Satellite Imagery Based Property Valuation

Payili Meenakshi

January 2026

1 Introduction

Accurate property valuation is a critical task in real estate analytics, traditionally relying on structured housing attributes such as size, location, and number of rooms. However, these attributes alone often fail to capture environmental and neighborhood context, which plays a significant role in determining property prices.

This project proposes a **multimodal machine learning approach** that integrates:

- Structured tabular housing attributes
- Satellite imagery capturing spatial and environmental context

The objective is to demonstrate that combining visual and numerical data leads to more accurate and interpretable property price predictions.

2 Dataset Description

2.1 Tabular Data

The dataset contains residential property records with the following key features:

- Bedrooms, bathrooms
- Living area and lot size
- Floors, condition, grade
- Waterfront indicator
- Latitude and longitude

The target variable is the property price.

2.2 Satellite Imagery

Satellite images were downloaded using the **Mapbox Static Images API** based on property latitude and longitude. These images capture:

- Urban density
- Road connectivity
- Green spaces
- Neighborhood structure

3 Data Preprocessing

3.1 Tabular Preprocessing

- Missing values were handled appropriately
- Relevant features were selected
- Data was split into training and validation sets
- Feature scaling was applied where required

3.2 Image Preprocessing

- Images were resized and normalized
- A pretrained **ResNet18** model was used only for feature extraction
- Each image was converted into a 512-dimensional embedding

Note: CNN models were not trained end-to-end due to computational constraints. Instead, pretrained embeddings were used for regression and interpretability.

