THE RELATIONSHIP BETWEEN STREET VISUAL FEATURES AND PROPERTY VALUE USING DEEP LEARNING

WEI LI
March 2020

SUPERVISORS:
Mila Koeva
Claudio Persello

ADVISER:
Dr. M. Kuffer

THE RELATIONSHIP BETWEEN STREET VISUAL FEATURES AND PROPERTY VALUE USING DEEP LEARNING

WEI LI
Enschede, The Netherlands, March 2020

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Urban Planning and Management

SUPERVISORS:
Mila Koeva (ITC-PGM)
Claudio Persello (ITC-EOS)

ADVISER:
Dr. M. Kuffer

THESIS ASSESSMENT BOARD:
Chair: Prof.Dr. R.V. Sliuzas
External examiner: Dr. M. Wang

ABSTRACT

Recent studies on property valuation models have been using a growing number of factors to improve their accuracies, such as physical characteristics, location, accessibility, and environmental factors. However, beyond such 'hard' location factors, also 'soft' factors such as the aesthetic of appearance and street visual features, have an impact on housing prices. From an economic perspective, a place with good perceptual value will bring more value for users since it has a positive impact on achieving the goal of diverse health, social, economic, and environmental public policy. Thus, residents are willing to pay more to have better conditions. Hence, the issue of the street perceptual value is important, but it is not used in property valuation models (e.g., hedonic price models) due to its complexity to be modelled. In recent years, street view image as a new data has been widely used to explore the relationship between street visual features or street visual quality and socio-economic variations such as crime rate, income, population density, etc.

Inspired by the mentioned above, this study aims to explore the impact of street visual features extracted from the street view images on housing prices in Xi'an. To achieve this goal, the study first uses Fully Convolutional Networks to extract 17 categories features from the street view image. At the same time, for comprehensively analyze key factors affecting housing prices and improve the accuracy of the property valuation model, the auxiliary geospatial data, which constituted the main independent variables in the traditional research (such as location characteristics, house characteristics, and surrounding infrastructure characteristics), also contained in this work. Then, to test the importance of particular variables with respect to the model accuracy, the study using random forest builds three property valuation models with different data sources. The results show that the street visual features can explain the majority of the variance of the house price. By comparing the results of three models, the model using geospatial data performs better than the model using street view image data. More specifically, the results show that there are non-linear relationships between different street visual features and property value. In addition, compared with the hedonic model, this study shows that the random forest regression model can more accurately estimate the housing prices.

Keywords: property valuation, street view image, satellite image, deep learning, hedonic price model

ACKNOWLEDGEMENTS

Upon the completion of this thesis, I would express sincerest gratitude to those who have offered me encouragement and support during the course of my study.

My deepest gratitude goes first and foremost to my supervisors, Mila Koeva, Claudio Persello and Dr. M. Kuffer for their numerous valuable comments and guidance with incomparable patience. During the process of selecting the research topic, improving the outline and the argumentation, writing the thesis, they give me enlightening guidance, advice, and help. Without they consistent and illuminating instruction, this thesis would have been impossible to reach its present form. Their strong sense of responsibility and professional dedication also left a deep impression on me, which not only taught me how to conduct academic research, but also set an example for my future work.

I would like to express my heartfelt thanks to ITC and all teachers who have helped and taught me in this university that greatly broadened my horizon and enriched my knowledge. Their lively and enlightening lectures have provided me with a firm basis in terms of linguistics, academic writing, statistics, and professional knowledge. It will always be of great value to my future academic research.

Finally, I'd like to give my special thanks to my friends and classmates, especially Jiaxin Sun, Xin Tian, Yu Li, Violla Okoth, from whom I get tremendous love and encouragement. Besides, they have helped me and shared with me my worries and happiness, wishing our friendship last long.

Last but not least, I am deeply indebted to my family for their loving considerations and supporting me to do whatever I want to all through these years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Optimizing property valuation models has been an attractive objective, and many scholars have made sustained endeavours to achieve this goal. Through literature review, a growing number of factors have been considered to improve the accuracy of the property valuation model. However, the relationship between street space quality and property value has not been investigated since many methodological and technical challenges have to be solved. Using state-of-the-art quantitative methods, this study attempts to analyse the impact of street space quality captured by street view images on property values.

1.1. Background and Justification

The housing issue is not only an economic problem but also affects social stability (Case & Shiller, 2004). Especially in China, housing investment has made a significant contribution to China's gross domestic product (GDP) growth in the past decades (Deng & Chen, 2019) . But at the same time, soaring housing prices have not only prevented most urban residents from buying new homes but have also increased inequalities among urban residents, leading to possible social and political instability (P. Li & Song, 2012). Focusing on the phenomenon that prices and transactions continue to surge, the Chinese government has issued policies on 'speculative' home purchases to rein in rising property prices, while implementing relevant measures to improve long-term housing supply(McFarlane, 2019). In this situation, the rationality of the property pricing mechanism has attracted more attention. Analysis of the property valuation is therefore crucial in the process of making the housing policies and achieving sustainable urbanization (Yuan, Wu, Wei, & Wang, 2018a). In this context, property valuation models can provide important information for housing management, policy-making, and economic analysis.

Commonly, the difference in prices between two houses is due to the difference in their attributes such as location, size (Chin & Chau, 2003). The attributes affecting housing prices are, therefore, a research hotspot and investigated by many scholars. Li (2019) categorized the indicators of property price into three types: physical characteristics, location characteristics, and environment characteristics at the city level and built a 3D property value model which shows the spatial heterogeneity of the property market in Xi'an. Ying (2019) showed that view quality, sky view factor(SVF), sunlight and property orientation are effecting property values at the neighborhood level and proved that building height has a negative impact on property value. Zhang (2019) focused on an individual building and combines physical characteristics of the apartment, accessibility characteristics, and a cultural indicator such as Fengshui as an index affecting property value to investigate the property value mode model, among others. To sum up the above discussion, the authors investigated indicators from different perspectives and multi-scale. However, some amenities value such as the prestige of a neighborhood, the aesthetic of appearance, and visual features such as graffiti also have an impact on property value. In theory, a place with good perceptual value has a positive impact on achieving goals related to diverse health, social, economic, and environmental public policy (Carmona, 2019). From an economic perspective, residents are willing to pay more to have better conditions such as air quality improvement and a positive neighborhood environment (Lavaine, 2013). Hence, the issue of the street perceptual value is important, but it is also an overlooked aspect of property valuation since the difficulty of data collection at large scale is not solved. In this work, it is assumed that the street visual features extracted from street view images reflect urban environment quality and resident perception. Figure 1. 1 shows the visual appeal of some neighborhoods and the average housing price of the surrounding area.

Figure 1. 1: The map illustrates the appeal of the street across Xi'an

In recent years, growing attention has been given to the systematically quantifying perceptual-based visual features (Naik, Raskar, & Hidalgo, 2016; Dubey et al, 2016). This benefits from the formation of new data environments such as Google street view images, as well as the boost of advances in urban research methods such as deep learning, big data mining and computer vision (Long & Ye, 2016).

Literature(P. Zhang et al., 2019) shows that deep learning based techniques have achieved the State-of-the-art performance in the field of semantic segmentation. Odgers et al (2012) showed that street view images are a reliable and cost-effective data source and provide new opportunities to measure neighborhood features. For instance, Naik, Raskar, and Hidalgo (2016) measured the urban appearance using street view images and found that the variation of urban appearance will widen income inequality. Ibrahim, Haworth, and Cheng (2019) used street view images, combine with CNNs and computer vision to detect slums, pedestrian, and transport modes.

In property valuation aspect, Law, Paige, & Russell (2018) used street view image and satellite image data and adopted neural network as an approach to extract features from images to estimate property value of London, but the result indicates that neighborhood space quality is more important than street space quality for buyers. However, in China, there is no literature using street view images to estimate street space quality as an indicator of the property valuation model. In general, the urban environment in Chinese cities differs from the one in London in terms of social, economic, and cultural aspects. Thus, it is expected that the results and conclusion will differ. On the other hand, since the results show the street view has little impact on housing prices in London, the authors did not explore the relationship between automatically extracted features and the property value. Therefore, the general idea is to take advantage of the capacity of the deep learning-based techniques to extract detailed information about street space and estimate property value. At the same, this work expects to provide more evidence for generalization of methods that using deep learning-based techniques extracting features from street view images.

Street space, as one component of public space, can be regarded as an imperative complement for affecting housing prices. But to systematically analyze key factors affecting housing prices and improve the accuracy of the property valuation model, only explore the street visual attributes is not enough. Therefore, I simultaneously use the auxiliary geospatial data, which constituted the main independent variables in the traditional research (such as location characteristics, house characteristics, and surrounding infrastructure characteristics) to improve the predictive power of property valuation models. The following section will briefly introduce the relevant works and methods in this area and the research gap.

1.2. Research gap identification

This study aims to investigate the effect of the street visual features base on street view images on housing prices in Xi'an, China. Literature indicates that the amenity value from visual features is difficult to assess, as a consequence property developers rarely take these factors into consideration (Jim & Chen, 2006). However, a current study certificate that the street view images can capture the street visual attributes reflecting street visual quality (Salesses, Schechtner, & Hidalgo, 2013). Law, Paige, and Russell (2018) demonstrated that using CNN extracting features from street view images combined with the hedonic model can obtain an encouraging result. Currently, there are few studies systematically quantifying street visual features in property valuation in China. Therefore, there are potentials for further investigation. Exploring the value of street visual features can give us a better understanding of the variation in property value and provide better insight for decision-makers in terms of urban planning and management.

The second gap is related to street visual attributes. To be specific, past studies were trying to analyze and correlate the street visual features with property value, but it still exists insufficient points. Some studies estimated the street green space ratio and demonstrated its positive impact on property value. As an example, Zhang and Dong (2018) developed an index of street green space using a hedonic model to measure the influence of street greenery. However, street greenery cannot reflect a comprehensive view of street visual characteristics. However, through the literature review, some studies analysed the street perception value from multi-dimensions but not connected with property price variation. For instance, Jingxian and Ying (2017) from the enclosure, human scale, transparency, tidiness, and imageability estimate the street visual value. Yaotian (2016) proposed using multi-year street view images to estimate the variation of street space and recognized the impact factors. These studies investigated the important indicators affecting users' perception of street space and tried to develop a framework or index to characterize the street visual features, but not correlated it with property value. Extracting discriminative features and correlating them with property value can improving property valuation models.

1.3. Research Objectives and Research Questions

1.3.1 General objective

To develop a deep learning model for modeling housing prices in Xi'an city by adding the street visual features to general geospatial indicators.

1.3.2 Specific objectives:

1. To identify the street visual features that affect property values.
2. To identify geospatial indicators that affect property values in Xi'an.
3. To evaluate the street visual features using deep learning.
4. To develop a property values estimation model with street visual feature indicators and geospatial indicators.

1.3.3 Research questions

1. To identify the street visual features that affect property value.
   - What street visual features affect housing prices?
   - To what extent of can the street visual features explain housing price variation?

2. To identify geospatial indicators that affect housing prices in Xi'an.
   - What are the current indicators that affect property value in Xi'an?
   - To what extent does the current indicators explain the variation in house prices?
3. To evaluate the street visual features using deep learning.
   - What is the optimal architecture of an FCN in terms of accuracy and efficiency?
   - What is the best strategy for selecting train, validation, and test sets?
4. To develop a housing price model for prediction with street visual quality indicators and geospatial indicators.
   - To what extent the housing prices model can predict variations of the property value?
   - To what extent the street visual features can explain variations of the property value?
   - What indicators influence more the property value in Xi'an?

1.4. Hypothesis

The central hypothesis in this study is that the street visual features have an impact on housing prices, whether positive, negative, or complex, and street view images can capture the visual features. The anticipated results are through analyzing the street view images using deep learning quantify the amenity value of street visual features and improve the accuracy of the property valuation model.

1.5. Thesis structure

The thesis structure is as follows:

Chapter 2 reviews related literature in this field of index affected property value, the method used to estimate and predict housing prices and induce the street view image dataset and the semantic segmentation used in this work.

Chapter 3 introduce the study area from urbanization and talent policy, following the data description.

Chapter 4 starts with the data pre-processing and potential variables selection description. Then, elaborate on the random forest algorithms along with feature selection, variable importance, and model evaluation.

Chapter 5 presents the majority of results. First, the result of the visual features extracted from the street view image is shown. After the correlation analysis, the main results of the RF model for housing prices estimation are illustrated, end up with discussion and limitation.

Chapter 6 is the conclusion. It mainly answers the research questions proposed in chapter 1.

2. LITERATURE REVIEW

This chapter elaborates on the related work to provide an overview of the topic and clarifies research directions. First, the previous work in terms of indices that are influencing housing prices is presented. Second, an overview of methods related to this research is presented.

2.1. Traditional Indicators that Affecting Housing Prices

The housing price is strongly linked with social-economic development and citizens' quality of life. Understanding the dynamic changes in housing prices is crucial for formulating effective housing policies and promoting the equitable development of society. Commonly, the difference in prices between two house is due to different their attributes such as the number of rooms, size, elevators(Tung Leong Chin & Chau, 2003). The attributes affecting housing prices are a research hotspot and investigated by many scholars. Through literature review, features in location, neighborhood, structure, and environmental aspects are extensively identified determinants.

The topic of housing price estimation has been well explored by numerous researchers from a different aspect. Same as heterogeneous goods, housing prices are the capitalization of utility-bearing components in the market. The conventional method is the hedonic price modelling that is the application of regression analysis to measuring the influence of factors that premise identified and affected housing prices. In the literature, the determinants of housing value are diverse and complicated, while usually attributed to three aspects, that is *structure characteristics, location characteristics, surrounding environmental characteristics* (Wittowsky et al., 2020).

Structure characteristics refer to the attribute of house units itself, such as the number of bedrooms, land area, age of the house, and other physical characteristics. Opoku & Abdul-Muhmin (2010) found that in Saudi, the size of the kitchen, the number of bedrooms, and the size of the bedroom are the major factors affected housing value. T. H. Tan (2012) point out that the number of bedrooms and private living space is an important factor for first-time homebuyers. Lu (2018) explored the relationship between view orientation of the dwelling units and property value in Shanghai estate market, China, and proved that the dwelling units with south view orientation have higher premium about 14% on property value compared with other orientations because of the better aesthetic effect of scenic views combined with sunlight. Xiao et al (2019)reveal that the amenity value of landscapes exists vertical heterogeneity at different floor levels within a building since the landscape proximity influences the interaction of floor level and housing price.

Location characteristics are not only referred to the geographic location of house units, but also closely concern the accessibility of household to the job, surrounding facilities, other social networks, and urban amenities. Transportation infrastructure, education facilities, and commercial facilities are the most selected factors. First, public transportation plays an important role in urban dweller's daily travel in terms of higher accessibility and opportunity for activities throughout the city. There is extensive literature on housing prices premiums for closer to **transportation infrastructure**. For example, Xu et al. (2015)reveal that people are more willing to pay for subway proximity while driving restriction policy imposed. The influence of transfer stations is greater than that of non-transfer stations and this difference is more obvious in a suburb than in inter-city (Dai et al., 2016; R. Tan et al., 2019). Zheng et al. (2016)proved that the metro station promotes the consumer amenities development, and further induced the housing prices rising. **Educational facilities** are another factor and have been capitalized into property value. Hansen's (2014) research shows that parents will move to education-related houses to make sure their children have good educational quality and willing to pay more property value. The school with an excellent reputation brings a premium of about 16%-20% to property value in Coventry, U.K (Leech & Campos, 2003). Similarly, school quality has a significant impact on property value, and potential buyers willing to pay 27%-39% additional premium in Hong Kong(Jayantha & Lam, 2015). Due to the "nearby enrolment" policy lead to school district effect, Wen et al. (2019) affirm that the quality of the secondary school has the most remarkable impact on house prices, then is the quality of the primary school, university, senior high school in Hangzhou. **Commercial facilities,** including shopping malls, supermarkets, financial institutions, are one of the urban amenities providing substantial convenience for dwellers. Previous studies have proved that the shopping mall positively affected the property value of nearby communities, while with the distance to shopping mall increasing, the beneficial effect will decrease (L. Liu et al., 2019). In addition, the large shopping centers have higher attractiveness since owing to diverse and full range commercial activities. Based on the gross leasable area, the number of shores, parking space and construction year, François Des Rosiers, Antonio Lagana, Marius Thériault (1996) selected three types of shopping centers from the neighborhood, community, and regional level, and the results confirmed that large size of shopping center did have a greater positive contributory effect on the surrounding residential property value. Retail as one of the widely spread commercial facilities deeply related to the convenience of residents' lives. The study of Song & Sohn (2007) calculated the accessibility of residential 'units to retail stores by developing accessibility index in the city level, and further using the gravity-based model measured the relationship of retail accessibility and housing prices. The results support the hypothesis that greater spatial accessibility brings a premium to nearby housing prices.

Environmental characteristics refer to the urban landscape, such as rivers, mountains, parks. The urban landscape provides aesthetic, recreational, and ecological functions, and benefits to residents' mental, emotional, and physical wellbeing. Luttik (2000) collected nearly 3000 property transaction data in eight towns to explore the external effect of environmental attributes on housing prices. The salient result is housing with a pleasant view will have a higher transaction price, more specifically, a river view increased the housing prices 8%-10% and open-space view increase 6%-12%. In China, with improvements in education, income, and living quality, the urban resident is becoming to pursue a high-quality living environment. Wen et al. (2015) verified the positive effect of the landscape of inner-city such as mountains, lakes, rivers, and parks on housing prices in Hangzhou. Specifically, every 1% increase in the distance of West Lake and the nearby park will lead to a 0.229% and 0.052% decreased in housing prices, respectively. G. Liu et al.(2019) using Chinese geomantic omen theory, proved that the river land mountain landscapes drive up the housing prices almost 15% and much higher than the value of accessing river or mountains independently by taking Chongqing, China as a case study.

2.2. Visual features that affect housing prices

Excepting these 'hard' factors, the 'soft' factors or intangible assets such as visual characteristics of a community, which reflect the safety, lively, depressing, or beautify of neighborhood environment, also have an impact on housing prices.

Through the literature review, some studies have proved the importance of visual feature value on housing prices. In the field of housing prices estimation, Poursaeed et al. (2018) investigated the impact of visual characteristics of a house on transaction prices by developing a deep convolutional neural network on a large image dataset, including interior and exterior of a home. The result shows that by adding visual characteristics significantly improved the performance of the housing prices estimation model. Y. Zhang & Dong (2018) explored the impact of street greenery at block level using green view index (GVI). The indicator of GVI is the percentage of green vegetation area from the perspective of human eyes, which reflect the perception of pedestrians on street greenery (Yang et al., 2009). The finding demonstrated that higher street-greenery brings additional value for surrounding properties, and homebuyers are willing to pay the premium for a house with higher street-greenery. Arietta et al. (2014), using street view image and Support Vector Regression, explored the relationship between a set of visual elements of a city and its social-economic attributes such as crime rate, property value, population density. One of the conclusions is that the hedges, gable roof, and tropical plants are highly related to the higher property value.

There are several related works validated that city visual characteristics can be used to estimate social-economic activities. Although these studies do not estimate housing prices directly but provide evidence that visual elements of a city are an overlooked aspect of housing prices. For instance, Kevin Lynch identified the five foremost important visual elements, which are paths, edges, districts, nodes, and landmarks, which have on effect on the perception of city's visitors or residents(Kevin lynch, 1960).

Yin & Wang (2016) extracted the sky element by applying Artificial Neural Network and Support Vector Machine on Google street view imagery (GSVI) image. They found that the proportion of the sky in GSVI can reflect the visual enclosure of the street. Furthermore, the visual enclosure of the street negatively related to pedestrian counts and walkability. Ewing et al. (2016) from the perspective of urban space design measuring the 20 streetscape features identified that street furniture, the number of shops, restaurants, public space in the street, the area of the window of ground floor façade are significant features that positively related to pedestrian volume.

2.3. Method to Estimate Housing Prices

2.3.1. Hedonic price model

The hedonic price model is the most common method used to the scientific investigation of various aspects of the real estate market, which inspired by CS Lancaster (1966)consumer theory and Rosen's theoretical (2019) model. The term hedonic is derived from Greek, refer to the sense of pleasure of buyers obtained from an attribute of a specific commodity. In 1974, economist Rosen(2017) introduced the hedonic model to calculate the contribution of different factors on wages, such as the cost of living, education, work experience, and further compared the quality of life in some American cities. Rosen think that goods are valued for its utilitarian attributes or characteristics. Hedonic price is the aggregation of individual implicit prices associated with specific attributes. In the real estate field, the objective of the hedonic pricing model is to measure the implicit value of the house's attributes or characteristics base on the transaction price.

