Enhancing Geospatial House Price Prediction in Greater Jakarta Using XGBoost and ResNet18 Feature Fusion

Alif Biuti Anastasya[1] , Hadi Santoso[2*]

[1,2] *Faculty of Computer Science, Study program of Informatics, Universitas Mercu Buana, Indonesia*

*corr-author: hadi.santoso@mercubuana.ac.id

Abstract - Precise house price predictions play a vital role in shaping housing policies and informing investment decisions in urban regions such as Greater Jakarta, encompassing Jakarta, Bogor, Depok, Tangerang, and Bekasi. Most models rely exclusively on structured data, ignoring spatial and environmental factors that influence property prices. This study proposes a multimodal machine learning framework that integrates structured property data with Sentinel-2 satellite imagery within a geospatial context. The baseline dataset consists of 17 tabular variables. The ResNet-18 algorithm extracts visual environmental information from satellite imagery. The integration of both modalities through a late fusion strategy results in improved predictive performance. The baseline XGBoost achieved R^2 scores of 0.81 (log scale) and 0.79 (Rupiah), with an error of about 184 million Rupiah. The image-only model achieved an R^2 of 0.43, indicating moderate explanatory capability. Late fusion further improved performance, achieving R^2 values of 0.94 (log scale) and 0.93 (Rupiah), while reducing prediction error by over 40%.

Keywords: House price prediction; XGBoost; ResNet-18; Sentinel-2 satellite imagery; multimodal analysis.

I. INTRODUCTION

Predicting house prices is a central topic in urban and economic analysis because property values reflect broader socioeconomic dynamics, infrastructure cycles, and spatial planning priorities [1]. In large metropolitan metropolitan regions such as Jakarta, Bogor, Depok, Tangerang, and Bekasi (Jabodetabek), price formation is shaped by location, accessibility, public facilities, and environmental quality [2], and accurate prediction supports market actors as well as local and national policy decisions[3]. Early studies relied on hedonic and multiple regression over structured attributes (e.g., land area, building area, room counts) [4], which provide interpretability but are limited in capturing nonlinear interactions and spatial heterogeneity [5]. Ensemble approaches like Random Forest, Gradient Boosting, and Extreme Gradient Boosting (XGBoost) have established good baselines for housing appraisal due to their resistance to multicollinearity and complicated feature interactions [6-8]. Ensembles consistently achieve higher R^2 and lower MAE/RMSE than classical regression on diverse urban datasets [9, 10], and remain competitive across regions and data regimes [11, 12].

Beyond structural attributes, urban prices are systematically influenced by environmental and locational context green/blue space, density, road networks, and amenity proximity [13-15]. Multi-source geo-data and explainable ML have shown that environmental quality and access can explain a non-trivial share of price variation and appreciation [16, 17]. Remote sensing (RS) and GIS now enable continuous observation of land-use/land-cover (LULC) and spatial heterogeneity at the city scale, improving real-estate analytics and planning[18-21].

Parallel to tabular modeling, computer vision (CV) and deep learning (DL) leverage imagery to encode neighborhood desirability. Convolutional Neural Networks (CNNs) learn visual features of vegetation, texture, building patterns from street-view and satellite data that correlate with prices and perceptions (e.g., greenness, openness, walkability) [22-26]. Image-only models can already explain a meaningful portion of price variance and provide complementary signals to tabular features [22, 25]. Deep learning approaches using pre-trained CNN models such as ResNet and DenseNet have demonstrated effectiveness in visual feature extraction [27-29], providing a foundation for multimodal integration. Recent works, therefore, move toward multimodal fusion, combining structured variables with visual representations to capture both physical and perceptual dimensions of place [24, 30 , 31]. Late fusion designs train each modality separately before combining, which improves generalization and interpretability and often yields 8–12% gains over single-modality baselines [25, 30]. In parallel, unsupervised representations from

satellite imagery help delineate sub-markets and capture intra-urban scale effects relevant to price formation [32]. Despite worldwide development, Indonesian studies have mainly focused on tabular models without using satellite or street-view signals [12]. Open data such as Sentinel-2 (10 m) can enrich models with environmental and infrastructural context while remaining practical for local deployment [19-21].

This study proposes a multimodal, late fusion methodology for Greater Jakarta that uses XGBoost for tabular predictors and ResNet-18 for visual feature extraction from Sentinel-2 imagery. By including structural and spatial information, the strategy intends to increase forecast accuracy and interpretability, giving data driven evidence to support urban housing policy and valuation practice in Greater Jakarta [14, 24, 30, 31].

