required to eliminate implicit model bias and simulates prospective real-world use of the NN. Typically, available data with known outcomes is divided into a training set and a validation or test set, each with unique examples. Training data should also be adjusted to reflect balanced proportions of the classes to be predicted in order to assist the NN with learning all classes evenly and not overlearning more prevalent classes (Wei and Dunbrack Jr, 2013).

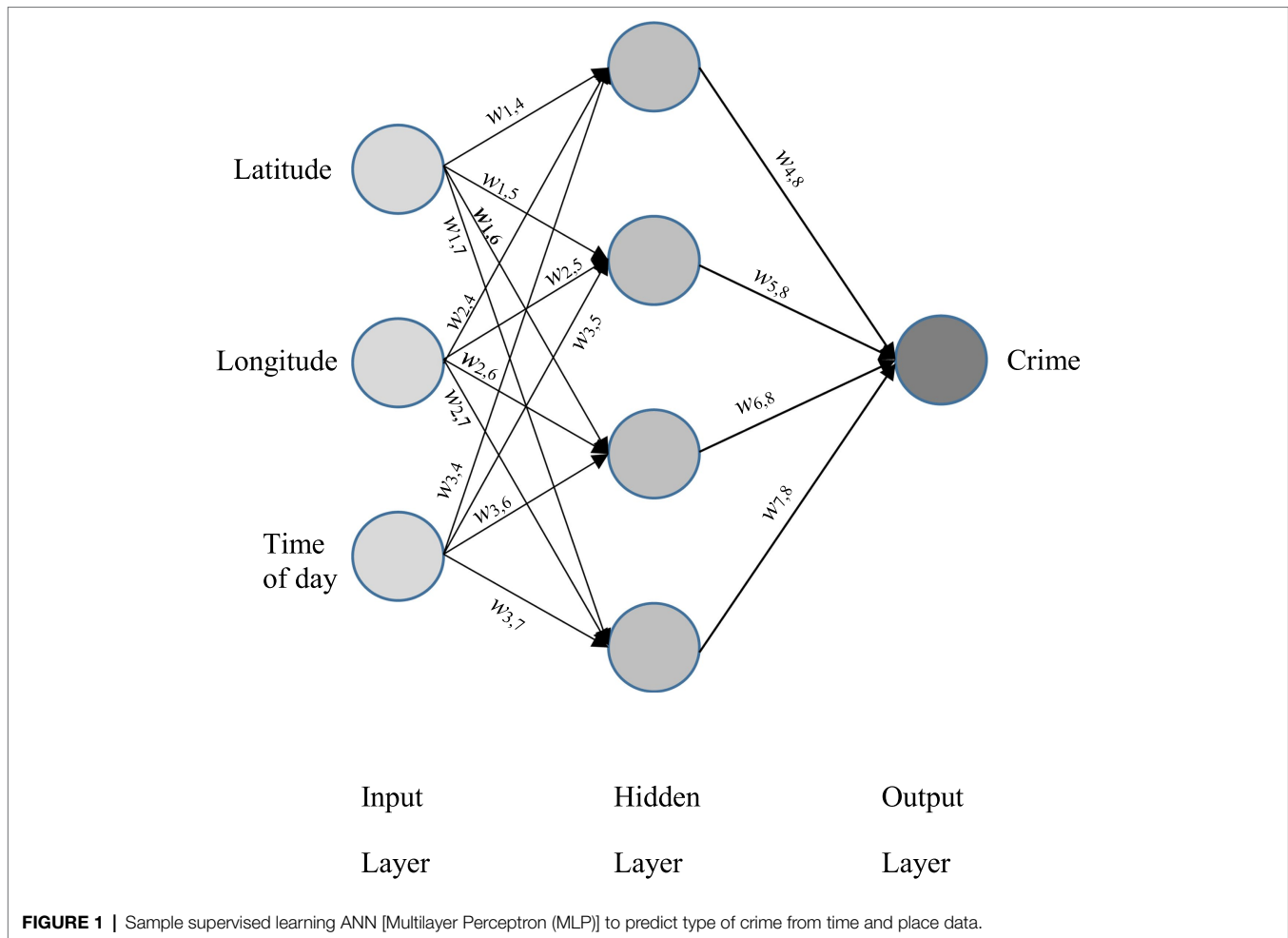
Numerous tutorials exist on how to build and utilize NNs across various domains. The reader is directed to the references for further clarification: Gallo (2015), Jain et al. (1996), Walczak and Cerpa (1999), and Wythoff (1993).

## NNS USED IN CRIME: LITERATURE REVIEW

An examination of how NNs are being used in Crime Sciences is done by performing a literature review. The purposes of literature reviews may be manifold and serve to provide an archive and analyses of commonalities across various studies. The current literature review uses the search terms: “neural network” and “crime.” Searches are performed using Google Scholar™ and other library search engines when the documents were not available through Google Scholar™. The review initially considers the top 50 articles returned, excluding patents and citations, without using any date qualifier were examined and the search was re-performed for articles from 2017 to 2020 and the top 25 were examined. Some overlap occurred between the two literature searches. The use of Google Scholar relies on Google Scholar’s relevance ranking for classifying research articles and may be a possible shortcoming of this literature review strategy.

The criteria for the literature review search is to identify NN research be applied to either crime solving or police decision making. The initial search returned 34,800 articles and the recent date search returned 14,000 articles. Four articles meeting the literature review purpose were all from the same book describing different aspects of a single project and so the quantity of articles considered in the initial literature review was extended to the top 55 articles to overcome any bias introduced by this cluster of articles. A total of 68 articles out of the possible 80 that fit the criteria for the current research.

Two categories of NN research were not included, though there exists a plethora of research, because they did not focus on police problem solving and practice. The first category of articles removed is for NNs focused on jurisprudence since these articles are about the application and discovery of relevant laws for legal prosecution, with examples including Bochereau et al. (1991), Fernandes et al. (2019), and Stranieri and Zeleznikow (2012). The other category removed from consideration is NNs



applied to cybersecurity, since cybersecurity other than cyberbullying, is not typically bounded geographically. A few examples of NN research in cybersecurity include Li and He (2020) and Siddiqui et al. (2018).

Kounadi et al. (2020) have recently performed a review and their findings indicate that there is significant growth in spatial crime forecasting research and furthermore claim that the most predominant type of crime forecasting research is the prediction of crime hotspots. Other research has confirmed the emphasis on predicting crime occurrence in specific locations (Loeffler and Flaxman, 2018). The size of the area for which predictions are made varies from study to study. The current literature review presented in this article confirms this with 36 of the 69 (52%) reviewed articles using NNs to predict crime hotspots and shown in **Table 1**. Neural network predictions from the literature may be for a specific type of crime or the frequency of occurrence of a specific type of crime, or the general location for a specific type of crime. **Table 1** defines the type of crime involved in each prediction, the location and size of the area specified for the location, and the type of neural network used to make the prediction. The types of NNs listed in **Table 1** are distinguished by the specific type of learning employed in each NN and are defined as:

- Multi-layer perceptron (MLP), a supervised learning feedforward network, most commonly trained using backpropagation or variant thereof;
- Recurrent neural network (RNN), a MLP-like architecture with backward links to previous layers enabling better time series modeling;
- SOM, an unsupervised learning method; and
- Convolutional neural network (CNN), a hybrid unsupervised and supervised learning method that typically contains many hidden layers, frequently used for image classification. This type of NN is frequently referred to as a deep learning neural network.

