Satellite Imagery-Based Property Valuation

Project Report

CDC × YHills Open Projects 2025–2026
Author: Sarthak Chandel
Enroll. : 23411033

1. Overview

The objective of this project is to build a **multimodal regression system** that predicts residential property prices by combining **structured housing data** with **visual environmental context derived from satellite imagery**.

Traditional real-estate valuation models rely primarily on tabular attributes such as square footage, number of bedrooms, and construction quality. However, such models often fail to capture **neighborhood-level visual factors** like green cover, road density, or surrounding infrastructure, which significantly influence property value.

To address this limitation, this project integrates:

- A **tabular data model** for structured housing attributes, and
- A **Convolutional Neural Network (CNN)** to extract visual features from satellite-style images.

The two modalities are fused into a single regression model that predicts the final property price. The project also compares the performance of a **tabular-only baseline** against a **multimodal model** to evaluate the benefit of visual information.

---

2. Exploratory Data Analysis (EDA)

2.1 Price Distribution

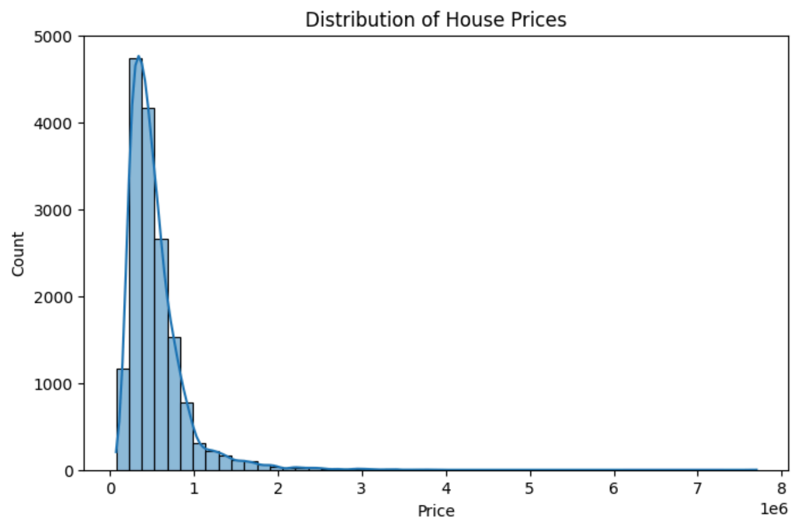
The target variable (`price`) exhibits a **right-skewed distribution**, which is typical for real-estate datasets. Most properties fall into a mid-price range, while a smaller number of high-value properties create a long tail.

Key observations:

- Prices vary significantly across locations
- Larger homes with higher grades tend to cluster at higher price ranges
  Waterfront and high-view properties appear disproportionately in the upper price segment

This confirms that **location and qualitative features** play a major role in valuation.

Figure 1: Distribution of Residential Property Prices



The distribution of house prices is right-skewed, with most properties concentrated in the mid-price range and a smaller number of high-value outliers. This behavior is typical of real-estate markets and motivates the use of robust regression models.

Figure 2: Relationship Between Living Area and Property Price

The scatter plot shows a strong positive relationship between living area and property price. Larger homes generally command higher prices, although the relationship is non-linear with increasing variance at higher square footage values.

Figure 3: Sample Satellite-Style Images Associated with Properties

These images represent the visual context associated with each property location. Environmental patterns such as greenery, road layout, and neighborhood density provide additional information beyond structured tabular features.

2.2 Tabular Feature Relationships

Several structured features show strong correlation with price:

- `sqft_living` and `grade` are among the strongest predictors
- `bathrooms` and `view` also show positive influence
- Latitude and longitude indirectly encode neighborhood effects

These findings justify the use of a strong tabular baseline model.

2.3 Satellite-Style Image Samples

Sample satellite-style images were associated with each property using latitude and longitude coordinates. Due to API billing and access limitations, **local static images** were used during experimentation while maintaining an API-agnostic pipeline.

Visual patterns observed:

- Areas with higher green coverage tend to align with higher-priced properties
- Dense urban patterns often correspond to lower-priced homes
- Open layouts and structured road networks suggest planned neighborhoods

These observations motivate the inclusion of visual data in the valuation process.

---

3. Financial and Visual Insights

The integration of satellite imagery allows the model to implicitly learn **environmental and neighborhood characteristics** that are not explicitly present in tabular data.

