World Journal of Advanced Engineering Technology and Sciences

eISSN: 2582-8266

Cross Ref DOI: 10.30574/wjaets

Journal homepage: https://wjaets.com/

(Research Article)

Check for updates

House price prediction with Convolutional Neural Network (CNN)

Mohit Jain [1,*] and Arjun Srihari [2]

[1] *University of Illinois at Urbana-Campaign, United States of America.*

[2] *M.S. Ramaiah Institute of Technology, India.*

World Journal of Advanced Engineering Technology and Sciences, 2023, 08(01), 405-415

Publication history: Received on 04 January 2023; revised on 17 February 2023; accepted on 20 February 2023

Article DOI: https://doi.org/10.30574/wjaets.2023.8.1.0048

Abstract

In this paper, we examine the applicability of Convolutional Neural Networks (CNNs) for predicting the cost of houses with an inclusion of visual and non-visual elements. Traditional end-to-end patterns of supervised machine learning usually incorporate a finite set of numerical or categorical features, ignoring spatial and aesthetic information. CNNs, which can parse complex patterns in unstructured data, represent a way to improve predictive accuracy at the price of including property images with numerical and geographical coordinates. This variety of inputs is integrated within the proposed model to reflect the visual and contextual aspects of real estate valuation. The training and testing set used property images, georeferenced information, and typical numerical features. Incorporation of architectural features, including design, quality and neighbourhood aesthetics, showed a higher variance than the baseline machine learning model. Analysis of the outcomes obtained in experiments showed that CNN-based models performed better than the benchmark models, such as regression, gradient boosting, and random forest regression, in identifying the subtle interconnection between attractiveness and property values. The results further support the ability of CNNs to use multimodal information to solve diverse prediction problems within real estate. This research emphasizes the potential value of disseminating multiple data sources and merging relatively sophisticated neural structures into real estate applications, indicating the future viability of such paradigms in broader property market technologies. Subsequent studies can look at topics such as increasing the capacity of model structures and expanding the dataset to incorporate time dependency and other regions.

Keywords: Convolutional Neural Networks; House Price Prediction; Visual Features; Multimodal Data Integration; Real Estate Valuation; Spatial Analysis; Machine Learning

1. Introduction

Real estate price prediction, in particular, has a critical role in determining its business and future impact on buyers, sellers, investors and policymakers. Several factors determine market value: location, size, facilities, and the nature of the neighbourhood. Over the last decades, the increase in real estate turnover and the large amount of available data in this sphere contributed to the emergence of predictive modelling as an important tool facilitating decision-making. The right price forecasting optimizes the property investment process, minimizes identified financial risks, and oversees urban development projects. Yet the entanglement of factors like economic cycles, local conditions and a property's objective/ subjective characteristics presents major challenges for a model follower approach to housing market forecasting.

1.1. Traditional Approaches

In the past, house price forecasting used statistical models, including linear Regression and hedonic models, which endeavour to work with numerical and categoric data. While useful for determining basic relationships between

* Corresponding author: Mohit Jain

variables, these approaches are limited by their inability to handle data in formats other than vector space – for example, images – and because they inherently assume certain structures of variable interactions. In the same way, decision trees, random forests, and gradient boosting machines' families are an enhanced solution compared to traditional statistical models; however, they are also reserved in the ability to analyze multiple patterns in multiple-modal data. He noted that these methods fail to capture the more important aspects, such as the property's physical appearance, such as the architectural design of the structure or its appearance.

1.2. Motivation for CNN

With the help of CNN, image analysis has become much easier as it allows models to capture complex data and eventually get meaningful features. CNN has been used in areas like medical image analysis, self-driving cars, and even biometrics, and all these areas require or process high dimensional data. For real estate, property images contain information related to the state of the house, the design, and the house's surroundings, which defines buyers' preferences and market prices. This unstructured data is not exploited through conventional approaches, which leads to a gap in predictive precision. CNNs present a rich approach to supplement visual information with numerical and geographical data to explain drivers for property valuations comprehensively.

