Received: 15 November 2017 | Revised: 5 February 2018 | Accepted: 10 February 2018

DOI: 10.1111/tgis.12330

RESEARCH ARTICLE

WILEY Transactions in GIS

Mapping fine-scale urban housing prices by fusing remotely sensed imagery and social media data

Yao Yao[1] | Jinbao Zhang[2] | Ye Hong[2] | Haolin Liang[2] | Jialv He[2]

[1] School of Information Engineering, China University of Geosciences, Wuhan, China

[2] School of Geography and Planning, Sun Yat-sen University, Guangzhou, China

Correspondence
Yao Yao, School of Information Engineering, China University of Geosciences, No. 388 Lumo Road, Wuhan 430074, China.
Email: yaoy33@mail2.sysu.edu.cn

Funding information
National Key R&D Program of China, Grant/Award No.: 2017YFA0604402; National Natural Science Foundation of China, Grant/Award No.: 41671408; Nature Science Foundation of Hubei Province, Grand/Award No.: 2017CFA041

Abstract

The accurate mapping of urban housing prices at a fine scale is essential to policymaking and urban studies, such as adjusting economic factors and determining reasonable levels of residential subsidies. Previous studies focus mainly on housing price analysis at a macro scale, without fine-scale study due to a lack of available data and effective models. By integrating a convolutional neural network for united mining (UMCNN) and random forest (RF), this study proposes an effective deep-learning-based framework for fusing multi-source geospatial data, including high spatial resolution (HSR) remotely sensed imagery and several types of social media data, and maps urban housing prices at a very fine scale. With the collected housing price data from China's biggest online real estate market, we produced the spatial distribution of housing prices at a spatial resolution of 5 m in Shenzhen, China. By comparing with eight other multi-source data mining techniques, the UMCNN obtained the highest housing price simulation accuracy (Pearson $R = 0.922$, OA = 85.82%). The results also demonstrated a complex spatial heterogeneity inside Shenzhen's housing price distribution. In future studies, we will work continuously on housing price policymaking and residential issues by including additional sources of spatial data.

1 | INTRODUCTION

In the last decade, the contradiction between the housing demand from residents and high housing prices has become a top issue in the economy and livelihood of China, especially in metropolitan cities such as Beijing, Shanghai, and Shenzhen (Chen, Guo, & Wu, 2011; Du & Zhang, 2015; Wen & Goodman, 2013). Previous studies show that the main reason for the rising housing prices in Chinese cities originates from the rapid increase in the urban population, which is caused by the largest flow of rural–urban migration in the world due to China's sustained and rapid economic development and urbanization (Chen et al., 2011; Chen, Liu, Li, Liu, & Xu, 2016; Wu et al., 2016; Yiu, Yu, & Jin, 2013). As the world's largest developing country entering the middle-income club, China's housing prices in cities play a very important role in economics at both the macro and micro scales in factors such as gross domestic product (GDP) and household housing/non-housing consumption (Li & Wu, 2014). Moreover, a fine-scale spatial distribution map of urban housing prices can provide the real estate market with valuable information for urban policymaking in order to finely regulate urban housing prices and determine a reasonable level of residential subsidies (Chen et al., 2016), which just have the effect of raising housing prices by the level of the subsidy.

Based on previous studies and the consideration of the complicated spatial heterogeneity of China's urban housing prices (Bitter, Mulligan, & Dall Erba, 2007; Wu, Deng, & Liu, 2014), this study aims to build an effective framework to map a more accurate high-resolution distribution map of urban housing prices by fusing multi-source geospatial datasets via deep neural networks (DNNs) at a fine scale. First, we apply a multi-scale stochastic sampling method to build a housing price correlated spatial dataset from high spatial resolution (HSR) images, points-of-interest (POIs), and basic geographical information. Second, we design a convolutional neural network for united mining (UMCNN) to mine and fuse multi-source spatial data into convolutional neural network (CNN)-based features. Third, we simulate the spatial distribution of housing prices via a random forest (RF)-based fitting model at high resolution. Fourth, our UMCNN-based model is applied to simulate a fine-scale spatial distribution of urban housing prices in Shenzhen, one of the most developed metropolitan areas in China (and even in the world). By analyzing the results of the data mining model input with various features, we have obtained the most highly optimized fused features and urban housing price map.

The remainder of this article is organized as follows. Section 2 reviews the related literature about macro-scale or fine-scale housing price mapping. Section 3 describes the proposed fine-scale housing price mapping approach. Section 4 introduces the study area and dataset used in this study. Sections 5 and 6 report the experimental results and discuss these using different methods. Finally, we offer conclusions in Section 7.

2 | RELATED WORK

Most previous studies about urban housing or renting problems focus on the macro scale, with study data mostly coming from official statistical data and manual surveys (Basu & Thibodeau, 1998; Dub & Legros, 2014; Granziera & Kozicki, 2015; Osland, 2010; Punzi, 2013; Rondinelli & Veronese, 2011). For example, Rondinelli and Veronese (2011) use census data and detailed rental price data provided by real estate developers to estimate the change in rental prices; Feng, Li, and Zhao (2011) use surveying to obtain original data on commercial residential buildings in Beijing to study housing prices; and Wu, Deng et al. (2014) gather newly built housing transaction data from a real estate management system for 35 prime cities in China to build a housing price index and note that China's current housing market has been suffering a greater risk of mispricing than reported by the existing official metrics. These studies adopt the classical hedonic price-regression models, while some researchers also suggest that the spatial autocorrelation and spatial heterogeneity of housing prices require further consideration (Basu & Thibodeau, 1998; Bitter et al., 2007; Dub & Legros, 2014). Classical hedonic price regression models could not handle the spatial factors, so some researchers improve this model and introduce a spatial econometrics model (Osland, 2010), although this model still uses the official statistical data and transaction data provided by developers. In addition, Bitter et al. (2007) also consider the spatial variance of housing prices and conduct geographically weighted regression on house sales data to predict housing prices and obtain higher accuracy than the hedonic housing price model; Kuntz and Helbich (2014) utilize statistical data combined with the kriging interpolation method with consideration of houses' structural and neighborhood characteristics to map real estate prices. Considering the existing spatiotemporal heterogeneity and autocorrelation of housing prices, Wu, Deng et al. (2014) and Fotheringham, Crespo, and Yao (2015) adopted detailed housing statistical data to simulate housing price distribution, which fully takes into account that different influence factors should be considered in different areas. As we know, these data sources have several main problems: first is their high cost, especially high labor cost and long update cycle; moreover, these data are discrete and have difficulty demonstrating the fine spatial patterns of city housing price distributions (Chen et al., 2016; Wu et al., 2016).

To solve the above mentioned problems of study data sources, some research suggests analyzing and obtaining fine-scale housing/rental price data or related data on residents' economic conditions by combining statistical data and remotely sensed imagery mined from natural physical information (Anselin & Le Gallo, 2006; Duque, Patino, Ruiz, &

Pardo-Pascual, 2015; Yu, Wei, & Wu, 2007). Anselin and Le Gallo (2006) use an interpolation method to fit the air quality of the entire area and feed it into the prediction model, obtaining good results; Yu et al. (2007) suggest, based on normal housing attributes, adding remote-sensing images into the housing price study and find that remote-sensing image-generated patterns work well for simulating housing price distributions. As we know, people who have similar social and demographic characteristics may reside in urban areas with similar physical housing conditions. Duque et al. (2015) propose a very high spatial resolution (VHSR) image-based model of texture and structural features extracted from VHSR images and the Slum Index measure of intra-urban poverty to obtain reasonable inter-censal and inter-survey estimates of intra-urban Slum Index maps; the results indicate that housing conditions obtained by remote sensing have strong correlations with the livelihoods of inner-city residents. To the best of our knowledge, the distribution of housing prices is related to socioeconomic factors, such as land-acquisition cost, development cost, marketing fees, and developer profit (Wen & Goodman, 2013), and not all these factors can be reflected via remote-sensing images alone.

In recent years, with the rapid development of Internet techniques, the online real estate market is able to provide massive real-time housing information about real property transactions and rental housing (Chen et al., 2016; Hogan & Berry, 2011; Rae, 2015). Additionally, with the swift development of spatial big data, Liu et al. (2015) and Zheng, Capra, Wolfson, and Yang (2014) introduce the concepts of "social sensing" and "urban computing," thus enabling economic simulations based on social media data. Chang, Lu, Yue, and Li (2014) utilize several large-scale social network datasets of spatial distributions of economic activities, showing that social network data such as human activity is a good indicator to model high-resolution economic activity. Wu et al. (2016) apply check-in data and housing price data from Sofang.com, combined with a geographical weighted regression model, to extend the study of housing prices even deeper. Chen et al. (2016) use rental price data gathered from Anjuke.com and apply an ensemble learning method to predict overall rental prices in Guangzhou, obtaining high accuracy at the level of the neighborhood committee (NC).

In the research described above, we can identify two main problems. First, the prediction models are just regular statistical or machine learning models, which are too simple to fully mine hidden semantic information from complicated geospatial datasets (Huang & Zhang, 2013; Yao et al., 2016; Zhong, Zhu, & Zhang, 2015). Additionally, each of the above studies only utilizes a single type of data, such as manual survey data, social media data, or remote-sensing data, instead of fusing multi-source spatial data to mine the natural physical and socioeconomic factors that affect housing prices, which is helpful for simulating the complex heterogeneity of a housing price distribution, due to the lack of an effective data fusion model (Chen et al., 2016; Duque et al., 2015). In recent years, DNNs (Ciregan, Meier, & Schmidhuber, 2012), including CNNs and recurrent neural networks (RNNs), have been used fully in the fields of computer vision, data mining, and data fusion, and have achieved great results (Simonyan & Zisserman, 2014; Yao et al., 2016; Zhong, Fei, & Zhang, 2016).

The state-of-the-art DNN model also appears from time to time in related studies of remote sensing and mining of spatial big data. For example, Zhong et al. (2016) have solved a land-use classification problem with multi-scale effects via a multi-scale stochastic sampling method and large patch convolutional neural network (LPCNN), in which LPCNN fully mined the context information of ground components in remote-sensing images (Zhong et al., 2016). Meanwhile, a study published in *Science* introduced a novel data-based CNN model to predict and map consumption expenditures and asset wealth in poor African countries, which indicates that the CNN model can extract patterns well in remote-sensing images and work as a descriptor of economic conditions (Jean et al., 2016). In contrast, studies combining deep learning with geospatial social media data in data mining have begun to appear in recent years, and have obtained relatively good results (Chen & Lin, 2014; Yao et al., 2016). The aforementioned studies hold possibilities for effectively fusing multi-source spatial data and data mining, which is one of the main issues in this study.

