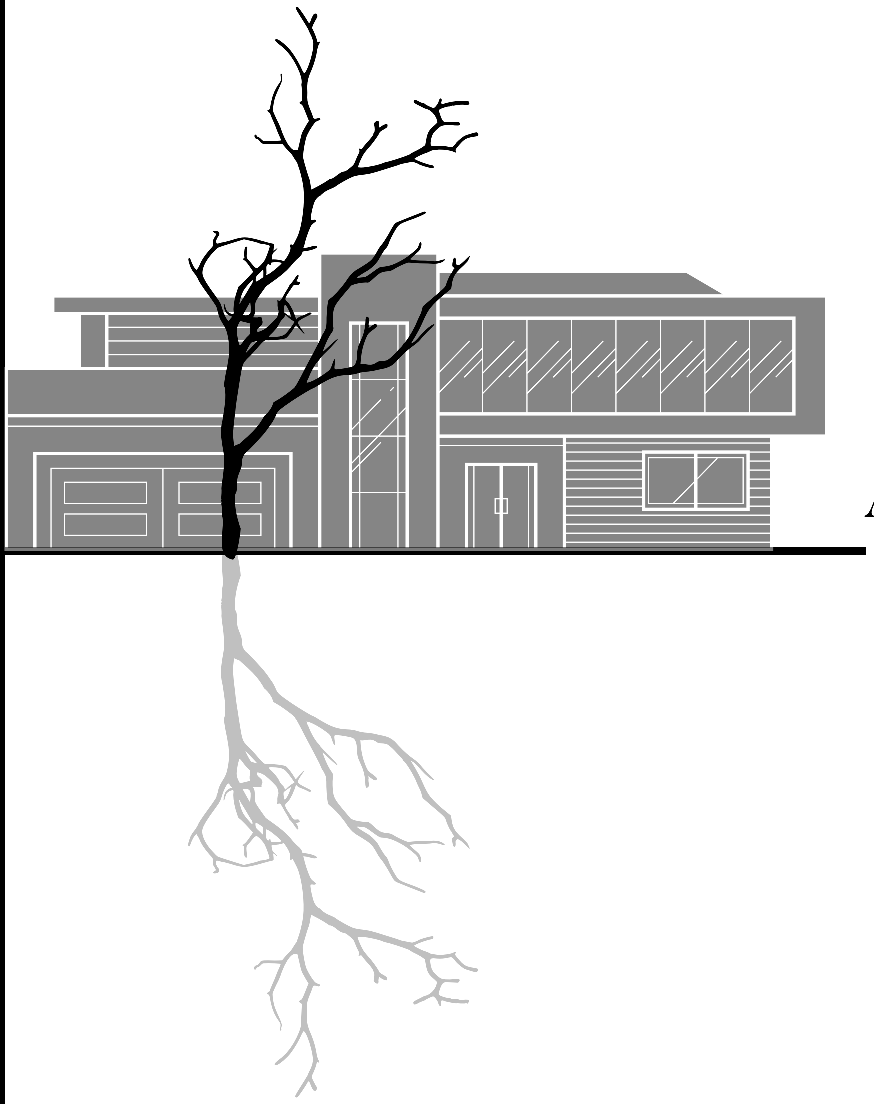
Satellite Imagery-Based Property Valuation

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?

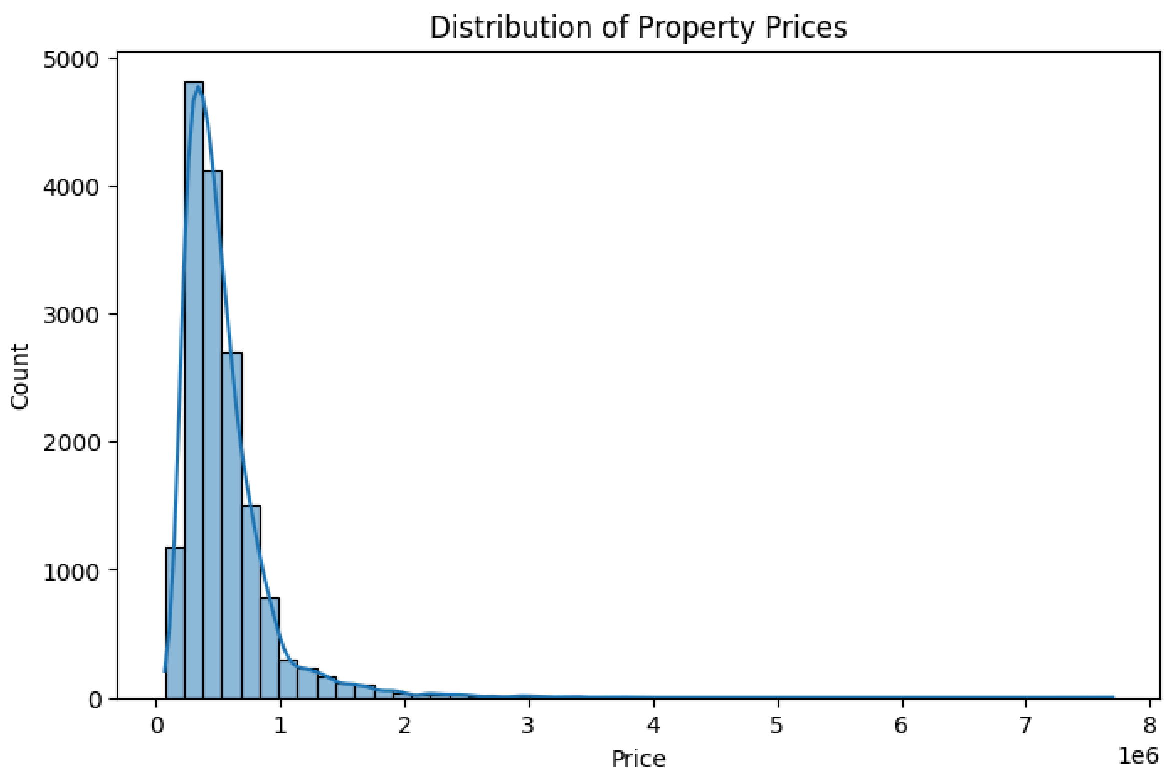
- Tabular features explain *what* the property is.
- Satellite imagery explains *where* the property is situated.
- The combination improves interpretability and captures contextual value drivers.