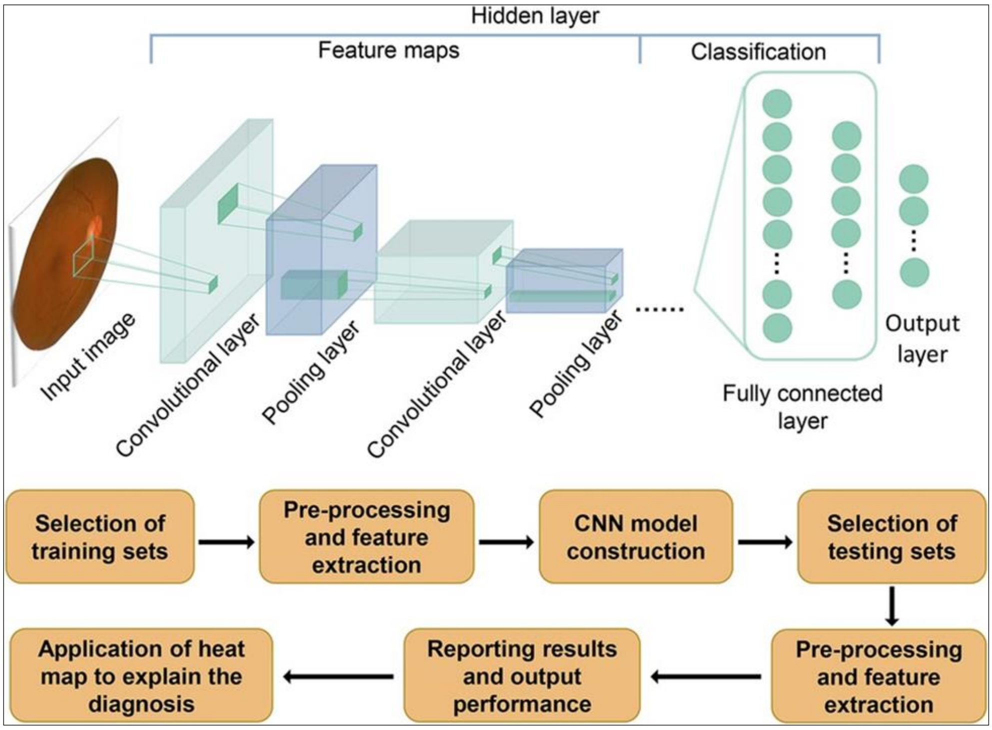

Figure 1 A Conceptual Diagram Showing The Flow Of CNN-Based House Price Prediction

Objective

This paper seeks to propose and assess a CNN-based House price prediction model that includes visual and non-visual inputs. The current study proposes a new model for enriching property images with structured information like area and place coordinates, foreclosure data, and market trends to improve its ability to estimate accurately compared to more traditional machine learning models. The approach then analyzes the consequences of CNNs on the model's predictive performance and discusses the further extended applicability of the approach in real estate market analysis. By so doing, this study aims to fill a gap where unstructured data and structured data differ in property valuation to ensure the property business makes informed decisions.

2. Literature Review

2.1. Existing Models

This research problem has been considered a very difficult one in house analysis and has long received a lot of focus from researchers and practitioners in real estate analysis. In the past, different Statistical and machine-learning models have been used to solve this problem. Linear Regression was traditionally used due to a concept's inherent ease of explanation. This technique is based on the assumption that predictors are linearly correlated with the target variable. They include property size, location, age of the house, etc. However, since this approach is linear, its generalization ability to consider complicated feature interconnection is confined.

Different decision tree methods, such as CART (Classification and Regression Trees), represent a more flexible modelling framework for expressing non-linear dependencies. These models are easy to interpret and do not require assumptions about the interconnectivity of variables or factors. Therefore, starting with Random Forests and extending to Gradient Boosting Machines (GBMs), predictive accuracy was increased using multiple decision trees to reduce the overfitting problem. Such approaches have found their applicability in real estate analysis and modelling due to their inherent strength in dealing with complex data forms. However, as strong as those traditional machine learning models are, they fail at handling unstructured data, such as images, which contain useful context for properties.

2.2. Role of Deep Learning

Deep learning is likely to provide rich opportunities for development in diverse fields of application as the technique has stepped forward as a solution of choice due to its efficiency in handling complex data. As a class of artificial neural networks, deep learning models, especially CNNs, have been introduced for house price prediction since they can input and analyze visual data and extract complex structures distorted by feed-forward methods. In contrast to most forms of machine learning that tend to work on formatted numerical data, CNNs can assess raw image data to understand important attributes like architectural styles, materials quality, and neighbourhood appearance.

The present studies also indicate that CNNs can be used in real estate evaluations. For instance, when comparing property images, some scholars have employed CNNs to connect property appearance and house value, and while doing so, they have experienced increased prediction accuracy. These models expand the existing frameworks, VGGNet, ResNet, or Inception, to decipher the differences within image data and utilize derived insights on real estate. Furthermore, the integration of CNN with other forms of information, such as geographical coordinates and other numerical values, has provided good features for formulating an overall model. This approach enables several of them to be processed simultaneously, limiting the model's ability to deal with the visual information and the context of the price.

