Improving Real Estate Appraisal with POI Integration and Areal Embedding⋆

Sumin Han[0000−0002−4071−8469]1, Youngjun Park[0000−0002−4254−2268]1, Sonia Sabir[0009−0001−8485−0701]1, Jisun An[0000−0002−4353−8009]2, and Dongman Lee[0000−0001−5923−6227]1

[1] Korea Advanced Institute of Science and Technology (KAIST){hsm6911,youngjourpark,sonia,dlee}@kaist.ac.kr
[2] Indiana University Bloomington jisunan@iu.edu

Abstract. Despite advancements in real estate appraisal methods, this study primarily focuses on two pivotal challenges. Firstly, we explore the often-underestimated impact of Points of Interest (POI) on property values, emphasizing the necessity for a comprehensive, data-driven approach to feature selection. Secondly, we integrate road-network-based Areal Embedding to enhance spatial understanding for real estate appraisal. We first propose a revised method for POI feature extraction, and discuss the impact of each POI for house price appraisal. Then we present the Areal embedding-enabled Masked Multihead Attention-based Spatial Interpolation for House Price Prediction (AMMASI) model, an improvement upon the existing ASI model, which leverages masked multi-head attention on geographic neighbor houses and similar-featured houses. Our model outperforms current baselines and also offers promising avenues for future optimization in real estate appraisal methodologies.

Keywords: House price appraisal · Point of Interest (POI) · Areal Embedding · Masked multi-head attention

1 Introduction

Real estate appraisers play a crucial role in delivering essential property valuations that serve diverse purposes, including buying and selling, securing financing, determining taxation, and establishing insurance coverage. Conventional real estate appraisal often relies on subjective methods such as the sales comparison approach, where a property's value is assessed by comparing it to similar properties in the same market. While the Hedonic Price Model [9] remains foundational in real estate appraisal, positing that a property's value is shaped by its individual characteristics, the inherent subjectivity and context dependency in establishing comparability makes property valuation a complex task. Recently,

there has been a surge in deep learning-based methodologies, aiming to determine land prices that reflect spatial understanding and correlations among houses in large-scale datasets [3, 6, 7, 10, 15].

Despite the significant strides made in real estate appraisal methods, we mainly focus on two major challenges in this paper:

1. **POI Integration:** While numerous studies have focused on physical house feature selection in property valuation models, a common issue is the lack of explicit identification of which POI is truly pivotal in determining property values [2]. This challenge highlights the need for a more comprehensive and data-driven approach to feature selection, which can help pinpoint the most influential factors in property valuation.
2. **Areal Embedding:** The recent study demonstrates that employing trainable regional embedding can improve real estate appraisal by incorporating comprehensive spatial knowledge beyond latitude and longitude [5]. However, these methods often demand sophisticated and comprehensive data, such as mobility data or satellite images, posing a challenge for the integration of scalable areal embedding techniques.

To address these challenges, we employ a simplified method for POI feature extraction and introduce the Areal Embedding-based Masked Multihead Attention-based Spatial Interpolation for House Price Prediction (AMMASI) model as a refinement of the existing ASI [10] model. Applying masked multihead attention to both geographic neighbor houses and those with similar features, our model outperforms current baselines. We summarize our key contributions as follows:

- We introduce a novel method named AMMASI, building upon the foundation of the ASI model, with publicly available code implementation [3].
- We propose integrating Point of Interest (POI) data with publicly accessible OpenStreetMap by engineering features with Gaussian proximity.
- Our model utilizes road network-based Areal Embedding to incorporate geographical insights, presenting a simplified approach in comparison to existing methods.

2   Related Work

2.1   Recent Deep Neural Networks on Real Estate Appraisal

Recent advancements in real estate appraisal have placed a strong emphasis on innovative approaches such as the selection of reference houses or community construction. PDVM [1] employs K-nearest similar house sampling for sequence generation, while ASI [10] focuses on estimating house prices through a hybrid attention mechanism. MugRep [15] incorporates a hierarchical heterogeneous community graph convolution module. ReGram [7] utilizes a neighbor relation graph

---

[3] https://github.com/cdsnlab/AMMASI

with an attention mechanism, and ST-RAP [6] adopts a hierarchical model, integrating temporal and spatial aspects alongside amenities. Additionally, DoRA [3] introduces a domain-based self-supervised learning framework with pretraining on geographic prediction. Collectively, these methodologies comprehensively address the complexities of real estate appraisal by considering spatial, temporal, and community factors.