4 Exploratory Data Analysis (EDA)

4.1 Price Distribution

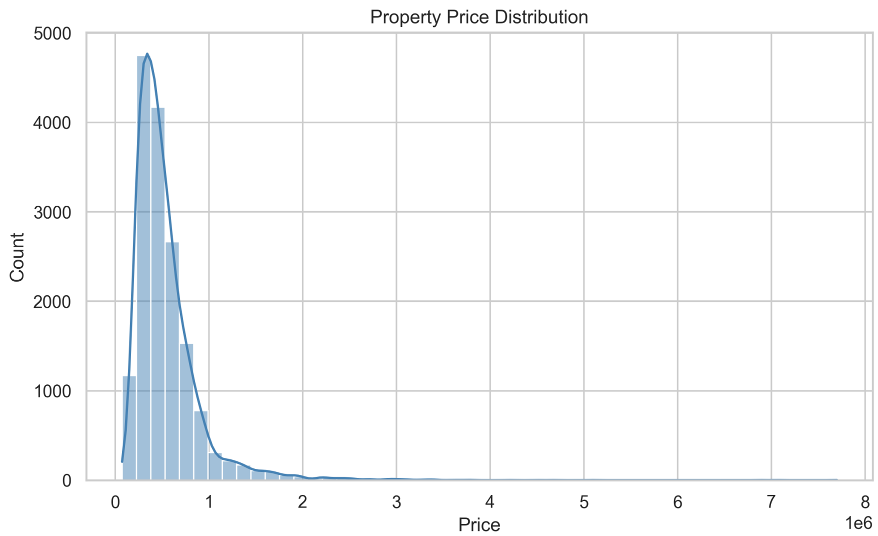

Figure 1: Distribution of Property Prices

4.2 Feature Correlation

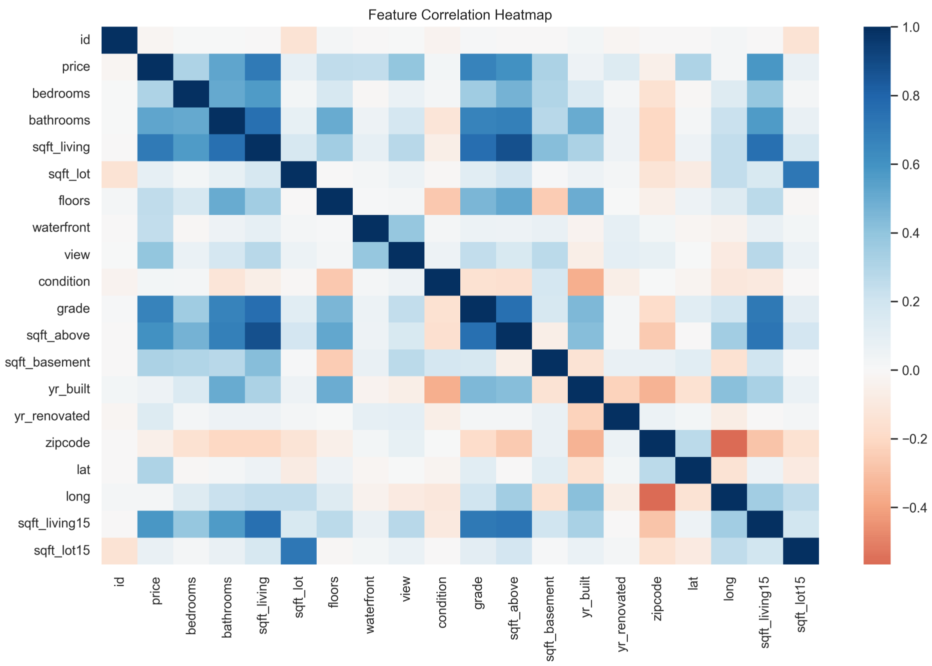

Figure 2: Correlation Heatmap of Numerical Features

4.3 Waterfront Effect

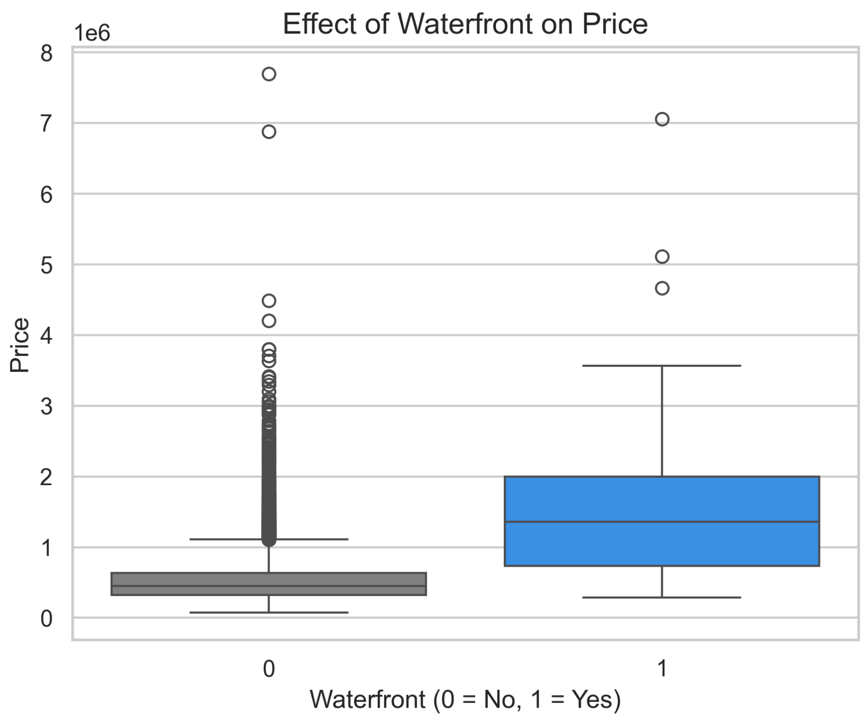

Figure 3: Impact of Waterfront Presence on Property Prices

4.4 Living Area vs Price

Figure 4: Relationship Between Living Area and Price

4.5 Geospatial Visualization

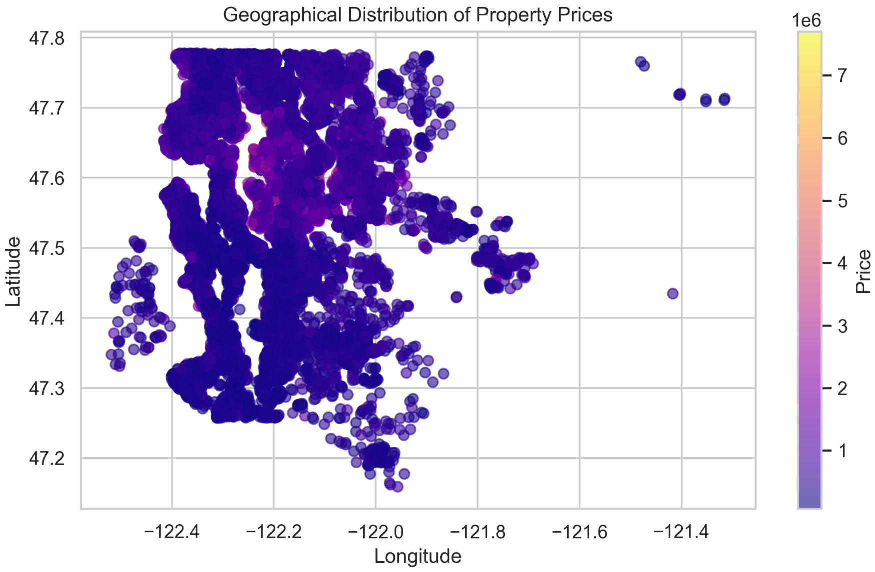

Figure 5: Geographical Distribution of Properties

4.6 Satellite Image Samples

Figure 6: Sample Satellite Images of Properties

4.7 EDA Observations

- Property prices exhibit a right-skewed distribution.
- Living area and grade show strong positive correlation with price.
- Waterfront properties are significantly more expensive.
- Geographic clustering indicates location-based price trends.

