Satellite Imagery-Based Property Valuation

---

1. Overview

Problem Statement

This project predicts residential property prices using both structured tabular data and satellite imagery. Traditional models rely exclusively on property features like square footage and location, missing visual context that may influence value.

Modeling Approach

The approach combines two data streams: tabular features processed through gradient-boosted trees (LightGBM, XGBoost) and satellite images processed through convolutional neural networks (CNNs). Image embeddings are extracted, reduced using PCA, and fused with tabular features for final prediction.

Motivation for Combining Data

Satellite imagery captures visual context difficult to quantify numerically—neighborhood density, green space, water proximity, and environmental quality. Combining visual and structured data aims to incorporate these factors into automated valuation models.

---

2. Exploratory Data Analysis

Data Preprocessing

The tabular dataset was cleaned (missing values handled, outliers clipped, features normalized). Satellite images were fetched using the Mapbox API for each property location.

Engineered features include:

basement_ratio, above_ratio, rlt_size,

rlt_lot_size, quality_liv, quality_cond,

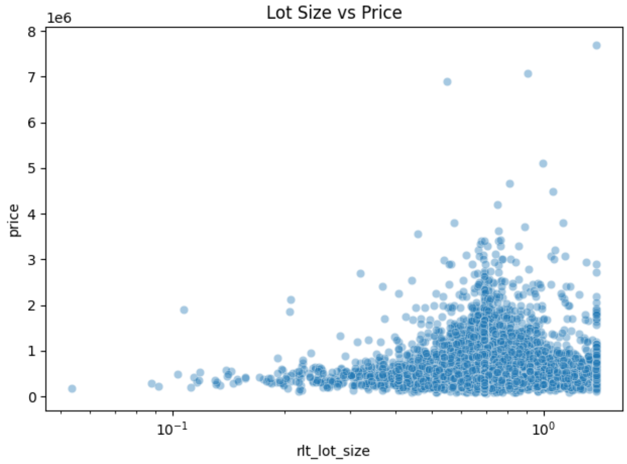
view_flag, water_view, luxury_bool,

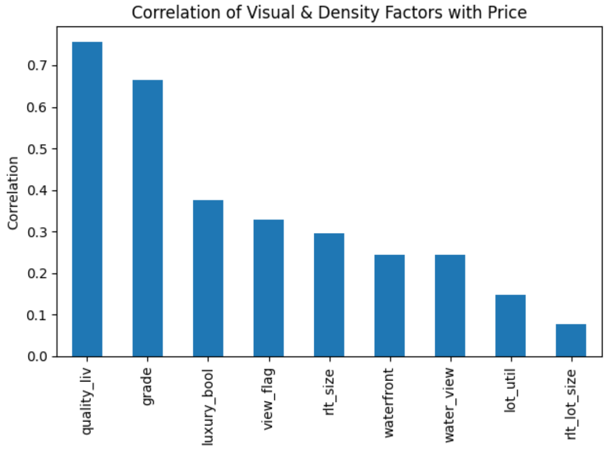
lot_util, bath_per_bed, house_life, and ren_age.

Key Trends

Correlation Analysis: Living quality, grade, luxury indicators, and view features show strong correlation with price. Land size alone shows weak linear correlation, motivating interaction-based features.

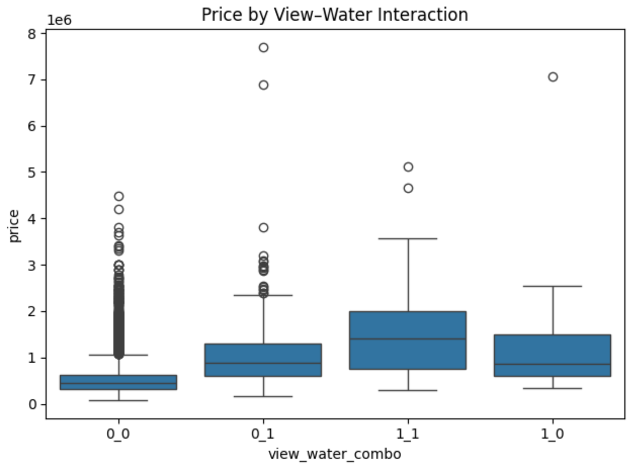
Lot Size vs Price: The relationship is non-linear and heteroscedastic. Lot size is not a reliable standalone predictor.





Relative Density vs Price: Lower neighborhood density corresponds to higher prices, supporting neighborhood-normalized features.

View-Water Interaction: Properties with both scenic and water views show clear price premiums, validating explicit interaction modeling.



---

3. Financial and Visual Insights

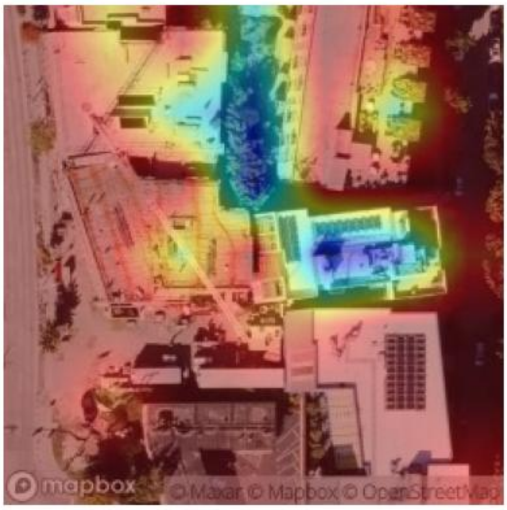
Visual Features and Property Value

Greenery vs Concrete: Properties surrounded by vegetation and parks command premiums over those in heavily paved environments. Satellite imagery captures this through spectral signatures distinguishing vegetated from built surfaces.