2.2   Using POI Features for Real Estate Appraisal

The utilization of Points of Interest (POI) in real estate appraisal has a long-standing history dating back to the hedonic model. These attributes, encompassing schools, parks, shopping centers, and transportation hubs, exert diverse impacts on property values, posing a challenge in quantifying their relative significance. Ottensmann et al. [8] introduced the use of distance to the Central Business District (CBD) as an additional house attribute, while Xiao et al. [14] proposed accessibility indices using a distance-based metric $(1 - d_{ij}/D)$ to measure proximity to amenities. ASI [10] took a different approach by leveraging the number of POIs in each dataset through the crawling of external APIs, although their POI dataset remains non-public. On the other hand, DoRA [3] adopted a tabular format, distinguishing between Yes In My Back Yard (YIMBY) facilities, like parks and schools, and Not In My Back Yard (NIMBY) facilities, such as power stations and landfills. The calculation of the number of POIs for the real estate property was performed using the Euclidean distance.

2.3   Areal Embedding

Addressing the Learning an Embedding Space for Regions (LESR) problem, RegionEncoder [5] presents a holistic approach to jointly learning vector representations for discrete urban regions, leveraging diverse data sources such as satellite images, point-of-interest data, human mobility patterns, and spatial graphs. Hex2vec [13] introduces an innovative technique for learning vector representations of OpenStreetMap (OSM) regions, incorporating considerations of urban functions and land use within a micro-region grid. Shifting focus to Representation Learning for Road Networks (RLRN), RN2Vec [11] leverages the neural network model to obtain embeddings of intersections and road segments. Node2Vec [4] is often referenced as a baseline for graph-based node representation learning in this domain.

3   Problem Formulation

Let $\mathbf{f}_i \in \mathbb{R}^{N_f}$ represent the physical features of house i, such as the number of beds and rooms. Longitude and latitude are denoted as loc_i^x and loc_i^y, respectively, and are not included in the house features. Furthermore, we define global geometric features of Points of Interest (POIs) and roads, that every house can leverage. Polygon geometries are defined for N_P types of Points of

Interest (POIs), such as commercial land uses and parks. Each POI type has corresponding regions or buildings represented as polygon geometries, and the union geometry for each POI type is denoted as $GEO^{POI}_{1,...,N_P}$. Line geometries of roads are given and denoted as $GEO^{ROAD}_{1,...,N_R}$. The goal is to train a model h that appraises the house price as $\hat{y}_i$ using the provided input defined above. The model function is defined as:

$$h(\mathbf{f}_i, loc^x_i, loc^y_i; GEO^{POI}_{1,...,N_P}, GEO^{ROAD}_{1,...,N_R}) \rightarrow \hat{y}_i \tag{1}$$

4   Method

4.1   POI Feature Extraction

Using the loc^x_i and loc^y_i for each house, we extract the POI feature by calculating the proximity to each POI geometry $GEO^{POI}_{1,...,N_P}$ using the gaussian filter based on distance measure as $\mathbf{p}_i = [\text{Prox}_{i,\tau}]$ $\forall \tau \in \{1,...,N_P\}$, where $\text{Prox}_{i,\tau} = \exp\left(-\text{Dist}^2_{i,\tau}/2\beta^2\right)$. We calculate the Euclidian distance of a i-th house to closest τ-th geometry $\text{Dist}_{i,\tau}$ using a GIS software[4], and experimentally determine the appropriate β value in Sec. 6.1.

4.2   Areal Embedding

We propose two primary approaches to leverage areal embedding for the location of houses, diverging from the conventional use of latitude and longitude coordinates as extra house attributes as described in [10]. Our method involves dividing an area into $M = M_x \times M_y$ grids, each comprising areas denoted as $\mathscr{A}_{1,...,M}$. For each of these areas, we construct a D-dimensional embedding. Each house corresponds to an area, enabling us to later use the respective areal embedding to infer its price in that specific location.

2D Positional Encoding An initial avenue we can explore involves the application of sinusoidal 2D positional encoding, inspired by the work of [12], as an alternative to using raw latitude and longitude coordinates. By incorporating this spatial information, the model can discern the correlation between house prices and location tendencies, resulting in an embedding in $\mathbb{R}^{M_x \times M_y \times D}$.