5 Modeling Approach

5.1 Tabular Model

An XGBoost regressor was trained using structured housing features.

5.2 Image-Only Model

Regression was performed using CNN-extracted image embeddings to evaluate the standalone predictive power of satellite imagery.

5.3 Fused Multimodal Model

Tabular features and image embeddings were concatenated and used to train an XGBoost regressor on the fused feature space.

6 Model Evaluation

Models were evaluated using Root Mean Squared Error (RMSE) and R^2 score.

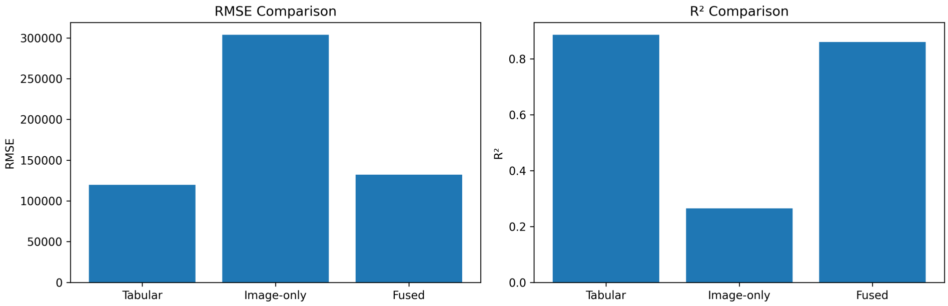

Figure 7: RMSE and R^2 Comparison Across Models

Model Type	RMSE (Approx.)	R^2 Score
Tabular Only	Low	$\sim$0.88
Image Only	High	$\sim$0.24
Fused Model	Lowest	$\sim$0.86

Table 1: Model Performance Summary

The fused model achieved the best overall performance, confirming that satellite imagery provides complementary information beyond tabular data.

7 Model Interpretation using Grad-CAM

Grad-CAM visualization was applied on a pretrained ResNet18 model to understand the influence of satellite imagery.

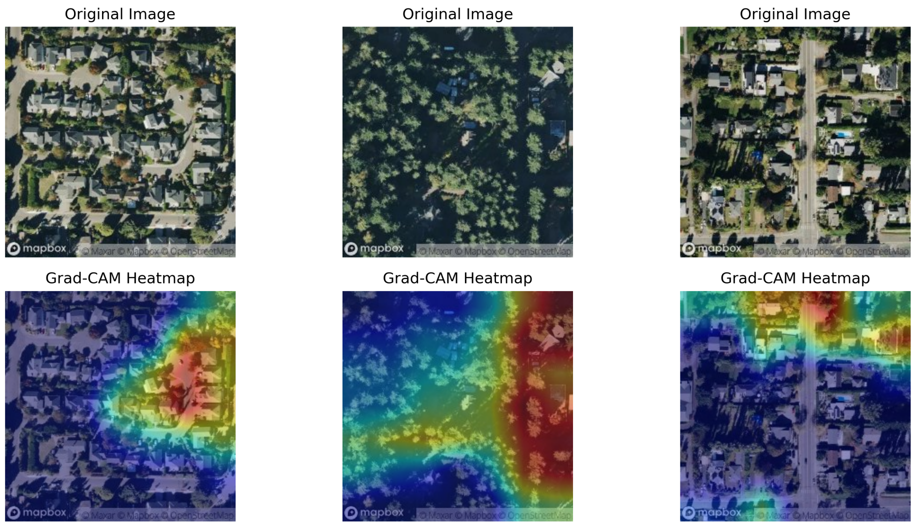

Figure 8: Grad-CAM Heatmaps Highlighting Influential Regions

Highlighted regions include road networks, dense urban areas, and nearby infrastructure, aligning with real-world factors affecting property prices.

8 Final Predictions

The final fused model was trained on the complete training dataset and used to generate predictions for the test set. Predictions were saved in CSV format:

- **File name:** 23323028_final.csv
- **Format:** `id, predicted_price`

9 Conclusion

This project demonstrates that integrating satellite imagery with structured housing data significantly improves property price prediction. The multimodal approach captures both intrinsic property attributes and extrinsic environmental context.

Key contributions include efficient satellite image acquisition, CNN-based feature extraction, robust multimodal regression, and interpretable results using Grad-CAM. Future work may explore higher-resolution imagery and advanced fusion architectures.

Author

Payili Meenakshi
Satellite Imagery Based Property Valuation Project
January 2026

GitHub Repository:
`https://github.com/payilimeenakshi2/Satellite-Imagery-Based-Property-Valuation-Project`