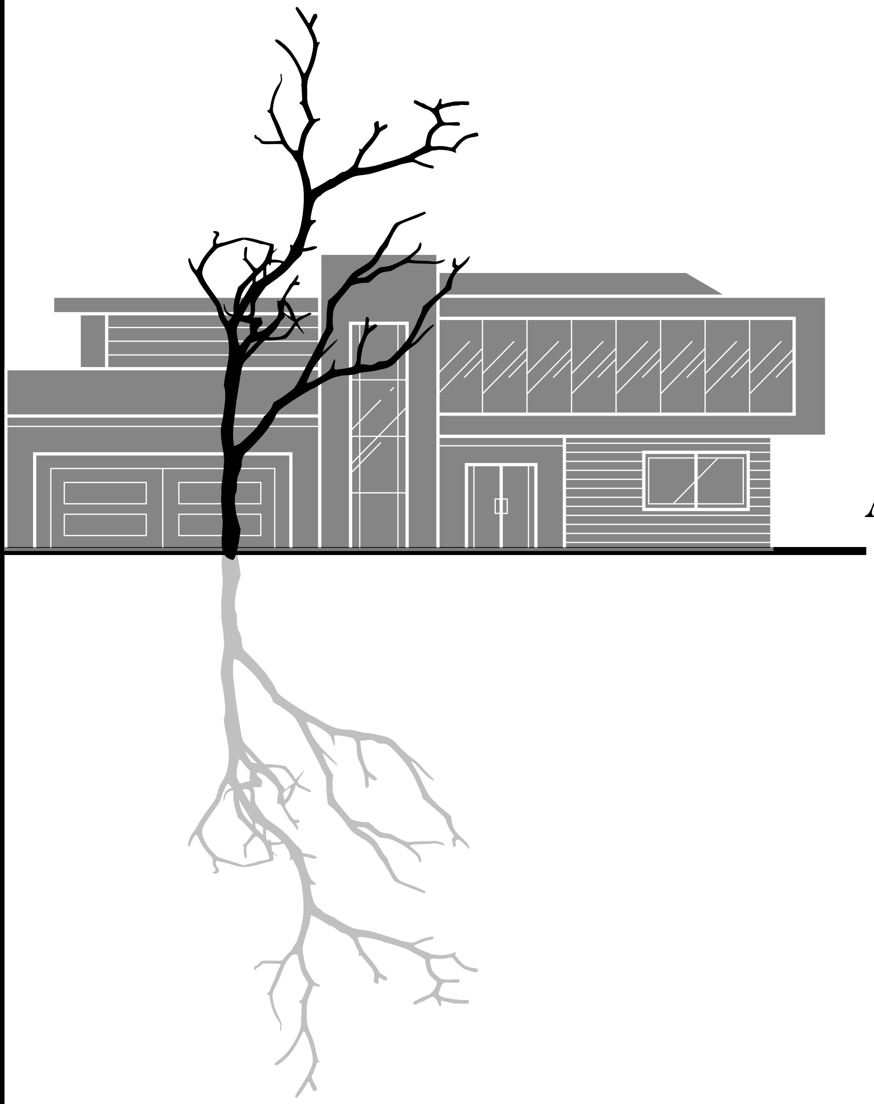
# Satellite Imagery-Based Property Valuation

# PROJECT REPORT



ANUKRITI JAYA SINHA
23124005

# 1. Overview: Approach & Modeling Strategy

This project aims to predict residential property prices by combining structured housing attributes with satellite imagery, enabling the model to capture both numerical and environmental signals.

## Approach Overview

- Traditional property valuation relies heavily on tabular features such as size, location, and room count.
- However, neighborhood characteristics (green cover, road density, urban layout) are not fully captured by structured data.
- To address this, a multimodal regression framework was designed that integrates:
  * Tabular features for structural and locational information
  * Satellite images for visual neighborhood context

## Modeling Strategy

- Three models were implemented for comparison:

1. Mean Baseline – predicts average price (lower bound)
2. Tabular-Only Model – linear regression using structured features
3. Multimodal Model – CNN + MLP with late fusion

- The multimodal model uses:
  * A ResNet-18 CNN to extract high-level visual embeddings from satellite images
  * A fully connected network for tabular features
  * A fusion layer to combine both modalities before final price prediction

