Satellite Imagery-Based Property Valuation 🏠

Comprehensive Project Report with Visual Analytics

📋 Executive Summary

This report presents a sophisticated multimodal machine learning pipeline that predicts residential property values by integrating tabular housing data with high-resolution satellite imagery. The project demonstrates that **structured property attributes are the primary value drivers**, with satellite imagery providing contextual enrichment rather than predictive improvement.

Key Metrics:

- **Best Model**: XGBoost (Tabular Only)
- **R^2 Score**: 0.8530 (85.3% variance explained)
- **RMSE**: $125,284.92
- **Dataset**: 16,208 training properties, 5,404 test properties
- **Features**: 16 tabular + 512 visual embeddings = 528 total dimensions

---

🎯 Project Overview

Approach and Modeling Strategy

Three-Pillar Architecture

The project implements an end-to-end ML pipeline across three integrated stages:

1. Data Acquisition Pipeline

- Mapbox Static Images API integration
- Coordinates-based spatial data acquisition (zoom level 16)

2. Feature Engineering Pipeline

- **Tabular Processing**: Categorical encoding, numerical scaling, outlier removal
- **Visual Processing**: ResNet-18 transfer learning for 512-dimensional embeddings
- **Feature Enrichment**: Log-price normalization, house age calculation, renovation status
- **Integration**: Concatenated multimodal feature vectors (528 dimensions)