The length of time in the future that crime forecasts use in making predictions affects how those predictions can be used. Advanced hot spot forecasting times by NNs vary widely from hourly (Wang et al., 2019), to 1 day (Corcoran et al., 2003; Duan et al., 2017), to 2 weeks (Nair and Gopi, 2020), to 1 month (Yu et al., 2011; Cortez et al., 2018; Krishnan et al., 2018; Lin et al., 2018), and even up to 1 year (Feng et al., 2019; Yi et al., 2019). Annual predictions are primarily useful for determining adequate budget allocations and performing force recruitment, while hourly and daily future forecasts may help

**TABLE 1** | NN research predicting crime hot spots.

Type of crime	Type of NN	Location	Area of prediction	References
All	MLP	Various	N/A	Altameem and Amoon, 2019
All	CNN	Rio de Janeiro	Street	Andersson et al., 2017
All	MLP	United States	N/A	Anuar et al., 2015
Motor vehicle accident	MLP	Taiwan	0.1–4.2 km	Chang, 2005
All	Supervised learning	City in NE United States	Partial city region	Corcoran et al., 2003
Violent crimes	RNN	Guatemala City	Partial city region	Cortez et al., 2018
Felonies	CNN	New York city	0.18 km <sup>2</sup>	Duan et al., 2017
All	CNN	Chicago	Partial city region	Esquivel et al., 2019
All	RNN	San Francisco	City	Feng et al., 2019
Burglary, robbery, assault, and larceny	RNN	New York	Partial city region	Huang et al., 2018
Drug	MLP	London	Neighborhood	Intraliqi and Buscema, 2013
12 Crimes	MLP	Poland	N/A	Jankowski et al., 2018
All (aggregate)	Supervised learning	India	Country	Jha et al., 2020
Larceny/theft, assault, other	CNN	San Francisco	Partial city region	Jin et al., 2020
All	CNN	Chicago	Partial city region	Kang and Kang, 2017
Crime around transportation hubs	MLP	Chicago	Transportation hubs	Kouziokas, 2017
All	RNN	Various	N/A	Krishnan et al., 2018
Car theft	CNN	Taoyuan City, Taiwan	PCR	Lin et al., 2018
All	MLP	N/A	N/A	Mao and Du, 2019
All	CNN	Los Angeles	N/A (25 pixels)	Nair and Gopi, 2020
Drugs (crack)	Supervised learning	Pittsburgh	2,150 feet <sup>2</sup>	Olligschlaeger, 1997
Assault, robbery, and murder	Supervised learning	Boston	Neighborhood	Patil et al., 2020
Burglary, robbery, and battery		Belgium	200 meter <sup>2</sup>	Rummens et al., 2017
Gang related	Supervised learning	Los Angeles	Partial city region	Seo et al., 2018
All	N/A	Cebu City, Philippines	750–1,000 meters <sup>2</sup>	Tumulak and Espinosa, 2017
All	CNN	Los Angeles	17.8 km <sup>2</sup>	Wang et al., 2019
All	RNN,CNN	Los Angeles	Partial city region	Wang et al., 2017a
All	RNN	Chicago	Partial city region	Wang et al., 2017b
All	CNN	San Francisco	Partial city region	Wang et al., 2020
14 Crimes	RNN, CNN	N/A	City	Wawrzyniak et al., 2018
All	CNN	County in China	County	Wu et al., 2018
All	RNN	Chicago and New York	City	Yi et al., 2019
Property crimes	Supervised learning	Boston	Partial city region	Yu et al., 2011
All	MLP	N/A	N/A	Yu et al., 2012
All	RNN	Chicago	N/A	Zhu et al., 2019
All	RNN	Portland	N/A	Zhuang et al., 2017

with planning proper responses and re-routing current police resources to hot spot areas.

The second most popular research trend is to utilize NNs for datamining. As stated previously, police generate a large amount of arrest data. Datamining may be used to pull relevant data from natural language police reports, match previous crime data with current crime data to assist investigators, or analyze demographic or spatiotemporal patterns. The datamining analysis of patterns usage of NNs often complements or overlaps with hot spot forecasting, hence there exists some overlap in the reported NN research. Datamining NNs accounted for 23 out of 69 (33%) reviewed crime science NN studies and are displayed in **Table 2**. The most common NN datamining goal in the previous studies displayed in **Table 2** is to cluster similar crimes together from large databases, which may then be used to associate previous crimes with a current or unsolved crime. A secondary use of NN-based datamining is to identify trends in data, which may potentially be used for crime forecasting and preparation.

Prediction of criminal recidivism, or the likelihood that a criminal will either repeat an offense or be arrested for a different more serious offense, is another popular area of NN research. NNs have been identified as one of the most popular datamining

techniques applied to criminal recidivism over the last two decades (Hashim and Nohuddin, 2018). A total of six articles out of 69 (9%) articles reported NN applications to predict recidivism for various crimes (Babcock and Cooper, 2019) or races (Jain et al., 2019) or general prediction of recidivism (Caulkins et al., 1996; Palocsay et al., 2000; Liu et al., 2011; Chun et al., 2019). Prediction of likely recidivists is normally used to improve rehabilitation, but may also be used to identify possible crime perpetrators.

The final area of NN research that will be discussed is NN research meant to facilitate police investigative work and police response. NNs are a recommended tool for making criminal investigation recommendations and providing confirmatory analysis (Tastle, 2013). A datamining example that fits into this classification of NN applications in crime science and police decision making is the work by Kou and Tung (1994) who tried to identify a specific gun, as opposed to the type of gun, based on the markings on a bullet. This study is originally listed with the datamining NNs due to its need to analyze data from numerous databases containing specific gun-bullet markings. Other investigation decision support examples are: a NN that identifies what the type of glass is, out of a possible seven types, from a small piece of glass shard or fragment found at a crime scene (El-Khatib et al., 2019),