Numerous papers are using the hedonic model to capture the relationship between the property value with the characteristics associated with different houses. Through perusing the previous literature on the application of the hedonic model, T L Chin & Chau (2003) examined and identified the commonly used attributes in housing prices estimation. Then T L Chin summarized that the application of the hedonic model concentrated on estimating the contribution of location attributes, structural attributes, neighborhood attributes to housing prices. Online et al.(2010)take stock of most cited studies on the hedonic model and classified them into some categories base on the selected indicators. The classification shows that neighborhood attributes and environmental amenities are over-researched, while the social aspect such as the effect of racial segregation, the crime rate on housing prices is overlooked.

The hedonic model uses regression analysis to measure the importance of various indicators on housing prices. However, the economic theory does not assign an agreement on the functional relation between housing prices and characteristics. Therefore, many functional forms occurred in a related document such as the linear, the log-linear, the semi-log-linear, the Box-Cox form, looking forward to increasing the goodness of fit.

2.3.2. Machine learning

In recent years, machine learning algorithms (MLA) have been prevalent in the field of housing price estimation. Although the hedonic price model has been extensively used, it shows potential limitations in fundamental model assumptions and estimation, especially in uncovering the nonlinear relationship between housing price and housing characteristics and cannot deal with the problem of spatial autocorrelation and spatial heterogeneity. To pursue higher accuracy, ML has been regarded as an alternative approach to the hedonic model. Multiple studies have examined and identified that MLA has a remarkable performance in handling complicated data and regression analysis. For example, Fan et al. (2006)examined the usefulness of the decision tree approach through analyzing the relationship between the transaction price and housing characteristics using Singapore real estate market as a case study, and identified the significant variables of housing prices. Selim (2009) explored the determinants of housing prices in Turkey and compared the performance of the hedonic method and artificial neural network(ANN). They found that ANN has higher prediction results. Gu et al. (2011) employed Support vector machine (SVM) to forecast housing prices and proved that SVM is a robust and competent algorithm in regression analysis. Y. Chen et al. (2016) selected six MLA including Gaussian process regression (GPR), k-nearest neighbour algorithm (k-NN), backpropagation neural networks (BP-NN), radial basis function neural network (RBF-NN), fast decision-tree (FDT) and support vector regression (SVR) to form an ensemble learning approach for housing prices estimation.

Among multiple machine learning algorithms, random forest (RF) has been considered an effective regression analysis method(Segal, 2003) and has been employed in many fields. RF is one of the bagging algorithms in Ensemble Learning. Breiman (2001) combined the bagging sampling approach and random selection of features to develop RF. An RF is an ensemble of simple individual regressor/classifiers, which perform regression by averaging predictions made by each classifier. In the RF algorithm, each classifier is built independently and prediction using a sample selected from the training data. The concept of RF is taking account of the strength of individual classifiers and the correlation among them to reduce the generalization error. Numerous literature highlight the advantage of RF in handling multiple dimensional data, multicollinearity, and less sensitive to noise and the overfitting problem. Breiman (2001) applied RF on the data sets selected from a different domain, including 13 small datasets, 3 large datasets, and 4 synthetic data sets, the results attest that the performance of RF is superior to other algorithms. By comparison with Adaboost, Breiman argued that RF is "favorably" comparable and saving computing time.

In recent studies, there have some examples using RF to estimate and predict housing prices. Yoo et al.(2012)applied RF to variable selection and housing prices modeling. By comparison with the traditional hedonic model, RF achieved the highest prediction accuracy. Hu et al (2019)using a six-machine learning algorithm to build housing prices prediction model including random forest regression (RFR), extra-trees regression (ETR), gradient-boosting regression (GBR), support vector regression (SVR), multi-layer perceptron neural network (MLP-NN) and k-nearest neighbor algorithm (k-NN). They found that RF shows the best prediction accuracy than other models. Antipov & Pokryshevskaya (2012) verified that RF is a competitive method in property value mass appraisal. In the analysis process, RF demonstrates the stability of outliers, the ability to work with a target data set that include missing value, and multi-level categorical variables. Besides, they compared RF with other 9 algorithms(Multiple regression, CHAID, Exhaustive CHAID, CART, k-Nearest Neighboors (2 modifications), Multilayer Perceptron neural network (MLP) and Radial Basis Function neural network (RBF), Boosted Trees) and the RF model give a best prediction results.

2.4. Street View Image

The street is an important component of the urban physical environment. Multiple studies have verified that street environment has a direct or indirect effect on residents' behavior, life expectancy, obesity, and mental health(Clarke et al., 2010; Mujahid et al., n.d.). Measuring street space attributes have attracted many scholars' attention, but relevant approaches such as in-person audit, interviewing methods or questionnaires are expensive and time-consuming (Rundle et al., 2011).

In recent years, street view (SV) image has been regarded as an effective alternative data. It records many cities at high resolution, thus provides a new opportunity for the recognition of street physical features on a large scale. SV images are available from applications such as Google street view (GSV), Baidu street view (BSV), Tencent Street View (TSV). Besides, map applications provide the API for users to download SV images for free. Compared to the traditional methods, the street view image contains rich urban physical information such as infrastructure, landscape, building elevation features. This data source lays the foundation for observing, perceiving, and assessing street environment.

Different from remote sensing images that present urban information from the version of the sky to ground, SV image is recording the urban environment from a similar vision of human eyes. It captures more rich information on street 3D space. Through perusing related literature, the SV image has been extensively used to extract visual features in the urban environment. One common application direction is quantifying street green landscapes, green view, an indicator that quantitative reflect the percentage of green area in the view of the human eye, is been calculated using street view image to estimate the street-level green landscape quality(Wu et al., 2019; J. Chen et al., 2020). Nguyen et al. (2019) extracted neighborhood environment characteristics from SV, and further estimate the link between the built environment and resident health (chronic disease, premature mortality) at the county level. F. Zhang et al. (2019) demonstrated the spatial-temporal urban mobility pattern by train deep convolutional neural network (DCNN) on SV image, and results show that high-level visual feature extract from SV image can explain the taxi trips variation. Kang et al. (2018) built a framework for classifying the functionality of individual buildings from the street view image. Runge et al. (2016) proposed a system that uses Google street view images to generate scenic routes and classify them according to their visual characteristics to enhance the driving experience.

In addition, considerable literature highlights the advantages of SV image in street scene quantitative representation (F. Zhang, Zhang, et al., 2018), measuring changes of the urban physical environment (Naik et al., 2017), mapping resident perception of built-up area (F. Zhang, Zhou, et al., 2018).

2.5. Semantic Segmentation

Semantic segmentation is the process of assigning each pixel in the input image to a category of what is being represented, belonging to dense pixel-wise prediction. In terms of semantic segmentation on street view image, these labels could include person, traffic light, car, tree, etc. Semantic segmentation is a crucial foundation for robots and other unmanned systems to completely understand the context in the environment. At present, semantic segmentation has been widely employed in the field of autonomous vehicles and medical image diagnostics.

Semantic segmentation is one of the key applications in computer vision and has been studied for many years. With the development of Deep Learning, Semantic segmentation has achieved tremendous progress. In 2014, Long et al. (2014) first proposed the Fully convolutional network(FCN) for image segmentation area. It is an extension of classical CNN but replaces the fully connected layer of CNN by a fully convolutional layer. More specifically, different from CNN that uses a fully connected layer with a 1*1 convolutional layer obtaining a fixed-length feature vector for classification, FCN only has a convolutional layer and pooling layer which enable it an ability to handle arbitrary-sized images. Besides, FCN adopts a deconvolutional layer to up-sample the feature map generated by the last convolutional layer to make the size of the output image the same with the input. In this way, FCN achieved the prediction for each pixel on the up-sampled feature map and preserved the spatial information of the original input image simultaneously.

On the basis of FCN, scholars have proposed a set of advanced deep networks that have significantly accelerate the progress of semantic segmentation.

Segnet was developed by the University of Cambridge(Badrinarayanan et al., 2017). The main insight is the upsampling layer in the decoder stage that uses pooling indices recorded in the max-pooling step of the encoder stage to upsample feature maps. The upsampled map is convolved with a set of trainable filters and then generate dense feature maps, where the feature maps have restored to the initial resolution. Comparing Segnet with FCN, DeepLab-LargeFOV, DeconvNet, the results showed that Segnet is more efficient in terms of computational time and memory(Badrinarayanan et al., 2017).

Yu & Koltun (2016) proposed the dilated convolutions that supports expanding receptive fields and aggregating multi-scale contextual information without a decrease of spatial dimensions. As part of the work, a new network structure containing dilated convolutions has been designed and reliably increased the accuracy of the semantic segmentation system.

Considering the fact that the Deep Convolutional Neural Networks (DCNNs) has limitations to get accurate object segmentation since the very invariance properties. L.-C. Chen, Papandreou, Murphy, et al.(2018) research overcome this condition by combining the response of the final layer in DCNN with a fully connected Conditional Random Field (CRF) improved the ability of models to capture details and edge information. On this basis, L.-C. Chen, Papandreou, Member, et al.(2018) optimized it on three aspects: first, using 'atrous convolution' to effectively control the resolution of the calculated feature response in deep convolutional networks. This simultaneously allows us to effectively expand the field of view of the convolution kernel to integrate more context information without adding additional parameters or computational effort. Second, the authors propose atrous spatial pyramid pooling (ASPP) to segment images robustly at multiple scales. ASPP uses convolution kernels with multiple sampling rates and effective fields of view to detect incoming convolutional features, thereby capturing the contextual content of targets and images at multiple scales. Third, combining DCNN and probabilistic graphical models improved the localization of object boundaries.

L.-C. Chen, Papandreou, Schroff, et al. proposed "DeepLabv3+" model. In order to solve the challenge of multi-scale segmentation of objects, in this research, a cascade or parallel atrous convolution module with different atrous rates is designed to capture multi-scale context information. In addition, it extends the previously proposed atrous spatial pyramid pooling module, which detects convolution features at multiple scales, encodes global context features at the image level, and further improves performance. In the end, the authors tested its model on the PASCAL VOC 2012 dataset and the results show that 'DeepLavb3' achieved the best performance compared with other state-of-art models in the same dataset.

The success of various algorithms related to semantic segmentation has created an opportunity for researchers to understand cities in a more accurate and precise fashion. For instance, Gebru et al. (2017) demonstrates a method that estimates the socioeconomic attributes such as income, race, education, and voting patterns of US cities inferred from cars extracted from Google street view image using deep learning-based computer vision techniques. Goodfellow et al. (2013) proposed a unified deep convolutional neural network that combines localization, segmentation, and recognition to recognize arbitrary multi-digit numbers from Street View imagery. Wojna et al. (2017) presented a neural network model to recognize scene text, including street names, business names. Middel et al. (2019) Zeng et al.(2018) employed a deep learning approach to segment the Google street view image for exploring the form and composition of cities, and further evaluate the relationship between street-level morphology, micro-climate with urban features.

2.6. Research Concepts

Considering the literature review and the above-mentioned discussions, the core concept of this study and their relations are briefly shown below (Figure 2. 1). In traditional research, the main indicators are physical characteristics, location characteristics, surrounding infrastructure characteristics. In recent years, the development of street view images and image segmentation techniques provide an effective way for the research about surrounding environment characteristics, especially street visual features (Law et al., 2018). Simultaneously, the correlated state of the art method is Deep Convolutional Neural Networks as it shows a perfect performance on image recognition, segmentation, detection, and has been regarded as the best techniques for extract features from images(Khan et al., 2019). In this research, the core is combining the current indicators used in property valuation in China and street space quality to build a comprehensive index and to develop a deep learning model for property valuation in Xi'an city.

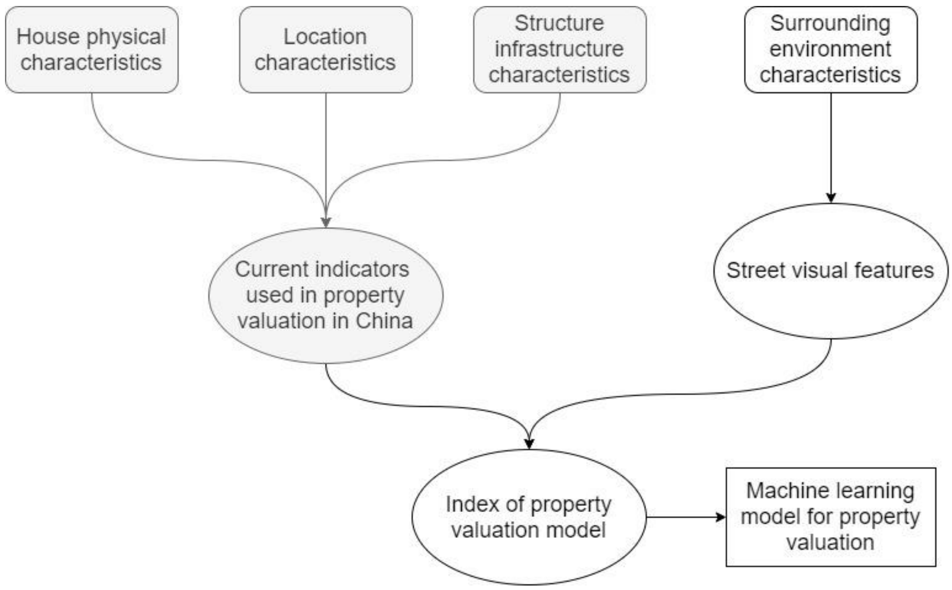


Figure 2. 1: Research concept

3. STUDY AREA AND DATA DESCRIPTION

3.1. Study area

This section, starts with an introduction of the study area, followed by explanation of the used available data. Xi’an is the capital of Shaanxi province and the political, economic, and cultural center of northwest China. It covers a total area of 10752 square kilometers and has jurisdiction over 11 administrative regions and 2 counties (Figure 3. 1). Considering the status and scale of urban development, Xi'an is a representative study area.

Fuelling by the Belt and Road and Building National Central City policies, Xi’an is developing rapidly in recent years. Guanzhong Plain City Group Plan was approved by the central government in 2018 in which it was proposed to construct Xi 'an as the national central city. In 2019, the government published the Xi'an 2050 Space Development Strategic Plan, in which promoting coordinated development of Xi’an with surrounding cities and building a national city cluster. It accelerated the development of industrial agglomeration and the concentration of population in the urban area. According to the bureau of statistics, Xi'an's GDP in 2019 is 93.219 billion yuan, ranking 24th in the country. Besides, in the 2017 year, to promote economic development, the government puts forward policies about Census Register and Talent Attraction which has obtained significant results. The population increased from 824.93 ten thousand in 2016 to 986.87 ten thousand in 2018.

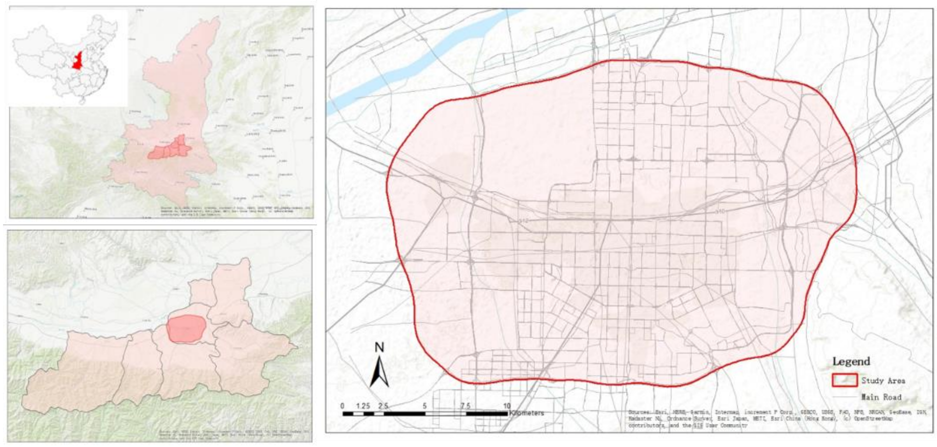


Figure 3. 1: Study area location

Urbanization and population increasing have also promoted real estate market development. As shown in Table 3. 1, the housing price keeps a stable rising trend before 2016, but the price grows rapidly after 2016. Especially, the house price in 2018 increased by 56% compared to 2016. In addition, the per capita building area of urban residents is increasing steadily because of the improvement of living standards and the change of family structure caused by carrying out the two-child policy. Data from the Xi'an Bureau of Statistics shows that people prefer to buy the house over 144 m^2. In 2019, the transaction volume of the house of 90 m^2 or less fall 37.7%, but the sale volumes of the house of 90 to 144 m^2 and the house of more than 144 m^2 increased by 3.6% and 19.5%, respectively, compared to the same quarter a year earlier(Bureau, n.d.).

On the other hand, with the development of the property market, a series of problems are exposed such as overheated investments and a housing price bubble. The property market development is closely related to people's basic living conditions, sustainable economic development, and social stability. Therefore, to limit speculation in the housing market, stabilize price fluctuations, and keep the housing market in a healthy state of development, since 2017, the Xi’an government has introduced a series of administrative policies. One effect is that the investment by enterprises for real estate development tends to be flat after experiencing continuous rapid growth.

Table 3. 1: Selling Price Indices of Residential Buildings from 2008 to 2018

year	Investment by enterprises for real estate development(100 million yuan)	Average Selling Price of Commercialized Residential Buildings(yuan/sq.m)	Total Population (year-end) (10000 persons)	Per capita building area of urban residents
2018	1446.51	9984.54	986.87	34. 4
2017	1505.03	8166	905.68	33. 7
2016	1337.35	6385	824.93	33. 4
2015	1304.6	6221	815.66	32. 1
2014	1321.91	6105	815.29	32. 06
2013	1226.28	6435	806.93	33. 43
2012	1003.85	6224.03	795.98	32. 98
2011	836.05	5829.79	791.83	28. 9
2010	670.41	4341	782.73	28. 7
2009	566.77	3749	781.67	28. 4
2008	421.13	3769	772.3	26.32

3.2. Data description

In this work, we used five types of data to estimate housing prices as shown in Table 3. 2**.**

Table 3. 2: Overview of available data

Data	Time	Source
Housing Prices	September 2019	Homelink website (https://xa.lianjia.com/)
Street View Images	2019	Baidu Maps (https://lbs.qq.com/panostatic_v1/index.html)
POIs	October 2019	Amap website (https://lbs.amap.com/api/webservice/summary)
Road Network	2019	Open Street Map
Remote sensing images	2017	GaoFen-1(16m).

The house prices data. It is the second-hand price of apartments on different floors. This data is collected from the Homelink website. Homelink is one of the largest real estate brokerage companies in China (https://bj.lianjia.com/). In this work, the house prices data includes 34198 samples (Figure 3. 2) and each sample contains the house information that is administrative district, coordinate, house unit price, year of construction, building area, house type, community name, number of floors, the number of living room, the number of bedrooms, the number of kitchens, elevator(if there is or no).

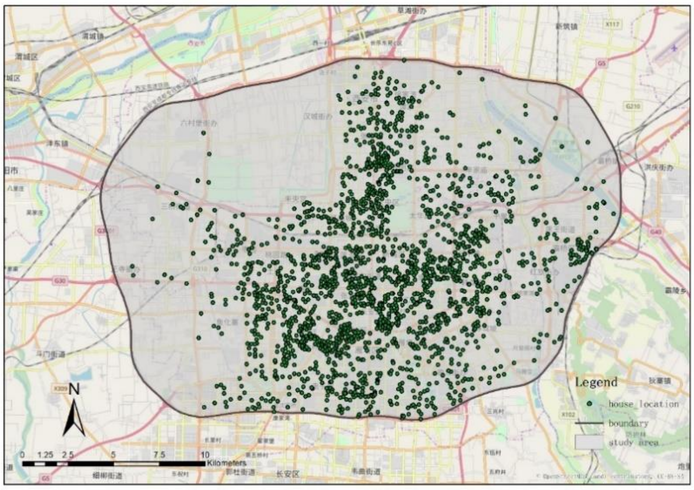


Figure 3. 2: House price sample data.

Street view data and road network. The road network data is downloaded from the open street map including the main road, minor road, and branch. The road data is used to download street view images. The street view images are from Baidu Maps which is one of the most popular maps platform in China. The process is: first, breaking road lines into points at 50-meter intervals and then generating latitude and longitude coordinates information for each point in ArcGIS. Finally, based on these coordinate points to retrieve the street view images of all the streets of Xi 'an. Therefore, the location of the street view image is the same as the coordinate points.