3 | METHODOLOGY

The flowchart of the model proposed in this study is illustrated in Figure 1. The purpose of our study is to map the fine-scale spatial distribution of Shenzhen's housing prices by fusing multi-source geospatial information extracted

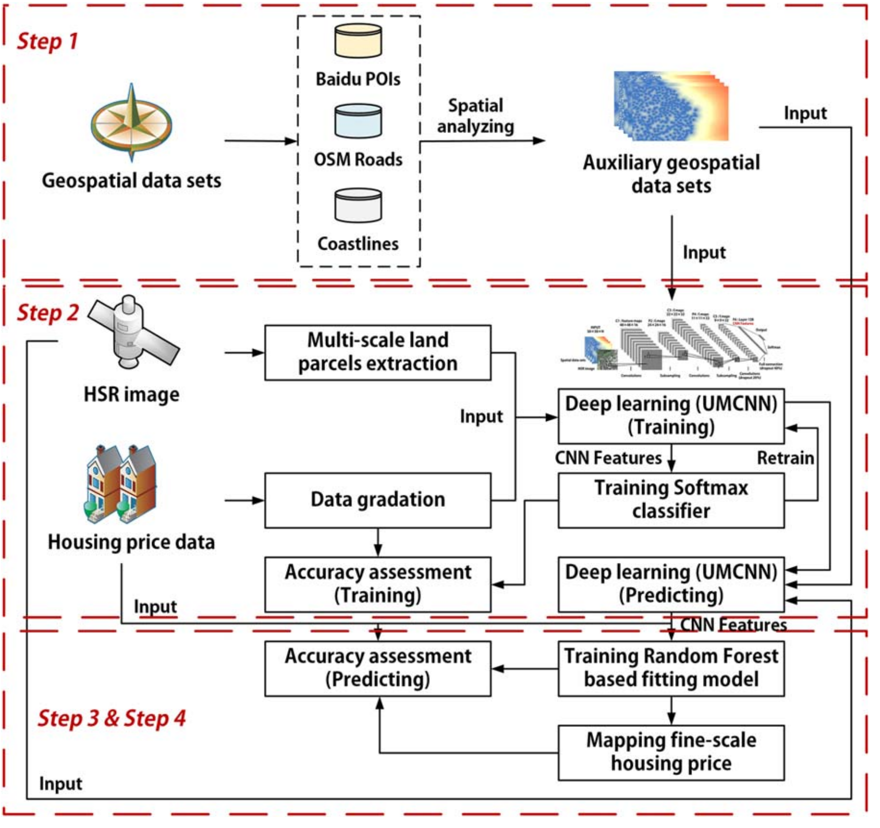

FIGURE 1 The flowchart of mapping housing prices by united mining of multi-source geospatial datasets via UMCNN

from both HSR remote-sensing and social media datasets via a deep learning model. This study uses four steps to obtain a housing price map: (1) via data preprocessing and multiple-scale window sampling, we construct multi-source and multi-scale spatial datasets of impact factors of urban housing prices as training and validation datasets; (2) we train a lightweight convolutional neural network named UMCNN with the aforementioned datasets and graded housing prices, iterate back propagation based on the error calculated from the softmax layer, and obtain the optimal deep learning model; (3) we remove the softmax layer of the pre-trained UMCNN model and take its output, the high-dimensional vectors, as training features to build a RF fitting model with original housing price data; and (4) based on the RF-based fitting model, we calculate the housing price of each pixel with a certain window size and acquire the final housing price results at a fine scale. Moreover, we also build several fusion models with different spatial information extracted from the proposed training dataset in this study and conduct accuracy assessment and uncertainty analysis of the results of different fusion models.

3.1 | Geospatial data preprocessing

As illustrated in Figure 1, auxiliary geospatial data consist of distance or density raster datasets computed from the social media and basic geographical information. The selection of the data covers several main factors that influence housing prices in Chinese cities, including living environment, traffic conditions, and the convenience of life (Li & Wu, 2014; Wu, Deng et al., 2014). In this study, the bandwidth of the Gaussian function-based kernel density analysis is automatically determined according to the mean integrated squared error (MISE) criterion (Wand & Jones, 1994; Yuan, Zheng, & Xie, 2012).

To our knowledge, the CNN model goes through a softmax layer, which can be seen as a multi-class classifier, obtains the classification error, and uses back propagation to adjust the local parameters of each neuron. The

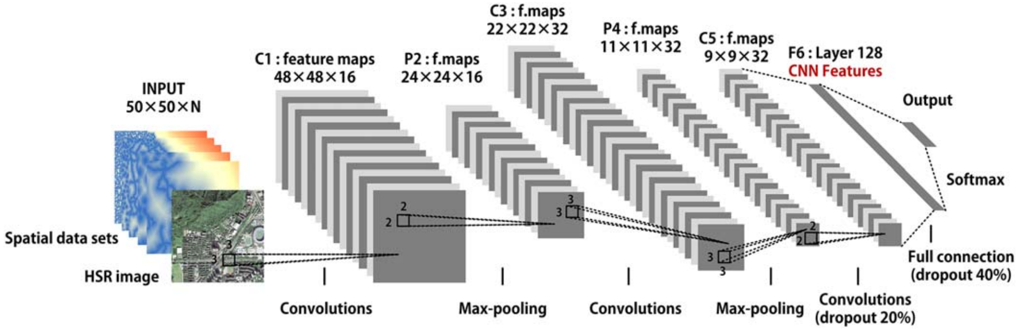

FIGURE 2 The computational framework of proposed UMCNN, used to fuse multi-source datasets

appearance of the best validation accuracy indicates that the optimal CNN model has been generated (Krizhevsky, Sutskever, & Hinton, 2012). The housing price data in this study are continuous, but we discretize them by grading them with different standard-deviation ranks (Jean et al., 2016). First, we assume that the mean and standard deviation of housing prices per square meter in the study area are μ_h and σ_h, respectively. During the data preprocessing, in order to ensure the reliability of data, we remove the housing price samples that lie below 3,000 RMB yuan/m^2 or above the value $\mu_h+3\sigma_h$. Then, we grade the original housing price data with a step of $0.25 \cdot \sigma_h$ within $[\mu_h-3\sigma_h,\ \mu_h+3\sigma_h]$, reducing the noise contained in the original housing price data; the graded data are used as the input parameters of the proposed CNN model.

To solve the scale difference issues of the ground components, the multi-scale stochastic land patch sampling strategy is proposed in this study. First, we set a sampling window centered on each of the training housing price data points with width W. Within each window, we then randomly sample W/s samples, where s is a size that is smaller than W and s gradually increases with a certain step until it is equal to W. A previous study indicates that by multi-scale stochastic patch sampling, the amount of CNN training data can be enlarged, which is helpful to avoid the overfitting problem and solve for the multi-scale effects of ground components in the HSR image (Zhong et al., 2016), improving the classification accuracy of our proposed CNN model.

3.2 | Mapping housing price via UMCNN

The structure of CNN for united mining (UMCNN) designed in our study is shown in Figure 2, which is a relatively less deep model with fewer convolution layers considering the state of computation resources and training datasets. On the one hand, our proposed model is launched via a CPU, where, with limited computation capability, a great deal of time will be needed to generate the optimal model; on the other hand, with regard to the actual housing price data, a complex model will easily cause the overfitting phenomenon. Meanwhile, previous studies also note that with a small dataset, it is essential to reduce the complexity of the CNN network to avoid the overfitting problem (Zhong et al., 2016). Therefore, in consideration of the above factors, we manually tune the optimal CNN network structure.

As demonstrated in Figure 2, the proposed UMCNN contains seven layers, including three convolution layers, two max-pooling layers, a fully connected layer, and a softmax layer (Krizhevsky et al., 2012). The activation function we use in the convolution layers and the fully connected layer is rectified linear units (ReLu), which is a linear activation function with better effect and faster training speed compared with the traditional Sigmoid function (Krizhevsky et al., 2012). We also apply a dropout operation in the third convolution layer and the final fully connected layer, which disables the weights of some neurons randomly, effectively preventing model overfitting (Hinton, Srivastava, Krizhevsky, Sutskever, & Salakhutdinov, 2012). During the convolution operation, we choose a convolution kernel with size 3×3 pixels. Previous studies show that compared with a larger convolution kernel size, the parameters of each kernel are fewer, render a better training result, and have a faster training speed (Simonyan & Zisserman, 2014).

During the training process, the initial size of the input data patch is $50 \times 50 \times N$ (where N depends on whether auxiliary data are used; if we only input the triple-band remote-sensing image, $N = 3$, and $N = 11$ if we add some other auxiliary data, as proposed in the previous section), and bilinear interpolation is conducted on data with an invalid size. Our proposed UMCNN model not only supports mining remote-sensing data, but also supports united mining with multi-source geospatial datasets. As illustrated in Figure 2, the first convolution layer convolutes the input data with 163 × 3 kernels and a step length of 1 pixel (three of the convolution layers have the same step length), generating feature maps of size 48 × 48 × 16. The second layer is the max-pooling layer, where feature maps generated in the first convolution layer are max-pooled with a kernel of size 2 × 2 and a step length of 2 pixels, generating feature maps of size 24 × 24 × 16. The third layer convolutes with 32 convolution kernels, and the output size is 22 × 22 × 32. The adjacent fourth layer is another max-pooling layer, producing 11 × 11 × 32 feature maps. The fifth layer is the same as the third layer, using 32 convolution kernels and generating 9 × 9 × 32 feature maps. The next layer is the fully connected layer with 128 neurons, where the weights of this layer are the housing price fitting features and, finally, a softmax layer is used to classify the housing prices into several classes, which indicates that the grade of housing price is divided by data gradation. Moreover, to avoid the overfitting problem in our proposed UMCNN model, we apply a dropout operation in the last two layers before the fully connected layer by disabling 20 and 40% of the connections among neurons. This operation can reduce complex co-adaptations of neurons by incorporating stochastic error, aiming to ensure the generalization ability (Abadi et al., 2016).

