Analyzing Non-linearity and Threshold Effect between Street-Level Built Environments and Local Crime Patterns: An Interpretable Machine Learning Approach

Sugie Lee*, Donghwan Ki**, John R. Hipp***, and Jae Hong Kim****

*Sugie Lee, Ph.D. Corresponding author
Department of Urban Planning & Engineering,
Hanyang University
222 Wangshimni-ro, Sungdong-gu, Seoul, Korea
Email: sugielee@hanyang.ac.kr

**Donghwan Ki
Department of City and Regional Planning
The Ohio State University
Email: ki.17@osu.edu

***John R. Hipp, Ph.D.
Department of Criminology, Law and Society and Department of Sociology
University of California, Irvine
Email: hippj@uci.edu

****Jae Hong Kim, Ph.D.
Department of Urban Planning and Public Policy
University of California, Irvine
206E Social Ecology I, Irvine, California 92697
Phone: 949.824.0449
Fax: 949.824.8566
Email: jaehk6@uci.edu

April 10, 2024

1. Introduction

Urban scholars and criminologists are interested in building safe and sustainable built environments. During the past several decades, researchers have examined the relationship between crime patterns and social and physical dimensions of the surrounding environment through the lens of environmental criminology (Brantingham and Brantingham, 1981), crime pattern theory (Brantingham and Brantingham, 1984), or routine activity theory (Cohen and Felson, 1979). Some scholars have also explored ways to achieve crime prevention through environmental design (CPTED). Recent studies have found support for the positive effects of CPTED strategies, such as surveillance, access control, or territorial enforcement, in reducing certain types of crime and increasing public safety and highlighted the importance of the visual aspects of the built environment beyond the conventional physical elements that attract crimes (Cozens and Love, 2015). In other words, eye-level three-dimensional built environment characteristics have increasingly been viewed as a key determinant of criminal activities.

On the other hand, it is challenging to measure the environmental features precisely using conventional approaches. As a consequence, scholars in various fields, including criminology, have attempted to audit fine-scale environmental features using street images and computer vision techniques (Gong et al., 2018; He et al., 2017; Hipp et al., 2022; Lu, 2018). This approach has several advantages, including that it facilitates the measurement of detailed built environment features at the pedestrian level and it is cost effective at a large scale. However, this promising approach has only rarely been applied in environmental criminology studies.

Of particular interest to us here is the potential nonlinear relationships between some measures of the built environment and crime levels. There is growing attention to non-linearity in the criminology field (e.g., Chamberlain et al., 2021; He et al., 2020; Walker, 2007; Zhang et al., 2022), and below we highlight theoretical reasons why there might be nonlinear relationships between certain built environment features and crime. We use a machine learning strategy for this question given that the nonparametric strategy of machine learning is particularly well suited to this research question. A variety of academic fields have adopted machine learning models and reported their superiority in the identification of relationships when relaxing the linear and parametric assumptions of common estimation strategies. In addition, the recent development of interpretable machine learning (IML) techniques makes it feasible to understand the nature of the black box of machine learning algorithms. This methodology interprets the model's architecture, which provides a deeper understanding of the relationship between variables as well as credibility to the model's results, and allows us to detect possible nonlinear relationships (Zhang et al., 2022).

Thus, this study aims to investigate the nonlinear relationships and threshold effects between crime patterns and street-level neighborhood environments using machine learning models and the IML technique. To measure the neighborhood environment at the eye level, this study utilizes semantic segmentation techniques for GSV images. This approach enables us to effectively capture street-level built environments and quantify them numerically. Then, we build several machine learning models such as Random Forest(RF), Support Vector Machine(SVM), XGBoost, Artificial Neural Network(ANN), and Deep Neural Network(DNN) for crime prediction and compare their predictive accuracy with a traditional statistical model to identify the best-performing model. Finally, the SHAP interpretable machine learning algorithm is applied to show the implications of the best-performing model. Through this approach, we explore the nonlinear relationships and threshold effects between crime patterns and the built environment, which can provide insights into the complex relationships between them and the policy implications for public safety.

2. Literature Review

2.1. Crime and the Built Environment

Environmental criminology theories focus on the surrounding environment in which crime events occur (Brantingham and Brantingham, 1981; Wortley and Mazerolle, 2008). According to Brantingham and Brantingham (1981), criminal events are the consequences of complicated interactions among offenders, victims, law, and places. They proposed crime pattern theory, which states that the specific setting of places in time and space plays a fundamental role in initiating criminal activities. Wortley and Mazerolle (2008) provided three key elements of environmental criminology: the importance of the immediate environment, non-random characteristics of criminal activities in time and space, and the importance of criminogenic environments for crime prevention and control.

Environmental criminology shares a theoretical background with the concepts of "CPTED" formulated by the early contributions of Jeffery (1971) and Newman (1972). Newman's (1972) defensible space theory emphasized the importance of architectural and environmental design to promote a sense of territoriality and social responsibility that controls criminal behaviors. He argued that defensible spaces limiting access to strangers or potential criminals are more effective at reducing criminal activities. Jacobs (1961) also pointed out that the lack of natural guardianship on the streets is associated with criminal activities. She argued that permeable streets would give safe environments with natural surveillance from "eyes on the street." According to her theory, built environments, such as small block sizes, high-density mixed-use residential environments, and sidewalks are more likely to increase urban vitality that promotes urban safety from criminal activities.

Based on the CPTED concept, several studies have investigated the relationship between specific neighborhood environments and crime (He et al., 2017; Langton and Steenbeek, 2017). However, the effects of environmental factors are inconsistent in the literature. For instance, physical barriers, such as walls and fences, are known to mitigate crime by blocking access to potential offenders (He et al., 2017; Langton and Steenbeek, 2017), but they can impair surveillance opportunities and create opportunities for crime via blocking visibility (Cozen and Love, 2015). In this regard, some have noted that the overreliance on physical barriers can be detrimental to another concept of crime deterrence, "eyes on the street" (Jacobs, 1961), by forming a "fortress mentality" (Cozen and Love, 2015).

It is also important to note that the relationship between the presence of vegetation and crime is debated. Some studies demonstrated that the greenspace of parks facilitates crime, as it increases the influx of potential targets with less informal social control (Groff and McCord, 2012). Furthermore, vegetation can hide perpetrators and criminal activities from bystanders (Wolfe and Mennis, 2012). In terms of perceptions, considerable vegetation can increase the level of fear as people feel a sense of visual closure (Baran et al., 2018). On the other hand, vegetation can serve as a place to gather people, increasing natural surveillance to prevent crime (Kondo et al., 2016; Troy et al., 2016). Additionally, natural elements, especially vegetation, suppress crime by reducing the mental stress that can incite crime, especially in low-income neighborhoods (Burley, 2018).

While these equivocal findings are, in part, attributable to the different data sources and measurement strategies used in the previous studies, the mixed findings may suggest that the relationship between the environment and crime is much more complex than the way it is traditionally portrayed. In terms of data collection, most studies utilized public or private archival data (Groff and McCord, 2012; Troy et al., 2016), but some of the studies employed satellite images, field surveys, and street images for manual audits (He et al., 2017; Langton and Steenbeek, 2017; Wolfe and Mennis, 2012).

As to the reason for the conflicting results, we can also infer the possible existence of nonlinear relationships between built environments and crime. Walker (2007) pointed out the adherence to linear modeling as one of the reasons for ambiguity in the relationship between crime patterns and neighborhood environments. Several empirical studies also have demonstrated nonlinear relationships between important variables, including those focusing on racial composition, land use, and vitality (Browning et al., 2010; He et al., 2020; Hipp et al., 2019; Wheeler and Steenbeek, 2020). As an example, Browning et al. (2010) substantiated an inverted U-shape relationship between mixed land use and violent crime occurrence, which can be interpreted by drawing potential victims or regulatory effects due to vibrancy. Hence, focusing on nonlinearity has the potential to clarify the complicated relationships between crime occurrence and neighborhood environments.

2.2. Scale of Analysis and Street Image Applications

Benchmarking the seminal study of Weisburd et al. (2004) that used longitudinal crime data at the level of street segments, Groff et al. (2010) found that many individual street segments have heterogeneous trajectories that are unrelated to the immediately adjacent streets. They demonstrated the importance of micro-level analysis of crime events. This finding is consistent with existing studies that pointed out the importance of using a micro-level unit of analysis in criminology studies (Sherman et al, 1989). The micro-level unit of analysis of the street segment has theoretical and methodological advantages and can better suggest specific policy implications over aggregated units of analysis, such as block groups, census tracts, or other units of spatial aggregation (Groff et al., 2010; Sherman et al., 1989). From the methodological perspectives of criminology studies, small units of analysis, such as the street segment, are more likely to reduce spatial heterogeneity and the ecological fallacy and produce more appropriate statistical analysis results. Furthermore, the small unit of analysis can identify hot spots of criminal activities and suggest policing strategies to mitigate potential criminal activities.

Also, it should be emphasized the importance of fine-scale physical features among crime determining environmental factors (He et al., 2017; Macro et al., 2017; Vandeviver, 2014). For instance, physical disorder (e.g., cigarette butts, empty bottles, and graffiti) and physical decay (e.g., vacant or abandoned houses and commercial buildings) have been established as crime determinants (He et al., 2017; Macro et al., 2017). However, the conventional approach is limited in collecting fine-scale quantitative environmental features at a small unit of analysis. In particular, few criminology studies have been conducted in a small unit of analysis, especially the street segment, due to the lack of built environment data. Also, field audit is limited in collecting fine-scale quantitative environmental data for large-scale areas.

