

حسن خليفة

Paper Clarivision System-3.docx

 Midocean University

Document Details

Submission ID

trn:oid::2945:404562957

Submission Date

Aug 3, 2026, 5:35 PM GMT+3

Download Date

Aug 3, 2026, 8:36 PM GMT+3

File Name

Paper Clarivision System-3.docx

File Size

3.1 MB

28 Pages

10,318 Words

60,020 Characters

10% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text

Match Groups

-  **109 Not Cited or Quoted 10%**
Matches with neither in-text citation nor quotation marks
-  **5 Missing Quotations 1%**
Matches that are still very similar to source material
-  **0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 4%  Internet sources
- 5%  Publications
- 9%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 109 Not Cited or Quoted 10%**
Matches with neither in-text citation nor quotation marks
- 5 Missing Quotations 1%**
Matches that are still very similar to source material
- 0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 4% Internet sources
- 5% Publications
- 9% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

- 1** Internet
mission.sfgov.org <1%
- 2** Internet
www.mdpi.com <1%
- 3** Submitted works
National Institute of Business Management Sri Lanka on 2026-03-31 <1%
- 4** Internet
thesai.org <1%
- 5** Submitted works
Tishk International University on 2026-06-22 <1%
- 6** Submitted works
National Institute of Business Management Sri Lanka on 2025-10-15 <1%
- 7** Submitted works
BB9.1 PROD on 2026-07-23 <1%
- 8** Submitted works
University of Hertfordshire on 2026-03-30 <1%
- 9** Internet
www.ijnrd.org <1%
- 10** Submitted works
Islington College,Nepal on 2026-01-21 <1%

11	Submitted works	BB9.1 PROD on 2026-07-29	<1%
12	Internet	cutscholar.cut.ac.za	<1%
13	Internet	github.com	<1%
14	Submitted works	Coventry University on 2025-12-03	<1%
15	Publication	Bhabani Prasanna Pattanaik, Dharendra Kumar Shukla, Sambit Satpathy, Kamala...	<1%
16	Publication	Yousif, Amira. "Automated House Valuation With Multimodality Learning", Univer...	<1%
17	Internet	mbrenndoerfer.com	<1%
18	Submitted works	Faculty of Engineering, Chulalongkorn University on 2026-06-16	<1%
19	Submitted works	University of East London on 2026-04-27	<1%
20	Publication	Gege Dong, Lu Qiu, Yanzhe Jiao. "Boosting Urban Housing Price Forecasting with ...	<1%
21	Submitted works	Islington College,Nepal on 2026-01-23	<1%
22	Publication	Richard Shan. "A Hybrid Ensemble Approach for Depression Detection: Combinin...	<1%
23	Submitted works	University of Nottingham on 2025-09-04	<1%
24	Submitted works	BB9.1 PROD on 2025-08-26	<1%

25	Submitted works	University of Wollongong on 2026-07-08	<1%
26	Submitted works	Asia Pacific University College of Technology and Innovation (UCTI) on 2026-06-05	<1%
27	Publication	Ian Lenaers, Lieven De Moor. "Analyzing the impact of energy efficiency on Belgi...	<1%
28	Submitted works	University of Sheffield on 2025-11-12	<1%
29	Internet	mro.massey.ac.nz	<1%
30	Publication	"Digital Technology for an Innovative Energy Transition", Springer Science and B...	<1%
31	Internet	1www.jemi.edu.pl	<1%
32	Submitted works	Asia Pacific University College of Technology and Innovation (UCTI) on 2025-10-05	<1%
33	Publication	Gege Dong, Lu Qiu, Yanzhe Jiao. "Boosting urban housing price forecasting with a...	<1%
34	Submitted works	PSG Institutions on 2025-05-29	<1%
35	Submitted works	Sunway Education Group on 2025-06-26	<1%
36	Submitted works	UNICAF on 2026-02-22	<1%
37	Submitted works	University of East London on 2025-09-09	<1%
38	Submitted works	University of New South Wales on 2025-08-03	<1%

39	Internet	ideas.repec.org	<1%
40	Publication	Amiri, Ahmed. "Rental Property Demand and Supply Analysis Using Machine Lear...	<1%
41	Publication	Amit Kumar Tyagi, Shabnam Kumari. "Artificial Intelligence and Digital Twin-Base...	<1%
42	Submitted works	Asia Pacific University College of Technology and Innovation (UCTI) on 2023-07-11	<1%
43	Submitted works	Blackboard on 2026-06-08	<1%
44	Submitted works	Capella University on 2025-08-12	<1%
45	Submitted works	College of Public Health on 2026-03-12	<1%
46	Publication	Eltayeb, Abdelwahid A.. "Integrating Ensemble Learning With QASR for Scalable S...	<1%
47	Publication	Nasser, Mohammad. "Predicting Residential Real Estate Prices in Dubai Using Int...	<1%
48	Submitted works	Trine University on 2025-09-01	<1%
49	Internet	journals.ksauniv.ks.ua	<1%
50	Internet	scholar.ppu.edu	<1%
51	Submitted works	American Public University System on 2026-04-19	<1%
52	Submitted works	Berlin School of Business and Innovation on 2025-06-16	<1%

53	Submitted works	Chapman University on 2024-12-02	<1%
54	Submitted works	De Montfort University on 2026-05-28	<1%
55	Submitted works	Islington College,Nepal on 2026-01-07	<1%
56	Submitted works	King's College on 2025-04-03	<1%
57	Submitted works	Liverpool John Moores University on 2026-01-30	<1%
58	Submitted works	Loughborough University on 2024-05-02	<1%
59	Submitted works	Manchester Metropolitan University on 2024-09-26	<1%
60	Submitted works	Napier University on 2025-12-05	<1%
61	Submitted works	Queen's University of Belfast on 2023-07-28	<1%
62	Submitted works	Rochester Institute of Technology on 2025-09-16	<1%
63	Submitted works	Sunway Education Group on 2025-12-05	<1%
64	Submitted works	The Robert Gordon University on 2025-12-15	<1%
65	Submitted works	University College London on 2026-04-24	<1%
66	Submitted works	University of East London on 2023-05-05	<1%

67	Submitted works	University of Hertfordshire on 2022-12-22	<1%
68	Submitted works	University of Malaya on 2026-06-05	<1%
69	Submitted works	University of Pretoria on 2023-12-12	<1%
70	Submitted works	University of Salford on 2023-05-04	<1%
71	Submitted works	University of Sunderland on 2026-06-18	<1%
72	Internet	ris.utwente.nl	<1%
73	Submitted works	universiteknologimara on 2026-08-03	<1%
74	Internet	www.frontiersin.org	<1%
75	Internet	www.itu.int	<1%
76	Publication	"Proceedings of Fourth International Conference on Computing and Communicat...	<1%
77	Submitted works	Dundalk Institute of Technology on 2023-09-25	<1%
78	Submitted works	National Institute of Business Management Sri Lanka on 2024-09-28	<1%
79	Submitted works	Universitas Esa Unggul on 2026-07-26	<1%
80	Submitted works	Crown Institute of Business and Technology on 2025-09-16	<1%

81	Publication	Harvinder Singh, Priyanka Kaushal, Sarabjeet Kaur. "Advanced Computing and AI..."	<1%
82	Publication	Jaskaran Singh, Meenu Gupta, Rakesh Kumar. "Smart Technologies and Intelligen..."	<1%
83	Submitted works	Rochester Institute of Technology on 2025-12-10	<1%
84	Submitted works	Strathmore University (Main Account) on 2026-03-26	<1%
85	Publication	Ebrat, Danial. "Toward Intelligent Personalization: Integrating Large Language M..."	<1%

Clarivision AI: An Explainable Hybrid Framework for Residential Property Valuation and Similar Property Retrieval Using Structured Data and ResNet50-Based Deep Visual Features

Hassan Mohammed Ahmed Alamer Khalifa

ORCID ID :0009-0006-0611-1563

Collage of Informatics, Midocean University, Comoros

Email: hassanalamer2019@gmail.com

Supervisor Name: Dr / Amjad Qtaish

Abstract

Objectives: This study develops Clarivision AI, an explainable hybrid framework for residential property valuation that integrates structured property attributes with deep visual features extracted using ResNet50 and enhances model interpretability through Explainable Artificial Intelligence (XAI).

Study Location: The study was conducted using the Real Estate with Images dataset, comprising 15,474 residential property records with corresponding structured attributes and property images.

Methodology: The dataset was divided into 80% training data and 20% testing data using a fixed random seed of 42. Property prices were transformed using $\log(1 + \text{price})$ to reduce target skewness. Structured property attributes were preprocessed and combined with deep visual embeddings extracted using ResNet50 to construct hybrid features. An XGBoost regression model was then developed for property valuation, while SHAP was used to interpret the contributions of the structured features. A similar-property retrieval mechanism based on cosine similarity and an interactive decision-support interface developed using Streamlit were also incorporated.

Results: The proposed hybrid model achieved promising predictive performance, with an MAE of 0.20985, an RMSE of 0.29112, and an R^2 value of 0.67375. The MAE and RMSE values were calculated on the log-transformed property price scale. After back-transforming the predictions to the original price scale, the model achieved an MAE of 151,078.31 and an RMSE of 238,044.55, expressed in the original price units of the dataset. The results indicate that the model explained approximately 67.4% of the variance in the log-transformed property prices. SHAP provided explanations of the contributions of the structured property features, while the developed platform supported similar-property retrieval and interactive visualization.