**TABLE 2** | NN research performing crime datamining.

Type of crime	Datamining purpose	References
Sex crimes	Determine behaviors and serial behaviors of sex offenders	Adderley and Musgrove, 2001
All	Data mines four databases to determine crime count patterns	Altameem and Amoon, 2019
All	Analyze data to classify crimes into low, medium, or high severity crimes	Anuar et al., 2015
Drugs	Examine police reports to identify related cases	Chau et al., 2002
All	Examine social media posting to identify criminal authors	Chen et al., 2004
All	Examine images on social media to identify ones that could be related to crimes	Chitrakar et al., 2016
Armed robbery	Mine police databases to identify serial crimes	Dahbur and Muscarello, 2003
All	Cluster crime reports into similar categories	Das et al., 2019
Drugs	Examine police notes and databases to determine if drugs were a cause of a crime	Duardo-Sanchez, 2011
Murder (serial)	Identify similarities between crimes in Federal databases	Helbich et al., 2013
Drugs	Identify neighborhood patterns of drug crimes	Intraligi and Buscema, 2013
Four crimes	Looks for trends based on population and other census factors across 6,100 United States cities	Kaikhah and Doddameti, 2006
Burglary	Identify similarities between previous crimes and determine if any similarity to a current crime	Keyvanpour et al., 2011
All	Identify a specific gun from bullet evidence	Kou and Tung, 1994
14 Crimes	Identify seasonality and location trends	Li et al., 2010
Drugs	Evaluate ethnicity effects in drug crimes	Massini, 2013
Murder	Classify murders and identify serial patterns	Memon and Mehboob, 2003
Burglary	Link known offenders to unsolved crimes	Oatley and Ewart, 2003
Burglary	Link known offenders to unsolved crimes	Oatley et al., 2006
Assault, robbery, murder	Determine hour of day and day of week patterns	Patil et al., 2020
Burglary, battery, and robbery	Match behavior evidence of a crime to known criminals	Rummens et al., 2017
All	Examine relations between demographics and crimes	Wu et al., 2018
All	Identify objects in crime scene images	Zhang et al., 2020

and a NN that identifies blood or other threatening objects in crime scene photos (Nakib et al., 2018). Strano (2004) describes how to develop a NN to assist investigators in performing deductive murder profiling.

Similar to investigative support, but more directly affecting police response, NN models have been developed to assist police in identifying the likelihood of a type of crime being

committed. Buscema and Intraligi (2013) report on a NN that attempts to distinguish crack based drug crimes from other types of drug crimes. To assist officers responding to an unknown crime type, a NN identifies likely crimes based on the image of a building or street scene (Fu et al., 2018). The last NN reported in this literature review is a NN that classifies anonymous tips from police hot lines into one of 20 possible categories of crime (Pinho et al., 2017), though the crime type may have been mentioned in the anonymous tip.

The NN models reviewed in this section are composed of both supervised learning ( $N=35$ ) and unsupervised learning ( $N=29$ ) NN models. The count of NN models trained using unsupervised learning methods includes the CNN hybrid method. The quantity of each learning method used in prior literature indicates that neither method is dominant, but supervised learning is used slightly more frequently. Nineteen of the 35 supervised learning examples used MLP or a derivative thereof. The difference between the type of training used for the various NN models may have to do with the type of problem being modeled. SOM and CNN NNs typically are used in other domains for image or audio classification, while supervised learning methods are used for classification and forecasting problems, and supervised learning RNN models are used for problems that are based on time, such as time-series forecasting.

The review of existing literature has shown that while NN research in criminal science is expanding, with the focus of existing research being on staffing or budget planning months or even years in the future. The research presented in this article examines utilization of data that could be collected from an initial crime report (*via* 911 in the United States or 112 in the European Union) to provide either type of crime or crime location information immediately to emergency responders. The intended immediate use of NN predictions from a minimal amount of information, distinguishes the presented research from prior research.

## NN MODELS TO PREDICT CRIME TYPE AND CRIME LOCATION

To demonstrate NN model usage in police crime predicting and preparedness, two NN prediction models are developed. Both models utilize supervised learning NN architectures. All NN models are developed using NeuralWare's® Professional II Plus® neural network shell tool.

### Data and Method

Data for the NN models were obtained from the open access RMS Crime Incidents crime data available from the city of Detroit Michigan (City of Detroit, 2020). A total 272,623 records were available at the time the research models were created. The records as grouped by Detroit criminal charge number are clustered into 38 different crimes, which are spread across 30 zip code regions of the city. A zip code is a postal code used to define geographic regions for postal delivery. The crime codes associated with each cluster are defined in the **Appendix**.

Other data included a date-time stamp, the day of the week, the hour of the day, and the longitude and latitude of the incident. The data covered a period of just over 4 years from 2016 to the first quarter, 2020.

For all NN models described, a variety of architectures are attempted, and we follow the NN research guidelines of Walczak and Cerpa (1999) and Zhang (2007), including using one or two hidden layers and varying the number of NN nodes within each hidden layer. The number of nodes in any hidden layer starts at one half the number of nodes in the previous layer and is increased incrementally until performance starts to drop. This method helps to prevent over-learning which can decrease resultant generalization performance. The incremental increase for the first hidden layer is two nodes at a time and if a second hidden layer is used, then that layer is increased a single node at a time for all possible first hidden layer architectures.

Two different training methods are used: backpropagation and radial basis function (RBF). All NN models are trained for a minimum of 150,000 training epochs and training stops when the root mean squared error becomes less than 0.05 or remains constant for 1,000 training epochs. Backpropagation is used because it is the most common NN training method and thus enables better cross study comparisons. Additionally, the backpropagation learning algorithm has been shown to be a reliable universal classifier (Hornik, 1991). RBF training sometimes outperforms backpropagation when extrapolation from data is required (as opposed to interpolation) or when data is more limited (Walczak and Cerpa, 1999). RBF NN require an initial hidden layer that utilizes a radial basis function as the transfer function, which effectively determines the proximity of an input value to a desired output value. While a large number of data examples are available for training and validation/testing, the limited number of variables used as input (independent variables) for the NN models may impose data limitations and hence the desire to also utilize RBF trained NN models. The best performing architecture of all attempted NN architectures is reported for each NN crime forecasting problem.

## NN to Predict Type of Crime

The first NN model's purpose is to predict what crime is occurring when given only location and time information. This is to simulate a 911 call response when the caller is unable to describe the crime that is occurring. It may also be viewed as a preparatory model for police officers as they are assigned to different neighborhoods on a specific day of the week and time and consequently hear a call for help. The goal of the NN is to predict the most likely crime that is occurring in order to better prepare police response.

Data were originally divided evenly into a training set and a validation set, with extra examples due to an odd number of samples for a specific crime cluster placed in the validation set. Backpropagation trained NNs with both a single hidden layer and two hidden layers, which would allow for greater nonlinearities in the solution surface (Walczak and Cerpa, 1999), are developed. The predictions from all of these NNs however only correctly classified crimes from 3 to 4 of the 38 possible

crime clusters. This is due to the very large difference in the quantities of any type of crime compared to others. For example, crimes clustered as assault crimes had 71,931 total cases, whereas the harassment and stalking crime cluster only had 41 total cases. Over represented clusters in the training data is a cause of poor classification and prediction performance by machine learning methods including NNs (Lin et al., 2013).

