

# Real estate valuation with multi-source image fusion and enhanced machine learning pipeline

Lin Deng

Published: May 19, 2025 • <https://doi.org/10.1371/journal.pone.0321951>

## Abstract

The automated valuation model (AVM) has been widely used by real estate stakeholders to provide accurate property value estimations automatically. Traditional valuation models are subjective and inaccurate, and previous studies have shown that machine learning (ML) approaches perform better in real estate valuation. These valuation models are based on structured tabular data, and few consider integrating multi-source unstructured data such as images. Most previous studies use fixed feature space for model training without considering the model performance variation brought by various feature configuration parameters. To fill these gaps, this study uses Hong Kong as a case study and proposes an enhanced ML-based real estate valuation framework with feature configuration and multi-source image data fusion, including exterior housing photos, street view and remote sensing images. Eight ML regressors, namely, Random Forest, Extra Tree, XGBoost, Light Gradient Boosting Machine (LightGBM), K-Nearest Neighbors (KNN), Support Vector Regression (SVR), Multilayer Perceptron (MLP), and Multiple Linear Regression (MLR) are used to formulate ML pipelines for training. The SHapley Additive exPlanations (SHAP) method is used to examine the effects of images on housing prices. The experimental results show that the model performances using different feature configuration parameters are significantly different, indicating the necessity of feature configuration to obtain more accurate and reliable predictions. Extra Tree performs significantly better than other models. Half of the top 10 significant features are image features, and incorporating multi-source image features can improve property valuation accuracy. Nonlinear associations exist between image features and housing prices, and the spatial distribution patterns of image feature values and corresponding SHAP main effects vary significantly from the city centre to the suburbs. These findings contribute to a better understanding of AVM development with image fusion and the nonlinear associations between image features and housing prices for public authorities, urban planners, and real estate developers.

<b>Citation:</b> Deng L (2025) Real estate valuation with multi-source image fusion and enhanced machine learning pipeline. PLoS One 20(5): e0321951. <a href="https://doi.org/10.1371/journal.pone.0321951">https://doi.org/10.1371/journal.pone.0321951</a>
<b>Editor:</b> Zhenggeng Ye, Zhengzhou University, CHINA
<b>Received:</b> October 29, 2024; <b>Accepted:</b> March 13, 2025; <b>Published:</b> May 19, 2025
<b>Copyright:</b> © 2025 Lin Deng. This is an open access article distributed under the terms of the <a href="#">Creative Commons Attribution License</a> , which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
<b>Data Availability:</b> All files are available at <a href="https://github.com/Linhkust/Novel-ML-framework-for-residential-property-valuation">https://github.com/Linhkust/Novel-ML-framework-for-residential-property-valuation</a>
<b>Funding:</b> The author(s) received no specific funding for this work.
<b>Competing interests:</b> The authors have declared that no competing interests exist.

## Introduction

The automated valuation model (AVM) is a mathematically based computer program designed to estimate the prices of properties [1]. To develop an AVM, the developers must first identify the influencing factors of housing prices and then build a statistical model to mathematically describe the relationships between the factors and housing prices based on big housing transaction data. Compared with traditional human-based valuations, AVM is more accurate and much cheaper and, therefore, has been commonly used in the real estate industry for property valuation, supplementing or replacing the work of human valuers.

Two groups of factors mainly influence housing prices: macro-level and micro-level [2]. Macro-level factors are mostly time-dependent housing market factors, including inflation, interest, unemployment, and gross domestic product (GDP) [3]. Micro-level factors include property characteristics and locational features [4]. The property characteristics refer to view orientation, number of rooms, floor level and area, building age, etc. Locational features are mainly extracted based on the geographic location of the house, such as geographic coordinates, distance to the city centre or central business districts (CBDs), density, diversity, quality and accessibility of neighbourhood amenities, sociodemographic status, crime rate, and built and natural environment.

Traditional property valuation models include cost, income, and sale comparison approaches [1]. The cost approach is an indirect method, calculating the value of a property as replacement cost less depreciation plus the market-derived land value. The income approach estimates the market value of a property as the net operating income divided by the capitalization rate. The sale comparison approach includes: (1) **comparable sale model** that estimates property values based on small samples of properties with similar characteristics. (2) **direct market model** that describes property values as a function of a property's location and physical attributes. The hedonic pricing model (HPM) is often used to depict the relationship between the housing price and its determinants [5] using elasticity with semi-log or log-log transformations. Traditional models are well-explanatory and interpretable but have limited capability [6]. The HPM excels in explaining the elasticity of real estate economics, but it has low property valuation accuracy and cannot capture the nonlinear associations [7].

With the exponentially increasing housing transaction volume, machine learning (ML) and deep learning (DL) models have been applied to property valuation [8] in a neighbourhood, city, or country. ML and DL models outperform traditional models in extracting information from big data, measuring the nonlinear relationships between property value and its characteristics, and feature

selection [9, 10]. Two primary strands with ML and DL applications emerge: **(1) Method-oriented studies** that adopt novel methods to showcase improved prediction accuracy [6, 11–14], and **(2) Features-oriented studies** that explore the significance of specific features in enhancing price predictions and delve into the economic rationale behind their relevance [15–21]. The most common ML models used in previous studies include tree-based models, k-nearest neighbour (KNN), support vector machine (SVM), multilayer perceptron (MLP), deep artificial neural networks (DANN), etc. A detailed overview using ML and DL models for housing price prediction is provided in the review papers [22–24]. Among these ML and DL models, tree-based models usually perform the best due to ensemble learning strategies.

Among previous property valuation studies, most features are structured data and easy to quantify, except for the built and natural environment that affects homebuyers' willingness to pay for the properties and are complex to convert into numerical values. Given recent advancements in deep learning (DL) based computer vision techniques, researchers can now extract built and natural environment features from multiple types of images [25, 26]. These studies show that image-based features contribute to the accuracy of housing price predictions. Integrating various types of images into AVM development to achieve higher valuation accuracy is deemed imperative. Despite the significant insights of the existing studies, some research gaps remain. First, studies on AVM development with multi-source image fusion and comparing the effects of different image features on housing prices are still limited. Second, most studies use single feature configuration space for model training, and little is known about how one should make decisions regarding generating the best-performing features and how the resultant image-based features may affect the model performance based on the choices made. The effects of feature configuration parameters on model performances should be thoroughly investigated.

In this study, an enhanced ML-based residential property valuation framework is proposed by (1) fusing the multi-source images of exterior estate photos, street view images, and remote sensing images, (2) identifying the feature configuration parameters to formulate a series of ML pipelines, (3) using a server-client based distributed computing strategy to accelerate the ML pipeline training process, (3) evaluating the ML pipelines' performances to analyze the effects of feature configuration and selection, and ML models, (4) enhancing the interpretability of the ML-based approach by analyzing the model-based global feature importance and the SHAP-based local feature importance.

## Literature review

This section critically reviews studies on real estate valuation with multi-source image fusion, focusing specifically on multi-source image data and features, and image feature extraction. [Table 1](#) summarizes the studies regarding research scope, image data, image sampling configuration, image processing, configuration, prediction model, and research results.