Node2Vec Embedding In this method, we consider each line geometry of the r-th road, denoted as GEO^{ROAD}_r. We assume that the areas traversed by this road can be represented by the set $\mathscr{A}^{(r)} \subset \mathscr{A}_{1,...,M}$. Assuming interconnectedness between these areas, we count each ordered pair $(u,v) \in \mathscr{A}^{(r)} \times \mathscr{A}^{(r)}$, where $u \neq v$, as one connection. By applying this process to all N_R roads, we construct the adjacency matrix $Adj \in \mathbb{N}^{M \times M}$, where $Adj_{[i,j]}$ denotes the number of roads connecting $\mathscr{A}_i$ and $\mathscr{A}_j$. Subsequently, we employ Node2Vec [4] with Adj to learn

---
[4] Using GeoPandas (https://geopandas.org/) without CRS transformation (WGS84).
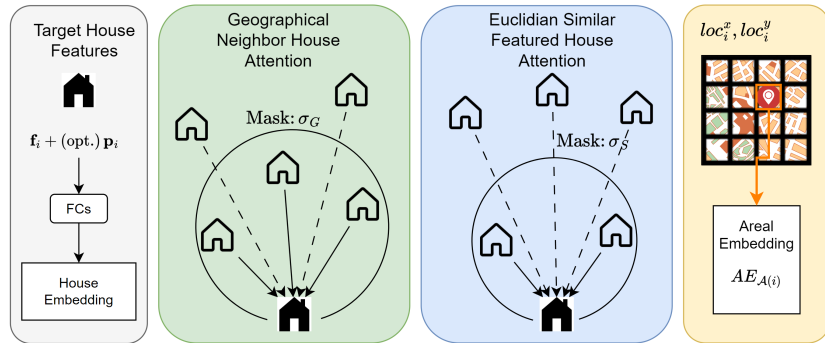
Fig. 1: Proposed model architecture of AMMASI before the final regression layer.

areal embeddings and visualize as Fig. 2. Note, that the normalized weights of *Adj* can be considered as the probability of random walks between areas, aligning with the fundamental concept of Node2vec.

4.3 Deep Neural Network (AMMASI)

We introduce **A**real embedding-enabled **M**asked **M**ultihead **A**ttention-based **S**patial **I**nterpolation (AMMASI) for house price prediction. Our proposed model builds upon the foundational framework of ASI [10], specifically leveraging geographical neighbor houses and similar-featured houses. The comprehensive architecture is illustrated in Fig. 1.

House Feature Embedding Layer To encode the house features $\mathbf{f}_i$, we employ two stacked fully connected layers (FCs), resulting in D-dimensional embedding. Optionally, POI features $\mathbf{p}_i$ can be concatenated to $\mathbf{f}_i$ to enrich the house features before ingested to FCs.

Masked Multi-head Attention on Reference Houses Similar to the architecture proposed in [10], our model employs a house attention mechanism operating on two types of house knowledge. First, we leverage information from geographically nearby houses and their prices, as they offer insights into the surrounding area. Secondly, we consider houses with similar features regardless of geographical distances and their corresponding prices. While building upon the foundational framework of ASI, we have identified a scalability limitation in implementing the attention mechanism, particularly when using the widely adopted dot product-based approach. To overcome this challenge, we enhance our model by introducing a masked multi-head attention mechanism tailored specifically for these two types of house attention.

Let $\mathbf{Gidx}^{(i)} = \{\text{Gidx}_1^{(i)}, ..., \text{Gidx}_{N_G}^{(i)}\}$ and $\mathbf{GDist}^{(i)} = \{\text{GDist}_1^{(i)}, ..., \text{GDist}_{N_G}^{(i)}\}$ represent the indexes and geographical distances of geographically neighboring

houses based on location. Additionally, let $\mathbf{Sidx}^{(i)} = \{\text{Sidx}_1^{(i)}, ..., \text{Sidx}_{N_S}^{(i)}\}$ and $\mathbf{SDist}^{(i)} = \{\text{SDist}_1^{(i)}, ..., \text{SDist}_{N_S}^{(i)}\}$ denote the indexes and distances of houses with similar features, measured by Euclidean distances of house features. Here, N_G and N_S are the numbers of houses for attention of each module. As we leverage the same index and distance-measured legacy dataset provided by the authors of ASI, readers can readily refer to their definition in [10] for further details. Note, that the reference houses are found among the houses that only exist in the training dataset, that is $(\mathbf{Gidx}^{(i)} \cup \mathbf{Sidx}^{(i)}) \subset \text{TrainDataset}$. Therefore, we can look up their house attributes and prices during the inference phase. Let concatenation of i-th house attributes with its price as $\mathbf{f}_i^\star$. Then we can conduct the attention mechanism as follows:

$$emb^{(query)} = f_q(\mathbf{f}_i) \in \mathbb{R}^{1 \times d} \tag{2}$$

$$emb^{(key)} = f_k([\mathbf{f}_{\text{Gidx}_1^{(i)}}^\star, ..., \mathbf{f}_{\text{Gidx}_{N_G}^{(i)}}^\star]) \in \mathbb{R}^{N_G \times d} \tag{3}$$

$$emb^{(value)} = f_v([\mathbf{f}_{\text{Gidx}_1^{(i)}}^\star, ..., \mathbf{f}_{\text{Gidx}_{N_G}^{(i)}}^\star]) \in \mathbb{R}^{N_G \times d} \tag{4}$$

Here, f_q, f_k, f_v are two-sacked FCs with ELU activation.

$$score_j = \frac{\exp\left(\langle emb^{(query)}, emb_j^{(key)}\rangle/\sqrt{d} + Mask_j^{(i)}\right)}{\sum_{k=1}^{N_G} \exp\left(\langle emb^{(query)}, emb_k^{(key)}\rangle/\sqrt{d} + Mask_k^{(i)}\right)} \in \mathbb{R}^{1 \times N_G} \tag{5}$$

where $\langle \bullet, \bullet \rangle$ denotes inner product operator, and $Mask_j^{(i)} = -\infty$ if $\text{GDist}_j^{(i)} > \sigma_G$ else 0. Finally, we leverage $output_G = \sum_{j=1}^{N_G} score_j \cdot emb_j^{(value)} \in \mathbb{R}^{1 \times d}$ as our final output embedding. Suppose there are K different attention heads that conduct the same function, then we can introduce K-attention head as $Output_G = f_G\left(\Big\|_m^{1,...,K} output_G^{(m)}\right)$, where $\|$ denotes the concatenation and f_G as two-stacked FCs. We apply attention similarly to houses with Euclidean similar features, yielding $Output_S$ as a result. The values of σ_G and σ_S are determined empirically, and their discussion will be elaborated in Section 6.3.

Areal Embedding Lookup Every house is associated with a specific area, denoted as $\mathscr{A}(i)$, where the location of the i-th house resides. Leveraging the pre-trained embedding values for each area introduced in Sec. 4.2, we incorporate areal embeddings before the final output layer in our model.

Final Output Layer In the concluding steps, we concatenate the previously mentioned embeddings, including (1) house feature embedding, (2) attention output with geographically neighboring houses, (3) attention output with similar-featured houses, and (4) areal embedding, resulting in a $4D$-dimensional embedding. Subsequently, we apply a final output dense layer, consisting of a two-stacked FCs with ELU activation, to generate the final output.
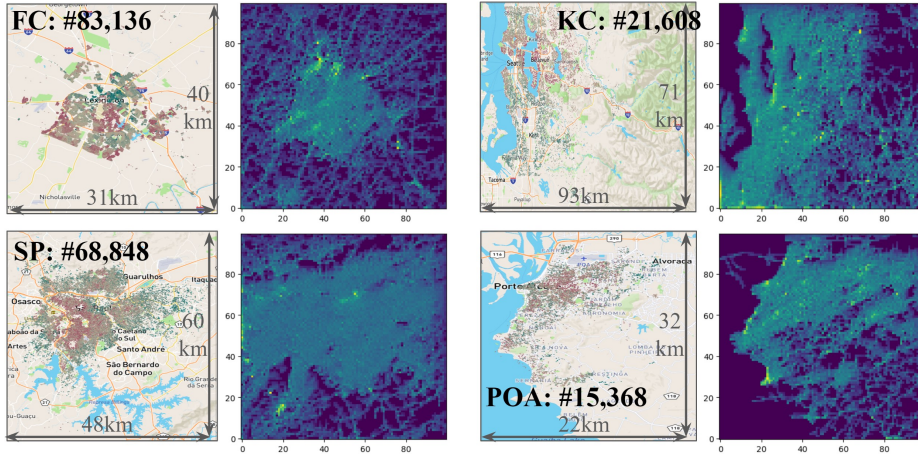
Fig. 2: Visualization of house prices (green < red, with the house count #) and Node2Vec areal vector magnitudes ($M_x \times M_y = 100 \times 100$).