To compensate for the severe imbalance between the crime cluster representations, the data is pre-balanced so that only clusters that have at least 600 total samples are used, which reduces the number of clusters from 38 to 27 [removing the crime clusters for minor in possession of alcohol ( $N=10$ ), animal cruelty ( $N=3$ ), extortion ( $N=147$ ), gambling ( $N=6$ ), harassing and stalking ( $N=41$ ), health and safety ( $N=30$ ), invasion of privacy ( $N=34$ ), liquor violations ( $N=340$ ), personal protection order ( $N=6$ ), probation violation ( $N=8$ ), non-sex solicitation ( $N=25$ )]. From the remaining 27 clusters, a sample of 350 crime reports were used unless the total cases for that cluster was less than 700, then half of the cases were used. To add a very small influence for crime clusters with a very large number of cases a proportional number of extra crime samples for those clusters was added to the training set to a maximum of 370 samples for any single crime cluster. This created a training set with 9,503 training samples approximately evenly distributed between all 27 crime type clusters. The remaining 262,473 crime report samples served as the validation set for the fully trained NN. It should be noted that no validation sample is ever used for training the NN and is therefore representative of real-world prospective use of the NN.

The input variables for the NN models consisted of the day of the week and the hour of the day and one of three variables for location. Location variables could be either a categorical variable representing the 30 different zip codes within Detroit, or the longitude and latitude (LL), or a combination of both zip code and LL. The output vector for all NN models is a 27 variable set, one for each of the 27 crime clusters. The output of the NN is evaluated with respect to both prediction accuracy and cluster coverage, to avoid cases of overfitting on a small number of clusters with large sample size. Backpropagation trained NNs outperformed the RBF NN models. The best performing NN for each of these evaluation measurements is reported along with the best performing RBF model since it reflects a middle position.

## NN to Predict Crime Location

The second NN model's purpose is to predict the location of a crime when only the general type of crime (i.e., cluster) and time of day and day of week are known. This simulates the situation when an emergency call comes in and the crime reporter, victim or otherwise, is unsure of their location. Such could occur if the reporter were a visitor to the city or was in some other way unable to determine their location, such as a kidnap victim in the trunk of a vehicle. The goal of this NN is to predict the most likely zip code. The choice of the zip code area for crime location prediction follows the common practice reported by other hot spot crime estimation NNs as

shown in **Table 1**, where an area of prediction (e.g., city) is used instead of spot identification of a precise geographic location (i.e., longitude and latitude).

Both the NN to predict a crime given the location and time and the one to predict the zip code region of a crime given the crime and time focus on spatiotemporal relationships. Both NNs are ambitious in attempting to predict a specific crime happening at a specific time or a generalized location of a crime. A more precise location prediction is not attempted since prior research has shown that there is no basis for short term crime forecasting due to evidence from data showing that spatial heterogeneity and time lag cannot accurately be reflected in short-term prediction (Zhu et al., 2019).

The input values for the crime location prediction NN consists of a 27 value categorical variable for the crime cluster being reported and the day and time. The output consists of 30 variables, one representing each of the 30 zip codes present in the city of Detroit. Again, both backpropagation and RBF training are used, but the backpropagation NN models outperform the RBF models. Since location is important, and the boundaries of zip code areas are irregular, in addition to reporting exact matches to the zip code area, a fuzzy algorithm is implemented to identify when the NN predicted zip code location is contiguous with the actual zip code of the location, meaning that if the predicted and actual zip code locations share a border, then a near miss is recorded. The NN with the highest exact match as well as the highest near miss prediction percentage are reported.

## RESULTS AND DISCUSSION

The crime type NN prediction models performed best when using the zip code only for location information. The results are displayed in **Table 3**. The two-hidden layer MLP had 12 nodes in its first layer and four nodes in its second layer. The single-hidden layer MLP had 12 nodes in its hidden layer. Lastly the RBF NN had an association layer of 54 nodes and a hidden layer of 12 nodes.

Results for the NN models predicting the location, by zip code, are displayed in **Table 4**. The two-hidden layer MLP NN had 27 nodes in its first layer and six nodes in its second layer, while the single-hidden layer MLP NN had 36 nodes in its hidden layer.

Recall that for both sets of results, the prediction accuracies are over 262,799 reported crime arrests. This means for the NN that predicted 16.4% of the crimes when given just the zip code, time of day, and day of week, it was able to predict over 43,100 specific crime clusters. However, methods to improve this performance should be investigated. Much of the hot spot NN prediction research shown in **Table 1**, when not predicting a hot spot for all crimes, predicts hot spots likelihood for only a subset of all possible crimes, from 1 to 12 types of crime. If the number of crimes predicted is limited to five or six crimes, with both burglary with forced entry and without forced entry combined into the single cluster burglary, this would cover from 72.3 to 79.2% of all reported crimes. However, only evaluating the top six crimes, even though it covers almost 80% of the crimes, is

**TABLE 3** | NN prediction of crime type results (27 crime clusters).

NN model	Crime prediction accuracy (%)	Number of crime clusters predicted
Two-hidden layer backpropagation trained MLP	16.4	11
Single-hidden layer backpropagation trained MLP	7.9	24
RBF trained NN	12.9	17

**TABLE 4** | NN prediction of crime location (30 zip code regions).

NN model	Exact location prediction (%)	Within 1 (near miss and exact) location prediction (%)
Two-hidden layer MLP	8.2	28.4
Single-hidden layer MLP	7.6	31.2

still not useful to police, since they would need to know when the other 31 types of crimes are being committed.

Therefore, a NN model is developed using the top six clusters with a seventh variable representing all other types of crime. The variables are balanced as before with approximately 1,000 samples per cluster for a total of 7,607 training samples and all other data used for validation/testing with 265,016 crime samples. The same research process is duplicated for the new NN crime prediction model as used for the original crime prediction NN. Results for the new seven cluster NN crime prediction models are displayed in **Table 5**, which shows a 65% overall increase in performance at the six cluster level predicting over 27% of crimes and over a 45% increase in performance at the seven cluster prediction level predicting almost 24% of crimes across all crime clusters. This indicates that trying to be too precise concerning which crime causes some classification problems with NNs. This may be because the principle crime could be classified as many different types of crime or a crime may be misclassified with respect to the degree or severity of the crime. It is beneficial for police to be able to have advanced knowledge if a crime is violent or non-violent and future research will further examine this distinction to create new sets of clusters for prediction based on the presence of violence in a criminal situation.

Knowing the location of a crime before the actual location is determined can help police response time for these type of cases. The NN model given only the crime cluster, time of day, and day of week is able to determine the approximate location, within one contiguous zip code area (out of 30) almost one third of the time. Faster response times may improve the chance of catching the criminal and preserving crime scene evidence.