2.3. Research Gaps

Yet, there are still some gaps, even in the case of house price prediction, after considering the new approach to machine learning – deep learning. One more limitation of prior research is that the incorporated image information is scarce. A portion of previous research has looked into the use of property pictures. However, many works concentrate solely on data fields, including size, number of rooms and location of facilities. It emphasizes items that continue to say little about property conditions, style and even the area surrounding a property.

Also, the integration of multimodal data, including structured, unstructured, and geospatial data, is relatively unknown. While there has been growing interest in what the various implementations of multimodal data can offer to improve FERS, the integration of data types in a unified method remains rather unsystematic. Some, like the data alignment, the preprocessing and model optimization processes, have been a major cause of straightening a broad use of these techniques.

Table 1 Summarization related work, comparing traditional models

Aspect	Traditional Models	Deep Learning Models
Examples	Regression, Decision Trees, Random Forests	Convolutional Neural Networks (CNNs), Transformers
Accuracy	Moderate to high accuracy for simple datasets	High accuracy, especially with large and complex datasets
Features	Relies on manual feature engineering	Automatic feature extraction through layers
Input Data	Structured numerical data	Structured and unstructured data (e.g., images, text)
Complexity	Simple models, easy to interpret	Complex models, often harder to interpret
Scalability	Limited scalability with large datasets	Highly scalable with large datasets and compute power

They include inadequate consideration of spatial correlation, unexplored cross-sectional dependencies, and scarce analysis of temporal changes in property prices. Static data offers a description of the market conditions. However, housing markets always change due to economic cycles, policy changes, or seasons. Further, published articles that include temporal data or examine how deep learning models are suited for changing market conditions over time are scarce.

Finally, issues related to the suitability of CNN-based models for scale and their applicability remain questions. Previous methods used in the literature are based on relatively small samples. Therefore, they do not fully mimic real markets. Such a limitation gives rise to doubts concerning the universality of these models for other geographical areas, economic environments, and cultural tastes. It is to fill these gaps that this paper contributes to furthering the field and exploring the possibilities of deep learning in real estate applications.

3. Methodology

This section describes the approach to constructing a predictive model for house prices using CNNs and shows how it was evaluated. It embraced precise phases in data acquisition, initial preparation, selecting and building models or predictors to achieve strength and reliability of the outcomes.

3.1. Data Collection

This was achieved by allowing visual and non-visual features in the data set used in the study to boost the predictions and improve their authenticity. Open data databases and datasets like Zillow or Kaggle were used, while where available and relevant, closed or restricted datasets were used. These sources provided a good understanding of real estate properties to reduce variance in the depicted properties and the market conditions across the regions.

The features extracted for this work are divided into numerical characteristics and image characteristics. The numerical characteristics included gross living areas, number of bedrooms and bathrooms, property age and sexagenary coordinates on the map. Such attributes gave the basic knowledge regarding the quantitative data on property appraisal. Image-based features were characterized by high-quality interior and exterior photos of houses. These visual aspects incorporated aesthetics and architectural characteristics, defining buyer preferences and house prices in the model.

3.2. Data Preprocessing

Data preprocessing was crucial because the data should meet the requirements of the model. Preprocessing for image data entailed scaling the images to a standard size since the convolutional neural network's input layers require standardized dimensions. Normalization of the pixel intensity was done to range the pixel values from 0-1, which is crucial for training the model. The preliminary data augmentation techniques used in this work included flipping, rotation, and thickness adjustment to expand the variability of the training inputs.

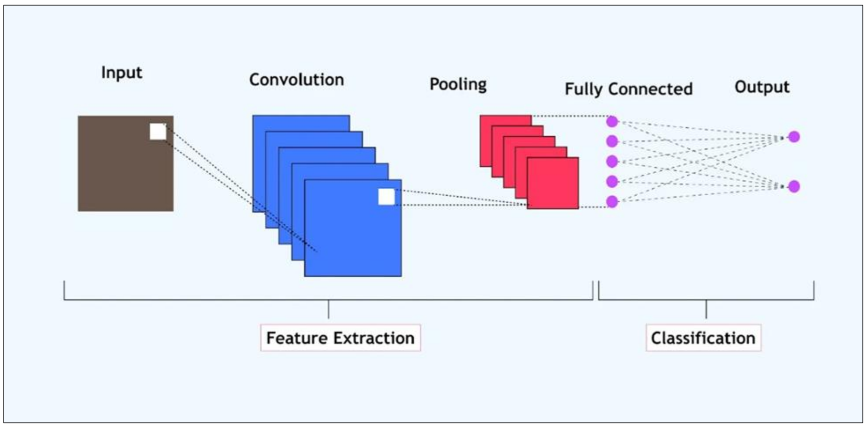

Figure 2 CNN architecture diagram showing layers

The preprocessing of numerical data was also conducted using a different set of operations. To handle missing values, the authors used the data imputation method either through simple mean or simple median since most data were collected through self-generated questionnaires. Feature scaling was used to bring the numerical input data to a common range and make the emphasis on all features comparable. The categorical data, like the properties' types, were also encoded into the form that could be inputted into the neural network through one hot encoding.

