The Annals of Regional Science (2024) 72:579–599
https://doi.org/10.1007/s00168-023-01212-7

ORIGINAL PAPER

Machine learning approach to residential valuation: a convolutional neural network model for geographic variation

Hojun Lee[1] · Hoon Han[1,2] · Chris Pettit[2] · Qishuo Gao[2] · Vivien Shi[2]

Received: 20 March 2022 / Accepted: 24 January 2023 / Published online: 27 February 2023

Abstract
Geographic location and neighbourhood attributes are major contributors to residential property values. Automated valuation models (AVM) often use hedonic pricing with location and neighbourhood attributes in the form of numeric and categorical variables. This paper proposed a novel approach to automated property valuation using a machine learning model with a convolutional neural network (CNN), fully connected neural network layers with numeric and categorical variables. In this study we compare the results of a fused model, which treat geographical data as an input with the performance of the baseline neural network model with only numerically or categorically represented data. Furthermore, the residential valuation by the proposed fused model was tested with actual sold price data in Greater Sydney, Australia. The study found that the fused model produced valuations with a significantly lower mean absolute percentage error (MAPE) (8.71%) than the MAPE of the baseline model (11.59%). The results show that the fused model with CNN significantly improves the accuracy for residential valuation, reducing spatial information loss by data manipulation and distance calibration.

JEL Classifcation C31 · O18 · C31 · R32

1 Introduction

Residential property accounts for three-quarters of the total global real estate value of US$217 trillion (Barnes and Tostevin 2016). In Australia, the context of this research, this value exceeds US$7 trillion. An accurate residential valuation is crucial for a diverse range of stakeholders, both the supply side (e.g. financial

✉ Hoon Han
h.han@unsw.edu.au

[1] School of Built Environment, University of New South Wales, Sydney, Australia

[2] City Futures Research Centre, University of New South Wales, Sydney, Australia

institutions) and the demand side (e.g. developers and purchasers). Owner-occupied dwellings in Australia were the largest asset held by households accounting for 42% of their total assets in the financial year of 2017–18 (Australian Bureau of Statistics 2019). Major banks rely on property valuation to make lending decisions on mortgages (Scheurwater 2017), and they are a major user of real estate valuation services. Warren and Elliott (2005) reported that about 60% of property valuation services in Australia occur for banks and mortgage companies. For better decision making, real estate agents and potential homeowners require accurate and reliable property valuation. Furthermore, government agencies assess annual residential values for taxation purposes, and policymakers can measure the impact of new infrastructure provision such as public transport facilities on 'value uplift' (Lieske et al. 2019; Leao et al. 2021).

Given the importance of accurate residential valuation, studies have focused on the determinants of house prices and quantified the magnitude of each determinant on house prices using hedonic price models (HPMs) (Mulley et al. 2016; Lieske et al. 2019). Location and neighbourhood attributes are perceived as major determinants of residential prices and have been included as numeric and categorical variables in HPMs (Bartholomew and Ewing 2011). However, such variables may lead to information loss and compromised valuation when geographical data is arbitrarily quantified as numeric and categorical variables. Kopczewska (2021) pointed out that while classic statistical approaches such as HPMs are commonly used in urban and regional studies, along with an increasing number of geographically weighted regression driven models (Lu et al. 2014; Wu et al. 2019), machine learning (ML) based on a black-box testing approach is less frequently studied in the field. In contrast, machine learning approaches have been rigorously tested and applied in areas such as epidemiology, ecology, geology and climate change.

Recently, several studies have investigated a machine learning approach to urban modelling, and its predictive capability has extended to other fields of regional science such as land surface temperature (Osborne and Alvares-Sanches 2019), classifying land use (Jochem et al. 2018), evaluating the built environment (Liu et al. 2017), evaluating streetscape (Lieu et al. 2021), understanding patterns of gentrification(Reades et al. 2019) and predicting property prices (Lock et al. 2020). Despite the limited interpretability of machine learning models, machine learning has an advantage of accurate prediction from its ability to consider multilevel interactions and nonlinear relationships between input data and its capability to process various forms of data (Hu et al. 2019). A convolutional neural network (CNN) is a subset of machine learning, and its deep learning algorithms are designed for processing structured arrays of big data such as images, speech, or audio signal inputs. This paper examines the potential of convolutional neural networks (CNNs) in property valuation using geographical data as an input in the machine learning approach.

Existing CNN models for property valuation have used image data such as photographs of the house, street views, or birds-eye views as inputs (Kang et al. 2020; Kostic and Jevremovic 2020). However, geographical data has not been used as input data for CNNs in residential property valuation models.

First, geographical layers of the surrounding area of the target house, such as distribution of points of interest (POI), demographics and land control, are fed into an

optimal CNN to extract the spatial information. Traditional housing features, such as structural characteristics of the house, neighbourhood, and land features, are all processed by a multilayer perceptron (MLP). Finally, the information extracted by the CNN and MLP is combined to predict the house price. The novel contribution of this paper is twofold: (1) we demonstrate the superiority of neural network-based methods in the accuracy of residential valuation; and (2) we reduced the information loss by feeding a diverse range of geographical data directly into the model instead of considering it in the form of numeric and categorical variables. To the best of our knowledge, this is the first study to apply geographical layers in residential valuation using a CNN model.

2 Literature review

2.1 Residential valuation models

The hedonic price model (HPM) is the most widely used model in residential valuation because of its consistency and straightforwardness (Bartholomew and Ewing 2011). The hedonic price model is extended from Alonso's (1964) bid-rent theory, which shows the trade-off relationship between accessibility to the central business district (CBD) and expected rental yields (Evans 1995). Rosen (1974) proposed the HPM to estimate the impact of various housing characteristics on residential prices beyond the impact of accessibility to the CBD. The traditional HPM is based on ordinary least squares (OLS) regression, and its coefficients show how much value changes by each unit of an attribute increasing or decreasing. House price in the HPM is generally regarded as a set of the characteristics of houses that contributes to an equilibrium price between sellers and buyers (Seo 2019). In particular HPMs need statistical assumptions on data distribution and linear relationships between dependent and independent variables. Statistical bias and incorrect inferences can be produced when these assumptions are violated (Li et al. 2012). HPMs also have the risk of multicollinearity, which exists when there is a high correlation between several independent variables (Grover 2016). Multicollinearity issues can lead to unstable hedonic prices of independent variables, which reduces the advantage of the HPM in interpretability (Powe et al. 1995).