Objective Function We train our model based on $L(\theta) = \frac{1}{n} \sum_{i=1}^{n} |log(y_i) - log(\hat{y}_i)|$ which is MAE of log value of the house prices.

5   Evaluation Settings

5.1   Dataset

We utilize four real estate datasets, namely FC, KC, SP, and POA, published by [10][5]. These datasets represent King County (KC) in Washington State, USA; Fayette County (FC) in Kentucky, USA; São Paulo (SP), Brazil; and Porto Alegre (POA), Brazil. We leverage the same **Gidx**, **GDist**, **Sidx**, **SDist** which are the indexes and the distance for the referencing houses for attention as the original paper for a fair comparison between ASI and AMMASI in the experiment. Furthermore, we gather the geometry information for 15 types of OpenStreetMap points of interest (POI), as depicted in the columns of Fig. 4. We also collect road network data from OSM as well. We applied the processed POI features $\mathbf{p}_i$ on both ASI and AMMASI[6].

5.2   Parameters

We have chosen an embedding dimension of $D = 64$ for the house feature embedding, the two attention embedding outputs, and the areal embedding. For areal embedding, we split the region into 100×100 grid, resulting in $M = 10,000$,

---

[5] https://github.com/darniton/ASI

[6] That is, we do not leverage the POI features in [10] as it was not publicly available.

and also illustrated in Fig. 2. In the case of two-stacked FCs, we utilize a D-dimensional (c.f. d-dimensional for f_q, f_k, f_v) output of each layer. The first layer employs the ELU activation function, while the second layer has no activation. The dataset is split into training, validation, and test sets with a ratio of 0.72, 0.08, and 0.2, respectively as the same setting of [10]. For the multi-head attention, we have set $d = 8$ and $K = 8$. We leverage Adam optimizers, batch size of 250, initial learning rate of 0.008, conduct early stopping on 10 patients, and reduce the learning rate to $1/10$ after 5 patients.

5.3   Error Measures

We evaluate the performance of the models by using three different error measures[7] described as $\mathrm{MALE} = \frac{1}{n} \sum_{i=1}^{n} |log(y_i) - log(\hat{y}_i)|$, $\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$, and $\mathrm{MdAPE} = \mathrm{median} \left(\frac{|y_1 - \hat{y}_1|}{|y_1|}, \frac{|y_2 - \hat{y}_2|}{|y_2|}, \ldots, \frac{|y_n - \hat{y}_n|}{|y_n|} \right)$.

6   Results

6.1   POI Feature Significance
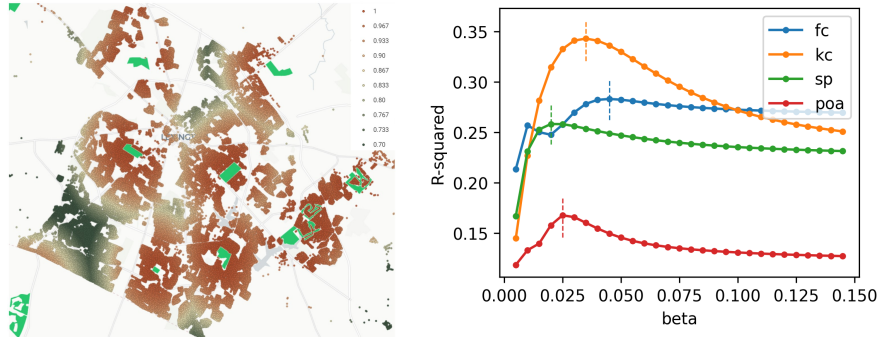
(a) Sample proximity heatmap (green < red) of each house to a POI (leisure-golf_course) on FC.

(b) R-squared value of linear regression with varying β: (best at FC: 0.045, KC: 0.035, SP: 0.020, POA: 0.025)

Fig. 3: Proximity-based POI feature extraction.