Recommendations: To include more structured and visual property attributes. To test framework on datasets from other countries. To apply more XAI methods. To extend the framework with property classification and cloud deployment of AI models.

Keywords: Real Estate Valuation, Hybrid Model, Explainable Artificial Intelligence, XGBoost, ResNet50, Clarivision AI

الكلمات المفتاحية: التقييم العقاري، النموذج الهجين، كلاريفيجن، الذكاء الاصطناعي القابل للتفسير

1. Introduction.

14 The real estate industry is critical to the growth of an economy and in making decisions. The accurate valuation of residential properties is critical for the buyers, sellers, financial institutions, and the government organizations.

10 Nonetheless, the conventional methods of property valuation are majorly dependent on human knowledge and the structured data, hence making the process subjective, inconsistent and time-consuming because of the dependence on the quality of the information regarding the property.

2 However, recent developments in Artificial Intelligence (AI) and Machine Learning (ML) have greatly enhanced the accuracy of the Automated Valuation Models (AVMs). Some of the latest prediction algorithms like Random Forest, XGBoost, and Deep Learning have shown great performance in terms of their ability to identify non-linear relations between the attributes of the properties and market prices. However, most Automated Valuation Models (AVMs) still depend mainly on the structured attributes of the property while ignoring the visual information available from residential property pictures. Moreover, many of the AI valuation systems remain black boxes hence reducing the understanding of the valuation process.

In order to overcome the limitations described above, this paper puts forth Clarivision AI, a novel explainable hybrid approach for fusing structured property attributes with deep visual representations obtained via the ResNet50 convolutional neural network. Structured numerical property attributes and deep visual representations are fused in the hybrid learning architecture based on XGBoost, and SHapley Additive explanations (SHAP) are used for explaining predictions made by the model. In addition, Clarivision AI includes a Similar Property Retrieval component, where visually similar residential property listings are retrieved via cosine similarity measure over deep image embeddings.

Even though recent research showed significant improvements in residential property valuation by applying machine learning, multimodal learning, and XAI techniques, most of the available solutions consider these components separately. There are very few papers that introduce the framework combining all the discussed aspects into one decision-support solution. It is exactly this research gap that motivates the introduction of Clarivision AI hybrid framework.

3 The key contribution of the study is the creation of an integrated explainable decision support system which integrates hybrid property valuation, visual similarity search, model interpretability, and interactive reporting in one intelligent system. The experiment shows that the developed framework is not only better at valuating residential properties but also increases the transparency, usability, and trustworthiness of real estate decision making.

2. Research Problem

Despite significant achievements in the use of Artificial Intelligence (AI) in Automated Valuation Models (AVMs), some problems exist concerning the accuracy, transparency, and usability of existing models for residential real estate valuation. For example, although most models incorporate structural aspects such as size and number of rooms, they do not take into account some other features that can influence the appearance of the property. Furthermore, the current deep learning models used for the purpose of valuation of properties are black boxes, which makes it difficult for users to understand how the decision-making process works in the valuation system. As a result, there is a need to develop the technology of Explainable Artificial Intelligence (XAI). Even though current models provide accurate price predictions, they do not provide enough explanations of how their decisions are reached. In addition, previous research efforts do not cover other aspects apart from price prediction.

Thus, it becomes necessary to have a hybrid explainable valuation system, which could incorporate structured property features, deep vision-based visual features, predictability and explainability, similar property searches, and decision support in an all-in-one intelligent system. The Clarivision AI has been created in order to resolve these limitations and create exact and understandable valuation of residential properties.

Therefore, the aim of this research paper is to solve the problems of existing residential property valuation models and develop an integrated Explainable Artificial Intelligence model.

3. Importance of the Study.

2.1 Academic Significance

This study, will help to enhance the understanding of the use of artificial intelligence for the valuation of property. It discusses the application of a novel approach that involves integrating data about the house such as its measurements and the location with images of the property to obtain an enhanced valuation. Moreover, the research applies the application of SHAP technique that aids in interpreting the decision-making process made by the machine learning model.

2.2 Practical Significance

The above-discussed Clarivision AI solution provides a tool for making decisions on residential properties. The prediction, interpretability of models, similar property search, visualization and PDF reporting capabilities have been incorporated to assist the real estate professionals and other stakeholders in making proper valuations.

4. Study objectives

Key Goals of the Study:

- Develop a hybrid residential property valuation model by integrating structural and visual property features using ResNet50.
- Enhance the accuracy of residential property valuation through an XGBoost regression model.
- Improve the transparency and interpretability of the models using Shapley Additive Explanations (SHAP).
- Create a Similar Property Retrieval module that fetches properties from a selected city and uses cosine similarity on ResNet50 visual embeddings to identify the most visually comparable properties for comparative analysis.
- Build an interactive decision support platform, Clarivision AI, incorporating residential property valuation capabilities.

5. Theoretical Literature of the Study.

Real estate valuation has changed remarkably during recent decades from conventional valuations methods towards intelligent Artificial Intelligence (AI) approaches. The conventional approaches to real estate valuation heavily rely on professional appraisers who calculate property values using their market experience and property characteristics. However, despite the widespread use of these valuation methods, they are typically time-consuming and inconsistent.

The rapid advancement in Machine Learning (ML) has resulted in remarkable improvements in Automated Valuation Model (AVM) performance. Such ML algorithms as Decision Trees, Random Forest, Gradient Boosting Machines, and XGBoost have shown remarkable ability to model complicated nonlinear relations between property characteristics and market prices. They have been constantly showing superiority in capturing complicated nonlinear relations between property characteristics and market prices which resulted in more accurate property valuations than conventional statistical approaches.

In more recent times, however, Deep Learning approaches have brought new possibilities to the domain of property valuation in that they use visual information derived from property pictures. ResNet50 CNNs, for example, have demonstrated outstanding capability in capturing architectural specifics, visual appearance of the house, its external state, which is impossible to encode in numeric form. Structured data combined with deep visual features can thus be viewed as an effective solution to improving the accuracy of residential property valuation models.

However, even with the development of new solutions, many AI-based valuation models lack interpretability. In order to solve this problem, a new approach called Explainable Artificial Intelligence (XAI) was proposed. It is an approach that provides clear explanations for the results obtained from model outputs. The most widely used XAI technique is the SHapley Additive exPlanation (SHAP). It makes people more confident about using these AI-driven models. When it comes to valuation of residential properties, SHAP makes it possible to get the exact contribution of every attribute of a property to the house price prediction.

In this regard, recent research focuses on the development of hybrid explainable valuation models by means of the use of structured data and deep visual features. Furthering the line of research, this study presents the development of Clarivision AI model that includes structured data, deep visual features based on ResNet50 architecture, prediction with the help of XGBoost algorithm, SHAP explanation, similar property retrieval and decision support capabilities.

6. Review of Previous Studies

In recent years, developments in AI and ML have changed the valuation process of houses, and gone beyond the conventional methods of valuation. The literature review presented here can be classified into four different areas of research: (1) Structured-data valuation; (2) Deep learning and multimodal valuation; (3) Explainable artificial intelligence, and (4) Intelligent decision support systems. Even though all these methods have contributed to the betterment of valuations, their differences in terms of data modalities, model design, explainability, and functionalities are considerable.

6.1 Structured-Data Valuation Models

The application of machine learning to house valuation has been in practice for a while already. Traditionally, it accounted for such features as the size of a property, its location, number of bedrooms and bathrooms, and the details of the sale process. However, several researches suggested that it was possible to improve machine learning models by including geographical data about the area of the property. For instance, the recent study done by Wang et al. offers a new model for valuing houses in Shenzhen, China. It utilizes a unique geographical deep neural network, which accounts for the location of a property, and proves to be extremely effective. At the same time, Mora-Garcia et al. applied various machine learning algorithms for predicting housing prices. The results of their experiments demonstrate that modern techniques may be as successful, as the traditional ones. This fact is especially important, since it implies that machine learning may serve as a powerful tool for house valuation.

Balila and Shabri conducted a more detailed analysis on how different algorithms of machine learning are able to forecast property prices in Dubai. The authors collected an extensive amount of structured data regarding properties to find out how different approaches work. They concluded that preprocessing of the data, selection of the necessary features, and transformation of such features are crucial. In addition, Balila and Shabri made comparisons of different models in order to understand what approach should be selected to obtain the most accurate results. It has been found out that one of the major aspects which affects property valuation is its size. However, there are several limitations in the research of Balila and Shabri. Namely, they have considered only structured data without including images of properties and different features at the same time.

Most structured data-based models are capable of producing practical and comprehensible estimations of the property's value. However, there are certain limitations regarding the model's ability to reflect the visual characteristics of the property, which include its exterior appearance, interior quality, state of repair, as well as finishing. Moreover, they cannot incorporate any additional visible characteristics that can be captured in an image. This is the reason behind the need to come up with new methods of property valuation based on pictures and data combination.

6.2 Deep Learning and Multimodal Property Valuation

The recent research has included the analysis of property images in the automated valuation model. It is possible to get visual representations from the property images and integrate them with structured property data with the help of deep learning methods.

Yazdani and Raissi (2023) examined the application of self-supervised vision transformers in real estate valuation and demonstrated the benefits of visual representations in property valuation. The authors Hasan et al. (2024) suggested multimodal deep learning technique for predicting the price of a house based on integrating property images and structured information. In turn, Deng et al. (2025) developed multimodal machine learning pipeline, including several information sources and explainable artificial intelligence.