Ref.	Research	Image data	Image sampling configuration	Image processing configuration	Prediction model	Research results
[17]	England and Wales: 43,803 properties since 2010	70,000 GBV images	Sampling Interval: 100 m	SP: BPRN	MUR	SP variables impact house prices more than some observable price environment attributes.
[18]	Taipei, Japan: 17,362 properties for sale and 137,881 properties for rent from 2006 to 2015	Two satellite images on April 13, 2006, and October 13, 2015	Spatial resolution: 1.5 m	RS: NDVI	MUR	A 1% increase in average satellite imagery value increases the price of a property by 2 to 2.5%.
[19]	Shanghai, China: 53,445 records in 2019	25,726 BBV images	Sampling Interval: 100 m Sampling range: 1000 m Horizon view: parallel to the street Vertical view: 0° FOV: 120° Image size: 640×360	SP: RF, DT, VS, V: Pretrained Parallel Deep Fusion Network (PDFNet) based on ADE20K dataset	MUR, SAC	Permutation variables and view index can increase the R squared by 0.001 and 0.003 from the baseline (0.791) with OLS model
[20]	Amsterdam: 12,382 records on July 8, 2020; 11,798 records on November 3, 2020	10,903 GBV images	Sampling Interval: 100 m Sampling range: 100 m Horizon view: 0°, 90°, 180°, 270° Image size: 640×360	V: Pretrained DeepNet on the Cityscapes dataset	MUR, GAM	View index variables increase the R squared by 0.001–0.003 (MLP) and 0.005–0.01 (GMR)
[21]	Hanjiang city, Jiangsu: 832 satellite property images (17 images per property) from 2016 to 2019	9620 GBV images of the satellite property images	Horizon view: every 30 degrees from 0° to 360° Time: August 2021	V: Pretrained Hierarchical Multi-Scale Fusion Network (HMFNet) on the Cityscapes dataset	MUR	View index variables increase the R squared by 0.007 (OLS)
[22]	Shanghai, China: 53,445 records in 2019	25,726 BBV images	Sampling Interval: 50m Sampling range: 1000m Horizon view: parallel to the street Vertical view: 0° FOV: 120° Image size: 640×360	V: Pretrained Parallel Deep Fusion Network (PDFNet) based on ADE20K dataset	MUR, SAC, GAM	SP and V variables increase the R squared by 0.001 and 0.013 (OLS), 0.002 and 0.043 (MLP) and 0.168 (GMR)
[23]	Netherlands: 1,584 records from April to June in 2020	Over 40,000 photos of the rental offers	1	V: Pretrained CNN for classification	MUR, RF	Image and text variables increase the R2 by 0.02 (MLR) and 0.03 (DT)
[24]	London: 197,132 records from 2013 to 2015	1,154,848 geo-tagged Flickr images	Sampling range: 800 m	V: Pre-trained PlaceNet CNN based on Flickr dataset	MUR, RF, GB	Image variables increase the R2 by 0.001 (MLR), 0.006 (RF) and 0.019 (GBM)
[25]	Zhengzhou, China: 499 records on 20 December, 2020	7,984 BBV images	Sampling Interval: 50 m and 200 m Sampling range: 800/1600/3200/6400 m Horizon view: 0°, 90°, 180°, 270° Time: 27 July, 2020	V: Fully Convolutional Neural Network (FCN-8s)	1	The index of green space in each urban area is higher than that in the new development zone of eastern Zhengzhou.
[26]	Shenzhen, China: 2,246 residential communities	1,100,832 800 images at 287,713 locations, 83,775 photos from online real estate websites	Sampling Interval: 50 m Sampling range: 1000 m Horizon view: 0°, 90°, 180°, 270° Vertical view: 0°	V: BagNet	MUR, RF, GAM	Visual context with green space has positive impacts on housing prices
[27]	Beijing, China: 25,786 records in May 2021; 25,098 records in May 2021	64,968 BBV panorama images 38,967 BBV images in Chengyang	Sampling Interval: 50m Sampling Time: September 2019	V: DeepLabV3+	MUR, MGRM	The green view index is the most critical visual cue affecting house prices and is positively correlated with house prices
[28]	Shanghai, China: 40,138 records in 2019	25,726 BBV images	Sampling Interval: 50 m Sampling range: 1000 m Horizon view: parallel to street Vertical view: 0° FOV: 120° Image size: 640×360	SP, SVM, RF, DT, V: PSPNet	MUR	SP variables and V increase the R squared by 0.008 and 0.004 from the baseline model
[29]	Shenzhen, China: 12,137 records in all January 2021	80V images, London: 197,132 records in all January 2021	Sampling Interval: 50 m Sampling range: 1000 m Horizon view: 0°, 90°, 180°, 270° Image size: 640×360	VS: BagNet, VGG-16, ResNet101 in buffer	MUR, XGBoost	The green view index has significant and positive effects on housing prices
[30]	Taipei City and New Taipei City: 80,887 records from 2017 to 2018	30,887 Google satellite images	Sampling Interval: 50 m Sampling range: 1000 m Horizon view: 0°, 90°, 180°, 270° Image size: 640×360	FW: Spatial Transformer Network (STN)	MUR, XGBoost, LightGBM	Image features extracted from satellite maps with STN can improve model performance
[31]	New York: 1,464 records from 2010 to 2017	90,382 GBV images	Sampling range: 100 m Horizon view: north, south, east, and west Time: between April and September	V: Python 3.7, RandomCV library	MUR	A positive and statistically significant impact exists on commercial building transaction prices and rents
[32]	Guji City, South Korea: 3,307 property photos	Image size: 1024×1024	Image size: 1024×1024	FW: CNN	MLP	It is useful to use both photographs and metadata to enhance the accuracy of house price estimation
[33]	Greater Boston Area: 1,328 records	126,000 house images at about 47,000 locations	Sampling Interval: 100 m Sampling range: 50 m Horizon view: 10° angles from 0° to 360°	FW: ResNet101	MUR, GBM	House photos and street view images increase R2 by 0.01 and 0.04, respectively
[34]	16 large and small cities in China: 73,717 records in November 2019	218,757 GeoCode house images	Sampling Interval: 100 m Sampling range: 1000 m Horizon view: 0°, 90°, 180°, 270° Image size: 640×360	FW: VGG16, InceptionV3 for classification	MUR, SVM, MLP	Satellite images increase the R2 by 0.040 (MLR), 0.044 (SVM), and 0.0395 (MLP)
[35]	Los Angeles: 158,571 records; Boston: 16,862 records	110,996 GBV images Images are labeled from 50 countries	Sampling Interval: 50 m Horizon view: four directions	SP: pretrained CNN	MUR, SAC, GAM	SP variables increase the R2 by 0.02 (MLR) in Los Angeles and 0.01 in Boston (MLR)
[36]	Philadelphia: 15,810 records	13,815 GBV images	Horizon view: only contains the front view of the house	V: pretrained PIPNet, pre-trained VGGNet of Places365 and deep association	BRIT	Fusing metadata and extracted levels of property values derived by image features has the highest R2
[37]	London: 6,110 records in 2011	112,850 front-facing street images from GBV API and Google Maps	Sampling Interval: median of the street edge between two junctions Sampling range: 100 m Horizon view: front view Image size: 256×256	V: pretrained VGG-16, ResNet101, InceptionV3, Attention on the-glob-top	MUR, spatial regression model	The quality of urban street heritage can be an important housing price determinant
[38]	New York and Massachusetts: 20,000 records in 2016	Interior images, outdoor images (property photos and Google maps), Google satellite images	1	FW: pretrained ResNet101, InceptionV3, architecture on imageNet and ImageNet	MUR, XGBoost, LightGBM, CATBoost	Image features contribute to the overall predictive power
[39]	Shanghai: 2047 properties in 2019	64,520 data	Sampling Interval: 50 m Image size: 1024×1024 Time: August and September 2017 Sampling range: 100 m Horizon view: four directions Vertical view: 0° FOV: 120° Image size: 640×360	V: pretrained ResNet101 on the Cityscapes dataset, RS: 100m for urban green (GG) and water (WW) coverage rate calculation within a 400 m radius buffer	MUR, XGBoost, RF, GBR	View index increase R2 from 0.721 to 0.897, and VGG-UN 0.8722 to 0.887
[40]	Mid-rise and low-rise areas in Shanghai: China: 1365 private leasehold apartments in late 2015	Approximately 230,000 BBV images from China: 1365 private leasehold apartments in late 2015	Sampling Interval: 40 m Sampling range: 100 m Horizon view: 0°, 90°, 180°, 270° Vertical view: 0° Image size: 640×360	V: BagNet	MUR	Human-scale daily movement patterns of great significance for housing price prediction
[41]	London, UK: 136,887 records	111,721 images for both the street image dataset and the aerial image dataset	Sampling Interval: median of the street edge Sampling range: 120° Image size: 256 × 256	FW: CNN	MUR, GAM, XGB	Image features increase the R2 by 0.002 (MLR), 0.008 (GMR), and 0.008 (XGB)
[42]	Ohio, US: 86,382 records in 2020	58,352 GBV images	Sampling Interval: median of the street edge Sampling range: 100 m Horizon view: back the home and surrounding at 90° interval Image size: 800 × 200	V: SVM	MUR	The saliency of street view images can better capture variations in home value price
[43]	Wuhan, China: 2349 records	11 Luojia-1-01 satellite images	Spatial resolution: 130 m resolution	RS: digital values	MUR, GAM	A strong linear positive correlation exists between Wuhan's housing prices and night time lights
[44]	Beijing, China: 3,917 records; China: 2,915 records	71,002 and 91,739 BBV images in Beijing and Shanghai	Sampling Interval: 100 m Sampling range: 800 m Time: between March and October 2015	V: pretrained PIPNet based on KITTI-ADE20K dataset	MUR	The green view index and sky view index can significantly affect housing prices
[45]	Zibo: 9,000 records	More than 200k images of house interiors and exteriors	1	V: DenseNet	SVM with 10-fold cross-validation	Visual information of the real estate photos reduces the median error rate by 2.4% compared to the benchmark with only metadata
[46]	Beijing, China: 22,331 records in 2014	226,746 Tencent SVT images	Sampling Interval: 100 m Sampling range: 400 m Horizon view: six directions from 0° to 360° Image size: 600 × 600	V: pretrained SegNet on the CamVid dataset (photo dataset)	MUR	An improvement in GVI can significantly increase the nearby housing prices
[47]	Wuhan, China: 400 records from February 23 to October 31, 2020	Two ETM+ images on September 17 and October 15, 2020	Spatial resolution: 30m	RS: mathematical equations	SLM	A 1% increase in the thermal environmental index and vegetation coverage index decreases the house price index by 0.45 and 0.79, respectively
[48]	San Jose, US: 2048 records; Rochester, US: 1000 records	2048 property photos in San Jose; 1000 property photos in Rochester	1	FW: GoughNet	Bidirectional LSTM	Location is relatively more important than the visual features in the realtor business

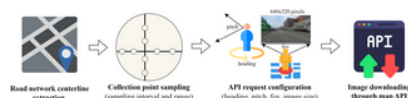
Note: SP=Single perspective; V=View index; VTI=View type; RS=Remote sensing index; FW=Feature vectors; DT=Decision tree; BPRN=Bayesian Personalized Ranking Neural Network; VSI=Visual selection; ML=Multiple linear regression; SAC=Spatial autocorrelation; GMR=Geographically weighted regression; RF=Random Forest; GB=Gradient boosting; GBM=Gradient boosting machine; XGBoost=Machine geographically weighted regression; MLP=Multiple layer perceptron; GBM=Gradient boosting machine; SVM=Support vector machine; GAM=Generalized additive model; SVM-Support vector regression; SLM=Support by model; LSTM=Long short-term memory

**Table 1. Summary of studies on image-based property valuation.**  
<https://doi.org/10.1371/journal.pone.0321951.t001>

#### Multi-source image data and features

An image is worth a thousand words. Three types of images are most often used for hedonic analysis: interior and exterior housing photos, street view images, and remote sensing images. User-generated images are also explored for housing price estimation [34]. There are two types of street view images in a sampling point: (1) single-view images that have specific headings and (2) panoramic images that provide a 360-degree view of the surrounding environment [58]. Fig 1 shows street view image collection procedures, including road network centerline extraction, collection point sampling (sampling interval and range), API request configuration (heading, pitching, field of view, image size), and image downloading through map API. Remote sensing images

include satellite images [28,43,56] and nighttime light (NTL) images [52]. They can reflect the landscape metrics and human activities near the property. Interior and exterior housing photos show the houses' inner furnishing and outer appearance, which is significant in formulating the market value [54].