II. METHOD

A. Research Design

This study predicts housing values in Greater Jakarta using a supervised machine learning framework that integrates tabular data and Sentinel-2 satellite imagery. XGBoost models tabular data and ResNet-18 captures visual features, integrated through late fusion. The process includes data prep, training, fusion, and evaluation, as shown in Fig. 1.

B. Data Collection

This study uses two primary data sources: tabular property data and Sentinel-2 satellite imagery.

1) Tabular Data (Internal): The dataset from Public Appraisal Service Office (KJPP) Sapto, Kasmodiard, and Partners includes residential property valuation reports from 2020 to March 2025, covering variables like property price, land and building area, rooms, floors, electricity, certificate, condition, and geographic coordinates. All data were anonymised and used under legal consent.

2) Satellite Image Data (Public): Google Earth Engine obtained Sentinel-2 Level-2A images (European Space Agency). RGB bands (B4, B3, B2) at 10 m resolution were used. Cloud masking was applied, and surrounding areas cropped into 192 × 192 pixel images representing built-up density, vegetation, and roads. Latitude and longitude linked each image to property records.

C. Preprocessing, Data Exploration, and Visualization

Preprocessing was conducted separately for both data modalities.

1) Tabular Data: Duplicate records were removed, missing values imputed, and outliers eliminated with IQR. Categorical variables were one-hot encoded, and numeric variables standardized using Z-scores. New features like building_land_ratio, total_area, and price_per_m² were created. EDA included correlation heatmaps, price distribution, and geospatial plots to explore feature relationships.

2) Satellite Imagery: Sentinel-2 Level-2A imagery for 2021-2024 was obtained from Google Earth Engine. The Scene Classification Layer (SCL) filtered out cloud, cirrus, and shadow pixels to ensure cloud robustness. Only scenes with less than 20% cloud probability within the crop region were valid. When multiple suitable scenes existed, the closest to the appraisal date within ±3 months was chosen for temporal consistency. If no cloud-free scenes were available, a median composite from nearby dates minimized noise from illumination and atmospheric differences. Since physical features like roads, vegetation, roofs, and urban density change slowly, the composite represents the built environment during 2020-2025. Each property coordinate was cropped into a 192 × 192 RGB patch for feature extraction.

D. Modeling

The model was developed in three scenarios:

1) Tabular Data Based Model (XGBoost): An XGBoost regressor predicted log_price from features, optimizing parameters via Grid Search with 10-fold cross-validation.

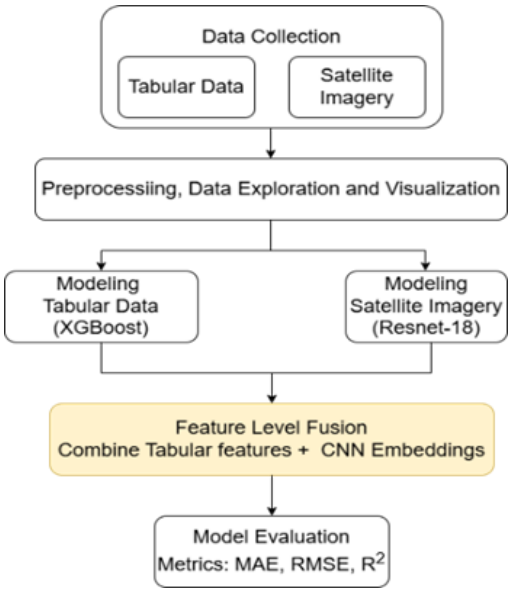

Fig. 1 Research design

2) Satellite Image Based Model (ResNet-18): A ResNet-18 model pre-trained on ImageNet was employed to extract visual features from each 192 × 192 Sentinel-2 RGB patch. The final layer of the model was restructured into a linear output unit, allowing the network to predict log_price rather than class labels. During training, the convolutional backbone was kept frozen, and only the final layers were fine-tuned using Huber Loss and the Adam optimizer with a learning rate of 1×10^{-4}. Despite ImageNet's collection of natural RGB images, its initial convolutional filters, which detect edges, textures, and spatial gradients, transfer well to medium-resolution satellite imagery. The final 512-dimensional embedding, obtained via global average pooling (GAP), served as the feature representation for both the image-only model and the multimodal fusion stage.

3) Multimodal Fusion: The fusion mechanisms, including early fusion and late fusion, are described in Section II.E (Feature Fusion).