This proves that the use of multiple ways of learning provides more information on property than only numerical values. However, it should be noted that most of the research analyzed aimed at the improvement of predictions

and did not include the analysis of similar properties based on property images, interactive data visualization and automated reporting within a single platform for property valuation.

Huang et al., for example, demonstrated that combining various sources of data, including numerical and graphical, may be useful in determining the value of the property. It is considered essential for researchers to combine various data sources to gain an improved understanding of the value of a property. Unfortunately, existing methods are not transparent enough and cannot be used effectively in practical situations.

6.3 Explainable Artificial Intelligence in Property Valuation

As models for machine learning become more complicated, model transparency becomes crucial for assessing properties. The methods of Explainable Artificial Intelligence (XAI) could help detect the influential factors for property value estimation.

64

In their study, Gunes (2024) analyzed model-agnostic interpretability methods of the residential property valuation and revealed the significance of explanation of model predictions. Moreover, Maselli and Nesticò (2025) stressed the importance of explaining AI models in terms of property valuation. Furthermore, the methods of XGBoost and Shapley values were used by Kee and Ho (2025) to explore the influence of property characteristics on predicted values.

This research proved the possibility of enhancing the transparency of the automated valuation models. Nevertheless, explainability was considered as an additional tool of analysis instead of a part of the whole decision support system, which includes price prediction, visual similarity retrieval, interactive visualization, and automated reporting.

6.4 Intelligent Decision-Support Systems

21

Recent research has explored intelligent systems that extend beyond numerical property price prediction. Geerts et al. (2025) examined the use of large language models in real estate appraisal, showcasing how language-based AI can support appraisal-related reasoning. However, their focus was primarily on language model capabilities rather than a comprehensive multimodal valuation system.

39

Rodriguez-Serrano (2025) studied prototype-based learning for real estate valuation, highlighting the benefits of example-based explanations. While prototype-based methods enhance interpretability, they do not offer image-based Similar Property Retrieval or a fully interactive decision-support environment.

On the whole, the above research literature demonstrates substantial advancements in the domain of structured data valuation, multimodal learning, explanation of artificial intelligence, and intelligent decision-making support systems. Nonetheless, such aspects are researched separately from each other. Very few academic papers consider price prediction, Similar Property Retrieval based on images, explanation of the models, interaction with visualizations, and automatic report generation simultaneously.

Research Direction	Dataset / Data Type	Modality	Model Type	Explainability	Similar Property Retrieval	Main Limitation
Structured-Data Valuation	Residential property attributes and geographic information	Structured data	GNNWR, ML, and ensemble models	Not emphasized	No	Limited consideration of property visual characteristics
Deep Learning and Multimodal Valuation	Property images combined with structured property data	Visual and multimodal data	Vision Transformers, CNN-based models, and hybrid ML models	Limited or study-dependent	No	Mainly focused on predictive performance rather than integrated decision support
Explainable AI in Property Valuation	Primarily structured property data	Structured data	Interpretable and model-agnostic ML approaches	SHAP and other XAI methods	No	Focused mainly on explaining predictions without visual retrieval or integrated services
Intelligent Decision-Support Systems	Structured and textual real-estate information	Structured and textual data	LLM-based and intelligent appraisal approaches	Model-dependent explanations	Limited	Limited integration of visual features and comprehensive valuation services
Clarivision AI — This Study	Real Estate with Images dataset containing 15,474 residential properties	Structured and visual data	ResNet50 visual embeddings + XGBoost regression	SHAP for structured feature contributions	Yes — city-filtered visual retrieval using cosine similarity	Explainability is currently focused primarily on structured features

Figure 1: Critical Comparison of Previous Research Directions (Source: Developed by the author based on the reviewed literature)

6.5 Critical Synthesis

According to the comparative analysis, the structured data-based valuation models have great predictive ability, although they tend to use only the numeric and geographic features of the properties without taking into account the visual properties. The deep learning and multimodal methods try to overcome this challenge by introducing the property images into the structure. Still, some of those studies emphasize predictive accuracy. The XAI approach tries to increase the level of transparency and explanation of the contribution of the features; still, those studies usually do not consider similar-property retrieval and decision-making assistance. The intelligent decision support studies go beyond the simple price prediction and create the valuation systems; however, they can operate mostly based on the structured or text-based data and ignore the visual features. Hence, it could be concluded that an integrated system is required for hybrid valuation, explanation, similar-property retrieval, visualization, and report generation.

6.6 Research Gap

From the reviewed literature, there has been a lot of success in the field of valuation of residential properties through the use of machine learning techniques for structured data, deep learning, multimodal methods, and Explainable Artificial Intelligence (XAI). Unfortunately, only very few of the studies in the literature seem to combine the use of structured property features, deep visual features, hybrid property valuation, SHAP-based explainability, similar property retrieval, visualization, and automatic reporting.

In order to fill this gap, this study proposes Clarivision AI; a hybrid explainable AI framework that combines the use of structured property features together with visual features generated using ResNet50, through an XGBoost regression algorithm. This framework also includes SHAP-based explanation, city-filtered Similar Property Retrieval using cosine similarity, visualization, and automatic PDF reporting.

This study adopted an experimental research methodology to develop, implement, and evaluate Clarivision AI, an explainable hybrid framework for residential property valuation. The proposed methodology integrates structured property attributes with deep visual representations extracted from property images to support residential property price prediction and improve the transparency of model outputs. The methodology consists of five sequential phases: (1) data collection and preprocessing, (2) deep visual feature extraction, (3) hybrid model development and evaluation, (4) model interpretability, and (5) similar-property retrieval and decision-support system development.

7.1 Dataset Source and Characteristics.

The dataset employed in this analysis is House Prices and Images - SoCal (ted8080, n.d.), which is openly available on Kaggle. It contains 15,474 observations related to residential properties in the southern part of California, where each observation contains one image of the property exterior using the image_id column. It comprises eight columns: image_id, street, city, n_citi, bed, bath, sqft, and price.

In developing the predictive model, four variables were chosen: city, the number of bedrooms, the number of bathrooms, and property square footage (sqft). The target variable was the property price. The variable n_citi was employed as the numerical value for the categorical variable of the city. The variable image_id was used to connect each observation with the respective image. The variable street was not included in the predictive model because of its high cardinality and textual type.

The dataset was sourced from a publicly available Kaggle page with an unknown license status. Therefore, it was used solely for academic research and experimental purposes, with proper acknowledgment of the original source.

7.2 Structured Data Preprocessing.

In order to make sure that we will not encounter problems associated with missing or duplicated entries in the process of the development of our prediction model, we decided to review the entire dataset for such issues. To our luck, after reviewing all eight attributes of the dataset, we have established that none of them contained any missing values and no duplicates could be found. Therefore, it means that there is no need for us to pay any attention to the problem of missing values and record duplication.

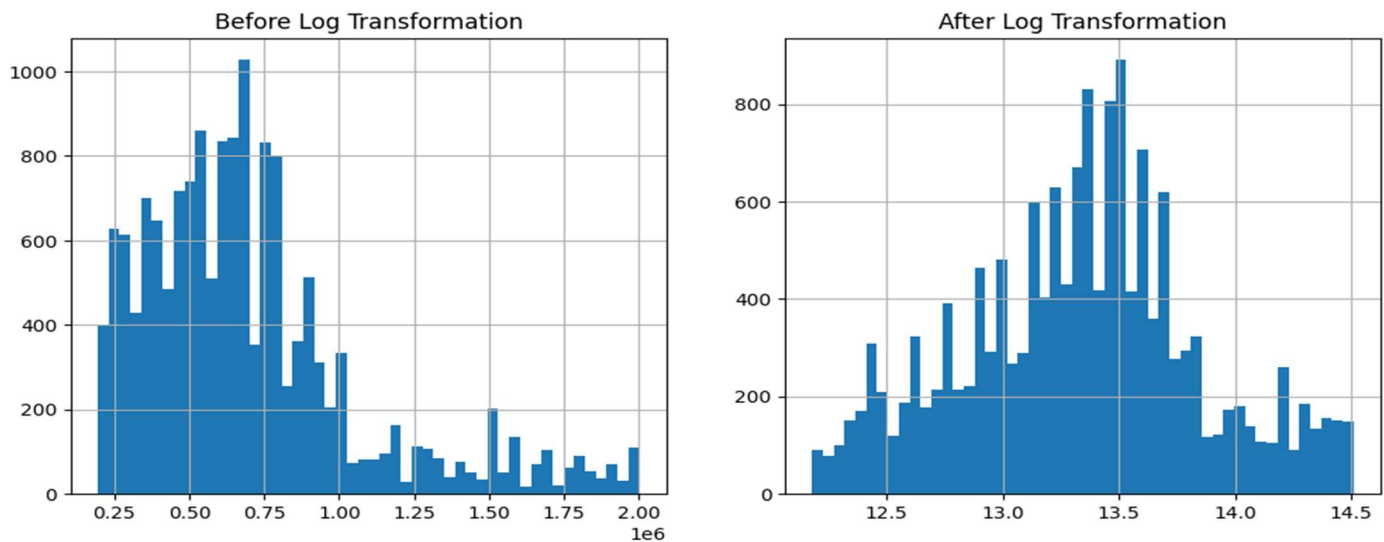
When talking about the prediction of the value of a property, there are certain factors that should be considered. That is why we have chosen the following four features as the main ones, the analysis of which will allow us to estimate the value of each house: the geographical location within the city, the number of bedrooms and bathrooms and the overall size of the house in square feet. These features are essential, as they show the location, the size and the characteristics of the property.