Gaussian parameter (β) for POI proximity feature In Fig. 3a shows a heatmap of proximity to a specific POI type. When calculating proximity, β decides how distant POIs will be assumed to be considered important, as high β value causes the proximity to consider POIs at greater distances as important. Although, it is true that β for each POI might be different, however, in this study we find the most optimal beta for each dataset region. In Fig. 3b, we empirically find optimal β as $\beta_{\mathrm{FC}} = 0.045$, $\beta_{\mathrm{KC}} = 0.035$, $\beta_{\mathrm{SP}} = 0.020$, and $\beta_{\mathrm{POA}} = 0.025$ which shows highest R^2 values in linear regression with $\mathbf{p}_i$ and y_i.

---

[7]   [10] incorrectly referred to MdAPE as MAPE, unlike their code implementation.

		amenity				landuse					leisure			natural		aero
		hospital	university	school	worship	cemetery	commercial	industrial	retail	railway	golf_course	park	sports_centre	water	wood	aerodrome
USA	F	-4.2	-1.8	10.0	-1.5	2.7	*0.5*	-17.1	5.5	8.0	*-1.0*	-8.5	-1.3	19.6	0.9	*-0.2*
USA	K	-0.2	0.9	1.7	0.7	-0.7	-0.9	-6.2	-0.9	-0.3	0.4	*-0.4*	1.3	-8.0	3.5	-0.8
Brazil	S	8.9	6.4	-31.7	-9.3	1.4	4.3	-9.7	-1.3	-1.0	0.7	-7.0	-3.4	-5.5	-4.8	1.0
Brazil	P	-8.9	*-1.5*	-120	*1.7*	11.9	10.9	3.8	10.6	-3.2	4.9	-21.4	2.4	27.7	-5.3	1.5

Fig. 4: Coefficients from linear regression for each POI proximity. Non-significant values (p-value > 0.05) are highlighted in gray.

Impact of different types of POIs. Fig. 4 illustrates the coefficients derived from ordinary least squares regression for $\mathbf{p}_i$ and y_i, excluding $\mathbf{f}_i$ to emphasize the generalized trend of POI proximity and its impact on house prices. Notably, industrial land use demonstrates a negative influence on house prices. The impact of waterfront proximity is pronounced in FC and POA datasets, as also indicated by the Fig. 2 heatmap. However, these observations often deviate from conventional wisdom. Surprisingly, parks do not show a significant enhancement in house values in any region. Additionally, commercial areas particularly have a positive influence in Brazil. Moreover, while the presence of schools positively influences in the USA, it exerts a notably strong negative influence in Brazil.

The results can be interpreted in two ways. Firstly, the observed heterogeneity in the impact of features could be ascribed to regional variations in culture, policies, and other human factors. This diversity proves valuable for comprehending the distinctive characteristics of each region. Secondly, the disparities in impact, even among POIs of the same type, might be attributable to a lack of data engineering and parameter tuning. This highlights the necessity of filtering out unimportant POIs for more nuanced approaches. Additionally, the possibility that different β values may be optimal for different POI types (i.e., $\beta_1, ..., \beta_\tau$) suggests opportunities for improvement.

6.2 Performance Comparision

Table 1 presents a performance comparison relative to baseline models. The baseline models include linear regression (LR), random forest (RF), XGBoost, and ASI [10]. Here, HA refers to using only House Attribute features, while HA+P indicates the inclusion of POI features with HA. Overall, our proposed AMMASI consistently shows the best performance in different datasets.

LR, RF, and XGBoost models demonstrate substantial improvement in predicting house prices when incorporating POI features along with house attributes. This underscores the significance of utilizing POI information, aligning with findings in existing research.

Table 1: Performance comparison. HA: House Attribute, HA+P: HA + POI. *
denotes the model leverages neighbor and similar house attention. Areal Embedding of AMMASI – FC & SP: Node2Vec, KC & POA: sinusoidal.