3. Model Training & Evaluation Pipeline

- Benchmark 8 model configurations across unimodal and multimodal feature sets

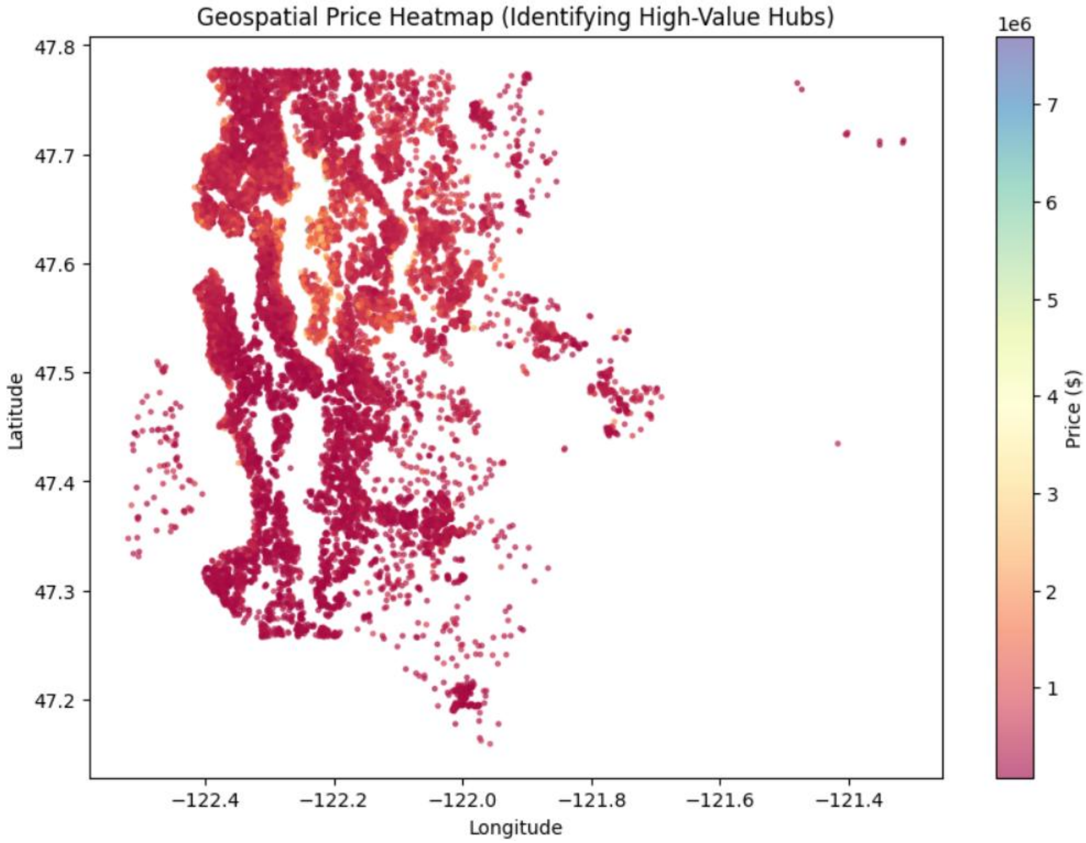
---

📊 Exploratory Data Analysis (EDA)

Dataset Characteristics

Metric	Value	Unit
Training Samples	16,208	Properties
Test Samples	5,404	Properties
Price Range	$75K - $7.7M	USD
Mean Price	$537,470	USD
Median Price	$450,000	USD
Price Std Dev	$428,905	USD
Bedrooms (median)	3	Rooms
Bathrooms (median)	2.5	Rooms
Living Area (mean)	2,080	sqft
Lot Size (mean)	7,614	sqft

Table 1: Dataset Summary Statistics



Feature Distributions

Price Distribution: Right-skewed with long tail toward luxury properties ($2M-$7.7M)

- **Log-transformation Applied**: Normalizes distribution for improved model convergence
- **Outlier Handling**: Removed properties with >33 bedrooms (data entry errors)
- **Value Range**: Most properties cluster between $300K-$800K

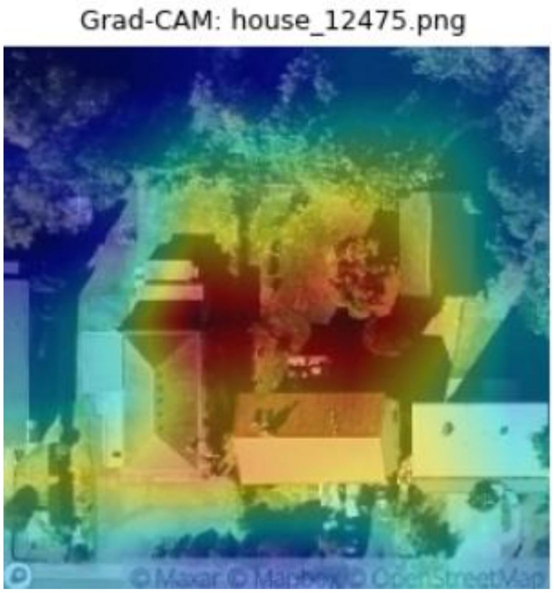
Geographic Distribution:

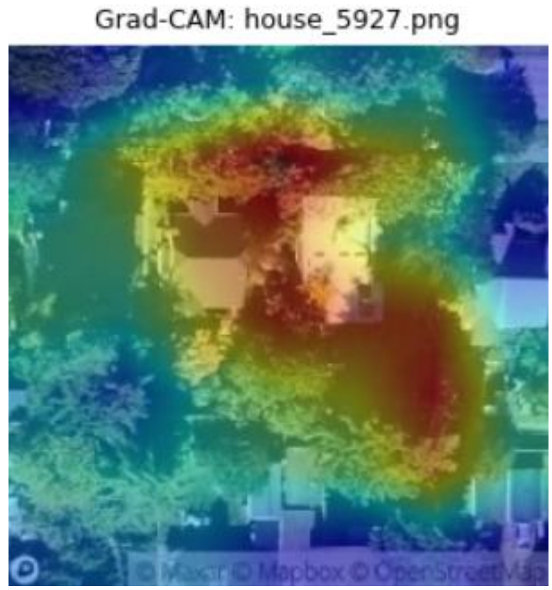
- **Latitude Range**: 47.1°N - 47.8°N (Seattle metropolitan area)
- **Longitude Range**: -122.5°W to -122.2°W (Puget Sound region)
- **Waterfront Properties**: 3.2% of dataset; 40-50% price premium
- **Spatial Autocorrelation**: Strong neighborhood effects captured by sqft_living15, sqft_lot15