E. Feature Fusion

Two fusion techniques were assessed. The early fusion baseline used a single XGBoost regressor trained on combined features, concatenating normalized tabular data with the 512-dimensional ResNet-18 embedding from global average pooling, forcing the model to learn from heterogeneous features. The late fusion method trained the tabular model (XGBoost) and image model (ResNet-18) separately, then concatenated their outputs to a secondary XGBoost regressor for final prediction.

F. Model Evaluation

The model's accuracy and robustness were evaluated with three common metrics: mean squared error (MSE), root mean squared error (RMSE), and the coefficient of determination (R^2).

The mean squared error (MSE) quantifies the average of the squared differences between actual and predicted values, as shown in (1):

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(yi - \hat{y}i)^2 \quad (1)$$

Furthermore, the root mean squared error (RMSE) is employed to quantify the prediction error in the same unit as the target variable, as expressed in (2):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - \hat{y}i)^2} \quad (2)$$

Additionally, the coefficient of determination (R^2) is used to assess how much of the variance in the dependent variable can be predicted from the independent variables, as shown in (3):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (3)$$

where yi represents the actual value, $\hat{y}i$ is the predicted value, $\bar{y}_i$ is the average actual value, and n is the number of data samples.

III. RESULTS AND DISCUSSION

A. Dataset Characteristics and Preprocessing

This study uses a dataset of 1,730 residential properties from the Greater Jakarta area, covering Jakarta, Bogor, Depok, Tangerang, and Bekasi, obtained from professional property valuation reports by KJPP Sapto, Kasmodiard, and Partners from 2020 to 2025, covering the COVID-19 pandemic years. According to certified appraisers, the pandemic mainly affected market liquidity and transaction volume, not property values. Since the data are from appraisal reports rather than listing or transaction prices, they better reflect long-term fundamentals. The data are combined with Sentinel-2 Level-2A satellite images from Google Earth Engine, where each property matches a Sentinel-2 RGB image (10 m resolution) based on coordinates. The full set of features is summarized in Table I.

Table I shows all dataset attributes, but only predictive and target-related variables are used for model training. Table II presents a sample of the first five filtered records.

After preprocessing, the remaining 17 tabular variables used for model training are listed in Table III.

Table III presents the final set of 17 tabular features used as model inputs. These features encompass structural attributes, facility I ndicators, spatial distances, and engineered ratios that describe each residential property. Each property entry is paired with a Sentinel-2 RGB image cropped from its geographic coordinates, offering visual context of the surrounding environment. The images correspond to specific properties in the dataset, with the number on the image (e.g., ID 17, 25, and 47) indicating the property ID linked to the tabular record. These examples illustrate the variety in residential environments captured by Sentinel-2, from densely built-up areas to mixed vegetation and residential zones. Selected samples are shown in Fig. 2.

B. Exploratory Analysis and Data Visualization

An exploratory analysis examined the distribution, variability, and relationships among the main numerical features and log_price. It also identified outliers and assessed variable suitability for regression. Table IV shows descriptive statistics of the key features.

TABLE I
DESCRIPTION OF TABULAR AND SATELLITE IMAGE FEATURES

Category	Feature Name	Type / Unit	Description
Main Features	ID, latitude, longitude, land_area, building_area, price	Integer / Numeric (m², IDR, WGS84)	Basic identifiers and physical attributes of each property, including coordinates, size, and appraised price.
Structural & Facility	certificate, floor, water_supply, water_well, road_width	Categorical / Numeric (meter)	Building characteristics and facility indicators such as ownership type, number of floors, water source, and access road width.
Spatial Features	dist_bus_stop, dist_campus, dist_market, dist_hospital, dist_school, dist_station	Numeric (km)	Euclidean distance of each property to nearby public facilities and transportation nodes.
Derived Features	building_land_ratio, price_per_m2, total_area, log_price	Numeric (ratio, IDR/m², log)	Engineered variables derived from main attributes to improve model stability and interpretability.
Satellite Imagery	Sentinel-2 RGB Patch (192×192, Bands B4–B3–B2 @10 m)	Image (RGB)	Visual representation of the surrounding environment linked by property ID.