Advanced HPMs have been used to control spatial effects such as spatial dependency and heterogeneity. Conventional HPMs have spatial autocorrelation among residuals, and it violates the assumption of independent residuals among observations and may lead to inaccurate results (D'amato 2017). Spatial autocorrelation occurs because relationships between neighbouring areas are spatially dependent as neighbouring zones are likely to have similar external factors such as neighbourhood and attributes (Bourassa et al. 2007). To overcome the issue of spatial dependency, spatial lag or spatial error terms have been introduced in HPMs (Feng and Humphreys 2012; Brennan et al. 2014).

Another spatial effect is spatial heterogeneity which refers to the spatially non-stationary distribution of the coefficients of HPMs (D'Amato 2017). To deal with spatial heterogeneity, moving window regression (MWR), geographical weighted

regression (GWR), and eigenvector spatial filtering (ESP) have been used. MWR estimates local regression and coefficients of attributes using neighbouring observations. In MWR, neighbouring observations are included with equal weight (Páez et al. 2008). GWR, as proposed by Fotheringham et al. (2003), added distance-based weights to MWR to more effectively address the heterogeneity issue (Helbich and Griffith 2016) and has been applied in residential valuation studies (Hu et al. 2016; Mulley et al. 2016). However, the reliability of coefficients due to multicollinearity among independent variables is consistently noted as a weakness of GWR (LeSage 2004).

2.2 Considering external factors of residential price in valuation models

Determinants of residential value can be characterised into three categories: 1) structural attributes, 2) location attributes, and 3) neighbourhood attributes (Dubin 1988; Lieske et al. 2019). Structural attributes are variables related to the structure of the house itself and often include variables such as the number of bedrooms, bathrooms, garages, floor area, built year, and presence of guest parking.

Location and neighbourhood attributes can be grouped as external factors related to the location of the house and its surroundings (Bartholomew and Ewing 2011). Location attributes are typically defined as accessibility to the amenities that considerably impact residential value. Most studies have used one of the measures of accessibility: 1) distance to the closest amenity (AlQuhtani and Anjomani 2019; Hu et al. 2019), 2) existence of the amenity within a certain distance (McIntosh et al. 2014; Mulley and Tsai 2017), and 3) density of the amenities (Li et al. 2019; Yang et al. 2019) (see Fig. 1). These studies found a significant relationship between residential price and accessibility to amenities, including transportation, shopping

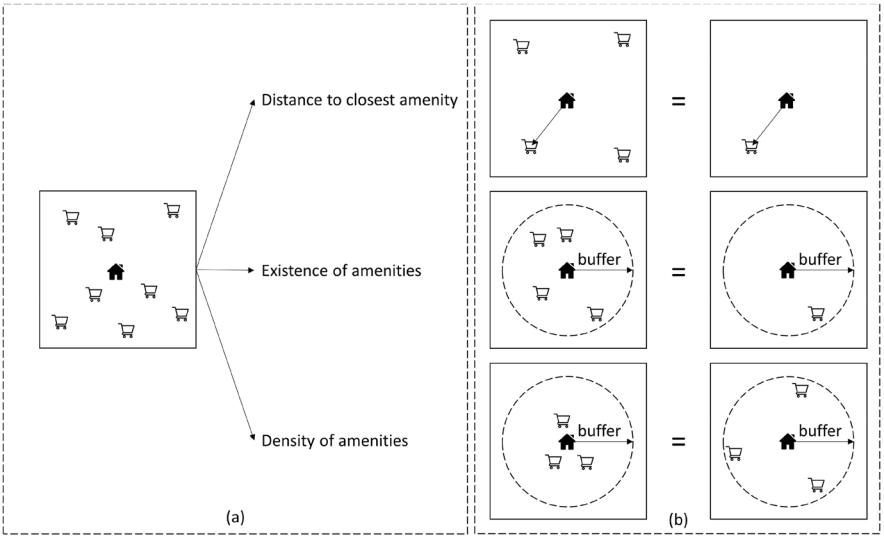

Fig. 1 Example of various measurements. **a** and information loss **b** of location attributes

centres, medical services, education facilities, and the city centre (AlQuhtani and Anjomani 2019, Hu et al. 2019; Yang et al. 2019).

Despite this, some studies have criticised the limitations of quantifying location attributes as numeric and categorical variables. Sadayuki (2018) and Hu et al. (2019) pointed out that each measurement of locational attributes represents different information about amenities such as distance to and distribution of amenities. Therefore, they can have a divergent impact on the hedonic price model. Similarly, Yuan et al. (2020) addressed that these measurements of location attributes could not provide the comprehensive effects of amenities on residential price but only part of them. For example, in measuring the distance to the closest amenity, additional amenities beyond the closest one are assumed to have no contribution to residential price. Furthermore, buffers for density or using dummy variables for the existence of amenities is usually determined in an arbitrary and heuristic manner (Sadayuki 2018; Lieske et al. 2019).

Furthermore, predefined geographical boundaries such as jurisdictional or statistical areas may misinterpret the magnitude of neighbourhood attributes (e.g. household income, employment, education and crime) to the property price. Neighbourhood features used in the previous studies often consider only the neighbourhood the house is in, even though residential value can be affected by the characteristics of surrounding neighbourhoods. For a house located at the edge of a neighbourhood, this problem is more prominent. Some studies on residential valuations considered neighbourhood characteristics such as socioeconomic and demographic status using a spatial Durbin model (Mussa et al. 2017; Copiello 2020).

2.3 CNNs for residential valuation

Recent studies have applied machine learning models to residential valuation. Generally, machine learning models do not need strict statistical assumptions and are more capable than HPMs of analysing the nonlinear relationship between residential price and its attributes (Li et al. 2012; Chen et al. 2017). In particular, several studies showed the improved performance of the residential valuation model by employing a convolutional neural network (CNN) which is one type of deep neural network. A CNN specialises in capturing spatial patterns (LeCun et al. 2015) and has an advantage in the efficient processing of 2-dimensional data such as images. Studies using CNNs used outdoor and indoor images of houses (Kang et al. 2020, Kostic and Jevremovic 2020), street view images of the neighbourhood (Law et al. 2019; Kang et al. 2020) and birds-eye view images of the surrounding area (Bency et al. 2017, Kostic and Jevremovic 2020) as input for the neural network.