Recently, various studies have attempted to measure environmental features using street-view images with computer vision techniques. These studies created measures assumed to be crime determinants (He et al., 2017; Hipp et al., 2022; Marco et al., 2017) or public-health-associated environmental features (Ki and Lee, 2021; Lu, 2018). This approach has several advantages. First and foremost, this approach enables us to assess fine-scale environmental features (e.g., streetscapes) in a small unit of analysis. Second, once the specific algorithm is created, computer vision techniques can be applied to numerous street images and thus can be applied to a large-scale target site. Third, using street images has relatively high accuracy compared to a field survey in measuring street trees and buildings (Gong et al., 2018), and is highly efficient in evaluating the neighborhood environment on a large scale (He et al., 2017; Vandeviver, 2014). Forth since the street view contains an actual landscape at the street level, it is possible to measure the neighborhood environment in detail from pedestrians' perspectives. At present, however, only a few empirical studies have measured the neighborhood environment using this approach to analyze the associations with crime (e.g., Hipp et al., 2022).

2.3. Big Data and Machine Learning Applications

Big data and machine learning algorithms have attracted increasing attention, as they can provide researchers with new data sources and predictive analytical tools for creating safer environments (Kang and Kang, 2017; Rummens et al., 2017). Rummens et al. (2017) demonstrated the use of predictive analysis in spatiotemporal crime forecasting in an urban context. Focusing on three types of crime, they tested three analysis models: logistic regression, a neural network model, and an ensemble model. The analysis results indicated that each model had its advantages in terms of performance measures based on spatial and temporal resolutions. Similarly, Wheeler and Steenbeek (2020) conducted long-term predictions of crime at micro places using a machine learning algorithm with RFs. For robberies in Dallas at 200 by 200 ft^2 grid cells, this study concluded that the RF greatly outperformed risk terrain models and kernel density estimation in terms of forecasting future crimes. This study also illustrated one strategy to generate interpretable model summaries of RFs that could be helpful in understanding the predictive importance for crime at micro places.

Despite recent studies applying machine learning algorithms to criminology, a critical limitation of machine learning algorithms is the nature of the black box in the process of prediction. Zhang et al. (2018) pointed out that although artificial neural networks (ANNs) are powerful tools for modeling associations between variables for the best prediction of an outcome, they are limited in their problem-solving ability by the nature of the "black box model", as one does not obtain a set of coefficients associated with the variables in the model, as is typical in common estimation strategies. As Das and Tsapakis (2020) pointed out, interpretability is crucial to solving problems in the machine learning approach. Their study focused on the interpretable machine learning approach in estimating traffic volume and described the advantages and disadvantages of several IML techniques, such as partial dependence plots (PDFs), individual conditional expectation (ICE), LIME, and SHapley Additive exPlanations (SHAP).

2.4. Research Gaps and Research Directions

In quantitative criminology studies, conventional statistical models, especially negative binomial and Poisson regression, are the dominant methods used to examine the impacts of explanatory variables on crime. However, these traditional statistical models have limited ability to elucidate the complex relationship between crime and built environments. Furthermore, when used in a predictive setting, they have generally exhibited lower accuracy than machine learning models (Das and Tsapakis, 2020; Kang and Kang, 2017). Thus, using a machine learning approach can clarify the relationship between neighborhood environments and crime with high accuracy.

Recently, a few studies (Deng et al., 2022; Hipp et al., 2022) have delved into using GSV images to assess micro-level built environments and crime patterns. These studies employed the semantic segmentation technique of GSV images but predominantly utilized traditional regression methods to explore the associations between neighborhood built environments and crime patterns. Recent studies have also highlighted the intricate relationships between built environments and crime, emphasizing the necessity for novel techniques capable of analyzing non-linear relationships and their threshold effects (Hipp et al., 2022; Wheeler & Steenbeek).

On the other hand, understanding the built environment-crime relationship requires precise measurement of the fine-scale built environment features (He et al., 2017; Marco et al., 2017; Vandeviver, 2014). Accordingly, a growing number of studies have utilized street images to measure the visual aspects of the built environment at eye level. Previous research, however, has often relied on manual audits. Using an advanced computer vision algorithm can give us new opportunities to take full advantage of street images and measure various dimensions of the built environment more efficiently.

Finally, the applications of machine learning algorithms in criminology studies have been criticized because of the inability to understand the directions and strengths of explanatory variables on crime due to the black-box nature of machine learning algorithms. In other words, it is important to have a means to interpret what machine learning models capture when using them. Explanatory variables and their potentially nonlinear impacts on crime patterns are highly critical issues in criminology studies that demand policy implications for crime prevention. To fill this gap, we utilize machine learning models and interpret the model results through IML techniques to elucidate the complex relationships between crime patterns and the built environment. Furthermore, we identify non-linearity and thresholds effects between street-level neighborhood built environmental variables and crime patterns and suggests policy implications to promote urban safety.

3. Methodology

3.1. Study Area and Data

To investigate the relationship between crime patterns and neighborhood environments, this study focuses on the City of Santa Ana, California (Figure 1). Santa Ana is located in Orange County, California, with a current size of 70.85 km^2 and a population of nearly 330,000. The unit of analysis is a road segment, and there are 5,343 segments in Santa Ana, with an average length of 136 m, ranging from a minimum of 20 m to a maximum of 749 m. Among the 5,343 segments, we excluded 826 segments for which GSV images were not available.

The dependent variable is the number of crimes that occurred on a road segment from 2017 to 2019 (36 months). The number of crimes per year is 7,227, 6,615, and 6,782, respectively. The crime dataset was provided by the local police agency and includes the location of each crime. We geocoded the crime points to the nearest segment[1], and a total of 20,624 Part 1 crimes (aggravated assaults, robberies, burglaries, motor vehicle thefts, and larcenies) were analyzed.

In general, each crime type has different determinants (e.g., Hipp et al., 2022), so it could be appropriate to build a model for each crime type. Notwithstanding, it should be noted that there is a methodological challenge to doing so. About 80-90% of street segments have zero value across the five crime types in our study site (e.g., larcenies – 94.4% and burglaries – 70.1%). This left-skewed distribution can definitely affect the performance of model for a specific crime type, whereas the distribution of aggregating all crime types is more suitable as only about half (52.1%) of the training samples have zero values for all Part 1 crimes.

Also, there is a spatial homogeneity in the distribution of the crime types, which mitigates the potential fallacy of combining all crime types together. To demonstrate this, we conducted a principal component factor analysis of these five separate crime types, and the results showed just one single factor. Furthermore, all five crime types quite strongly loaded on this single factor, with factor scores almost all at 0.72 (ranging from larceny at 0.78 to aggravated assault at 0.66). Thus, we see that the spatial patterns of these crime types are quite similar.

[1] The geocoding match rate is 95.23%

Figure 1. Location of Santa Ana

3.2. Research Framework

As mentioned above, this study utilizes machine learning models to better capture the relationship between the built environment and crime and to draw insights from the best-performing model through the interpretable machine learning (IML) approach. To this end, this study is conducted in the following steps (Figure 2). First, this study constructs a dataset with the dependent variable being the number of all Part 1 crimes, and independent variables being demographic, socio-economic, and built environment measures. Specifically, eye-level streetscape features are measured through Google Street View (GSV) and the semantic segmentation for each of the 4,517 road segments is explained in Section 3.4.

Second, we train several machine learning models using the training dataset (n=3,613) and conduct a comparative evaluation of models through the test dataset (n=678) to identify the best performing model. We tune the models' hyperparameters during the training process using the validation dataset (n=678) to select optimal parameters for the machine learning model (we describe the logic of these sample sizes in the next section). Additionally, we assess the performance of the machine learning models compared to the traditional statistical model (negative binomial regression).

Figure 2. Research framework

Finally, the model with the highest accuracy in the previous step is interpreted using an IML algorithm. The IML algorithm can improve the reliability of the model by interpreting the model (Ribeiro et al., 2016), and it is possible to determine the relationships between predictor variables and the response variable. Specifically, this study employs SHAP, one of the most promising IML approaches. This is because the SHAP methodology can not only provide global interpretation, but also local interpretation, which lends itself toward determining non-linearity (García and Aznarte, 2020; Lundberg et al., 2020), as explained in Section 3.7 below.

3.3. Dataset Split Strategy

This study splits all of the samples into training, validation, and test samples based on two dimensions: spatial and temporal. First, the street segments in the study area are divided into three subsets at a ratio of 7:1.5:1.5 based on a random process that prioritizes spatial evenness. Specifically, to avoid spatial autocorrelation in the resulting dataset splits, we performed sampling by maximizing the physical distance between segments[2]. Thereby, the number of observations in training, validation, and test datasets are 3,161, 678, and 678 street segments, respectively, and within each subset, segments are spatially distributed evenly.

Second, this study takes into account the temporal dimension in the data split process to identify whether our crime prediction model trained from past data can be used to predict future crime. We set the dependent variables of the training, validation, and test datasets to crimes in 2017, 2018, and 2019, respectively. Namely, the training set is 3,161 street segments with the crime in 2017 as a dependent variable, and the validation and test sets are 678 street segments with the crime in 2018 and 2019, respectively. Thus, different street segments are in each sample. The independent variables for all subsets are the neighborhood environment and socio-economic characteristics of street segments at a point in time. This partitioning strategy allows us to verify the potential of our model to predict future crimes, which can provide implications for crime-prevention practices.