Instead of using the original city data, we used another data form, known as n_citi, because this is numerical data that the machine learning algorithm understands. Image_id was retained so that we could relate each property with its images. Street data, on the other hand, was omitted since this is a textual data which contained several values and we didn't use it for value prediction of the properties.

We also examined properties that appeared to be outliers. We determined if there were any abnormalities between number of bedrooms, bathrooms, property area, and prices of the properties. There were two particular properties that appeared highly questionable since they contained 36 and 13 bathrooms respectively while having just four bedrooms

and small property area. We therefore removed these two from our data set. Other properties that had very high values were retained only if they correlated to the expected value of a large residential property.

We took a closer look at how the original property prices were spread out, using something called the skewness statistic. What we found was that the prices were skewed to the right, with a skewness value of 1.349 - this means that there were some really high prices that were pulling the average up. To make the numbers more stable and easier to work with, we used a logarithmic transformation, which is like a math formula that helps reduce the impact of those really high prices. We added 1 to the price, and then took the log of that - this helped to even out the distribution and make it more normal. We used this transformed price variable when we were training and testing our model, because it made the whole process more reliable.



We were left with 15,472 property records after checking the data for quality and removing two records that didn't seem right. These records will be used to develop our model further.

Figure 2 Distribution of Property Prices Before and After Log Transformation.

7.3 Image Preprocessing and Deep Visual Feature Extraction

Images were loaded and resized to a fixed size of 224 x 224 pixels, with three RGB color channels. Images were converted to a numerical array and a batch dimension was added to generate an input tensor with the dimensions (1, 224, 224, 3).

Preprocessing was carried out using the `preprocess_input` function from the TensorFlow/Keras ResNet50 library. The preprocessing process carries out the transformation of the input according to the requirements of the pre-trained ResNet50 model, providing uniformity between property images and the model based on the ImageNet dataset.

A pre-trained ResNet50 convolutional neural network served as a fixed deep visual feature extractor. The model was pre-initialized using the weights of the ImageNet dataset, and the layers responsible for classifying images were excluded using the parameter `include_top=False`. Global average pooling was used by setting `pooling='avg'`, generating a 2048-dimensional fixed visual embedding of each property image. No further training of the layers was done for this pre-trained ResNet50 model.

The extracted visual embeddings represented high-level visual characteristics of the residential properties and were subsequently combined with the selected structured property features to construct the hybrid feature representation used in the valuation model.

7.4 Hybrid Feature Construction and XGBoost Model Development

To get a sense of how stable our model is and to have something to compare it to, we did something called five-fold cross-validation on our training data. We had already set aside 20% of our data as a test set, which we didn't use for this part. When we did the cross-validation, we made sure to shuffle the data and used a fixed random seed, which was 42. We looked at two kinds of models: XGBoost and Random Forest regression. Both of these models used the same features, like the number of cities, bedrooms, bathrooms, and square feet. We didn't use the hold-out test set for the cross-validation, we saved it for when we finally evaluated our hybrid model. This way, we could get a good idea of how well our model would work on data it had never seen before.

In analyzing the performance of the XGBoost model, one notes that it performed quite impressively, with a mean score of 0.6620 and standard deviation of 0.0118. On the other hand, Random Forest algorithm which served as a benchmark model performed slightly better, with a mean score of 0.6765 and standard deviation of 0.0101. The interesting aspect of this analysis is the fact that both models performed consistently across the data set, as seen through the low standard deviations of the scores. All in all, the Random Forest algorithm was found to perform marginally better than the XGBoost model in this case, based on their respective mean scores.

7.5 Model Interpretability Using SHAP

In order to explain the proposed approach of calculating property value prediction better and more comprehensible, Shapley Additive Explanations (SHAP) technique was utilized to determine the influence of different features of the property on the price prediction. As the visual features obtained using ResNet50 are complicated and not understandable by people, the explanation of property values was based on the properties' attributes.

To conduct SHAP analysis of property prices, we trained an additional XGBoost regression model based on several important features of properties. Such important features include a numerical feature of a city, the number of bedrooms, bathrooms, and property area (square feet). Our target variable is log-transformed property price. For this purpose, we divided our dataset into training and testing samples. 80% of the dataset were used as a training sample, while the rest 20% - as testing sample. Random seed 42 was fixed to achieve the same split every time. Therefore, we got 12,379 property observations in the training sample and 3,095 observations in the testing one.

The structured XGBoost model was configured with 300 estimators, a maximum tree depth of 6, a learning rate of 0.05, and a fixed random seed of 42. After training the model using the structured training data, SHAP values were calculated for the hold-out test set using the TreeExplainer method. TreeExplainer was selected because it is specifically designed to efficiently explain predictions generated by tree-based machine-learning models, including XGBoost.

For the SHAP results, the output provided a matrix consisting of 3095 rows and 4 columns, indicating that there were 3095 test samples and 4 structured features. The SHAP analysis was done in order to have an overview of the importance and influence of these structured features. Meanwhile, it also allowed the understanding of the situation from the local level, such as how these structured features influenced the pricing of single real estate objects.

The current model, although it allows explaining its decision based on certain factors such as the property size or location, it does not allow for understanding why it chose one particular picture over another. This is due to the fact that the current model uses a specific mathematical method for analysis of the pictures, called ResNet50, where the images are broken down into extremely small parts, that do not have any meaning on their own. While the current model has a great ability to analyze both pictures and numbers in order to determine the price of a certain house, it

cannot point to a specific feature of the picture and state, that this particular picture was liked by the model or disliked. The model could provide the explanation that the property was overvalued since it had a large backyard or it was undervalued due to the fact that it is located near a main street.

One future research direction would be to make the explanation component of the framework even more visual by incorporating image-based explanation approaches such as the Gradient-weighted Class Activation Mapping (Grad-CAM). Using such an approach would yield additional visual insights into the image areas that are most crucial in determining property prices. One may also be interested in exploring how the visual explanations generated are combined with the rest of the framework components to get a more comprehensive insight into property value estimation.

7.6 Similar Property Retrieval

To extend the functionality of Clarivision AI beyond residential property price prediction, a Similar Property Retrieval module was developed. The module identifies residential properties that are visually similar to a property image uploaded by the user and provides relevant comparison properties to support real-estate valuation and decision-making.

The system uses a database of property features that's stored in a file called property_features.pkl. This database has lots of information about each property, like the image identifier, city, street, price, number of bedrooms and bathrooms, and the size of the property. It also includes a special set of features that describe what the property looks like, which is called the deep visual feature representation. When someone uploads a new property image, the system uses the same process it used to build the database to extract the visual features from the image. It then changes the features into a simple format that can be used to compare the image to others in the database.

The location of a property is an important criterion while searching for the right property to buy or rent. It plays an important role in determining the value of a property. Since cities have different markets, different attributes and prices, it becomes difficult for the search to provide relevant results. Thus, in order to improve the relevancy of the results of the search process, only the properties within your selected city will be considered by the system. This would ensure that you get comparable and appropriate properties for comparison. Properties outside your selected city are discarded as they may be irrelevant.

The visual similarity of the image of an uploaded property and each of the candidate property images is computed using cosine similarity. Cosine similarity between an uploaded-property feature vector ($\mathbf{u}$) and a candidate-property feature vector ($\mathbf{v}$) is calculated as:

$$\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$$

The cosine similarity score measures the similarity between the directions of the two deep visual feature vectors. A higher similarity score indicates that the uploaded property image and the candidate property image have more similar visual feature representations, whereas a lower score indicates weaker visual similarity. The calculated similarity scores are stored with the corresponding property information.

The property candidates are sorted in decreasing order based on their cosine similarity scores. The duplicate property entries are filtered out by using the criteria of city, street name, price, property area, number of bedrooms, and number of bathrooms. The top five properties (Top-K with K=5) are returned as output by the system.

The returned properties are shown using the Streamlit app interface. Each property is displayed along with its image, price, visual similarity score, city, number of bedrooms, number of bathrooms, and property area (in square feet). The similarity score is transformed into a percentage and is shown as a percentage and a progress bar that intuitively represents the degree of similarity.

In the current version of the Similar Property Retrieval module, there is an interactive qualitative assessment of similarity between property images in the form of returning the top 5 visually similar properties from the chosen city. A quantitative evaluation based on metrics like Precision@5, Recall@5, or NDCG@5 was not performed due to the absence of labeled ground-truth data indicating visually similar properties.