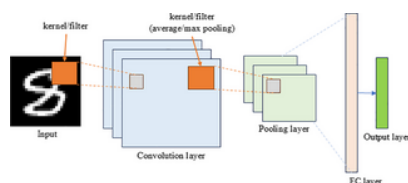
**Fig 1. Procedures of SVI collection.**

<https://doi.org/10.1371/journal.pone.0321951.g001>

Image features can be subjective and objective. **Subjective features** refer to human perceptions of the surrounding environment [59], including greenness, walkability, safety, imageability, enclosure, and complexity [38], lively score [44], and safety score [60]. A crowdsourced dataset called Place Pulse is commonly used as the training dataset to derive the perception score [61]. Since the Place Pulse dataset does not include cities in mainland China, some researchers built the online survey and invited participants to select a preferred photo from two random street view images in response to questions such as "Which place looks greener?" [32]. The pairwise preference is then transformed into a perception score using the Microsoft TrueSkill rating algorithm [62]. The labelled images will be split into training and test sets to derive the pre-trained model to predict the perceptions of the remaining images [29]. **Objective features** include view index, view type, remote sensing index, and feature vector. View index is represented as the proportion of specific elements using semantic segmentation techniques, such as green view index, building view index, and sky view index [31,36]. View type refers to the classification results of images. View types of street view images include housing price levels [10,45], kind of view [46,33], and scene categories [34,51]. View types of remote sensing images can be high or low property price zones [43]. Housing photo-based view types include indoor attribute categories [47] and luxury level [54]. Remote sensing index includes normalized difference vegetation index (NDVI) [28,63], urban green and water coverage rate [48], thermal environmental index and vegetation coverage index [56], and NTL intensity [52,64]. Feature vectors are numerical vectors extracted by convolution neural networks (CNNs) and are usually combined with other numerical features to formulate the final feature space [50,42]. Despite the substantial contributions of these studies, they do not thoroughly examine a house's built and natural environment from the perspective of exterior housing appearance, street and aerial view.

#### Image feature extraction

Image feature extraction aims to transform raw image data into numerical features while keeping the essential information. These features are necessary for various downstream tasks, such as image classification and semantic segmentation. As one of the deep learning neural network architectures used in computer vision, CNN is often used for multi-source image extraction because it has demonstrated effective and efficient performance in image processing tasks [65]. As is shown in Fig 2, CNN has three layers [66]: (1) a convolutional layer that learns feature representations of the inputs; (2) a pooling layer that seeks to reduce the image size while preserving important characteristics using average pooling or max pooling; and (3) a fully connected (FC) layer that performs high-level reasoning and is connected to the output layer (e.g., softmax layer for classification tasks). Kernels or filters are small matrices that perform convolution operations on the input data.



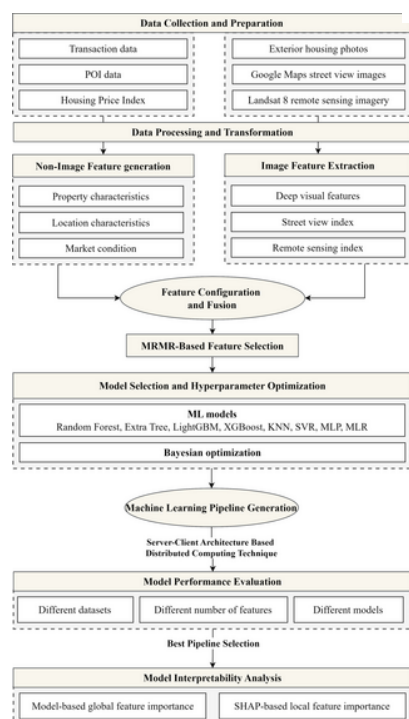
**Fig 2. Typical CNN architecture.**

<https://doi.org/10.1371/journal.pone.0321951.g002>

For street view images, the CNN-based semantic segmentation models are often used to analyze and classify the image pixels into different object categories. Common semantic segmentation models include PSPNet [67], Place365 [68], FCN-8s [69], SegNet [70], DeepLabv3 [71], and DeepLabv3+ [72]. The elements extracted from images and used for housing price prediction include buildings, vegetation, sky, sidewalks, pedestrians, etc. For housing photos and remote sensing images, CNN is applied for image classification by using the images as inputs and predetermined labels as outputs, such as housing price levels (cheap/expensive) and view types (urban/suburban areas). Researchers first train a CNN model with minor manually labelled images and then apply the pre-trained model to classify remaining images with transfer learning and fine-tuning. Some studies use the feature vectors of CNN's specific layers as input features, which are combined with other numerical feature vectors to formulate the final feature space [57,73].

#### Overall research methodology

Fig 3 shows our proposed ML framework. The process starts with collecting non-image data and image data. After data processing and transformation, feature generation and extraction are applied to identify the non-image and image features, followed by feature configuration and fusion. The minimum redundancy maximum relevance (MRMR) is used for feature selection. Bayesian optimization is applied to optimize the hyperparameters of eight ML models. A series of ML pipelines are generated by considering different feature configuration parameters, the number of selected features by MRMR, and ML models. The pipelines are then executed using a server-client architecture-based distributed computing technique. The model performances are evaluated using six criteria and statistical tests. The best-performing ML pipeline is selected and used for model interpretability analysis regarding model-based global feature importance and SHAP-based local feature importance.



**Fig 3. Overall research methodology of this study.**

<https://doi.org/10.1371/journal.pone.0321951.g003>

## Study area and data

### Study area

There are three regions (Hong Kong Island, Kowloon, and New Territories) with 18 districts altogether in Hong Kong. Private apartments sold in the secondary residential market of Hong Kong are selected due to (1) large valuation demands. Private apartments are the most common property type in the open market of Hong Kong. More than 60% of the yearly transaction volume is from the secondary residential market between 2020 and 2022 [74]. (2) sufficient transaction data. There are many mature real estate agencies in Hong Kong, such as Centaline Property, Midland Realty, 28Hse, etc. They provide detailed transaction data lists for individual apartments on their websites. (3) abundant geoinformation database. Hong Kong government has developed an online platform called Common Spatial Data Infrastructure (CSDI) to provide different spatial data types.

### Data collection and preparation

The multi-source datasets in this study consist of (1) non-image data: transaction data, POI data, and housing price index; and (2) image data: exterior housing photos, Google Maps Street view images, and Landsat 8 remote sensing imagery. Before collecting the data, we double-checked the terms and conditions for the data source to ensure the data collection and analysis methods fully meet the protocols.

Transaction data, POI data, and housing price index.

A web crawler via Python was developed to collect the apartment transaction records from 28hse.com. The data attributes include transaction date, the district where the apartment is located, estate name, unit address, floor level, gross floor area (GFA) of the apartment, and the transaction price. A total of 26,377 transaction records from July 2021 to December 2021 were collected. POI data in Hong Kong were collected from the CSDI platform. Centa-city leading (CCL) index reflects the housing price fluctuations in Hong Kong (Fig 4). It is a weekly index generated by Centaline Property using the transaction prices of properties with large transaction volumes in Hong Kong. CCL index is expressed as:



**Fig 4. CCL index in Hong Kong.**

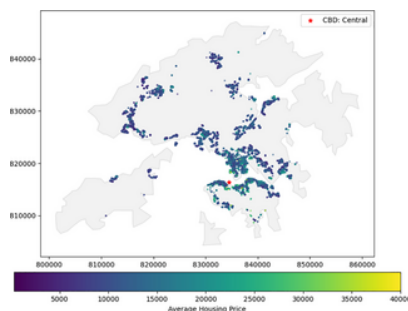
<https://doi.org/10.1371/journal.pone.0321951.g004>

(1)

where  $\bar{P}_w$  is the CCL in week  $w$ ;

$\bar{P}_w$  is the total market value of representative properties in week  $w$

We first identified the columns related to the transaction date, property address, property characteristics, and transaction price. Other columns were discarded. Data cleaning was applied to address the inconsistent and missing values and outliers. Then, the CCL index was inserted into the transaction records by matching the transaction date with the index's release date. Lastly, the property address was formulated as a string to obtain geographic coordinates using the CSDI platform's geocoding API. A total of 22,888 transaction records with the CCL index and geographic coordinates were finalized. The pseudo-code is provided in Algorithm 1. The spatial distribution of average housing prices is provided in Fig 5. The red marker refers to Central, the central business district (CBD) of Hong Kong. The housing prices near the CBD area are relatively higher than in other regions.



**Fig 5. Spatial distribution of average housing prices in Hong Kong.**

<https://doi.org/10.1371/journal.pone.0321951.g005>

Algorithm 1 Data preparation procedures for raw transaction records

Input: Raw data transaction records ; CCL index data

Output: Clean data transaction records with CCL index and geographic coordinates

*/\*Data cleaning\*/*

1: Identify the columns related to the transaction date, property addresses, property characteristics, and transaction price. Other columns are discarded.

2: Delete data rows with missing values */\*3207 samples are removed\*/*

3: Remove the string of units in the GFA column (i.e., ft<sup>2</sup>)

*/\*Outliers are defined as data samples that are not within the interval:  $[Q1 - 1.5 * (Q3 - Q1), Q3 + 1.5 * (Q3 - Q1)]$ , Q1 is 25th percentile and Q3 is 75th percentile]*

4: Use boxplot to detect outliers in the GFA column and delete rows with outliers. */\* 122 samples are removed\*/*

5: Unify the units in the transaction price column */\* (i.e., Million HK\$ to HK\$)\*/*

6: Use boxplot to detect outliers in the transaction price column and delete rows with outliers */\*160 samples are removed\*/*

*/\*Assign CCL index based on the transaction date\*/*

7: Identify the time string format in the transaction date column. */\*i.e., year-month-day\*/*

8: Extract the year, month, and day. Assign each record the latest CCL index before the transaction date from

*/\*Get geographic coordinates based on property address\*/*

9: Create a new column of property address string that includes district name, estate name, and building block name (if any).

10: Use the location search API of the CSDI platform to derive the HK1980 Grid (northing, easting) and the WGS84 (latitude, longitude) coordinates based on the address string.

Exterior housing photos.

A web crawler was designed to collect the exterior images of all estates from 28hse.com, and then we matched them with transaction records using the unique estate ID. A total of 2,372 estate photos were finalized.

Google Maps street view images.