In these studies, CNN has two purposes. First, most of these studies use a CNN to extract visual features of the house and neighbourhood, or specific concepts such as neighbourhood safety (De Nadai and Lepri 2018) and luxury home features (e.g. landscaping) (Poursaeed et al. 2018), which are usually not considered in a hedonic price model. For example, Kang et al. (2020) extracted the visual features from internal house photos and street view images using a CNN and included them in their residential valuation model. Poursaeed et al. (2018)

showed that the performance of residential valuation is improved by including visual luxury home features measured by a CNN using house photos from Zillow, a leading online real estate valuation company in USA.

Secondly, other studies used a CNN to gain information on more comprehensive features of the surrounding area of the house. Bency et al. (2017) included features extracted from a CNN using satellite images of the surroundings in their valuation model of residential property, rent and sales price in UK cities, arguing that satellite images can provide contextual information on the neighbouring area. They showed the model is superior to the spatial autoregressive (SAR) model, and using a wider surrounding area of the target house improved model performance. Similarly, Bin et al. (2019) included street maps in their CNN-based valuation model to consider latent geographical features from the open street map and a birds-eye view image of the neighbourhood.

In summary, previous studies on residential valuation have shown significant relationships between housing price and external factors, including location attributes and neighbourhood attributes. However, information on these external factors might be lost while measuring them as numeric or categorical variables in conventional valuation models, and thus lead to a less accurate valuation model. Meanwhile, despite CNNs being recently applied to residential valuation models, the models have focused on including visual features extracted from images, such as outdoor and indoor images of houses, street view images and birds-eye view images rather than geographical layers, especially the surrounding amenities which are closely related to house prices. Our research approach focused on the application of CNN for modelling geographically relationships for improving residential value prediction.

3 Study area

This study applies the proposed framework to Greater Sydney, NSW, Australia. The Sydney metropolitan area is approximately 12,300 sq km, with a population of 5.3 million in 2019. According to Demographia's housing affordability survey, Sydney is one of the most expensive cities in the world and the most expensive city in Australia (Cox and Pavletich 2020). Population growth in Greater Sydney is one of the main contributors to its dynamic property market. Between 2001 and 2019, the population in Greater Sydney has risen by about 30% from 4,102,580 to 5,312,163, and it is one of the liveable cities in the world (The Economist Intelligence Unit 2019). Various geographical and big data is available for Sydney. For example, APM, a Sydney-based company, provides historical sales data with various property attributes. This study uses residential property transaction data in 2018 for the Greater Sydney area acquired from APM under a research license. The data includes sales price in Australian dollars, sales date, and the geographic location of houses. Of the total 44,128 records, 1793 records were excluded because their surrounding area was located outside Greater Sydney, where data is limited.

4 Analytical framework

To alleviate the information loss discussed in previous sections, this paper proposes a fused model based on MLP and CNNs to comprehensively characterise the location and neighbourhood information of a target house from geographical data.

Figure 2 shows the proposed framework for data manipulation and modelling in the fused model. First, residential sales data (Australian Property Monitors, APM) and various geographical data including Australian Bureau of Statistics(ABS) Census, land zoning, point of interests(POIs) are processed into numeric and categorical variables about structural, location, neighbourhood characteristics of the house as input data for MLP. Geographical data is also manipulated into multiple layers of geographical raster layers as input data for CNN. The fused model structure consists of three parts: 1) an MLP, 2) an CNN, and 3) networks of fully connected layers. The CNN which is main novelty of this paper extracts the features from

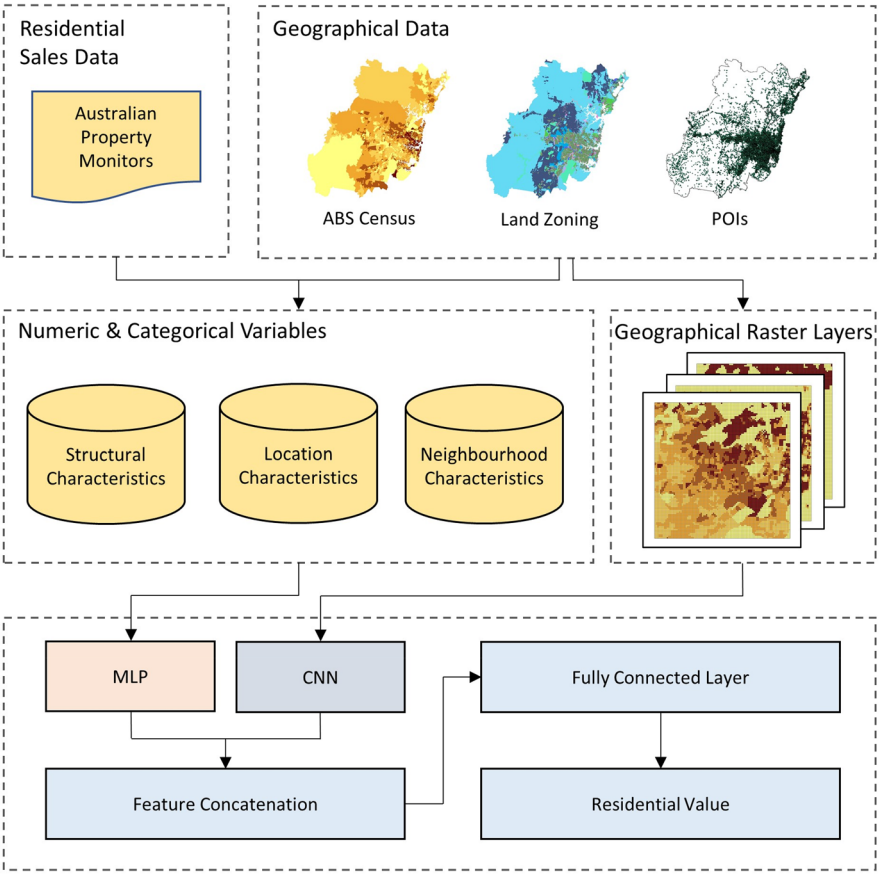

Fig. 2 Modelling framework

multilayers of geographical data. Multiple geographical data were processed to raster shape geographical layers to unify the form of data and to have similar structure with imagery data which is usual input data of CNN. These raster geographical data include location and neighbourhood attributes of surrounding rectangle area of the target house. MLP uses numeric and categorical variables, which have been widely used in conventional residential valuation models, as inputs. MLP includes variables about location and neighbourhood characteristics as well as structural characteristics of the house. Then, the features extracted by MLP and CNNs are concatenated and input into the fully connected layers.