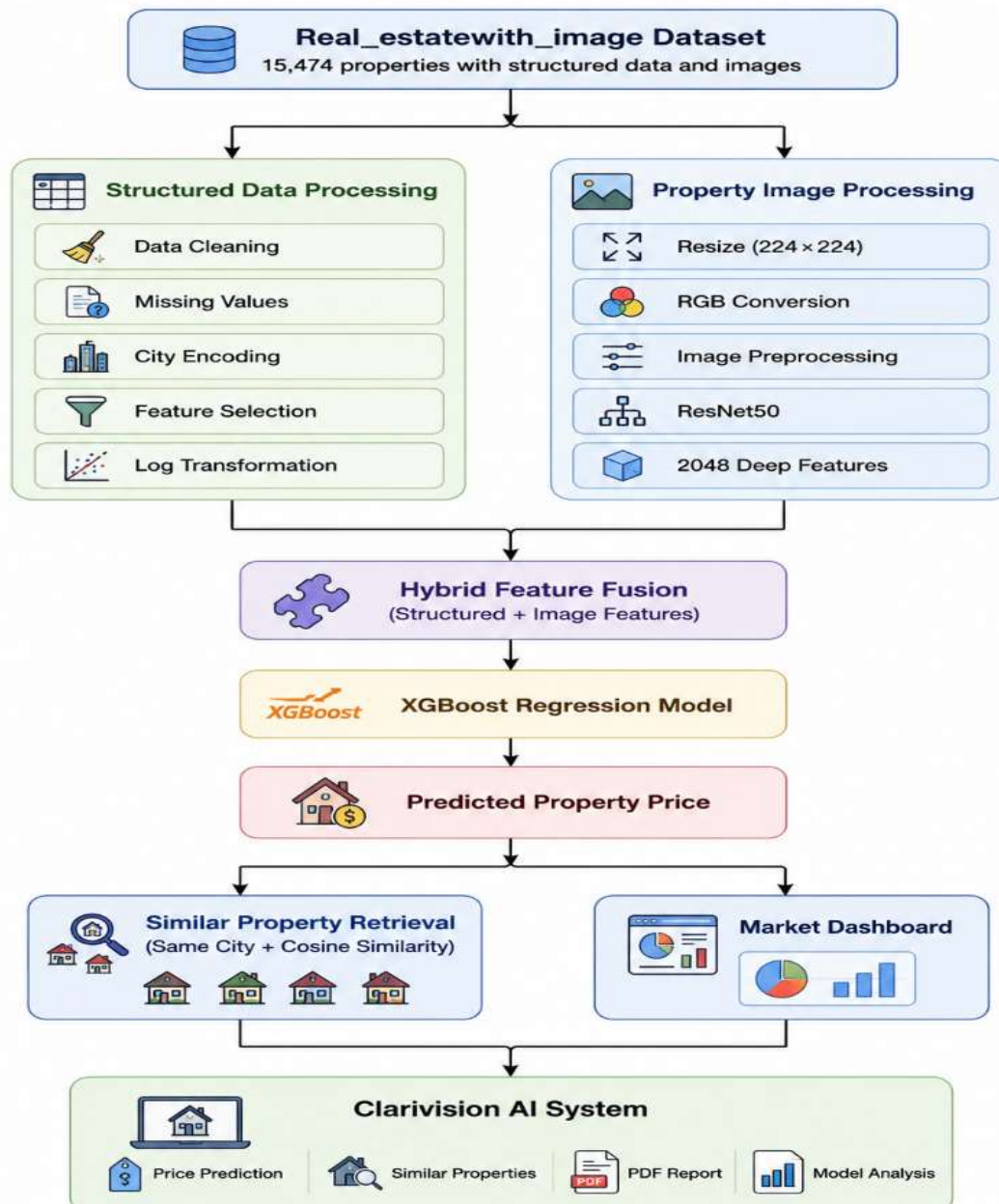


Figure 3: Overall Architecture of the Proposed Clarivision AI Framework

Figure 3 presents the architecture of the overall design of the proposed Clarivision AI framework. The framework starts from the real estate dataset in multimodal form that has structured property features and residential property images. The structured data goes through the process of data cleaning, data preprocessing, and feature engineering, and the images go through the processing by the ResNet50 convolutional neural network to extract the visual deep features. These two types of feature sets are combined in a unified hybrid form to be used as training data for the XGBoost regression model of residential property valuation. The resulting valuation is further used by the modules of Similar Property Retrieval and interactive visualization dashboard in the process of decision-making. Lastly, all the modules are combined in one Clarivision AI platform.

8. Results

The Clarivision AI framework was put to the test using a dataset that had 15,474 residential properties, each with its own set of details and pictures. This test looked at how well the new hybrid valuation model could predict things, and compared it to a different model that only used structured data. It also checked how stable the models were by trying them out multiple times, and used a technique called SHAP to see which property details were the most important. Additionally, it evaluated how well the Similar Property Retrieval module worked, which is a key part of the framework. The goal was to see if the new framework could provide more accurate valuations by combining different types of data.

Our dataset has been divided into two datasets - 80% training and 20% testing. This division was carried out based on a random seed value (which equals 42 in our case). The latter allowed us to obtain the same split of the dataset each time. Our test dataset was stored separately and used for evaluation only once all other tests were complete. As for our further results, it is important to discuss the process step by step - starting from the way the data was split and the way we have configured the hybrid model, its performance in terms of predicting property prices, comparisons with baseline model (structured data), and so on. Moreover, we examined the performance of our model under multiple testing and utilized such technique as SHAP for further interpretation of results. Finally, we also analyzed the similarity performance and Clarivision AI system's results.

8.1 Dataset Split and Model Setup

The experimental dataset included 15,474 real estate property entries after preprocessing and quality validation of data. The dataset was split into training data set 80% and testing data set 20% based on a fixed random seed value 42. Property price was transformed using natural logarithm $\log(1+\text{price})$ before applying machine learning models to make the target variable less positively skewed.

The proposed hybrid model combined property structured attributes with deep features of property photos obtained through ResNet50 architecture. XGBoost regression algorithm was selected as a base model for this task. XGBoost model parameters included 300 estimators, max tree depth equal to 6, learning rate 0.05 and random seed 42.

The baseline XGBoost regression model based only on structured property attributes was used to estimate the effect of deep visual features. Property structured attributes used for structured baseline were four: numerical city representation (n_city), number of bedrooms (bed), number of bathrooms (bath) and property area in square feet (sqft).

8.2 Performance Evaluation of the Proposed Hybrid Model

The predictive power of the proposed Clarivision AI Hybrid model was tested on the basis of MAE, MSE, RMSE, and the coefficient of determination (R^2). This combination of criteria provides additional information regarding the average prediction error, the magnitude of prediction error, and the percentage of explained variance by the model.

Table 1. Performance Evaluation of the Proposed Clarivision AI Hybrid Model

Metric	Value
Mean Absolute Error (MAE)	0.20985
Mean Squared Error (MSE)	0.08475
Root Mean Squared Error (RMSE)	0.29112
Coefficient of Determination (R^2)	0.67375

The hybrid model gave an MAE of 0.20985, MSE of 0.08475, RMSE of 0.29112, and (R^2) value of 0.67375. The MAE and RMSE figures have been calculated on the log-transformed property-price scale. The (R^2) value represents the extent to which the variance in the log transformed property price values was explained by the model.

It looks like a very promising figure, but let's not be too overconfident about it. With an R^2 value of 0.67375, it can be said that the results obtained were moderate to promising. But the performance cannot be termed as excellent as the figure implies that 32.6% of variance was still unexplained.

8.3 Comparison with the Structured-Data Baseline Model

In order to assess the importance of the features from deep vision generated by ResNet50, the Clarivision AI Hybrid model was compared to an XGBoost model, which utilized only the structured property characteristics.

Table 2. Performance Comparison Between the Structured XGBoost Model and the Proposed Clarivision AI Hybrid Model

Model	MAE	RMSE	(R^2)
XGBoost (Structured Only)	0.22924	0.31369	0.62749
Clarivision AI (Hybrid)	0.20985	0.29112	0.67375

From this comparison, it is clear that the hybrid model performs better than the structured data only XGBoost model in all the metrics used. For instance, the model reduces the MAE from 0.22924 to 0.20985 and reduces the RMSE from 0.31369 to 0.29112. Moreover, the (R^2) metric rises from 0.62749 to 0.67375.

These findings suggest that the inclusion of the deep features learned by ResNet50 together with the structured property data helps in enhancing the property valuation performance by providing more information. This is evident from the fact that the (R^2) metric increases by 0.04626.

8.4 Cross-Validation and Structured-Model Comparison

To check how well the models for structured data work, a test called five-fold cross-validation was done. This was only done on the part of the data used for training, after 20% of it was set aside for testing. When doing the cross-validation,

the parts of the data were mixed up and a fixed random seed, 42, was used to make sure the results were the same each time.

We compared how well two models worked: one that used structured data with XGBoost and another that used Random Forest. To do this, we looked at the average and spread of the scores from a test called R^2 , which we ran five times with different sets of data.

Table 3. Five-Fold Cross-Validation Results for Structured-Data Models

Model	Mean (R^2)	Standard Deviation
XGBoost	0.66200	0.01175
Random Forest	0.67651	0.01011

The structured XGBoost model achieved a mean (R^2) value of 0.66200 with a standard deviation of 0.01175. The Random Forest model achieved a slightly higher mean (R^2) value of 0.67651 with a standard deviation of 0.01011.

The relatively low standard deviations indicate that both structured-data models demonstrated stable performance across the five validation folds. Random Forest achieved slightly higher average performance in this structured-data cross-validation experiment. However, these results should be interpreted separately from the final evaluation of the proposed hybrid model because the cross-validation experiment used only the four structured property attributes and did not include the ResNet50 visual features.