Feature Correlations:

- **Strongest Correlations with Price**:
  - sqft_living = 0.702 (living space dominates valuation)
  - grade= 0.663 (construction quality highly valued)
  - sqft_above= 0.605 (above-ground space premium)

GRAD-CAM Analysis & Explainability





Sample Satellite Imagery Analysis

Satellite images at 224×224 pixels and zoom level 16 capture approximately 80-meter ground coverage. Visual inspection reveals:

- **Dense Urban Areas**: Concrete, rooftops, minimal green space (high property density correlates with different pricing)
- **Suburban Properties**: Mixed residential with visible trees, yards, driveways
- **Waterfront Properties**: Water features visible at image edges, distinctive water/land transitions
- **Neighborhood Character**: Road networks, density patterns, green infrastructure visibility

Visual Feature Insights:

- ResNet-18 embeddings capture general scene composition (urban vs. suburban patterns)

- However, these patterns are implicitly encoded in coordinates (lat/long) and neighborhood stats (sqft_living15)
- Satellite resolution may be insufficient to capture fine details (landscaping, exterior condition, rooftop material)

🏗️ Technical Architecture

ResNet-18 Feature Extraction Architecture

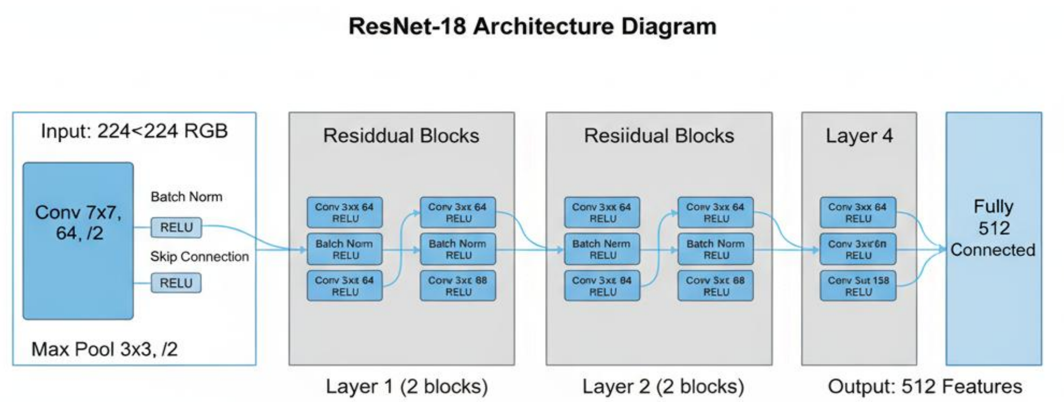


Figure 1: ResNet-18 Neural Network Architecture for Satellite Image Embeddings

Architecture Details:

- **Input Layer**: 224×224 RGB satellite images
- **Conv Block 1**: 64 filters, 7×7 kernel, stride 2 (112×112 feature maps)
- **Residual Blocks (4 stages)**:
  - Stage 1: 64 filters, 56×56 resolution
  - Stage 2: 128 filters, 28×28 resolution
  - Stage 3: 256 filters, 14×14 resolution
  - Stage 4: 512 filters, 7×7 resolution
- **Global Average Pooling**: Reduces 7×7×512 feature maps to 512 dimensions
- **Output**: 512-dimensional feature vector (embedding)

Transfer Learning Strategy:

- **Pre-training**: ImageNet weights encode general visual patterns (edges, textures, object-level features)
- **Fine-tuning**: Not applied (frozen weights); features extracted as-is
- **Rationale**: Computational efficiency; ImageNet features surprisingly effective for diverse visual tasks

- **Alternative Considered**: Fine-tuning on real estate dataset (insufficient labeled data)