3.3. Model Architecture

This work designed the predictive model as a multimodal neural network that combined CNNs to handle images and fully connected layers to cater for numerical data. This architecture allowed it to incorporate various data types and exploit the CNNs' ability in pattern recognition and the dense layers in numerical computation.

Hence, the base models used for image processing were ResNet and VGG in a pre-trained manner. These architectures were chosen because they demonstrated the ability to learn high-level semantic features from the imagery. These base models were trained on the property image dataset to modify them towards domain-specific tasks of house pricing in real estate. The fully connected layers of the CNNs extracted features were flattened and then fed to a fusion layer.

Regarding the Converts, the numerical data were passed through a sequence of fully connected layers to identify structural patterns and relations among the features. These layers, for instance, applied the Rectified Linear Units (ReLU) activation functions to introduce non-linearity into the model and increase the model learning feat.

The CNN and numerical pathways were combined through the fusion layer to produce a single output of the multimodal input data. The features below were passed through more dense layers, making the final output layer suitable for regression problems. The element of the output layer employed a linear activation function to make real-value predictions regarding the house prices.

The Adam optimizer was chosen throughout model optimization due to its capacity to adjust learning rates on the fly. To rectify this, a learning rate was scheduled to be scaled down as training was done, ensuring the model did not overfit and had converged. The loss function will be used as a mean squared error (MSE) since it fits regression problems and is more punitive to large errors than small ones.

3.4. Training and Validation

The data was then split into training, validation, and test data using conventional ratios in machine learning. All the training sets were used to update the models' weights, the validation sets were used to select the optimum hyperparameters, and the test sets were used to determine the generality of the models. This format was chosen to avoid the risk of the subsets being skewed in some way when selecting properties for model evaluation.

Various data augmentation strategies were employed during training to increase model resistance. By using augmented versions of the original images, the model was exposed to different situations, making it more capable of dealing with unforeseen circumstances.

The performance of the built model was evaluated using various assessment measures to give a full view of the strengths and weaknesses of the system. The others include using Root Mean Square Error (RMSE) to estimate the mean of the squared differences between predicted and actual prices, focusing on the effect of outsized error bars. As a measure of the accuracy of the prediction, the mean absolute error gave a clear idea of the average error made in terms of its magnitude. Also, the R-squared measure ofริตรว was employed in the assessment of the capability of the model in accounting for the variability of the house prices in the best way possible.

By following a strict development and testing process, the paper's authors establish the usefulness of using CNNs in enhancing house price prediction using available multimodal data. This approach emphasizes the possibility of uniting graphical and numeric characteristics to solve challenging real-life issues in real estate.

4. Experimental Results

This section will describe the experiment findings and lessons learned when using convolutional neural networks (CNNs) to estimate the price of a house. The accuracy of the proposed model is compared against traditional machine learning models. Then, a focused analysis of CNN is performed to determine how the model can use features of the images. Finally, an ablation study looks at the importance of different data types and analyses the effect of different architectural decisions on the model.

4.1. Baseline Comparison

Consequently, a relative benchmarking was set when the dataset was compared to traditional machine learning models such as Random Forest, XGBoost, and Linear Regression. These models were trained with numerical and categorical features typical in property appraisal, such as number of rooms, area size, coordinates of the property and age of the property. To achieve feature consistency across all models, basic Figure 10 Normalization of numerical variables and one hot encoding of the categorical attributes were applied.