3.4. Measuring Streetscape Features

3.4.1. Google Street View

In this study, street images were obtained along street segments to audit the streetscape features. We utilized the Google Street View API to acquire the static images with a 640 x 640 size which is the maximum available size. As for the GSV acquisition criteria, previous studies have used various intervals such as 20, 50, and 100 m between images (Kim et al., 2021). Because the average length of a road segment in Santa Ana is 136 m, we acquired images on a 20 m interval in order to measure representative streetscape features (n=28,257). We only collected GSV images taken between 2017 and 2020 to secure enough images for each street segment as 94.5% of the available images were from this period. Additionally, GSV images

[2] To conduct spatially even sampling, the SpatialBlock package in R is applied (https://rdrr.io/cran/blockCV/man/spatialBlock.html)

near intersections were excluded, as they contain the features of more than one segment. As a result, this study obtained GSV images from 26,703 points.

We acquired four images (front, right, back, left) for each location to capture the environment in all directions (see Figure 3). That is, 106,812 images of 26,703 points were used for analysis. This study used 4,517 segments out of 5,343 segments, excluding segments where GSV images do not exist, and where images were taken before 2017 and after 2020. Thus, we utilized an average of 5.91 GSV points and 23.65 images per segment, respectively.

3.4.2. Semantic Segmentation

To quantify the streetscape features through the acquired GSV images, this study utilized semantic segmentation, a computer vision technique. Semantic segmentation can classify each pixel as an image component, such as greenery, vehicle, or building. Recently, several studies have used this approach to extract the neighborhood environment from GSV images (Gong et al., 2018; Ki and Lee, 2021).

Among the semantic segmentation models, the present study employed Deeplabv3+ (Chen et al., 2018), which is based on the deep convolutional neural network (DCNN). This model was pre-trained with the Cityscapes dataset containing a collection of streetscape images similar to GSV images (Cordts et al., 2016). Using this pre-trained semantic segmentation model, we identified 13 streetscape elements from the imagery: buildings, humans, sidewalks, pavement, vehicles, fences, walls, vegetation, terrain, sky, traffic lights, traffic signs, and poles (refer to Figure 3). Among these built environment features, we narrowed our focus to a subset expected to be linked with crime patterns. Specifically, we re-categorized these variables, emphasizing six features representing dimensions crucial to crime ecology theory: vibrancy (buildings, humans), greenery (vegetation, terrain), and defensible space (walls, fences). Streetscape elements like buildings and humans align with Jacobs's 'eyes on the street' theory (Jacobs, 1961). Conversely, variables such as walls, fences, vegetation, and terrain are associated with Newman's 'defensible space' theory (Newman, 1972). Aligning with theoretical concepts linking streetscape features to crime, this study selected key features most likely associated with criminal activities. As depicted in Figure 3, we calculated the ratio of elements within each street view image. Feature variables for each segment were constructed by averaging the elements present in the segment's images.

Figure 3. Example of Google Street View images and results of semantic segmentation

3.5. Control Variables

To determine the relationship between the street-level built environment and crime patterns, this study controlled for the effects of crime generators and other potential crime determinants. These were constructed as a ½ mile buffer with an exponential decay ($\beta = -0.5$) centered on the focal segment. Whereas some measures from the Census are in blocks, for others in block groups we needed to impute them from block groups to blocks. Therefore, 2010 blocks and the American Community Survey (ACS) 5-year estimates for 2008-2012 are used as this is the Census data available to us despite the time gaps with the other data sources.

First, we considered measures of the ambient population, which includes residential population, total employees, retail/food employees, and percent aged 16 to 29 (to capture the possible presence of offenders). Although these Census variables have limitations in directly reflecting the ambient population and vibrancy (He et al., 2020), they were used as proxy variables. Furthermore, the vibrancy variables measured through GSV (specifically humans and buildings) can complement the limitations of these Census variables.

Given that social disorganization theory posits that disadvantaged neighborhoods will have less informal social control capability, and therefore more crime, we included several measures to capture this theory (Sampson and Groves 1989). Concentrated disadvantage was measured as a factor score from a principle factor analysis of percent at or below 125% of the poverty level, average household income, percent with at least a bachelor's degree, and percent single-parent households. Residential stability was calculated as a factor score combining percent owners, percent in the same house five years ago, and average length of residence. Racial/ethnic heterogeneity measures racial diversity, and was measured as the Herfindahl index based on five ethnic categories (White, Black, Asian, Latino, others). To capture the presence of racially disadvantaged groups, we included the percentage of the population that was Latino, Asian, or Black. Given that vacant units can create both physical and social disorder in the neighborhood and increase crime, we created a measure of percent vacant units. Numerous studies have shown that these measures are consistent predictors of crime (Kubrin and Weitzer 2003).

Finally, studies have shown that it is important to account for length of the street segment to account for crime opportunities (Kim and Hipp, 2017). Longer street segments provide more crime opportunities, and therefore including this in the model translates the outcome to a measure of crime density.

3.6. Machine Learning Model

To predict the number of crimes, we utilized the following five machine learning models (i.e., Random Forest, Support Vector Machine, XGBoost, Artificial Neural Network, and Deep Neural Network). We used these five techniques given that prior research has shown that they are typically among the most effective models in studies that compare different machine learning strategies. The explanation of models and error indicators to compare the models' performances were described in detail in Appendix A. Briefly, four major error indicators were used to evaluate the models and thus select a final model. This study also compared these models with the conventional statistical model, negative binomial (NB) regression, to assess whether the machine learning models outperform it.

3.7. Interpretable Machine Learning (IML)

As mentioned above, machine learning has the advantage of prediction power, but a limitation is the lack of information about how specific measures in the model are related to the outcome measure. Recently, work on interpretable machine learning (IML) has attempted to decode the "black box" architecture of earlier neural networks (Molnar, 2020). The SHAP method can allow one to interpret the model in both global and local senses (García and Aznarte, 2020; Molnar, 2020). Given this advantage, an increasing number of studies have employed this approach in various fields, including criminology (García and Aznarte, 2020; Wheeler and Steenbeek, 2020).

The SHAP algorithm is based on coalitional game theory that is, how much each individual contributed to collaborative outcomes (Shapley, 1953). SHAP calculates the feature importance (Shapley value) by comparing the change in outcomes depending on the presence or absence of each variable. As mentioned earlier, SHAP enables not only global interpretation but also local interpretation. To be more specific, the global interpretation determines the overall influence (global Shapley value) and direction that each independent variable has on the estimated outcomes. Local interpretation yields a Shapley value for each observation in the sample, which makes it possible to investigate the nonlinear relationships of the variables (Lundberg et al., 2020).

4. Analysis Results

4.1. Comparative Evaluation of Machine Learning Models

This study constructed crime prediction models by setting the dependent variable as the number of crimes and the independent variables as the built environment and the four groups of other determinants, as described above. All independent variables were normalized using min-max normalization due to the feature scales issue (Lin et al., 2018). The model specification and hyperparameter tuning of the machine learning model were delineated in detail in Appendix B.

The error of each model after the hyperparameter tuning is shown in <Table 1>. As hypothesized, all machine learning models overall showed higher accuracy than the traditional statistical method (i.e., NB). In particular, the DNN showed the lowest error across each of the four indicators, and this result is consistent with Lin et al. (2018)'s study comparing various machine learning model performances, including the DNN. Additionally, it was found that the MSE value of the NB model was much higher than those of the machine learning models, indicating that its prediction was far off for some street segments. This highlights that interpreting the machine learning model, especially a DNN, can enable us to more completely understand the relationship between crime and the neighborhood environment than with the conventional statistical approach.

Table 1. Evaluation of models' errors

4.2. Interpreting the DNN using the SHAP Algorithm

In this section, we utilized the SHAP approach for the DNN that showed the highest accuracy in the previous step. As mentioned earlier, the SHAP algorithm can decode the machine learning models in two respects: global and local interpretation. Through global interpretation, we can verify the overall impacts of each variable on the outcome. Local interpretation enables us to detect nonlinear relationships between features.

4.2.1. Global Model Interpretation

<Figure 4> shows the global Shapley value for each variable. The size of each bar represents the contribution of that independent variable to the predicted value. The color can be interpreted as the direction of the variable; red is a positive and blue is a negative relationship with the dependent variable.

Figure 4. Global Shapley value

For instance, the segment length has the greatest impact on the number of crimes, and it can be interpreted that crimes increase as the segment length increases [3]. This result is not surprising as we included segment length in our analysis model to capture the presence of more crime opportunities when assessing the impact of the built environment on crime incidents. In the next section, we describe the influence of these variables in more depth, paying special attention to their nonlinear relationships with crime.

4.2.2. Local Model Interpretation

Given that the purpose of this study is to investigate the non-linear relationship between streetscape variables and crime, local interpretation enables us to further explore it. As mentioned above, the local interpretation provides the Shapley value for each sample in the test dataset.

Figure 5. Local interpretation and examples for defensible space

In <Figure 5- A>, the x-axis and y-axis represent the ratio of walls and the Shapley value, respectively. Each dot represents each sample in the test dataset (n=678), and the blue line is the locally weighted scatterplot smoother (LOWESS) line based on these points. Note that this figure plots the *derivatives* of this relationship. Therefore, for a more intuitive understanding, <Figure 5-A`> plots the relationship between the dependent variable (number of crimes) and the ratio of the wall variables itself. This is a more conventional non-linear plot

[3] Please note that our dependent variable is the number of crimes per segment, which is not normalized.

that readers will be familiar with and is based on the formula from the LOWESS graph. In other words, in <Figure 5-A`>, the y-axis is the number of crimes, and the x-axis is the proportion of walls in the street segment.