We first downloaded the road centerline network from the CSDI Platform of Hong Kong. Then, we imported the clean transaction records and the road centerline network into ArcGIS Pro to sample the image collection points at 50 m intervals. A total of 46,147 sampling points within 1 km of the properties were kept and exported to an Excel file with point ID, latitude, and longitude columns. Google Street View API was used to identify the pano\_id parameter of each sampling point to collect panoramic street view images. The image tiles were collected based on the pano\_id and stitched into a full-view panoramic image with a size of 1664x832 pixels. For sampling points without available panoramic images, we turned to collect the single-view images with four different headings: front view (heading=0), right view (heading=90), rear view (heading=180) and left view (heading=270). The FOV parameter was set

to 90 degrees, and the *pitch* parameter was set to 0. The image size was set to 640x360 pixels, and the collection date was set to July 2021. A total of 42,398 panoramic images and 13,324 (3331x4) single-view images were collected, and 418 sampling points were discarded because there were no available street view images in these points.

Landsat 8 remote sensing imagery.

The Earth Explorer of the United States Geological Survey was used to collect Landsat 8-based GeoTIFF files. We first identified the search criteria by drawing the polygon around the target area (Hong Kong), selecting the Landsat 8 data set, and determining the range of data generation date as 2021/07/01 to 2021/12/31. Among the forty-six search results, we chose one GeoTIFF file generated on 2021/12/05, which covers all housing transaction samples and has the best image quality without cloud interference.

## Feature engineering

### Non-image feature generation

Feature generation creates new features from the primary feature space or raw data to increase the robustness and generalizability of the model [75]. We identify numerical feature spaces as property characteristics, location characteristics, and market conditions. Property characteristics include floor level and floor area. Location characteristics include geographic coordinates, Wi-Fi hotspot density, POI density, POI diversity, and POI accessibility. The POI database has 8 POI categories, 18 POI classes, over 100 POI types, and over 38,000 geocoded POI places in Hong Kong (Table 2). POI density is identified as the number of POI places within 1 km of the property. POI diversity includes the number of POI classes and types within 1 km of the property, the entropy-based diversity index of POI class and types within 1 km of the property. The entropy-based diversity index is expressed as:

POI Categories	POI Class	No. of POI types	No. of POI places			Total
			Hong Kong	Kowloon	New Territory	
Community Facilities	Commercial Facility (CMF)	7	627	1331	1127	3085
	Community Facility (COM)	4	80	144	424	648
	Cultural Facility (CUF)	11	271	216	676	1163
	Municipal Facility (MUF)	7	564	694	2958	4216
	Total	29	1538	2385	5385	9512
Education	Educational (ECH)	9	797	1132	1915	3642
	Aquatics (ACQ)	2	18	0	39	57
	Competition (PAC)	7	9	0	48	56
	Leisure Facility (PSF)	21	1382	1889	4827	7958
	Tourism and Hiking (THH)	3	45	10	80	135
	Total	39	1453	1993	4954	8176
Medical	Health Care (PHC)	5	251	369	479	1099
	Recreational (RSC)	4	39	29	102	169
Public Services	Public Service (PSV)	29	886	1311	947	2144
	Administration (ATT)	5	518	545	1915	2574
	Religious (REL)	7	281	545	864	1693
	Total	38	1542	1922	2966	5344
Religion	Religious (REL)	7	281	545	864	1693
	Bus Stop (BST)	2	296	252	797	1345
Transportation	Transportation Furniture (TFP)	5	128	170	627	925
	Transportation (TRN)	16	528	867	1660	3055
	Accommodation (AAC)	4	288	1321	278	1885
	Total	145	6712	3625	1967	10585

<https://doi.org/10.1371/journal.pone.0321951.t002>

Table 2. POI database in Hong Kong.

<https://doi.org/10.1371/journal.pone.0321951.t002>

(2)

(3)



where  $H_i$  and  $H_j$  are the diversity index of POI

classes and types for property  $i$  and

are total number of POI classes and types.

and  $n_{ij}$  are the number of POI places that belong to class

and type  $j$  within 1 km of the property

is the total number of POI places within 1 km

of the property

For POI accessibility, we identified 12 POI types according to the previous research results regarding the influencing factors of housing prices, including mall/shopping centre/commercial complex (MAL) and supermarket (SMK) [76], kindergarten (KDG), primary school (PRS) and secondary school (SES) [77,78], park (PAR), playground (PLG) and minor open space (RGD) [79,80], bus terminus (BUS), green minibus terminus (MIN) and railway station entrance (MTA) [81], and car park (CPO) [82]. POI accessibility is represented with a binary variable indicating whether a POI is within a predetermined circular distance range of a property. Market condition is proxied using the latest CCL index before the transaction date.

#### Image feature extraction

In this study, three types of image features are extracted from exterior housing photos, street view and remote sensing images, including deep visual features after dimension reduction, street view and remote sensing index. Detailed procedures are introduced in the following three subsections.

##### Deep visual features after dimension reduction.

CNN is used to extract deep visual features because it leverages local receptive fields, shows computational effectiveness with sharing parameters, and efficiently and effectively captures the spatial and structural information in images [83]. Specifically, CNN is first used to extract the feature vectors of the last FC layer before the output layer. As the FC layer is usually high-dimensional, directly integrating the high-dimensional features will cause the curse of dimensionality, which can increase computational complexity and degrade model performance. Therefore, a nonlinear t-distributed stochastic neighbor embedding (t-SNE) is then used to reduce the FC layer's dimension into low-dimensional space to avoid the curse of dimensionality, retain the most relevant features and achieve higher computational efficiency. The t-SNE is an unsupervised nonlinear dimensionality reduction technique

that visualizes high-dimensional data by giving each data point a location in a two- or three-dimensional space [84]. The deep visual features after dimension reduction are extracted from exterior housing photos, and each property is attached with corresponding photo-based features.

Semantic segmentation-based street view index.

A review study benchmarks the semantic segmentation models, and DeepLabv3+ performs best on the Cityscapes and PASCAL VOC datasets [85]. The two public datasets are widely used for benchmarking in computer vision tasks such as object detection, semantic segmentation, and classification. As the Cityscape dataset is related to the semantic understanding of urban street scenes, the DeepLabv3+ model pre-trained on the Cityscapes dataset was used to extract the view index from street view images. Three types of view index are used in this study, including the building view index (BVI), sky view index (SVI), and vegetation view index (VVI). These three indices have demonstrated significant impacts on housing prices, and the formulas are expressed as [53]:

(4)

where  $\alpha$  is the percentage of visual element class

$\alpha$  in the SVI.  $N$  is the total number of pixels

associated with the visual element class  $\alpha$ .  $N$  is

the total number of pixels in the street view image. A property's view index is represented by averaging the view index of sampling points within a predetermined radius of the property.

Remote sensing index.

Landsat 8 has eleven bands, namely, different ranges of frequencies along the electromagnetic spectrum. Band 3 (green), band 4 (red), band 5 (near infrared, NIR), and band 6 (shortwave infrared 1, SWIR 1) of remote sensing images were used to calculate three types of remote sensing index, namely, the Normalized Difference Vegetation Index (NDVI), the Normalized Difference Water Index (NDWI), and the Normalized Difference Built-up Index (NDBI). The formulas are expressed as follows:

(5)

(6)

(7)

Similar to view index, a property's remote sensing index is represented by averaging the index within a predetermined radius of the property.

#### Feature configuration and fusion

Fig 6 shows the proposed feature configuration and fusion framework. First, feature configuration is applied to several features by selecting different feature configuration parameters to create various feature combinations. The features to be configured include geographic coordinates, POI accessibility, deep visual features with DCNN and after dimension reduction, average view and remote sensing index. Non-image features extracted from tabular data include property, location and market condition. Image features extracted from exterior housing photos, street view images, and remote sensing images correspond to deep visual features after dimension reduction, semantic segmentation-based view index, and Landsat 8-based remote sensing index. The non-image and image features are then concatenated and fused into the finalized feature space.

**Fig 6. Feature configuration and fusion framework.**

<https://doi.org/10.1371/journal.pone.0321951.g006>

#### Feature configuration.

Table 3 presents the feature configuration parameter descriptions and candidates. The geographic coordinates can be represented by either easting and northing in Hong Kong 1980 grid system or latitude and longitude in The World Geodetic System 1984 (WGS84). Four DCNN types are tested to extract deep visual features from housing photos, including GoogleNet [86], AlexNet [65], VGG16 [87], and ResNet-101 [88]. The number of reduced dimensions using t-SNE is set as either 2-dimensional or 3-dimensional. Two distance radiuses are used to calculate POI accessibility, i.e., 300 and 500 m; Two distance radiuses are used to derive view index and remote sensing index, i.e., 500 and 1000 m. A total of 128 feature combinations are generated after feature configuration.

**Table 3. Feature configuration parameters.**

<https://doi.org/10.1371/journal.pone.0321951.t003>

#### Non-image and image feature fusion.

Tabular data is used to generate 23 non-image features that are related to property, location, and market conditions. Exterior housing photos are processed with DCNN and t-SNE to generate two or three features after 2D or 3D t-SNE. Street view images are segmented with a pre-trained DeepLabv3+ model to generate three features related to the view of building, sky and vegetation. Remote sensing images are processed to produce three features related to the NDVI, NDWI and NDBI. The non-image and image features are concatenated and fused to formulate the finalized feature space with 31 or 32 features. The feature space containing all potential features is summarized in Table 4, where the last two columns show each feature's mean value and standard deviation (SD).

**Table 4. Feature space with all potential features.**

<https://doi.org/10.1371/journal.pone.0321951.t004>

#### MRMR-based feature selection

Feature selection aims to build a feature subset based on the original feature set to reduce effects from data noise or irrelevant variables and still provide good prediction results [89]. This study uses the MRMR method to select the most efficient and distinctive features. MRMR uses mutual information to find the optimal feature set by minimizing the redundancy between selected features and maximizing the relevance between selected features and target variable simultaneously [90] expressed as:

(8)

where

is the relevance between the selected feature set

and target variable

is the redundancy between the selected features in

is the mutual information operator. The

procedures of the MRMR algorithm are shown in Algorithm 2.

Algorithm 2 Minimum Redundancy Maximum Relevance

Input: Dataset , feature set , number of

selected features

Output: A feature subset

1:

2: for do

3:

4: Add to

5: end for

6: return S

#### Model selection and hyperparameter optimization

Model generation aims to choose a learning algorithm automatically and simultaneously set its hyperparameters to optimize model performance [91]. Previous studies on residential property valuation have shown that tree-based ML models perform the best [16,92,93]. Therefore, four tree-based ML models were selected: Random Forest, Extra Tree, XGBoost, and LightGBM.