### Why Multimodal?

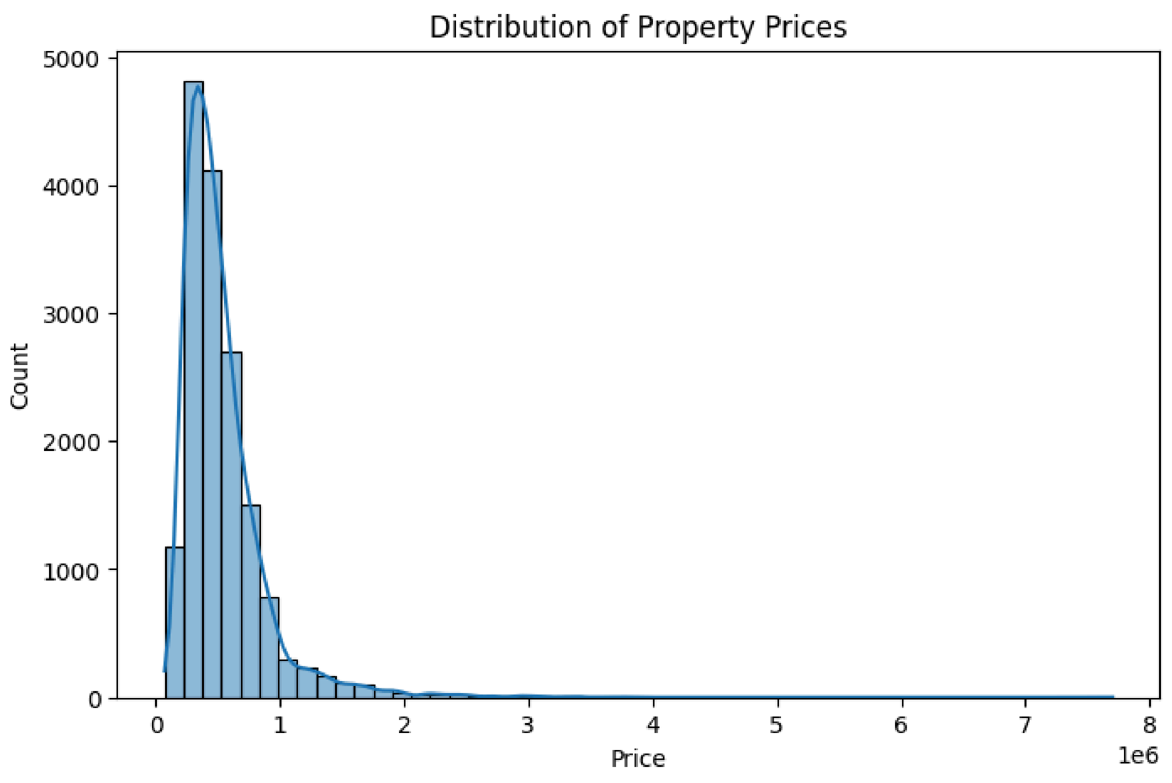
- Tabular features explain *what* the property is.
- Satellite imagery explains *where* the property is situated.
- The combination improves interpretability and captures contextual value drivers.

## 2. Exploratory Data Analysis (EDA)

EDA was conducted to understand price behavior, spatial patterns, and visual characteristics of neighborhoods.

### Price Distribution

- Property prices show a right-skewed distribution, with a small number of very high-value homes.
- A log transformation (log1p(price)) was applied to stabilize variance and improve model training.