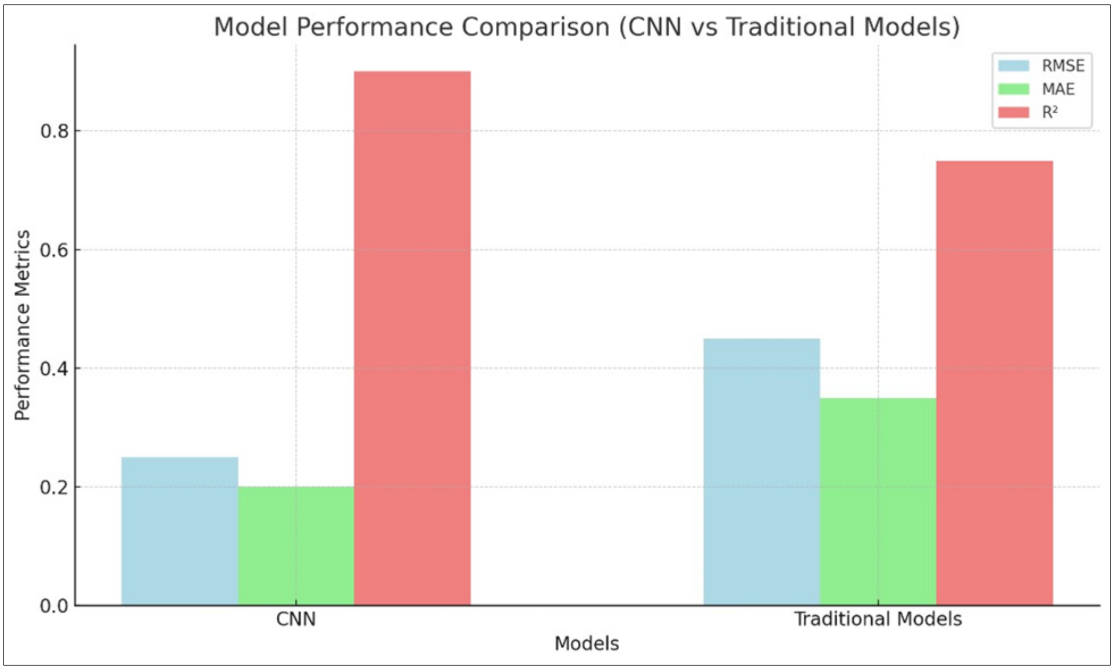

Figure 3 A Graph Comparing Model Performance Metrics

The results have suggested that while basic models offer sufficiently good performance, they fail to accurately represent the spatial and aesthetic characteristics innate to property images. The results found that Random Forest and XGBoost have slightly better accuracy than the Linear Regression model because they are suited for modelling non-linear datasets. However, these methods could not readily capture the aspects of architectural style, structural integrity or attractiveness of the surrounding environment. Such limitations highlighted the need for a more advanced powered framework that can incorporate regional and multimodal spatial data.

4.2. CNN Model Performance

Another strength established from the CNN-based model was the enhanced results in the predictive accuracy compared to traditional approaches. This means that the proposed model improved solution quality by incorporating important aspects of property evaluation previously neglected due to the complexity of property images. Specifically, we found that using the CNN improved the baseline models for all primary evaluation measures, such as MAE and RMSE, when tested on the dataset containing the selected test pictures.

The flexibility of the CNN in identifying extra-small features of images, such as the quality of the construction materials, presence of modern amenities, and [neatness], helped attain better prediction. Moreover, the CNN model demonstrated that numerical and non-numerical data could be combined and analyzed synergistically since structured and unstructured information bring about different productivity determinants of prices.

To arrive at this conclusion, a deeper analysis of the contribution of the CNN showed that specific structures that showed different architectural surfaces or accommodation structures located in visually attractive regions received a substantial boost from the model's capabilities. For example, estimating the probabilities of houses with high image features, such as curb appeal, was done with much less error than traditional models that only relied on numerical features. Thus, these observations support the hypothesis that it is crucial to analyze visual data to rectify and enhance an accurate identification of the property's value.

4.3. Ablation Study

An ablation study compared and evaluated which features influence the model most. This analysis systematically lesioned inputs of the CNN model, that is, features such as numerics or images, and assessed the impact of such lesioning on the model's predictive capability. Evaluation studies showed that the model gave reasonable predictions when only the numerical vector was used. Still, when connected to image data, the accuracy improved, especially for properties with distinguishing features when photographed. On the other hand, the performance was significantly reduced when the image features with low discriminative capability were omitted, thereby proving their significance in the CNN architecture.

Table 2 Performance Metrics for Various CNN Architectures

CNN Architecture	Accuracy (%)	Training Time (hrs)	Model Size (MB)	Number of Parameters	Inference Time (ms/image)	FLOPs (Floating Point Operations)	CNN Architecture
ResNet-50	75.0	5	98	25.6 million	10	4.1 billion	ResNet-50
VGG-16	71.4	10	528	138 million	25	15.5 billion	VGG-16
VGG-19	72.1	12	549	143 million	30	19.6 billion	VGG-19
InceptionV3	79.4	7	92	23.8 million	15	5.7 billion	InceptionV3
MobileNetV2	70.6	4	14.5	3.4 million	7	0.6 billion	MobileNetV2
DenseNet-121	74.9	6	33	8.0 million	12	3.3 billion	DenseNet-121
Xception	79.0	8	88	22.9 million	13	8.7 billion	Xception
EfficientNet-B0	76.3	5	20	5.3 million	9	1.4 billion	EfficientNet-B0

This work also benchmarked different CNN architectures to determine the best setting for this task. To compare the performance of the proposed mixture of experts, architectures from the plain convolutional layers up to architectures like ResNet and EfficientNet were tested. All the architectures showed gains concerning the baseline architectures. However, the deeper network with residual connections performed well, underscoring the importance of capturing deeper hierarchical features for this task. However, what was also observed was that the performance dropped when architectures were very large and comprised multiple layers, perhaps due to the small size of the samples.