8.5 SHAP-Based Feature Importance Analysis

SHapley Additive exPlanations (SHAP) were used to improve the transparency of the structured valuation component and to examine the contributions of the structured property attributes to the model predictions. A separate XGBoost model was trained using the four structured features: numerical city representation (n_citi), number of bedrooms (bed), number of bathrooms (bath), and property area in square feet (sqft).

The resulting SHAP value matrix had a shape of ((3095, 4)), corresponding to 3,095 test observations and four structured property attributes. The SHAP analysis indicated that property area (sqft) was the most influential structured feature, followed by the numerical city representation (n_citi), number of bathrooms (bath), and number of bedrooms (bed).

The relative feature importance values were approximately 50.16% for property area, 28.56% for city, 14.28% for the number of bathrooms, and 7.00% for the number of bedrooms.

FEATURE IMPORTANCE OF STRUCTURED PROPERTY ATTRIBUTES

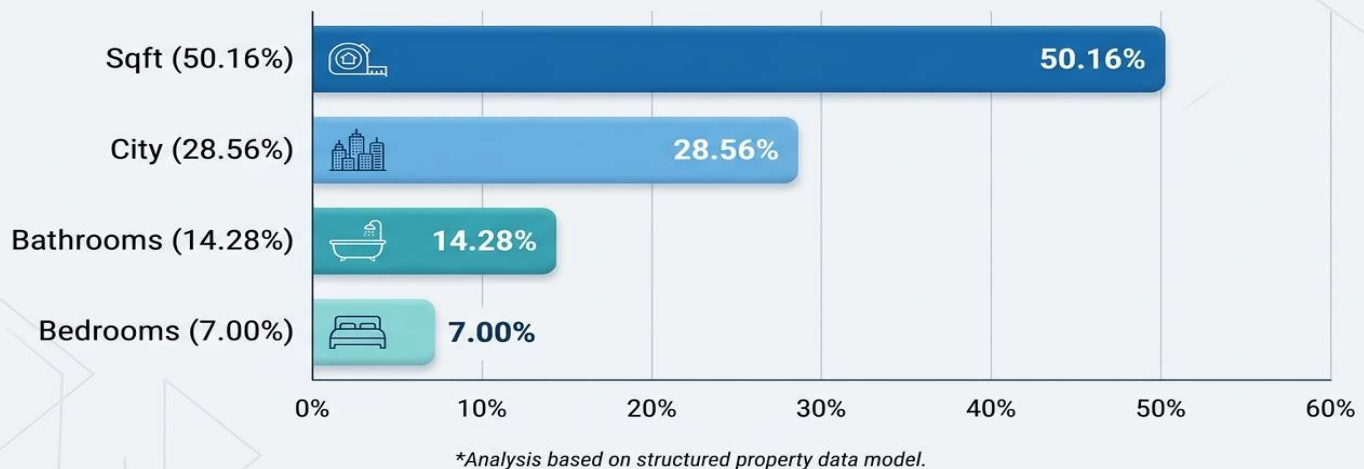


Figure 4. Relative Importance of Structured Property Features

- 18 The results demonstrate that for the structured XGBoost model, the mean score is equal to 0.66200 with the standard deviation of 0.01175. At the same time, the Random Forest model demonstrated a slightly higher result – the mean score was equal to 0.67651, while the standard deviation equaled 0.01011. Therefore, we can conclude that the Random Forest model is not only more accurate but also more consistent.
- 14 The results prove that both models, based on the structured data, performed well during all experiments. The Random Forest model performed a bit better than the other one in this particular experiment. However, it is crucial to distinguish the results of this experiment from the final results concerning the evaluation of our hybrid model since this experiment included just four attributes of the structured data but no visual attributes from ResNet50 model.

8.6 Similar Property Retrieval

The Similar Property Retrieval system was developed for the purpose of decision-making regarding the pricing of residential properties. In order to accomplish its purpose, the system uses cosine similarity algorithm, which measures visual similarities between certain images. These are extracted through ResNet50 model – pre-trained artificial intelligence system that specializes in analyzing images. Each property picture is characterized by a vector of numbers, 2,048 to be precise. Thus, the system measures similarities based on these vectors of numbers and allows for finding visually similar property pictures. This helps to analyze residential properties based on their visual characteristics, which can be highly useful for establishing their price.

For the purpose of ensuring that the search results will be more location-specific to the users, we narrowed down the search to only include the properties located in the same city as the user had specified. This is important because despite the fact that two properties could be identical, they may vary in terms of their market price because of the fact that they would be in two different cities, which would mean two different markets for them. Thus, this helps us provide more relevant data for the users when they are trying to estimate the price of a residential property.

This system selected the property images that matched our search request the most and displayed the top five results. For each of the found properties, it is possible to see a preview image, how much this property matches the request, its price, the city it is located in, the number of rooms and bathrooms, and its area. We tried this system, and it performed well selecting and ranking the properties that are similar to our request. In case the property we searched for was

included in the system, it will appear with the perfect match score. Other properties will have lower match scores based on their similarity. For measuring the similarity between properties, the system uses so-called "deep visual representations", and performs rather well at it.

The similarity between two property pictures is measured by a certain score. It indicates to which degree the features of these pictures are similar to each other. The higher the score is, the more similarities there are. However, it does not indicate the probability of similarity or equal market value of the properties. It only compares the features seen by a certain program – ResNet50.

As opposed to the classical Automated Valuation Models that offer a figure of the estimated price of a property, Clarivision AI integrates the valuation of the property with the photographs of the similar properties. Thus, the user would be able to compare his/her object of interest with those objects that have similar visual appearance and thus gather more data regarding the value of the property.

The present study focused on process evaluation and quality of results as we did not have any ground truth to be able to assess similarities among objects. Thus, we could not use any typical measures of performance like Precision, Recall, and NDCG. However, in the future, we need to engage some experts or users to assess the level of similarity between different objects and thus use such results to evaluate the effectiveness of the system. We require some kind of labeled data to test the effectiveness of our system.

Here's how the process works for finding similar properties using Clarivision AI. First, an image of a property is uploaded and then processed using a special tool called ResNet50. This tool helps create a unique set of features for the image, kind of like a fingerprint, but for pictures. This fingerprint is made up of 2,048 different points that describe what the property looks like. Next, the system compares this fingerprint to a big database of other property images to find the ones that look the most similar. It does this by calculating something called cosine similarity, which is a way of measuring how alike two things are. To make sure the results are relevant, the system only looks at properties in the same city as the one in the uploaded image. After comparing all the images, the system ranks them based on how similar they are to the original property. It then shows the top 5 properties that look the most similar. This helps users find properties that are similar in style or design to the one they're interested in, which can be really useful for things like real estate or design inspiration.

The Similar Property Retrieval module was created in order to aid in the visual comparison of properties whose visual attributes were similar to that of the query property. Cosine similarity was used in calculating the similarity using the feature vector extracted from the pre-trained ResNet50 convolutional neural network. In order to enhance the context of the retrieved data, the search was narrowed down to only properties that are in the same city. Up to five visually similar properties were supposed to be retrieved by the module. However, in some instances, fewer than five similar properties could be retrieved due to the small number of properties that qualified for search within the selected city after filtering out duplicate properties.