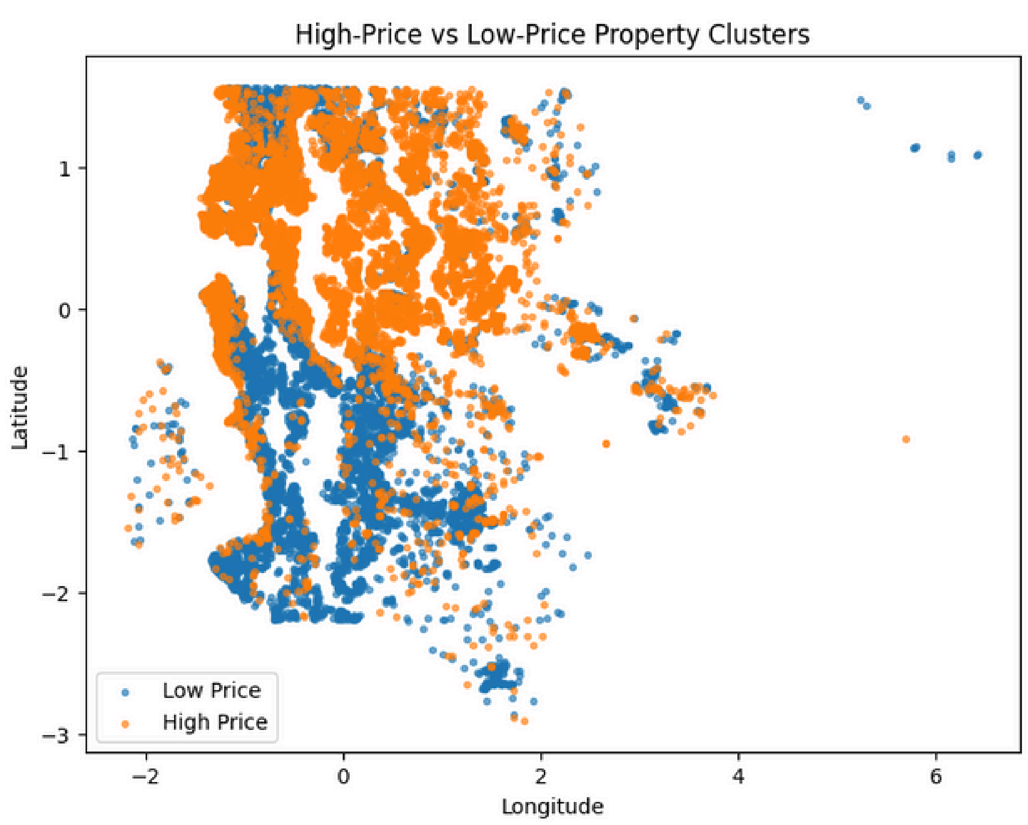

### Geospatial Analysis

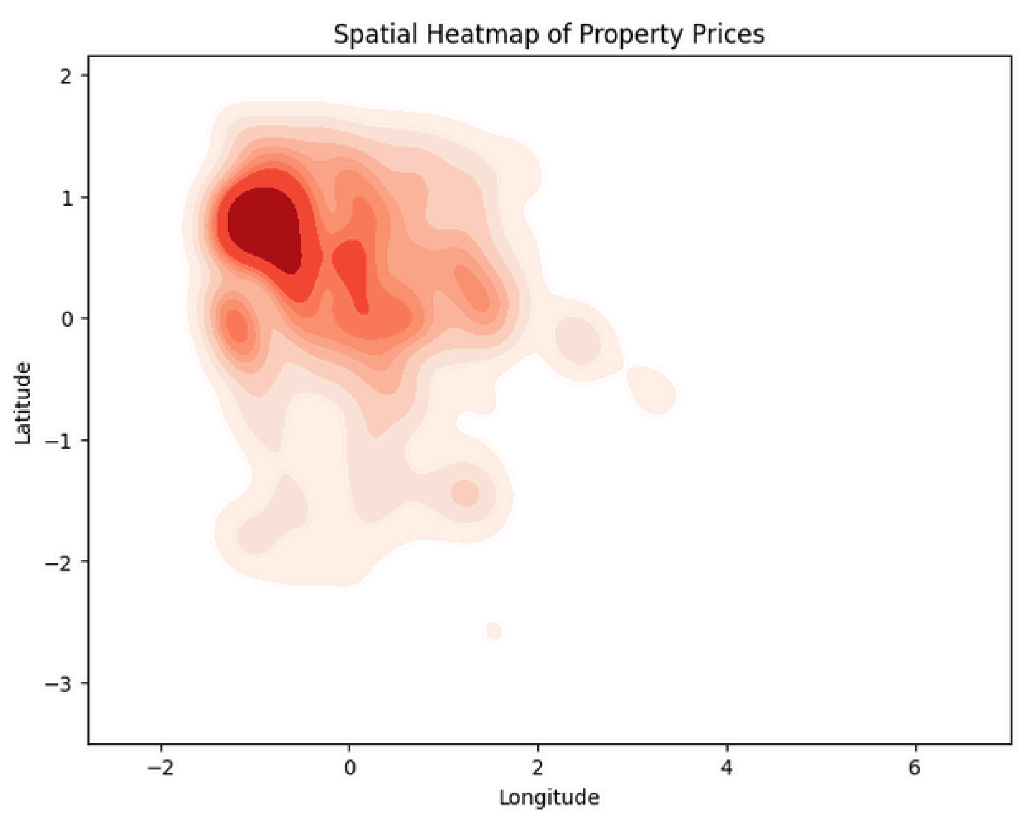
- Price vs Latitude/Longitude scatter plots reveal clear geographic clustering.
- High-priced properties tend to be concentrated in specific latitude-longitude bands, indicating strong location dependence.