Predicting the location of a crime could be combined with other NN models performing hotspot predictions to determine when it is likely that certain types of crimes are more likely to occur and then the location NN would be able to tell the general vicinity of the crime almost. Although the reported research used the 30 zip code regions to define location and proximity, other measurements could also be used, but it should be kept in mind that as the regions get smaller, this introduces

**TABLE 5** | NN prediction of crime type for largest 6 clusters and other (clusters=7).

NN model	Crime prediction accuracy (%)	Number of crime clusters predicted
Single hidden layer MLP	27.1	6
Single-hidden layer MLP	23.8	7

a greater degree of error due to the larger number of locations that must be correlated to the different crimes. Square regions that are X blocks by Y blocks could serve as an alternate location prediction paradigm, which would eliminate need to map police response to zip code regions.

Recall that prior research has shown that there is no theoretical evidence to support short term forecasting of crimes, including both the crime and location (Zhu et al., 2019). This is why the majority of the research reported in **Table 1** predicted hot spots from 1 day to 1 year in the future. To further examine this claim, a key is created for all of the original Detroit crime data for the years 2017–2020, a total of 284,869 crime records. The keys were then searched to identify duplicate records. Only 3.4% of all crimes were repeated at the same location on the same day and at the same hour, which left 275,192 total crimes without repeats for these 39 months. Only the hit-and-run, disorderly conduct, and homicide clusters had repeat offenses at greater than 1%. For the 27 clusters defined for the crime prediction NN, nine of these clusters had no similar offenses, defined as a repeat of a crime cluster offense at the same location and at the same day and hour, but for different months and years. This does not account specifically for repeat victimization (Farrell and Sousa, 2001), since the day of week and hour of day must also match. Similar crimes with regard to geospatial and temporal data does not play a role in facilitating crime or location predictions, since there was only a 1.26% duplication between the training data and the test/validation data. The duplicate crime analysis confirms that trying to predict crimes based solely on geospatial and temporal information is problematic due to the very low repetition of crimes with similar time and location.

However, the NN models reported in this paper do just that: predicting a crime or location at the immediate time of occurrence. This is an advancement over traditional hot spot predictions which are normally for crimes overall and for much longer periods of time in the future.

## SUMMARY AND CONCLUSION

The utilization of machine learning in general and NNs in particular in crime sciences is both prevalent and expanding. Traditional NN research in crime science focuses on future hot spot predictions, datamining large corpora of data to connect relevant information together regarding a crime or serial crime or criminals, determining the probability of criminal recidivism, and analysis of crime scene objects and data to assist with criminal investigations. A small literature review of these four types of NN research in crime science is presented, focusing on more recent research. NNs have additionally been used to help identify cybercrime activities and also in a

datamining sense to identify relevant laws and decisions for jurisprudence. Both of these utilizations of NNs have not been reviewed as the focus for this article is on police practice.

Machine learning enables the automatic identification of complex and usually nonlinear associations between crimes and geospatial and temporal factors of those crimes, which are beyond the scope of traditional parametric statistical approaches. These associations further enable systems that can then predict the type of crime or crime category and also the probable location of a crime as demonstrated in the NN examples presented in this article.

The only data used for these NNs are the general type of crime, time of day, and day of the week. The other data item available to each NN model depends on what is being predicted with the zip code of occurrence as an independent variable for predicting the dependent crime cluster or the independent crime cluster variable for predicting the dependent location.

NN models are presented that can predict a specific crime cluster 16.4% of the time over 27 crime clusters or if the number of clusters is limited to seven to better focus on the most prevalent crimes, then a prediction accuracy of 27.1% is achieved. This prediction accuracy is significantly better than just randomly guessing, which would produce an estimated prediction accuracy of 14.3% with a value of  $p < 0.001$  (standard Z-Test). This type of on-time knowledge of the type of a potential crime may increase police response preparedness.

Another NN model identifies the location of a crime within one neighboring zip code, when the location is not immediately known, 31.2% of the time. This result indicates that specific crime types are at least partially geographically clustered.

These NN models show an advancement over current NN research that is more limited to forecasting overall crime for a much larger period and further into the future. Future research is needed to further investigate the use of NNs, possibly using additional temporal or different sized geospatial cues to improve near-term (immediate) crime predictions.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: [https://data.detroitmi.gov/datasets/0825badfe6304620a998d162be0e135e\\_0/data?geometry=-83.503%2C42.264%2C-82.695%2C42.442](https://data.detroitmi.gov/datasets/0825badfe6304620a998d162be0e135e_0/data?geometry=-83.503%2C42.264%2C-82.695%2C42.442).

## AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.587943/full#supplementary-material>



## REFERENCES

- Adderley, R., and Musgrove, P. B. (2001). "Data mining case study: modeling the behavior of offenders who commit serious sexual assaults." in *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; August 26, 2001 (New York: ACM), 215–220.
- Altameem, T., and Amoon, M. (2019). Crime activities prediction using hybridization of firefly optimization technique and fuzzy cognitive map neural networks. *Neural Comput. Appl.* 31, 1263–1273. doi: 10.1007/s00521-018-3561-7
- Andersson, V. O., Birck, M. A., and Araujo, R. M. (2017). "Investigating crime rate prediction using street-level images and siamese convolutional neural networks." in *Latin American Workshop on Computational Neuroscience*; November 22, 2017 (Switzerland: Springer), 81–93.
- Anuar, S., Selamat, A., and Sallehuddin, R. (2015). "Hybrid artificial neural network with artificial bee colony algorithm for crime classification." in *Computational Intelligence in Information Systems*. eds. S. Phon-Amnuaisuk, and T. Somnuk (Switzerland: Springer), 31–40.
- Babcock, J. C., and Cooper, J. (2019). Testing the utility of the neural network model to predict history of arrest among intimate partner violent men. *Safety* 5:2. doi: 10.3390/safety5010002
- Baxt, W. G. (1991). Use of an artificial neural network for the diagnosis of myocardial infarction. *Ann. Intern. Med.* 115, 843–848. doi: 10.7326/0003-4819-115-11-843
- Bochereau, L., Bourcier, D., and Bourguine, P. (1991). "Extracting legal knowledge by means of a multilayer neural network application to municipal jurisprudence." in *Proceedings of the 3rd International Conference on Artificial Intelligence and Law*; May 1, 1991 (New York: ACM), 288–296.
- Bose, B. K. (1994). Expert system, fuzzy logic, and neural network applications in power electronics and motion control. *Proc. IEEE* 82, 1303–1323. doi: 10.1109/5.301690
- Buscema, M., and Intraligi, M. (2013). "Auto-identification of a drug seller utilizing a specialized supervised neural network," in *Intelligent Data Mining in Law Enforcement Analytics*. eds. M. Buscema and W. J. Tastle (Dordrecht, Netherlands: Springer), 167–175.
- Caulkins, J., Cohen, J., Gorr, W., and Wei, J. (1996). Predicting criminal recidivism: a comparison of neural network models with statistical methods. *J. Crim. Just.* 24, 227–240. doi: 10.1016/0047-2352(96)00012-8
- Chalfin, A., and McCrary, J. (2018). Are US cities underpoliced? Theory and evidence. *Rev. Econ. Stat.* 100, 167–186. doi: 10.1162/REST\_a\_00694
- Chang, L. Y. (2005). Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network. *Saf. Sci.* 43, 541–557. doi: 10.1016/j.ssci.2005.04.004
- Chau, M., Xu, J., and Chen, H. (2002). "Extracting meaningful entities from police narrative reports." in *Proceedings of the National Conference for Digital Government Research*; May 19, 2002 (New York: ACM), 1–5.
- Chen, H., Chung, W., Xu, J. J., Wang, G., Qin, Y., and Chau, M. (2004). Crime data mining: a general framework and some examples. *Computer* 37, 50–56. doi: 10.1109/MC.2004.1297301
- Chitrakar, P., Zhang, C., Warner, G., and Liao, X. (2016). "Social media image retrieval using distilled convolutional neural network for suspicious e-crime and terrorist account detection." in *2016 IEEE International Symposium on Multimedia (ISMM)*; December 11, 2016; 493–498.
- Chun, S. A., Paturu, V. A., Yuan, S., Pathak, R., Atluri, V., and Adam, N. R. (2019). "Crime prediction model using deep neural networks." in *Proceedings of the 20th Annual International Conference on Digital Government Research*; June 18, 2019 (New York: ACM), 512–514.
- City of Detroit (2020). *RMS Crime Incidents*. Available at: <https://data.detroitmi.gov/datasets/rms-crime-incidents> (Accessed June 2, 2020).
- Corcoran, J. J., Wilson, I. D., and Ware, J. A. (2003). Predicting the geotemporal variations of crime and disorder. *Int. J. Forecast.* 19, 623–634. doi: 10.1016/S0169-2070(03)00095-5
- Cortez, B., Carrera, B., Kim, Y. J., and Jung, J. Y. (2018). An architecture for emergency event prediction using LSTM recurrent neural networks. *Expert Syst. Appl.* 97, 315–324. doi: 10.1016/j.eswa.2017.12.037
- Costa, A. M. (2010). The economics of crime: a discipline to be invented and a Nobel prize to be awarded. *J. Policy Model* 32, 648–661. doi: 10.1016/j.jpolmod.2010.07.010
- Dahbur, K., and Muscarello, T. (2003). Classification system for serial criminal patterns. *Artif. Intell. Law* 11, 251–269. doi: 10.1023/B:ARTI.0000045994.96685.21
- Das, P., Das, A. K., Nayak, J., Pelusi, D., and Ding, W. (2019). Group incremental adaptive clustering based on neural network and rough set theory for crime report categorization. *Neurocomputing*. doi: 10.1016/j.neucom.2019.10.109 (in press).
- Duan, L., Hu, T., Cheng, E., Zhu, J., and Gao, C. (2017). "Deep convolutional neural networks for spatiotemporal crime prediction." in *Proceedings of the International Conference on Information and Knowledge Engineering*; July 17, 2017, 61–67.
- Duardo-Sanchez, A. (2011). "Criminal law networks, Markov chains, Shannon entropy and artificial neural networks," in *Complex Network Entropy: From Molecules to Biology, Parasitology, Technology, Social, Legal, and Neurosciences*. eds. H. González-Díaz, F. J. Prado-Prado and X. García-Mera (Kerala, India: Transworld Research Network), 107–114.
- El-Khatib, M. J., Abu-Nasser, B. S., and Abu-Naser, S. S. (2019). Glass classification using artificial neural network. *Int. J. Acad. Res.* 3, 25–31.
- Esquivel, N., Peralta, B., and Nicolis, O. (2019). "Crime level prediction using stacked maps with deep convolutional autoencoder." in *2019 IEEE Chilean Conference on Electrical, Electronics Engineering, Information and Communication Technologies (IEEE)*; November 13, 2019; 1–7.
- Fajnzylber, P., Lederman, D., and Loayza, N. (2002). What causes violent crime? *Eur. Econ. Rev.* 46, 1323–1357. doi: 10.1016/S0014-2921(01)00096-4
- Farrell, G., and Sousa, W. (2001). Repeat victimization and hot spots: the overlap and its implications for crime control and problem-oriented policing. *Crime Prevention Studies* 12, 221–240.
- Feng, M., Zheng, J., Ren, J., Hussain, A., Li, X., Xi, Y., et al. (2019). Big data analytics and mining for effective visualization and trends forecasting of crime data. *IEEE Access* 7, 106111–106123. doi: 10.1109/ACCESS.2019.2930410
- Fernandes, W. P. D., Silva, L. J. S., Frajhof, I. Z., Konder, C. N., Nasser, R. B., de Carvalho, G. R., et al. (2019). Appellate court modifications extraction for portuguese. *Artif. Intell. Law* 28, 327–360. doi: 10.1007/s10506-019-09256-x
- Fu, K., Chen, Z., and Lu, C. T. (2018). "Streetnet: preference learning with convolutional neural network on urban crime perception." in *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*; November 6, 2018 (New York: ACM), 269–278.
- Gallo, C. (2015). "Artificial neural networks tutorial," in *Encyclopedia of Information Science and Technology*. 3rd Edn. ed. M. Khosrow-Pour (Hershey, PA: IGI Global), 6369–6378.
- Hashim, E. N., and Nohuddin, P. N. (2018). Data mining techniques for recidivism prediction: a survey paper. *Adv. Sci. Lett.* 24, 1616–1618. doi: 10.1166/asl.2018.11121
- Helbich, M., Hagenauer, J., Leitner, M., and Edwards, R. (2013). Exploration of unstructured narrative crime reports: an unsupervised neural network and point pattern analysis approach. *Cartogr. Geogr. Inf. Sci.* 40, 326–336. doi: 10.1080/15230406.2013.779780
- Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. *Neural Netw.* 4, 251–257. doi: 10.1016/0893-6080(91)90009-T
- Huang, C., Zhang, J., Zheng, Y., and Chawla, N. V. (2018). "DeepCrime: attentive hierarchical recurrent networks for crime prediction." in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*; October 17, 2018 (New York: ACM), 1423–1432.
- Intraligi, M., and Buscema, M. (2013). "Application of the constraint satisfaction network" in *Intelligent Data Mining in Law Enforcement Analytics*. eds. M. Buscema and W. J. Tastle (Dordrecht, Netherlands: Springer), 231–313.
- Jain, B., Huber, M., Fegaras, L., and Elmasri, R. A. (2019). "Singular race models: addressing bias and accuracy in predicting prisoner recidivism." in *Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*; June 5, 2019 (New York: ACM), 599–607.
- Jain, A. K., Mao, J., and Mohiuddin, K. M. (1996). Artificial neural networks: a tutorial. *Computer* 29, 31–44. doi: 10.1109/2.485891
- Jankowski, S., Szymański, Z., Wawrzyniak, Z., Cichosz, P., Szczechla, E., and Pytlak, R. (2018). "Selection of neural network for crime time series prediction by virtual leave-one-out tests." in *International Conference on Time Series and Forecasting*; September 19, 2018 (Switzerland: Springer), 117–133.