TABLE II
EXAMPLE OF THE FIRST FIVE RECORDS (SELECTED COLUMNS)

id	latitude	longitude	land_area	building_area	certificate	floor	road_width	price
1	-6.348136	106.732040	128.0	96.0	1	0	8.0	940800000
2	-6.317640	106.692040	162.0	119.0	1	1	8.0	252580000
3	-6.169644	106.500482	119.0	55.0	1	0	8.0	715100000
4	-6.202304	106.655564	90.0	57.0	1	0	8.0	1060500000
5	-6.264736	106.672007	72.0	96.0	1	1	6.0	113280000

TABLE III
TABULAR FEATURES USED IN THE MODEL

Features					
latitude	building_area	water_supply	dist_campus	dist_school	building_land_ratio
longitude	certificate	water_well	dist_market	dist_bus_stop	total_area
land_area	floor	road_width	dist_hospital	dist_station	

Fig. 2 Example Sentinel-2 RGB

TABLE IV
DESCRIPTIVE STATISTICS OF MAIN NUMERICAL FEATURES

Feature	Min	Maximum	Mean	Std. Deviation
Land Area (m²)	26.0	265.00	104.64	41.85
Building Area (m²)	20.00	230.00	88.46	42.46
Road Width (m)	0.80	22.00	6.51	1.80
Price (IDR)	139,170,000	2,796,840,000	1,054,453,154	579,648,514
Building–Land Ratio	0.11	3.53	0.89	0.39
log_price	18.75	21.75	20.63	0.56

Table IV shows significant variation in land and building sizes across Greater Jakarta, with wide property price dispersion. The log_price transformation stabilizes regression analysis, though large properties cause high-value outliers in size features. Building_area (r = 0.71) and land_area (r = 0.54) are the main price determinants, with road_width having a moderate positive effect. The negative correlation between building_land_ratio and land_area indicates typical suburban patterns. These are visualized in Fig. 3, showing the correlation matrix of key numerical variables.

C. Tabular Data Modeling

Tabular modeling was performed using six regression algorithms: Linear Regression, Ridge, Lasso, Random Forest, Gradient Boosting, and XGBoost. The 10-fold cross-validation (CV) evaluation results are presented in Table V.

1) Best Model Evaluation: The tuning process used RandomizedSearchCV (10-fold) to identify the optimal parameter combinations for learning_rate, max_depth, n_estimators, subsample, and colsample_bytree. The best configuration achieved a cross-validation score of $R^2 = 0.8335$. Further evaluation on the test data, which comprised 20% of the total dataset, is shown in Table VI. Performance was assessed using three standard regression metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). These metrics are computed as in Table VI.

The R^2 value of 0.8108 in the log domain shows that the model accounts for about 81.1% of the variation in house prices in unseen test data. When converted to the Rupiah scale, the model's performance remains stable with $R^2 = 0.7872$, MAE = IDR 184 million, and RMSE = IDR 265 million. The slight difference between the cross-validation score ($R^2 = 0.8335$) and the test-set result suggests that the model maintains strong generalization capability with minimal indication of overfitting.

2) Feature Importance Analysis: Feature importance analysis highlights key factors influencing house prices. The top three are total_area, building_area, and floor, representing main physical traits. Road_width, longitude, and land_area have moderate impact, while environmental and accessibility factors also contribute, as shown in Fig. 4.

Fig. 5 shows the relationship between actual prices and model predictions, with points mostly around the diagonal line (y = x). The evaluation metrics $R^2 = 0.787$, MAE = IDR 184 million, and RMSE = IDR 265 million confirm strong predictive accuracy across property values.

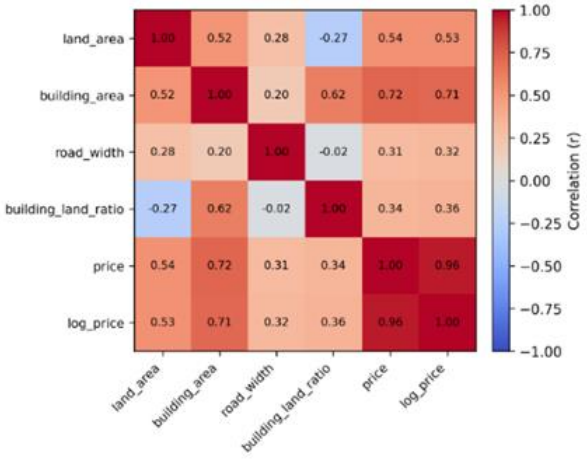

Fig. 3 Correlation heatmap among selected numerical features

TABLE IV
CROSS-VALIDATION PERFORMANCE (R^2 ON LOG_PRICE)

Model	CV_R^2_Mean	CV_R^2_Std
XGBoost	0.831	0.0342
Gradient Boosting	0.8043	0.0396
Random Forest	0.7959	0.0418
Lasso Regression	0.6946	0.0383
Linear Regression	0.6947	0.0385
Ridge Regression	0.6948	0.0385