Model	LR		RF		XGBoost		ASI*		AMMASI*	
Feature	HA	HA+P	HA	HA+P	HA	HA+P	HA	HA+P	HA	HA+P
F MALE	0.228	0.210	0.197	0.113	0.184	0.117	0.098	0.097	0.097	**0.096**
RMSE	52084	48716	41791	27035	38736	26087	23177	**22678**	23451	22918
MAPE	15.60	14.58	13.17	7.14	12.59	8.00	6.49	6.29	**6.09**	6.12
K MALE	0.245	0.189	0.161	0.141	0.132	0.129	0.113	0.118	0.113	**0.112**
RMSE	246865	198921	174138	161233	145130	138499	115763	133543	108748	**104582**
MAPE	20.19	14.90	11.32	9.87	9.64	9.39	8.00	8.40	**7.82**	7.95
S MALE	0.271	0.227	0.243	0.153	0.241	0.160	0.135	0.144	0.135	**0.133**
RMSE	267317	235776	250505	170105	244170	172915	155585	162520	156099	**154483**
MAPE	23.12	18.91	19.72	11.34	19.73	12.60	9.80	10.82	9.56	**9.51**
P MALE	0.271	0.236	0.251	0.163	0.246	0.172	0.143	0.151	**0.139**	0.142
RMSE	154878	140237	147384	102851	139401	104024	94492	98615	**93246**	94522
MAPE	23.26	20.33	20.65	11.40	20.60	13.98	9.58	10.89	**8.89**	9.36

Moreover, AMMASI consistently outperforms ASI when we compare the best results from ASI and AMMASI on different datasets. However, both ASI and AMMASI show only marginal performance gains or even degradation when incorporating POI features. This can be attributed to the fact that ASI and AMMASI have access to information on prices of neighboring or similarly featured houses, minimizing the need for additional knowledge gained from including POI data. Furthermore, ASI frequently demonstrates a decline in performance with the addition of POI features compared to AMMASI, particularly in KC, SP, and POA. This suggests that the existing methods of integrating POI (which is concatenation) may easily lead to overfitting, while for AMMASI this negative effect is not very significant as the masked multi-head attention mechanism captures sophisticated feature understanding.

6.3   Parameter and Ablation Test

Threshold on masked attention (σ_G, σ_S). We empirically determine the optimal pair of σ_G and σ_S[8]. Initially, we identify the most effective σ_S values, wherein our model with areal embedding exhibits the least RMSE error. The parameter σ_S influences how attention is assigned to houses with similar features, potentially influenced by Euclidean distance measures.

Simultaneously, Fig. 5 illustrates the impact of σ_G on model performance. The parameter σ_G represents the location-based distance threshold for attention. The varying values of σ_G and σ_S for each region highlight the dataset's heterogeneity across different geographic areas.

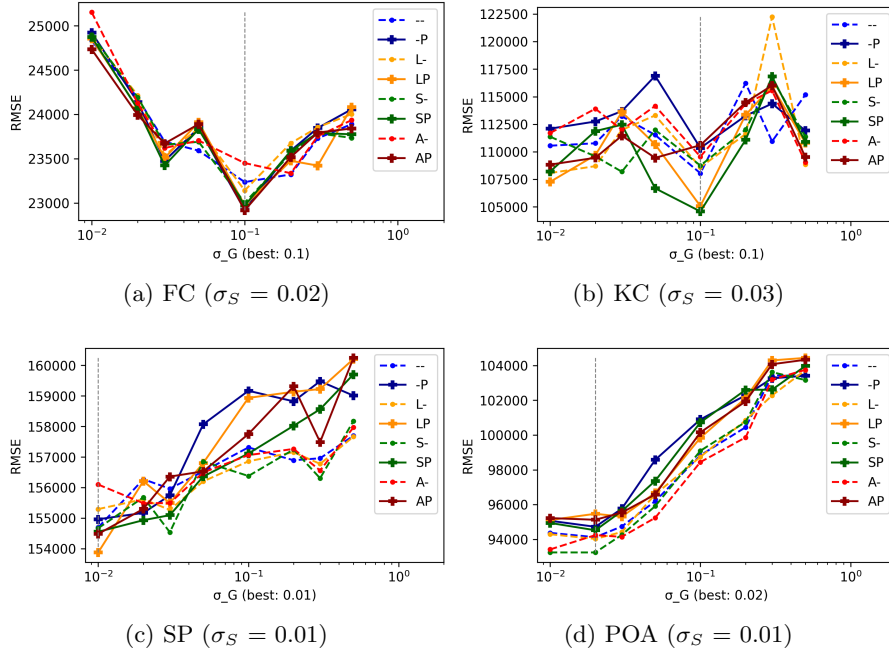[8] Testing every case of $(\sigma_G, \sigma_S) \in \{0.01, 0.02, 0.03, 0.05, 0.1, 0.2, 0.3, 0.5\}^2$

Fig. 5: Study on σ_G (with experimentally best σ_S) and an ablation study.