5. Discussion

5.1. Interpretation of Results: Insights into Why CNN Outperforms Traditional Models

The findings of this research conclusively emphasize that CNNs outperform conventional statistical or machine learning models in real estate valuation-related tasks. This has stemmed from the fact that CNNs can deal with high-level features and high dimensionality data. Different from traditional approaches that heavily embrace quantitative data of, for example, form-based inputs in quantitative and qualitative categories, CNNs perform exceptionally well when it comes to feature extraction on irregular data like images and geography, which play a crucial role in real estate evaluation.

One of the groundbreaking insights that could be deduced from the work is CNNs ability to analyze and integrate disparate forms of visuals and locational data that affect property value considerations. Image data can capture things like architectural design, view, or the relative closeness to a famous structure or park well, thus providing a more integrative picture. Older methods, which rely on hand-crafted features, tend to overlook these nuances as they are constructed based on feature engineering. CNNs, however, can self-normally learn different levels of features that include edges and texture patterns, building styles, and urban layouts, among others.

However, CNNs are far better at managing spatial dependencies, which are important in real estate settings. As location aspects, such as neighbourhood characteristics, accessibility and urban density, are location aspects, they are hard to quantify using tabular data. In this way, CNNs manage to conserve and extract spatial structures within the data, making them more suitable for tasks that require understanding the dependence of locality and (or) its influence on property values. This enhancement in handling and processing multiple forms of information, including images, maps and recordations of numbers, puts them at an advantage beyond their traditional equivalents.

5.2. Practical Implications: Real-World Applications in Real Estate Valuation and Decision-Making

The real-life applications of CNNs to real estate valuation are significant in their consequences. In the case of power players like property developers, investors, and government bodies, the ability of CNNs to offer a more accurate – and live – evaluation will be highly beneficial for decision-making. For example, property developers can use CNN-based models to locate underserved properties or comprehend how proposed structures affect property markets in specific neighbourhoods. These predictive accuracies can dramatically decrease real estate investing and financial marketing risks.

To a real estate agency, it brings a competitive advantage as it does give a definite valuation of the property, taking into consideration visual and spatial characteristics of the property, all due to CNN-powered valuation tools. These models can investigate the effectiveness of particular aesthetic or locational characteristics that other approaches tend to eliminate. For example, things like a well-maintained facade or being close to a park can significantly affect property value; information cannot be easily obtained without the impressive feature extraction attributes of CNN.

It's just as bright for public sector applications. Urban analysts and decision-makers can apply CNN-based approaches to simulate the housing supply-demand systems, implement fair-zoning policies or estimate the social impacts of urban redevelopment projects. For instance, knowledge of how green areas or the provision of transport facilities affect property prices in the adjacent regions can help in efficient urban growth planning.

In addition, these models could provide real estate valuation at a median price point when integrated into consumer-facing applications. Homebuyers could use other AI-powered platforms to make instant and correct property value estimations by submitting photographs or describing the location. This accessibility could eliminate the need for human appraisers, making the process more efficient and providing a look into the real estate market.

5.3. Challenges: Computational Requirements, Data Availability, and Model Interpretability

CNNs have obvious benefits; at the same time, it is important to mention several issues that can hinder the appropriate use of CNNs to solve real estate valuation problems. First, one must examine the computational demands for training and deploying CNN models. These networks consume many computer resources, especially when handling vast amounts of data and high-definition images. Implementing a CNN requires, in many cases, specialized hardware like GPU, which is still quite expensive for small businesses or even individuals. Therefore, the basic results of model optimization, hyperparameters tuning, and updating also elevate the resources required.

Another vital challenge is the availability of data. The downside is that if CNNs must be implemented to their optimal potential, they need huge quantities of structured and unstructured data, including records such as property prices squared footage and inputs like images, maps and street views. In many cases, especially in developing markets, such detailed data may not be available or may exist in a piecemeal fashion in several places. Another challenge is data quality and consistency; that is, data inaccuracy of certain formats can have a massive impact on typical ML model quality.