In the case of the wall, the global Shapley value is -0.176, but the local value is widely distributed from -1.747 to 0.858. More specifically, walls have a very modest positive relationship with the number of crimes at low levels (wall<0.1). Beyond a threshold (wall=0.1), however, its relationship turns negative and becomes stronger, as seen in <Figure 5-A`>. Although walls show a non-linearity with the crime occurrences, it is worth mentioning the validity of the values before the threshold. Given that the threshold is close to 0 and the Shapley values of the observations located in the positive section are small, the positive effect of walls is negligible. <Figure 5, A`> showing the relationship with the dependent variable displays this immaterial positive effect.

Similarly, the Shapley value of fences becomes stronger in the negative direction as the fence proportion increases (Figure 5 - B, B`). Fences and walls, which are related to the notion of defensible space and CPTED, can prevent crime by restricting the access of potential offenders and reinforcing territoriality (Cozens and Love, 2015; Langton and Steenbeek, 2017). Additionally, the likelihood of preventing access increases as the two features become more prevalent, and thus crimes rapidly decrease.

The two elements related to urban greenery, vegetation and terrain, have different relationships with crime patterns. The vegetation measure has a robust U-shape relationship with crime (see Figure 6- C, C`). Thus, most crime occurs in segments with either very low or very high concentrations of vegetation. On the contrary, the Shapley value for terrain changes from positive to negative as the percentage of terrain increases (see Figure 6- D, D`), which means that terrain exhibited an inverted U-shape relationship with crime. In terms of specific greenery, vegetation includes vertical vegetation, such as trees, and hedges, whereas terrain includes horizontal vegetation, such as grassy areas, and soil[4]. In other words, both elements can act as attractive places to gather people, but if the vegetation is too dense at a location, it

[4] Cityscapes dataset labeling policy and class definitions: https://www.cityscapes-dataset.com/dataset-overview/#labeling-policy

may give perpetrators a chance to hide from the view of guardians (Wolfe and Mennis, 2012).

Figure 6. Local interpretation and examples for urban greenery

Additionally, dense vegetation creates a visually closed environment, thereby contributing to a high level of fear (Baran et al., 2018). However, since terrain is an open green space, it cannot provide offenders with opportunities to hide, even if there is a large area of terrain in a specific place.

Finally, GSV variables that are related to vibrancy – buildings and humans – show a similar relationship with crime (see Figure 7- E, F). They exhibit a relatively robust positive relationship with crime, and the impact becomes stronger as their presence increases. These results may be attributable to the fact that vibrant places have more targets and offenders, thereby creating more crime opportunities (see Figure 7). Also, unlike the Census variables, the GSV human variable also includes non-local residents. Thus, many non-local residents may lead to a loss of informal social control, creating more crime opportunities (He et al., 2020). For these reasons, several studies have reported a positive relationship between vibrancy and crime (He et al., 2020; Hipp et al., 2019).

Similar to walls, the validity of the values before the threshold should be noted. For both vibrancy variables, the Shapley values are negative at low levels (building < 4.7 and humans < 0.07), but the threshold is close to zero and the Shapley value in this section is also very small. As shown in <Figure 7 – e`, f`>, the crime deterrence effect of the vibrancy variables can be seen as trivial in this range.

It is important to note that the human variable in the scene is rare and dynamic (Kim et al., 2021), which indicates that the presence of humans is likely to fluctuate depending on the time when the GSV was collected. In this regard, using street images to directly capture vibrancy may not perfectly capture the concept of interest. Notwithstanding, this approach has been adopted in several previous studies (Chen et al., 2020; Yue et al., 2022). They noted its potential and reliability compared to an on-site pedestrian count survey (Chen et al., 2020). Furthermore, Yue et al. (2022) theoretically underlined the advantages of this novel approach in terms of availability and reliability compared with other methods for measuring vibrancy, such as using social media or mobile phone data, in the criminology field. On the one hand, we operationalized the ambient population with the buildings as well as humans, which is a more reliable and less dynamic variable.

Figure 7. Local interpretation and examples for vibrancy

5. Discussion and Conclusion

This study attempted to reveal the association between streetscape features and crime patterns at the street segment level using GSV and several machine learning techniques. The use of street images and semantic segmentation with deep learning techniques enables us to quantify the micro-level built environment features that cannot be easily measured with conventional data collection methods. This approach is particularly promising for large areas, allowing one to capture the neighborhood environment at the micro-scale level in an efficient manner. In addition, the micro-level of the street segment unit is most likely to reduce spatial heterogeneity and the ecological fallacy in criminology studies (Groff et al., 2010; Hipp et al., 2022). Given that few criminology studies have employed this method despite its advantages, more attention should be paid to the possibility of using it to further investigate the relationship between fine-scale environmental features and crime patterns.

Many studies utilizing traditional regression models to explore the relationship between urban environmental elements and crime incidence have yielded mixed outcomes (Hipp et al., 2022; Lee and Contreras, 2021). The relationships between neighborhood environments and crime are ambiguous due to the possible existence of non-linearity between them in the criminology literature (Walker, 2007; Browning et al., 2010; Hipp et al., 2019; Wheeler and Steenbeek, 2020). Although non-linearity analysis between neighborhood environments and crime patterns is possible in the conventional regression models with polynomial functions, the applications of machine learning models with interpretability techniques are very efficient to detect possible nonlinear relationships and threshold effects between built environment features and crime. Tao et al. (2020) have highlighted the flexibility of machine learning models in analyzing urban environmental nonlinearity. While the machine learning models, especially DNN, outperformed negative binomial regression in predicting future crime events, DNN also outperformed other machine learning models such as RF, SVM, XGBoost, and ANN. By using the SHAP algorithm, we were able to interpret the non-linear relationships between variables and crime yielded by the DNN model. This approach is particularly advantageous because it allowed us to obtain a deeper understanding of the complex relationships between crime patterns and environmental factors.

The non-linearities revealed in this study can shed light on the factors associated with crime patterns. Additionally, it is important to note that the IML methodology gives credibility to the model and makes it possible to formulate a problem (Ribeiro et al., 2016). For instance, our results indicated that walls and fences, in line with the defensible space theory (Newman, 1972) and CPTED (Jeffery, 1971), effectively deterred criminal activities in the studied area. Yet, it is essential to highlight that the connection between walls and crime demonstrates a threshold effect, displaying a positive association below the threshold. In general, as these two factors surpass certain thresholds, they exhibit increased crime deterrence effects, potentially due to their ability to obstruct the access of potential offenders (Langton and Steenbeek, 2017).

Our findings indicated that vegetation and terrain variables exhibited more evident nonlinear patterns and threshold effects. We observed a robust U-shaped relationship between vegetation and crime, while terrain showed an inverted U-shaped relationship with crime patterns. These differences could be attributed to the inclusion of various types of greenery. As mentioned earlier, vegetation demonstrated a dual relationship with crime. It can encourage crime by concealing perpetrators and criminal activities, increasing potential targets (Wolfe and Mennis, 2012). On the other hand, it can act as a gathering place, enabling natural surveillance that deters crime (Kondo et al., 2016; Troy et al., 2016). The nonlinear link between vegetation and crime in Santa Ana might reflect these conflicting influences. In essence, low levels of vegetation may be negatively associated with crime by enhancing human presence, supporting Jacob's (1961) theory of natural surveillance by "the eyes on the street." However, beyond a certain threshold, vegetation forms a visually enclosed landscape, offering hiding spots for offenders and leading to increased crime.

We explored the link between crime and GSV variables related to urban vibrancy. In our study area, the association between vibrancy and crime initially showed a negative correlation below a specific threshold, shifting to a positive connection beyond that point. Additionally, we observed that this positive association was more pronounced with a higher percentage of these vibrancy indicators. Vibrant areas suggest numerous potential guardians but can also present a higher number of potential targets. Notably, the inability to distinguish between local and non-local residents within the GSV-based human variable might imply that larger numbers of non-local residents could potentially contribute to increased crime (He et al., 2020). Considering these perspectives, our study's findings align with research demonstrating a positive relationship between vibrancy and crime patterns (He et al., 2020; Hipp et al., 2019; Hipp et al., 2022). However, it's crucial to note the negative associations between urban vibrancy indicators and crime events below the threshold. A certain level of vibrant places with potential guardians appears to act as a deterrent against criminal activities.

This study also has a few limitations. First, GSV images were collected by vehicles, which differs slightly from a pedestrian perspective. This is an inherent limitation of street images because it collects images over a large area. Second, we cannot ignore the time gap between the datasets used in this study. The disparity between street images and actual crime occurrences suggests potential differences in the independent variables measured and the neighborhood conditions when the crimes took place. Additionally, the study's use of block-level data sourced from the 2010 Census and ACS 5-year estimates for 2008-2012, despite time gaps with other sources, assumes a relatively stable population in Santa Ana. However, we acknowledge that this assumption may not fully account for differences in demographic and socioeconomic characteristics. In addition, as highlighted in a few studies (He et al., 2020; Lan et al., 2019), Census variables might not fully capture the ambient population, encompassing residents, employees, and visitors. To address this, we supplemented the Census variables with GSV variables like 'human' and 'building.' However, we acknowledge that these GSV variables also have constraints in measuring the ambient population.