After training the proposed UMCNN model in the previous step, we remove the last softmax layer and the UMCNN model has become a feature extractor, which can extract high-dimensional semantic features using a fully connected layer (Jean et al., 2016). By inputting the UMCNN-based features into a RF-based classifier, a fitting model between features and actual housing prices is built to generate fine-scale housing price distribution results. The RF model has been proved to be an outstanding state-of-the-art machine learning model, which obtains relatively better results in many classification and regression tasks (Biau, 2012; Fernández-Delgado, Cernadas, Barro, & Amorim, 2014). To clarify, RF is an aggregation of decision-tree classifiers. A new sub-dataset is generated by extracting random samples from the original training dataset via the bagging method (Biau, 2012; Breiman, 2001). In the process of random feature selection, individual decision trees are constructed from each training sub-dataset, and these decision trees are not pruned during the growth process so we can obtain an out-of-bag (OOB)-based estimation error report for each decision tree. The generalization error of RF can be calculated by averaging the errors of the decision trees via OOB estimation. The RF-based fitting model introduced in previous studies overcame the multiple correlative problems among spatial variables, especially in higher-dimensional fitting situations (Palczewska, Palczewski, Marchese Robinson, & Neagu, 2014).

3.3 | Accuracy assessment and uncertainty analysis

To evaluate the final accuracy of our proposed model in predicting housing price distributions, this study randomly split the original housing price data into training and validation parts, which can be used to cross-validate the proposed UMCNN model. When mapping fine-scale housing prices, OOB-based estimation is used to estimate the accuracy of CNN-based features and related actual housing price data. Previous studies have indicated that the OOB-estimated fitting error has been proved to have better effects than cross-validation (Biau, 2012; Fernández-Delgado et al., 2014). To ensure the reliability of the classification result, we apply the bagging method in this study by randomly dividing part of the data into OOB data, iterating the training and predicting process 100 times, and obtaining the average prediction accuracy; the specific parameters will be discussed in the next section. Moreover, we adopt several accuracy assessment methods to identify the performance of the proposed RF-based fitting model, including Pearson's correlation coefficient (Pearson R), the standard coefficient of determination (Standard R^2), the root mean squared error (RMSE) and its percentage (%RMSE), and the mean absolute error (MAE) and its percentage (%MAE):

$$\text{Pearson } R = \frac{\sum_{i=1}^{n}\left(r_{i,o}-\bar{r}_o\right)\left(r_{i,s}-\bar{r}_s\right)}{\sqrt{\sum_{i=1}^{n}\left(r_{i,o}-\bar{r}_o\right)^2}\sqrt{\sum_{i=1}^{n}\left(r_{i,s}-\bar{r}_s\right)^2}} \tag{1}$$

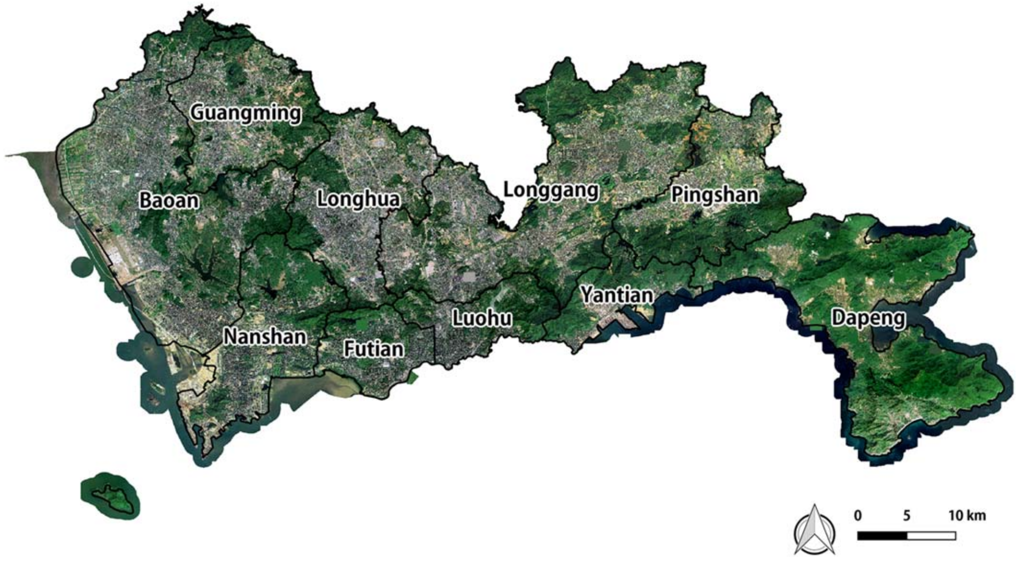

FIGURE 3 Study area: Shenzhen, Guangdong province. The background data are the HSR remote-sensing image provided by Tianditu.cn with a spatial resolution of 5 m (size: 13,976 × 22,514)

$$\text{Standard } R^2=\frac{n\sum_{i=1}^{n} r_{i,o}r_{i,s}-\sum_{i=1}^{n} r_{i,o}\sum_{i=1}^{n} r_{i,s}}{\sqrt{n\sum_{i=1}^{n} r_{i,o}^2-\left(\sum_{i=1}^{n} r_{i,o}\right)^2}\sqrt{n\sum_{i=1}^{n} r_{i,p}^2-\left(\sum_{i=1}^{n} r_{i,p}\right)^2}} \tag{2}$$

$$\text{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(r_{i,o}-r_{i,s}\right)^2} \tag{3}$$

$$\%\text{RMSE}=\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(r_{i,o}-r_{i,s}\right)^2}}{\bar{r}_o} \tag{4}$$

$$\text{MAE}=\frac{1}{n}\sum_{i=1}^{n}\left|r_{i,o}-r_{i,s}\right| \tag{5}$$

$$\%\text{MAE}=\frac{1}{n}\sum_{i=1}^{n}\frac{\left|r_{i,o}-r_{i,s}\right|}{r_{i,o}} \tag{6}$$

where $r_{i,o}$ and $r_{i,s}$ are the actual and simulated housing prices (unit: RMB yuan/m^2) for the ith building, and n is the total number of buildings in the housing price data obtained from Fang.com. After the evaluation of RF-based fitting models, we also compare the predicted housing price results with multi-source official statistical data at different scales to evaluate the applicability of the results.

4 | STUDY AREA AND DATA

This study area is located in Shenzhen (Figure 3), Guangdong province, with a total area of 1,996.850 km^2 and a permanent population of approximately 10.779 million. Shenzhen is considered one of the largest international metropolitan cities and economic centers in China, with a total GDP of 1,750.299 billion RMB yuan in 2015 (http://www.sztj.gov.cn/xxgk/tjsj/tjnj/). As illustrated in Figure 3, Shenzhen has 10 administrative county-level districts (Futian, Luohu, Nanshan, Yantian, Baoan, Guangming, Longhua, Longgang, Pingshan, Dapeng), within 60 street-level divisions and 734 basic NCs. Among all these districts, Futian, Luohu, Nanshan, and Yantian are the earliest special economic zones in Shenzhen, and still the most populated and developed districts of Shenzhen, accounting for 52.48% of the total GDP of Shenzhen. As one of four first-tier cities in China, and the city with the highest immigrant ratio, Shenzhen's urban land-use patterns, migrant population distribution, and economic structure are very complex, which has caused complicated and diverse socioeconomic problems, such as urban sprawl, environmental deterioration, traffic congestion,

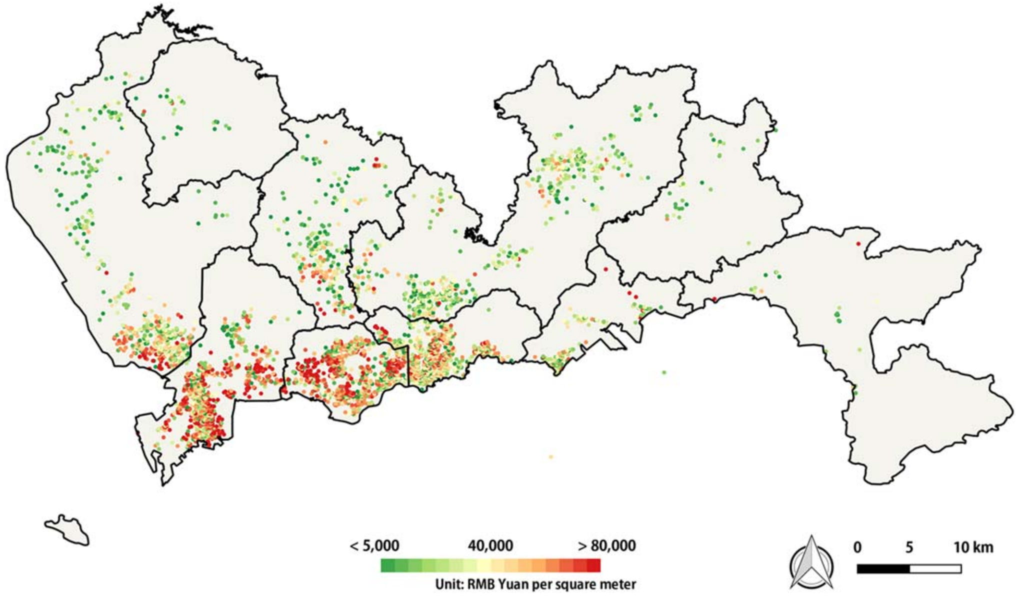

FIGURE 4 The acquired housing price data from Fang.com, China's biggest online housing market website

soaring housing prices, etc. (Chen et al., 2011; Hao, Sliuzas, & Geertman, 2011; Wang, Wang, & Wu, 2010). Therefore, this study aims to develop an effective method to map housing prices at a fine scale, offering references and recommendations for the relevant governmental departments to adjust the housing prices in Shenzhen.

Housing price data are the most important data in this study, provided by Fang.com, the most popular and largest online real estate market website in China. Fang.com provides real-time rental and sale prices of newly built and second-hand houses with abundant information for 651 cities in China, and is already influencing online real estate agencies in Shenzhen. To our knowledge, most existing housing price index construction models rely mainly on transaction data from the second-hand housing market (Rondinelli & Veronese, 2011; Wu, Deng et al., 2014), but previous studies have proposed that housing markets in most Chinese cities are currently dominated by the newly built sector as a direct result of the large volume of new supply (Wu, Deng et al., 2014). In this study, we have compiled a program of web crawlers to collect the newly built and second-hand housing prices in the study area from Fang.com, along with the attributes of identity number, latitude, longitude, price/m^2, residential quarter name, size of house, and number of rooms. After correcting geographic coordinates, eliminating outliers, and cleaning data, Figure 4 displays 4,331 sets of valid data. The number of housing price data of each district is displayed as follows: Futian (1,091), Luohu (811), Nanshan (951), Yantian (102), Baoan (474), Guangming (37), Longhua (359), Longgang (450), Pingshan (44), and Dapeng (12). By analyzing the data obtained from Fang.com statistically, it is determined that housing prices in Shenzhen follow a log-normal distribution (kurtosis = 2.724, skewness = −0.774), where the average housing price is 36,571.771 RMB yuan/m^2, with a standard deviation of 36,515.693 RMB yuan/m^2; more detailed analysis will be discussed in the following sections.