- Jayalakshmi, T., and Santhakumaran, A. (2011). Statistical normalization and back propagation for classification. *Int. J. Comput. Theory Eng.* 3, 1793–8201.
- Jha, S., Yang, E., Almagrabi, A. O., Bashir, A. K., and Joshi, G. P. (2020). Comparative analysis of time series model and machine testing systems for crime forecasting. *Neural Comput. Appl.* 33, 10621–10636. doi: 10.1007/s00521-020-04998-1
- Jin, G., Wang, Q., Zhu, C., Feng, Y., Huang, J., and Zhou, J. (2020). “Addressing crime situation forecasting task with temporal graph convolutional neural network approach.” in *12th International Conference on Measuring Technology and Mechatronics Automation (IEEE)*; February 28, 2020; 474–478.
- Kaikhah, K., and Doddameti, S. (2006). Discovering trends in large datasets using neural networks. *Appl. Intell.* 24, 51–60. doi: 10.1007/s10489-006-6929-9
- Kang, H. W., and Kang, H. B. (2017). Prediction of crime occurrence from multi-modal data using deep learning. *PLoS One* 12:e0176244. doi: 10.1371/journal.pone.0176244
- Keyvanpour, M. R., Javideh, M., and Ebrahimi, M. R. (2011). Detecting and investigating crime by means of data mining: a general crime matching framework. *Procedia Comput. Sci.* 3, 872–880. doi: 10.1016/j.procs.2010.12.143
- Kou, C., and Tung, C. T. (1994). “FISOFM: firearms identification based on softm model of neural network.” in *1994 Proceedings of IEEE International Carnahan Conference on Security Technology*; ed. L. D. Sanson; October 12, 1994; IEEE; 120–125.
- Kounadi, O., Ristea, A., Araujo, A., and Leitner, M. (2020). A systematic review on spatial crime forecasting. *Crime Sci.* 9, 1–22.
- Kouziokas, G. N. (2017). The application of artificial intelligence in public administration for forecasting high crime risk transportation areas in urban environment. *Transp. Res. Proc.* 24, 467–473. doi: 10.1016/j.trpro.2017.05.083
- Krishnan, A., Sarguru, A., and Sheela, A. S. (2018). Predictive analysis of crime data using deep learning. *Int. J. Pure Appl. Math.* 118, 4023–4031.
- Li, J., and He, P. (2020). Detection and prevention of cyber crime based on diamond factor neural network. *J. Phys. Conf. Ser.* 1437:012011
- Li, S. T., Kuo, S. C., and Tsai, F. C. (2010). An intelligent decision-support model using FSOM and rule extraction for crime prevention. *Expert Syst. Appl.* 37, 7108–7119. doi: 10.1016/j.eswa.2010.03.004
- Lin, M., Tang, K., and Yao, X. (2013). Dynamic sampling approach to training neural networks for multiclass imbalance classification. *IEEE Trans. Neural Netw. Learn. Syst.* 24, 647–660. doi: 10.1109/TNNLS.2012.2228231
- Lin, Y. L., Yen, M. F., and Yu, L. C. (2018). Grid-based crime prediction using geographical features. *ISPRS Int. J. Geo Inf.* 7:298. doi: 10.3390/ijgi7080298
- Liu, Y. Y., Yang, M., Ramsay, M., Li, X. S., and Coid, J. W. (2011). A comparison of logistic regression, classification and regression tree, and neural networks models in predicting violent re-offending. *J. Quant. Criminol.* 27, 547–573. doi: 10.1007/s10940-011-9137-7
- Loeffler, C., and Flaxman, S. (2018). Is gun violence contagious? A spatiotemporal test. *J. Quant. Criminol.* 34, 999–1017. doi: 10.1007/s10940-017-9363-8
- Mao, L., and Du, W. (2019). A method of crime rate forecast based on wavelet transform and neural network. *Int. J. Embed. Syst.* 11, 731–737. doi: 10.1504/IJES.2019.103990
- Massini, G. (2013). “Self-organizing maps: identifying nonlinear relationships in massive drug enforcement databases.” in *Intelligent Data Mining in Law Enforcement Analytics*. eds. M. Buscema and W. J. Tastle (Dordrecht, Netherlands: Springer), 193–214.
- Memon, Q. A., and Mehboob, S. (2003). “Crime investigation and analysis using neural nets.” in *7th International Multi Topic Conference (IEEE)*; December 8, 2003; 346–350.
- Mena, J. (2011). *Machine Learning Forensics for Law Enforcement, Security, and Intelligence*. Boca Raton, FL: CRC Press.
- Nair, S. N., and Gopi, E. S. (2020). “Deep learning techniques for crime hotspot detection.” in *Optimization in Machine Learning and Applications*. eds. A. Kulkarni and S. Satapathy (Singapore: Springer), 13–29.
- Nakib, M., Khan, R. T., Hasan, M. S., and Uddin, J. (2018). “Crime scene prediction by detecting threatening objects using convolutional neural network.” in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IEEE)*; February 8, 2018; 122–125.
- Oatley, G. C., and Ewart, B. W. (2003). Crimes analysis software: ‘pins in maps’, clustering and Bayes net prediction. *Expert Syst. Appl.* 25, 569–588. doi: 10.1016/S0957-4174(03)00097-6
- Oatley, G., Ewart, B., and Zeleznikow, J. (2006). Decision support systems for police: lessons from the application of data mining techniques to “soft” forensic evidence. *Artif. Intell. Law* 14, 35–100. doi: 10.1007/s10506-006-9023-z
- Olligschlaeger, A. M. (1997). “Artificial neural networks and crime mapping,” in *Crime Mapping and Crime Prevention*. eds. D. Weisburd and T. McEwen (Cincinnati, OH: Criminal Justice Press), 313–348.
- Palocsay, S. W., Wang, P., and Brookshire, R. G. (2000). Predicting criminal recidivism using neural networks. *Socio Econ. Plan. Sci.* 34, 271–284. doi: 10.1016/S0038-0121(00)00003-3
- Patil, A. P., Nawal, D. J., and Jain, D. (2020). “Crime prediction application using artificial intelligence.” in *Proceedings of ICETIT 2019*; Switzerland: Springer, 238–245.
- Pinho, R. A., Brito, W. A., Motta, C. L., and Lima, P. V. (2017). “Automatic crime report classification through a weightless neural network.” in *Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*; Aril 26, 2017; 165–170.
- Rummens, A., Hardyns, W., and Pauwels, L. (2017). The use of predictive analysis in spatiotemporal crime forecasting: building and testing a model in an urban context. *Appl. Geogr.* 86, 255–261. doi: 10.1016/j.apgeog.2017.06.011
- Seo, S., Chan, H., Brantingham, P. J., Leap, J., Vayanos, P., Tambe, M., et al. (2018). “Partially generative neural networks for gang crime classification with partial information.” in *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*; New York: ACM, 257–263.
- Siddiqui, M. Z., Yadav, S., and Husain, M. S. (2018). Application of artificial intelligence in fighting against cyber crimes: a review. *Int. J. Adv. Res. Comput. Sci.* 9, 118–121. doi: 10.26483/ijarcs.v9i0.6149
- Stebbins, S. (2019). *The Midwest is Home to Many of America’s Most Dangerous Cities*. USA Today. Available at: <https://www.usatoday.com/story/money/2019/10/26/crime-rate-higher-us-dangerous-cities/40406541/> (Accessed July 15, 2020).
- Stranieri, A., and Zeleznikow, J. (2012). “Insights from jurisprudence for machine learning in law,” in *Machine Learning Algorithms for Problem Solving in Computational Applications: Intelligent Techniques*. ed. S. Kulkarni (Hershey, PA: IGI Global), 85–98.
- Strano, M. (2004). A neural network applied to criminal psychological profiling: an Italian initiative. *Int. J. Offender Ther. Comp. Criminol.* 48, 495–503. doi: 10.1177/0306624X04265284
- Sundararaj, A., Ravi, R., Thirumalai, P., and Radhakrishnan, G. (1999). Artificial neural network applications in electrochemistry: a review. *Bull. Electrochem.* 15, 552–555.
- Tastle, W. J. (2013). “Introduction to artificial networks and law enforcement analytics,” in *Intelligent Data Mining in Law Enforcement Analytics*. eds. M. Buscema and W. J. Tastle (Dordrecht, Netherlands: Springer), 1–9.
- Tumulak, J. A. U., and Espinosa, K. J. P. (2017). “Crime modelling and prediction using neural networks.” in *Theory and Practice of Computation: Proceedings of Workshop on Computation: Theory and Practice*. 218–228.
- Vellido, A., Lisboa, P. J., and Vaughan, J. (1999). Neural networks in business: a survey of applications (1992–1998). *Expert Syst. Appl.* 17, 51–70. doi: 10.1016/S0957-4174(99)00016-0
- Walczak, S., and Cerpa, N. (1999). Heuristic principles for the design of artificial neural networks. *Inf. Softw. Technol.* 41, 107–117. doi: 10.1016/S0950-5849(98)00116-5
- Wang, Q., Jin, G., Zhao, X., Feng, Y., and Huang, J. (2020). CSAN: a neural network benchmark model for crime forecasting in spatio-temporal scale. *Knowl.-Based Syst.* 189:105120. doi: 10.1016/j.knosys.2019.105120
- Wang, B., Yin, P., Bertozzi, A. L., Brantingham, P. J., Osher, S. J., and Xin, J. (2019). Deep learning for real-time crime forecasting and its ternarization. *Chinese Ann. Math. Ser. B* 40, 949–966. doi: 10.1007/s11401-019-0168-y
- Wang, B., Zhang, D., Zhang, D., Brantingham, P. J., and Bertozzi, A. L. (2017a). “Deep learning for real time crime forecasting.” in *2017 International Symposium on Nonlinear Theory and Its Applications*; July 9, 2017.
- Wang, K., Zhu, P., Zhu, H., Cui, P., and Zhang, Z. (2017b). “An interweaved time series locally connected recurrent neural network model on crime forecasting.” in *International Conference on Neural Information Processing*; November 14, 2017 (Switzerland: Springer), 466–474.
- Wawrzyniak, Z. M., Jankowski, S., Szczechla, E., Szymański, Z., Pytlak, R., Michalak, P., et al. (2018). “Data-driven models in machine learning for