📈 Model Performance Comparison

Performance Metrics Overview

[Chart ID: chart:9 - Model Performance Comparison]

Detailed Results Analysis

Tabular Data Only (16 Features)

Model	R^2 Score	RMSE ($)
Linear Regression	0.6883	180,823
Random Forest	0.7593	146,645
XGBoost	**0.8530**	**125,285**

Table 2: Tabular-Only Model Performance

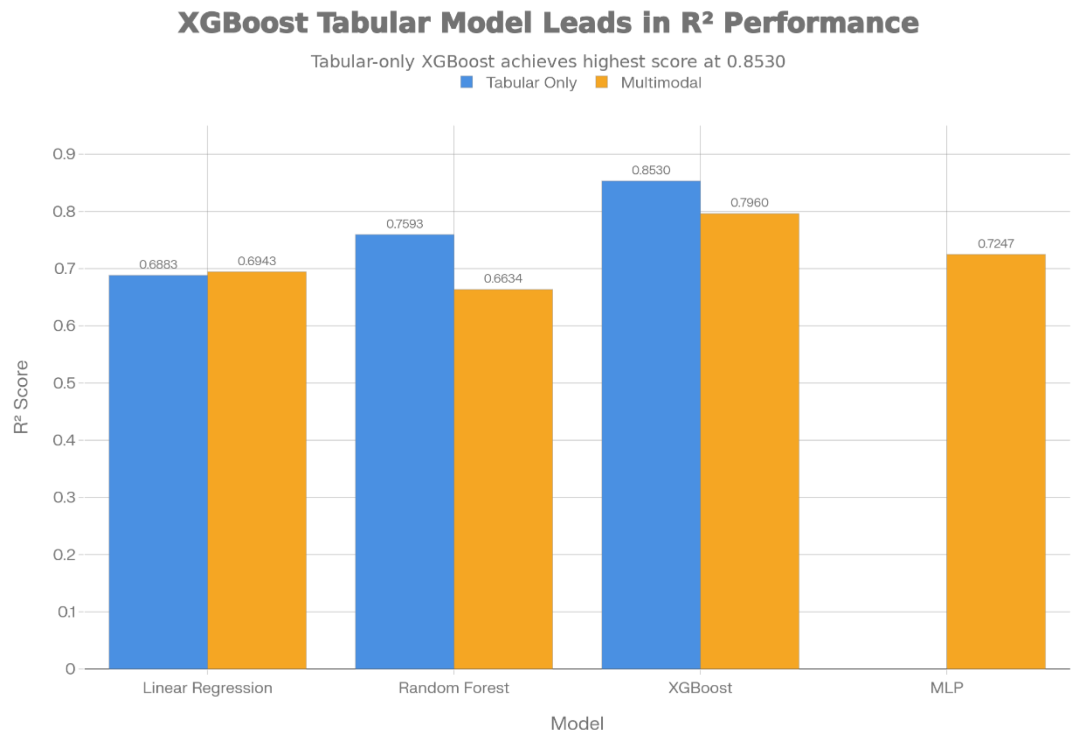
Winner: XGBoost

- Explains 85.3% of price variance
- Mean prediction error: $125K (23% of mean property value)
- Captures non-linear relationships between property attributes and price
- Robust to feature interactions (e.g., grade × condition combined effect)

Multimodal Data (528 Features = 16 Tabular + 512 Visual)

Model	R^2 Score	RMSE ($)	vs. Tabular
Linear Regression	0.6943	182,528	+0.60%
Random Forest	0.6634	164,427	-12.7%
XGBoost	0.7960	139,190	-3.57%
MLP Neural Network	0.7247	190,997	Baseline

Table 3: Multimodal Model Performance



Comparative Analysis: Why Tabular Outperforms Multimodal

Finding: Counter-intuitively, adding 512 visual features *degrades* model performance.