In addition to the HSR Worldview-2 remote-sensing image, which has been downscaled to 5 m from Tianditu.cn (Figure 3), several basic geographic and social media datasets, including POIs and OpenStreetMap (OSM) road nets, are also applied in our study. As shown in Figure 4, the POIs in our study are provided by Gaode Map Services (http://lbs.amap.com/), one of the most popular and biggest web map service providers in China. Approximately 211,076 records with 432 categories in the study area were acquired with web crawlers via Gaode Map APIs (Figure 5), including business establishments, commercial sites, educational facilities (kindergartens, primary schools, middle schools, etc.), residential communities, clinical facilities, and scenic locations. In the Chinese real estate market, the most influential factors in housing prices are traffic, medical treatment, infrastructure, and the convenience of basic living facilities (Basu & Thibodeau, 1998; Wu, Deng et al., 2014). Therefore, we extracted housing price-related ground components

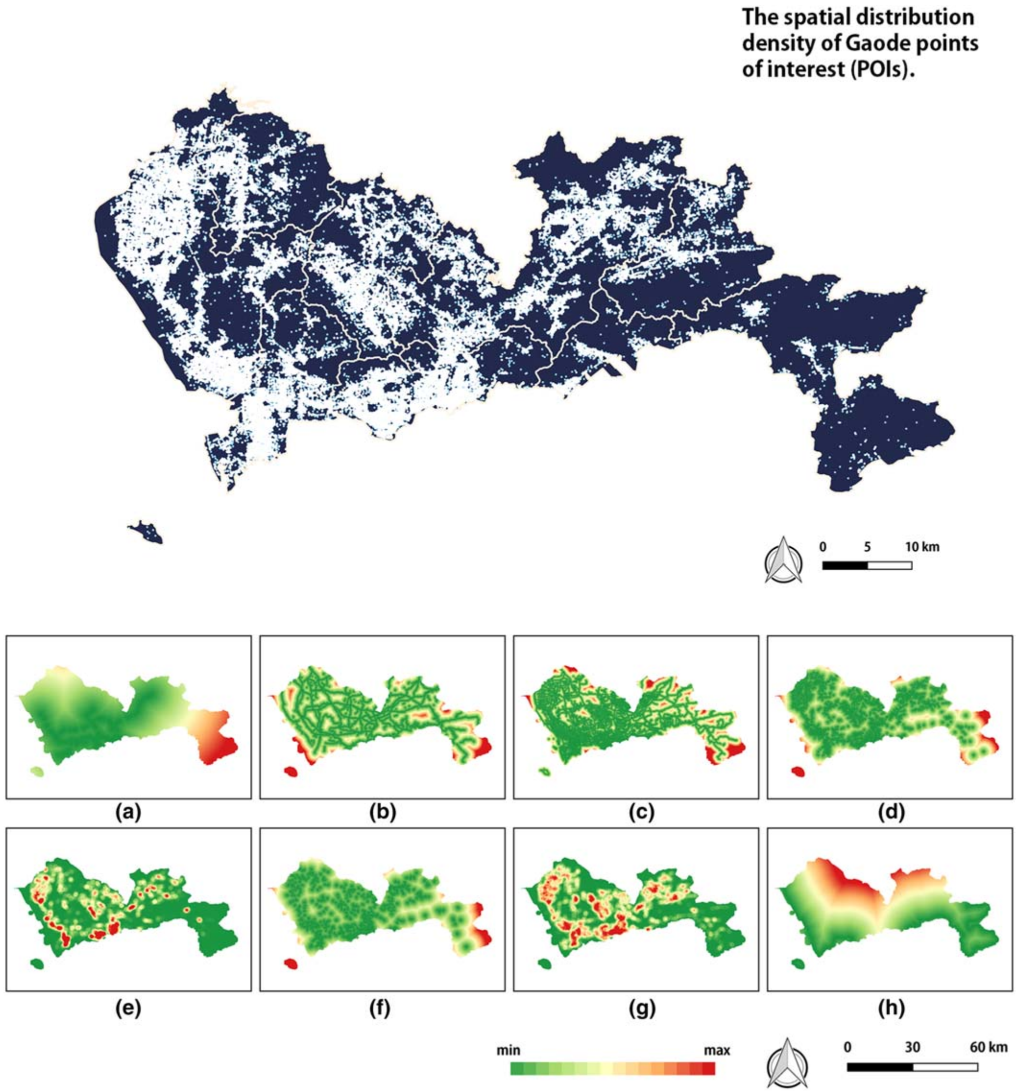

FIGURE 5 The spatial distribution density of Gaode POIs and the auxiliary geospatial datasets: (a) distance to subway stations; (b) distance to main roads; (c) distance to other roads; (d) distance to medical facilities; (e) density of living facilities; (f) distance to primary and secondary schools; (g) density of bus stations; and (h) distance to coastlines

from POIs, such as subway stations, bus stations, primary and secondary schools, and medical and living facilities, to form the distance- or density-based auxiliary geospatial datasets (Figure 5). Previous study has recommended that several explanatory variables for mapping housing price in Shenzhen (Wu, Deng et al., 2014), such as land area, green space, building structure type, quality, etc., can be well reflected from HSR images. Moreover, considering the impact of traffic and environmental condition on housing prices (Feng et al., 2011; Wang et al., 2010), we also take the density of road nets and distances to main roads and coastline into consideration, as shown in Figure 5.

5 | RESULTS

Our research team built a software application and realized the model proposed in Section 3 using C++ on Windows Server 2008 with a 32-core CPU. Several open-source C/C++ libraries, such as the CGAL (http://www.cgal.org), GDAL (http://www.gdal.org/), and SHARK (http://image.diku.dk/shark/) libraries, were applied to this project for

TABLE 1 The methods of mapping housing prices via different fusion or combination methods used to fuse multi-source features

Expr. ID	Expr. description	Label
A	Using original features extracted from HSR images via CNN only	CNN (HSR)
B	Using lower-dimensional features extracted from HSR images via CNN only	PCA-CNN (HSR)
C	Using spatial datasets extracted from SD only	SD
D	Using original features extracted from HSR images and spatial datasets, and fusing by CNN	CNN (HSR & SD)
E	Using lower-dimensional features extracted from HSR images and spatial datasets, and fusing by CNN	PCA-CNN (HSR & SD)
F	Combination of original features extracted from HSR images via CNN and original features extracted from SD	CNN (HSR) & SD
G	Combination of lower-dimensional features extracted from HSR images via CNN and original features extracted from POIs	PCA-CNN (HSR) & SD
H	Using original features extracted from HSR images via SD only	CNN (SD)
I	Combination of features extracted from HSR and SD via CNN respectively	CNN (HSR) & CNN (SD)

mapping the spatial distribution of housing prices using UMCNN. The source codes of the proposed UMCNN platform are implemented in C++ with OpenMP and run on a multi-processor computation server. The results of mapping fine-scale housing prices in the study area have been released on the GeoSOS website (http://geosimulation.cn/ydoc/shenzhen_hp/shenzhen_hp_5m.zip).

5.1 | Mapping housing prices via different feature combinations

By the aforementioned multi-scale land parcel sampling method, we set the limited size of the sampling window as 50 pixels and obtain a multi-scale spatial dataset D containing more than 25,000 records. Then, we split the dataset D into the training dataset D_T, which is fed into the proposed UMCNN model, and the testing dataset for evaluating the final accuracy, with ratios of 80 and 20%, respectively. During the training process of the UMCNN model, we randomly pick 80% of the data in D_T for model training, and the remaining 20% of the data are used for validation and error back propagation.

After the training of the CNN has been completed, an RF-based fitting module is employed to replace the original softmax layer, and then this module is sent for training in order to fit the actual housing prices. We split the original housing price data into two parts: 60% (2,598 records) for training data and 40% (1,732 records) for validation data, to evaluate the fitting accuracy of this module. Previous studies indicate that RF is an effective and novel non-parametric machine learning model, which is widely used to solve high-dimensional non-linear fitting and classification issues, with relatively good effect (Fakhraei, Soltanian-Zadeh, & Fotouhi, 2014; Fernández-Delgado et al., 2014). We set up 100 decision trees and 20% OOB data, then cross-validate the result with boosted random sampling and iterate 100 epochs to obtain the average accuracy for the most reliable result.

As illustrated in Table 1, we designed nine different experiments in order to highlight the advantage of our proposed model. Within these nine experiments, Groups A and C only use features extracted from an HSR image via UMCNN and several spatial datasets, respectively, while Groups D and F use two different data fusion methods: Group D manages to apply UMCNN for united mining of the deep features of multi-source spatial datasets, while Group F directly combines the CNN-based features via HSR images with spatial data. Furthermore, a previous study adopted principal component analysis (PCA) to reduce the dimension of UMCNN-based features to obtain better classification results (Jean et al., 2016). Therefore, in Groups B, E, and G, we take the top 10 principal components of the CNN-

TABLE 2 The accuracy assessment results of different combination methods: (A) CNN (HSR), (B) PCA-CNN (HSR), (C) SD, (D) CNN (HSR & SD), (E) PCA-CNN (HSR & SD), (F) CNN (HSR) & SD, (G) PCA-CNN (HSR) & SD, (H) CNN (SD), (I) CNN (HSR) & CNN (SD)

Expr.	A	B	C	D	E	F	G	H	I
Pearson R	0.905	0.878	0.824	0.922	0.882	0.915	0.900	0.869	0.916
Standard R^2	0.724	0.665	0.586	0.745	0.664	0.752	0.724	0.606	0.739
RMSE	19.180	21.135	23.484	18.456	21.172	18.172	19.189	20.781	18.658
%RMSE	15.73%	17.33%	19.25%	15.13%	17.36%	14.90%	15.73%	17.04%	15.30%
MAE	9.568	10.056	14.492	9.757	10.279	9.523	10.308	12.907	9.901
%MAE	24.54%	25.70%	28.71%	26.20%	28.65%	24.19%	27.19%	28.39%	26.11%
t Test	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

based high-dimensional features to build fitting models and evaluate their accuracy and validity. Additionally, we designed Groups H and I to verify the potential validity of the proposed UMCNN in multi-source spatial dataset mining.