TABLE VI
TEST-SET EVALUATION OF THE TUNED XGBOOST MODEL (LOG AND RUPIAH SCALES)

Scale	MAE	RMSE	R^2
log(price)	0.1786	0.2365	0.8108
IDR	184,311,730	264,803,908	0.7872

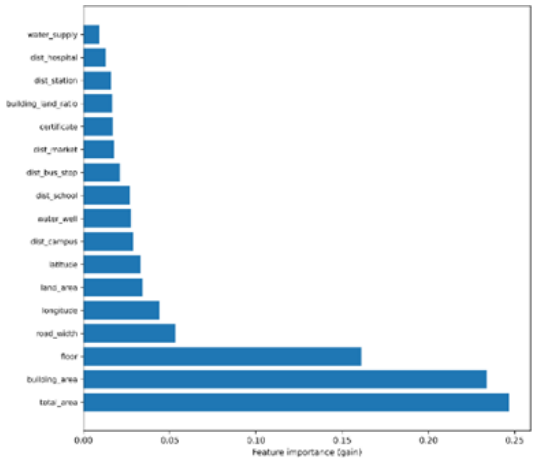

Fig. 4 Feature importance of the XGBoost model

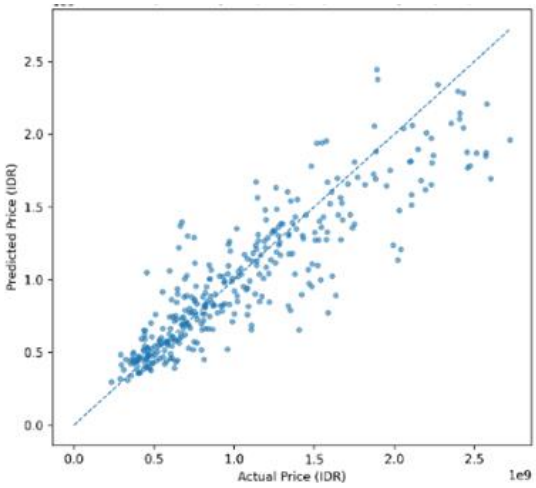

Fig. 5 Actual vs predicted house prices

D. Satellite Image Data Modeling

Image modeling used a pre-trained ResNet-18 (ImageNet) for house price regression. Sentinel-2 images (192×192 RGB) underwent normalization, augmentation (RandomResizedCrop 0.85–1.0), and batch training. The model was trained with Huber Loss, Adam optimizer, ReduceLROnPlateau scheduler, and an 80/20 data split. Performance is in Table VII.

As shown in Table VII, the pure image model yielded $R^2 = 0.46$, indicating about 46% of house price variation can be explained by the visual environment. Although less accurate than the tabular model, the image-based approach remains relevant because satellite imagery captures spatial characteristics such as building density, vegetation coverage, and road patterns that influence property values.

In this study, the ResNet-18 model emphasizes macro-scale visual features in Sentinel-2 imagery, such as built-up density, road patterns, and vegetation coverage. These cues serve as proxies for neighborhood structure, accessibility, and environmental quality, which are not always fully captured by tabular data. Satellite imagery thus provides supportive contextual information for the overall prediction when combined with structural and locational features.

E. Multimodal Fusion Model

The multimodal model combines the 512-dimensional ResNet-18 embedding with the tabular feature vector, and the fused representation is then processed by an XGBoost stacker as the final regressor. The performance comparison of all evaluated models is presented in Table VIII.

Table VIII shows the late fusion model has the highest accuracy, with $R^2 = 0.9425$ in the log domain and $R^2 = 0.9289$ on the Rupiah scale. The early fusion baseline, which concatenates all standardized features with a 512-dimensional ResNet-18 embedding and trains one XGBoost regressor, performs worse ($R^2 = 0.7673$). Late fusion effectively combines structural and spatial environmental information, surpassing early fusion or single modality models.