Studies have shown that the field of view of the human eye is 80°-160° in the horizontal direction, 130° in the vertical direction (Figure 3. 3). Especially, the comfortable vision limit of one eye is about 55° in the vertical direction (Vasilios & Gasteratos, 2006). The website of the street view image provides the related parameters that can adjust to make sure the vision in the street view image is the same as the human eye's vision. At the same time, the website of Baidu Maps provides related parameters that can adjust parameters to make sure the vision in the street view image is the same as the human eye's vision. The parameters are shown in Table 3. 3. In this work, using Baidu Application Programming Interface (API) and set the parameters to a vertical angle of -10° -45° and a horizontal Angle of 120° , 4 images with a different orientation per each coordinate point were selected. Finally, 279663 images with a size of 850×680 pixels were selected in total. Figure 3. 4 shows an example of these street view image data, in which 10102 means the identification number of a coordinate point. The orientation information of SVI depends on the driving orientation of the street view car. In Figure 3. 4, 1 means the left side of the driving direction of the street view car, 2 is right, 3 is front and 4 is behind.

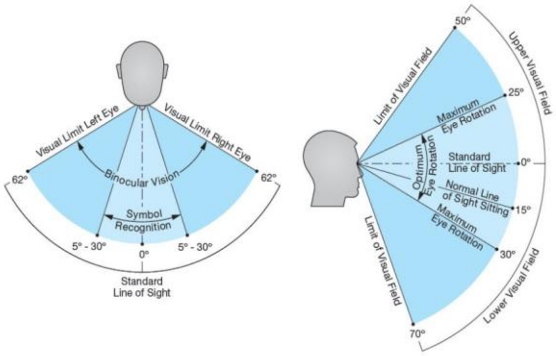


Figure 3. 3: Average field of view of the eye

Image source: https://www.quora.com/What-is-the-maximum-human-field-of-vision

Table 3. 3: the parameters for download street view image

Parameter	Description
Width	The width of the image ranging from 10 to 1024.
Height	The height of the image ranging from 10 to 1024.
Location	It shows the coordinates of the position. The format is: ling<latitude>, lat<longitude>.
Heading	This is the horizontal vision ranging from 0 to 360° .
Pitch	This is the vertical vision ranging from 0 to 90° .
Fov	This is the horizontal vision ranging from 10 to 360. If Fov = 360, it will show a panorama.

10102_1 10102_2 10102_3 10102_4

Figure 3. 4: Example of street view image data.

POI. Point of interest, or POI, is a term in cartography to represent a feature that located a point("Points of interest - OpenStreetMap Wiki," n.d.). In geographic information systems, it can be a specific point location. In this research, it is the location point of different facilities such as retail stores, shops, and banks. The data of POIs is obtained from the Amap website which is one of widespread used map software in China. In the Amap official documents of POIs, it contains contain 22 big categories, 264mid categories, 869 sub-categories. (map Api Poi classification table, 2014). This data used for assessing the neighborhood characteristics and location characteristics of each house.

Remote sensing images. This data was provided by Dr. Li from Chang'an University. In this work, it used to calculate the green rate of the surrounding area of the sample.

4. METHODOLOGY

This section elaborates on the research methodology, which contains five major procedures as shown in Figure 4. 1. It starts with data pre-processing, including removing duplicate data, dealing with missing values, data standardize. After that, the first part is to select multiple potential determinants from POI data and SVI based on similar studies. The second part is to train a regression model for estimating the housing prices of Xi'an. Finally, the housing prices estimation model is evaluated, and the relative importance of determinants are analyzed.

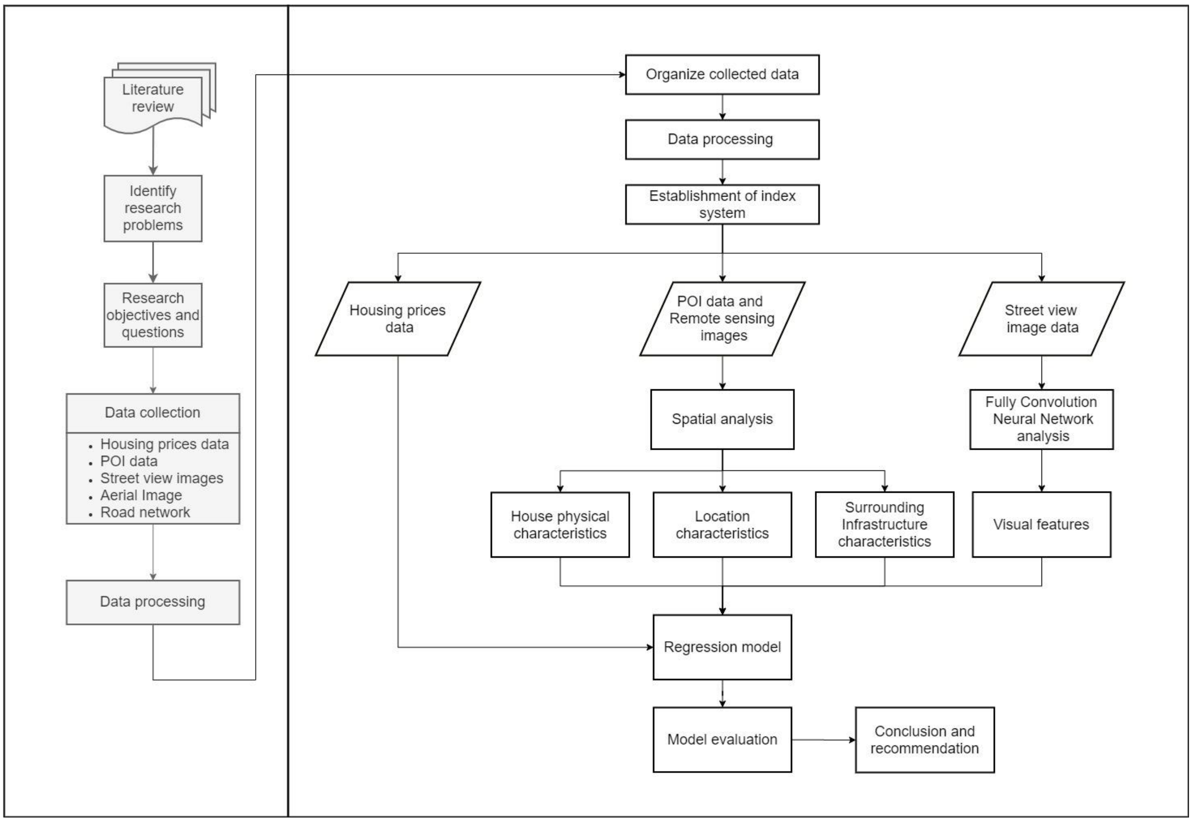


Figure 4. 1: Methodology

4.1. Data pre-processing

Data pre-processing is an important step for the training regression model. In this work, the major step is data cleaning. For housing prices data, there are three steps for data cleaning. Firstly, removing duplicate data in terms of the situation where the landlord publishes the same advertisement multiple times. Secondly, through data inspection and data profiling, removing abnormal records with obvious errors in the dataset. More specifically, filter and delete disturbing data, such as parking spot, shop. Thirdly, removing the data that is missing major important attributes such as total area, elevator, structure. Besides, removing the housing price sample in which no SVI has been collected within its surrounding area. As a result of the data cleaning process, 1784 records were deleted and 34573 records remained for further estimating housing prices.

The statistics of processed housing prices data are shown in Table 4. 1. Figure 4. 2 demonstrates the spatial distribution of original housing prices data using Kernel density analysis in ArcGIS, where the compact deep red region means the housing prices are highest, and the light red region indicates the housing prices are lowest. The imbalanced distribution of housing prices is as expected.

Table 4. 1: Statistics of processed housing price (unit: yuan/m^2).

The number of housing prices	Min	1st Quantile	Median	Mean	3rd Quantile	Max
34573	4619	12759	14964	15995	18041	81081

Figure 4. 2: Housing prices thermodynamic diagram

Second, by observing the distribution of housing prices, a logarithmic transformation has been performed to convert it to a normal distribution as shown in Figure 4. 3.

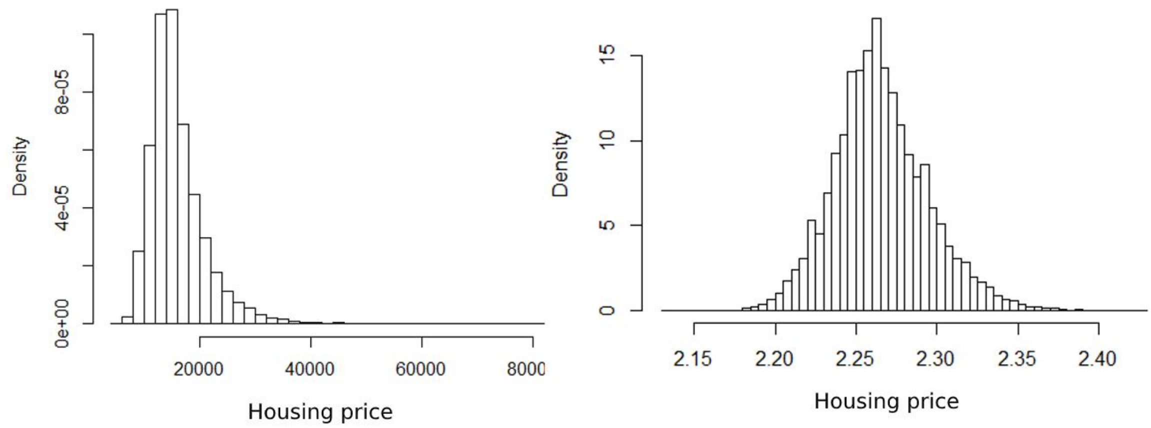


Figure 4. 3: The histogram of housing price data distribution(yuan/m^2)

Third, creating a dummy variable for the category variable. The information of dataset contain category data such as elevator (has or no), building height (high, middle, low), build type(slab-type apartment building, slab-type apartment building and tower, tower, and bungalows)and decoration(excellent, good, average, poor). To process data more conveniently and efficiently, in this work, dummy variables were created for the category as shown in Table 4. 2.

Table 4. 2: Category Data transform into numeric data

Variable	Category	Value
Building height	High	3
	Middle	2
	Low	1
Elevator	Have	1
	No	0
Type of building structure	Slab-type apartment building	4
	Slab-type apartment building and tower	3
	Tower	2
	Bungalows	1
Decoration	Excellent	1
	Good	2
	Average	3
	Poor	4

For street view images (SVI), where cars are visible, cropping was done, as shown in Figure 4. 4. In addition, images that were too dark we also removed (e.g., under the bridges).

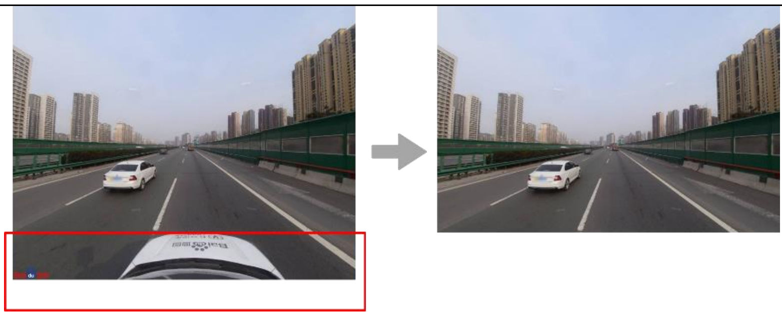


Figure 4. 4: an example of the street view image data process

4.2. Selection of potential determinants

Based on the data availability, analyzing the study area, and related research, this work identifies the four aspects to choose the potential determinants for housing prices.

Firstly, a buffer with a radius of 500 meters was used around each residential property to capture the spatial pattern effects of the neighbourhood environment and street view features. The buffer method with a fixed distance around the resident property presents the actual preference of residents for the surrounding environment such as amenities, social ties, and other infrastructure (Yoo et al., 2012). Through perusing associated studies, researchers have not reached a consensus for defining the neighborhoods in this issue. For example, (Sander et al., 2010)used the buffer with a radius of 100m, 250m, 500m, and 1000m exploring the urban tree values and its impact on property price. (Acharya & Lewis, n.d.) choose 1 mile and 1/4-mile radius as buffer size to capture the effect of environmental variables including open space, land-use diversity crime rate. According to the Standard for Urban Residential Area

Planning and Design proposed by Ministry of Housing and Urban-Rural Development of the People's Republic of China (MOHURD), one of strategy is taking 15-minutes walking time(800m-1000m) as the interval distance to plan the public infrastructure (China, n.d.). 15-minute walking distance is a suitable distance that residents are willing acceptable and has been regarded as a standard for building a liveable community in China(Su et al., 2019). Therefore, following the above-mentioned reasonings, the buffer with a radius of 500 meters was selected to analyze the impact of environmental characteristics on property value.

4.2.1. Structure characteristics

Based on the previous related research and data availability and exploring the actual living conditions of residents of Xi'an, 8 variables about housing structure were selected. Among them, the number of bedrooms, bathrooms are highly related to the utility and comfort of a house. Especially, more than one bathroom has been regarded as an improvement of living standards and will satisfy the needs of buyers' (Jim & Chen, 2006). The house floor level is an indicator that has been explored by other scholars and proved that floor level has an impact on housing prices since it associated with the proximity to landscape and noise (Xiao et al., 2019; Wen et al., 2020). In terms of the type of building structure, generally, the slab-type building does not exceed 12 floors and has less than 4 units. This house has good ventilation and lighting. By contrast, the tower is a high-rise building between the 12th floor and the 35th floor with a wide view, while have more unites compared with a slab-type apartment building. Therefore, the price of the slab-type building is higher than the tower. Lastly, the decoration is one of the indicators that affects housing costs. Similarly, the elevator indicates the convenience of living.

4.2.2. Location characteristics

For location characteristics, transportation infrastructure is an essential aspect affecting housing value because surrounding residents will benefit from lower commuting costs and better accessibility. Public transportation has become a main way for residents to travel in their daily life. Some researchers have provided evidence proving public transport, which has an effect on the housing value. (AlQuhtani & Anjomani, 2019; Cervero & Kang, 2011; Mulley, 2014). In Xi'an, the major public transport types are rail transit, bus, and bicycle. The bus route layout has covered the entire city and bus stations meet the demand adequately. Rail transit system covers a considerable part of the city, especially on the city center. Until September 2019, there are 5 rail transit routes in operation with a total of 107 stations, including 7 transfer stations. Therefore, for rail transport, choosing the shortest route of each house to the rail station as a variable is important. For the bus system, the number of bus stations were used within a radius of 500 meters as the variable. The number of bus stops suggests more bus lines also have higher accessibility and convenience. In addition, public bicycles are relatively flexible in terms of leasing policy and parking rules. Therefore, it was assumed that it has little impact on property value and were not considered.

Some previous studies used the Euclidean distance from a house to the nearest metro station as a variable. Considering the fact that residents are sensitive for the distance from their home to the metro station, this study takes the real distance (walking distance) from a house to the nearest metro station into account using Baidu Route Matrix API. The Route Matrix function allows for a quick and easy calculation of the distances and driving times between a set of points. It is needed to specify the start point and endpoints.

The steps of calculating the proximity to the metro station:

1. Calculating the Euclidean distance of each house to all metro stations.
2. Selecting the top 5 nearest metro station for each house.
3. Using Route Matrix API calculate the real distance.
4. Selecting the minimum distance for each house.

Lastly, as shown in Figure 3. 1, the built-up area boundary of Xi'an is a circle highway. The variable of the distance to the highway indicates the distance of each house to the urban boundary. The highway exit/entrance usually has a higher traffic flow with noise and air pollution. Through looking at the map, there are some communities located in the surrounding area of highway exit/entrance. Thus, it is assumed that the highway exit/entrance has an effect on housing prices and take it into analysis to further capture the preference or attitude of residents.

4.2.3. Neighborhood characteristics

For neighborhood characteristics, the first category is the **educational resource**. In Xi'an, children are assigned to designated primary or secondary schools within their living region. But the educational quality is diverse, parents prefer to take their children to a good school. Therefore, the school with good educational quality has a positive effect on surrounding housing prices (Jayantha & Lam, 2015; Jayantha & Lam, 2015). This work first focus on selection of school with a good reputation and then identification of the corresponding "schoolhouse" in the dataset based on the 'school district system' implemented in 2019 followded. Finally, there are 1053 'schoolhouses' in the dataset. In addition, the concept of "education real estate" or "school district housing" has to become popular in the real estate market. Therefore, the number of schools within 500m radius as a variable were taken for the calculation. Meanwhile, the kindergarten facilities are necessary for most homebuyers. Universities provide academic development but also have some sports facilities and landscapes. Therefore, in this study the number of kindergartens and the proximity to the university as potential variables.

The second category is the **convenience facilities**. To measure the effect of convenience facilities, five sub-categories have been considered. First, the hospital provides the health care service for residents and residents prefer a hospital with higher medical services. Thus, the AAA class hospital and the original hospital with the clinic have been separately evaluated. AAA class hospital belongs to the regional hospital that provides high-level specialized medical and health services and performs higher education and scientific research tasks to several regions. Employment is the number of employment places including government agencies, small and medium enterprises. The main sub-categories of the cultural facility, sports facility, and convenience facility are shown in table 4.4.

Table 4. 3: sub-categories of convenience facilities

Category	Sub-categories
Cultural facility	Cinema, Concert hall, Theatre, Museum, Exhibition hall, Convention &Exhibition center, Art gallery, Library, Science & Technology museum, Planetarium, Culture place, Archives hall
Hospital	AAA class hospital
Sport facility	Sports center, Bowling Hall, Tennis court, Basketball Stadium, Football field, Skating rink, Natatorium, Gym center, Table tennis hall, Pool room, Squash court, Badminton court, Taekwondo venue
Convenience facility	Restaurant, Convenience store, Hospital and Clinic, Bank,
Employment	Provincial Level Government and Institution, Prefecture Level Government and Institution, District & County Level Government and Institution, Company

For the natural category, city parks are associated with the need for residents for entertainment and leisure. I calculate the proximity to the park as a factor related to the landscape. Besides, the Normalized difference vegetation index (NDVI) of each house within a radius of 500m for quantificational analyzing the environment difference of each neighborhood was calculated, following the formula below:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

Where NIR is near-infrared (NIR) band, and RED is the red band.

For the commercial facilities aspect, shopping is an important activity of residents in daily life. With the number of shopping centers increasing, it has become an integral part of residential life. The quality of the shopping mall is evolving as a crucial aspect of influencing housing value. According to the standard Classification of Retail Formats(*retail formats - MBA*, n.d.), it classified the shopping center into three levels as shown in Table 4. 4. This work chooses the regional shopping center as a potential factor, which contains multiple commercial activities or called one-stop shopping, large parking lots, and higher quality in the shopping environment and after-sales service aspect. In terms of the supermarket, the chain supermarkets including Carrefour, Wal-Mart, METRO AG and China Resources Vanguard as possible factors were selected, because they offers a wide variety of food, beverages, and household products with a relatively lower price. In addition, through reading related studies, the distance to CBD and the distance to government were also included as potential variables.

Table 4. 4: The classification of the shopping center.

Level	Location	Area
Community shopping center	The regional commercial center of the city	<=50000 m^2
Regional shopping center	The commercial center of the city	<=100000 m^2
Super-regional shopping center	the outskirts of the city	>100000 m^2

All of the selected variables using geospatial data are calculated in ArcGIS and the analysis procedure as shown in Figure 4. 5. More specifically, consider the multilevel effects, all of them are calculated using two methods: (1) distance from each house to the nearest facility; (2) the number of facilities within a buffer with a radius of 500m. In Figure 4. 5, using an example to illustrate these two types of methods.

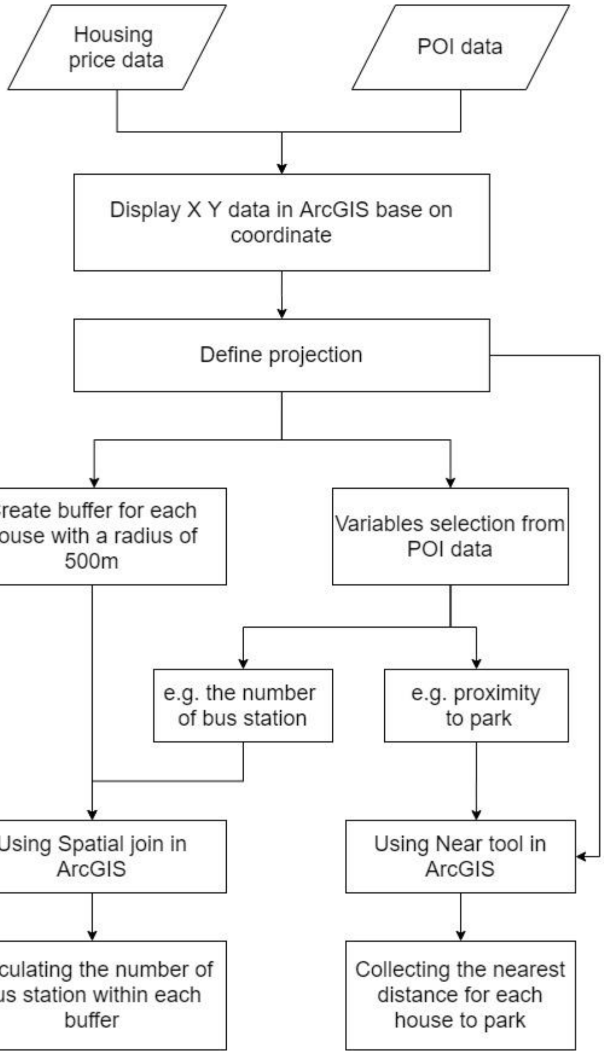


Figure 4. 5: the analysis procedure of geospatial variables

4.2.4. Street visual feature characteristics

The last characteristic is street visual features. There are many methods for semantic segmentation such as DeepLabv3, SegNet, Atrous Convolution, etc. DeepLab-v3 + model was developed by Google, which was added a decoder module for refining segmentation results based on DeepLabv3. (Dataset Overview – Cityscapes Dataset, n.d.). The authors replaced the existing model with depthwise separable convolutions in neural computer vision architectures and showed a large improvement on the JFT dataset compared to Inception V3(Chollet, 2017). In addition, Google shared its TensorFlow model training and evaluation code, as well as models pre-trained on the Cityscapes benchmark semantic segmentation task. The performance of the test set on Cityscapes reached 82.1% (Chen et al., 2018).