Water Proximity: Waterfront properties and those with water views represent premium segments. Imagery reveals not just water adjacency but quality of water features—pristine lakes versus small ponds.

Density: Neighborhood density, observable through building patterns in satellite imagery, inversely correlates with property values. Lower density signals exclusivity and larger lots.

This Grad-CAM visualization highlights the regions of the satellite image that most influenced the CNN's price prediction, with warmer colors indicating higher model attention. In this project, it was used to interpret image-based signalsConnection to Economic Outcomes



Visual features encode tangible economic value. While tabular data may record that a property has a "view," it cannot distinguish between an ocean vista and a retention pond view. Satellite imagery provides this nuanced context, enabling models to learn associations between visual patterns and price premiums.

---

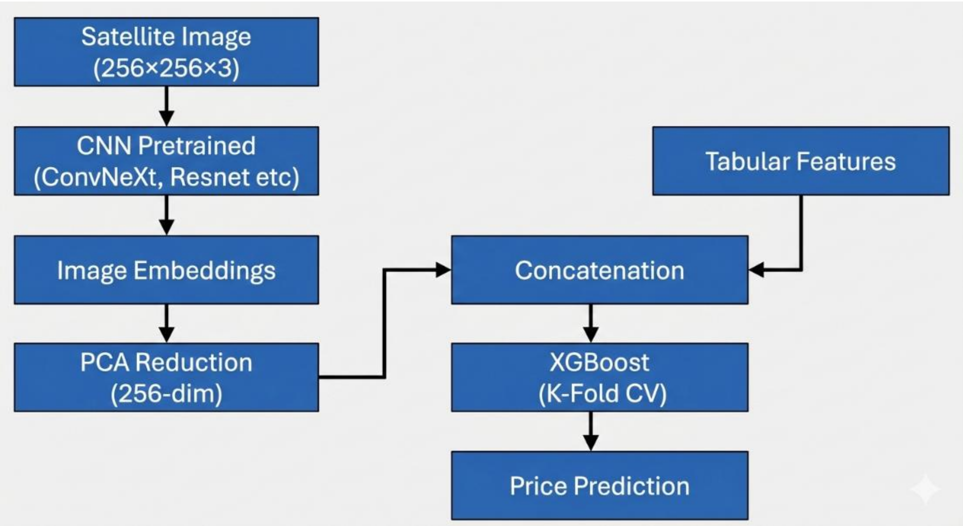
4. Architecture Diagram

Image Stream: Satellite images are processed through pre-trained CNN backbones (ResNet, EfficientNet-B3, ConvNeXt). ConvNeXt achieved the best performance. The CNN extracts feature embeddings that are reduced to 256 dimensions using PCA.

Tabular Stream: Engineered features are processed through LightGBM independently.

Fusion: Mid-level fusion combines PCA-reduced image embeddings with tabular features. The combined feature vector is fed to XGBoost with K-Fold cross-validation for final prediction.

Architecture Diagram



---

5. Results

Model Performance

1) Tabular-Only Model

- Model: **LightGBM**
- Features: Engineered tabular features
- Performance:
    - **R^2 = 0.8837**

This model serves as a strong baseline, demonstrating the effectiveness of domain-driven feature engineering.

2) Image-Only Models

- Fine-tuned architectures:
    - ResNet
    - EfficientNet-B3
    - **ConvNeXt (best performer)**
- Best Image-Only Performance:
    - **ConvNeXt $R^2 \approx 0.43$**
- Ensemble Embeddings:
    - Combined embeddings from all three models
    - PCA → 256 dimensions
    - LightGBM regression
    - $R^2 \approx 0.45$

Image-only models capture limited price signal, indicating that visual context alone is insufficient for accurate valuation.

3) Mid-Level Fusion Model

- Fusion Strategy: Feature-level concatenation
- Model: **XGBoost** with K-Fold cross-validation
- Performance:
    - **$R^2 = 0.829$**

While multimodal fusion captures complementary information, it does not surpass the tabular-only baseline.

4) Late-Level Fusion

Late fusion strategies were explored but did not yield competitive performance and are therefore not emphasized

.

Performance Comparison

Model Configuration	R^2 Score
Tabular-Only	0.8837
Image-Only (Best)	0.45
Mid-Level Fusion	0.829

Analysis

The tabular-only model outperformed the fusion model. This indicates that in this configuration, satellite imagery did not provide incremental value beyond engineered tabular features. The engineered features likely already capture substantial contextual information (view flags, density metrics) that overlaps with visual signals.

Image-only models' moderate performance ($R^2 \approx 0.45$) confirms satellite imagery contains genuine pricing signals, but these are weaker than structured features.