Table 2 presents the accuracy assessment results of the fitting models with different input features shown in Table 1, where we can observe some interesting phenomena. In general, each experimental group reaches a relatively high total accuracy, where Pearson's correlation coefficients are higher than 0.8. Through comparison among the four groups of experiments with the highest accuracy (Groups D, I, F, and A), it can be ensured that when the input data contained features from both an HSR image and auxiliary geospatial data, the RF-based fitting models (Groups D, F, and I) obtain higher accuracies than fitting models that use only a single type of input data (Groups A, C, and H). This indicates that united mining on multi-source spatial datasets—or just simply applying feature combination—is able to fuse the natural physical and socioeconomic information from multiple sources effectively. In contrast, CNN works well for mining high-level semantics and context information (Chen & Lin, 2014; Simonyan & Zisserman, 2014; Yao et al., 2016; Zhong et al., 2016), which helps achieve high classification accuracy. In addition, we apply a multi-scale sampling method and reduce the dimensions of UMCNN-based features, which solves the problem that exists in Jean et al.'s (2016) model, where the number of feature dimensions is much larger than the size of the training dataset. PCA is used to simplify the CNN-based features, but it also misses part of the feature information, resulting in some loss of fitting accuracy (Geiger & Kubin, 2012; Lu, Wang, Wang, Yan, & Lam, 2004). At the same time, as far as we know, single-band spatial data (SD) possesses very little structural and textural information, but multi-band SD is equipped with spatial features of rich object attributes. Through comparing the results of Groups C and H, where Group H obtains relatively higher accuracy, the ability of the proposed UMCNN to effectively exploit the potential spatial structure and high-level semantic features from the multi-source SD is certified.

As shown in Figure 6, we compare the results of different fitting models with actual housing price data at the NC level. Traditional approaches that fit the model via auxiliary geospatial datasets (Chen et al., 2016; Wu et al., 2016), such as that shown in Figure 6c, can only obtain relatively high accuracy in areas with concrete human activities, such as downtown areas, but lower accuracy in regions with sparse human activities and economic backwardness due to a lack of online real estate market data. Zhong et al. (2016) notes that CNN can extract the local features and generalize the global information directly from the ground component level instead of the sub-object parts. However, even using the features obtained from mining spatial data via UMCNN to fit the housing prices (Figure 6h), the resulting error distribution is still similar to the conventional approaches (Figure 6c).

With only local high-level semantic features mined from remote-sensing images by UMCNN, the fitting model (Figure 6a) obtains results with better accuracy than mining only with auxiliary geospatial data; after united mining with auxiliary geospatial data and remote-sensing image features, we observe an apparent increase in the prediction accuracies of the housing prices in the suburban areas near downtowns. These results prove one of our assumptions: housing price is strongly correlated with building location, as well as with many potential neighborhood environmental

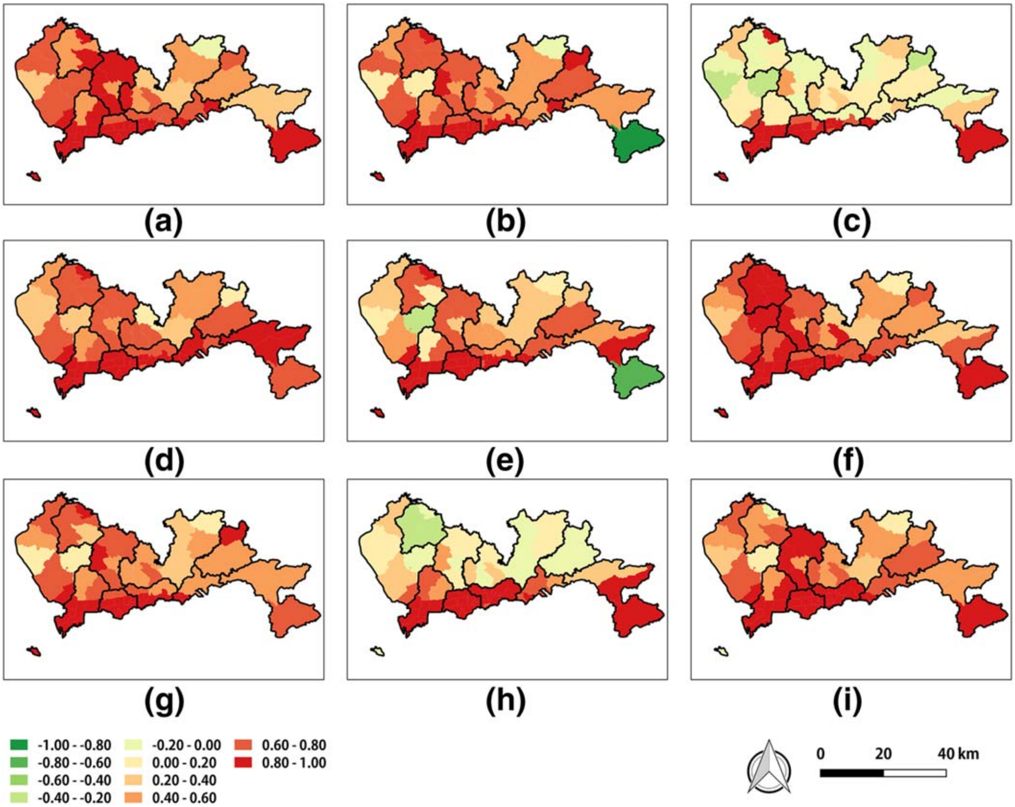

FIGURE 6 The correlation analysis results of different combination methods between actual housing price data and simulation results at the NC level: (a) CNN (HSR); (b) PCA-CNN (HSR); (c) SD; (d) CNN (HSR & SD); (e) PCA-CNN (HSR & SD); (f) CNN (HSR) & SD; (g) PCA-CNN (HSR) & SD; (h) CNN (SD); and (i) CNN (HSR) & CNN (SD)

factors (such as the distance to the metro station and the density of living facilities), which can be extracted well from the proposed auxiliary geospatial datasets.

Additionally, Figures 6d, f, and i blend the information well from multi-source spatial datasets and have similar error distribution patterns. However, we can also find that Figure 6f shares a similar error distribution with Figure 6a, with corrections in some of the areas, and therefore improves the fitting accuracy. To our knowledge, the number of dimensions extracted by CNN is far larger than the number of features extracted from auxiliary geospatial data. This indicates that the feature combination method cannot fuse the multi-source spatial data properly because all fitting models, when applied to a high-dimensional problem, have a high tendency to only select some of the features from some dimensions, which is called “the class imbalance problem” and may cause overfitting errors (Wasikowski & Chen, 2010). The methods adopted by Groups I and D can effectively avoid the class imbalance problem, but with similar overall accuracy (Table 2), while the training time and parameter adjustment requirements in Group D, which employs a joint mining strategy, are only half those in Group I. To sum up, the proposed united mining approach based on HSR images and geospatial big data (Group D) can effectively and reasonably fuse multi-source spatial data and obtain the best prediction accuracy.

To further explain the advantages of our proposed CNN (HSR & SD) model, 20 input parcels were randomly selected from the study area, and trained with adjusted CNN (HSR), CNN (SD), and CNN (HSR & SD), outputting the features of the first convolution layer. Through inspecting the randomly selected samples in the study area, we ensure that the chosen sample covers the house price distribution in the vast majority of locations. By employing the PCA dimensionality reduction method to the features of the first convolution layer, the first three principal components are synthesized into RGB false color images, as shown in Figure 7. In Figure 7, areas with higher hue values indicate the

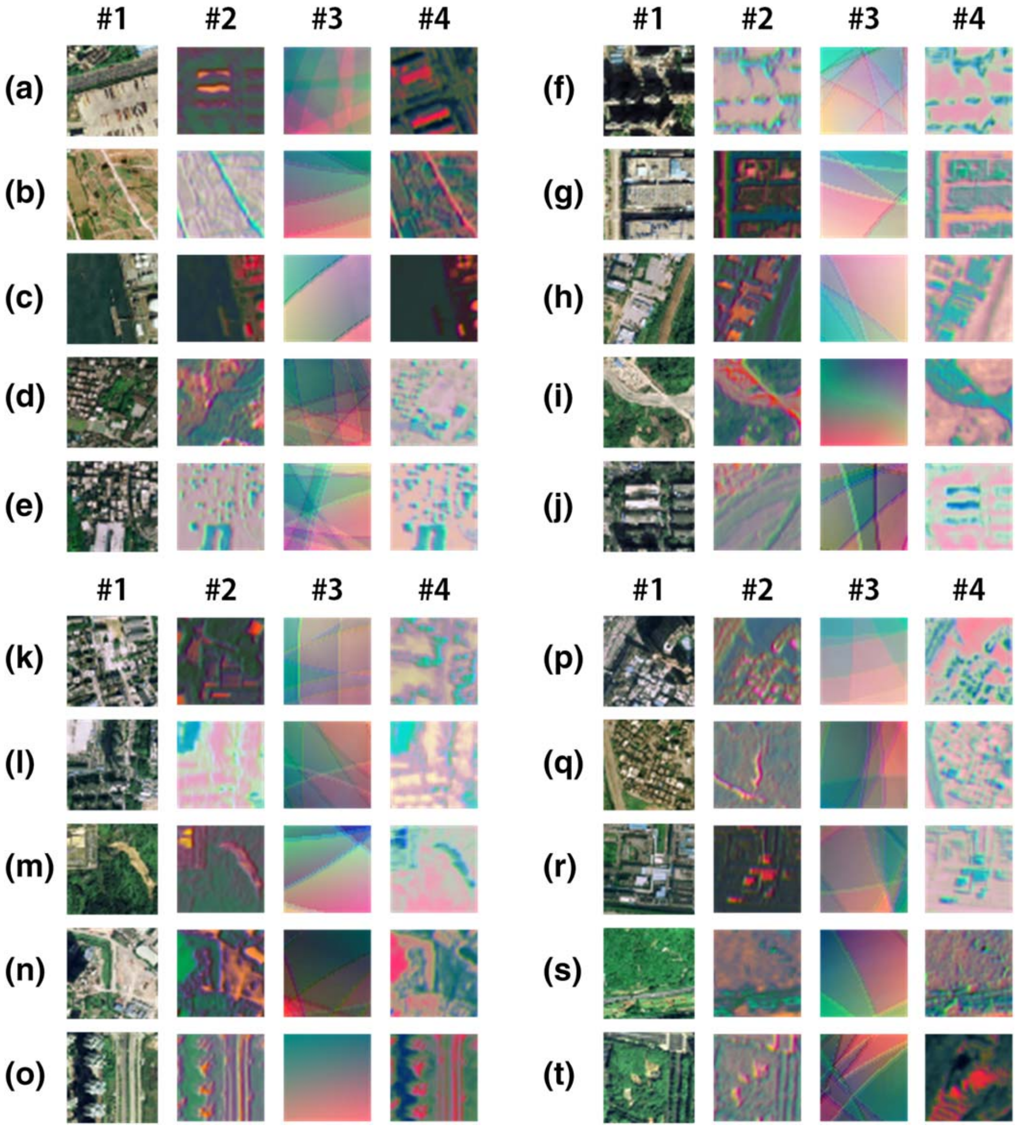

FIGURE 7 UMCNN first-layer convolution output graph of the random samples: (#1) original remote-sensing image parcels; (#2) dimensionality reduction output of CNN (HSR); (#3) dimensionality reduction output of CNN (SD); (#4) dimensionality reduction output of CNN (HSR & SD)

higher importance in the subsequent determination. It can be found that UMCNN excavates not only the structural features of high-resolution remote-sensing images (Figure 7, #2), but also the structural features of multi-source spatial variable sets (Figure 7, #3). The spatial function structure and land use of the land parcels are strengthened and highlighted through combined mining of both feature sets (Figures 7a–j), which demonstrates that CNN (HSR & SD) is able to fuse the physical and economic features quite well, and more accurately identify the housing structure, urban land use, and their relation to housing prices.