Another big issue is model interpretability. That's why CNNs have high predictive capabilities, but what they do with this data is unclear, as they are often criticized for being 'black boxes'. The imperative for a high level of transparency of parameters in this model can hamper trust and take-up by stakeholders used to the more linear mechanics of conventional frameworks. For instance, property appraisers or investors may not use the model if it does not show outright why the valuation of a property rose due to some attributes. This study attempts to mitigate this problem by presenting methods for increasing the interpretability of CNNs, including visualization of feature maps or using explainable AI (XAI).

Finally, some ethical or regulatory concerns arise when applying CNNs to real estate. In light of this, the application of unstructured data, especially images, raises issues of privacy and consent; the privacy and consent issues; social media platforms or any other source while obtaining consent. Furthermore, the training data set may contain biases that can either exacerbate or create unfair differentiation in property assessment among groups of people belonging to minorities. For instance, historical redlining patterns in the spatial data may predispose the model to yield discriminating results unless the problem is actively woven.

6. Future Directions

Though the current process has shown a high accuracy in the predictions, there are some possibilities to improve the model to eliminate its shortcomings and adapt it to the constantly changing conditions of the real estate market. One of these is collecting more diverse data types, as described in the previous section. Increasing training data sizes with properties from different areas, prices, and types of buildings' architecture can improve the model's performance. The second reason is that the model could benefit from a more diverse data set for figuring out cultural, economic and environmental characteristics influencing the value of assets in various markets.

Figure 4 A Roadmap Chart Detailing Future Advancements

In this sense, the first opportunity concerns the potential of using a different deep learning architecture, such as transformers. Initially applied to NLP, transformers have excelled at problems where there is a need to understand relevant relationships between elements of high dimensional data. To address the issue of incorporating heterogeneous inputs such as textual and image descriptions alongside structured numerical data, applying transformers or combining CNNs and transformers as in hybrid structures could enhance the model's ability to process and analyze the inputs. These architectures may also be able to model longer-range dependencies of the data more effectively–thus possibly yielding more precise predictions.

Also, extending the model to enable the incorporation of temporal characteristics could open up new opportunities for temporal price prediction. The nature of real estate prices includes specific time factors in its prices, including economic factors, policy changes, and cycles. Apart from the current property prices, the situational data, including the historical price changes and transaction rate, the model could predict the future trends of property prices. This capability would be particularly useful to investors and policymakers for future market forecasts.

From the current work, improvements could be made to extend the model's complexity and increase its interpretability and usefulness. As demonstrated, CNNs outperform competing methods in prediction accuracy; they suffer from the "black-box" problem that often hinders end-users who wish to understand how the prediction arrived at. Methods to better show and understand how the model is made would lead to increased trust and take up among stakeholders. Likewise, fine-tuning the model through the integration of optimized GUI and deployment of the model as a live application has the potential to expand its audience to technophobe individuals.

Lastly, ethical and fairness issues should be addressed in the future real estate market. It is, therefore, important that the model leaves no possibility of compounding the discrimination observed in the training data set by discriminating against those looking for dwelling units based on prohibited characteristics. More studies can be conducted on using fairness-aware treatments in machine learning algorithms to minimize risk and guarantee that the model is accurate and fair.

7. Conclusion

Here, conclusions revolve around CNNs for accurate house price prediction. As the method that can identify spatial features from image and structure data, CNNs will provide a stable framework for embodying the relations between property attributes and their corresponding value. The study proves that CNNs perform better than statistical and machine learning models with the right tuning and complete data inputs. This success demonstrates the ability of CNNs to handle high-dimensional data while stressing the importance of these networks in approaching problems in real estate forecasting in the future.

The results obtained by this model confirm that CNNs provide effective solutions in cases where visual or structured information is present, for instance, for evaluating the property characteristics obtained from images, georeferenced data, and numerical descriptors. Using such data types makes the model more precise and realistic, bringing significant benefits to potential buyers and sellers of real estate and financial organizations. However, it has been seen that there are still some ways to refine and enlarge the model to drive more preciseness and flexibility.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed

References

[1] Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I., & Angel, S. (1977). A pattern language. Oxford University Press.

[2] Anselin, L. (1988). Spatial econometrics: Methods and models. Kluwer Academic.

[3] Bengio, Y., LeCun, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

[4] Cheshire, J., & Sheppard, S. (1995). On the price of land and the value of amenities. Economica, 62(246), 247–267. https://doi.org/10.2307/2554903

[5] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. arXiv preprint arXiv:1610.02357. https://arxiv.org/abs/1610.02357

[6] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Li, F. F. (2009). ImageNet: A large-scale hierarchical image database. In Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 248–255). https://doi.org/10.1109/CVPR.2009.5206848

[7] Doersch, C., Singh, S., Wu, C., & Hui, W. (2012). What makes Paris look like Paris. ACM Transactions on Graphics, 31(4), Article 101. https://doi.org/10.1145/2185520.2185564

[8] ESRI. (2013). CityEngine. Retrieved from http://www.esri.com/software/cityengine (Accessed in 2013).