4.1 Multilayer perceptron model

The MLP is a type of neural network that has multiple hidden layers. It has an input layer, multiple hidden layers, which are fully connected layers, and output layers. Each hidden layer consists of multiple hidden nodes, which are neurons. Each neuron multiplies values from previous layers with weights and feeds them to the next layer. The function of each neuron can be described as follows:

$$a = \sigma(fx + b) \tag{1}$$

where a is the output of the neuron, f is the weight, x is the input of the neuron from the previous layer, b is bias, and $\sigma(\cdot)$ denotes the activation function.

4.1.1 Input data for multilayer perceptron (MLP)

The input for MLP is numeric and categorical variables related to structural, location and neighbourhood characteristics, as shown in Table 1. Structural attributes include information on land size, the number of bedrooms, the number of bathrooms, and the existence of parking from APM data. Structural characteristics are only included in the input of MLP and not included in the input of CNNs. Location attributes were calculated in measurements of distance to amenities and the data about the existence of amenities, using points of interest, transmission lines, and easement in land use. The geographical data is extracted from the New South Wales Land and Property Information government agency (NSW LPI). We used a total of 45 variables as location attributes for amenities, including city centres, schools, shopping centres, transit stations, and hospitals—See Table 1.

Neighbourhood attributes can be grouped into demographics, crime, education quality, and land control. The demographic data was sourced from the 2016 Australian Census at the Statistical Areas Level1 (SA1s). SA1s are irregular shaped geographic areas with a population of 200 to 800 people, and there are 61,845 SA1 in Australia (ASGS, 2021). Majority of SA1 is residential area with a density of over 200 per square kilometre in urban area and residential area with a density of below 200 per square kilometre in rural area. Some SA1 areas having nil population such as airports, ports, and natural area including national parks and lakes. Three neighbourhood attributes were measured at SA1 level: median weekly family income, Percentage of the population aged 65 above, and Percentage of

Table 1 List of input variables for MLP

Type		Variables	Source
Structural characteristics		Area size, the number of bedrooms, the number of bathrooms, the existence of parking	Australian Property Monitors (2018)
Location characteristics	Distance to amenities	CBD, city centres, coastlines, highway ramp, light rail, bus interchanges, primary school, high school, university, hospital, swimming pool, library, shopping centre, transmission lines	NSW Land & Property Information (2017)
	Existence of amenities	primary easement, main road, railway, electricity transmission lines, primary school, high school, shopping centre, hospital, university, light rail	
Neighbourhood characteristics	Demographics	Median weekly family income, percentage of population aged 65 above, Percentage of the population born overseas	Australian Bureau of Statistics Census (2016)
	Crime	Crime rate	NSW Bureau of Crime Statistics and Research (2018)
	Education quality	NAPLAN score of nearest primary school and high school within the school catchment	NSW Department of Education (2018)
	Land control	Land use (rural, residential, environment, business, recreation, industrial, and special)	NSW Department of Planning & Environment (2018)

the population born overseas. To adequately measure local education quality, the National Assessment Program—Literacy and Numeracy (NAPLAN) scores of public primary and high schools and relevant school catchments were collected. The quality of education is measured as the highest NAPLAN score of the primary school and high school within the school catchment of the house. Lastly, we also considered crime rates (number of crimes divided by population) of the Suburbs and Localities (SAL) areas where the house is located. Boundary of SALs are defined by State and Territory governments of Australia, and total 15,353 SALs cover the whole of Australia. The land use at the location of the target house is also considered.

4.2 Convolutional neural network

Compared to other neural network algorithms, a CNN is characterised by convolutional layers and pooling layers. In convolutional layers, neurons compute the output value that is connected to the spatial location (i, j) of the input. The output of neurons in a convolutional layer $a_{i,j}$ can be defined as follows:

$$a_{i,j} = \sigma\left((F \otimes X)_{i,j} + b\right) \tag{2}$$

F is a kernel with weights, X is the input of the neuron, which are input geographical data or feature maps from the previous convolutional layer, b is bias, and $\sigma(\cdot)$ denotes the activation function, which is a nonlinear function. The sigmoid function is applied as an activation function throughout this study. The notation of the sigmoid function is as follows:

$$\sigma(x) = 1/(1 + e^{-x}) \tag{3}$$

The pooling layer performs down sampling feature maps from convolutional layers to reduce competitive cost and over-fitting. Max pooling, which is the most common, is applied in this study. Max pooling divides a features map into a local rectangle region and makes a summarised feature map with maximum values of each region.

Neural network models, including CNNs and MLP, minimise loss function $\mathcal{L}$ by updating weights and bias of neurons by repeating training. In this study, the following loss function was applied to reduce the errors of residential price and standard deviation of residential price errors:

$$\mathcal{L}\left(Y_i, \widehat{Y}_i\right) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\widehat{Y}_i - Y_i}{Y_i}\right)^2} \tag{4}$$

where $\widehat{Y}_i$ is predicted price, Y_i is the actual price, and N is the number of observations.

In each iteration of training, weight and bias are updated to reduce the loss by gradient descent optimisation algorithm and backpropagation algorithm, which calculate gradient descent backwards through the network to minimise the cost of the network (Goodfellow et al. 2016). Adaptive moment estimation (Adam) optimiser, which is the most frequently used optimisation algorithm, is applied throughout this study. Adam

is an optimiser that applies momentum and an adaptive learning rate to prevent falling into a local minimum, which is a limitation of the basic gradient descent optimiser (Roodposhti et al. 2019).

4.2.1 Input data for the convolutional neural network (CNN)

The input for the CNN is the multilayers of geographical data about location and neighbourhood attributes within a 9.6 km×9.6 km surrounding rectangle area. The size of the input data of CNN was selected based on the network architecture of the fused model and available computational resources. To unify different forms of geographical data such as point and various shapes of polygons, geographical layers were processed to raster shape geographical layers consisting of 9216 raster cells of 100 m×100 m. Table 2 lists the types of variables included in each cell, and Fig. 3 illustrates an example of the proposed input of the CNN.