It can be found from Figures 7k–t that, in comparison with CNN (HSR), CNN (HSR & SD) identified other core affected areas (such as water and roads, etc.) on the basis of strengthening the spatial functional structure within the land parcel, which shows that in the training process of CNN (HSR & SD), various driving factors affecting housing prices were explored based on remote-sensing images and multi-source spatial datasets, so the result of convolution is not limited to the spatial structure of housing obtained from remotely sensed imagery. In the prediction process, CNN

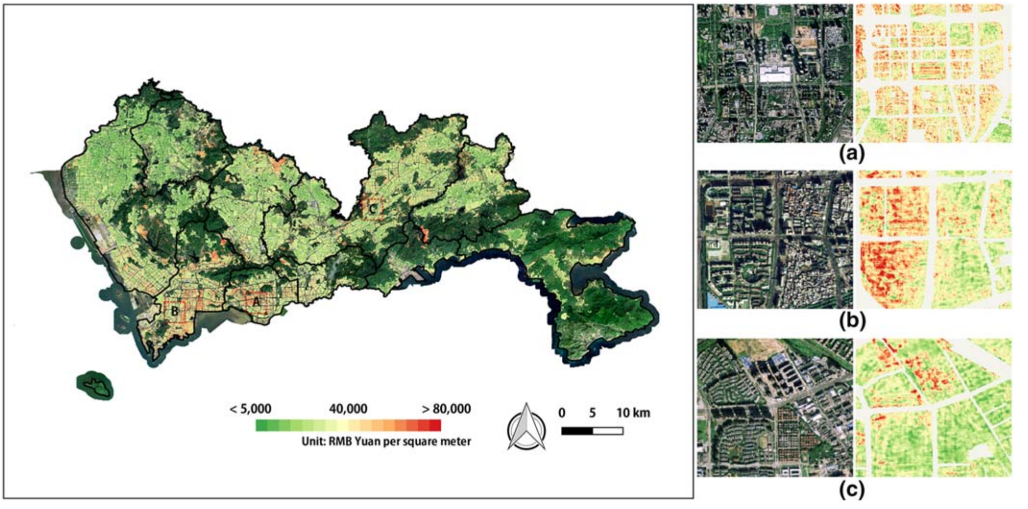

FIGURE 8 Some details of the spatial distribution of Shenzhen's housing prices simulated via CNN (HSR & SD) in Figure 9: (a) Shenzhen's public administration center in Futian district; (b) Nantou in Nanshan district; and (c) Zhonghai real estate in the Longgang district

(HSR & SD) adopts diverse rules to establish a spatial non-linear housing price model. In general, the properties of the house itself play a leading role in the development of housing prices, while the peripheral traffic, economic, natural environment, and other factors also occupy a very important impact ratio.

5.2 | Spatial distribution of housing prices at a fine scale

Based on the above result and analysis, we adopt a UMCNN trained via CNN (HSR & SD) to predict the housing prices, in units of "RMB yuan/m^2," per pixel in the study area, with a spatial resolution of 5 m; the results are shown in Figure 8. Generally, we can observe that the high housing price areas ($\geq$ 40,000 RMB yuan/m^2) are mainly located in the center of Futian, in Nanshan district, along the Shenzhen coastline, and at the junction area of Yantian and Futian districts, where the overall accuracies of these areas are higher than 90%, as demonstrated in Table 3. Longgang, Pingshan, Dapeng, and Guangming districts have lower overall accuracies (<70%), because these regions are newly developed, resulting in a lack of housing price samples and large amounts of "shanty towns" (Chen et al., 2016), causing fitting errors in the UMCNN model.

Figure 9 illustrates the overall histograms of housing price data from Fang.com and simulated fine-scale housing prices, where the overall histograms are similar to the standard normal distribution. The housing price data provided by Fang.com were mainly sampled from the newly built and second-hand houses in Shenzhen, ignoring most of the fully residential communities and relatively cheaper real estate in rural areas, causing a low average housing price in the simulated results; however, an overall accuracy of 85.82% is still achieved. In addition, we also find something interesting in Figure 9, where a peak value exists between 40,000 and 50,000 RMB yuan/m^2 in both the actual and simulated data. By comparison with remote-sensing images, we found that the areas within a certain range centered at this peak value are filled with newly developed high-rise residential buildings, which always have stable housing prices. In general, via the united mining method proposed in this study, we can map Shenzhen's housing prices at a fine scale, especially in the economically well-developed regions.

To further illustrate the fine scale and rationality of our proposed model, in Figure 8 we present the detailed housing price distributions of three typical areas in Shenzhen, and the overall results are also available online. The area in Figure 8a includes the public administration center of Futian district, Shenzhen, which is also the location of the Shenzhen convention center and the largest shopping mall; thus, it has the highest housing prices in Shenzhen. An urban village named Lianhua village, located in the northeast corner of Figure 8a, has relatively lower housing prices compared

TABLE 3 The average values, standard deviations, and overall accuracies of different districts in Shenzhen

	Actual		Simulation		
District	Mean	St. dev.	Mean	St. dev.	OA
Futian	44,114.810	31,617.661	42,006.874	16,030.465	95.22%
Luohu	36,556.088	36,729.468	37,334.489	13,845.462	97.87%
Nanshan	49,573.861	27,338.270	43,176.652	18,074.621	87.10%
Yantian	30,785.033	20,562.288	30,415.693	16,429.687	98.80%
Baoan	29,170.748	19,937.242	27,192.262	14,432.516	93.22%
Longhua	25,491.366	20,098.939	30,331.736	14,186.353	81.01%
Longgang	23,726.156	14,630.571	31,079.579	14,513.756	69.01%
Pingshan	18,516.147	12,633.294	30,040.577	16,023.564	37.76%
Dapeng	21,401.600	22,374.080	32,588.309	15,594.998	47.73%
Guangming	15,220.643	15,559.348	24,120.337	13,330.930	41.53%
Shenzhen	36,571.771	20,515.693	31,387.255	15,922.300	85.82%

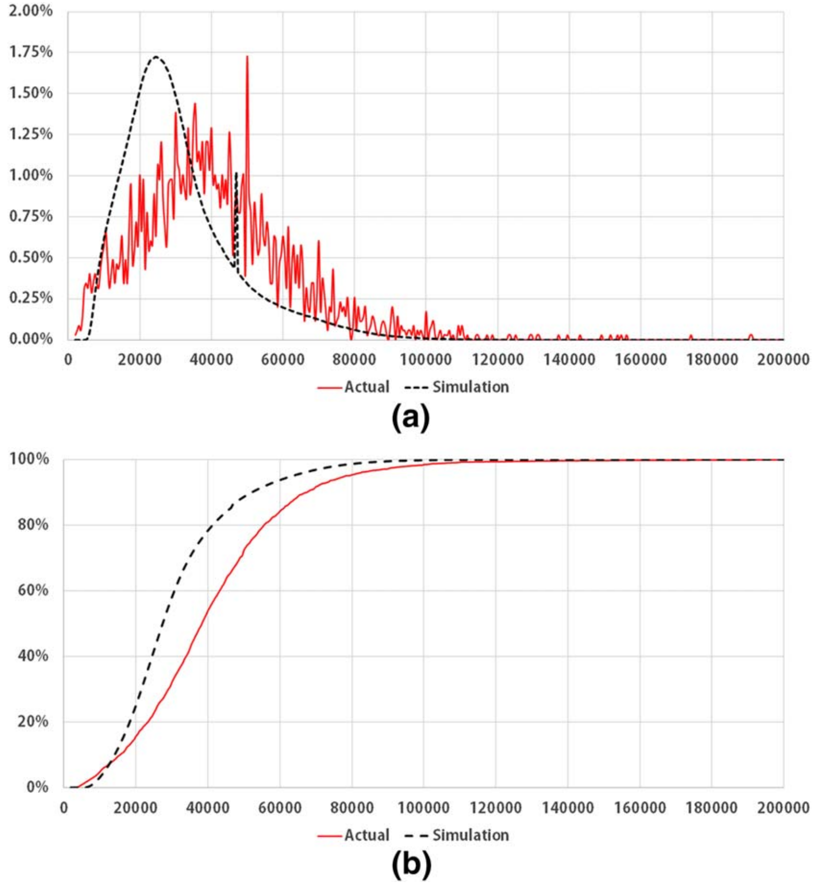

FIGURE 9 The histogram distribution of Shenzhen's housing prices obtained via UMCNN trained by CNN (HSR & SD): (a) frequency histogram; and (b) accumulated frequency histogram

TABLE 4 Training and prediction accuracies of mapping housing prices via UMCNN trained by CNN (HSR & SD) using different sampling window sizes

Window size		25	50	75	100	125	150
Training accuracy	Training	73.24%	90.63%	90.60%	91.97%	90.91%	90.91%
	Validating	74.61%	80.29%	80.70%	83.31%	79.01%	82.30%
	Time (s)	3,125	3,777	5,067	7,188	10,336	12,898
Predicting accuracy	Pearson R	0.764	0.922	0.920	0.919	0.907	0.911
	Standard R^2	0.530	0.745	0.743	0.742	0.738	0.740
	RMSE	26.524	18.456	19.887	19.606	20.066	19.141
	%RMSE	21.75%	15.13%	16.31%	16.08%	16.45%	15.69%
	MAE	13.534	9.757	10.578	9.839	9.959	9.817
	%MAE	36.15%	26.20%	28.93%	26.32%	27.31%	26.18%
	t Test	0.269	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

with other areas, which was well displayed in our simulated result. At the junction area near Shenzhen customs, such as Nantou in Nanshan district (Figure 8b), the luxury residential area and urban village located on opposite sides of the road might have similar location properties, but the spatial heterogeneity of these two types of land use can also be observed in our simulated results. The spatial heterogeneity of the housing price distribution cannot be observed well at a very fine scale if only auxiliary geospatial datasets are used, as in previous studies (Chen et al., 2016; Kuntz & Helbich, 2014; Wu et al., 2016). The area shown in Figure 8c includes the Longgang district, with lower overall housing prices and a high housing price area inside this area consisting of luxury residential areas newly built by Zhonghai Investment Co., one of the biggest real estate development companies in China. To sum up, the proposed method not only simulates the housing price distribution at a fine scale, but also adopts a multi-source spatial data-based united mining method to display the spatial heterogeneity well at a fine scale, which has not been fully addressed in previous studies.