[9] Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E., & Li, F. (2017). Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. Proceedings of the National Academy of Sciences, 114(50), 13108–13113. https://doi.org/10.1073/pnas.1700035114

[10] Gehl, J. (1971). Life between buildings: Using public space. The Danish Architectural Press.

[11] Gehl, J. (2010). Cities for people. Island Press.

[12] Girshick, R. (2015). Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 1440–1448). https://doi.org/10.1109/ICCV.2015.169

[13] Google. (2017). Google Maps. Retrieved from https://www.maps.google.com (Accessed in 2017).

[14] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385. https://arxiv.org/abs/1512.03385

[15] Heffernan, E., Heffernan, T., & Pan, W. (2014). The relationship between the quality of active frontages and public perceptions of public spaces. Urban Design International, 19(1), 92–102. https://doi.org/10.1057/udi.2013.30

[16] Historic England. (2017). Historic England Listed Parks and Gardens dataset. Retrieved from https://historicengland.org.uk/listing/the-list/data-downloads/

[17] Jacobs, J. (1961). The death and life of great American cities. Random House Inc.

[18] Kickert, C. (2016). Active centers - Interactive edges: The rise and fall of ground floor frontages. Urban Design International, 21(1), 55–77. https://doi.org/10.1057/udi.2015.26

[19] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (NIPS). Retrieved from https://papers.nips.cc/

[20] Land Registry. (2017). House price transaction dataset. Retrieved from https://www.gov.uk/search-house-prices

[21] Law, S., Shen, Y., & Seresinhe, C. (2017). An application of convolutional neural network in street image classification: The case study of London. In Proceedings of ACM GeoAI '17 Workshop on Artificial Intelligence and Deep Learning for Geographic Knowledge Discovery.

[22] Law, S., Paige, B., & Russell, C. (2018). Take a look around: Using street view and satellite images to estimate house prices. arXiv preprint. https://arxiv.org/abs/1807.07155

[23] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324. https://doi.org/10.1109/5.726791

[24] Li, L., Tompkin, J., Michalatos, P., & Pfister, H. (2017). Hierarchical visual feature analysis for city street view datasets. In Proceedings of the IEEE VIS 2017 Workshop on Visual Analytics for Deep Learning.

[25] Li, Y., Paluri, M., Rehg, J., & Dollar, P. (2016). Unsupervised learning of edges. In Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR).

[26] Li, X. J., Santi, P., Courtney, T. K., Verma, S. K., & Ratti, C. (2018). Investigating the association between streetscapes and human walking activities using Google Street View and human trajectory data. Transactions in GIS, 22(4), 1–16. https://doi.org/10.1111/tgis.12442

[27] Liu, L., Silva, E., Wu, C., & Hui, W. (2017). A machine learning-based method for the large-scale evaluation of the urban environment. Computers, Environment and Urban Systems, 65, 113–125.

[28] Llewelyn Davies Yeang & Homes Communities Agency. (2013). Urban Design Compendium. English Partnerships and The Housing Corporation.

[29] Mirowski, P., Grimes, M.K., Malinowski, M., Hermann, K.M., Anderson, K., Teplyashin, D., Simonyan, K., Kavukcuoglu, K., Zisserman, A., & Hadsell, R. (2018). Learning to navigate in cities without a map. arXiv preprint. arXiv:1804.00168.

[30] Moran, P. (1948). The interpretation of statistical maps. Biometrika, 35, 255–260.

[31] Naik, N., Philipoom, J., Raskar, R., & Hidalgo, C.A. (2014). StreetScore – Predicting the perceived safety of one million streetscapes. CVPR Workshop on Web-scale Vision and Social Media.

[32] Nase, I., Berry, J., & Adair, A. (2013). Hedonic modelling of high street retail properties: A quality design perspective. Journal of Property Investment & Finance, 31(2), 160–178.

[33] Nationwide. (2012). House price dataset. Permission granted from LSE.

[34] Office of the Deputy Prime Minister (ODPM). (2005). Safer Places: The Planning System and Crime Prevention. Home Office.

[35] Palmquist, R.B. (1984). Estimating the demand for the characteristics of housing. Review of Economics and Statistics, 66(3), 394–404.

[36] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. arXiv preprint. arXiv:1506.01497.

[37] Ridker, R.G., & Henning, J.A. (1967). The determinants of residential property values with special reference to air pollution. Review of Economics and Statistics, 49(2), 246–257.