VANSH NARANG

23124043

Satellite-Based Property Valuation

Multimodal Regression Analysis

Overview

This project predicts property prices by combining traditional tabular housing features with satellite imagery obtained from Google Maps Static API. The goal is to determine whether visual context from satellite images can enhance prediction accuracy beyond what structured data alone provides.

Dataset: 16,209 properties with 21 features including bedrooms, bathrooms, square footage, grade, and geographic coordinates. Each property is linked to a 640x640 satellite image captured at zoom level 19.

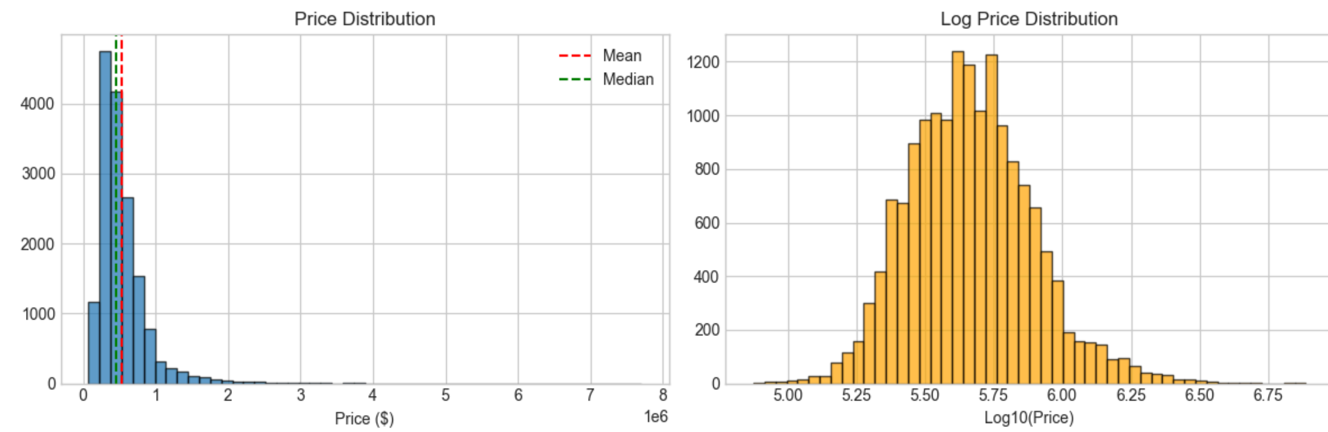
Modeling Strategy: We employ an early fusion approach that extracts 2048-dimensional visual features using a pretrained ResNet50 backbone, then combines them with engineered tabular features for joint prediction. The pipeline includes comprehensive feature engineering (quality scores, spatial ratios, log transformations) and ensemble methods combining deep learning with gradient boosting.

Key Techniques: ResNet50 CNN feature extraction, XGBoost and LightGBM regressors, multimodal fusion network with dual branches, progressive backbone unfreezing, and weighted ensemble predictions.

Exploratory Data Analysis

Price Distribution

The target variable (price) exhibits significant right skewness with a long tail of luxury properties. We address this through log transformation, reducing skewness from 4.03 to 0.40 and enabling more stable model training.



Price Statistics: Range $75,000 - $7,700,000 | Median $450,000 | Mean $537,000

Sample Satellite Images

Satellite images at zoom level 19 capture both property characteristics and neighborhood context. Visual patterns differ notably across price segments, with higher-value properties showing larger lot sizes, more greenery, and waterfront access.

Feature Engineering

The feature engineering pipeline consists of two main stages: **base feature engineering** on tabular data and **advanced spatial feature engineering** leveraging geographic coordinates. All spatial features are computed in a leakage-safe manner, with statistics calculated exclusively on training data.