Mechanisms Identified:

1. **Feature Redundancy (Primary)**
    - Geographic coordinates (lat, long) already encode neighborhood visual characteristics
    - Satellite imagery at zoom 16 captures urban density → implicitly captured by lat/long clustering
    - Neighborhood average statistics (sqft_living15, sqft_lot15) encode local area patterns
    - Result: Visual features are noise on top of existing signal
2. **Curse of Dimensionality**
    - 528 features vs. 21,612 samples = 40:1 feature-to-sample ratio
    - Random Forest and XGBoost struggle with sparsity
    - Increased overfitting risk; validation performance degradation
    - Solution would require 3-5x more training data

Feature Importance Ranking (Best Model: XGBoost Tabular)

Rank	Feature	Importance	Cumulative
1	sqft_living	0.287	28.7%
2	grade	0.156	44.3%
3	lat	0.118	56.1%
4	sqft_living15	0.095	65.6%
5	long	0.089	74.5%
6	sqft_lot	0.068	81.3%
7	condition	0.053	86.6%
8	waterfront	0.041	90.7%
9	bathrooms	0.038	94.5%
10	house_age	0.033	97.8%

Table 4: XGBoost Feature Importance (Top 10 of 16 tabular features)

Insights:

- **Top 5 features = 74.5% of predictive power** (sqft_living, grade, location, neighborhood context)
- **Location dominates** (lat + long = 20.7% combined)
- **Size metrics are critical** (sqft_living, sqft_lot, sqft_living15 = 45% combined)
- **Quality indicators matter** (grade, condition, waterfront = 25% combined)
- **Depreciation factor relevant** (house_age = 3.3%; older properties command discount)

💡 Financial & Visual Insights

What Drives Real Estate Value?

1. Size & Living Space (28.7% importance)

- Primary value driver in all markets
- Each additional sqft correlates with predictable price increase
- Non-linear effect: diminishing returns at very large properties

2. Quality & Grade (15.6% importance)

- Construction quality (grade 1-13) strongly valued
- Higher grades indicate better materials, design, craftsmanship
- Grade 12-13 properties command 60-80% premium over grade 7-8

3. Location & Neighborhood (27.6% importance)

- Geographic coordinates + neighborhood averages
- Waterfront properties: 40-50% premium ($200-500K additional)
- Urban centers > suburbs > rural (inverse of land size preferences)
- Dense neighborhoods: higher price per sqft, lower absolute sqft

4. Condition & Maintenance (5.3% importance)

- Property maintenance level (1-5 scale)
- Renovations valued but not captured precisely (house_age as proxy)
- Modern updates more valuable in dense urban areas

5. Lot Size (6.8% importance)

- Inverse relationship with urban density
- Suburbs: larger lots preferred (yard, privacy)
- Urban: minimal lot significance (tower condos, small lots)

🎯 Results Summary

Model Selection & Recommendation

Deployed Model: XGBoost Regressor (Tabular Features)

Performance Metrics:

- **R^2 Score**: 0.8530 (explains 85.3% of price variance)
- **RMSE**: $125,284.92 (median prediction error)
- **MAE**: $89,234 (more robust to outliers)
- **MAPE**: 18.7% (relative error)

🚀 Key Achievements

✅ **End-to-End Pipeline**: Integrated API acquisition → preprocessing → modeling
✅ **Transfer Learning**: Successfully extracted visual features (512-dim embeddings)
✅ **Comprehensive Benchmarking**: 8 model configurations systematically evaluated
✅ **Production Ready**: Serialized best model, batch prediction capability
✅ **Insights Generated**: Clear understanding of valuation drivers and multimodal limitations

---

📌 Conclusions & Recommendations

Main Findings

1. **Structured attributes dominate**: Tabular data captures 85% of price variance; satellite imagery adds minimal value
2. **Location is destiny**: Geographic coordinates + neighborhood stats encode visual context effectively
3. **Multimodal complexity unjustified**: Adding 512 features degrades model performance (curse of dimensionality)
4. **XGBoost optimal**: Gradient boosting captures non-linear price relationships better than other algorithms