- crime prediction.” in *2018 26th International Conference on Systems Engineering*; December 18, 2018; IEEE; 1–8.
- Wei, Q., and Dunbrack, R. L. Jr. (2013). The role of balanced training and testing data sets for binary classifiers in bioinformatics. *PLoS One* 8:e67863. doi: 10.1371/journal.pone.0067863
- Wu, S., Wang, C., Cao, H., and Jia, X. (2018). “Crime prediction using data mining and machine learning.” in *The 8th International Conference on Computer Engineering and Networks*; eds. Q. Liu, M. Misir, X. Wang and W. Liu; August 17, 2018; Switzerland: Springer.
- Wythoff, B. J. (1993). Backpropagation neural networks: a tutorial. *Chemom. Intell. Lab. Syst.* 18, 115–155. doi: 10.1016/0169-7439(93)80052-J
- Yi, F., Yu, Z., Zhuang, F., and Guo, B. (2019). “Neural network based continuous conditional random field for fine-grained crime prediction.” in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI*; August 10, 2019; 10–16.
- Yu, H., Liu, F., and Zou, K. (2012). Improved fuzzy bp neural network and its application in crime prediction. *J. Liaoning Tech. Univ. (Nat. Sci.)* 2:025.
- Yu, C. H., Ward, M. W., Morabito, M., and Ding, W. (2011). “Crime forecasting using data mining techniques.” in *2011 IEEE 11th International Conference on Data Mining Workshops*; December 11, 2011; IEEE; 779–786.
- Zhang, G. P. (2007). Avoiding pitfalls in neural network research. *IEEE Trans. Syst. Man Cybern. Syst.* 37, 3–16. doi: 10.1109/TSMCC.2006.876059
- Zhang, X. Q. (2016). The trends, promises and challenges of urbanisation in the world. *Habitat Int.* 54, 241–252. doi: 10.1016/j.habitatint.2015.11.018
- Zhang, Q., Liu, Y., Wang, F., Lu, J., and Li, D. (2020). “Fusion CNN based on feature selection for crime scene investigation image classification,” in *Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery*; August 9, 2020. eds. Y. Liu, L. Wang, L. Zhao and Z. Yu, ICNC-FSKD 2019. *Advances in Intelligent Systems and Computing* (Switzerland: Springer), 773–780.
- Zhu, Q., Zhang, F., Liu, S., and Li, Y. (2019). An anticrime information support system design: application of K-means-VMD-BiGRU in the city of Chicago. *Inf. Manag.* 2019. doi: 10.1016/j.im.2019.103247
- Zhuang, Y., Almeida, M., Morabito, M., and Ding, W. (2017). “Crime hot spot forecasting: a recurrent model with spatial and temporal information.” in *2017 IEEE International Conference on Big Knowledge*; August 9, 2017; IEEE; 143–150.
- Conflict of Interest:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2021 Walczak. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.