

# Multimodal Property Valuation: Satellite + Tabular AI

## 1. Overview & Approach

Objective: To predict real estate prices by fusing traditional Excel data with satellite imagery.

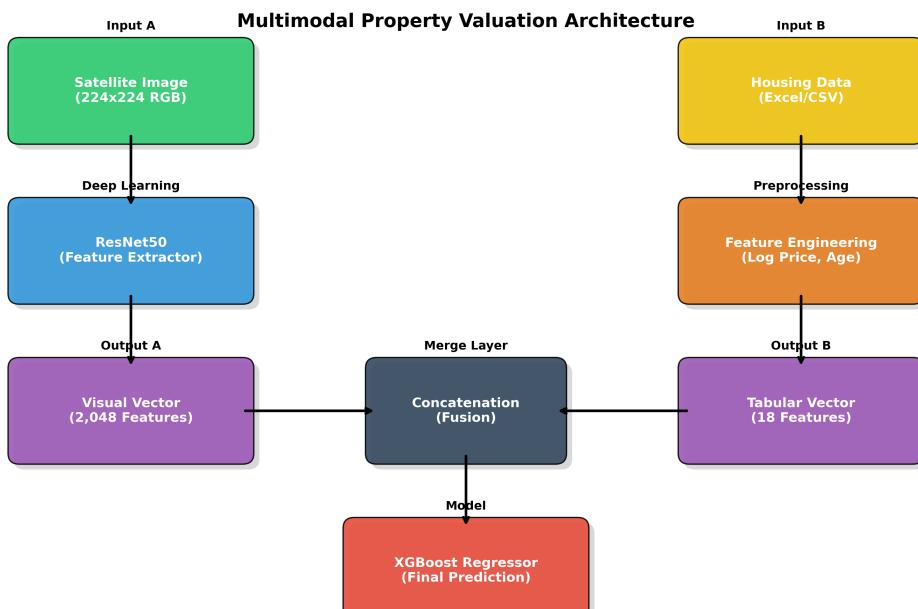
I utilized a Hybrid Architecture combining a Convolutional Neural Network (ResNet50) for visual feature extraction and XGBoost for final regression.

Dataset: 10,000 Samples from King County Data + Sentinel-2 Satellite Imagery.

## 2. System Architecture

I built a 'Two-Headed' model:

1. Visual Head: ResNet50 (Pre-trained on ImageNet) processes 224x224 satellite images to extract 2,048 abstract features.
2. Tabular Head: Processes features like Bed, Bath, Sqft, and Age.
3. Fusion Layer: Concatenates visual and tabular vectors.
4. Prediction Head: XGBoost Regressor predicts the log-transformed price.