Random Forest (RF) builds a new dataset with a replacement (bootstrap sampling) from an existing dataset and trains several decision tree models with randomly selected features on the new dataset [94]. The predictions of each decision tree model are aggregated into the final prediction. Extra Trees (ET) splits nodes by choosing cut points fully at random and uses the whole learning sample (rather than a bootstrap replica) to grow the trees [95]. XGBoost is a gradient boosting tree (GBDT) based algorithm that constructs the objective function of model deviation and regularization term to prevent over-fitting [96]. LightGBM is also a GBDT base algorithm that proposes two novel techniques, i.e., gradient-based one-side sampling (GOSS) and exclusive

feature bundling (EFB), to realize faster training efficiency and lower memory usage. Four non-tree-based models are also included for benchmark comparison, including k-nearest neighbours (KNN), support vector regression (SVR), multiple layer perceptron (MLP), and multiple linear regression (MLR).

Bayesian optimization (BO) is an iterative stochastic optimization framework for the CASH problem [97]. It first builds a probabilistic surrogate model (Gaussian process or tree-based model) mapping from the hyperparameters to the objective metrics. Then, it defines an acquisition function to decide which hyperparameter configuration to evaluate next, balancing the exploration and exploitation during the search process [98]. The hyperparameter search spaces of tree-based and benchmark models are presented in Table 5.

**Table 5. Hyperparameter search spaces.**

<https://doi.org/10.1371/journal.pone.0321951.t005>

## ML pipeline generation and evaluation

### Pipeline generation and execution

A total of 128 datasets are created with various combined features and split into training and test sets with a ratio of 80:20. As there are 31 features in total, and the number of selected features using the MRMR method is set as 1, 6, 11, 16, 21, 26, 31. Each dataset with selected features is trained with four ML models. A total of pipelines are generated and executed by server-client based distributed computing technique.

The distributed computing technique runs programs across several computers on a network to achieve high-performance scientific computing. The server-client architecture requires that the server divides the tasks of ML pipeline execution to the clients, which run the tasks and send the pipeline performance results back to the server. The experiments have been conducted by selecting one laptop as a server and three desktops as clients. A 10-core Intel i7-12650H processor (2.30 GHz) with 16 GB RAM is used as the server that sends the pipeline training tasks to: (1) desktop client 1: an 8-core AMD Ryzen 7 5700X processor (3.40 GHz) with 32GB RAM; (2) desktop client 2: a 6-core Intel i5-10500 processor (3.10 GHz) with 16GB RAM; and (3) desktop client 3: a 4-core Intel i7-4770 processor (3.40 GHz) with 32GB RAM. The network condition is configured as Wi-Fi with 1000Mbps uplink bandwidth and 1000Mbps downlink bandwidth. The proposed machine learning and distributing computing experiments are implemented in Python3 using *scikit-learn* and *multiprocessing* package.

### Pipeline evaluation and interpretation

#### Evaluation criteria.

The pipelines are measured with multiple evaluation criteria to identify the best performing one: root mean squared error (RMSE) (Eq. 9), percentage of RMSE (Eq. 10), mean absolute error (MAE) (Eq. 11), percentage of MAE (Eq. 12), R squared (Eq. 13), and coefficient of dispersion (COD) (Eq. 14).

(9)

(10)

(11)

(12)

(13)

(14)

where  $\hat{y}_i$  = the actual value,  $y_i$  = the predicted

value,  $\bar{y}$  = the mean of the actual values, COD = coefficient of dispersion,

, and  $med(y)$  = the median of

in a dataset with  $n$  numbers of samples.

#### Statistical evaluation.

The statistical evaluation framework includes three steps: (1) calculate the performances of different pipelines according to the six evaluation criteria to obtain average rankings for each pipeline; (2) use the Wilcoxon signed-rank test to test whether two groups of pipelines perform equally; (3) use Friedman test to test whether multiple groups of pipelines perform equally.

The pipeline performances are first evaluated using the six evaluation criteria.

refers to

the ranking of the

pipeline using the

evaluation criterion. A pipeline's ranking

is calculated by averaging

:

(15)

Where

and

are the total number of pipelines

and evaluation criteria, respectively. The Wilcoxon signed-rank test is a non-parametric alternative of the paired t-test [99], and it aims to perform paired comparisons and test whether one pipeline group (e.g., pipelines using Extra Tree) performs significantly better than the other one (e.g., pipelines using Random Forest). The Friedman test is also a non-parametric statistical test method [100], and it is used to test whether there are statistically significant performance differences among multiple pipeline groups (e.g., pipelines using different datasets).

Model interpretation.

The optimal pipeline will be interpreted by the SHAP method, which explains the prediction of an instance by computing the contribution of each feature to the prediction by computing Shapley values from coalitional game theory [101]. The SHAP value for

feature

of observation

,

, is defined as:



(16)

where  $F_{\text{sub}}$  is a subset of the features with feature

excluded.  $F_{\text{total}}$  is the total number of features.  $f$  is

the model's prediction function.

## Results and discussions

A total of 7168 ML pipelines are generated and trained with the distributed computing technique. After all pipelines are trained, a leaderboard of their performances will be created, including dataset ID, pipeline ID, ML model with optimized hyperparameters, selected features, model performances and corresponding average rankings.

### Model performance analysis

Model performances using different datasets.

Fig 7 illustrates the normalized average model performance of all 128 datasets with six evaluation criteria, and the performances vary among different datasets. The Friedman test is used to test whether there are significant differences between the performances of these datasets, and the p-value is smaller than 0.01, indicating the performances of these datasets are significantly different at the 99% confidence level. Most previous studies used only one dataset to train ML models; however, there is no guarantee that the dataset will lead to the best-performing one. For instance, many distance ranges have been used by previous studies to calculate the street view index of a property, and simply following their parameter settings may not work well for our case. Formulating a series of datasets and searching for the optimal feature configuration parameters is necessary to improve the overall model performance.

**Fig 7. Model performance of different datasets.**

<https://doi.org/10.1371/journal.pone.0321951.g007>

Model performances selecting different number of features.

Fig 8 presents the mean and distributions of model performances with different numbers of features selected. Selecting more features leads to more accurate and stable prediction results. Table 6 shows the model performance results with varying numbers of features. The Wilcoxon test is used to test whether there are significant differences when selecting different numbers of features, and the model performances with more features selected are statistically better than those with fewer features.

**Table 6. Model performance results of different number of selected features.**

<https://doi.org/10.1371/journal.pone.0321951.t006>

**Fig 8. Model performance of different selected features.**

<https://doi.org/10.1371/journal.pone.0321951.g008>

Model performances using different models.

Fig 9 presents the model performance distribution of eight ML models using violin plots, which can visualize the distribution of numeric data in a hybrid of box plots and kernel density plots. The extreme values (maximum and minimum) and mean values are shown in blue dashes. Overall, the performances of Random Forest and Extra Tree are more stable than those of other models. The model performance distributions of XGBoost and LGBM are similar with large variance. KNN and MLR perform the best and worst of all the benchmark models, respectively. Table 7 shows the performance results of different ML models. The eight pipeline groups using eight models (RF, ET, XGBoost, LGBM, KNN, SVR, MLP, and MLR) are selected for paired comparisons to find the best-performing model. Each pipeline group has 896 pipelines. The Wilcoxon signed-rank test checks whether one pipeline group performs significantly better than the other. The test results show that Extra Tree performs significantly better than other models among all six evaluation criteria and average rankings.

**Table 7. Model performance results of different models.**

<https://doi.org/10.1371/journal.pone.0321951.t007>

**Fig 9. Model performance of different ML models.**

<https://doi.org/10.1371/journal.pone.0321951.g009>

Best pipeline selection.

The pipeline with the highest average ranking is selected as the best; its details are presented in Table 8. The best pipeline selects 31 features and is formulated based on the Extra Tree model. The optimized hyperparameters after BO are also listed in the table. The pipeline performs the best among the six evaluation criteria, with the rankings listed in the bracket. Furthermore, the pipeline is used for model interpretability analysis.

**Table 8. Best model pipeline.**

<https://doi.org/10.1371/journal.pone.0321951.t008>

### Model interpretability analysis

Model based global feature importance.

Fig 10 shows the relative feature importance of selected features. The feature importance is calculated based on the Extra Tree

model, and then the relative importance of feature

is calculated with the equation:

**Fig 10. Relative feature importance.**

<https://doi.org/10.1371/journal.pone.0321951.g010>

(17)

where

is the Extra Tree-based feature importance of feature

and

are the maximum and minimum feature importance, respectively. The floor area is the

most significant feature, and the reason could be that floor area is closely related to the living comfort of residences, especially in densely populated cities like Hong Kong. Northing ( $y$ ) and easting ( $x$ ) are ranked 6<sup>th</sup> and 9<sup>th</sup>, respectively. The geographic location of properties matters in determining the prediction accuracy, as suggested by [92]. POI class and type diversity are ranked 7<sup>th</sup> and 11<sup>th</sup>, indicating that diversified distributions of POI places around the property could affect the housing price. Walking accessibility to MAL, MTA, and SES are the top 3 important features among all POI accessibility variables. Half of the top 10 significant features are image features, including building view (53.4%), NDVI (45.3%), NDBI (44.8%), NDWI (41.0%), and sky view (26.9%). The results show that images play significant roles in predicting housing prices.

We use the best pipeline for model performance comparison to further distinguish between the roles of different image features. The baseline experiment excludes all non-image features. Three experiments are conducted by feeding only one image type into the baseline experiment step by step. Finally, we train the pipeline with all image features. Due to the randomness of the Extra Tree model, the number of experiments for each type is set as 100. The average performances and rankings of each experiment type are provided in Table 9. According to the average ranks, the remote sensing (RS) image contributes the most to prediction accuracy among the three image types. The model incorporating all image features has the highest average ranks. It proves that the three types of images can characterize housing prices from different perspectives, and it is necessary to consider incorporating images to improve housing price prediction accuracy.