Thirdly, despite being based on a concrete theory like CPTED, the study's reliance on a cross-sectional design cannot statistically guarantee the causal relationship between neighborhood environments and crime patterns. We included a set of “control variables” that are common in the criminology literature in an effort to account for possible confounders to the relationships we observed. Nonetheless, similar to existing cross-sectional studies we are limited in our causal inferences, and some elements of the built environment might show complex association patterns with crime as shown in this study not only because their presence can enable or prevent crime occurrence but also because they serve as a proxy for another factor that is difficult to measure and control for. Finally, due to the aggregation of all Part 1 crime types, this study cannot identify specific factors related to each crime type, potentially biasing the model toward predicting more prevalent crime types, such as burglary.

References

1. Baran PK, Tabrizian P, Zhai Y, et al. (2018) An exploratory study of perceived safety in a neighborhood park using immersive virtual environments. *Urban Forestry & Urban Greening*, 35, 72-81.
2. Brantingham PJ, Brantingham PL (1981) Environmental criminology. Beverly Hills, CA: Sage.
3. Brantingham PJ, Brantingham PL (1984) Patterns in crime. New York: Macmillan.
4. Breiman L (2001). Random forests. *Machine learning*, 45(1), 5-32.
5. Browning CR, Byron RA, Calder CA, et al. (2010) Commercial density, residential concentration, and crime: Land use patterns and violence in neighborhood context. *Journal of Research in Crime and Delinquency*, 47(3), 329-357.
6. Burley BA (2018) Green infrastructure and violence: Do new street trees mitigate violent crime? *Health & Place*, 54, 43-49.
7. Chen L, Lu Y, Sheng Q, et al. (2020) Estimating pedestrian volume using Street View images: A large-scale validation test. *Computers, Environment and Urban Systems,* 81, 101481.
8. Chen LC, Zhu Y, Papandreou G, et al. (2018) Encoder-decoder with atrous separable convolution for semantic image segmentation. *In Proceedings of the European Conference on Computer Vision*, 801-818.
9. Chen, T., Guestrin, C (2016) Xgboost: A scalable tree boosting system. *In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.
10. Cohen LE, Felson M (1979) Social change and crime rate trends: A routine activity approach. *American Sociological Review*, 588-608.
11. Cordts M, Omran M, Ramos S, et al. (2016) The cityscapes dataset for semantic urban scene understanding. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3213-3223.
12. Cozens P. (2014) Think crime! Using evidence, theory and crime prevention through environmental design (CPTED) for planning safer cities. Quinns Rock Perth, WA: Praxis Education.
13. Cozens P, Love T (2015) A review and current status of crime prevention through environmental design (CPTED). *Journal of Planning Literature*, 30(4), 393-412.
14. Das S, Tsapakis I (2020) Interpretable machine learning approach in estimating traffic volume on low-volume roadways. *International Journal of Transportation Science and Technology*, 9(1), 76-88.
15. García MV, Aznarte JL (2020) Shapley additive explanations for NO2 forecasting. *Ecological Informatics*, 56, 101039.

16. Gong FY, Zeng ZC, Zhang F, et al. (2018) Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Building and Environment*, 134, 155-167.

17. Groff E, McCord ES (2012) The role of neighborhood parks as crime generators. *Security Journal*, 25(1), 1-24.

18. Groff ER, Weisburd D, Yang SM (2010) Is it important to examine crime trends at a local “micro” level? A longitudinal analysis of street to street variability in crime trajectories. *Journal of Quantitative Criminology*, 26(1), 7-32.

19. Gupta N (2013) Artificial neural network. *Network and Complex Systems*, 3(1), 24-28.

20. Hakim S, Rengert G., Shachamurove Y (2001) Target search of burglars: A revisited economic model, *Papers in Regional Science*, 80, 121–137.

21. He L, Páez A, Jiao J, et al. (2020) Ambient population and larceny-theft: A spatial analysis using mobile phone data. *ISPRS International Journal of Geo-Information*, 9(6), 342.

22. He L, Páez A, Liu D (2017) Built environment and violent crime: An environmental audit approach using Google street view. *Computers, Environment and Urban Systems*, 66, 83-95.

23. Hipp JR, Bates C, Lichman M, et al. (2019) Using social media to measure temporal ambient population: Does it help explain local crime rates? *Justice Quarterly*, 36(4), 718-748.

24. Hipp, JR, Lee, S, Ki, D, Kim, J H (2022) Measuring the built environment with Google street view and machine learning: Consequences for crime on street segments. *Journal of Quantitative Criminology*, 537-565.

25. Jacobs J (1961) The death and life of great American cities. New York: Random House.

26. Jeffery CR (1971) Crime prevention through environmental design. Beverly Hills, CA: Sage Publications.

27. Kang HW, Kang HB (2017) Prediction of crime occurrence from multi-modal data using deep learning. *PloS one*, 12(4), e0176244.

28. Ki D, Lee S (2021) Analyzing the effects of Green View Index of neighborhood streets on walking time using Google Street View and deep learning. *Landscape and Urban Planning*, 205, 103920.

29. Kim JH, Lee S, Hipp JR, Ki D (2021) Decoding urban landscapes: Google street view and measurement sensitivity. *Computers, Environment and Urban Systems*, 88, 101626.

30. Kim YA, Hipp JR (2017) Physical boundaries and city boundaries: Consequences for crime patterns on street segments?". *Crime & Delinquency,* 64(2):227-54

31. Kondo M, Hohl B, Han S, et al. (2016) Effects of greening and community reuse of vacant lots on crime. *Urban Studies*, 53(15), 3279-3295.

32. Kubrin CE, Weitzer R (2003). New directions in social disorganization theory. *Journal of Research in Crime and Delinquency,* 40(4):374-402.

33. Lan, M, Liu, L, Hernandez, A, Liu, W, Zhou, H, Wang, Z. (2019) The Spillover effect of geotagged tweets as a measure of ambient population for theft crime. *Sustainability*, 11, 6748.

34. Langton SH, Steenbeek W. (2017) Residential burglary target selection: An analysis at the property-level using Google street view. *Applied Geography*, 86, 292-299.

35. Lee, N, Contreras, C (2021) Neighborhood walkability and crime: Does the relationship vary by crime type?. *Environment and Behavior*, *53*(7), 753-786.

36. Lin YL, Yen MF, Yu LC (2018) Grid-based crime prediction using geographical features. *ISPRS International Journal of Geo-Information*, 7(8), 298.

37. Lu Y (2018) Using Google Street View to investigate the association between street greenery and physical activity, *Landscape and Urban Planning*, 191, 1-9.

38. Lundberg SM, Erion G, Chen H, et al. (2020) From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56-67.

39. Marco M, Gracia E, Martín-Fernández M, et al. (2017) Validation of a Google street view-based neighborhood disorder observational scale. *Journal of Urban Health*, 94(2), 190-198.

40. Mohler, GO, Short, MB, Brantingham, PJ, Schoenberg, FP, Tita, GE (2011) Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493), 100-108.

41. Molnar C (2020) Interpretable Machine Learning. Lulu. com.

42. Newman O (1972) Defensible space: people and design in the violent city. London: Architectural Press.

43. Pratt, TC, Cullen FT (2005) Assessing macro-level predictors and theories of crime: A meta-analysis. *Crime and Justice* 32:373–450.

44. Ribeiro MT, Singh S, Guestrin C (2016) “Why should I trust you?" Explaining the predictions of any classifier. *In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135-1144.

45. Rummens A, Hardyns W, Pauwels L (2017) The use of predictive analysis in spatiotemporal crime forecasting: Building and testing a model in an urban context. *Applied Geography*, 86, 255-261.

46. Sampson RJ, Groves WB (1989) Community Structure and Crime: Testing Social-Disorganization Theory. *American Journal of Sociology* 94 (4):774-802.

47. Shapley LS (1953) A value for n-person games. Contributions to the theory of games, 2(28), 307-317.

48. Sherman LW, Gartin P, Buerger ME (1989) Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*. 27, 27–55.

49. Srivastava N, Hinton G, Krizhevsky A, et al. (2014) Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929-1958.

50. Tao, T, Wu, X, Cao, J, Fan, Y, Das, K, Ramaswami, A (2023) Exploring the nonlinear relationship between the built environment and active travel in the twin cities. *Journal of Planning Education and Research*, 43(3), 637-652.

51. Taylor M, Nee C (1988) The role of cues in simulated residential burglary, *British Journal of Criminology*, 28, 396–401.

52. Town S, Davey C, Wooton A (2003) Design against crime: Secure urban environments by design, Salford: The University of Salford.

53. Troy A, Nunery A, Grove JM (2016) The relationship between residential yard management and neighborhood crime: An analysis from Baltimore City and County. *Landscape and Urban Planning*, 147, 78-87.

54. Vandeviver C (2014) Applying google maps and Google street view in criminological research. *Crime Science*, 3(1), 1-16.

55. Walker JT (2007) Advancing science and research in criminal justice/criminology: Complex systems theory and non-linear analyses. *Justice Quarterly*, 24(4), 555-581.

56. Weisburd D, Bushway S, Lum C, Yang SM (2004) Trajectories of crime at places: A longitudinal study of street segments in the city of Seattle. Criminology, 42(2), 283-322.

57. Wheeler AP, Steenbeek W (2020) Mapping the risk terrain for crime using machine learning. *Journal of Quantitative Criminology*, 1-36.

58. Wolfe MK, Mennis J (2012) Does vegetation encourage or suppress urban crime? Evidence from Philadelphia, PA. *Landscape and Urban Planning*, 108(2-4), 112-122.