As a measurement of location attributes, the number of amenities in each raster cell is calculated using POI data from NSW LPI. For each house, the input for the CNN is 9216 raster cells surrounding the target house. Therefore, the cells can provide geographical information related to amenities, including the proximity to amenities and distribution of amenities around the house. Additionally, each raster cell includes three demographic variables from the 2016 Australian Census: median weekly family income, percentage of population aged above 65 years, and percentage of the population born overseas. Lastly, we also considered the land control instrument as a neighbourhood attribute by employing average maximum building heights on the raster cell and the area of land zonings in the raster cell. The land control variables for each raster cell were sourced from the NSW Department of Planning and Environment (DPE).

5 Validation approaches

To evaluate the model performances, we employed performance measurements referring to previous studies that compared multiple models in terms of valuation accuracy (Hu et al. 2019; Tan et al. 2019). Model performance was evaluated by three measurements: mean absolute percentage error (MAPE), R^2, and percent predicted error (PPE) within 5%, 10% and 20%. Percent predicted error within % is the proportion of observations within a certain level of absolute percentage error (e.g. 5%, 10%, 20%). MAPE, R^2, PPE can be defined as follows:

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\hat{Y}_i - Y_i\right|}{Y_i} \tag{5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{N}\left(Y_i - \overline{Y}_i\right)^2} \tag{6}$$

Table 2 List of input variables for CNN

Type		Variables	Source
Location characteristics	Number of amenities	Railway station, transport interchange, swimming pool facility/ swimming pool, shopping centre, park, hospital, library, high school, primary school, including ocean area (dummy)	NSW Land & Property Information
Neighbourhood characteristics	Demographic	Median weekly family income, percentage of population aged 65 above, Percentage of the population born overseas	Australian Bureau of Statistics Census
	Land control	Area of 8 types of land zoning (rural, environment, business, residential, recreation, industrial, special, and waterway)	NSW Department of Planning & Environment

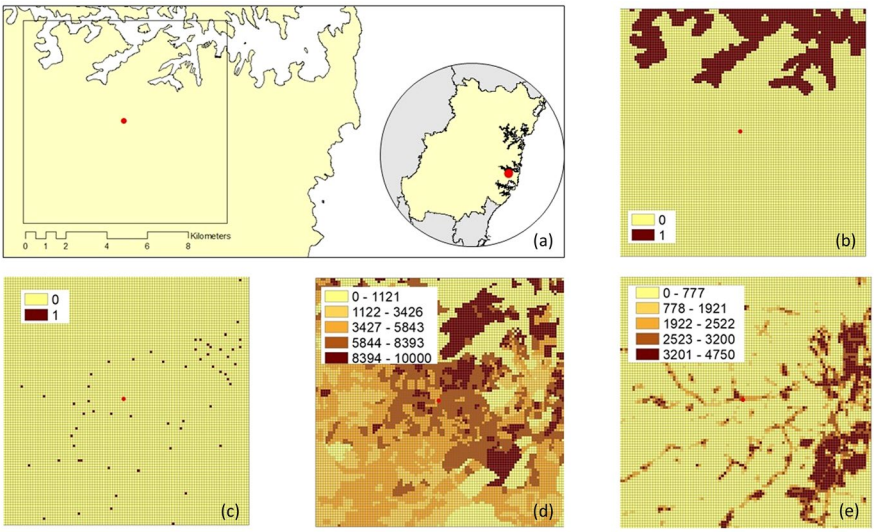

Fig. 3 Example of input for CNN: the location of house (**a**), existence of ocean (**b**), rail station (**c**), family income (**d**), business land use (**e**)

$$\mathrm{PPE}x = \frac{\mathrm{Len}(\frac{|\hat{Y}_i - Y_i|}{Y_i} < x\%)}{N} \tag{7}$$

where N is the total number of observations; Y_i and $\hat{Y}_i$ are the actual price and predicted price of residential property i; and $\overline{Y}_i$ is the average value of the actual price; Len represents the number of observations meet the requirements in the parenthesis;

We also considered the coefficient of dispersion (COD) and price-related differential (PRD), which are used as international benchmarks to measure (in) equity and uniformity of results. COD indicates horizontal uniformity of valuation result, which is equal regardless of the value of individual parcels, and PRD measures vertical uniformity considering systematic differences between low and high value properties (International Association of Assessing Officers 2017). COD and PRD are defined as follows:

$$\mathrm{COD} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\hat{Y}_i/Y_i - \mathrm{Median}(\hat{Y}_i/Y_i)\right|}{\mathrm{Median}(\hat{Y}_i/Y_i)} \tag{8}$$

$$\mathrm{PRD} = \frac{1}{N}\frac{\sum_{i=1}^{N}\left(\hat{Y}_i/Y_i\right)}{\sum_{i=1}^{N}\hat{Y}_i/\sum_{i=1}^{N}Y_{ii}} \tag{9}$$

6 Experimental design, results and analysis

6.1 Experimental setting

In this study, our proposed framework, which integrates MLP and CNNs, is compared to the single MLP methodology. The single MLP is used as a benchmark model, referred to as Model 1, and the proposed fused model is referred to as Model 2. Comparing Model 1 and Model 2 in terms of the accuracy of residential price estimation quantifies the contribution of additional information of location and neighbourhood features that are not included in numeric and categorical variables. During the experiments, the dataset is firstly randomly split into training and testing dataset (75:25), and the cross-validation technique is used for hyperparameter tuning. To avoid the issue of spatial autocorrelation in using spatial data, and to train a generalizable model, spatially stratified split is usually recommended in splitting training and validation datasets over random split. A spatially stratified split diminishes the probability that the observations which have autocorrelation with observation in the training set are included in the validation set. It is because observations in training set which are spatially correlated observations in validation set can provide "sneak preview" and it can result in an over-optimistic validation result and poor performance in testing (Lovelace et al. 2019; Salazar et al. 2022). However, the purpose of AVM in most cases is estimation of current price of houses in existing housing market rather than estimating housing prices in new market (e.g. other countries or cities). Also, most of related studies randomly split training and validation datasets (Bency et al. 2017, Bin et al. 2019, Hu et al. 2019, Kang et al. 2020, Kostic and Jevremovic 2020, Gao et al. 2022). In this context, we randomly split train and validation data instead of using spatially stratified split.

Both Model 1 and 2 are trained in Python using TensorFlow and Keras which are Python libraries for neural networks modelling.

Figures 4 and 5 illustrate the structure of the MLP (Model 1) and the fused model (Model 2). The applied MLP has a total of four hidden layers, and each layer has 32, 64, 32, and 16 hidden nodes. The CNN has three convolutional layers and pooling layers, and each convolutional layer includes 16, 32 and 64 kernels with the size of 3×3 for each kernel. Also, 2×2 max-pooling layers are used between convolutional layers along with batch normalisation to prevent 'gradient vanishing', which may downgrade the performance of the neural network (Ioffe and Szegedy 2015).