Experiment	RMSE	%RMSE	MAE	%MAE	R2	COO	Average is
Baseline (B)	2369.43 (3)	17.71 (3)	1324.14 (3)	9.98 (4.5)	0.809 (3)	16.50 (4)	4.8
B+SVI	2353.10 (4)	17.61 (4)	1303.76 (3)	9.89 (5)	0.813 (4)	16.38 (3)	5.5
B+RS	2313.15 (3)	17.44 (3)	1291.01 (1)	9.74 (1)	0.815 (3)	16.20 (1)	2.0
B+VF	2369.91 (3)	17.86 (3)	1303.91 (4)	9.98 (4.5)	0.814 (3)	16.37 (3)	3.3
B+SVI+RS+VF	2273.42 (1)	17.14 (1)	1269.90 (2)	9.60 (2)	0.821 (1)	16.38 (2)	1.5

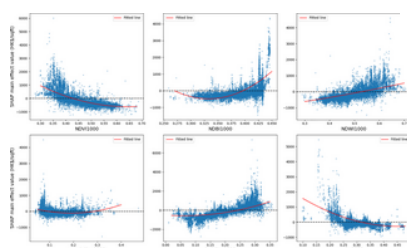
<https://doi.org/10.1371/journal.pone.0321951.t009>

**Table 9. Model performance in different combinations of image sources.**

<https://doi.org/10.1371/journal.pone.0321951.t009>

SHAP based local feature importance.

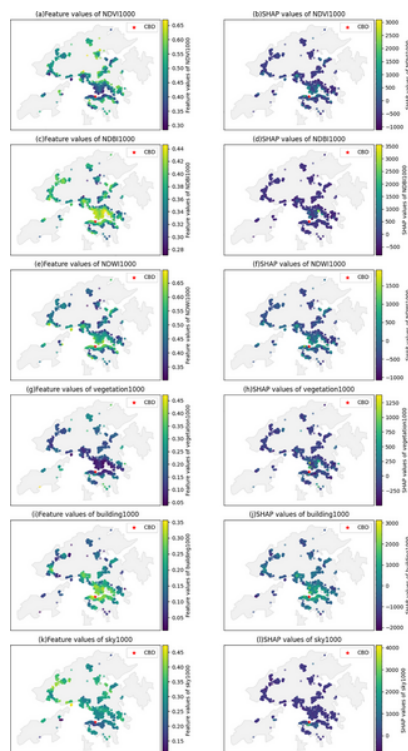
The relationships between image features and housing prices are depicted in Fig 11. Each blue point represents a sample of a housing unit, and the red lines are the fitted curve using the polynomial fitting method. The visual features of house photos do not have actual meanings and, therefore, are excluded from the analysis. The remote sensing and view index exhibit unique nonlinear associations with housing prices. The NDVI has a nonlinear and negative effect on housing prices, falling by 3000 HK\$/sqft within the range of 0.30~0.67. The NDBI shows a nonlinear and positive impact on the housing price, which increases rapidly when the NDBI is larger than 0.425. The housing price increases linearly by about 2000 HK\$/sqft when NDWI increases from 0.30 to 0.70. For VVI, the housing price is unchanged within the range of 0.1~0.3 and then increases when the index exceeds 0.3. BVI has a similar trend with NDBI, and the price increases by about 4000 HK\$/sqft when BVI is within the range of 0.05~0.35. SVI has a negative and nonlinear effect on housing prices. The price decreases when SVI increases from 0.15 to 0.25, followed by a slowly decreasing trend when SVI is larger than 0.25.



**Fig 11. Nonlinear associations between image features and housing prices.**

<https://doi.org/10.1371/journal.pone.0321951.g011>

To further explain the effects of remote sensing and view index on housing prices, the spatial distribution of feature values and corresponding SHAP main effects are illustrated in Fig 12. The location of CBD is identified as Central and marked as a red star. Housing units near CBD are characterized by lower NDVI and VVI, higher NDBI and BVI, higher NDWI, and lower SVI. The urban greenness distribution is unequal in Hong Kong (Figs 12a and 12g): the central city has less greenness than other areas. The effects of NDVI and VVI vary significantly from the central city to the suburbs (Figs 12b and 12h). Higher NDBI and building view index contribute to higher housing prices (Figs 12c, 12d and Figs 12i, 12j), possibly due to areas with high building density having more amenity facilities and job opportunities. As Hong Kong is a coastal city, housing units along the coastal lines have higher NDWI and enjoy more sea view premium on the housing prices (Figs 12e and 12f). The sky view of housing units near the CBD is lower than in other areas (Fig 12k); however, it positively contributes to housing prices (Fig 12l).



**Fig 12. Spatial distribution of feature value and SHAP values of image features.**  
<https://doi.org/10.1371/journal.pone.0321951.g012>

#### Practical implications of this study

Given the model interpretability analysis results, the implications for stakeholders involved in the real estate industry are provided: (1) **public authorities** can use the proposed housing price prediction model to formulate a formal property valuation model by considering multi-source image data fusion. The valuation model could be used in multiple government-led application scenarios: property tax estimation, dispute resolution about urban renewal compensation determination, and housing affordability assessment, etc. (2) **urban planners** can have an in-depth understanding of how urban infrastructure, as well as urban green and blue space, affects housing prices. The findings will enable the planners to make necessary adjustments to facilitate balanced urban development. For instance, POI diversity and accessibility are more important than POI density in determining housing prices in Hong Kong. Urban planners could be committed to increasing the diversity and accessibility of amenity facilities to meet the community's evolving needs. The unevenly distributed urban greenness has been identified in Hong Kong, and urban planners should prioritize vegetation development in the urban centre to ensure green justice. (3) **real estate developers** should pay more attention to site planning and architectural design (e.g., floor area ratio optimization) to improve residents' living conditions since the floor area is the most significant feature. Understanding the positive and negative effects of housing price determinants is necessary for real estate developers to help them evaluate potential development sites, conduct feasibility studies for new projects, and make data-driven decisions about investment opportunities. For instance, the project sites with lower NDBI and higher NDVI could be considered unpromising and lack investment potential in Hong Kong.

#### Limitations and future research directions

This study has limitations: (1) The feature configuration parameter types and candidates used in this study are limited. Future studies can use more feature configuration parameters, such as diversified accessibility measures and CNN types for image extraction. Multiple hyperparameter optimization methods could be used to find the best feature combination instead of the time-consuming grid search used in this study; (2) There is a large computational demand for image feature processing and pipeline execution, which could consume much time for large-scale image data and generated pipelines. The potential bottlenecks for improving the computational efficiency of the proposed approach include CPU/GPU capability and training speed. Therefore, the approach could be further improved by enhancing the scalability in terms of hardware (e.g., better and more high-performance CPU/GPUs) and software (e.g., more efficient computing strategies); (3) The effectiveness of multi-source image data may vary significantly across different regions, and it would make the findings more convincing by conducting more experiments across diverse areas to analyze the generalizability of the findings and adaptability of the proposed framework. Future studies could use datasets in multiple regions to examine and compare the effects of multi-source images on housing prices.

#### Conclusions

Using fine-scale housing transaction data in Hong Kong, this paper proposes an enhanced ML framework for residential property valuation with multi-source image fusion, including exterior estate photos, street view images, and remote sensing images. The research results show that different feature configuration parameters can significantly affect model performances. Formulating a series of datasets and searching for the optimal feature configuration parameters is necessary to improve the overall model performance. The MRMR-based feature selection method can effectively determine the optimal feature set. Extra Tree performs significantly better than others.

Model interpretability analysis is conducted based on the optimal machine learning pipeline, and the results prove that image features play significant roles in determining the prices. Half of the top 10 significant features are image features, including building view (53.4%), NDVI (45.3%), NDBI (44.8%), NDWI (41.0%), and sky view (26.9%). Incorporating image features into prediction

models can improve the accuracy of housing price prediction, increasing the R squared from 0.809 to 0.821. Nonlinear and positive associations exist between housing prices and NDBI, NDWI, vegetation and building view. NDVI and sky view have nonlinear and negative associations with housing prices. The spatial distribution patterns of image feature values and corresponding SHAP main effects vary significantly from the city centre to the suburbs. Housing units near the centre are characterized by (1) lower NDVI, sky view and vegetation view index and (2) higher NDBI, NDWI, and building view index. This study provides practical implications for real estate stakeholders, including public authorities, urban planners, and developers. Future research directions could emphasize diversified feature configuration parameters, efficient computing strategies, and more case studies for validation.