59. Wortley R, Mazerolle L (2008) Environmental criminology and crime analysis: Situating the theory, analytic approach and application. In R. Wortley & L. Mazerolle (eds). Environmental criminology and crime analysis. Cullompton, UK: Willan.

60. Yue H, Xie H, Liu L, et al. (2022) Detecting People on the Street and the Streetscape Physical Environment from Baidu Street View Images and Their Effects on Community-Level Street Crime in a Chinese City. *ISPRS International Journal of Geo-Information*, 11(3), 151.

61. Zhang Z, Beck MW, Winkler DA, et al. (2018) Opening the black box of neural networks: Methods for interpreting neural network models in clinical applications. *Annals of Translational Medicine*, 6(11).

62. Zhang X, Liu L, Lan M, et al. (2022) Interpretable machine learning models for crime prediction. *Computers, Environment and Urban Systems*, 94(101789), 1-9.

Tables

Table 2. Evaluation of models' errors

Accuracy	NB	RF	SVM	XGBoost	ANN	DNN
MAE	2.900	1.954	2.598	1.785	1.844	1.559
MSE	550.562	56.747	61.306	56.307	56.940	26.446
RMSLE	0.801	0.734	0.974	0.670	0.764	0.601
R^2	0.092	0.139	0.105	0.122	0.133	0.622

Note: NB (negative binomial); RF (Random Forest); SVM (Support Vector Machine); XGBoost (Extreme Gradient Boosting); ANN (Artificial Neural Network); DNN (Deep Neural Network)

Figures

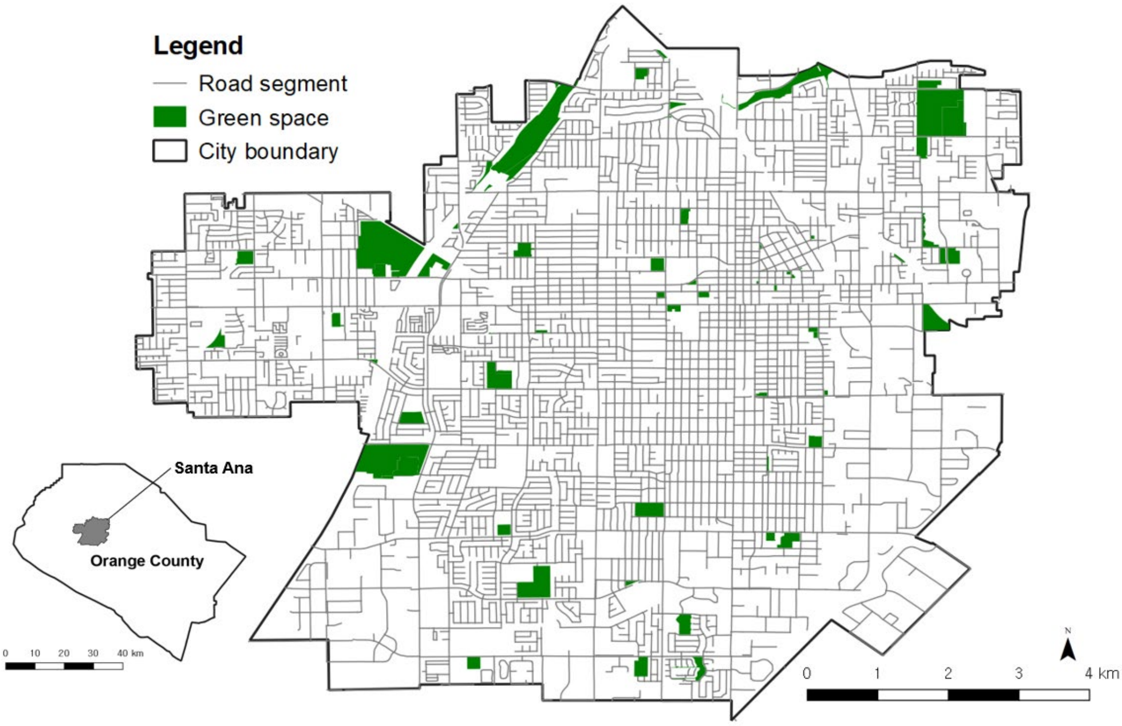


Figure 1. Location of Santa Ana

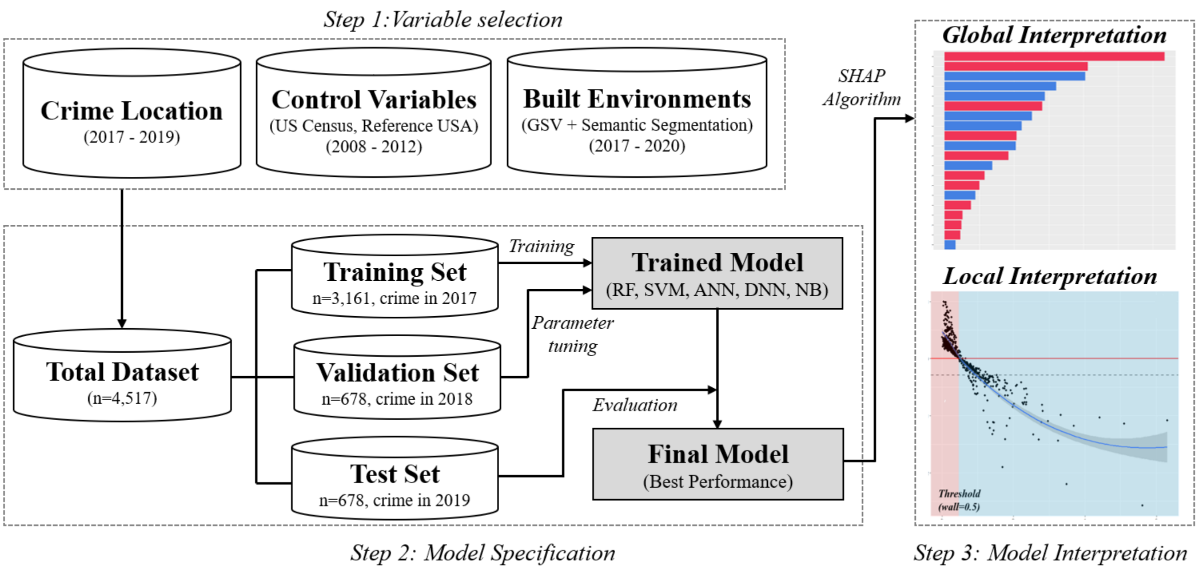
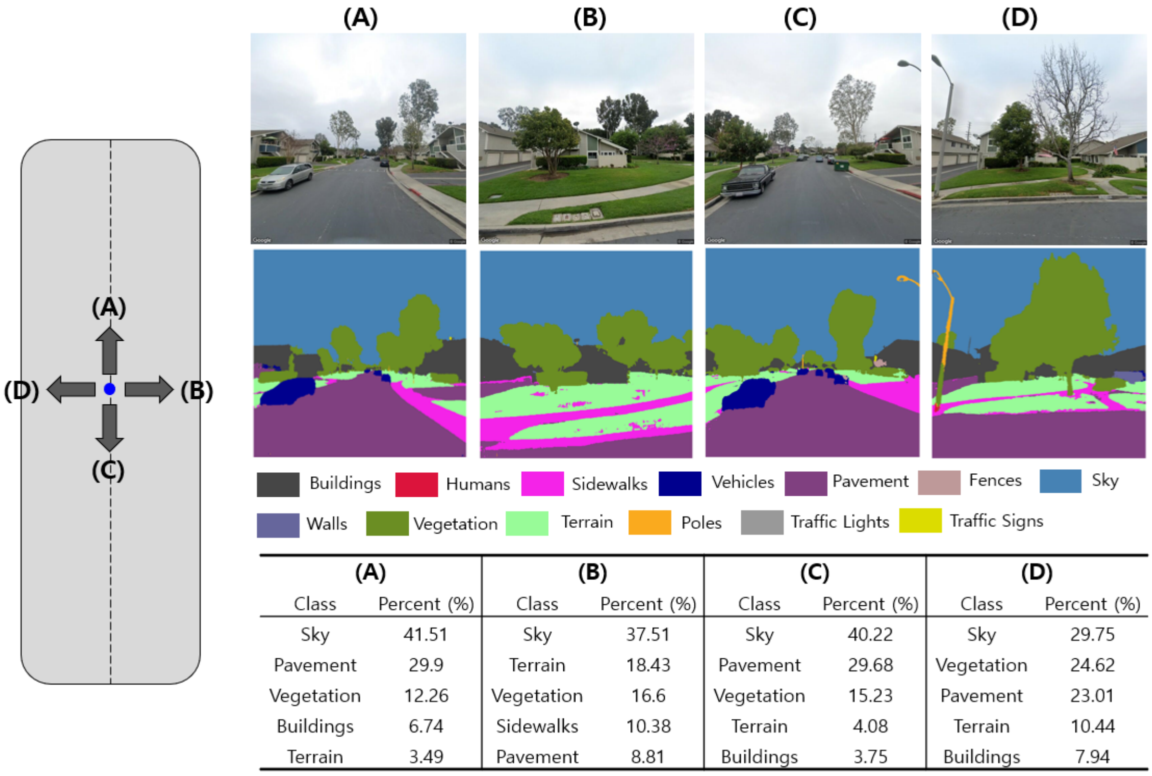