The sigmoid activation function is applied, and Adam optimiser is employed as a gradient descent algorithm throughout the two models in common. The learning rate of both models is 0.001, and the model was trained using mini-batch sizes 32.

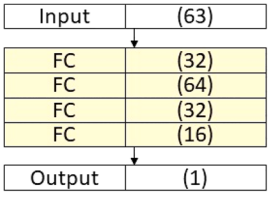

Fig. 4 Structure of Model 1

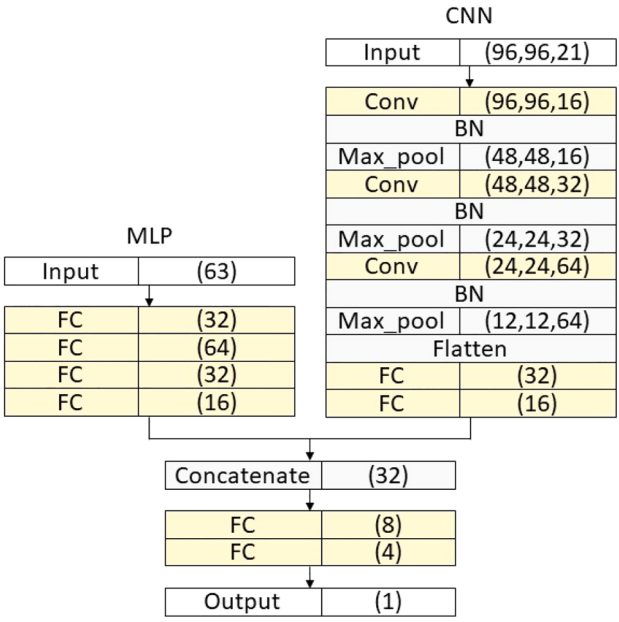

Fig. 5 Structure of Model 2

6.2 Experimental results and discussion

Table 3 shows the modelling results using the testing set. Model 2 performs better than Model 1 in measurements of residential valuation accuracy, based on MAPE (changes from 11.59 to 8.71%), R^2 (changes from about 0.83 to 0.88), PPE5 (changes from 33.57 to 49.89%), PPE10 (changes from 55.92 to 69.71%) and PPE20 (changes from 82.16 to 88.53%). In terms of accuracy of the models, Model 2 shows 2.88% less MAPE which is an improvement in accuracy compared to Model 1. Although it could be perceived as a minor improvement, it is an improvement that can reduce significant costs in the housing market considering the size of the residential real estate market. Also, recent studies related to this study show a similar level of improvement in their results. Bin et al. (2019) suggested an improvement of 2.6% in MAPE between the baseline model, a boosting regression model (20.1% in MAPE), and Attention-based Multi-Modal Fusion (AMMF) that a proposed method in this study. Poursaeed et al. (2018) trained CNN to measure the luxury level of the area at the house such as the bedroom, bathroom, living room, and kitchen using pictures of the house. In the results, by adding the CNN to the existing model, the median absolute percentage error (MdAPE) was decreased by 2.4% from 8.0 to 5.6%. The results of the related studies above can underpin that the improvement in this study is significant.

In terms of uniformity, both models show a COD value lower than 15%, which is the COD value for single-family homes and condominiums advised by the International Association of Assessing Officers (IAAO) (International

Table 3 Modelling result

Performance measurement	Model 1 (MLP)	Model 2 (Fused Model of MLP and CNN)	Improvement between Model 1 and Model 2
MAPE	11.59%	8.71%	−2.88p.p.
R^2	0.8276	0.8786	0.0510
PPE			
5%	33.57%	49.89%	16.39 p.p.
10%	55.92%	69.71%	13.80 p.p.
20%	82.16%	88.53%	6.28 p.p.
COD	11.67%	8.74%	−2.93 p.p.
PRD	1.04	1.02	−0.02

Association of Assessing Officers 2017). COD in Model 2 (8.74%) outperforms Model 1 (11.67%). Furthermore, PRD in Model 2 (1.02) also had a smaller PRD than Model 1 (1.04), which means the results of Model 2 have smaller variability than Model 1 and the PRD value of Model 2 falls in the acceptable range for PRD of 0.98–1.03 that provided by IAAO. On the other hand, the PRD value of Model 1 is higher than acceptable range. It means Model 1 has a tendency that its assessment ratios decline with higher price.

Figure 6 shows the distribution of the actual price and predicted price of each transaction using Models 1 and 2. In this figure, transactions are categorised into four groups by their absolute percentage error (APE). The distribution of results using Model 2 (Fig. 6b) is more correlated with actual price than the result using Model 1 (Fig. 6a). A correlation test between prediction and actual price shows Model 2 has a correlation of 0.9384 with the actual price, which is 0.0234 higher than the result from Model 1 (0.9150). Model 2 has 49.89% of the observations with absolute percentage error less than 5%, while Model 1 has 33.57%.

Figure 7 shows the relationship between the number of observations in the training set and the valuation performance of Models 1 and 2. Both models offer the best performance in houses valued under Aus$2 million, which is 88.36% of the validation observations, and MAPE of both models steadily increased in houses over A$2 million. It might be because the determinants of prices of high-priced houses and their magnitudes are different from those of relatively inexpensive houses, and more observations are needed to train the model. Comparing Models 1 and 2, Model 2 outperformed Model 1 in terms of MAPE, especially in houses valued under $4.5 million. However, for houses over A$4.5 million, the performance of Model 2 is similar to Model 1.

7 Conclusion

The study proposed a fused neural network model based on both the neural network of CNNs and MLP using geographical data with numeric and categorical variables. The performance of the proposed fused model was compared to the

MLP model, which is a baseline model. The model results show that the accuracy of residential valuation has been significantly improved by including a CNN with geographical data. The study found that a CNN in the fused model reduces information loss in transforming geographical information into numeric and categorical data. As the input data, the CNN uses a range of geographical data, such as the distribution of location and neighbourhood characteristics. This allows the CNN to capture the extended geographical patterns such as clusters of facilities and access to amenities. The fused model results show reliable and accurate valuation and improve automated valuation models (AVMs). The proposed framework also has applications for other research problems using geographical data and considering spatial externalities from the surrounding area.