TABLE VII
EVALUATION OF THE IMAGE-ONLY RESNET-18 MODEL (LOG AND RUPIAH SCALES)

Scale	MAE	RMSE	R^2
log(price)	0.3147	0.3948	0.4585
IDR	294,543,392	393,132,586	0.4261

TABLE VIII
PERFORMANCE COMPARISON OF ALL MODELS

Model	MAE (log)	RMSE (log)	R^2 (log)	MAE (IDR)	RMSE (IDR)	R^2(IDR)
XGBoost (Tabular only)	0.1786	0.2365	0.8108	184,311,730	264,803,908	0.7872
ResNet-18 (Image only)	0.3147	0.3948	0.4585	294,543,392	393,132,586	0.4261
Early Fusion	0.1886	0.2569	0.7673	192,645,614	290,026,935	0.7335
Late Fusion	0.1022	0.1277	0.9425	106,032,776	149,734,214	0.9289

F. Qualitative Error Analysis

A qualitative spatial analysis identified where the baseline model has the largest prediction errors. Fig. 6 shows heatmaps of absolute prediction error: purple for low, green for moderate, yellow for high, and orange for hotspots.

Fig. 6 (a) shows a tabular model with high error regions, especially in dense urban areas like Tangerang, South Jakarta, Depok, and Bekasi, due to heterogeneous property conditions and complex environments. In contrast, Fig. 6(b) shows that adding satellite image features reduces these errors, making the high-error regions smaller, weaker, and more localized. This indicates the fusion model effectively corrects systematic spatial errors, especially in highly urbanized regions.

G. Limitations of the Study

The dataset includes 1,730 residential properties, which is relatively small for deep learning models like ResNet-18. While this issue is partly addressed by using ImageNet pre-trained weights and data augmentation, the limited data size still restricts the ability to learn detailed visual features. Additionally, Sentinel-2 imagery at 10 m resolution captures broader environmental context but lacks the micro-level detail needed to depict specific property conditions. These limitations partially explain the performance gap between the image-only model (R^2 = 0.46) and the tabular model (R^2 = 0.81), suggesting that satellite imagery primarily serves as complementary contextual information rather than a standalone predictor.

IV. CONCLUSION

This study developed a multimodal house price prediction framework for Greater Jakarta by combining tabular property data with Sentinel-2 satellite imagery. The tabular XGBoost model performed strongly (R^2 = 0.81 on the log scale), while the image-only ResNet-18 model captured environmental patterns with moderate accuracy (R^2 = 0.46). When both modalities were integrated through late fusion, prediction accuracy improved substantially, achieving the best results (R^2 = 0.94 on the log scale) and reducing MAE and RMSE by more than 40% compared with single-modality models. Qualitative spatial analysis further revealed that multimodal fusion reduces high-error hotspots in dense urban areas, indicating that satellite imagery provides contextual information absent from tabular features. These findings demonstrate that integrating structural attributes with environmental visual cues yields more robust and spatially consistent house price predictions. While this study focuses on Greater Jakarta, the proposed framework is conceptually transferable to other urban regions. However, model performance may vary across cities with different environmental characteristics, and region-specific retraining would be required to maintain accuracy. Future work may incorporate higher-resolution imagery, larger study areas, or end-to-end deep fusion models to further enhance generalizability.

ACKNOWLEDGEMENT

The author acknowledges the Public Appraisal Services Office (KJPP) Sapto, Kasmodiard, and Partners for providing access to the residential property appraisal data used in this study. The author also acknowledges the European Space Agency (ESA) for granting open access to Sentinel-2 satellite imagery via Google Earth Engine (GEE) platform. Appreciation is extended to the academic supervisor and colleagues in the Informatics Engineering Study Program at Universitas Mercu Buana for their guidance, technical support, and constructive feedback throughout the research process.

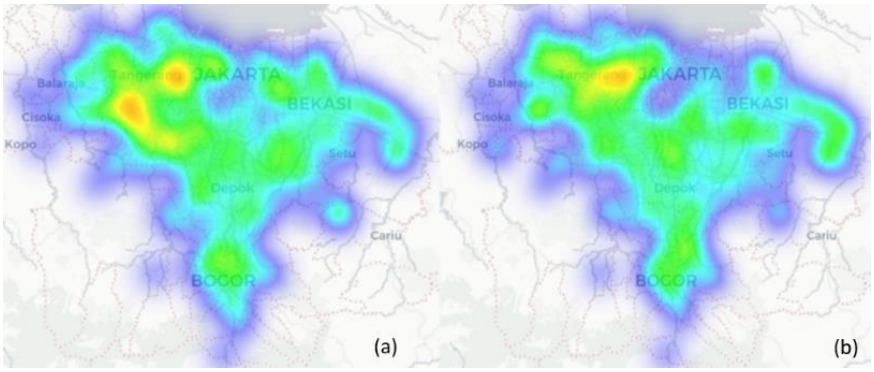

Fig. 7 Spatial error heatmaps for (a) the tabular only model and (b) the fusion model

REFERENCES

[1] R. Cellmer and K. Kobylińska, "HOUSING PRICE PREDICTION - MACHINE LEARNING AND GEOSTATISTICAL METHODS," *Real Estate Management and Valuation*, vol. 33, no. 1, pp. 1–10, Mar. 2025, doi: 10.2478/remav-2025-0001.