## References

1. International Association of Assessing Officers, "Standard on Automated Valuation Models (AVMs) International Association of Assessing Officers," 2018. [Online]. Available: [www.iaao.org](http://www.iaao.org).  
[View Article](#) • [Google Scholar](#)
2. Ma J, Cheng JCP, Jiang F, Chen W, Zhang J. Analyzing driving factors of land values in urban scale based on big data and non-linear machine learning techniques. *Land Use Policy*. 2019;94:104537.  
[View Article](#) • [Google Scholar](#)
3. Soltani A, Pettit CJ, Heydari M, Aghaei F. Housing price variations using spatio-temporal data mining techniques. *J Hous Built Environ*. 2021;36(3):1199–227.  
[View Article](#) • [Google Scholar](#)
4. Soltani A, Heydari M, Aghaei F, Pettit CJ. Housing price prediction incorporating spatio-temporal dependency into machine learning algorithms. *Cities*. 2021;131:103941.  
[View Article](#) • [Google Scholar](#)
5. Rosen S. Hedonic prices and implicit markets: Product differentiation in pure competition. *J Polit Econ*. 1974;82(1):34–55.  
[View Article](#) • [Google Scholar](#)
6. Yacim JA, Boshoff DGB. Impact of artificial neural networks training algorithms on accurate prediction of property values. *J Real Estate Res*. 2018;40(3):375–418.  
[View Article](#) • [Google Scholar](#)
7. Alfaro-Navarro J-L, Cano EL, Alfaro-Cortés E, García N, Gámez M, Larraz B. A fully automated adjustment of ensemble methods in machine learning for modeling complex real estate systems. *Complexity*. 2020;2020:1–12.  
[View Article](#) • [Google Scholar](#)
8. Su T, Li H, An Y. A BIM and machine learning integration framework for automated property valuation. *J Build Eng*. 2021;44:102636.  
[View Article](#) • [Google Scholar](#)
9. Yoo S, Im J, Wagner JE. Variable selection for hedonic model using machine learning approaches: A case study in Onondaga County, NY. *Landsc Urban Plan*. 2012;107(3):293–306.  
[View Article](#) • [Google Scholar](#)
10. Kang Y, Zhang F, Peng W, Gao S, Rao J, Duarte F, et al. Understanding house price appreciation using multi-source big geo-data and machine learning. *Land Use Policy*. 2019;111:104919.  
[View Article](#) • [Google Scholar](#)
11. Vargas-Calderón V, Camargo JE. Towards robust and speculation-reduction real estate pricing models based on a data-driven strategy. *J Oper Res Soc*. 2021;73(12):2794–807.  
[View Article](#) • [Google Scholar](#)
12. Wan WX, Lindenthal T. Testing machine learning systems in real estate. *Real Estate Econ*. 2023;51(3):754–78.  
[View Article](#) • [Google Scholar](#)
13. Chou JS, Fleschman DB, Truong DN. Comparison of machine learning models to provide preliminary forecasts of real estate prices, vol. 37, no. 4. Springer Netherlands, 2022. doi: 10.1007/s10901-022-09937-1.
14. Zhan C, Liu Y, Wu Z, Zhao M, Chow TWS. A hybrid machine learning framework for forecasting house price. *Expert Syst Appl*. 2023;233:120981.  
[View Article](#) • [Google Scholar](#)
15. Taecharungroj V. Google Maps amenities and condominium prices: Investigating the effects and relationships using machine learning. *Habitat Int*. 2021;118:102463.  
[View Article](#) • [Google Scholar](#)
16. Rico-Juan JR, Taltavull de La Paz P. Machine learning with explainability or spatial hedonics tools? An analysis of the asking prices in the housing market in Alicante, Spain. *Expert Syst Appl*. 2021;171:114590.  
[View Article](#) • [Google Scholar](#)
17. Iban MC. An explainable model for the mass appraisal of residences: The application of tree-based machine learning algorithms and interpretation of value determinants. *Habitat Int*. 2022;128:102660.  
[View Article](#) • [Google Scholar](#)

18. Deppner J, von Ahlefeldt-Dehn B, Beracha E, Schaefer W. *Boosting the Accuracy of Commercial Real Estate Appraisals: An Interpretable Machine Learning Approach*, no. 0123456789. Springer US, 2023. doi: [10.1007/s11146-023-09944-1](https://doi.org/10.1007/s11146-023-09944-1).
19. Stang M, Krämer B, Cajias M, Schäfers W. Changing the location game – improving location analytics with the help of explainable AI. *J Real Estate Res.* 2023;46(4):421–43.  
[View Article](#) • [Google Scholar](#)
20. Dou M, Gu Y, Fan H. Incorporating neighborhoods with explainable artificial intelligence for modeling fine-scale housing prices. *Appl Geogr.* 2023;158:103032.  
[View Article](#) • [Google Scholar](#)
21. Lorenz F, Willwersch J, Cajias M, Fuerst F. Interpretable machine learning for real estate market analysis. *Real Estate Econ.* 2022;51(5):1178–208.  
[View Article](#) • [Google Scholar](#)
22. Choy LHT, Ho WKO. The use of machine learning in real estate research. *Land.* 2023;12(4):740.  
[View Article](#) • [Google Scholar](#)
23. Geerts M, vanden Broucke S, De Weerd J. A survey of methods and input data types for house price prediction. *IJGI.* 2023;12(5):200.  
[View Article](#) • [Google Scholar](#)
24. Tekouabou SCK, Gherghina ȘC, Kameni ED, Filali Y, Idrissi Gartoumi K. AI-based on machine learning methods for Urban real estate prediction: A systematic survey. *Arch Comput Methods Eng.* 2023;31(2):1079–95.  
[View Article](#) • [Google Scholar](#)
25. Ibrahim MR, Haworth J, Cheng T. Understanding cities with machine eyes: A review of deep computer vision in urban analytics. *Cities.* 2020;96:102481.  
[View Article](#) • [Google Scholar](#)
26. Biljecki F, Ito K. Street view imagery in urban analytics and GIS: A review. *Landsc Urban Plan.* 2021;215:104217.  
[View Article](#) • [Google Scholar](#)
27. Yang S, Krenz K, Qiu W, Li W. The role of subjective perceptions and objective measurements of the urban environment in explaining house prices in greater London: A multi-scale urban morphology analysis. *IJGI.* 2023;12(6):249.  
[View Article](#) • [Google Scholar](#)
28. Kuroda Y, Sugawara T, *The Value of Scattered Greenery in Urban Areas: A Hedonic Analysis in Japan*, vol. 85, no. 2. Springer Netherlands, 2023.  
<https://doi.org/10.1007/s10640-023-00775-5>
29. Qiu W, Li W, Liu X, Zhang Z, Li X, Huang X. Subjective and objective measures of streetscape perceptions: Relationships with property value in Shanghai. *Cities.* 2023;132:104037.  
[View Article](#) • [Google Scholar](#)
30. Wang R, Rasouli S. Contribution of streetscape features to the hedonic pricing model using geographically weighted regression: Evidence from Amsterdam. *Tour Manag.* 2022;91:104523.  
[View Article](#) • [Google Scholar](#)
31. Suzuki M, Mori J, Maeda TN, Ikeda J. The economic value of urban landscapes in a suburban city of Tokyo, Japan: A semantic segmentation approach using Google Street View images. *J Asian Archit Build Eng.* 2022;22(3):1110–25.  
[View Article](#) • [Google Scholar](#)
32. Qiu W, Zhang Z, Liu X, Li W, Li X, Xu X, et al. Subjective or objective measures of street environment, which are more effective in explaining housing prices?. *Landsc Urban Plan.* 2022;221:104358.  
[View Article](#) • [Google Scholar](#)
33. Potrawa T, Tetereva A. How much is the view from the window worth? Machine learning-driven hedonic pricing model of the real estate market. *J Bus Res.* 2022;144:50–65.  
[View Article](#) • [Google Scholar](#)
34. Chen M, Liu Y, Arribas-Bel D, Singleton A. Assessing the value of user-generated images of urban surroundings for house price estimation. *Landsc Urban Plan.* 2022;226:104486.  
[View Article](#) • [Google Scholar](#)
35. Luo J, Zhai S, Song G, He X, Song H, Chen J, et al. Assessing inequity in green space exposure toward a “15-Minute City” in Zhengzhou, China: Using deep learning and urban big data. *Int J Environ Res Public Health.* 2022;19(10):5798. pmid:35627336  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
36. Wu C, Du Y, Li S, Liu P, Ye X. Does visual contact with green space impact housing prices? An integrated approach of machine learning and hedonic modeling based on the perception of green space. *Land Use Policy.* 2022;115:106048.  
[View Article](#) • [Google Scholar](#)
37. Liao X, Deng M, Huang H. Analyzing multiscale spatial relationships between the house price and visual environment factors. *Appl Sci.* 2021;12(1):213.  
[View Article](#) • [Google Scholar](#)

38. Xu X, Qiu W, Li W, Liu X, Zhang Z, Li X, et al. Associations between street-view perceptions and housing prices: Subjective vs. objective measures using computer vision and machine learning techniques. *Remote Sens.* 2022;14(4):891.  
[View Article](#) • [Google Scholar](#)
39. Li S, Jiang Y, Ke S, Nie K, Wu C. Understanding the effects of influential factors on housing prices by combining extreme gradient boosting and a hedonic price model (XGBoost-HPM). *Land.* 2021;10(5):533.  
[View Article](#) • [Google Scholar](#)
40. Wang P-Y, Chen C-T, Su J-W, Wang T-Y, Huang S-H. Deep learning model for house price prediction using heterogeneous data analysis along with joint self-attention mechanism. *IEEE Access.* 2021;9:55244–59.  
[View Article](#) • [Google Scholar](#)
41. Yang J, Rong H, Kang Y, Zhang F, Chegut A. The financial impact of street-level greenery on New York commercial buildings. *Landsc Urban Plan.* 2021;214:104162.  
[View Article](#) • [Google Scholar](#)
42. Lee C, Park K-H. Using photographs and metadata to estimate house prices in South Korea. *DTA.* 2020;55(2):280–92.  
[View Article](#) • [Google Scholar](#)
43. Lin RF-Y, Ou C, Tseng K-K, Bowen D, Yung KL, Ip WH. The Spatial neural network model with disruptive technology for property appraisal in real estate industry. *Technol Forecast Soc Change.* 2021;173:121067.  
[View Article](#) • [Google Scholar](#)
44. Kang Y, Zhang F, Gao S, Peng W, Ratti C. Human settlement value assessment from a place perspective: Considering human dynamics and perceptions in house price modeling. *Cities.* 2021;118:103333.  
[View Article](#) • [Google Scholar](#)
45. Bin J, Gardiner B, Li E, Liu Z. Multi-source urban data fusion for property value assessment: A case study in Philadelphia. *Neurocomputing.* 2020;404:70–83.  
[View Article](#) • [Google Scholar](#)
46. Law S, Seresinhe CI, Shen Y, Gutierrez-Roig M. Street-Frontage-Net: Urban image classification using deep convolutional neural networks. *Int J Geogr Inf Sci.* 2018;34(4):681–707.  
[View Article](#) • [Google Scholar](#)
47. Kostic Z, Jevremovic A. What image features boost housing market predictions?. *IEEE Trans Multimed.* 2020;22(7):1904–16.  
[View Article](#) • [Google Scholar](#)
48. Chen L, Yao X, Liu Y, Zhu Y, Chen W, Zhao X, et al. Measuring impacts of urban environmental elements on housing prices based on multisource data—A case study of Shanghai, China. *IJGI.* 2020;9(2):106.  
[View Article](#) • [Google Scholar](#)
49. Ye Y, Xie H, Fang J, Jiang H, Wang D. Daily accessed street greenery and housing price: Measuring economic performance of human-scale streetscapes via new urban data. *Sustainability.* 2019;11(6):1741.  
[View Article](#) • [Google Scholar](#)
50. Law S, Paige B, Russell C. Take a look around. *ACM Trans Intell Syst Technol.* 2019;10(5):1–19.  
[View Article](#) • [Google Scholar](#)
51. Yencha C. Valuing walkability: New evidence from computer vision methods. *Transp Res Part A: Policy Pract.* 2019;130:689–709.  
[View Article](#) • [Google Scholar](#)
52. Li C, Zou L, Wu Y, Xu H. Potentiality of using luojia1-01 night-time light imagery to estimate Urban community housing price-A case study in Wuhan, China. *Sensors (Basel).* 2019;19(14):3167. pmid:31323879  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
53. Fu X, Jia T, Zhang X, Li S, Zhang Y. Do street-level scene perceptions affect housing prices in Chinese megacities? An analysis using open access datasets and deep learning. *PLoS One.* 2019;14(5):e0217505. pmid:31145767  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
54. Poursaeed O, Matera T, Belongie S. Vision-based real estate price estimation. *Mach Vis Appl.* 2018;29(4):667–76.  
[View Article](#) • [Google Scholar](#)
55. Zhang Y, Dong R. Impacts of street-visible greenery on housing prices: Evidence from a hedonic price model and a massive street view image dataset in Beijing. *IJGI.* 2018;7(3):104.  
[View Article](#) • [Google Scholar](#)
56. Jiao L, Xu G, Jin J, Dong T, Liu J, Wu Y, et al. Remotely sensed urban environmental indices and their economic implications. *Habitat Int.* 2017;67:22–32.  
[View Article](#) • [Google Scholar](#)
57. You Q, Pang R, Cao L, Luo J. Image-based appraisal of real estate properties. *IEEE Trans Multimed.* 2017;19(12):2751–9.