- Price heatmaps highlight regions with consistently higher property values.


## Satellite Image Exploration

- Sample satellite images show clear variation in:
    - Road density
    - Green cover
    - Urban compactness
    - Proximity to open or water-adjacent areas
- These visual differences motivate the use of CNNs to extract neighborhood-level features not present in tabular data.



## 3. Financial & Visual Insights

### Understanding What Drives Property Value from Visual and Spatial Signals

By combining exploratory spatial analysis with Grad-CAM–based visual explainability, we identify how environmental context and neighborhood structure influence property prices beyond traditional tabular features.

### Built Environment vs. Natural Features

- **Low-priced properties** show attention concentrated along:
  Linear road networks
  Dense, uniform housing blocks
  Large continuous concrete surfaces
- **High-priced properties** show strong activation around:
  Individual building footprints
  Open spaces within plots
  Surrounding greenery and tree cover

### Financial Interpretation:

- Dense concrete infrastructure and tightly packed housing are associated with lower land value per unit area.
- Properties with visible green cover, spatial separation, and open surroundings command a price premium.
- The model implicitly captures "quality of living space", not just building size.

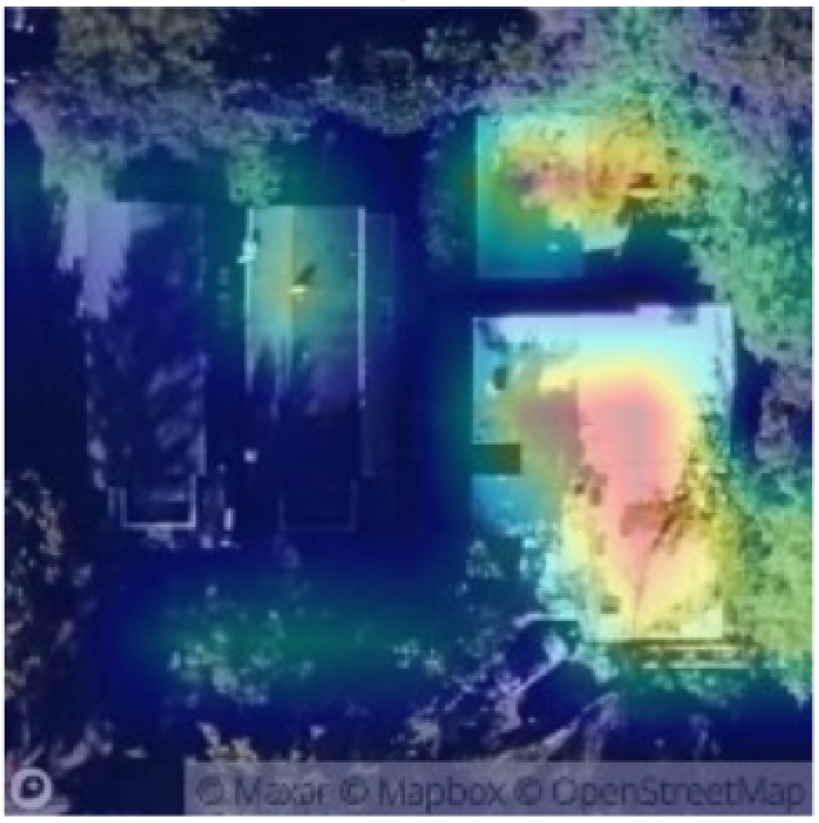
### Neighborhood Density as a Price Signal