5.3 | Parameter sensitivity analysis

The housing prices of a certain area are correlated not just with the architecture itself, but also with the overall environment, its neighborhood conditions, etc. (Basu & Thibodeau, 1998; Wu, Deng et al., 2014). Therefore, the variation of the sampling window is of vital importance in the training and prediction of our proposed UMCNN model. To cover the scale differences of ground components, we propose the multi-scale land patch sampling strategy in this study. In this section, we first clip samples with a certain window size and overlap the samples with a step of 25 pixels. The accuracy results are shown in Table 4. The results demonstrate that as the sampling window size increases, the validation accuracy shows a gradual increase and finally stabilizes at approximately 80%, which is similar to the result from a previous study (Zhong et al., 2016). When using our trained UMCNN model to simulate the distribution of housing prices, we also observe several similar trends in prediction accuracy. When the sampling window size is larger than 50 $\times$ 50 pixels, the simulation results tended to be stable, and the Pearson R value is approximately 0.920. Moreover, the UMCNN network model built in this study has not been accelerated via GPU computing; we will improve this in the future by refactoring the codes using the CUDA framework. In general, in this study, in consideration of computing accuracy and speed, we adopt a sampling window size of 50 pixels.

6 | DISCUSSION

Housing prices have been among the most pressing issues in China and have received considerable scholarly attention, especially in China's metropolitan cities (Du & Zhang, 2015; Wu, Deng et al., 2014; Wu, Li, & Huang, 2014; Yi & Huang, 2014). However, previous studies of housing prices focus mainly on the driving factors of housing prices instead of sensing the spatial distributions of urban housing prices at a fine scale due to the lack of available data and models (Chen et al., 2016; Kuntz & Helbich, 2014; Wu et al., 2016). This study proposed an effective deep learning-based framework for mapping Shenzhen's housing prices at a very high spatial resolution of 5 m via a united data mining method. To tackle the multi-scale problem of geospatial data, we adopt a multi-scale sample-based RF algorithm to build a UMCNN prediction model for housing prices, and with the help of this model, we obtain a fine-scale housing price distribution map of Shenzhen.

This study first introduces a deep learning method into fine-scale housing price prediction and obtains a high overall accuracy (Pearson $R = 0.922$, OA = 85.82%). We also use several different information fusion methods to mine the features from HSR images and geospatial data. Both of these methods can blend the natural physical and socioeconomic information well (via multi-source geospatial data). The proposed UMCNN-based united data mining on multi-source spatial datasets can avoid the overfitting problem caused by feature combination with the class imbalance problem (Wasikowski & Chen, 2010) and produce the housing price distribution results with the highest accuracy.

The proposed UMCNN model can not only map the distribution of housing prices at a fine scale, but also highlight the existence of spatial heterogeneity. In the field of image understanding, CNN is able to recognize different objects in the image by constructing complex convolution network and multiple recognition rules (Krizhevsky et al., 2012; Simonyan & Zisserman, 2014; Zhong et al., 2016). That is to say, through the training of multi-source spatial data (including HSR image and social media data) and housing price data, UMCNN is able to exploit the physical and socioeconomic driving factors that affect housing prices in different regions and the rules based on these factors. We can observe from the mapping result that even in the simulation of housing prices in different areas, through consideration of the spatial heterogeneity such as closed ground components and other multi-environment elements, the proposed model is still able to map the distribution of housing prices well. It is worth mentioning that based on the proposed UMCNN framework, further research on spatial data fusion and analysis from extra sources can easily be integrated, and therefore it is important to study the influence of each spatial factor well in the future.

Our proposed UMCNN-based data fusion model is more than just a mapping of urban housing prices; it incorporates many fields of urban research, such as human activity pattern recognition and classification of urban land-use conditions. To our knowledge, the distribution of housing prices can represent the economic conditions of residents to some extent (Chan, 2001); thus, the result can be the basic data for economics research in the future. By fusing multi-source temporal–spatial big data, such as social media check-in data (Hu, Wang, & Li, 2014; Zhou, Wang, & Hu, 2013), the GPS trajectory data of floating cars (Li, Zhang, Wang, & Zeng, 2011), the signal data of cell phones (Deville et al., 2014), and multi-source volunteered geographic information (Hultquist, Sava, Cervone, & Waters, 2017), we can carry out many different analyses on human activity pattern recognition of urban residents with different economic conditions. In addition, determining how to improve the accuracy of classifying urban city land use is a research focus in the remote-sensing field (Li, Wang, Wang, Hu, & Gong, 2014; Pei et al., 2014). The proposed model enables urban land-use classification at a fine scale via multi-source spatial data fusion.

This study chose the built-up area of Shenzhen to conduct our study without distinguishing the residential and commercial areas, which is mainly due to the inner complexity of the urban function structure at both area and building scale (Chen et al., 2017; Wang, Wang, & Wu, 2009; Wu, Li et al., 2016). With the input data collected via the Internet, the proposed UMCNN model obtains good housing price mapping accuracy without being given the urban function structure. In further studies, the proposed model can be extended to study the relationship between urban function structures and housing price spatial distributions, which is a very meaningful topic.

The land price and property management spend are also vital factors that affect housing prices (Wu et al., 2014). The reason for not applying land price data in this study comes from the difficulties in gathering timely official land price data of the study area. The timely housing prices we use for training the proposed model are collected via the Internet, which can reflect the timely land prices to some extent, making up for the deficiency in the actual land price. With sufficient real-time online housing price data and geospatial social media data, the proposed UMCNN model provides a possibility to study the spatial distribution of housing prices at a fine scale. Moreover, the proposed UMCNN model can be applied to predict housing price change of cities in different urban planning scenes. For example, it is possible to simulate the influence of the construction of a new subway station or a new road on nearby housing prices.

Determining how to automatically tune the parameters of deep learning models, including the total layer count and convolution kernel size, is still an open problem in the field of machine learning (Abadi et al., 2016; Chen & Lin, 2014; Krizhevsky et al., 2012; Simonyan & Zisserman, 2014), as well as our future concern. The main purpose of this study is to prove the effectiveness of a deep learning neural network model in fusing multi-source spatial data and mapping housing prices. The design of the UMCNN model mainly relied on human intuition and continuous attempts, which is a simple but effective approach that most studies use (Abadi et al., 2016; Krizhevsky et al., 2012). Moreover, the proposed UMCNN model possesses deep learning and predictive power to carry out house price mapping using multi-source spatial variables, but it does not explain the individual contributions of the independent variables. How to increase the interpretability of deep learning-based models is an issue currently being explored by the computer science community, and also a problem that we must face in the future.

The UMCNN model achieves high house price mapping accuracy when only inputting high-resolution remote-sensing images, so the model is also applicable to sparsely populated areas. However, we also need to go through and provide the UMCNN model with transfer learning capability to simulate and map the housing prices of any region with high accuracy and limited input data. How to extract the semantic features of the house internal structure data (such as number of rooms, type of structure, garage or not, attached or not, age of structure) provided on the network, inputting them into the proposed UMCNN and applying the model to a larger area, needs to be carefully studied. Finally, there is a strong correlation between housing price distribution and urban resident activities, so how to integrate human activities into the UMCNN model is a worthwhile question to discuss and study in the future.

7 | CONCLUSIONS

Determining how to curb the soaring housing prices has become an issue of great importance and urgency in big cities (Clayton, 1996; Du & Zhang, 2015; Wu, Deng et al., 2014). Therefore, it is fundamentally important to first map the housing price distributions at a fine scale. This study first proposes an effective deep learning-based framework named UMCNN for mapping fine-scale urban housing prices by fusing various geospatial datasets, including HSR images, social sensing data, and basic geographical data. Through comparing several simulated housing price results generated by multiple data mining methods, the proposed UMCNN-based multi-source data fusion method, which can effectively blend physical and socioeconomic information, obtained the highest housing price prediction accuracy (Pearson $R = 0.922$, OA = 85.82%). In addition, this study also adopts a multi-scale sampling method on multi-source spatial data, which solved the multi-scale problem well. The simulated results of Shenzhen's housing prices indicated that the proposed UMCNN can not only construct a housing price prediction model at a fine scale, but also is very robust, as the multi-source spatial data-based united mining method can consider well the spatial heterogeneity of urban housing price distributions.

We will carry out future work in three aspects. First, to increase the interpretability of the UMCNN model in urban spatial variability mapping; second, to consider semantic quantification of the house internal structures on the Internet, and input them into the UMCNN model to achieve housing price mapping and analysis at the building scale; and third, to enhance the learning capability of the model in order to map housing prices in a wider range, such as at national and global scale. Above all, the results of this study are relevant to issues such as governmental management, private business performance, housing price prediction, and urban planning. It also lays a solid foundation for further studies on fine-scale urban economics and residential activities.

ACKNOWLEDGMENTS

We really appreciate the three anonymous reviewers for their useful comments and suggestions. This study was supported by the National Key R&D Program of China (Grant No. 2017YFA0604402), the National Natural Science Foundation of China (Grant No. 41671408), and the Nature Science Foundation of Hubei Province (Grant No. 2017CFA041).

ORCID

Yao Yao http://orcid.org/0000-0002-2830-0377
Jinbao Zhang http://orcid.org/0000-0001-8510-149X
Ye Hong http://orcid.org/0000-0002-8996-3748
Haolin Liang http://orcid.org/0000-0001-6865-5650
Jialv He http://orcid.org/0000-0002-0997-558X

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... Zheng, X. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. Available at https://arxiv.org/abs/1603.04467

Anselin, L., & Le Gallo, J. (2006). Interpolation of air quality measures in hedonic house price models: Spatial aspects. *Spatial Economic Analysis*, *1*(1), 31–52.