Cityscape is a dataset focusing on the understanding of urban street scenes and contains large-scale stereo video sequences that recorded street scenes of 50 cities. It includes 5000 image frames with high-quality pixel-level annotation (2975 for training, 500 for validation, and 1525 for testing, respectively) and 20000 frames with weakly annotated (Cheng et al., 2019)(*Cityscapes Dataset – Semantic Understanding of Urban Street Scenes*, n.d.). Besides, it provides 8 types of 30 categories of semantic levels, instance levels, and dense pixel annotations (including flat surfaces, people, vehicles, buildings, objects, nature, sky, and space)(*Dataset Overview – Cityscapes Dataset*, n.d.).

Therefore, in this work, the pre-trained model to analyze street view images was used without manually annotating the visual features ant re-training the model parameters.

Based on the understanding of the study area and the category of fine annotation in Cityscape, this work selects 7 categories with 17 sub-categories street visual features in total to analyze the relationship with housing prices, as shown in Table 4. 5.

Table 4. 5: the list of visual features extracted from street view image

Category	Sub-category
Flat	Road, sidewalk,
Human	Person, rider
Vehicle	Car, truck, bus, train, motorcycle, bicycle
Construction	Building, wall, fence
Object	Pole, traffic sign, traffic light
Nature	Vegetation, terrain
sky	sky

The procedure of visual feature extraction shown in Figure 4. 6. More specifically, in Figure 4. 6 the calculation process, considering the number of street view images within each buffer are diverse, at the same time, to alleviate the impact caused by the problem of perspective is shown. This work calculates the average area of each feature within a house buffer (radius of 500m).

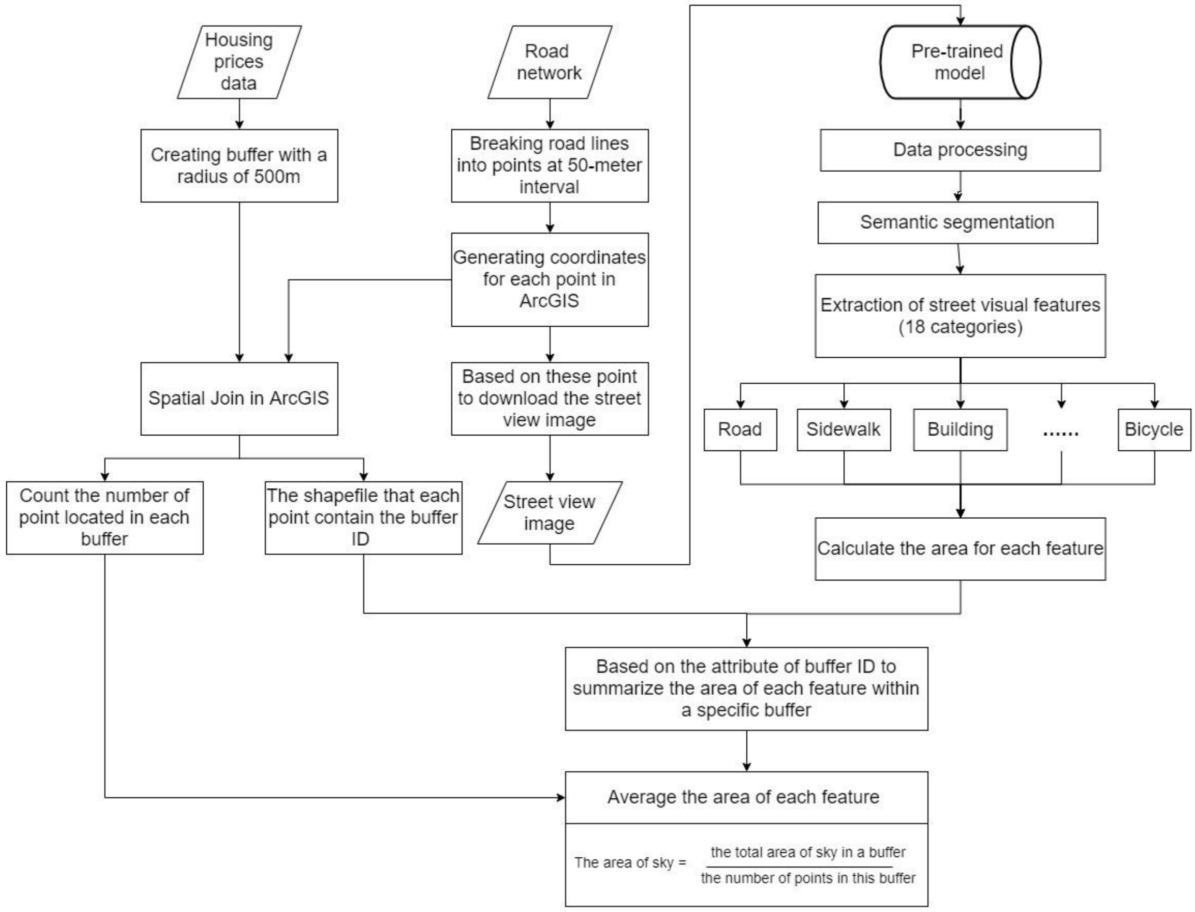


Figure 4. 6: Street view image processing and visual features extraction

4.3. Machine-learning algorithms / Random forests

Random forests are a combination of tree-structured classifiers proposed by **Breiman (2001)**. This algorithm is characterized by a supervised learning algorithm and evolves from a set of decision trees. Random Forest combines the two concepts of bagging and Random Selection of Features. The former is a sub-class of ensemble machine learning that contains some weak models and then merge the predictions of each weak model to make more accurate predictions. Bagging is used with decision trees. It designed to eliminates the challenge of overfitting and raise the stability of the regression or classification model in the reduction of variance and increase the accuracy. The latter refers to the Bootstrap sampling method that is a resampling method. More specifically, it independently takes samples with replacement from the original dataset with the same sample size.

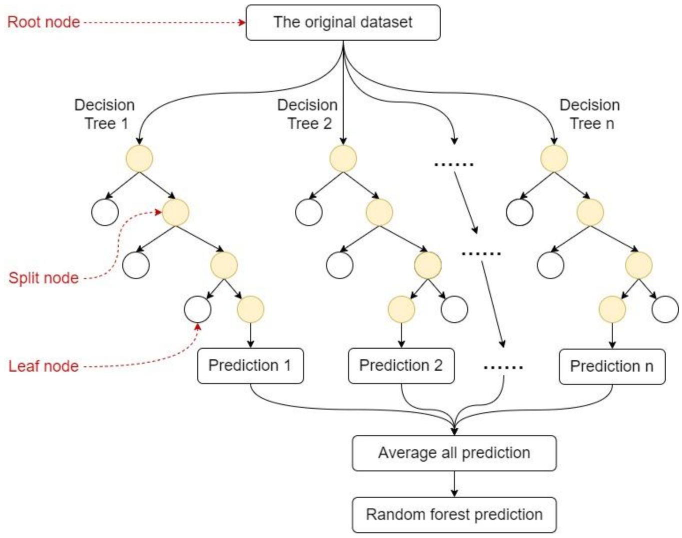


Figure 4. 7: Random Forest Structure

Figure 4. 7 shows the structure of a random forest. A random forest starts with the independent selection of many bootstrap samples from the original dataset, which further as the input dataset of each decision tree. Notice that the decision tree consists of the root node, split node, and leaf node. The root node is the topmost decision node that represents the entire sample. Split node is a process of dividing a node into two or more sub-nodes in which the certain feature used for the split based on a particular criterion, for example, Gini (for classification) or sums of squares (for regression) from the entire data set. Besides, the features considered for splitting at each node are a random subset of the previous set of features. The leaf node or terminal node contains a small subset of the observations at the end of branches and not split. In RF, each decision tree is different because the attribute selection for each node is randomized. Thus, the Random Forest models strive to reduce the generalization error of the decision tree model. Besides, to ensure that decision trees not heavily rely on a certain feature and use all potential predictive features fairly, The number of features that can be split on at each node is limited to some percentage of the total (which is known as the hyperparameter).

Steps for constructing random forest:

1. Assuming that there are N samples. Randomly N samples are taken repeatedly from the original data, in which each data element has an equal probability been selected and the size of each sample is the same as the original training set. The N samples are the input data at the root node of each decision tree.

2. If each sample has M variables, at the split node of a decision tree, randomly selecting m variables from the sample where m much smaller than M. Then based on a certain strategy such as information gain to determine the optimal node splitting feature.

3. During the growing process of a decision tree, based on step 2, selecting the best variable for each node of the decision tree.
4. Follow steps 1 ~ 3, building many decision trees will form a random forest.

One appealing feature of the RF model is that it does not require an external validation such as K-fold cv, withheld test. In RF, each tree is constructed using a bootstrap sample from the original dataset. The experiment found that the bootstrap samples contain about 66% of the features of the original dataset. There are 33% features of original data that are not used to train the RF model, which called the Out-Of-Bag (OOB) sample. The OOB data can be used to evaluate the performance of the decision tree in RF and calculate the prediction error rate of the model, which is called the out-of-bag error. In addition, OOB also used to estimate the importance of each determinant.
The steps of importance calculation:

1. Assuming that there are B decision trees in RF, for each decision tree, using the corresponding out of bag (OOB) data make a prediction and calculate the error, denote as error1, error2, ⋯, $error_b$.
2. The variable X is randomly permuted in b OOB samples to form new OOB samples. Using the RF make a prediction on the new OOB samples and calculate the error, denote as ERROR1, ERROE2,⋯, $ERROR_b$.
3. The importance of the determinant(X) is:

$$Importance(X) = \frac{\Sigma error_i - ERROR_j}{B}$$

Random forest not only can measure the importance of variables but also can calculate the partial dependencies that are the influence of input variables on the response variable. This method was been proposed by (statistics & 2001, n.d.). The formula is as follow:

$$\bar{f}(X_s) = \frac{1}{n}\sum_{i=1}^{n} f(X_s, x_i c)$$

Where X is the input data; $f(x)$ is the RF predictor model; Xs is a subset of the input dataset; χ_{ic} is the value of samples in Xc in the training data.
Besides, interpretation via variable importance measures and partial dependence plots has become a popular data mining tool and widely applied by many scholars(Berk et al., 2009; Girardello et al., 2010). In RF, a partial dependence plot gives a graphical depiction of the nonlinear dependencies between an input variable and the prediction, some with an apparent variable threshold. It shows how the predictions partially depend on the value of the input variable.

4.4. Model evaluation

Model evaluation is to evaluate the fitness of predicted prices with observed prices and estimate the generalization accuracy of a model on prediction data. The following statistical metrics are used to measure the RF model error.

1. R-squared/Adjusted R-Squared

R-squared (R2), or the coefficient of determination, is a statistical measure that represents the proportion of the variance in the dependent variable that is predictable from the independent variable or variables. The value of R-squared is range from 0 to 100% (or 0 to 1). Better the model, the higher the R-squared value. if the value equal to 100% or 1 means the best model with all correct predictions. The formula for R-Squared is given by:

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2}$$

Where y_i is the observed price; f_i is the predicted prices; $\bar{y}$ is the mean of the observed prices.

2. Root mean squared error (RMSE)

Root Mean Square Error (RMSE) is a measure of the differences between predicted prices and the observed prices. It is the standard deviation of the residuals and the RMSE metric is given by:

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - f_i)^2}$$

Where m is the total number of observations.

3. Mean Absolute Error (MAE)

Mean absolute error (MAE) is a measure of errors between predicted prices and observed prices, without considering the direction. It reflects the closeness of prediction value and observation value. The formula for MAE is as follows:

$$\frac{1}{m}\sum_{i=1}^{m}|(y_i - f_i)|$$

4. Pearson correlation coefficient

The Pearson correlation coefficient, or Pearson' R, is a method used to measure the correlation between two variables X and Y. Its value range between -1 and 1, where 1 is a completely positive correlation, and 0 is no correlation and -1 is a negative correlation. The formula is:

$$Pearson\ R = \frac{\sum_{i=1}^{n}(f_i - \bar{f})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(f_i - \bar{f})^2}}$$

Where $\bar{f}$ is the mean of the predicted prices.

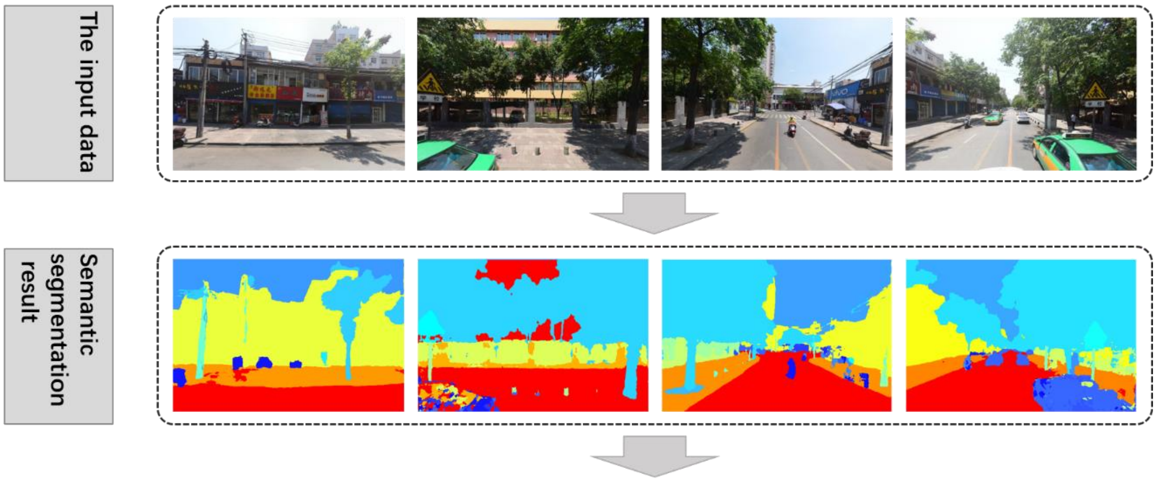
The building and analysis of housing prices estimation model were carried out with R, using the package RandomForest.

5. RESULT AND DISCUSSION

This chapter provides the results and discussion of the RF model. Section 5.1 illustrates sample results of street scene features extraction. Section 5.2 is a description of potential indicators selected from the structure, location, neighborhood environment, street visual features four aspects. Section 5.3 shows the correlation analysis and removes the variables having a high correlation. Section 5.4 elaborates on the results of the housing prices estimation model with three contrast experiments, feature selection in RF. Finally, section 5.5 is the limitation of this work.

5.1. The results of visual features extraction

Figure 5. 1 presents the visualization of the result of a sample street view image data. These four images are collected from one coordinate point. Semantic segmentation results show each visual feature such as sky, road, tree, etc. the statistical result of each feature area in the image is shown at the end.
It can be observed that the road, sky, sidewalk, and tree are the major features in each image. At the same time, motorcycles and cars are identified. In this work, there are 279663 images with a size of 850×680 pixels in total. It spent about one week for visual element segmentation.



Statistical analysis

road	0.158660963	road	0.105463	road	0.196006	road	0.156193
sidewalk	0.113625668	sidewalk	0.109153	sidewalk	0.103422	sidewalk	0.071012
building	0.423386096	building	0.082081	building	0.128036	building	0.118205
wall	0	wall	0.055955	wall	0.020605	wall	0.024045
fence	0	fence	0.117591	fence	0.021621	fence	0.006676
pole	0.015886631	pole	0.005906	pole	0.006631	pole	0.005099
traffic_light	0	traffic_light	0	traffic_light	0.000261	traffic_light	0
traffic_sign	0.001790374	traffic_sign	0.015653	traffic_sign	0.000798	traffic_sign	0.013386
vegetation	0.088211765	vegetation	0.488402	vegetation	0.291022	vegetation	0.404764
terrain	0	terrain	0	terrain	0.005897	terrain	0
sky	0.182660963	sky	0.002631	sky	0.200021	sky	0.10917
person	9.63E-05	person	0.001867	person	0.00468	person	0.005005
rider	0	rider	0	rider	0.000699	rider	0.001433
car	0	car	0.007677	car	0.006094	car	0.071042
truck	0	truck	0.006347	truck	0.000907	truck	0.008571
bus	0	bus	0	bus	0.003658	bus	2.14E-06
motorcycle	0.015328342	motorcycle	0.000116	motorcycle	0.006661	motorcycle	0.004958
bicycle	0.000352941	bicycle	0.001157	bicycle	0.001925	bicycle	0.000439

Figure 5. 1: The procedures of visual feature extraction

5.2. Description of the variables

In this study, to comprehensively estimate the housing prices in Xi'an, reviewing related literature, 44 potential indicators are selected from the structure, location, neighborhood environment, street visual features four aspects. Table 5. 1 shown these indicators with detailed statistics. To summarize, in my dataset, an average dwelling has 3 bedrooms, 2 bathrooms, a lift, owing good decoration, is located at the middle storey level in a residential building.

Table 5. 1: the list of selected indicators with respective descriptive statistics

Characteristic	Category	Variables	Max	Min	Mean	SD
Structure	Housing structure	The number of bedrooms	9	1	2.42	0.98
		The number of bathrooms	9	1	1.41	0.61
		Floor level	3	1	2.06	0.81
		Building height	40	1	23.85	9.06
		Type of building structure	4	1	2.02	0.71
		decoration	4	1	2.03	1.13
		Housing units	48	1	7.54	6.98
		elevator	1	0	0.94	0.24
Location	Transportation infrastructure	Proximity to metro station	15711	93	6139	3804.94
		Density of bus station	99	0	20.22	15.57
		Distance to highway entrance	9734.60	235.50	4650.20	2175.90
		Distance from highway	9734.63	2.33	4391.60	2293.58
Neighborhood	Educational facilities	Distance to university	9941.00	22.46	2587.15	1948.94
		Density of kindergarten	11	0	2.72	2.203
		The number of primary and middle school	8	0	0.94	1.116
		School house	1	0	0.029	0.17
	Convenience facilities	Density of cultural facilities	18	0	0.964	1.85
		Proximity to hospital	7730.56	19.88	1655.26	1045.134
		Density of sport facility	111.00	0.00	17.27	14.47
		Density of convenience facilities	1046	0	194	143.14
		Employment	1419.00	0.00	109.10	139.50
	Nature amenities	Proximity to park	3170.00	23.80	927.20	547.128
		NDVI	0.18940	-0.04520	0.03184	0.02697
	Commercial felicities	Proximity to shopping mall	4375.18	24.11	966.72	613.05
		Proximity to supermarket	7138.00	22.45	1455.66	1062.56
		Distance to CBD	13802.00	373.80	6870.10	2300.33
		Distance to government	12466.08	36.76	3793.66	1952.17
Street view features	Flat	Road feature	0.28275	0.06459	0.18621	0.02864
		Sidewalk	0.06927	0.00950	0.02869	0.00811
	Human	Person	0.032652	0.00018	0.00376	0.00256
		Rider	0.002361	0.000023	0.00057	0.00036

			76	62	
Construction	Building	0.5610	0.0120	0.2322	0.07692
	Wall	0.08306	0.00057	0.01459	0.00958
	Fence	0.047467	0.003511	0.01678	0.00726
Object	Pole	0.012559	0.002779	0.00656	0.00149
	Traffic light	0.00127	0.00001172	0.000257	0.00019
	Traffic sign	0.00637	0.00015	0.00155	0.00059
Nature	Tree	0.53721	0.02611	0.25344	0.07459
Sky	sky	0.45032	0.02429	0.18517	0.06730
Vehicle	car	0.090567	0.00235	0.04556	0.01427
	truck	0.10517	0.00026	0.00649	0.00725
	bus	0.02395	0.00003121	0.003556	0.00345
	motorcycle	0.009559	0.00006091	0.001324	0.00112
	bicycle	0.008816	0.00005407	0.00164	0.00109

5.3. Correlation analysis

Pearson's correlation is a technique for measuring the degree of association between two variables. The value of Pearson's correlation ranges from -1 to 1. The sign of value indicates the direction of the correlation and the magnitude of the value is the strength of the correlation. More, specifically, 1 means these two variables are completely correlated, while -1 means there is no association between the two variables. A high correlation means the two variables reflect the same phenomenon and using one of them can explain variations.