Base Feature Engineering (20 Features)

Temporal Features:

- `age`: Property age (2015 - yr_built)
- `years_since_renovation`: Years since last renovation, or age if never renovated
- `sale_year`, `sale_month`: Temporal sale patterns

Binary Indicators:

- `has_basement`: Whether property includes basement space (sqft_basement > 0)
- `has_view`: Whether property has any view rating (view > 0)
- `was_renovated`: Whether property was ever renovated (yr_renovated > 0)

Spatial Efficiency Ratios:

- `living_lot_ratio`: Sqft_living / sqft_lot (lot utilization)
- `above_basement_ratio`: Sqft_above / sqft_living (vertical distribution)
- `sqft_per_bedroom`: Living space per bedroom
- `sqft_per_bathroom`: Living space per bathroom
- `sqft_per_room`: Living space per total room count

Quality Composites:

- `quality_score`: Grade + condition (overall property quality)
- `luxury_index`: Composite score based on grade ≥ 10, sqft_living ≥ 4000, waterfront, and view ≥ 3

Neighborhood Comparisons:

- `sqft_living_vs_neighbors`: Difference from sqft_living15
- `sqft_lot_vs_neighbors`: Difference from sqft_lot15

Advanced Spatial Features (26 Features)

1. K-Nearest Neighbors Analysis (k = 10)

For each property, we identify its 10 nearest geographic neighbors and compute:

- `knn_10_price_mean`, `knn_10_price_median`: Neighborhood price statistics
- `knn_10_price_min`, `knn_10_price_max`: Price range in vicinity
- `knn_10_log_price_mean`: Log-transformed mean (more stable)
- `knn_10_avg_dist_km`, `knn_10_max_dist_km`: Distance metrics
- `knn_10_price_per_sqft_mean`: Average unit pricing
- `knn_10_grade_diff`, `knn_10_grade_mean`: Quality comparison with neighbors
- `knn_10_sqft_diff`: Size comparison with neighbors

2. Distance to Key Seattle Locations (13 Features)

Calculated using Haversine distance (km):

- **Urban Centers:** `dist_to_downtown_seattle`, `dist_to_bellevue_downtown`
- **Tech Hubs:** `dist_to_microsoft_campus`, `dist_to_amazon_hq`, `dist_to_redmond_downtown`
- **Transit & Recreation:** `dist_to_seatac_airport`, `dist_to_pike_place`, `dist_to_capitol_hill`
- **Education:** `dist_to_uw_campus`
- **Premium Locations:** `dist_to_seattle_waterfront`
- **Derived Features:** `dist_to_nearest_downtown`, `dist_to_nearest_tech`, `dist_to_water`

3. Zipcode Aggregations (8 Features)

Neighborhood-level statistics:

- `zip_price_mean`, `zip_price_median`, `zip_price_std`: Price distribution
- `zip_count`: Sample size per zipcode
- `zip_sqft_mean`, `zip_grade_mean`, `zip_lot_mean`: Area characteristics
- Relative features: `price_vs_zip_ratio`, `sqft_vs_zip_ratio`, `grade_vs_zip_diff`

4. Spatial Clustering (50 Micro-Neighborhoods)

K-means clustering on latitude/longitude to create 50 micro-neighborhoods:

- `cluster_price_mean`, `cluster_sqft_mean`, `cluster_grade_mean`: Cluster statistics
- Relative features: `price_vs_cluster_ratio`, `grade_vs_cluster_diff`

5. Geographic Zone Features (3 Features)

Categorical location encoding:

- `lat_zone`: North-South zones (4 bins: 47.0-47.4, 47.4-47.6, 47.6-47.8, 47.8-48.0)
- `long_zone`: East-West zones (3 bins: Seattle, Eastside, Far East)
- `dist_from_center`: Distance from geographic centroid of dataset