Basu, S., & Thibodeau, T. G. (1998). Analysis of spatial autocorrelation in house prices. *Journal of Real Estate Finance and Economics*, *17*(1), 61–85.

Biau, G. E. R. (2012). Analysis of a random forests model. *Journal of Machine Learning Research*, *13*(1), 1063–1095.

Bitter, C., Mulligan, G. F., & Dall Erba, S. (2007). Incorporating spatial variation in housing attribute prices: A comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems*, *9*(1), 7–27.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Chan, S. (2001). Spatial lock-in: Do falling house prices constrain residential mobility? *Journal of Urban Economics*, *49*(3), 567–586.

Chang, X., Lu, G., Yue, Y., & Li, Q. (2014). Estimating the distribution of economy activity: A case study in Jiangsu province (China) using large scale social network data. In *Proceedings of the 14th IEEE International Conference on Data Mining* (pp. 1126–1134). Shenzhen, China: IEEE.

Chen, J., Guo, F., & Wu, Y. (2011). One decade of urban housing reform in China: Urban housing price dynamics and the role of migration and urbanization, 1995–2005. *Habitat International*, *35*(1), 1–8.

Chen, X., & Lin, X. (2014). Big data deep learning: Challenges and perspectives. *IEEE Access*, *2*, 514–525.

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.

Chen, Y., Liu, X., Li, X., Liu, X., Yao, Y., Hu, G., ... Pei, F. (2017). Delineating urban functional areas with building-level social media data: A dynamic time warping (DTW) distance based k-medoids method. *Landscape & Urban Planning*, *160*, 48–60.

Ciregan, D., Meier, U., & Schmidhuber, J. U. R. (2012). Multi-column deep neural networks for image classification. In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3642–3649). Providence, RI: IEEE.

Clayton, J. (1996). Rational expectations, market fundamentals and housing price volatility. *Real Estate Economics*, *24*(4), 441–470.

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E., ... Tatem, A. J. (2014). Dynamic population mapping using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(45), 15888–15893.

Du, Z., & Zhang, L. (2015). Home-purchase restriction, property tax and housing price in China: A counterfactual analysis. *Journal of Econometrics*, *188*(2), 558–568.

Dub, E. J., & Legros, D. E. G. (2014). Spatial econometrics and the hedonic pricing model: What about the temporal dimension? *Journal of Property Research*, *31*(4), 333–359.

Duque, J. C., Patino, J. E., Ruiz, L. A., & Pardo-Pascual, J. E. (2015). Measuring intra-urban poverty using land cover and texture metrics derived from remote sensing data. *Landscape & Urban Planning*, *135*, 11–21.

Fakhraei, S., Soltanian-Zadeh, H., & Fotouhi, F. (2014). Bias and stability of single variable classifiers for feature ranking and selection. *Expert Systems with Applications*, *41*(15), 6945–6958.

Feng, C., Li, W., & Zhao, F. (2011). Influence of rail transit on nearby commodity housing prices: A case study of Beijing Subway Line Five. *Acta Geographica Sinica*, *8*, 7.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research*, *15*(1), 3133–3181.

Fotheringham, A. S., Crespo, R., & Yao, J. (2015). Exploring, modelling and predicting spatiotemporal variations in house prices. *Annals of Regional Science*, *54*(2), 417–436.

Geiger, B. C., & Kubin, G. (2012). Relative information loss in the PCA. In *Proceedings of IEEE Information Theory Workshop* (pp. 562–566). Lausanne, Switzerland: IEEE.

Granziera, E., & Kozicki, S. (2015). House price dynamics: Fundamentals and expectations. *Journal of Economic Dynamics & Control*, *60*, 152–165.

Hao, P., Sliuzas, R., & Geertman, S. (2011). The development and redevelopment of urban villages in Shenzhen. *Habitat International*, *35*(2), 214–224.

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. Available at https://arxiv.org/abs/1207.0580

Hogan, B., & Berry, B. (2011). Racial and ethnic biases in rental housing: An audit study of online apartment listings. *City & Community*, *10*(4), 351–372.

Hu, Q. W., Wang, M., & Li, Q. Q. (2014). Urban hotspot and commercial area exploration with check-in data. *Acta Geodaetica et Cartographica Sinica*, *43*(3), 314–321.

Huang, X., & Zhang, L. (2013). An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery. *IEEE Transactions on Geoscience & Remote Sensing*, *51*(1), 257–272.

Hultquist, C., Sava, E., Cervone, G., & Waters, N. (2017). Damage assessment of the urban environment during disasters using volunteered geographic information. In L. Shintler & Z. Chen (Eds.), *Big data for regional science* (pp. 443–457). Boca Raton, FL: CRC Press.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*, 790–794.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Proceedings of the 26th Annual Conference on Neural Information Processing Systems* (pp. 1097–1105). Lake Tahoe, NV.

Kuntz, M., & Helbich, M. (2014). Geostatistical mapping of real estate prices: An empirical comparison of kriging and co-kriging. *International Journal of Geographical Information Science*, *28*(9), 1904–1921.

Li, C., Wang, J., Wang, L., Hu, L., & Gong, P. (2014). Comparison of classification algorithms and training sample sizes in urban land classification with Landsat thematic mapper imagery. *Remote Sensing*, *6*(2), 964–983.

Li, L., & Wu, X. (2014). Housing price and entrepreneurship in China. *Journal of Comparative Economics*, *42*(2), 436–449.

Li, Q., Zhang, T., Wang, H., & Zeng, Z. (2011). Dynamic accessibility mapping using floating car data: A network-constrained density estimation approach. *Journal of Transport Geography*, *19*(3), 379–393.

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., . . . Shi, L. (2015). Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers*, *105*(3), 512–530.

Lu, W., Wang, W. J., Wang, X. K., Yan, S. H., & Lam, J. C. (2004). Potential assessment of a neural network model with PCA/RBF approach for forecasting pollutant trends in Mong Kok urban air, Hong Kong. *Environmental Research*, *96*(1), 79–87.

Osland, L. (2010). An application of spatial econometrics in relation to hedonic house price modeling. *Journal of Real Estate Research*, *32*(3), 289–320.

Palczewska, A. M., Palczewski, J., Marchese Robinson, R., & Neagu, D. (2014). Interpreting random forest classification models using a feature contribution method. In T. Bouabana-Tebibel & S. H. Rubin (Eds.), *Integration of reusable systems* (pp. 193–218). Berlin, Germany: Springer.

Pei, T., Sobolevsky, S., Ratti, C., Shaw, S.-L., Li, T., & Zhou, C. (2014). A new insight into land use classification based on aggregated mobile phone data. *International Journal of Geographical Information Science*, *28*(9), 1988–2007.

Punzi, M. T. (2013). Housing market and current account imbalances in the international economy. *Review of International Economics*, *21*(4), 601–613.

Rae, A. (2015). Online housing search and the geography of submarkets. *Housing Studies*, *30*(3), 453–472.

Rondinelli, C., & Veronese, G. (2011). Housing rent dynamics in Italy. *Economic Modelling*, *28*(1), 540–548.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. Available at https://arxiv.org/abs/1409.1556

Wand, M. P., & Jones, M. C. (1994). *Kernel smoothing*. Boca Raton, FL: CRC Press.

Wang, Y. P., Wang, Y., & Wu, J. (2009). Urbanization and informal development in China: Urban villages in Shenzhen. *International Journal of Urban and Regional Research*, *33*(4), 957–973.

Wang, Y. P., Wang, Y., & Wu, J. (2010). Housing migrant workers in rapidly urbanizing regions: A study of the Chinese model in Shenzhen. *Housing Studies*, *25*(1), 83–100.

Wasikowski, M., & Chen, X. (2010). Combating the small sample class imbalance problem using feature selection. *IEEE Transactions on Knowledge & Data Engineering*, *22*(10), 1388–1400.

Wen, H., & Goodman, A. C. (2013). Relationship between urban land price and housing price: Evidence from 21 provincial capitals in China. *Habitat International*, *40*, 9–17.

Wu, B., Li, R., & Huang, B. (2014). A geographically and temporally weighted autoregressive model with application to housing prices. *International Journal of Geographical Information Science*, *28*(5), 1186–1204.

Wu, C., Ye, X., Ren, F., Wan, Y., Ning, P., & Du, Q. (2016). Spatial and social media data analytics of housing prices in Shenzhen, China. *PLoS One*, *11*(10), e164553.

Wu, J., Deng, Y., & Liu, H. (2014). House price index construction in the nascent housing market: The case of China. *Journal of Real Estate Finance and Economics*, *48*(3), 522–545.

Yao, Y., Li, X., Liu, X., Liu, P., Liang, Z., Zhang, J., & Mai, K. (2016). Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *International Journal of Geographical Information Science*, *31*(4), 825–848.

Yi, C., & Huang, Y. (2014). Housing consumption and housing inequality in Chinese cities during the first decade of the twenty-first century. *Housing Studies*, *29*(2), 291–311.

Yiu, M. S., Yu, J., & Jin, L. (2013). Detecting bubbles in Hong Kong residential property market. *Journal of Asian Economics*, *28*, 115–124.

Yu, D., Wei, Y. D., & Wu, C. (2007). Modeling spatial dimensions of housing prices in Milwaukee, WI. *Environment & Planning B*, *34*(6), 1085–1102.

Yuan, J., Zheng, Y., & Xie, X. (2012). Discovering regions of different functions in a city using human mobility and POIs. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 186–194). Beijing, China: ACM.

Zheng, Y., Capra, L., Wolfson, O., & Yang, H. (2014). Urban computing: Concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems & Technology*, *5*(3), 38.

Zhong, Y., Fei, F., & Zhang, L. (2016). Large patch convolutional neural networks for the scene classification of high spatial resolution imagery. *Journal of Applied Remote Sensing*, *10*(2), 25006.

Zhong, Y., Zhu, Q., & Zhang, L. (2015). Scene classification based on the multi-feature fusion probabilistic topic model for high spatial resolution remote sensing imagery. *IEEE Transactions on Geoscience & Remote Sensing*, *53*(11), 6207–6222.

Zhou, M., Wang, M., & Hu, Q. (2013). A POI data update approach based on Weibo check-in data. In *Proceedings of the 21st International Conference on Geoinformatics*. Kaifeng, China: IEEE.

How to cite this article: Yao Y, Zhang J, Hong Y, Liang H, He J. Mapping fine-scale urban housing prices by fusing remotely sensed imagery and social media data. *Transactions in GIS*. 2018;00:1–21. https://doi.org/10.1111/tgis.12330