Based on the results of Pearson's correlation coefficient (Figure 5. 2), two variables are highly correlated (P= 0.98639): the distance to the highway and the distance to highway exit/entrance. Therefore, I delete the variable of the distance to the highway.

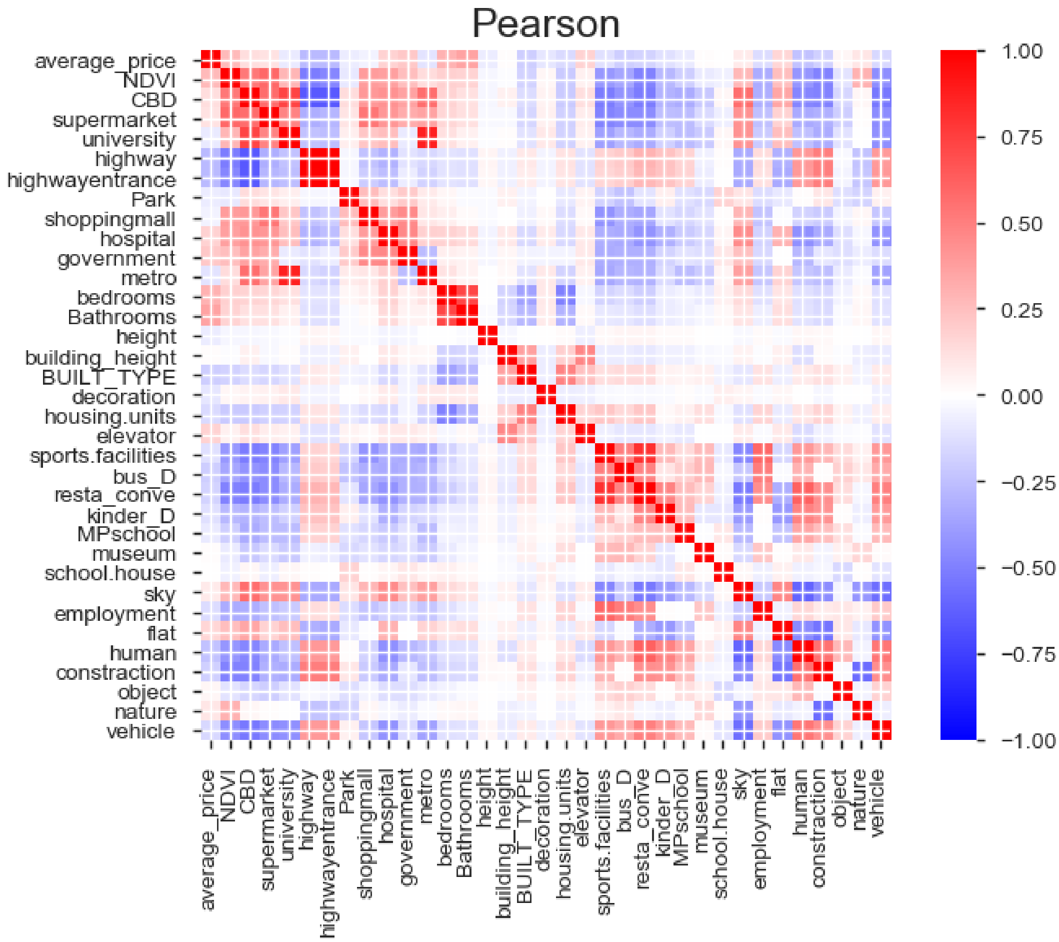


Figure 5. 2: the result of Pearson correlation analysis

5.4. Result of random forest

In order to explore the reliability and superiority of different datasets in housing prices estimation, this work considers three sets of experiments: the first experiment only uses the geospatial data; the second experiment only uses street view image data; the third experiment uses geospatial data and street view image data. For each experiment, 70% of the original data was randomly selected for training (23938 records), and the remaining 30% data are validation set (10260 records). The details of each experiment are as follows:

5.4.1. Model 1: only use geospatial data

This model only uses geospatial data source to explore what extent the traditional variables can explain the variance of housing prices.

Parameters play a central role in building modeling since the optimized parameters can improve the accuracy and control the capacity of the model. In RF, "mtry" and "ntree" are important hyperparameters. "mtry" is the number of variables randomly selected from original data to be sampled at each split. "ntree" refer to the number of decision tree contained in RF. In RF, the default value of "ntree" is 500 and the default value for mtry is about P/3 for regression, where p is the number of features. In most cases, the default values of parameters are not the best choice for getting a higher accuracy model. Figure 5. 3 shows that when the “mtry” value equals to 7, the corresponding “rsq” value is highest. According to the documentation of Package ‘randomForest’ in R language. “rsq” is pseudo-R-squared and the formula is:

$$rsq = \frac{1 - mse}{var(y)}$$

Where “mse” is the mean square error of OOB predictions with targets, vary(y) is the variance of targets. “rsq” value reflects how well the model explains the data.

Figure 5. 4 shows that when “ntree” closing to 200, the change of the error tends to be stable. In this experiment, I set “mtry” =7 and “ntree” = 200.

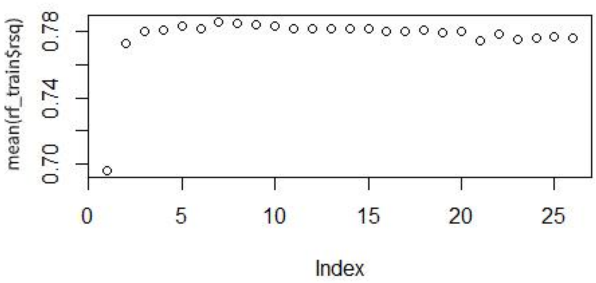


Figure 5. 3: The “mtry” screening of the RF model

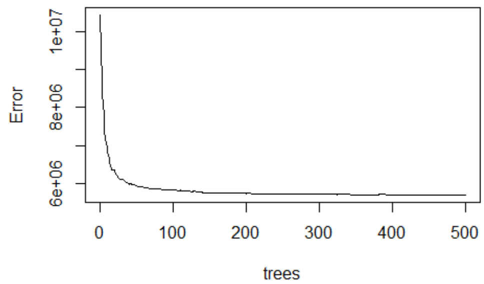


Figure 5. 4: The “ntree” screening of the RF model

Running the RF model and the result are obtained as follows: Mean of squared residuals: 5776218 and % Var explained: 78.09. “%explained variance” reflects the degree that OOB predictions explained the target variance of the training dataset.

The performance of the first experiment RF model is shown in Table 5. 2. It has a good performance on the training dataset with the R-square 0.91, and 0.80 on the validation dataset. It indicates that about 91% variance in the change of housing prices that the selected variables explain collectively. The correlation between the observed prices and the predicted prices in the training datasets is about 0.96 and 0.89 in the validation dataset. the value of RMSE is 1534.89 on the training dataset and 2253.647 on the validation dataset. Similarly, the MAE is 965.65 and 1438.24 on training data and validation data, respectively. RSE on training data and validation data are 0.089 and 0.197. Figure 5. 5 and Figure 5. 6 visually highlight the relationship between predicted value and real housing prices for training data and validation data.

Table 5. 2: The results of model performance evaluation

Matrix	**Results（Training data）**	**Results（Validation data）**
R-square	0.9106426	0.8014389
RMSE	1534.89	2253.647
RSE	0.0893574	0.1985611
MAE	965.6463	1438.24
Pearson R	0.95691	0.8896687

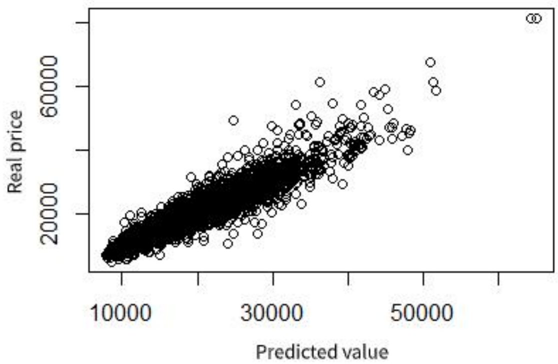


Figure 5. 5: Scatter plot of the predicted value vs. real price for training data

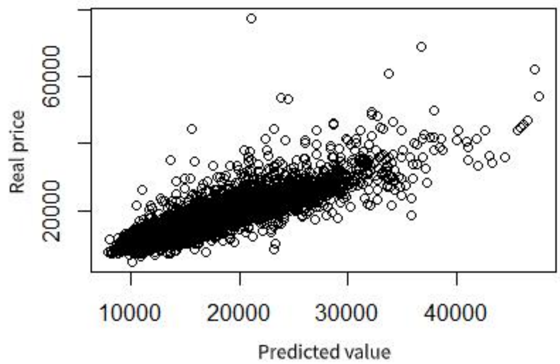
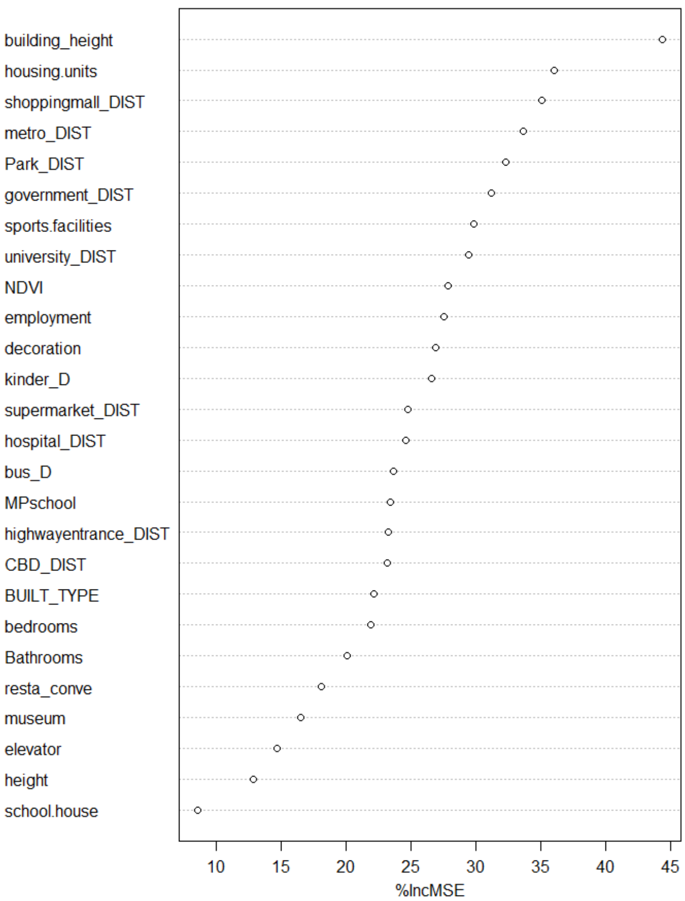


Figure 5. 6: Scatter plot of the predicted value vs. real price for validation data

The importance of each variable is estimated as shown in Figure 5. 7 by calculating how much the error increases based on OOB data randomly permuted each variable while keeping other variables unchanged. Combine with Table 5. 3, in this experiment, the most important first five variables are: build height, housing units, the proximity to shopping mall, the proximity to metro station, the proximity to the park. By contrast, schoolhouse, height, elevator, culture facilities, and convenience facilities are relatively unimportant.



Variables	%IncMSE
building_height	44.38943
housing.units	36.06009
shoppingmall_DIST	35.09066
metro_DIST	33.65729
Park_DIST	32.30542
government_DIST	31.20697
sports.facilities	29.8776
university_DIST	29.4849
NDVI	27.86757
employment	27.59722
decoration	26.93453
kinder_D	26.6179
supermarket_DIST	24.81428
hospital_DIST	24.5899
bus_D	23.70863
MPschool	23.44576
highwayentrance_DIST	23.25583
CBD_DIST	23.23571
BUILT_TYPE	22.18611
bedrooms	21.9515
Bathrooms	20.14202
resta_conve	18.12261
museum	16.56431
elevator	14.69965
height	12.87228

school.house	8.638787

Figure 5. 7: The importance ranking of selected variables

Table 5. 3: The importance of selected variables

5.4.2. Model 2: only using street view image data

This model only uses street view image data source to explore what extent the street visual features can explain the variance of housing prices. It observed from Figure 5. 8 and Figure 5. 9, when the “mtry” is 6, the “rsq” is highest, and when the “ntree” value close to 200, the error rate becoming flat. Therefore, in this experiment, I set the two parameters value to “mtry” = 6 and “ntree” = 200.

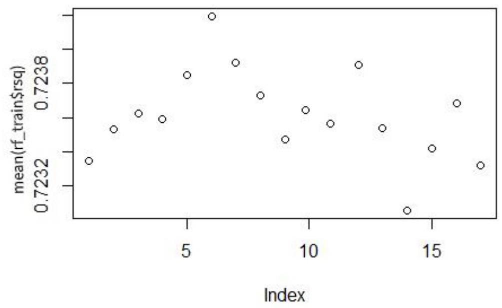


Figure 5. 8: The “mtry” screening of RF model

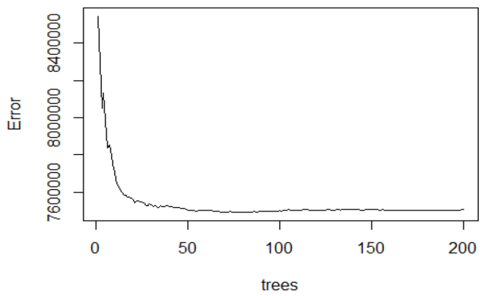


Figure 5. 9: The “ntree” screening of RF model

In this experiment, the RF model generated the results are: Mean of squared residuals is 7504173, and % Var explained is 71.54.

The evaluation results of model performance are shown in Table 5. 4. The R-square of this model is approaching 0.75 on the training dataset and 0.73 validation dataset. Pearson R is 0.87 and 0.86 on the training dataset and validation dataset, respectively. It indicated that there is a strong association between predicted prices and observed prices. Besides, RMSE, RSE, and MAE on the training dataset are 2552.163, 0.247054, 1585.99, respectively. Similarly, the corresponding value is 2590.168, 0.2622881, 1680.733 on the validation dataset. Figure 5. 10 and Figure 5. 20 and visually highlight the relationship between predicted value and real housing prices for training data and validation data.

Table 5. 4: The results of model performance evaluation

Matrix	Results（Training data）	Results（Validation data）
R-square	0.752946	0.7377119
RMSE	2552.163	2590.168
RSE	0.247054	0.2622881
MAE	1585.99	1680.733
Pearson R	0.86695	0.859408

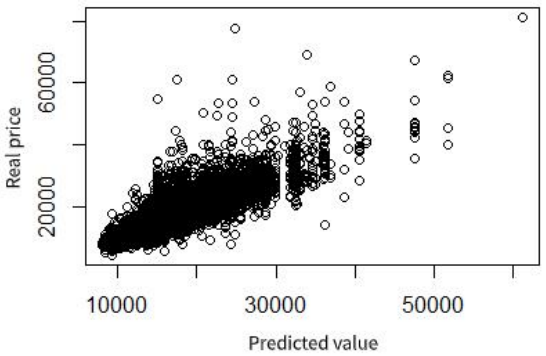


Figure 5. 10: Scatter plot of the predicted value vs. real price for training data

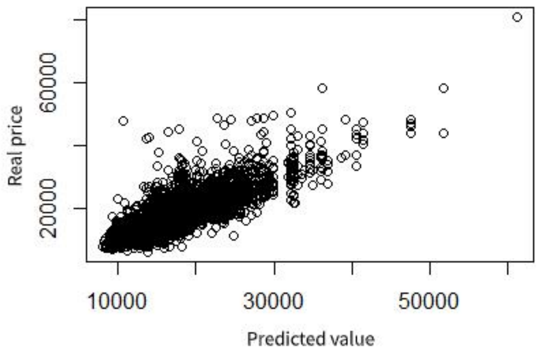


Figure 5. 11: Scatter plot of the predicted value vs. real price for validation data

Similarly, the importance of street visual features is estimated and shown in Table 5. 5 and Figure 5. 12. It shows that the traffic light, bus, pole, fence, and truck are the most important first five variables. The features person, rider, car, motorcycle, and traffic sign are relatively less important for housing prices.

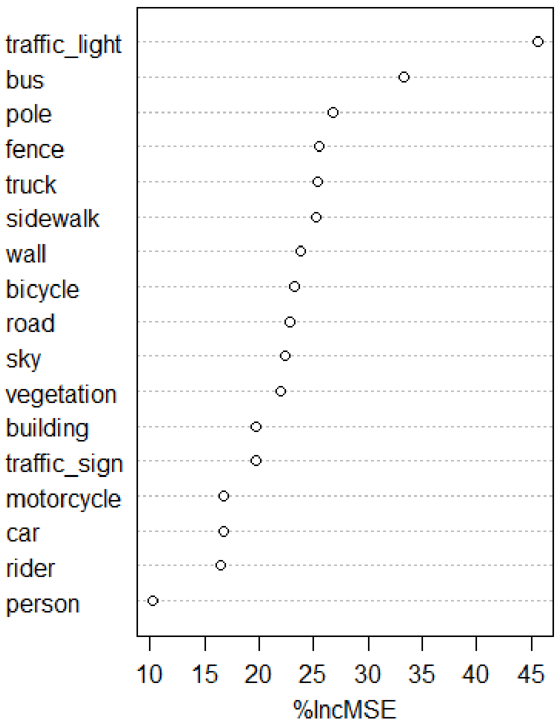


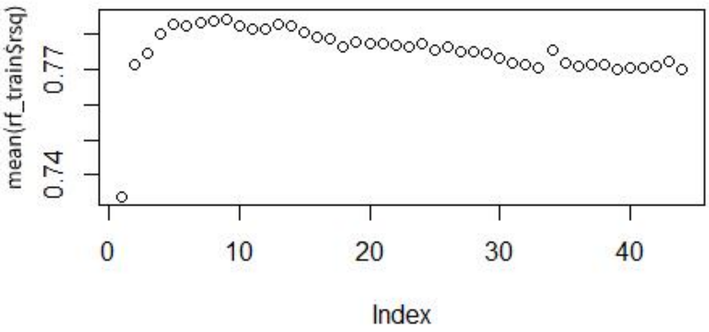
Figure 5. 12: The importance ranking of selected variables

Variables	**%IncMSE**
traffic_light	45.57143501
bus	33.30521642
pole	26.78953946
fence	25.53845079
truck	25.34145707
sidewalk	25.17431533
wall	23.75103353
bicycle	23.18697229
road	22.75117426
sky	22.44820907
vegetation	21.99672064
building	19.65845519
traffic_sign	19.65636973
motorcycle	16.79082207
car	16.76928884
rider	16.47081594
person	10.26533144

Table 5. 5:The importance of selected variables

5.4.3. Model 3: using geospatial data and street view image data

This model combines two types of data source, geospatial data source and street view image data source. As shown in Figure 5. 13 and Figure 5. 14, the best optimal “mtry” is 9 and “ntree” is 300



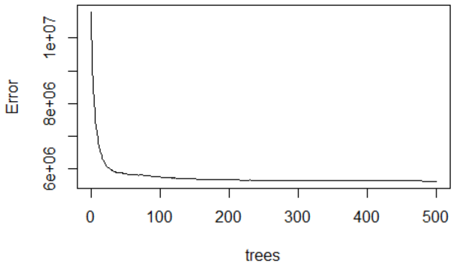


Figure 5. 13: The “mtry” screening of the RF model

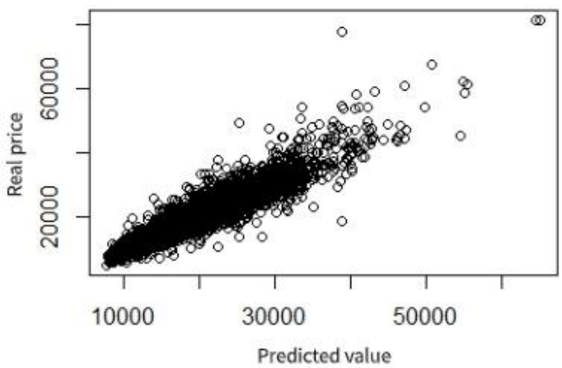
Figure 5. 14: The “ntree” screening of the RF model

This model was trained upon the dataset that combines geospatial data and street view image data. The results are as follows: Mean of squared residuals is 5647304 and % Var explained is 78.58.