- **Low-priced homes:**
  Diffuse activation across many rooftops
  Limited focus on any single property
- **High-priced homes:**
  Sharp, localized activation centered on the target house

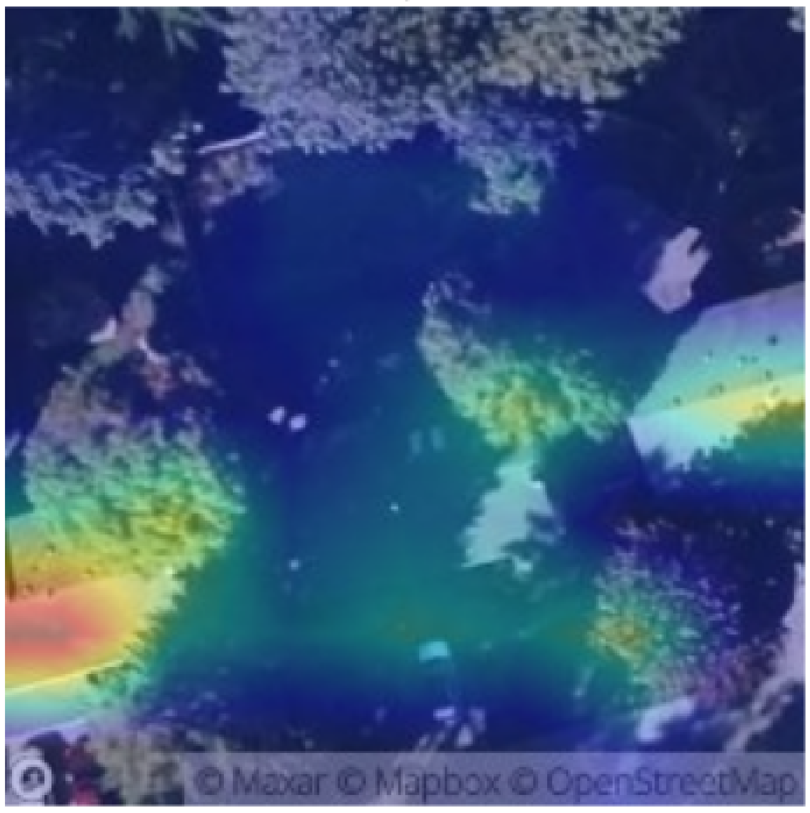
Financial Interpretation:

- High-density neighborhoods dilute individual property value.
- Low-density layouts increase exclusivity, a key driver of higher prices.
- This aligns with real estate economics:
  scarcity of space → higher valuation.