[2] L. Chen, X. Yao, Y. Liu, Y. Zhu, W. Chen, X. Zhao and T. Chi, "Measuring impacts of urban environmental elements on housing prices based on multisource data-a case study of Shanghai, China," ISPRS Int J Geoinf, vol. 9, no. 2, 2020, doi: 10.3390/ijgi9020106.

[3] X. Xu and Y. Zhang, "House price forecasting with neural networks," *Intelligent Systems with Applications*, vol. 12, p. 52, 2021, doi: 10.1016/j.iswa.2021.20.

[4] A. Kumar, "Recommendation of Regression Models for Real Estate Price Prediction using Multi-Criteria Decision Making," *Journal of Communications Software and Systems*, vol. 19, no. 3, pp. 220–229, 2023, doi: 10.24138/jcomss-2023-0102.

[5] I. Forys, "Machine learning in house price analysis: regression models versus neural networks," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 435–445. doi: 10.1016/j.procs.2022.09.078.

[6] T. Quang, N. Minh, D. Hy, and M. Bo, "Housing Price Prediction via Improved Machine Learning Techniques," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 433–442. doi: 10.1016/j.procs.2020.06.111.

[7] A. B. Adetunji, O. N. Akande, F. A. Ajala, O. Oyewo, Y. F. Akande, and G. Oluwadara, "House Price Prediction using Random Forest Machine Learning Technique," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 806–813. doi: 10.1016/j.procs.2022.01.100.

[8] H. Sharma, H. Harsora, and B. Ogunleye, "An Optimal House Price Prediction Algorithm: XGBoost," *Analytics*, vol. 3, no. 1, pp. 30–45, Mar. 2024, doi: 10.3390/analytics3010003.

[9] Z. Liu, "Real Estate Price Prediction based on Supervised Machine Learning Scenarios," *Highlights in Science, Engineering and Technology CMLAI*, vol. 39, 2023.

[10] N. H. Zulkifley, S. A. Rahman, N. H. Ubaidullah, and I. Ibrahim, "House price prediction using a machine learning model: A survey of literature," *International Journal of Modern Education and Computer Science*, vol. 12, no. 6, pp. 46–54, 2020, doi: 10.5815/ijmecs.2020.06.04.

[11] J. Avanija, G. Sunitha, K. Reddy Madhavi, P. Kora, R. Hitesh, and S. Vittal E A Associate, "Prediction of House Price Using XGBoost Regression Algorithm," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 2, pp. 2151–2155, Apr. 2021.

[12] R. Tanamal, N. Minoque, T. Wiradinata, Y. Soekamto, and T. Ratih, "House Price Prediction Model Using Random Forest in Surabaya City," *TEM Journal*, vol. 12, no. 1, pp. 126–132, Feb. 2023, doi: 10.18421/TEM121-17.

[13] H. Wang, Y. Hu, L. Tang, and Q. Zhuo, "Distribution of urban blue and green space in beijing and its influence factors," *Sustainability (Switzerland)*, vol. 12, no. 6, Mar. 2020, doi: 10.3390/su12062252.

[14] S. Chen, D. Zhuang, and H. Zhang, "GIS-Based Spatial Autocorrelation Analysis of Housing Prices Oriented towards a View of Spatiotemporal Homogeneity and Nonstationarity: A Case Study of Guangzhou, China," *Complexity*, vol. 2020, 2020, doi: 10.1155/2020/1079024.

[15] C. Xue, Y. Ju, S. Li, Q. Zhou, and Q. Liu, "Research on accurate house price analysis by using gis technology and transport accessibility: A case study of xi'an, china," *Symmetry (Basel).*, vol. 12, no. 8, pp. 1–21, Aug. 2020, doi: 10.3390/sym12081329.

[16] Y. Kang, F. Zhang, W. Peng, S. Gao, J. F. Duarte and C. Ratti, "Understanding house price appreciation using multi-source big geo-data and machine learning," Land use policy, vol. 111, Dec. 2021, doi: 10.1016/j.landusepol.2020.104919.

[17] S. Jin, H. Zheng, N. Marantz, and A. Roy, "Understanding the effects of socioeconomic factors on housing price appreciation using explainable AI," *Applied Geography*, vol. 169, Aug. 2024, doi: 10.1016/j.apgeog.2024.103339.