[View Article](#) • [Google Scholar](#)

58. Ye Y, Richards D, Lu Y, Song X, Zhuang Y, Zeng W, et al. Measuring daily accessed street greenery: A human-scale approach for informing better urban planning practices. *Landsc Urban Plan.* 2019;191:103434.  
[View Article](#) • [Google Scholar](#)
59. Yao Y, Wang J, Hong Y, Qian C, Guan Q, Liang X, et al. Discovering the homogeneous geographic domain of human perceptions from street view images. *Landsc Urban Plan.* 2021;212:104125.  
[View Article](#) • [Google Scholar](#)
60. Zhang F, Fan Z, Kang Y, Hu Y, Ratti C. "Perception bias": Deciphering a mismatch between urban crime and perception of safety. *Landsc Urban Plan.* 2021;207:104003.  
[View Article](#) • [Google Scholar](#)
61. MIT Media Lab, "Place Pulse," 2024. <https://www.media.mit.edu/projects/place-pulse-new/overview/> (accessed Jun. 03, 2024).  
[View Article](#) • [Google Scholar](#)
62. Minka T, Cleven R, Zaykov Y, "TrueSkill 2: An improved Bayesian skill rating system," *Microsoft.Com*, pp. 1–24, 2018, [Online]. Available: <https://www.microsoft.com/en-us/research/uploads/prod/2018/03/trueskill2.pdf>  
[View Article](#) • [Google Scholar](#)
63. Li W, Saphores J-DM, Gillespie TW. A comparison of the economic benefits of urban green spaces estimated with NDVI and with high-resolution land cover data. *Landsc Urban Plan.* 2015;133:105–17.  
[View Article](#) • [Google Scholar](#)
64. Zhang P, Hu S, Li W, Zhang C, Yang S, Qu S. Modeling fine-scale residential land price distribution: An experimental study using open data and machine learning. *Appl Geogr.* 2021;129:102442.  
[View Article](#) • [Google Scholar](#)
65. Krizhevsky A, Sutskever I, Alex GEH. ImageNet classification with deep convolutional neural networks. *Proceedings of the [Conference Name].* 2012.
66. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural networks. *Pattern Recognit.* 2018;77:354–77.  
[View Article](#) • [Google Scholar](#)
67. Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid Scene Parsing Network. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. <https://doi.org/10.1109/cvpr.2017.660>
68. Zhou B, Lapedriza A, Khosla A, Oliva A, Torralba A. Places: A 10 million image database for scene recognition. *IEEE Trans Pattern Anal Mach Intell.* 2018;40(6):1452–64. pmid:28692961  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
69. Long J, Shelhamer E, Darrell T, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, vol. 7, pp. 3431–3440. [Online]. Available: [https://openaccess.thecvf.com/content\\_cvpr\\_2015/html/Long\\_Fully\\_Convolutional\\_Networks\\_2015\\_CVPR\\_paper.html](https://openaccess.thecvf.com/content_cvpr_2015/html/Long_Fully_Convolutional_Networks_2015_CVPR_paper.html)
70. Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell.* 2017;39(12):2481–95. pmid:28060704  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
71. Chen L-C, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation. 2017, [Online]. Available: <http://arxiv.org/abs/1706.05587>.  
[View Article](#) • [Google Scholar](#)
72. Shen D, Yang B, Li J, Zhang J, Li Y, Zhang G, et al. The potential associations between acupuncture sensation and brain functional network: a EEG study. *Cogn Neurodyn.* 2025;19(1):49. pmid:40099217  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
73. Wang P-Y, Chen C-T, Su J-W, Wang T-Y, Huang S-H. Deep learning model for house price prediction using heterogeneous data analysis along with joint self-attention mechanism. *IEEE Access.* 2021;9:55244–59.  
[View Article](#) • [Google Scholar](#)
74. Midland Realty, "Statistics of Properties Transactions in Land Registry - 2022," 2022. <https://en.midland.com.hk/land-registry-record/2022.html> (accessed Jun. 03, 2024).  
[View Article](#) • [Google Scholar](#)
75. He X, Zhao K, Chu X. AutoML: A survey of the state-of-the-art. *Knowl-Based Syst.* 2021;212:106622.  
[View Article](#) • [Google Scholar](#)
76. Zhang L, Zhou J, Hui EC. Which types of shopping malls affect housing prices? From the perspective of spatial accessibility. *Habitat Int.* 2020;96:102118.  
[View Article](#) • [Google Scholar](#)

77. Wen H, Xiao Y, Hui ECM. Quantile effect of educational facilities on housing price: Do homebuyers of higher-priced housing pay more for educational resources?. *Cities*. 2019;90:100–12.  
[View Article](#) • [Google Scholar](#)
78. Wen H, Zhang Y, Zhang L. "Do educational facilities affect housing price? An empirical study in Hangzhou, China," *Habitat Int*. 2014;42:155–163.  
[View Article](#) • [Google Scholar](#)
79. Wu C, Ye X, Du Q, Luo P. Spatial effects of accessibility to parks on housing prices in Shenzhen, China. *Habitat Int*. 2017;63:45–54.  
[View Article](#) • [Google Scholar](#)
80. Breunig R, Hasan S, Whiteoak K. Value of playgrounds relative to green spaces: Matching evidence from property prices in Australia. *Landsc Urban Plan*. 2019;190:103608.  
[View Article](#) • [Google Scholar](#)
81. Jin T, Cheng L, Liu Z, Cao J, Huang H, Witlox F. Nonlinear public transit accessibility effects on housing prices: Heterogeneity across price segments. *Transp Policy*. 2022;117:48–59.  
[View Article](#) • [Google Scholar](#)
82. Manville M. Parking requirements and housing development. *J Am Plan Assoc*. 2013;79(1):49–66.  
[View Article](#) • [Google Scholar](#)
83. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–44. pmid:26017442  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
84. van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res*. 2008;9(11):2579–605. <http://jmlr.org/papers/v9/vandermaaten08a.html>  
[View Article](#) • [Google Scholar](#)
85. Yu Y, Wang C, Fu Q, Kou R, Huang F, Yang B, et al. Techniques and challenges of image segmentation: A review. *Electronics*. 2023;12(5):1199.  
[View Article](#) • [Google Scholar](#)
86. Szegedy AR, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015, pp. 1–9.
87. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc*. 2015, pp. 1–14.
88. He JSK, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 770–8.
89. Isabelle G, Elisseeff A. "An introduction to variable and feature selection," *J Mach Learn Res*. 2003;3(2):1157–1182, [Online]. Available: <https://www.jmlr.org/papers/volume3/guyon03a/guyon03a.pdf?ref=driverlayer.com/web>  
[View Article](#) • [Google Scholar](#)
90. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell*. 2005;27(8):1226–38. pmid:16119262  
[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)
91. Thornton C, Hutter F, Hoos HH, Leyton-Brown K. Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. vol. Part F1288, pp. 847–55. 2013.  
<https://doi.org/10.1145/2487575.2487629>
92. Tchuenté D, Nyawa S. Real estate price estimation in French cities using geocoding and machine learning, vol. 308, no. 1–2. Springer US, 2022.  
<https://doi.org/10.1007/s10479-021-03932-5>
93. Gao Q, Shi V, Pettit C, Han H. Property valuation using machine learning algorithms on statistical areas in Greater Sydney, Australia. *Land Use Policy*. 2022;123:106409.  
[View Article](#) • [Google Scholar](#)
94. Breiman L. Bagging predictors. *Mach Learn*. 1996;24(2):123–40.  
[View Article](#) • [Google Scholar](#)
95. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Mach Learn*. 2006;63(1):3–42.  
[View Article](#) • [Google Scholar](#)
96. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. vol. 13-17-Aug, pp. 785–94, 2016. <https://doi.org/10.1145/2939672.2939785>
97. Bischl B, Binder M, Lang M, Pielok T, Richter J, Coors S, et al. Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. *WIREs Data Min Knowl*. 2023;13(2).  
[View Article](#) • [Google Scholar](#)
98. Frazier PI, "A tutorial on bayesian optimization," no. Section 5, pp. 1–22, 2018, [Online]. Available: <http://arxiv.org/abs/1807.02811>

[View Article](#) • [Google Scholar](#)

99. Wilcoxon F. Individual comparisons of grouped data by ranking methods. *J Econ Entomol.* 1946;39:269. pmid:20983181

[View Article](#) • [PubMed/NCBI](#) • [Google Scholar](#)

100. Friedman M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J Am Stat Assoc.* 1937;32(200):675–701.

[View Article](#) • [Google Scholar](#)

101. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. in *Advances in Neural Information Processing Systems*. 2017, vol. 30.

<https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>.