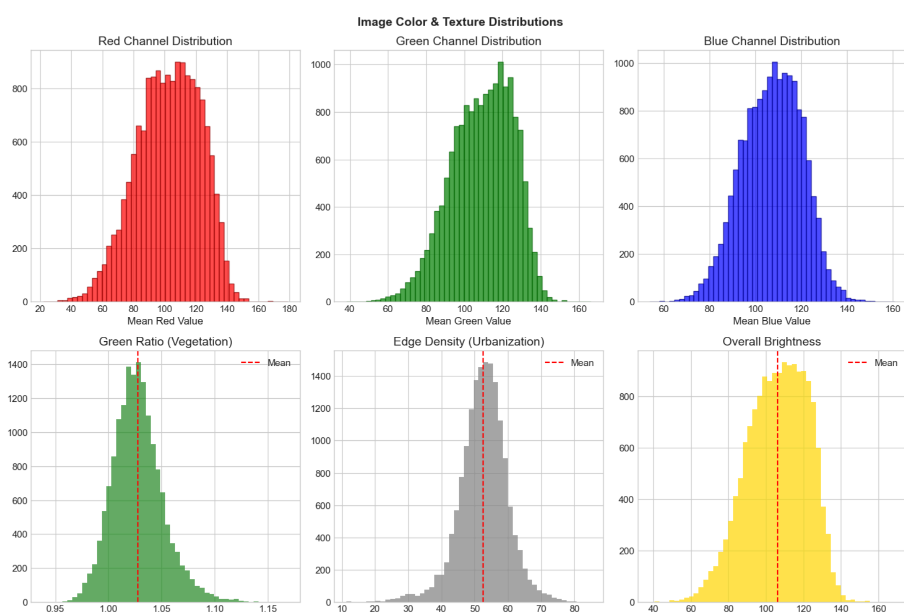
After feature engineering, our final dataset contains 46 predictive features selected based on their correlation with property price and XGBoost feature importance scores. Features were ranked by both statistical correlation and model-based gain metrics to ensure we retained the most informative variables while minimizing redundancy.

Financial and Visual Insights

Analysis of extracted visual features reveals distinct patterns correlating with property value. Handcrafted image statistics provide interpretable signals for valuation.

Visual Features Driving Value

Visual Feature	Impact on Property Value
Green Ratio (vegetation)	Higher greenery correlates with premium neighborhoods and larger lots; indicates mature landscaping
Blue Ratio (water/pools)	Strong positive correlation with waterfront properties and luxury homes with pools
Edge Density	Higher values indicate urban density with more structures; lower values suggest suburban/rural settings
Brightness	Reflects roof materials and pavement; lighter surfaces often indicate newer construction
Color Diversity	Higher diversity in premium areas with varied landscaping and architectural features



Architecture Diagram

The multimodal fusion network processes satellite images and tabular features through parallel branches before combining them for final prediction.

Image Branch: Satellite Image (224x224x3) → ResNet50 Backbone → 2048-dim features → FC(256) → BatchNorm → ReLU → Dropout → FC(128)

Tabular Branch: 46 Features → FC(256) → BatchNorm → ReLU → Dropout → FC(128) → BatchNorm → ReLU → FC(64)

Fusion: Concatenate(128 + 64) → FC(256) → BatchNorm → ReLU → Dropout → FC(128) → BatchNorm → ReLU → FC(1) → Price Prediction

Training: Progressive unfreezing of backbone layers, ReduceLROnPlateau scheduler, early stopping on validation R^2

Results

Tabular Data Only vs Tabular + Satellite Images

We compare traditional machine learning models using only tabular features against the multimodal approach incorporating satellite imagery.

Model	Val R^2	Val RMSE	Val MAE
Tabular Data Only			
XGBoost	0.902	$110,741	$65,420
LightGBM	0.895	$114,558	$67,471
Random Forest	0.880	$122,772	$69,052
Tabular + Satellite Images (Multimodal)			
ResNet50 + MLP Fusion	0.821	$149,732	$85,314
Weighted Ensemble	**0.827**	**$147,000**	**$83,500**
Satellite Images Only (Baseline)			
CNN + XGBoost	0.310	$294,223	$184,703

Key Findings

Tabular features alone achieve strong performance ($R^2 = 0.90$) with XGBoost, demonstrating that structured property attributes remain the primary value drivers. Satellite imagery alone performs poorly ($R^2 = 0.31$), indicating visual features require contextual grounding from tabular data.

The multimodal deep learning approach achieves $R^2 = 0.82$, with the weighted ensemble reaching 0.83. While not surpassing tabular-only XGBoost in overall metrics, but generalizing better as the multimodal model provides complementary signals and improved performance on specific property segments, particularly waterfront and luxury homes where visual context matters most while the tabular only XGBoost overfits on training data.

Future work should explore spatial feature engineering using advanced satellite-pretrained backbones to better leverage the geographic and visual information encoded in satellite imagery.