As a result of the evaluation of model ability (shown in Table 5. 6), R-square is 0.900 and 0.803 on training data and validation data. Pearson R for the training set is about 0.95, for the validation set is about 0.89. The RMSE for training data is 1618.898, the RSE is 0.0994065 and MAE is 1019.536. The validation RMSE is calculated as 2243.994, RSE is 0.1968638, and MAE is 1427.505. Figure 5. 15 and Figure 5. 16 visually highlight the relationship between predicted value and real housing prices for training data and validation data.

Table 5. 6: The performance of the third experiment

Matrix	**Results（Training data）**	**Results（Validation data）**
R-square	0.9005935	0.8031362
RMSE	1618.898	2243.994
RSE	0.0994065	0.1968638
MAE	1019.536	1427.505
Pearson R	0.9504118	0.8963358



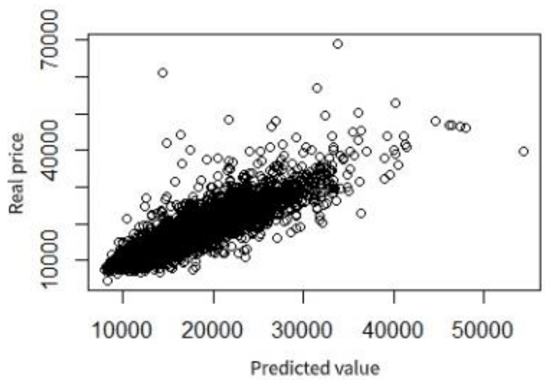


Figure 5. 15: Scatter plot of the predicted value vs. real price for training data

Figure 5. 16: Scatter plot of the predicted value vs. real price for validation data

The OOB estimates of error rate were used for measuring the importance of variables. The random forest model is also able to rank the order of variables based on the importance. For this model, the importance of each variable is presented in Figure 5. 17and Table 5. 7. The first five important variables are building height, housing units, truck, decoration, and proximity to the park. By contrast with the first experiment that only uses geospatial data, the biggest change in variable importance aspect is the proximity to metro station(ranked 4th in the first experiment) and the proximity to shopping(ranked 3rd) mall are ranking 6th and 9th, respectively. For street scene feature respect, the truck is more important and ranks 3rd. besides, the feature fence, vegetation, pole, traffic light, and traffic sign are ranking 11th, 12th,13th,14th, 15th, respectively out of 44 variables in total.

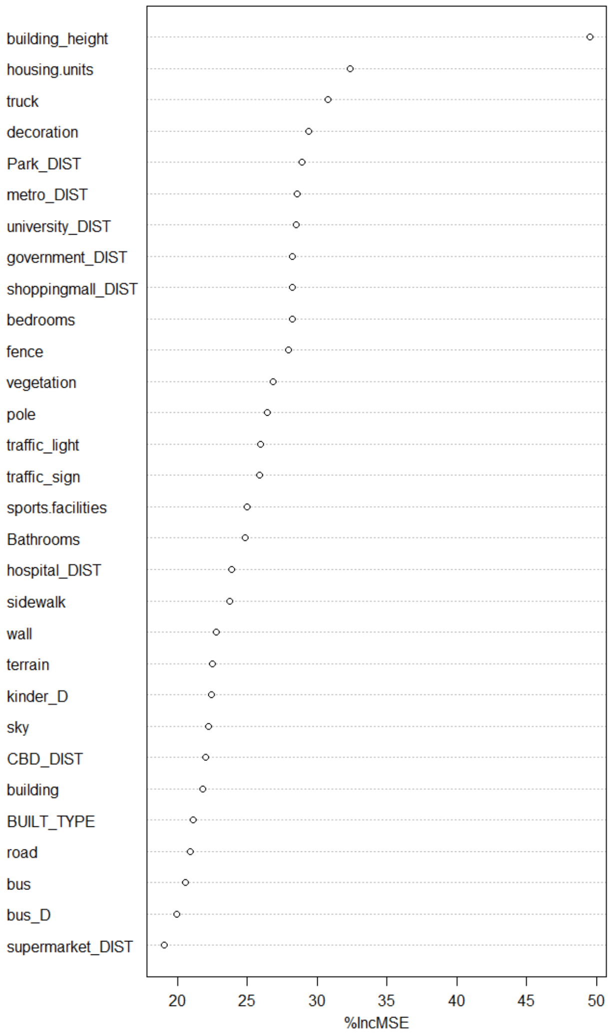


Figure 5. 17: The importance ranking of selected variables

Variables	%IncMSE
building_height	42.73553003
housing.units	31.08833717
decoration	29.3603375
government_DIST	28.5260942
fence	27.89315511
university_DIST	27.27874754
shoppingmall_DIST	27.04543132
truck	26.8084923
traffic_light	26.77251378
bedrooms	25.89240808
bus	25.04801241
car	24.7313876
metro_DIST	24.46205242
BUILT_TYPE	24.02072631
Park_DIST	23.992307
wall	23.8227124
vegetation	23.34522327
hospital_DIST	23.03580052
supermarket_DIST	23.02039231

Table 5. 7: The importance of selected variables

5.4.3.1. Model 4

Feature selection is a process that to select the most effective features from the original features to reduce the dimension, improve the model generalization ability and reduce the overfitting. The main purpose is to remove irrelevant features and redundant features and select the optimal feature subset. This model is using the new features/the most effective features to improve the performance of the property valuation model.

Through reading-related studies, there are no uniformed strategies for feature selection in the RF model. (Díaz-Uriarte & Alvarez de Andrés, 2006) proposed a strategy: first, using the original variables to build an RF model and compute each variable importance. Second, eliminating 20% variables having the smallest importance and build a new RF model using the remaining variables. Repeating the second step until select a set of variables that carry out the smallest OOB error rate.(Ghattas & Ben Ishak, 2008)adopted an ascendant strategy. First, computing the variable importance based on SVM. They then built a sequence of SVM models called k of the most important variables from step 1. Finally, they select a set of variables that led to the model with the lowest error rate. In this work, I use the strategy proposed by Díaz-Uriarte & Alvarez de Andrés to eliminate 20% of variables having the smallest importance. The feature selection is conduct at the basis of the model that uses geospatial data and street view image data and based on Figure 5. 17, 9 variables are selected. It includes the Density of bus station, Density of kindergarten, Density of convenience facilities, Motorcycle, Person, Density of cultural facilities, The number of primary and middle school, Rider, School house.

A new RF model is build using the remaining variables. According to Figure 5. 18 and Figure 5. 19, the parameters "mtry" and "ntree" are setting to 9 and 300, respectively.

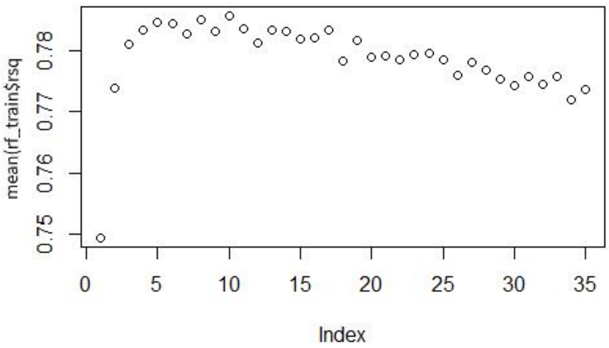


Figure 5. 18: The "mtry" screening of RF model

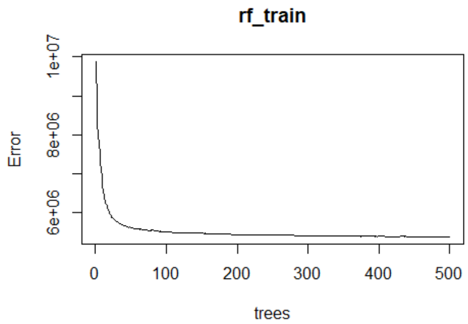


Figure 5. 19: The "ntree" screening of RF model

The new model generated the result as follows: Mean of squared residuals is 5412849 and % Var explained is 78.96. The evaluation for model ability is shown in Table 5. 8. The R-square is 0.91 for training data and 0.79 for validation data. Pearson is calculated as 0.96 on the training set and 0.89 on the validation set. Besides, RMSE is 1522.938 and 2411.592 for training data and validation data. RSE is calculated as 0.090174 on training data, 0.214804 on validation data. MAE is 976.4011 and 1431.126 for training and validation data.

By comparing with the original model, it observed that the new model has similar R-square, Pearson R, and RSE for both training data and validation data. RMSE and MAE of the new model are lower than the original model, which indicate the new model has a lower deviation between the predicted prices and the observed prices. Figure 5. 20 and Figure 5. 21 visually highlight the relationship between predicted value and real housing prices for training data and validation data.

In addition, stepwise regression model has been built. The performance of hedonic and random forest models is compared in Figure 5. 8. The R-square of stepwise regression model is 0.40, RMSE is 3956.745, RSE is0.599175, MAE is 2791.571 and Pearson R is 0.633107. The results report that the random forest perform better than linear model.

Table 5. 8: The performance of the third experiment

Matrix	**Results（Training data）**	**Results（Validation data）**	**Hedonic model**
R-square	0.909826	0.7851957	0.4008
RMSE	1522.938	2411.592	3956.745

RSE	0.090174	0.214804	0.599175
MAE	976.4011	1431.126	2791.571
Pearson R	0.955175	0.887549	0.633107

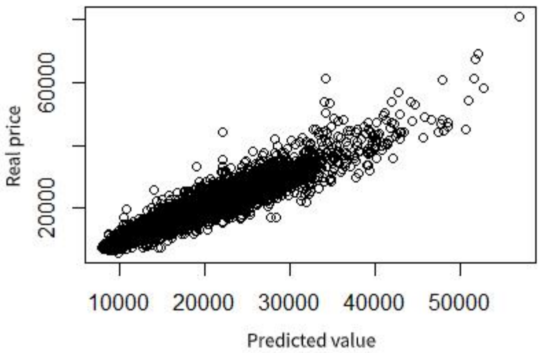


Figure 5. 20: Scatter plot of the predicted value vs. real price for training data

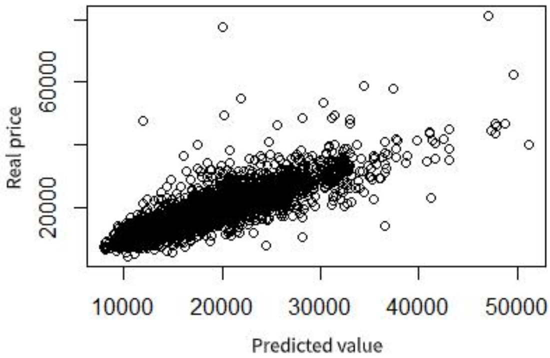


Figure 5. 21: Scatter plot of the predicted value vs. real price for validation data

The variable importance is shown in Figure 5. 22. The most two important variables are building height and housing unit. Besides, the importance of the number of bedrooms and decoration rank ninth and 12th, respectively. The number of bathrooms is standing 26th place. It is reflecting the housing structure characteristics have an important impact on housing prices. Further, in housing structure characteristic respect, building height and housing units are the crucial determinants that highly related to the comfort of a dwelling. For the location aspect, proximity to the metro station and the distance to highway exit/entrance is rank in 11th and 21st. Compared to buses, subways have a greater impact on housing prices. Highway exit/entrances have a negative effect on surrounding housing prices due to large traffic flow and negative environmental impact. For the education facilities, proximity to the university is ranking 10th. The university not only associated with social and cultural amenities but also provides a public place for recreation and cultural communication. In the aspect of convenience facilities, proximity to the hospital is the eighth most important variable. It indicates that buyers are sensitive to the convenience and quality of medical treatment. Besides, the importance of the density of sports facilities is 23rd and employment is on the 24th. For the natural attributes, proximity to urban parks and NDVI are ranking 5th and 15th. NDVI reflect the green area of the surrounding area of the dwelling unit. Natural amenities and environmental quality are highly associated with citizens' life satisfaction. It suggests that optimizing and building an attractive local landscapes should be considered in the future. The 10th variable is proximity to the university. For commercial facilities, the importance of proximity to the shopping mall stands in the 3rd place. It indicates that the premium provided by the shopping malls is larger than other variables. It also shows that commercial convenience plays an important role in the preference of home buyers. Proximity to the supermarket, distance to CBD are on the 17th and 20th. The importance of proximity to government ranks 4th. Usually, government agencies have an advantageous location with good infrastructure, convenient transportation, and facilities. In street visual feature respect, the importance of trucks and traffic lights ranks 6th and 7th. The importance of cars and traffic signs ranked 14th and 16th among 30 variables, while the building, terrain, wall, and road are relatively less important variables.

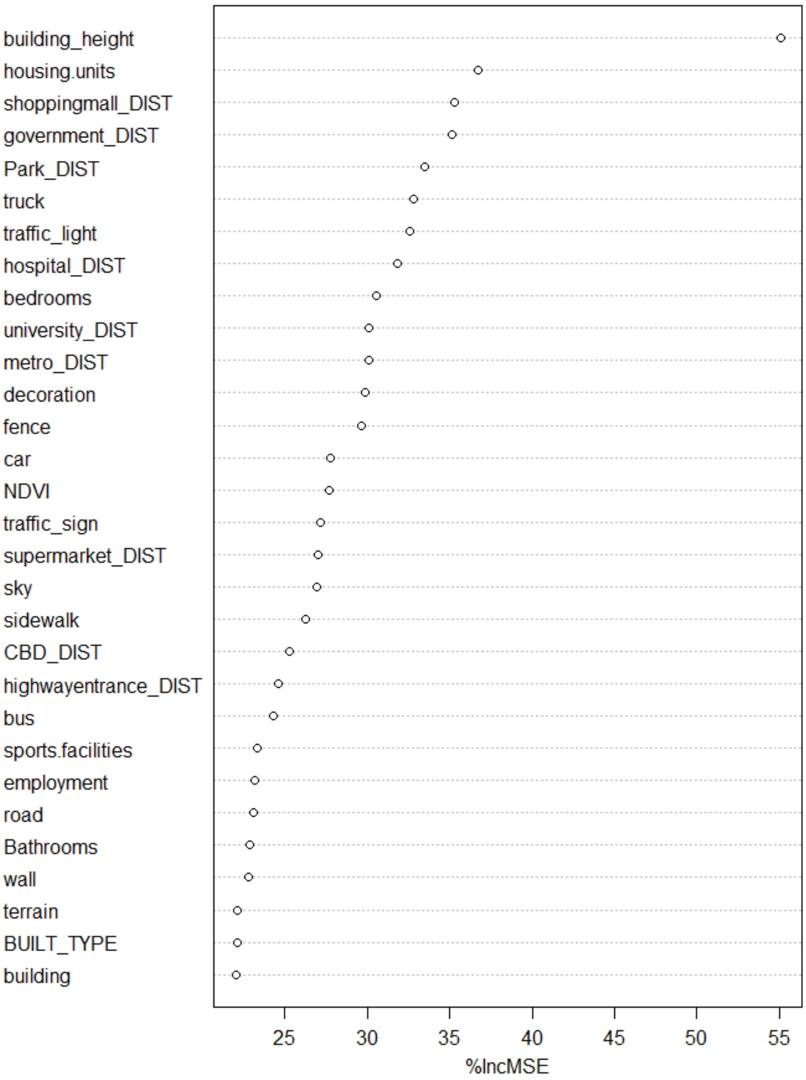


Figure 5. 22: The importance ranking of selected variables

Variables	%IncMSE
building_height	55.07867531
housing.units	36.72459815
shoppingmall_DIST	35.26443236
government_DIST	35.11899886
Park_DIST	33.45670414
truck	32.80153707
traffic_light	32.59421759
hospital_DIST	31.82328866
bedrooms	30.56273742
university_DIST	30.1105078
metro_DIST	30.06265404
decoration	29.85746588
fence	29.6119152
car	27.77565423
NDVI	27.67075559
traffic_sign	27.16057611
supermarket_DIST	26.97070193
sky	26.95416913
sidewalk	26.25093115
CBD_DIST	25.27401782
highwayentrance	24.5806602
bus	24.30465892
sports.facilities	23.3322669
employment	23.15444653
road	23.07439412
Bathrooms	22.86013727
wall	22.74917609
terrain	22.13879268
BUILT_TYPE	22.13660604
building	22.02115831
vegetation	21.54746859
bicycle	21.48416844
pole	16.23299573
elevator	15.31892596
height	13.45742095

Table 5. 9:The importance of selected variables

As shown in Figure 5. 23, the partial dependence plot reveals the presence of potential thresholds in the relationships between variables and housing prices. for each plot, the horizontal axis is the variable and the vertical axis is the housing prices when other predictors are held constant. The interpretation of the bivariate relationship between each variable and housing prices is given as follow (here I explained the first ten ranked variables):

The first plot in Figure 5. 23 present the nonlinear dependencies between building height and housing prices. Building height is the number of floors of the building that the house located in. It is apparent that the relationship between building height and housing prices is highly non-linear. The building with less than 3 floors usually is the villa. It is expansive than other types of residential buildings. The price of the multi-storied building is lower because it generally adopts the brick-concrete structure, which has a lower construction cost compared with other types. Another information reflected in the first plot is the average prices of the building with 10-30 floors are higher than the medium residential buildings and multi-storied buildings. The reason is that the construction cost is expensive in terms of the foundation, structure, fire protection. When the floor increase to the 30th floor, housing prices are decreasing. The possible explanation is related to residential population density. The more storey of a building, the more residents. It refers to a less comfortable living environment.

As a supplement, according to the Design Code for Residential Buildings of China, it divided the residential building into four categories based on the floors as shown in Table 5. 10.

Table 5. 10: Residential building categories in China

Floors	**Category**
1-3 floors	low-rise building
4-6 floors	Multi-storied building
7-9 floors	medium residential buildings
10-30 floors	high-rise residential buildings

The second variable is housing units that are the number of households within one floor. Usually, the smaller number of housing units, the living environment of the building is more comfortable. It can be observed from the second plot, there is a sharp decline from about 1 housing units to about 9 housing units and then the partial dependence plot begins a more gradual decrease until 19 housing units.

Shopping mall provides significant conveniences for surrounding residents and it plays an essential role in daily life. In this work, the importance of proximity to the shopping mall is the third important variable affecting housing prices out of 30 variables. It has been proved in the third plot with more details. With the distance to shopping mall increasing, the housing prices from 16600 yuan/m^2 slide fast to 15600 yuan/m^2. When the distance more than 1600, it has a little bit impact on housing prices.

The fourth plot in Figure 5. 23 shows that the distance to the government has a negative impact on housing prices when the distance is less than about 1000m. Then, with the distance increasing, the housing prices increase from about 15000 yuan/m^2 to about 16500 yuan/m^2. When the distance more than 5000m, the housing prices are remaining stable.

The impact of proximity to the park has an obvious positive impact on housing prices as shown in the fifth plot. With the increase of the distance to the urban park from 0m to 1500m, the housing prices decrease rapidly from about 16300 yuan/m^2 to 15800 yuan/m^2, and then the prices remain stable. It is consistent with expectation because parks are closely related to the amenity and health of citizens and people are increasing concern about the environmental quality in recent years.

The sixth plot shows partial dependence on trucks. First, there is a sharp decline from housing prices 16600 yuan/m^2 to about 15800 yuan/m^2 when the truck value increases from 0 to about 0.01. Then, the partial dependence plot remains stable. But when the truck area value close to 0.04, the housing prices gradually increase until the truck area about 0.06. It is no surprise that truck is a type of heavy vehicle and has been regarded as low-cost freight transportation, but it will generate a negative external cost such as air pollution, greenhouse gases, noise.

From the seventh plot, it is apparent that the traffic light promotes the housing prices rising. More preciously, when the traffic light area value growing from 0 to about 0.0008, housing prices increase from 15400 yuan/m^2 to 17000 yuan/m^2. The value of traffic light is larger indicate the road has good connectivity with others, also with large population flow, traffic flow, and many passageways. In this work, the study area belongs to the downtown area and base on the understanding of the study area, the road with better connectivity usually is a place where various commercial facilities clustering. Commercial facilities will drive the development of surrounding land use and further affect housing prices.

The ninth plot shows partial dependence on the proximity to the hospital. The housing prices dramatically decrease from 16300 yuan/m² to about 15800 yuan/m² with the distance to hospital increase from 0 to about 1000 m. Afterward, housing prices increase strongly up to about 16300 with the distance continue to increase to about 5000 m. the hospitals contained in this work belong to AAA level hospitals that have better medical facilities including medical level, health resource, and service quality. Usually, these hospitals are crowded with heavy traffic flow, which leads to a negative impact on the housing price of the surrounding area. However, the convenience provided by the hospital is an attractive factor for the buyers. The tenth plot shows partial dependence on the number of bedrooms. When the number of bedrooms is small than 4, the housing prices remain stable. It because a house with three bedrooms is the most common and economic house type considering the family structure of China that is most Chinese families have one or two children. When the number of bedrooms more than 3, the housing prices are gradually rising. Partly because the strategy for house type focuses on the comfortable instead of economic. One extreme example is the villa.