15

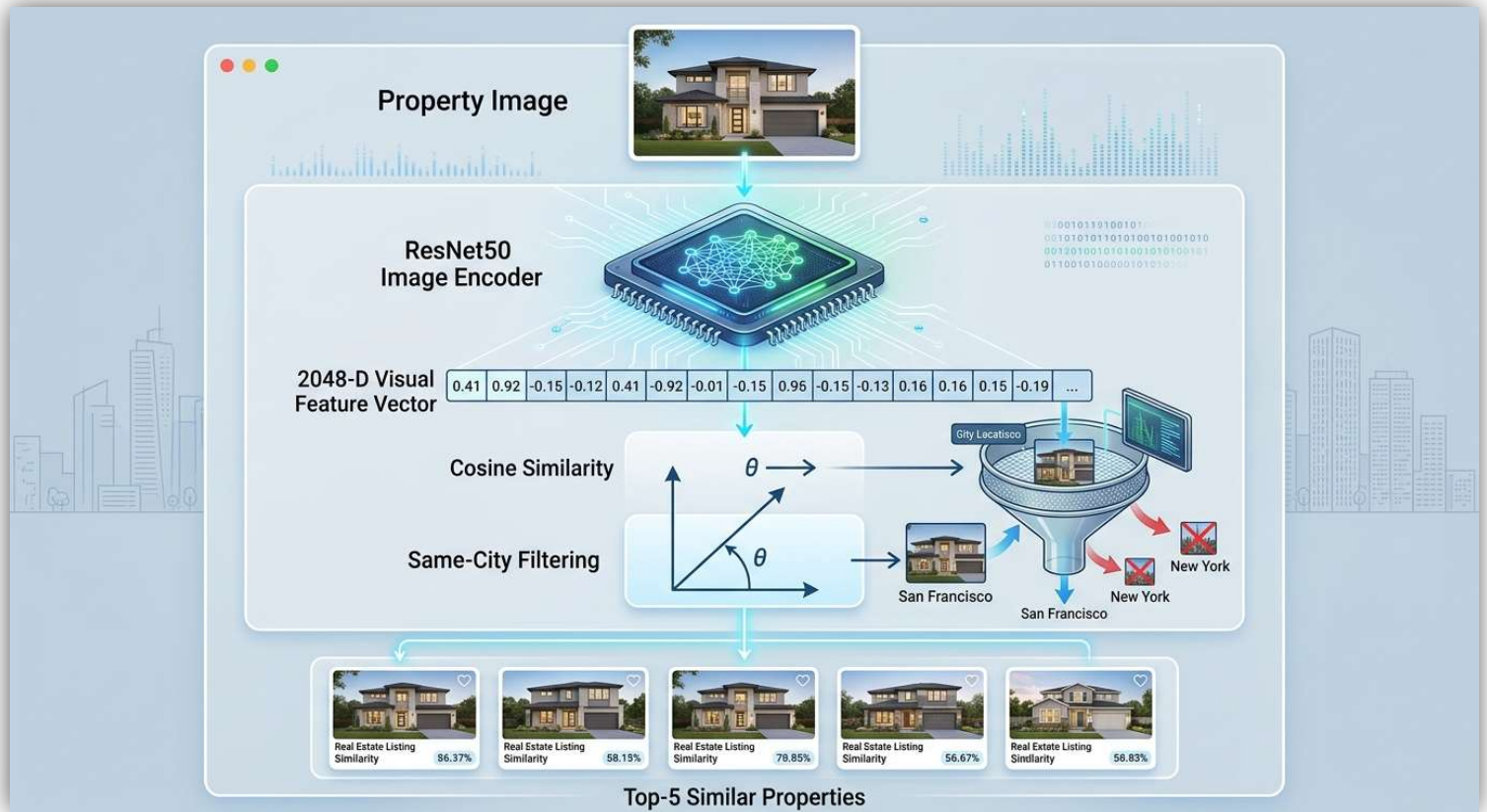


Figure 5. Workflow of the Similar Property Retrieval Module in Clarivision AI. Source: Developed by the author.

Figure 6 presents an example of the Similar Property Retrieval module. The retrieved properties are ranked according to their cosine similarity scores, with the highest-ranked properties displayed first. For readability, the retrieval results are presented in two parts: Figure 5(a) shows the first three retrieved properties, while Figure 5(b) shows the fourth and fifth retrieved properties.

Property Image



■ Top Similar Properties

Similar Property #1



Price: \$555,900

Similarity: 81.37%

City: Anaheim, CA

Bedrooms: 3

Bathrooms: 2.0

Square Feet: 1428

Similar Property #2



Price: \$739,000

Similarity: 80.39%

City: Anaheim, CA

Bedrooms: 5

Bathrooms: 3.0

Square Feet: 2406

Figure 6 (A).

Similar Property #3

	Price: \$669,000 Similarity: 78.98% City: Anaheim, CA Bedrooms: 3 Bathrooms: 2.0 Square Feet: 1603
---	---

Similar Property #4

	Price: \$567,000 Similarity: 78.50% City: Anaheim, CA Bedrooms: 3 Bathrooms: 2.0 Square Feet: 1358
---	---

Similar Property #5

	Price: \$939,000 Similarity: 78.00% City: Anaheim, CA Bedrooms: 5 Bathrooms: 3.0 Square Feet: 3001
---	---

Figure 6 (B).

Figure 6: Top five visually similar properties retrieved by Clarivision AI using cosine similarity: (a) the first three retrieved properties and (b) the fourth and fifth retrieved properties.

For testing the Similar Property Retrieval algorithm, a property query image taken from Anaheim, CA was used. It is observed that the top five visually similar residential properties were retrieved by filtering out the search based on the location of properties in the same city. It is observed that cosine similarity measures for these properties range between 78.00% to 81.37%.

The top-ranking property had a similarity score of 81.37%, followed by 80.39%, 78.98%, 78.50%, and 78.00% for the remaining retrieved properties. Even though the retrieved properties varied in terms of their prices, number of bedrooms, bathrooms, and sizes, they have visual features in common, captured by the ResNet50 feature representation. This illustrates that the retrieval process is mainly driven by visual similarity and not price similarity or similarity in any other structured feature.

Filtering the results to a single city ensures that the geographic relevance is preserved because residential property values and characteristics may differ substantially depending on the cities. Hence, the city filter limits the chance of

retrieving visually similar properties but from a geographically different location and which cannot be used as local reference property.

Cosine similarity scores should be treated as measures of similarity between the ResNet50 visual feature representations and not percentage similarity in physical/economic features. Higher scores mean that the visual representations of two property images are closer to each other in the learned feature space. The retrieved properties serve as additional visual references which could complement the numerical valuation performed by Clarivision AI.

8.7 Property Information Interface

The Property Information Interface is the place where users can enter the required structured data related to the property, including city, number of bedrooms, number of bathrooms, size of the property (square feet), and uploading the image of the residential property. After entering the required data, users can click on the “Predict Property Price” button to find the predicted price of the property using the proposed Clarivision AI framework.

The proposed interface serves as the primary interface to access the system and is convenient for performing the process of residential property valuations.

The Clarivision AI framework design contains all the elements of the proposed framework in a decision support environment that allows users to perform the following actions: perform residential property valuations, view the SHAP explanations, retrieve visually similar properties, access the visualization dashboards, segment the market of residential property, and generate PDF reports.

In addition to that, the combination of all those abilities into one system improves the ease of use by delivering an effective, clear, and informative way of making decisions regarding real estate. As opposed to the common practice when AVMs only deliver numeric values, Clarivision AI delivers an interactive system combining valuation prediction, visual analysis, and intelligent decision-making service.

Figure 7. Main interface of the proposed Clarivision AI decision-support platform.

The user interface used to enter structured property attributes and upload a residential property image before performing property valuation using the proposed Clarivision AI framework.

8.8 Property Information and Price Prediction Interface

The first step in the proposed Clarivision AI-based method is to gather both the structured attributes of the property as well as the image of the residential property using a user interface. The users provide all the necessary details regarding the city, number of bedrooms, number of bathrooms, size of the property in square feet, and the property image. After gathering the necessary information, the system then evaluates the input data and estimates the value of the residential property based on the proposed hybrid model.

Figure 8: Property Information and Price Prediction Interface

Figure 8: Property Information and Property Value Prediction Interface. Users can use the interface to enter information about properties in the form of city name, number of bedrooms, bathrooms, total square feet of the property, and a picture of the residential property. Clicking on Predict Property Value will result in prediction of the residential property value using Clarivision AI model.

8.9 Clarivision AI Prediction Dashboard

The Clarivision AI Dashboard provides an interactive visualization of the valuation results generated by the proposed framework. The output also shows the estimated market segment, confidence level of the prediction, details of the AI model used, and also the comparison of the assessed asset to the average market. This representation makes it easier for users to comprehend the results of the valuation process.

Prediction Confidence

94%

Market Segment

Premium

AI Model

Hybrid AI



Property Characteristics Compared with Market Average

	Feature	Your Property	Market Average
0	Bedrooms	6	3.5064
1	Bathrooms	9	2.4533
2	Square Feet	1250	2173.9132

Figure 9: Clarivision AI Prediction Dashboard.

As can be seen in Figure 9 below, the Clarivision AI prediction dashboard shows the predicted price of the residential property together with the prediction confidence rating, market category, information about the AI model used, and the comparison between the property to be valued and the average market characteristics. Through this dashboard, the user can visualize the value of the residential property.

8.10 Clarivision AI Report

The proposed Clarivision AI architecture will create an automatic PDF Valuation Report after the process of prediction is done. The report will summarize the results of the valuation process in an organized manner that enables users to be able to save, print, and even share their valuation result.

CLARIVISION SYSTEM

Artificial Intelligence Property Valuation Report

Report ID: CV-20260723105111
Generated: 23/07/2026 10:51
AI Engine: Hybrid CLARIVISION SYSTEM AI (ResNet50 + XGBoost)

Property Information

City	Alpine, CA
Bedrooms	5
Bathrooms	3
Square Feet	980 Sqft

Artificial Intelligence Prediction

Estimated Property Value	\$459,788
Prediction Confidence	94%
Market Category	Mid Range

Artificial Intelligence Decision

The Hybrid Artificial Intelligence model predicts that this property belongs to the "Mid-Range Market Segment". The prediction is supported by the combination of image analysis and structured property data. Main Factors: ✓ Property Area ✓ City Location ✓ Bedrooms and Bathrooms ✓ Image Features extracted using ResNet50 Over Confidence: Very Good.

Figure 10: Clarivision AI Report.

From the experiment findings, it is evident that incorporating structured property attributes and deep visual representations can enhance residential property price prediction as compared to structured data only, as used in this study. The proposed Clarivision AI hybrid model gave better prediction errors and higher (R^2) as compared to the structured data XGBoost model. This implies that the visual features captured from images through the ResNet50 were more complementary since they added value to structured numeric data.

26 In this case, the proposed Clarivision AI hybrid model generated an (R^2) of 0.67375, meaning that about 67.4% of the variation of the transformed property prices was explained by the model. In this case, the generated (R^2) is promising and moderate rather than complete or high. About 32.6% of the variation was unexplained, which could be due to various reasons that are not captured in the current dataset like age, quality of construction, location, latitude and longitude, market conditions, and other property-related attributes.

79 The superior performance of the hybrid model over the structured-data-only XGBoost model lends weight to the possibility that multimodal learning can be valuable for residential property valuation. The R^2 score for the structured data only model was 0.62749, while for the proposed hybrid model, it was 0.67375. The improved score by approximately 0.04626, in addition to the decreases in MAE and RMSE scores, implies that the inclusion of both structured data and visual data improved the performance of the valuation process. This finding aligns with those of previous studies which have shown that the inclusion of structured property data and image data improves the accuracy of residential property valuations. These findings must nonetheless be considered within the context of the data set and experiments carried out in this research study.