Figure 2. Research framework



(A)		(B)		(C)		(D)	
Class	Percent (%)	Class	Percent (%)	Class	Percent (%)	Class	Percent (%)
Sky	41.51	Sky	37.51	Sky	40.22	Sky	29.75
Pavement	29.9	Terrain	18.43	Pavement	29.68	Vegetation	24.62
Vegetation	12.26	Vegetation	16.6	Vegetation	15.23	Pavement	23.01
Buildings	6.74	Sidewalks	10.38	Terrain	4.08	Terrain	10.44
Terrain	3.49	Pavement	8.81	Buildings	3.75	Buildings	7.94

Figure 3. Example of Google Street View images and results of semantic segmentation

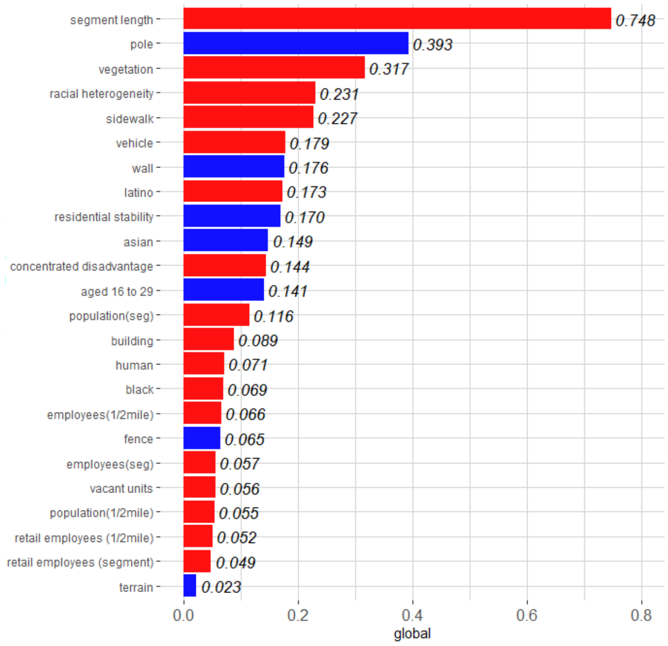


Figure 4. Global Shapley value

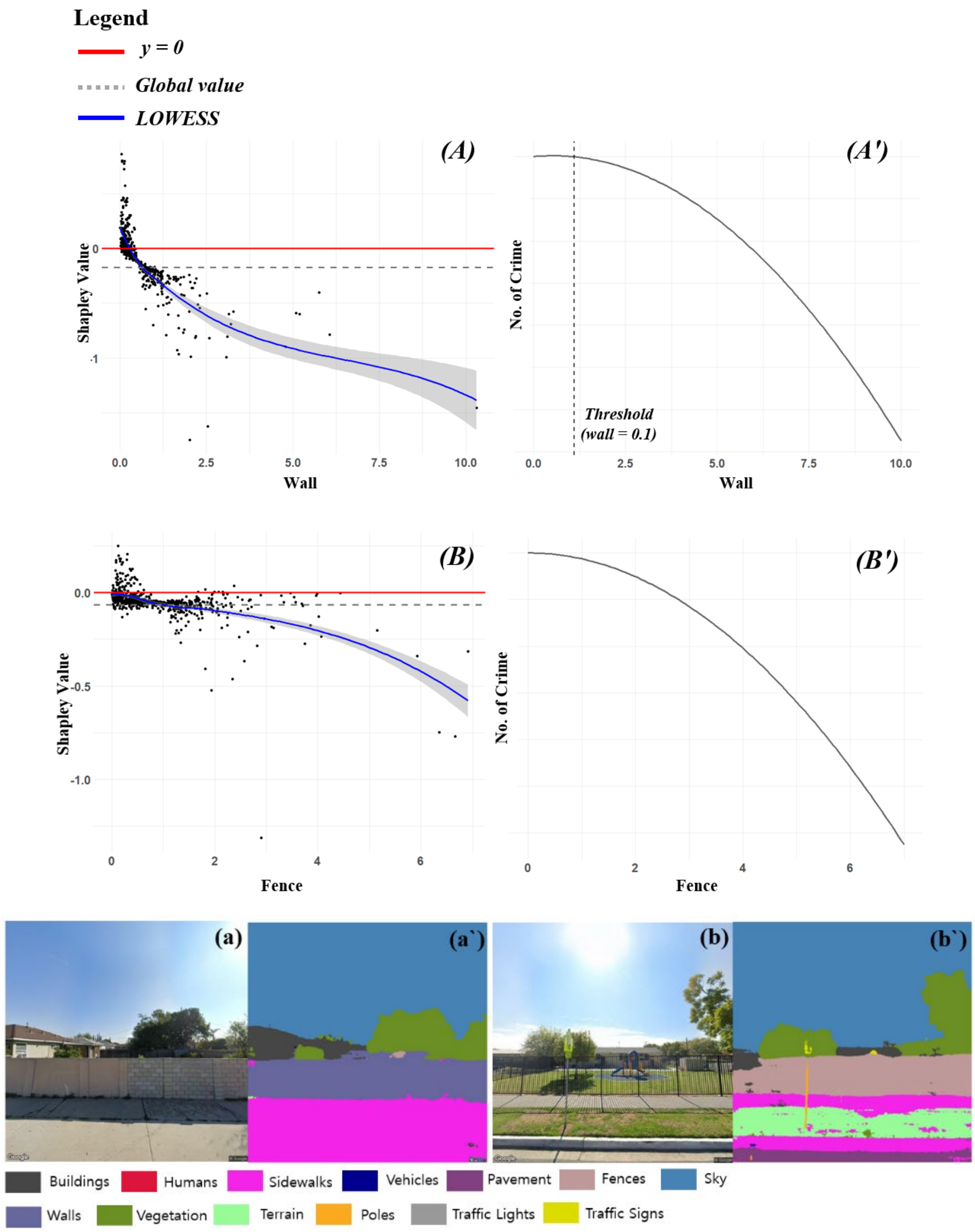


Figure 5. Local interpretation and examples for defensible space

(a, a`: high ratio of walls- 17.45% ; b, b`: high ratio of fences- 15.29%)

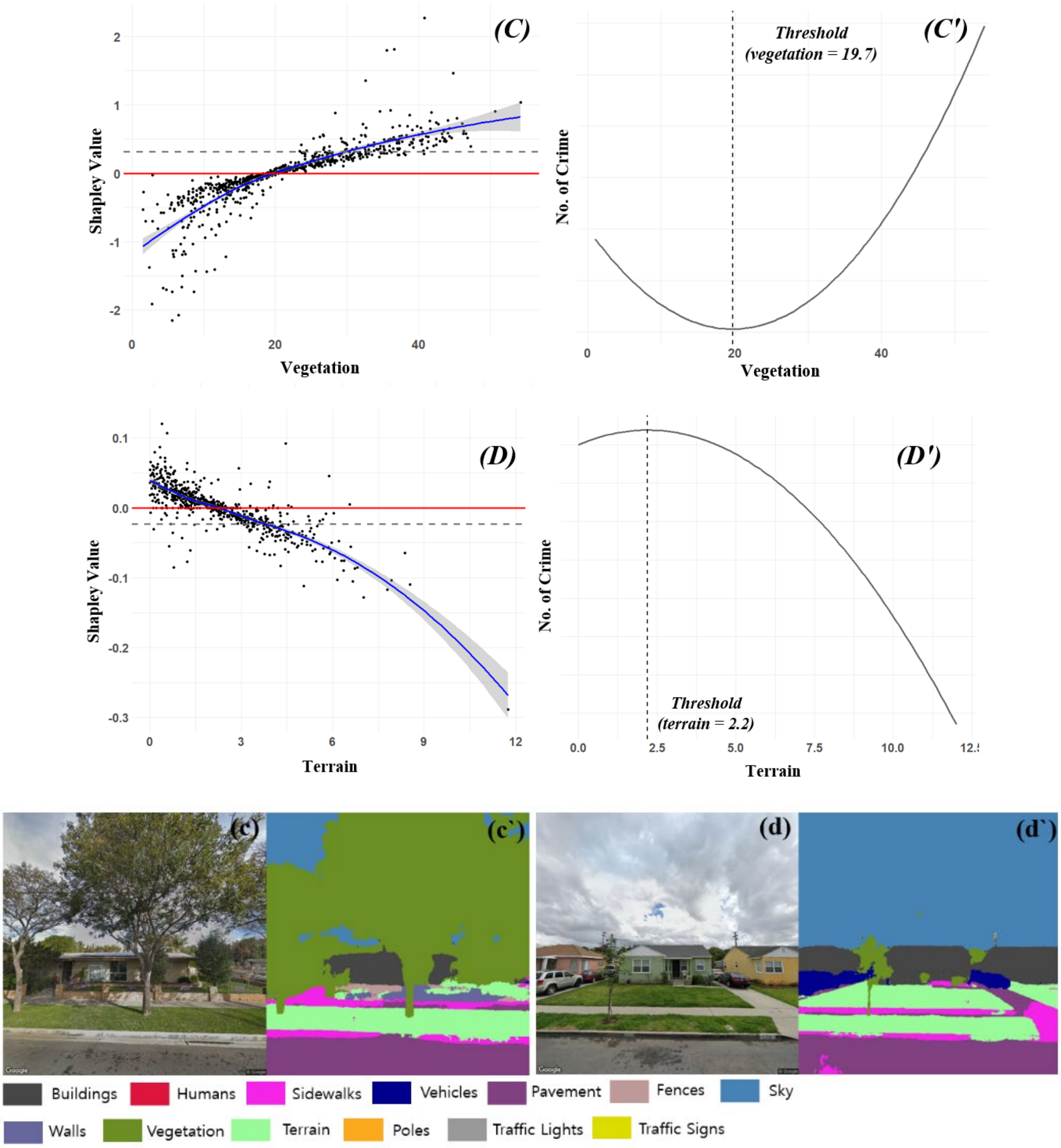


Figure 6. Local interpretation and examples for urban greenery

(c, c`: high ratio of vegetation- 56.36%; d, d`: high ratio of terrain- 13.74%)