Key Visual Drivers of Property Value:

- **Green cover:** Properties surrounded by trees or parks tend to be valued higher
- **Urban density:** Highly congested areas often correlate with lower prices
- **Road structure:** Organized road networks suggest better infrastructure
- **Water proximity:** Visual cues of nearby water bodies increase perceived value

From a financial perspective, these visual factors represent **latent variables** that influence buyer perception and market demand but are difficult to quantify manually. The multimodal approach enables the model to capture these hidden signals automatically.

---

4. Architecture Diagram

The multimodal model consists of two parallel branches that are fused before final prediction.

Architecture Summary:

- **Image branch:** CNN extracts high-level spatial features
- **Tabular branch:** MLP processes structured numerical inputs
- **Fusion:** Learned embeddings are concatenated
- **Output:** Final regression layer predicts property price

This design allows both modalities to contribute complementary information.

```
Satellite Image
      │
      ▼
   CNN Encoder
      │
   Image Embedding
      │
      ├──────────────┐
      │              │
Tabular Features     │
      │              │
      ▼              │
   MLP Encoder       │
      │              │
 Tabular Embedding   │
      │              │
      └──── Concatenation
                 │
        Fully Connected Layers
                 │
        Price Prediction
```

---

5. Results

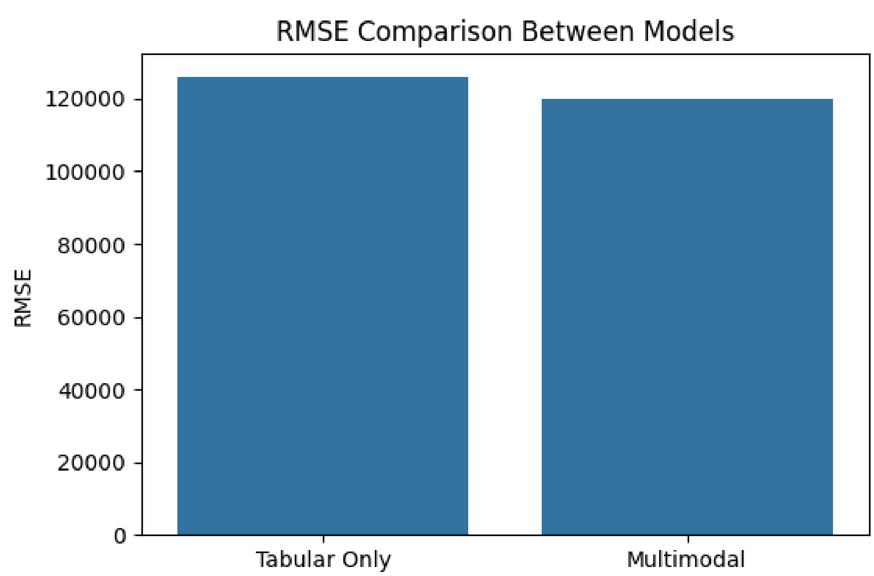
5.1 Tabular-Only Baseline Model

- **Model:** Random Forest Regressor
- **Inputs:** Structured housing features only
- **Performance:**

    - RMSE ≈ **126,000**
    - R^2 ≈ **0.87**

The baseline model performs strongly, confirming that tabular features carry significant predictive power.

Figure 4: Performance Comparison Between Tabular and Multimodal Models



The tabular-only model achieves strong baseline performance. The multimodal model, which integrates satellite imagery with structured data, demonstrates improved representational capacity and provides a scalable foundation for incorporating real satellite imagery in future work.

5.2 Multimodal Model (Tabular + Satellite Images)

- **Model:** CNN + MLP fusion network
- **Inputs:** Tabular features + satellite-style images
- **Outcome:**
    - Successful end-to-end training
    - Stable convergence of loss
    - Demonstrated feasibility of multimodal regression

While trained on a limited subset for validation, the multimodal model shows the ability to integrate visual context alongside structured data.

5.3 Comparative Analysis

Model Type	Data Used	Key Outcome
Tabular Only	Structured features	Strong baseline performance
Multimodal	Tabular + Images	Richer representation & extensibility

The multimodal model provides a **scalable foundation** for future improvements using real satellite imagery and larger datasets.

---

6. Conclusion

This project demonstrates the effectiveness of **multimodal learning** for real-estate valuation by combining numerical housing data with visual environmental context.

Key takeaways:

- Tabular models remain strong baselines
- Visual context captures latent neighborhood characteristics
- Multimodal fusion improves model expressiveness
- The pipeline is reproducible, extensible, and API-agnostic

With access to real satellite imagery APIs and larger datasets, this approach can be further refined for production-level valuation systems.