The partial dependence plots of other variables are shown in the Appendix.

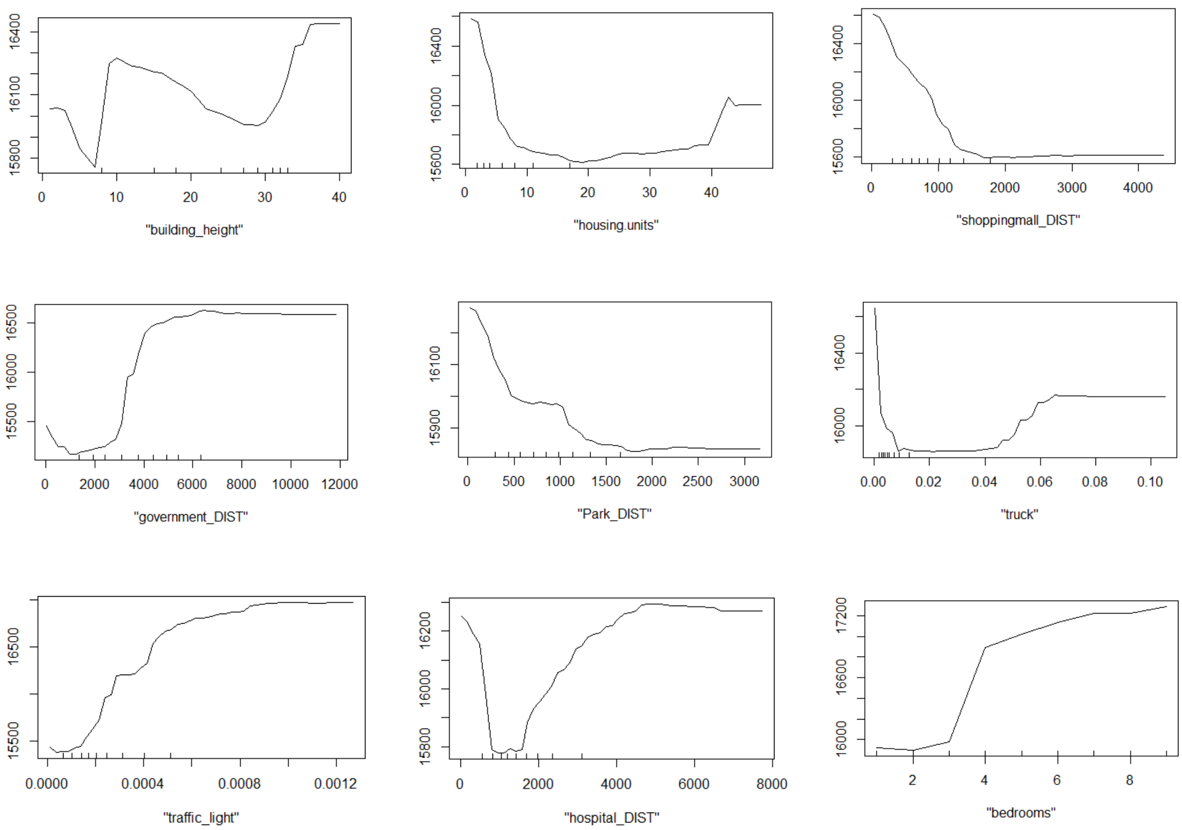


Figure 5. 23 : Partial dependence plot for variables

5.5. Discussion

This study builds a housing price prediction model in Xi'an by combining geospatial data, street view image data, and satellite data, generating good accuracy along with encouraging results.

First, in the determinants aspect, I find that traditional attributes including structure, location, neighborhood environment characteristics explain the majority of variance of housing value. Among the geospatial variables, the first five important variables are building height, housing units, proximity to shopping mall, and proximity to the metro station. To explore the impact of street visual features that reflect the street environment on housing prices. I collected the street view image within the study area and used semantic segmentation technique to extract 17 types of visual features. The results proved that the truck, traffic light have a significant impact on housing prices. According to the partial dependence plot for truck and traffic light, the truck is related to negative environments such as air pollution, noise, and heavy traffic, therefore it has an adverse influence on housing prices. However, the traffic light favorably impacts housing prices. By viewing the locations with more traffic lights in Baidu Panorama, the interesting founding is the place with more traffic light are bustling commercial district. It suggested that traffic light is related to the Bustling city scene to some extent. One unexpected result of the analysis is that NDVI and tree are not correlated. NDVI objectively quantifies the greening level of the sounding area of the dwelling units. Tree feature, or green view index, is to measure the green space seen by residents, belong to a subjective determinant. Previous studies suggested that the green view index has a positive impact on housing prices. it also reflects in my work, further, the importance of NDVI is higher than the green view index.

For the model, this work builds a RF model with three experiments that contain different data. The results show (Table 5. 11) that the third model that combines all dataset has the best performance both in % Var explained and Mean of squared residuals. I also find that the second model only using street view image data has a good prediction ability that the % Var explained about 72 and the R-square about 0.75.

Table 5. 11: Comparison of model results

	Geospatial data only	**Street view image data only**	**All data**
% Var explained	78.09	71.54	78.96
Mean of squared residuals	5776218	7504173	5412849
R-square	0.9106426	0.752946	0.9005935
RMSE	1534.89	2552.163	1618.898
RSE	0.0893574	0.247054	0.0994065
MAE	965.6463	1585.99	1019.536
Pearson R	0.95691	0.86695	0.9504118

5.6. Limitation

First, in China, some studies have proved that schoolhouse has a positive impact on housing prices. In this work, during the feature selection, the variable of schoolhouse has been eliminated because the importance of the schoolhouse is a relatively less important variable. The small amount of data makes it significantly less obvious. More preciously, there are 1053 records of schoolhouse among 34573 record housing price data.

Second, the variable height, or the storey, is a relatively less important determinant. Previous studies had proved that, from the view of buyers' preferences, they are willing to buy the middle story level partly related to the Doctrine of the Mean, a Chinese tradition. In this work, the impact of the storey is not obvious. The most likely reason is, in the same residential building, the change in housing prices does not change a lot. RF model does not capture the small float in housing prices.

Third, this work focuses on Xi'an single real estate market reducing the extent of generalization ability. Increasing the comparative research between cities helps to improve the generalization ability of the model and potentially reveal differences.

6. CONCLUSION

The previous chapter summarizes the major findings including determinants and models. This chapter will present the conclusion according to the objective and sub-objectives. Finally, recommendations are suggested for future research in the perspectives of research direction and relevant stakeholders.

6.1. conclusion

This study investigates the effect of the street visual features base on street view images on housing prices in Xi'an and builds a housing price estimation model simultaneously. Three models that using different data source are built in the analysis. The results show that the street visual features can explain the majority of the variance of the house price. By comparing the results of three models, the model using geospatial data performs better than the model using street view image data. More specifically, the results show that there are non-linear relationships between different street visual features and property value. In addition, compared with the hedonic model, this study shows that the random forest regression model can more accurately estimate the housing prices.

For the research questions, the answer is as follow:

1.1 What street visual features affect housing prices?

In this work, I selected 17 street visual features and explored its impact and non-linear relationship between each feature and housing prices. In the second model only using street view image data, by average relative importance, the first five features: traffic light, bus, pole, fence, truck.

The process of feature selection adopting the strategy that eliminates 20% variables having smaller importance, three features are been eliminated: Motorcycle, Person, Rider. In the new RF model, the remaining features are: truck, traffic light, fence, car, traffic sign, sky, sidewalk, bus, road, wall, terrain, building, vegetation, bicycle, pole.

To sum up, the first two important street view features are truck and traffic light. The former has a negative impact on housing prices, while the latter has a positive effect on housing prices.

1.2 What extent of current street visual features can explain housing price variation?

To answer this question, the model only using street view image data are been built. According to the results: "% Var explained" is 71.54, which means that OOB predictions explained the 71.54% variance of the housing prices.

2.1 What are the current variables that affect property value in Xi'an?

The current variables that affect property value in Xi'an are shown in Figure 5. 7. Based on the results of the model that only using geospatial data, the first five variables are building height, housing units, the proximity to shopping mall, the proximity to metro station and the proximity to urban park.

2.2 To what extent does the current indicator explain the variation in house prices?

To explore and answer this research question, the model only using geospatial data are been built. The model generates a result "% Var explained" is 78.09, which means the predictions explained the 78.09% variance of the housing prices using the current variables.

3.1 What is the optimal architecture of FCN in terms of accuracy and efficiency? / What is the best strategy for selecting train, validation, and test sets?

In this work, I used the pre-trained Deep lab v3 proposed by Google. Google shared its TensorFlow model training and evaluation code, as well as models pre-trained on the Cityscapes benchmark semantic segmentation task. The performance of the test set on Cityscapes reached 82.1%. The Cityscape is a dataset contain street scenes of 50 cities, at the same time, the accuracy of the pre-trained model is higher. Therefore, the optimal architecture is the architecture of Deep lab v3.

4.1 To what extent the housing prices model can predict variations of the property value?
Based on the evaluation results, the accuracy of the housing prices model on training data is 0.91 and 0.79 for validation data.
4.2 What indicators influence more the property value in Xi'an?
The first ten important variables are building height, housing units, shopping mall, the distance to government, proximity to urban park, truck, traffic light, the proximity to the hospital, the number of bedrooms, and the proximity to the university.

6.2. Ethical Considerations

The dataset contained in this work is download from websites. It does not refer to personal privacy, official protected information.

6.3. Recommendation

The recommendations are structured regarding different stakeholders.
First, for the government, the final variables with its importance ranking can be regarded as a reference in terms of the correctness and effectiveness of housing policymaking. Besides, the importance of the variables reflects the residents' preference. For example, the proximity to the park and truck is a relatively more important determinant. Increasing attention has focused on environmental quality. In the future work of urban management, promote urban environment quality including air quality, water quality, and acoustic environment quality and pollution reasons can improve people's quality of life.
For buyers, they can use the housing prices estimation model to estimate property value. The partial dependence plots show the threshold of the impact of each variable on housing prices. When choosing the residential location, it can be considered a reference to help them better understand the price/ performance trade-offs.
For real estate developers, understanding the more important variables that can help then to improve effectiveness when making site selection policies, projects, or investment decisions.
Besides, the analysis of street visual features gives an insight for urban planners to improve the street space quality. For example, the partial dependence plots for sky shows where the higher value of the sky area indicate higher housing prices. A similar trend occurred in the plot for road, wherewith the road area increasing, housing prices gradually higher. These reflect residents' preference on street environment.

6.4. Future work

In this work, only street view image was used to explore the relationship between environmental features and housing prices. In future work, taking the images of the property interior as complementary data extend, it will be helpful to quantitative analysis of the variable decoration.
The remote sensing image is been used to calculate the NDVI of each property surrounding the area. In recent years, some studies have proved that the remote sensing image is an effective and reliable data source for extracting socio-economic variables of slums. Inspired by this, future work can be using the housing prices data and remote sensing images to carry out the housing prices mapping.

LIST OF REFERENCES

Acharya, G., & Lewis, L. (n.d.). *Valuing Open Space and Land-Use Patterns in Urban Watersheds*.

AlQuhtani, S., & Anjomani, A. (2019). Do rail transit stations affect housing value changes? The Dallas Fort-Worth metropolitan area case and implications. *Journal of Transport Geography*, *79*, 102463. https://doi.org/10.1016/j.jtrangeo.2019.102463

Antipov, E. A., & Pokryshevskaya, E. B. (2012). Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics. *Expert Systems with Applications*, *39*(2), 1772–1778. https://doi.org/10.1016/j.eswa.2011.08.077

Arietta, S. M., Efros, A. A., Ramamoorthi, R., & Agrawala, M. (2014). City forensics: Using visual elements to predict non-visual city attributes. *IEEE Transactions on Visualization and Computer Graphics*, *20*(12), 2624–2633. https://doi.org/10.1109/TVCG.2014.2346446

Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(12), 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615

Berk, R., Sherman, L., Barnes, G., Kurtz, E., & Ahlman, L. (2009). Forecasting murder within a population of probationers and parolees: a high stakes application of statistical learning. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *172*(1), 191–211. https://doi.org/10.1111/j.1467-985X.2008.00556.x

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bureau, X. 'an statistics. (n.d.). *Regulation effect deepens, real estate market becomes more rational*. Retrieved March 27, 2020, from http://tjj.xa.gov.cn/tjsj/tjxx/5e258fcd65cbd81235581acf.html

Cervero, R., & Kang, C. D. (2011). Bus rapid transit impacts on land uses and land values in Seoul, Korea. *Transport Policy*, *18*(1), 102–116. https://doi.org/10.1016/j.tranpol.2010.06.005

Chen, J., Zhou, C., & Li, F. (2020). Quantifying the green view indicator for assessing urban greening quality: An analysis based on Internet-crawling street view data. *Ecological Indicators*, *113*, 106192. https://doi.org/10.1016/j.ecolind.2020.106192

Chen, L.-C., Papandreou, G., Member, S., Kokkinos, I., Murphy, K., & Yuille, A. L. (n.d.). *DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs*. Retrieved May 22, 2020, from http://liangchiehchen.com/projects/

Chen, L.-C., Papandreou, G., Murphy, K., & Yuille, A. L. (n.d.). *SEMANTIC IMAGE SEGMENTATION WITH DEEP CON-VOLUTIONAL NETS AND FULLY CONNECTED CRFS*.

Chen, Y., Liu, X., Li, X., Liu, Y., & Xu, X. (2016). Mapping the fine-scale spatial pattern of housing rent in the metropolitan area by using online rental listings and ensemble learning. *Applied Geography*, *75*, 200–212. https://doi.org/10.1016/j.apgeog.2016.08.011

Cheng, B., Collins, M. D., Zhu, Y., Liu, T., Huang, T. S., Adam, H., & Chen, L.-C. (2019). *Panoptic-DeepLab*. 1–4. http://arxiv.org/abs/1910.04751

Chin, T L, & Chau, K. W. (2003). A critical review of literature on the hedonic price model. In *International Journal for Housing and Its Applications* (Vol. 27, Issue 2). http://ssrn.com/abstract=2073594

Chin, Tung Leong, & Chau, K. W. (2003). A critical review of literature on the hedonic price model. *International Journal for Housing Science and Its Applications*, *27*(2), 145–165.

China, M. of H. and U.-R. D. of the P. R. of. (n.d.). *Standard for urban residential area planning and design [GB50180-2018]*. Retrieved May 27, 2020, from http://www.jianbiaoku.com/webarbs/book/1095/3781392.shtml

Cityscapes Dataset – Semantic Understanding of Urban Street Scenes. (n.d.). Retrieved January 13, 2020, from https://www.cityscapes-dataset.com/

Clarke, C. A., Miller, T., Chang, E. T., Yin, D., Cockburn, M., & Gomez, S. L. (2010). Racial and social class gradients in life expectancy in contemporary California. *Social Science and Medicine*, *70*(9), 1373–1380. https://doi.org/10.1016/j.socscimed.2010.01.003

Dai, X., Bai, X., & Xu, M. (2016). The influence of Beijing rail transfer stations on surrounding housing prices. *Habitat International*, *55*, 79–88. https://doi.org/10.1016/j.habitatint.2016.02.008

Dataset Overview – Cityscapes Dataset. (n.d.). Retrieved January 14, 2020, from https://www.cityscapes-dataset.com/dataset-overview/#class-definitions

Díaz-Uriarte, R., & Alvarez de Andrés, S. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, *7*(1), 1–13. https://doi.org/10.1186/1471-2105-7-3

Ewing, R., Hajrasouliha, A., Neckerman, K. M., Purciel-Hill, M., & Greene, W. (2016). Streetscape Features Related to Pedestrian Activity. *Journal of Planning Education and Research*, *36*(1), 5–15. https://doi.org/10.1177/0739456X15591585

Fan, G.-Z., Ong, S. E., & Koh, H. C. (2006). *Determinants of House Price: A Decision Tree Approach*.

François Des Rosiers, Antonio Lagana, Marius Thériault, M. B. (1996). *Empirical Modeling of Hte Relative Impact of Various Sizes of Shopping Centre and Value of the Surrounding Residential Properties.Pdf*.

Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L., & Fei-Fei, L. (2017). Using deep learning and google street view to estimate the demographic makeup of neighborhoods across the United States. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(50), 13108–13113. https://doi.org/10.1073/pnas.1700035114

Ghattas, B., & Ben Ishak, A. (2008). S ´ ELECTION DE VARIABLES POUR LA CLASSIFICATION BINAIRE EN GRANDE DIMENSION : COMPARAISONS ET APPLICATION AUX DONNÉESDONN´DONNÉES DE BIOPUCES. In *Journal de la société française de statistique* (Vol. 149, Issue 3).

Girardello, M., Griggio, M., Whittingham, M. J., & Rushton, S. P. (2010). Models of climate associations and distributions of amphibians in Italy. *Ecological Research*, *25*(1), 103–111. https://doi.org/10.1007/s11284-009-0636-z

Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., & Shet, V. (2013). Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks. *2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings*. http://arxiv.org/abs/1312.6082

Gu, J., Zhu, M., & Jiang, L. (2011). Housing price forecasting based on genetic algorithm and support vector machine. *Expert Systems with Applications*, *38*(4), 3383–3386. https://doi.org/10.1016/j.eswa.2010.08.123

Hansen, K. (2014). Moving house for education in the pre-school years. *British Educational Research Journal*, *40*(3), 483–500. https://doi.org/10.1002/berj.3092

Hu, L., He, S., Han, Z., Xiao, H., Su, S., Weng, M., & Cai, Z. (2019). Monitoring housing rental prices based on social media:An integrated approach of machine-learning algorithms and hedonic modeling to inform equitable housing policies. *Land Use Policy*, *82*, 657–673. https://doi.org/10.1016/j.landusepol.2018.12.030

Jayantha, W. M., & Lam, S. O. (2015). Capitalization of secondary school education into property values: A case study in Hong Kong. *Habitat International*, *50*, 12–22. https://doi.org/10.1016/j.habitatint.2015.07.011

Jim, C. Y., & Chen, W. Y. (2006). Impacts of urban environmental elements on residential housing prices in Guangzhou (China). *Landscape and Urban Planning*, *78*(4), 422–434. https://doi.org/10.1016/j.landurbplan.2005.12.003

Kang, J., Körner, M., Wang, Y., Taubenböck, H., & Zhu, X. X. (2018). Building instance classification using street view images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *145*, 44–59. https://doi.org/10.1016/j.isprsjprs.2018.02.006

kevin lynch. (1960). *The Image of the City - Kevin Lynch - Google Books*. https://books.google.nl/books?hl=en&lr=&id=_phRPWsSpAgC&oi=fnd&pg=PA1&dq=kevin+lynch++++The+Image+of+the+City&ots=jHGaaa4Bhl&sig=ApZYC5XmEQUrZ86OSSSqjqNnGUk&redir_esc=y#v=onepage&q=visual&f=false

Khan, A., Sohail, A., Zahoora, U., & Qureshi, A. S. (2019). *A Survey of the Recent Architectures of Deep Convolutional Neural Networks*. 1–67. http://arxiv.org/abs/1901.06032

Law, S., Paige, B., & Russell, C. (2018). *Take a Look Around: Using Street View and Satellite Images to Estimate House Prices*. http://arxiv.org/abs/1807.07155

Leech, D., & Campos, E. (2003). Is comprehensive education really free?: A case-study of the effects of secondary school admissions policies on house prices in one local area. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, *166*(1), 135–154. https://doi.org/10.1111/1467-985X.00263

Liu, G., Wang, X., Gu, J., Liu, Y., & Zhou, T. (2019). Temporal and spatial effects of a 'Shan Shui' landscape on housing price: A case study of Chongqing, China. *Habitat International*, *94*(July), 102068. https://doi.org/10.1016/j.habitatint.2019.102068

Liu, L., Wang, Q., & Zhang, A. (2019). The impact of housing price on non-housing consumption of the Chinese households: A general equilibrium analysis. *North American Journal of Economics and Finance*, *49*(April), 152–164. https://doi.org/10.1016/j.najef.2019.04.010

Long, J., Shelhamer, E., & Darrell, T. (n.d.). *Fully Convolutional Networks for Semantic Segmentation*.