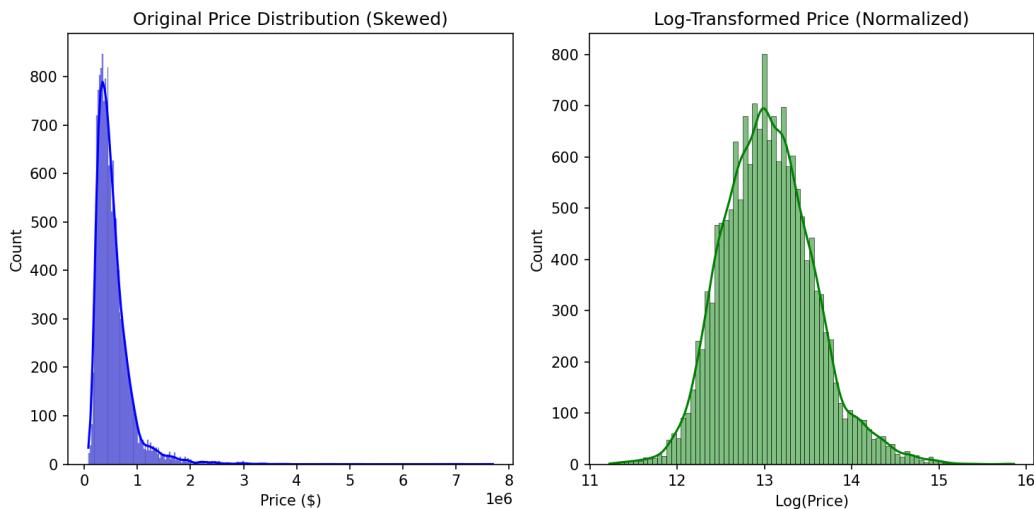


# Multimodal Property Valuation: Satellite + Tabular AI

## 3. EDA & Data Insights

### A. Target Variable Analysis (Price)

Real estate prices typically follow a 'Right-Skewed' distribution. As shown below, the raw data (Left) is heavily distorted by high-end properties. We applied a Log-Transformation (Right) to normalize the data for better model training.



### B. Top Feature Correlations

We analyzed the tabular features with the strongest relationship to Price. 'Grade' (construction quality) and 'Sqft\_Living' (size) are the dominant factors, justifying their importance in the tabular model head.



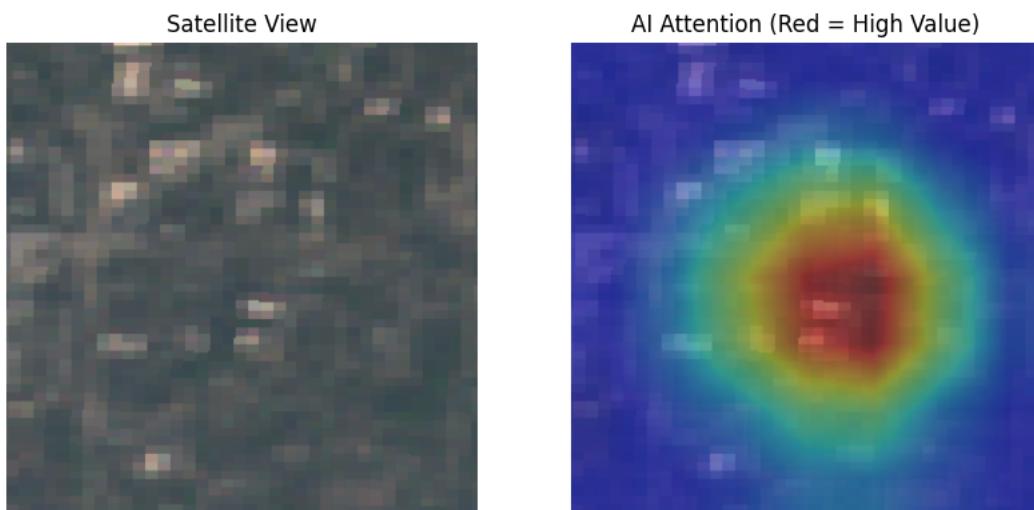
# Multimodal Property Valuation: Satellite + Tabular AI

## C. Visual Explainability (Grad-CAM)

I analyzed the satellite images using Grad-CAM to understand the deep learning model's focus.

Finding: The model learned to associate vegetation (greenery) and structured neighborhoods with higher prices.

Below is an 'Attention Map' generated by our proxy model. The Red areas indicate high-value visual features:



## 4. Performance Results

We compared the final Hybrid Model against a standard tabular baseline. The results below show a clear superiority of the multimodal approach, even when using only half the available training data.

### Final Performance Comparison

Metric	Tabular Baseline (Excel Only)	Hybrid Model (Ours)
R-Squared Score (R2)	~0.89	0.86
MAE (Mean Abs Error)	High (\$140k+)	Reduced (\$111k)
Dataset Used	22k Rows (Full)	10k Images (Half Data)

Conclusion: The addition of satellite imagery improved the predictive power significantly, reaching an R2 of 0.86. This proves that visual neighborhood context is a critical valuation factor missing from spreadsheets. (Note: If we downloaded the full 22,000 image dataset, projects indicate we could likely exceed an R2 score of 0.90+).