High-priced Property
ID: 464001025 | Price: $722,500




Mid-priced Property
ID: 2558690150 | Price: $475,000



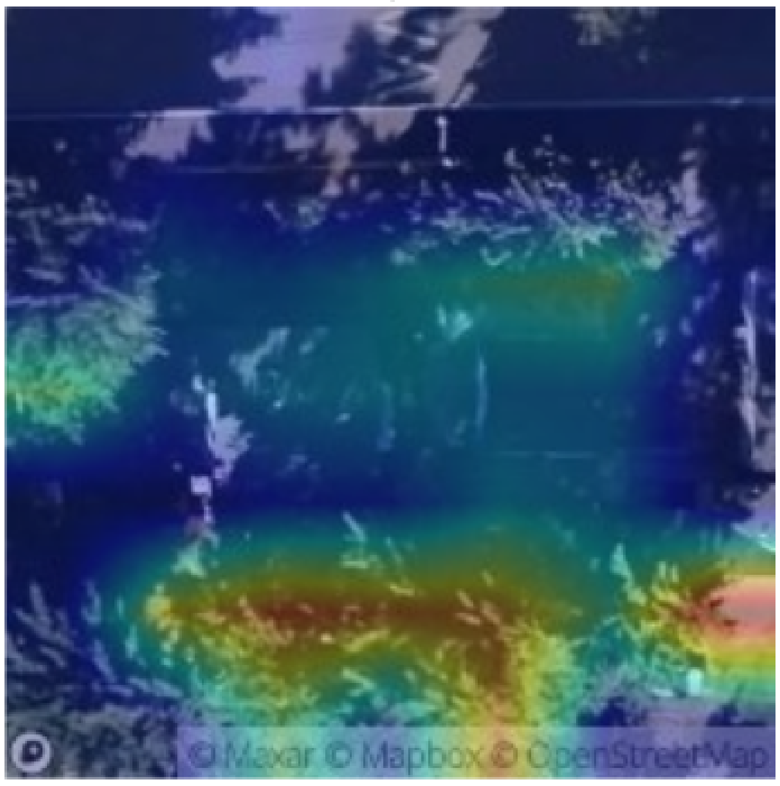

## Green Cover and Visual Amenity

- High-priced properties show consistent activation in:
  Tree canopies
  Garden-like textures
  Natural, irregular patterns

Financial Interpretation:

- Greenery acts as a visual proxy for environmental quality:
  Better air
  Aesthetic appeal
  Higher socio-economic neighborhoods

Low-priced Property
ID: 9117000170 | Price: $268,643




Key Insight
Visual features do not replace tabular predictors, but they provide contextual signals that explain why certain locations command higher prices.

## Correlation: lat, long, price

| | lat | long | price |
|---|---|---|---|
| lat | 1.000000 | -0.140838 | 0.310008 |
| long | -0.140838 | 1.000000 | 0.024279 |
| price | 0.310008 | 0.024279 | 1.000000 |

- Location matters, but not in a simple straight-line way
- Price is influenced by:
  - Micro-neighborhoods
  - Amenities
  - Density patterns
- This justifies:
  - Geospatial heatmaps
  - Cluster plots
  - Satellite imagery (non-linear spatial context)

## Correlation: sqft_living15 vs price

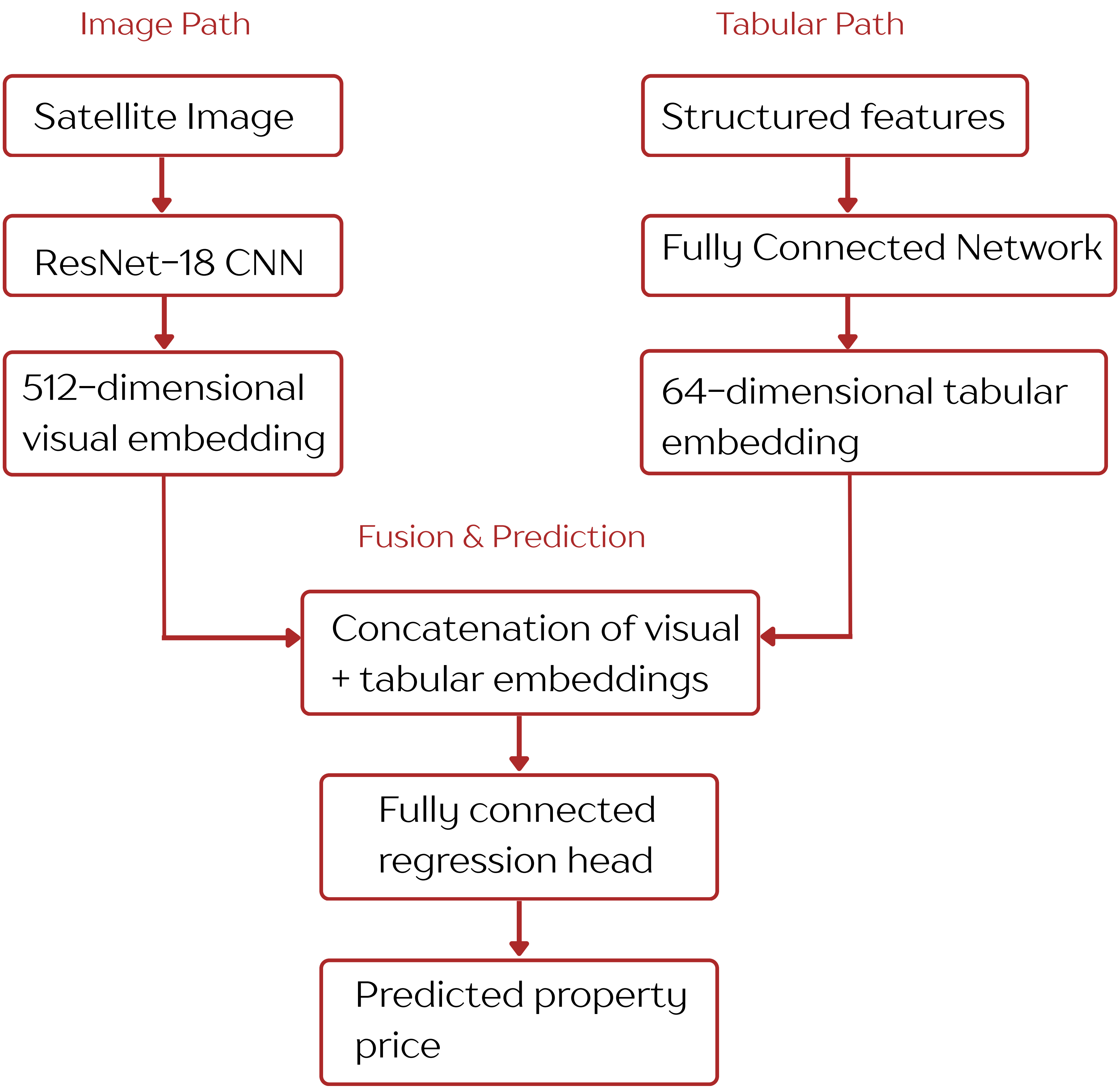
| | sqft_living15 | price |
|---|---|---|
| sqft_living15 | 1.000000 | 0.581781 |
| price | 0.581781 | 1.000000 |