Despite the above contribution of this study, there are limitations. The major limitation of this study is the interpretability of modelling results. The limitation of the machine learning model is caused by modelling the complex relationship between input data and a trade-off between interpretability and completeness (Gilpin et al. 2018). This study focuses more on reducing information loss within a given dataset to improve the performance of the residential valuation model. However, it is also important to mitigate the problem of unobserved information, including omitted location, neighbourhood features and visual features.

Also, there is a lack of exploration of the structure and hyperparameter of the model and input data selection. The proposed fused model requires a longer computing time to train than hedonic price and other machine learning models. To train the fused model, it takes about 30 h (1100 iterations at about 100 s per iteration) in a high-performance computing environment with 4 Tesla-V100 GPU, 32 CPU, and 184 GB memory.

Further studies are required to test various model structures, size of geographical input data, different grid sizes of geographical raster data, input data selections where the high-performance computing challenges the computing barriers.

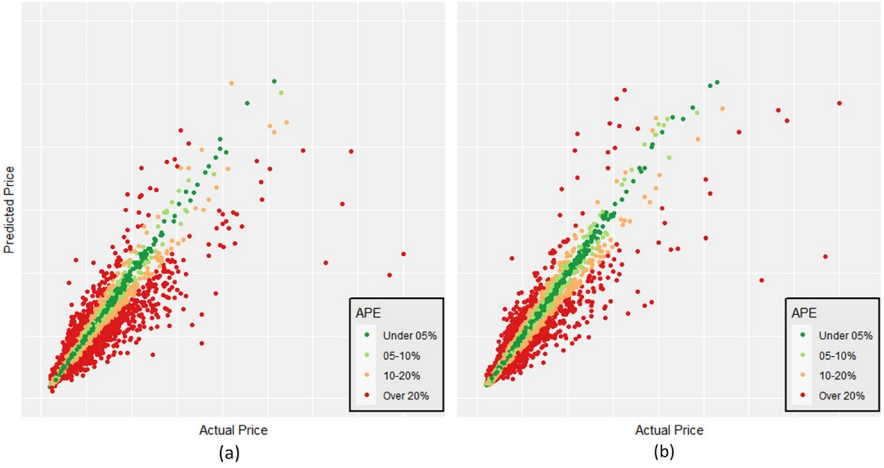

Fig. 6 Comparison between actual price and predicted price: **a** Model 1, **b** Model 2

Fig. 7 Relationship between the number of observations in training set and model performance

Moreover, visual features and image processing of house materials, landscape, front street view, and birds-eye view on neighbouring houses could further improve the performance of the fused model.

Acknowledgements The authors disclose receipt of the following financial support for the research and authorship of this article: This work has been supported by FrontierSI, a not-for-profit company that exists to deliver major benefits to governments, industry and the community in Australia and New Zealand through the application of spatial information. This research was funded through the Cooperative Research Centre Project—Value Australia

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Data and material availability Property sold recodes are proprietary (Australia Property Monitors), Census data is available by contacting Australia Bureau Statistical (ABS) Census, Crime data is available by contacting NSW Bureau of Crime Statistics and Research (BOCSAR), POI data is available by contacting NSW Land & Property Information (LPI), School catchment data is available by contacting NSW Department of Education (DET).

Code availability The codes area available by contacting the authors.

Declarations

Conflict of interst We declare that the work described has not been published before nor is under consideration for publication anywhere else.

References

AlQuhtani S, Anjomani A (2019) Do rail transit stations affect housing value changes? The Dallas Fort-Worth metropolitan area case and implications. J Transp Geogr 79:102463

Alonso W (1964) Location and land use. In: Location and land use. Harvard university press

Australian Bureau of Statistics (Jan 2019) household income and wealth australia 2017–18. ABo Statistics

Barnes Y, Tostevin P (2016) Around the world in dollars and cents 2016. Retrieved 25 April 2018. 198667–198660

Bartholomew K, Ewing R (2011) Hedonic price effects of pedestrian-and transit-oriented development. J Plan Lit 26(1):18–34

Bency AJ et al. (2017) Beyond spatial auto-regressive models: predicting housing prices with satellite imagery. In: 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE

Bin J, Gardiner B, Liu Z, Li E (2019) Attention-based multi-modal fusion for improved real estate appraisal: a case study in Los Angeles. Multimed Tools Appl 78(22):31163–31184

Bourassa SC, Cantoni E, Hoesli M (2007) Spatial dependence, housing submarkets, and house price prediction. J Real Estate Finance Econ 35(2):143–160

Brennan M, Olaru D, Smith B (2014) Are exclusion factors capitalised in housing prices? Case Stud Transp Policy 2(2):50–60

Chen J-H, Ong CF, Zheng L, Hsu S-C (2017) Forecasting spatial dynamics of the housing market using support vector machine. Int J Strateg Prop Manag 21(3):273–283

Copiello S (2020) Spatial dependence of housing values in Northeastern Italy. Cities 96:102444

Cox W, Pavletich H (2020) 16th annual demographia international housing affordability survey:2020

D'Amato M (2017) A brief outline of AVM models and standards evolutions. In: d'Amato M, Kauko T (eds) Advances in automated valuation modeling. Springer, Berlin, pp 3–21

Dubin RA (1988) Estimation of regression coefficients in the presence of spatially autocorrelated error terms. In: Dubin RA (ed) The review of economics and statistics. MIT Press, Cambridge, pp 466–474

Evans AW (1995) The property market: ninety per cent efficient? Urban Stud 32(1):5–29

Feng X, Humphreys BR (2012) The impact of professional sports facilities on housing values: evidence from census block group data. City Cult Soc 3(3):189–200

Fotheringham AS, Brunsdon C, Charlton M (2003) Geographically weighted regression: the analysis of spatially varying relationships. Wiley & Sons, New York

Gao Q, Shi V, Pettit C, Han H (2022) Property valuation using machine learning algorithms on statistical areas in Greater Sydney. Aust Land Use Policy 123:106409

Gilpin LH et al. (2018) Explaining explanations: an overview of interpretability of machine learning. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), IEEE

Goodfellow I, Bengio Y, Courville A, Bengio Y (2016) Deep learning. MIT press, Cambridge

Grover R (2016) Mass valuations. J Prop Invest Finance 34(2):191–204

Helbich M, Griffith DA (2016) Spatially varying coefficient models in real estate: eigenvector spatial filtering and alternative approaches. Comput Environ Urban Syst 57:1–11