Lu, J. (2018). The value of a south-facing orientation: A hedonic pricing analysis of the Shanghai housing market. *Habitat International*, *81*(April), 24–32. https://doi.org/10.1016/j.habitatint.2018.09.002

Luttik, J. (2000). The value of trees, water and open space as reflected by house prices in the Netherlands. *Landscape and Urban Planning*, *48*(3–4), 161–167. https://doi.org/10.1016/S0169-2046(00)00039-6

Middel, A., Lukasczyk, J., Zakrzewski, S., Arnold, M., & Maciejewski, R. (2019). Urban form and composition of street canyons: A human-centric big data and deep learning approach. *Landscape and Urban Planning*, *183*, 122–132. https://doi.org/10.1016/j.landurbplan.2018.12.001

Mujahid, M., Roux, A., … M. S.-A. journal of, & 2008, undefined. (n.d.). Relation between neighborhood environments and obesity in the Multi-Ethnic Study of Atherosclerosis. *Academic.Oup.Com*. Retrieved May 20, 2020, from https://academic.oup.com/aje/article-abstract/167/11/1349/131837

Mulley, C. (2014). Accessibility and Residential Land Value Uplift: Identifying Spatial Variations in the Accessibility Impacts of a Bus Transitway. *Urban Studies*, *51*(8), 1707–1724. https://doi.org/10.1177/0042098013499082

Naik, N., Kominers, S. D., Raskar, R., Glaeser, E. L., & Hidalgo, C. A. (2017). Computer vision uncovers predictors of physical urban change. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(29), 7571–7576. https://doi.org/10.1073/pnas.1619003114

Nguyen, Q. C., Khanna, S., Dwivedi, P., Huang, D., Huang, Y., Tasdizen, T., Brunisholz, K. D., Li, F., Gorman, W., Nguyen, T. T., & Jiang, C. (2019). Using Google Street View to examine associations between built environment characteristics and U.S. health outcomes. *Preventive Medicine Reports*, *14*, 100859. https://doi.org/10.1016/j.pmedr.2019.100859

Online, R., Herath, S., & Maier, G. (n.d.). *The hedonic price method in real estate and housing market research: a review of the literature*. Retrieved May 18, 2020, from http://ro.uow.edu.au/buspapers/971

Opoku, R. A., & Abdul-Muhmin, A. G. (2010). Housing preferences and attribute importance among low-income consumers in Saudi Arabia. *Habitat International*, *34*(2), 219–227. https://doi.org/10.1016/j.habitatint.2009.09.006

Poursaeed, O., Matera, T., & Belongie, S. (2018). Vision-based real estate price estimation. *Machine Vision and Applications*, *29*(4), 667–676. https://doi.org/10.1007/s00138-018-0922-2

Research, C. L.-A. S., & 1966, undefined. (n.d.). Reciprocity, redistribution and the male life cycle: variations in middle river Tonga social organization. *Africabib.Org*. Retrieved June 1, 2020, from https://africabib.org/rec.php?RID=189168722&DB=p

retail formats - MBA. (n.d.). Retrieved May 27, 2020, from https://wiki.mbalib.com/wiki/零售业态

Rosen, S. (2017). *Hedonic Prices and Implicit Markets : Product Differentiation in Pure Competition Author (s): Sherwin Rosen Published by : The University of Chicago Press Stable URL : http://www.jstor.org/stable/1830899 JSTOR is a not-for-profit service that helps scho*. *82*(1), 34–55.

Rosen, S. (2019). Hedonic prices and implicit markets: Product differentiation in pure competition. In *Revealed Preference Approaches to Environmental Valuation Volumes I and II* (pp. 5–26). Taylor and Francis. https://doi.org/10.1086/260169

Rundle, A. G., Bader, M. D. M., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using google street view to audit neighborhood environments. *American Journal of Preventive Medicine*, *40*(1), 94–100. https://doi.org/10.1016/j.amepre.2010.09.034

Runge, N., Samsonov, P., Degraen, D., & Schöning, J. (2016). No more autobahn! Scenic route generation using googles street view. *International Conference on Intelligent User Interfaces, Proceedings IUI*, *07-10-March-2016*, 147–151. https://doi.org/10.1145/2856767.2856804

Sander, H., Polasky, S., & Haight, R. G. (2010). The value of urban tree cover: A hedonic property price model in Ramsey and Dakota Counties, Minnesota, USA. *Ecological Economics*, *69*(8), 1646–1656. https://doi.org/10.1016/j.ecolecon.2010.03.011

Segal, M. R. (2003). *Machine Learning Benchmarks and Random Forest Regression*.

Selim, H. (2009). Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. *Expert Systems with Applications*, *36*(2 PART 2), 2843–2852. https://doi.org/10.1016/j.eswa.2008.01.044

Song, Y., & Sohn, J. (2007). Valuing spatial accessibility to retailing: A case study of the single family housing market in Hillsboro, Oregon. *Journal of Retailing and Consumer Services*, *14*(4), 279–288. https://doi.org/10.1016/j.jretconser.2006.07.002

statistics, J. F.-A. of, & 2001, undefined. (n.d.). Greedy function approximation: a gradient boosting machine. *JSTOR*. Retrieved June 1, 2020, from https://www.jstor.org/stable/2699986

Su, S., Zhou, H., Xu, M., Ru, H., Wang, W., & Weng, M. (2019). Auditing street walkability and associated social inequalities for planning implications. *Journal of Transport Geography*, *74*, 62–76. https://doi.org/10.1016/j.jtrangeo.2018.11.003

Tan, R., He, Q., Zhou, K., & Xie, P. (2019). The effect of new metro stations on local land use and housing prices: The case of Wuhan, China. *Journal of Transport Geography*, *79*(January 2017). https://doi.org/10.1016/j.jtrangeo.2019.102488

Tan, T. H. (2012). Meeting first-time buyers' housing needs and preferences in greater Kuala Lumpur. *Cities*, *29*(6), 389–396. https://doi.org/10.1016/j.cities.2011.11.016

Wen, H., Gui, Z., Zhang, L., & Hui, E. C. M. (2020). An empirical study of the impact of vehicular traffic and floor level on property price. *Habitat International*, *97*, 102132. https://doi.org/10.1016/j.habitatint.2020.102132

Wen, H., Xiao, Y., & Hui, E. C. M. (2019). Quantile effect of educational facilities on housing price: Do homebuyers of higher-priced housing pay more for educational resources? *Cities*, *90*(February), 100–112. https://doi.org/10.1016/j.cities.2019.01.019

Wen, H., Zhang, Y., & Zhang, L. (2015). Assessing amenity effects of urban landscapes on housing price in Hangzhou, China. *Urban Forestry and Urban Greening*, *14*(4), 1017–1026. https://doi.org/10.1016/j.ufug.2015.09.013

Wittowsky, D., Hoekveld, J., Welsch, J., & Steier, M. (2020). Residential housing prices: impact of housing characteristics, accessibility and neighbouring apartments – a case study of Dortmund, Germany. *Urban, Planning and Transport Research*, *8*(1), 44–70. https://doi.org/10.1080/21650020.2019.1704429

Wojna, Z., Gorban, A. N., Lee, D. S., Murphy, K., Yu, Q., Li, Y., & Ibarz, J. (2017). Attention-Based Extraction of Structured Information from Street View Imagery. *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR*, *1*, 844–850. https://doi.org/10.1109/ICDAR.2017.143

Wu, J., Feng, Z., Peng, Y., Liu, Q., & He, Q. (2019). Neglected green street landscapes: A re-evaluation method of green justice. *Urban Forestry and Urban Greening*, *41*, 344–353. https://doi.org/10.1016/j.ufug.2019.05.004

Xiao, Y., Hui, E. C. M., & Wen, H. (2019). Effects of floor level and landscape proximity on housing price: A hedonic analysis in Hangzhou, China. *Habitat International*, *87*(April), 11–26. https://doi.org/10.1016/j.habitatint.2019.03.008

Xu, Y., Zhang, Q., & Zheng, S. (2015). The rising demand for subway after private driving restriction: Evidence from Beijing's housing market. *Regional Science and Urban Economics*, *54*, 28–37. https://doi.org/10.1016/j.regsciurbeco.2015.06.004

Yang, J., Zhao, L., Mcbride, J., & Gong, P. (2009). Can you see green? Assessing the visibility of urban forests in cities. *Landscape and Urban Planning*, *91*(2), 97–104. https://doi.org/10.1016/j.landurbplan.2008.12.004

Yin, L., & Wang, Z. (2016). Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. *Applied Geography*, *76*, 147–153. https://doi.org/10.1016/j.apgeog.2016.09.024

Yoo, S., Im., J., & Wagner, J. E. (2012). Variable selection for hedonic model using machine learning approaches: A case study in Onondaga County, NY. *Landscape and Urban Planning*, *107*(3), 293–306. https://doi.org/10.1016/j.landurbplan.2012.06.009

Yu, F., & Koltun, V. (n.d.). *MULTI-SCALE CONTEXT AGGREGATION BY DILATED CONVOLUTIONS*.

Zeng, L., Lu, J., Li, W., & Li, Y. (2018). A fast approach for large-scale Sky View Factor estimation using street view images. *Building and Environment*, *135*, 74–84. https://doi.org/10.1016/j.buildenv.2018.03.009

Zhang, F., Wu, L., Zhu, D., & Liu, Y. (2019). Social sensing from street-level imagery: A case study in learning spatio-temporal urban mobility patterns. *ISPRS Journal of Photogrammetry and Remote Sensing*, *153*, 48–58. https://doi.org/10.1016/j.isprsjprs.2019.04.017

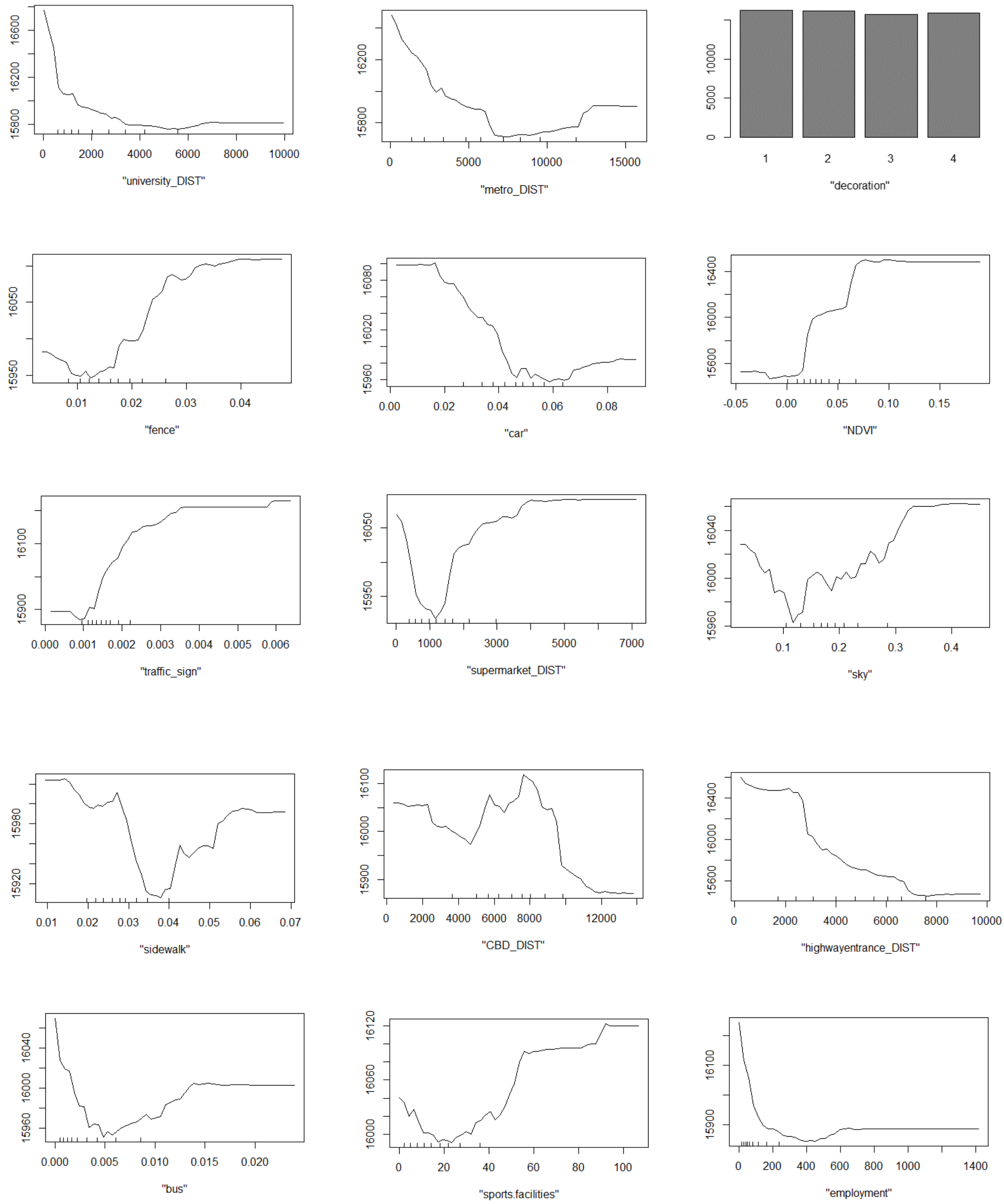
Zhang, F., Zhang, D., Liu, Y., & Lin, H. (2018). Representing place locales using scene elements. *Computers, Environment and Urban Systems*, *71*, 153–164. https://doi.org/10.1016/j.compenvurbsys.2018.05.005

Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning*, *180*, 148–160. https://doi.org/10.1016/j.landurbplan.2018.08.020

Zhang, Y., & Dong, R. (2018). Impacts of Street-Visible Greenery on Housing Prices: Evidence from a Hedonic Price Model and a Massive Street View Image Dataset in Beijing. *ISPRS International Journal of Geo-Information*, *7*(3), 104. https://doi.org/10.3390/ijgi7030104

Zheng, S., Xu, Y., Zhang, X., & Wang, R. (2016). Transit development, consumer amenities and home values: Evidence from Beijing's subway neighborhoods. *Journal of Housing Economics*, *33*, 22–33. https://doi.org/10.1016/j.jhe.2016.05.003

APPENDIX


"university_DIST"
"metro_DIST"
"decoration"
"fence"
"car"
"NDVI"
"traffic_sign"
"supermarket_DIST"
"sky"
"sidewalk"
"CBD_DIST"
"highwayentrance_DIST"
"bus"
"sports.facilities"
"employment"

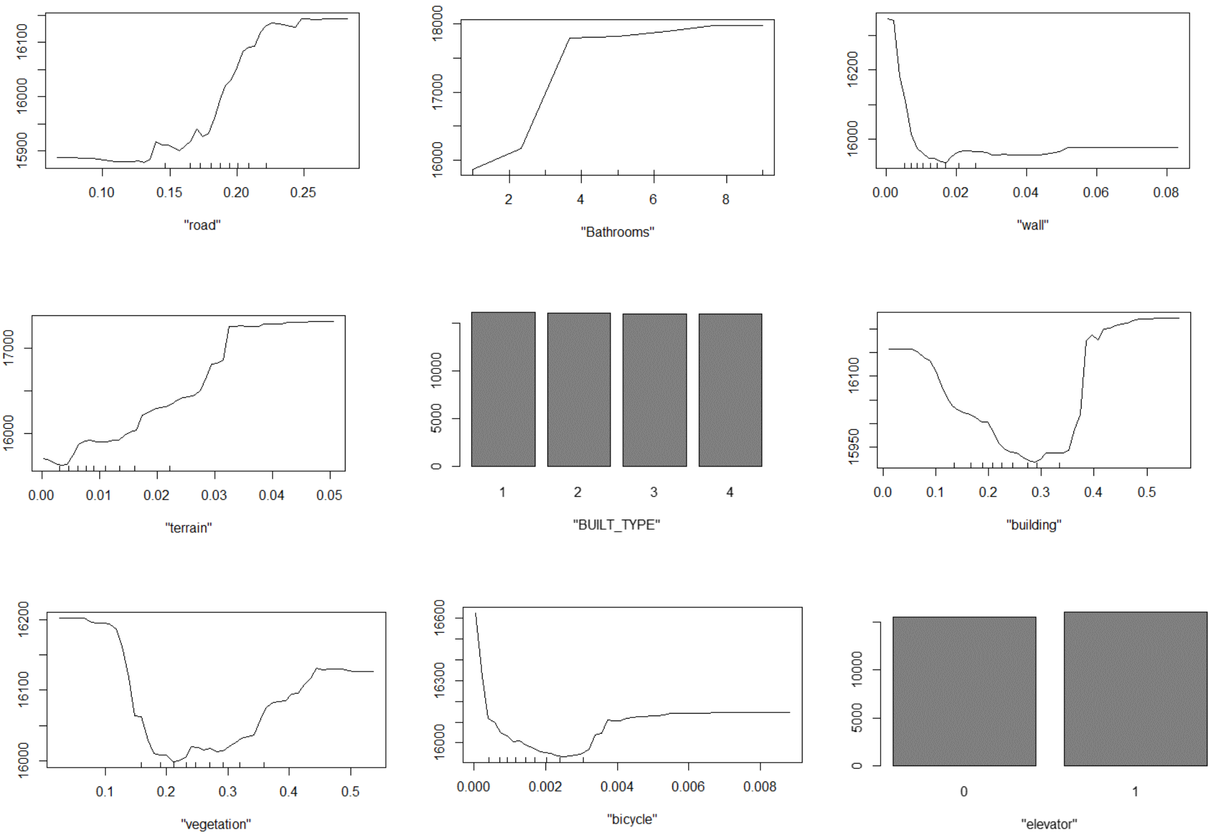


Figure 1: Partial dependence plot for variables

Table 1: Estimation results of the semi-parametric model

variable	Estimate	Std. Error	t value	Pr(>|t|)	
	(Intercept) 6.278e+04	4.06E+03	1.55E+01	< 2e-16	***
NDVI	2.05E+04	1.30E+03	1.58E+01	< 2e-16	***
CBD_DIST	-8.51E-01	3.72E-02	-2.29E+01	< 2e-16	***
supermarket_DIST	2.07E-01	3.39E-02	6.12E+00	9.69E-10	***
university_DIST	2.14E-01	3.05E-02	7.01E+00	2.50E-12	***
highwayentrance_DIST	-6.96E-01	2.31E-02	-3.01E+01	< 2e-16	***
Park_DIST	-3.71E-01	4.79E-02	-7.74E+00	1.06E-14	***
shoppingmall_DIST	-2.07E+00	4.93E-02	-4.20E+01	< 2e-16	***
hospital_DIST	7.68E-02	2.94E-02	2.61E+00	0.00897	**
government_DIST	7.95E-01	2.41E-02	3.30E+01	< 2e-16	***
bedrooms	4.87E+02	3.52E+01	1.38E+01	< 2e-16	***
Bathrooms	1.60E+03	5.22E+01	3.05E+01	< 2e-16	***
height	-2.59E+02	2.66E+01	-9.74E+00	< 2e-16	***
BUILT_TYPE	-7.90E+02	3.62E+01	-2.18E+01	< 2e-16	***
decoration	-4.81E+02	1.93E+01	-2.49E+01	< 2e-16	***
housing.units	-8.91E+00	4.02E+00	-2.22E+00	0.02651	*
elevator	3.37E+03	9.77E+01	3.45E+01	< 2e-16	***
sports.facilities	-5.95E+01	3.20E+00	-1.86E+01	< 2e-16	***

bus_D	-2.71E+01	2.19E+00	-1.24E+01	< 2e-16	***
resta_conve	4.34E+00	3.40E-01	1.28E+01	< 2e-16	***
kinder_D	3.40E+01	1.30E+01	2.61E+00	0.008989	**
Mpschool	-2.21E+02	2.24E+01	-9.87E+00	< 2e-16	***
school.house	3.26E+02	1.45E+02	2.24E+00	0.025028	*
road	-3.86E+04	4.07E+03	-9.48E+00	< 2e-16	***
sidewalk	-4.40E+04	5.18E+03	-8.50E+00	< 2e-16	***
building	-4.22E+04	4.27E+03	-9.90E+00	< 2e-16	***
wall	-4.89E+04	4.60E+03	-1.06E+01	< 2e-16	***
fence	-5.17E+04	5.29E+03	-9.79E+00	< 2e-16	***
traffic_light	3.55E+06	1.51E+05	2.36E+01	< 2e-16	***
traffic_sign	4.12E+05	4.73E+04	8.71E+00	< 2e-16	***
vegetation	-4.31E+04	4.37E+03	-9.86E+00	< 2e-16	***
sky	-5.11E+04	4.31E+03	-1.19E+01	< 2e-16	***
person	-6.60E+04	1.80E+04	-3.67E+00	0.000243	***
car	-5.83E+04	4.34E+03	-1.35E+01	< 2e-16	***
truck	-1.67E+04	5.17E+03	-3.23E+00	0.001238	**
bus	-1.71E+04	8.57E+03	-1.99E+00	0.046298	*
motorcycle	-5.95E+05	3.59E+04	-1.66E+01	< 2e-16	***
bicycle	1.65E+05	3.37E+04	4.90E+00	9.86E-07	***
employment	-2.84E+00	2.23E-01	-1.27E+01	< 2e-16	***