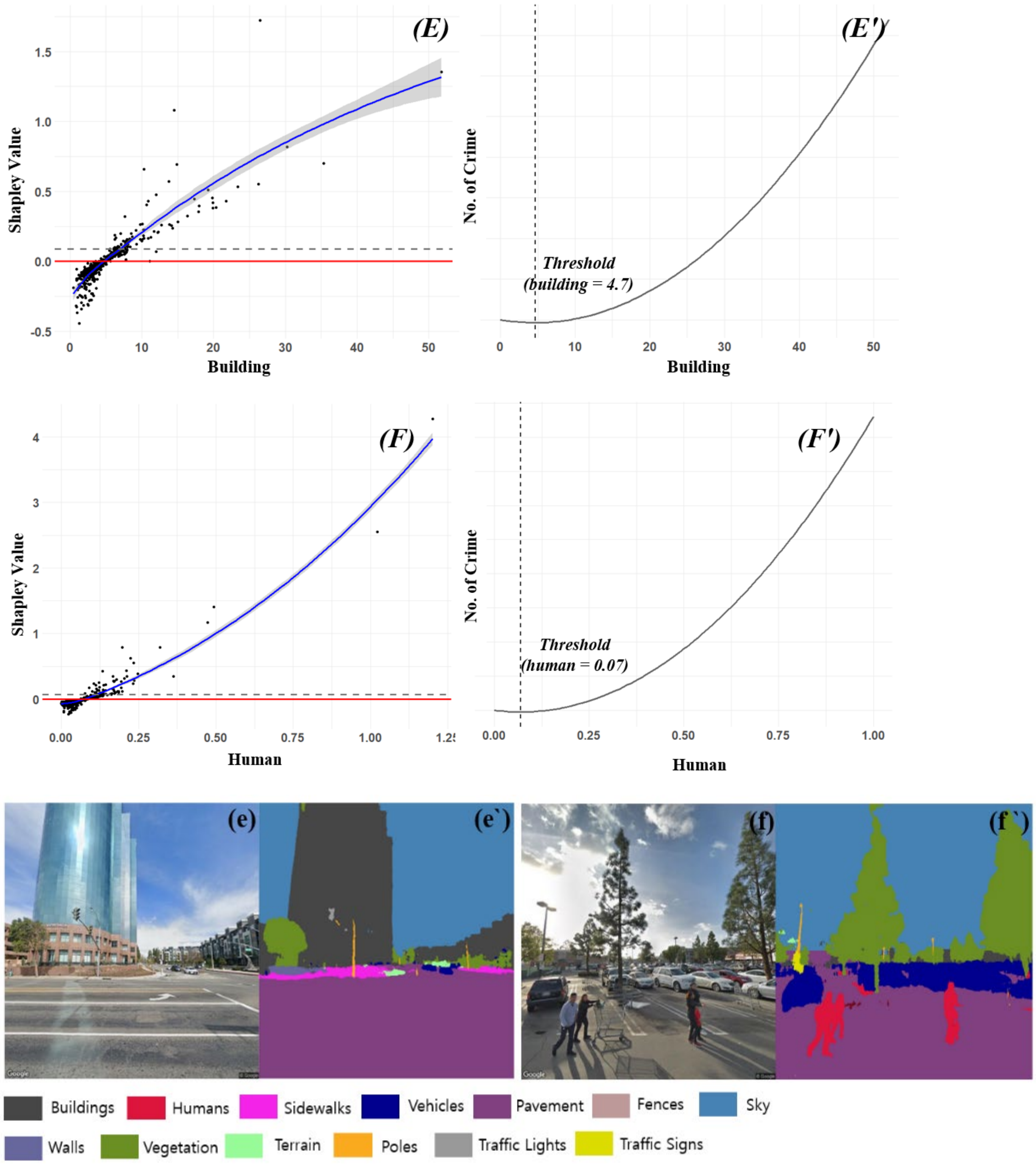


Figure 7. Local interpretation and examples for vibrancy

(e, e`: high ratio of building- 25.88%, f, f`: high ratio of human- 2.95%)

Appendix A: Machine learning models and evaluations

This study adopted five machine learning models for crime prediction.

- Random Forest (RF): A RF is a traditional ensemble model that combines various decision tree models. It enables more accurate and stable prediction than the decision tree model (Breiman, 2001).
- eXtreme Gradient Boosting (XGBoost): The XGBoost model is a tree-based ensemble model that is similar to Random Forest. The main difference is the use of gradient boosting instead of bagging. It has the advantages of fast training speed based on parallel tree learning and reliable accuracy for various tasks (Chen & Guestrin, 2016)
- Support Vector Machine (SVM): This algorithm is one of the supervised machine learning models used for classification, regression, and outlier detection. In this study, support vector regression (SVR) is used to predict the number of crimes (continuous variable).
- Artificial Neural Network (ANN): An ANN algorithm is based on the concept of a biological neural network. That is, it adjusts weights connecting various nodes through the training process and makes decisions based on the determined weights (Gupta, 2013).
- Deep Neural Network (DNN): A DNN has several (> 2) hidden layers, which is a model that extends the structure of the ANN. For this reason, it has a more complex architecture than the ANN model and is classified in the deep learning category, unlike the previous model.

We evaluated the performance of these models using four indicators: the mean absolute error (MAE), mean squared error (MSE), root mean square logarithmic error (RMSLE), and the coefficient of determination (R^2) (see the equations below). The MSE is one of the representative indicators in machine learning regression, which is the average squared error between the predicted and actual values. For this reason, it is highly sensitive to outliers. On the other hand, the RMSLE is robust to outliers and is calculated with a logarithmic scale. Except for the R^2, the lower the indicators, the more accurate the prediction.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} \quad \dots\dots \; eq(1)$$

$$MSE = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n} \quad \dots\dots \; eq(2)$$

$$RMSLE = \sqrt{\frac{\sum_{i=1}^{n} (\log(\hat{y}_i + 1) - \log(y_i + 1))^2}{n}} \quad \dots\dots \; eq(3)$$

$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2} \quad \dots\dots \; eq(4)$$

Where:

$\hat{y}_i$: Predicted number of crimes in the segment

y_i: Actual number of crimes in the segment

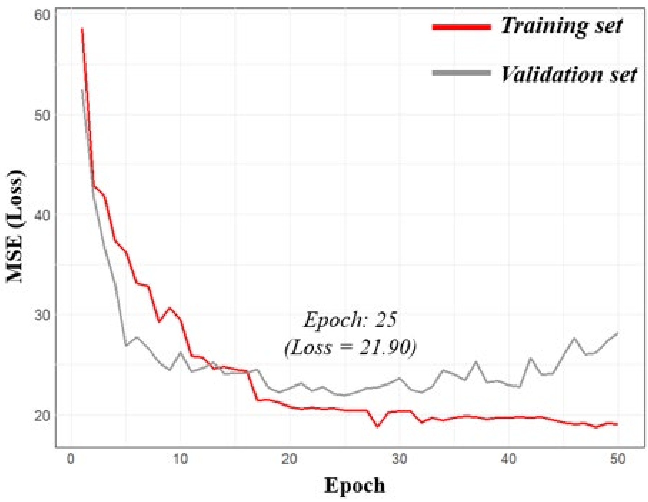
n: Number of street segments (sample size)

Appendix B: DNN hyperparameter tuning

The DNN model used here includes several hyperparameters, such as the number of hidden layers, the number of nodes for each layer, activation function, and dropout. A small number of hidden layers and nodes imply a simple architecture and can result in underfitting problems. On the contrary, when there are many elements, the model exhibits a complicated architecture, which can result in overfitting issues (Lin et al., 2018). To avoid the overfitting problem, we applied dropout to 0.2, which can prevent overfitting as 20% of nodes are randomly excluded at each layer (Srivastava et al., 2014).

In terms of model structure determinants, the number of nodes in the input layer is 26 (the same as the number of features), and the number of nodes in the output layer is 1 (in the regression model). Since there is no exact formula to determine the number of hidden layers and nodes (Karsoliya, 2012), this study found the optimal combination of hidden layers and nodes by evaluating the performance while changing them. In particular, given the amount of computing time and the number of observations in this study, this study explored the model's optimal parameters in the range of 1 to 5 hidden layers and 1 to 10 nodes for each layer. Through this process, the model's depth (number of hidden layers) was set to 4, and the hidden layers were established to 6, 4, 4, and 2 nodes, respectively. Additionally, this study set the activation function, learning rate, and epochs to ReLu, 0.01, and 25, respectively.

This study utilizes the Mean Squared Error (MSE) loss function of the DNN model. Appendix Figure 1 illustrates the loss function, specifically MSE, of the DNN model as per the aforementioned settings. The x-axis, represented by the epoch, signifies the number of times the complete training dataset is iterated. There's a consistent trend where both training and validation losses decrease as the epoch count rises, until reaching a particular point. However, beyond this threshold, the model begins overfitting the training set, resulting in an increase in validation dataset loss, while the training set's loss stabilizes. This pattern is also evident in another error metric, the Mean Absolute Error (MAE). Following an analysis of the validation loss function, the optimal epoch was identified to be 25.



Appendix Figure 1. MSE loss function of DNN model