Comparison on Areal Embedding The analysis presented in Fig. 5 investigates the impact of feature inclusion or exclusion, with a specific focus on areal embedding and POI. Four primary approaches are considered: (–) None, (L) lat/lon as additional attributes, (S) sinusoidal embedding, and (A) Node2Vec embedding. Furthermore, for each case, we explore the scenarios with or without the inclusion of POIs (–, P), resulting in a total of eight cases.

The results indicate that our areal embedding approach is notably effective, especially on SP:S– and POA:A–. Furthermore, POI features show effectiveness in the FC and KC scenarios but have a lesser impact on SP and POA. Plus, Node2vec does not promise performance improvement over simple 2D sinusoidal. We attribute these observations to inaccuracies in the recorded latitude and longitude of houses and the comparatively commercial-area-centered urban planning in Brazil compared to the USA. Furthermore, we believe that an improvement in the granularity of area embedding could provide nuanced differences in embedding for houses.

7   Conclusion

In conclusion, this paper addresses key challenges in real estate appraisal, emphasizing the integration of Points of Interest (POI) and the implementation of Areal Embedding for enhanced property valuation. The introduced AMMASI model

surpasses current baselines by incorporating POI data and utilizing a simplified road network-based Areal Embedding approach with the masked multi-head attention. Notably, areal embedding opens avenues for integrating diverse urban datasets, including census and geographic spatial features as an extra channel, suggesting exciting possibilities for future research endeavors.

References

1. Bin, J., Gardiner, B., Li, E., Liu, Z.: Peer-dependence valuation model for real estate appraisal. Data-Enabled Discovery and Applications **3**, 1–11 (2019)
2. De Nadai, M., Lepri, B.: The economic value of neighborhoods: Predicting real estate prices from the urban environment. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA) (2018)
3. Du, W.W., Wang, W.Y., Peng, W.C.: Dora: Domain-based self-supervised learning framework for low-resource real estate appraisal. In: Proc. of the ACM International Conference on Information and Knowledge Management (2023)
4. Grover, A., Leskovec, J.: node2vec: Scalable feature learning for networks. In: Proc. of the ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 855–864 (2016)
5. Jenkins, P., Farag, A., Wang, S., Li, Z.: Unsupervised representation learning of spatial data via multimodal embedding. In: Proc. of the ACM International Conference on Information and Knowledge Management (2019)
6. Lee, H., Jeong, H., Lee, B., Lee, K.D., Choo, J.: St-rap: A spatio-temporal framework for real estate appraisal. In: Proc. of the ACM International Conference on Information and Knowledge Management (2023)
7. Li, C., Wang, W., Du, W., Peng, W.: Look around! A neighbor relation graph learning framework for real estate appraisal. In: Proc. the AAAI Conference on Artificial Intelligence, MuFin Workshop (2022)
8. Ottensmann, J.R., Payton, S., Man, J.: Urban location and housing prices within a hedonic model. Journal of Regional Analysis and Policy **38**(1) (2008)
9. Rosen, S.: Hedonic prices and implicit markets: product differentiation in pure competition. Journal of political economy **82**(1), 34–55 (1974)
10. Viana, D., Barbosa, L.: Attention-based spatial interpolation for house price prediction. In: Proc. of the International Conference on Advances in Geographic Information Systems (2021)
11. Wang, M.X., Lee, W.C., Fu, T.Y., Yu, G.: On representation learning for road networks. ACM Transactions on Intelligent Systems and Technology (TIST) **12**(1), 1–27 (2020)
12. Wang, Z., Liu, J.C.: Translating math formula images to latex sequences using deep neural networks with sequence-level training. International Journal on Document Analysis and Recognition (IJDAR) **24**(1-2), 63–75 (2021)
13. Woźniak, S., Szymański, P.: Hex2vec: Context-aware embedding h3 hexagons with openstreetmap tags. In: Proc. of the ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery (2021)
14. Xiao, Y., Chen, X., Li, Q., Yu, X., Chen, J., Guo, J.: Exploring determinants of housing prices in beijing: An enhanced hedonic regression with open access poi data. ISPRS International Journal of Geo-Information **6**(11), 358 (2017)
15. Zhang, W., Liu, H., Zha, L., Zhu, H., Liu, J., Dou, D., Xiong, H.: Mugrep: A multi-task hierarchical graph representation learning framework for real estate appraisal. In: Proc. of the ACM SIGKDD on Knowledge Discovery & Data Mining (2021)