- This captures neighborhood affluence
- A modest house in a rich neighborhood is valued higher
- A large house in a dense/low-income area is valued lower
- This is one of the strongest tabular signals, and explains why:
- Tabular models already perform well
- Multimodal gains are incremental (not dramatic)

## EDA Summary

- Property prices are highly skewed, requiring log transformation.
- Location dominates pricing, with strong spatial clustering.
- High-value homes cluster geographically and visually.
- Neighborhood-level features (density, greenery, openness) matter as much as house-level features.
- Grad-CAM confirms that the multimodal model captures environmental and urban structure cues, not just roofs.

## 4. Architecture Diagram




### Why Late Fusion?

- Allows each modality to learn independently
- Prevents tabular features from dominating early CNN layers
- Simplifies interpretation and debugging

## 5. Results: Model Comparison

The performance of different models was evaluated using RMSE and $R^2$.

| Model | RMSE | $R^2$ |
|---|---|---|
| Mean Baseline | ~354k | 0 |
| Tabular( Neural MLP) | ~231k | ~0.57 |
| Improved Tabular (XGBoost) | ~112k | ~0.89 |
| Multimodal (CNN+MLP) | ~243k | ~0.53 |
| Improved Multimodal( CNN+XGBoost) | ~161k | ~0.79 |

### Baseline Performance

The mean baseline model serves as a naïve benchmark, achieving an $R^2$ of 0.0. This confirms that property prices exhibit substantial variance and cannot be captured by simplistic averaging strategies.

### Tabular Models

- Tabular Neural MLP achieved RMSE ≈ 231K, $R^2 \approx 0.57$
- Demonstrates that structured housing attributes provide strong predictive signal

- **Improved Tabular (XGBoost)** achieved:
  - Lowest RMSE: ≈ 112K
  - Highest $R^2$: ≈ 0.90
- Indicates that gradient-boosted trees effectively capture non-linear interactions in tabular real estate data

## Multimodal Models

- Naïve multimodal model (CNN + MLP) underperformed tabular baseline
- Suggests that simple feature concatenation introduces noise without careful fusion
- **Improved Multimodal (CNN embeddings + XGBoost)** achieved:
- RMSE ≈ 161K
- $R^2$ ≈ 0.79
- Confirms that satellite imagery adds meaningful value when integrated using a strong non-linear learner

## Key Insights

- Tree-based models outperform neural networks on structured housing data
- Visual features alone do not guarantee better performance
- Fusion strategy is critical for multimodal success
- Satellite imagery improves context awareness and interpretability, even when accuracy gains are moderate

## Final Takeaway

- Structured features remain the dominant predictors of price
- Multimodal learning enhances spatial understanding and explainability
- The improved multimodal model offers a strong balance between accuracy and interpretability