Hu S et al (2016) Spatially non-stationary relationships between urban residential land price and impact factors in Wuhan city, China. Appl Geogr 68:48–56

Hu L et al (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

International Association of Assessing Officers (2017) Standard on mass appraisal of real property. Kansas City, Missouri, USA, IAAO

Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167

Jochem WC, Bird TJ, Tatem AJ (2018) Identifying residential neighbourhood types from settlement points in a machine learning approach. Comput Environ Urban Syst 69:104–113

Kang Y et al (2020) Understanding house price appreciation using multi-source big geo-data and machine learning. Land Use Policy 111:104919

Kopczewska K (2021) Spatial machine learning: new opportunities for regional science. Ann Reg Sci. https://doi.org/10.1007/s00168-021-01101-x

Kostic Z, Jevremovic A (2020) What image features boost housing market predictions? IEEE Trans Multimed 22(7):1904–1916
Law S, Paige B, Russell C (2019) Take a look around: using street view and satellite images to estimate house prices. ACM Trans Intell Syst Technol (TIST) 10(5):1–19
Leao SZ et al (2021) A rapid analytics tool to map the effect of rezoning on property values. Comput Environ Urban Syst 86:101572
LeCun Y, Bengio Y, Hinton G (2015) Deep Learn Nature 521(7553):436–444
LeSage JP (2004) A family of geographically weighted regression models. In: Anselin L, Florax RJGM, Rey SJ (eds) Advances in spatial econometrics. Springer, Berlin, pp 241–264
Li Z, Liu P, Wang W, Xu C (2012) Using support vector machine models for crash injury severity analysis. Accid Anal Prev 45:478–486
Li H, Wei YD, Wu Y, Tian G (2019) Analyzing housing prices in Shanghai with open data: amenity, accessibility and urban structure. Cities 91:165–179
Lieske SN, van den Nouwelant R, Han JH, Pettit C (2019) A novel hedonic price modelling approach for estimating the impact of transportation infrastructure on property prices. Urban Stud. https://doi.org/10.1177/0042098019879382
Lieu S et al (2021) Analysis of street environmental factors affecting subjective perceptions of streetscape image in Seoul, Korea : application of deep learning semantic segmentation and YOLOv3 object detection. J Korea Plan Assoc 56(2):79–93
Liu L, Silva EA, Wu C, Wang H (2017) A machine learning-based method for the large-scale evaluation of the qualities of the urban environment. Comput Environ Urban Syst 65:113–125
Lock O, Bain M, Pettit C (2020) Towards the collaborative development of machine learning techniques in planning support systems–a Sydney example. Environ Plan B: Urban Anal City Sci 8(3):484–502
Lovelace R, Nowosad J, Muenchow J (2019) Geocomputation with R. Chapman and Hall/CRC, London
Lu B, Charlton M, Harris P, Fotheringham AS (2014) Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. Int J Geogr Inf Sci 28(4):660–681
McIntosh J, Trubka R, Newman P (2014) Can value capture work in a car dependent city? Willingness to pay for transit access in Perth, Western Australia. Transp Res Part a: Policy Pr 67:320–339
Mulley C, Tsai C-H (2017) Impact of bus rapid transit on housing price and accessibility changes in Sydney: a repeat sales approach. Int J Sustain Transp 11(1):3–10
Mulley C et al (2016) Residential property value impacts of proximity to transport infrastructure: an investigation of bus rapid transit and heavy rail networks in Brisbane, Australia. J Transp Geogr 54:41–52
Mussa A, Nwaogu UG, Pozo S (2017) Immigration and housing: a spatial econometric analysis. J Hous Econ 35:13–25
De Nadai M, Lepri B (2018) The economic value of neighborhoods: predicting real estate prices from the urban environment. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), IEEE.
Osborne PE, Alvares-Sanches T (2019) Quantifying how landscape composition and configuration affect urban land surface temperatures using machine learning and neutral landscapes. Comput Environ Urban Syst 76:80–90
Páez A, Long F, Farber S (2008) Moving window approaches for hedonic price estimation: an empirical comparison of modelling techniques. Urban Stud 45(8):1565–1581
Poursaeed O, Matera T, Belongie S (2018) Vision-based real estate price estimation. Mach vis Appl 29(4):667–676
Powe NA, Garrod G, Willis K (1995) Valuation of urban amenities using an hedonic price model. J Prop Res 12(2):137–147
Reades J, De Souza J, Hubbard P (2019) Understanding urban gentrification through machine learning. Urban Studies 56(5):922–942
Roodposhti MS, Aryal J, Bryan BA (2019) A novel algorithm for calculating transition potential in cellular automata models of land-use/cover change. Environ Model Softw 112:70–81
Rosen S (1974) Hedonic prices and implicit markets: product differentiation in pure competition. J Polit Econ 82(1):34–55
Sadayuki T (2018) Measuring the spatial effect of multiple sites: an application to housing rent and public transportation in Tokyo, Japan. Reg Sci Urban Econ 70:155–173
Salazar JJ, Garland L, Ochoa J, Pyrcz MJ (2022) Fair train-test split in machine learning: mitigating spatial autocorrelation for improved prediction accuracy. J Petrol Sci Eng 209:109885

Scheurwater S (2017) The future of valuations-the relevance of real estate valuations for institutional investors and banks–views from a European expert group. Report of Royal Institution of Charter Surveyors(RICS)
Seo W (2019) Comparing the Housing Implicit Prices of Restricted and Unrestricted Hedonic Price Models. J Korea Plan Assoc 54(6):80–88
Tan F, Cheng C, Wei Z (2019) Modeling and elucidation of housing price. Data Min Knowl Disc 33(3):636–662
The Economist Intelligence Unit (2019) The global liveability index 2019
Warren C, Elliott P (2005) The valuation profession in Australia: profile, analysis and future directions. Aust Prop J 38(5):362
Wu C, Ren F, Hu W, Du Q (2019) Multiscale geographically and temporally weighted regression: exploring the spatiotemporal determinants of housing prices. Int J Geogr Inf Sci 33(3):489–511
Yang L, Zhou J, Shyr OF (2019) Does bus accessibility affect property prices? Cities 84:56–65
Yuan F, Wei YD, Wu J (2020) Amenity effects of urban facilities on housing prices in China: accessibility, scarcity, and urban spaces. Cities 96:102433