65 The data-quality analysis showed that the dataset contained no missing values and no duplicate records. The property-price variable exhibited positive skewness, with a skewness value of approximately 1.35. Therefore, the target price was transformed using ($\log(1+\text{price})$) before model training. This transformation reduced the influence of the right-skewed price distribution and provided a more suitable target representation for the regression models.

5 The SHAP results yielded further explanations regarding the influence of structured property attributes. The property area (sqft) was found to be the most important structured attribute, while other important attributes include numerical city representation (n_citi), number of bathrooms (bath), and number of bedrooms (bed). The influence of property area and location variables is expected given the overall findings in the literature on residential property evaluation which considers property size and location as crucial factors influencing the value of the property. However, the SHAP explanation currently developed can explain only the structured attributes and not the individual visual attributes generated from ResNet50.

The five-fold cross-validation experiment revealed relatively little variability of the (R^2) metric scores in case of structured-data models. In particular, the mean (R^2) score of the structured XGBoost was found to be 0.66200 with the standard deviation of 0.01175, whereas that of the Random Forest model was equal to 0.67651 with the standard deviation of 0.01011. This suggests relatively high stability of the model's performance across the validation folds. It is worth mentioning that the cross-validation experiment has been carried out based on the structured property data only, and thus cannot be considered as the evidence of robustness of the hybrid model under consideration. Furthermore, the obtained results do not provide sufficient grounds to argue that the proposed hybrid model does not suffer from overfitting problem.

Apart from property price prediction, Clarivision AI also offers SHAP explanations, Similar Property Retrieval, visualization capabilities, market-oriented decision support outputs, and reporting in one platform. The Similar Property Retrieval system finds the Top-5 visually similar properties in the city chosen by the user based on cosine similarity of the deep visual features. This capability makes the system go beyond mere price prediction since it also gives the user some properties to compare. Nonetheless, retrieval was only assessed in terms of function and quality as there were no similarity ground truths for the data set used. Thus, the quantifiable measures of Precision@5, Recall@5, and NDCG@5 were not computed.

Combining the prediction of the valuation, explainability of structured features, similarity search via image retrieval, visualization, and reporting illustrates the real-world value of Clarivision AI as an Explainable Decision-Support Platform. However, the suggested framework should not be viewed as a substitute for professional property appraisal. The results are based on the properties of the dataset used, model setting and attributes of the property.

In general, the results presented above prove that multimodal learning and Explainable Artificial Intelligence represent a promising approach for enhancing property valuation systems. In particular, the suggested framework performed better than the structured-data-only approach in terms of predictive accuracy and provided additional functionality for decision-making. However, further studies with a larger variety of datasets, additional attributes of property, image-level explanation techniques, and quantitative retrieval assessment should be conducted before further generalizations can be made.

10. Conclusion

This paper presented the development of Clarivision AI, an interpretable hybrid model for residential property valuation that combined structured property data along with visual features obtained through ResNet50. The hybrid features have been used with the XGBoost regression algorithm while SHAP is used for the explanation of the contributions of the structured property features.

23 The proposed framework resulted in an MAE of 0.20985, RMSE of 0.29112 and (R^2) of 0.67375 on the transformed property-price scale. This indicates decent and promising predictive performance and shows that adding visual features increases the performance compared to the structured-data-only approach.

3 Apart from this, the Clarivision AI framework also includes Similar Property Retrieval, interactive visualization and automated reporting modules that can be helpful in real estate decision making. However, the framework performance depends on the quality of the dataset and its geographical diversity, availability of property features and the application of log-transformed prices. Hence, the framework must be seen as the decision support tool and not a replacement for professional property appraising.

Future work will involve evaluation of the framework based on larger and more geographically diverse datasets, additional property features, external validation, visual explanations and retrieval metrics.

11. Recommendations

Based on the findings and limitations of this study, the following recommendations are proposed:

- 2 1. For future research, structured attributes like property age, property type, environmental characteristics, socio-economic information, accessibility, etc., need to be added to enhance the valuation performance of residential property valuation.
2. In future versions of the framework, multiple images of the property, which can be indoor, outdoor, aerial, and street view images, will be used to offer a better visual representation of the properties.
3. The proposed framework needs to be tested on residential property datasets from different geographic regions and market situations to evaluate its robustness and regional validity.
4. To get explanations of visual and structured features of the property, other approaches to explainability like LIME, Grad-CAM, and Integrated Gradients need to be used.
- 85 5. Similar Property Retrieval module needs to be tested using quantitative retrieval metrics like Precision@K, Recall@K, and NDCG@K, supported by the validation of experts or users. In addition, the process of retrieval can be improved using visual similarity along with other factors like location, property type, construction year, market trends, and property condition.
6. Future developments can include live data from the market and online listings to give up-to-date information about the value of the property.
7. The future developments that can take place in Clarivision AI can include classification of the properties according to their market segmentations into low-value, medium-value, and high-value property classes.
8. The suggested framework can be developed into a cloud-based decision support system through web and mobile applications for real-estate professionals, financial institutions, investors, and other authorized users.
9. With further research in the future, possible biases and issues related to fairness in property valuations by the automated system will have to be investigated. The performance of the model will have to be investigated in various cities, neighborhoods, property classes, and price ranges to determine whether the prediction errors are targeting any particular places or markets.

10. The use of the automated valuation system in an ethical manner is important to consider when developing the framework. The framework will have to accommodate the use of data in an ethical manner, privacy, security of property information, and governance when using the automated prediction results in making financial decisions.
11. Transparent means of providing the users with information about the automated property price prediction system are essential. The system will have to provide users with the limitations of the model, uncertainty of predictions, main drivers of the predicted value, and proper human professional involvement in making the final decision.

References

- Balila, A. E., & Shabri, A. B. (2024). Comparative analysis of machine learning algorithms for predicting Dubai property prices. *Frontiers in Applied Mathematics and Statistics*, 10, Article 1327376. <https://doi.org/10.3389/fams.2024.1327376>
- Deng, L. (2025). Real estate valuation with multi-source image fusion and enhanced machine learning pipeline. *PLOS ONE*, 20(5), Article e0323351. <https://doi.org/10.1371/journal.pone.0323351>
- Despotovic, M., Koch, D., Stumpe, E., Brunauer, W. A., & Zeppelzauer, M. (2023). Leveraging supplementary modalities in automated real estate valuation using comparative judgments and deep learning. *Journal of European Real Estate Research*, 16(2), 200–219. <https://doi.org/10.1108/JERER-11-2022-0036>
- Du, W.-W., Wang, W.-Y., & Peng, W.-C. (2023). DoRA: Domain-based self-supervised learning framework for low-resource real estate appraisal. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM '23)* (pp. 1–7). Association for Computing Machinery. <https://doi.org/10.1145/3583780.3615470>
- Geerts, M., Reusens, M., Baesens, B., vanden Broucke, S., & De Weerd, J. (2025). *On the performance of LLMs for real estate appraisal* (arXiv:2506.11812). arXiv. <https://doi.org/10.48550/arXiv.2506.11812>
- Hasan, M. H., Jahan, M. A., Ali, M. E., Li, Y.-F., & Sellis, T. (2024). *A multi-modal deep learning-based approach for house price prediction* (arXiv:2409.05335). arXiv. <https://doi.org/10.48550/arXiv.2409.05335>
- Huang, C., Liang, B., Li, Z., & Chen, F. (2025). Multimodal machine learning for real estate appraisal: A comprehensive survey. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 345–361). Springer. https://doi.org/10.1007/978-981-96-8183-9_26
- Kee, T., & Ho, W. K. O. (2025). eXplainable Machine Learning for Real Estate: XGBoost and Shapley Values in Price Prediction. *Civil Engineering Journal*, 11(5), 2116–2133. <http://dx.doi.org/10.28991/CEJ-2025-011-05-022>
- Maselli, G., & Nesticò, A. (2025). Machine learning algorithms and explainable artificial intelligence for property valuation. *Land*, 14(1), Article 132. <https://doi.org/10.3390/land14010132>
- Mora-Garcia, R.-T., Cespedes-Lopez, M.-F., & Perez-Sanchez, V. R. (2022). Housing price prediction using machine learning algorithms in COVID-19 times. *Land*, 11(11), Article 2100. <https://doi.org/10.3390/land11112100>
- Rodríguez-Serrano, J. A. (2024). Prototype-based learning for real estate valuation: A machine learning model that explains prices. *Annals of Operations Research*, 344(1), 287–311. <https://doi.org/10.1007/s10479-024-06273-1>
- ted8080. (n.d.). *House prices and images – SoCal* [Data set]. Kaggle. <https://www.kaggle.com/datasets/ted8080/house-prices-and-images-socal>
- Wang, Z., Wang, Y., Wu, S., & Du, Z. (2022). House price valuation model based on geographically neural network weighted regression: The case study of Shenzhen, China. *ISPRS International Journal of Geo-Information*, 11(8), Article 450. <https://doi.org/10.3390/ijgi11080450>
- Yazdani, M., & Raissi, M. (2023). *Real estate property valuation using self-supervised vision transformers* (arXiv:2302.00117). arXiv. <https://doi.org/10.48550/arXiv.2302.00117>