[18] M. Li and A. Stein, "Mapping land use from high resolution satellite images by exploiting the spatial arrangement of land cover objects," *Remote Sens. (Basel).*, vol. 12, no. 24, pp. 1–21, Dec. 2020, doi: 10.3390/rs12244158.

[19] T. Zhang, J. Su, Z. Xu, Y. Luo, and J. Li, "Sentinel-2 satellite imagery for urban land cover classification by optimized random forest classifier," *Applied Sciences (Switzerland)*, vol. 11, no. 2, pp. 1–17, Jan. 2021, doi: 10.3390/app11020543.

[20] L. Liu, C. Zhang, W. Luo, S. Chen, F. Yang, and J. Liu, "New remote sensing image fusion for exploring spatiotemporal evolution of urban land use and land cover," *J. Appl. Remote Sens.*, vol. 16, no. 03, Aug. 2022, doi: 10.1117/1.jrs.16.034527.

[21] B. Wiatkowska, J. Słodczyk, and A. Stokowska, "Spatial-temporal land use and land cover changes in urban areas using remote sensing images and gis analysis: The case study of opole, poland," *Geosciences (Switzerland)*, vol. 11, no. 8, Aug. 2021, doi: 10.3390/geosciences11080312.

[22] C. Zhao, Y. Ogawa, S. Chen, T. Oki, and Y. Sekimoto, "Quantitative land price analysis via computer vision from street view images," *Eng. Appl. Artif. Intell.*, vol. 123, Aug. 2023, doi: 10.1016/j.engappai.2023.106294.

[23] X. Xu, "Associations Between Street-View Perceptions and Housing Prices: Subjective vs. Objective Measures Using Computer Vision and Machine Learning Techniques," *Remote Sens. (Basel).*, vol. 14, no. 4, Feb. 2022, doi: 10.3390/rs14040891.

[24] L. Yang, Z. Chen, and G. Li, "Satellite Image Price Prediction Based on Machine Learning," *Remote Sens.*

(Basel)., vol. 17, no. 12, Jun. 2025, doi: 10.3390/rs17121960.

[25] A. Nouriani and L. Lemke, "Vision-based housing price estimation using interior, exterior & satellite images," *Intelligent Systems with Applications*, vol. 14, p. 81, 2022, doi: 10.1016/j.iswa.2022.20.

[26] J. Chougale, A. Shinde, N. Deshmukh, D. Sawant, and V. Latke, "House Price Prediction using Machine learning and Image Processing," Journal of University of Shanghai for Science and Technology, vol. 23, no. 6, Jun. 2021.

[27] Adhinata, F. D. Fitriana, G. F. Wijayanto, and A. Putra, "Corn Disease Classification Using Transfer Learning and Convolutional Neural Network," *JUITA: Jurnal Informatika*, vol. 9, No. 2, Nov. 2021, doi: 10.30595/juita.v9i2.11686.

[28] F. D. Adhinata, D. P. Rakhmadani, M. Wibowo, and A. Jayadi, "A Deep Learning Using DenseNet201," *JUITA: Jurnal Informatika*, vol. 9, No. 1, May 2021, doi: 10.30595/juita.v9i1.9624.

[29] A. Nur, A. Thohari, L. Triyono, I. Hestiningsih, B. Suyanto, and A. Yobioktobera, "Performance Evaluation of Pre-Trained Convolutional Neural Network Model for Skin Disease Classification," *JUITA: Jurnal Informatika*, vol. 10, No. 1, May 2022, doi: 10.30595/juita.v10i1.12041.

[30] P. Y. Wang, C. T. Chen, J. W. Su, T. Y. Wang, and S. H. Huang, "Deep learning model for house price prediction using heterogeneous data analysis along with joint self-attention mechanism," *IEEE Access*, vol. 9, pp. 55244–55259, 2021, doi: 10.1109/ACCESS.2021.3071306.

[31] O. Poursaeed, T. Matera, and S. Belongie, "Vision-based real estate price estimation," *Mach. Vis. Appl.*, vol. 29, no. 4, pp. 667–676, May 2018, doi: 10.1007/s00138-018-0922-2.

[32] G. E. Kenyon, D. Arribas-Bel, and C. Robinson, "Extracting Features from Satellite Imagery to Understand the Size and Scale of Housing Sub-Markets in Madrid," *Land (Basel).*, vol. 13, no. 5, May 2024, doi: 10.3390/land13050575