2. Exploratory Data Analysis (EDA)

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

Price Distribution

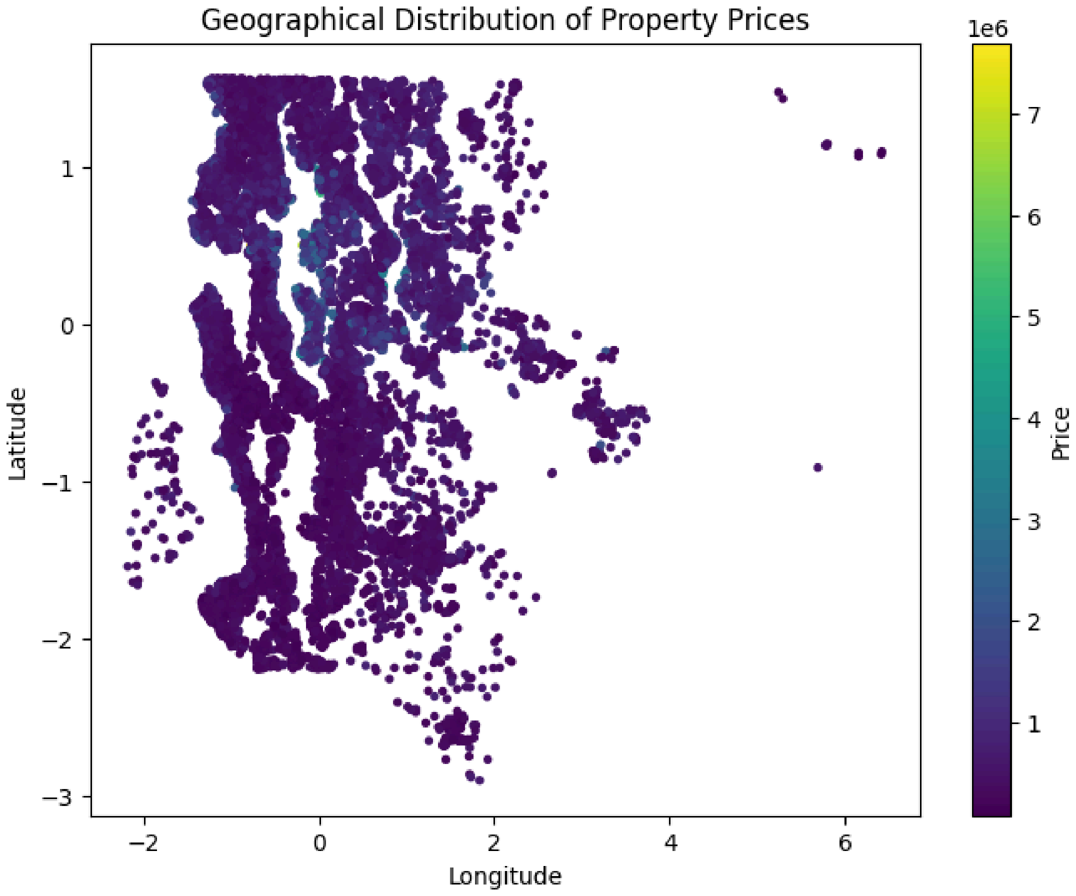
- Property prices show a right-skewed distribution, with a small number of very high-value homes.
- A log transformation (log1p(price)) was applied to stabilize variance and improve model training.



Geospatial Analysis

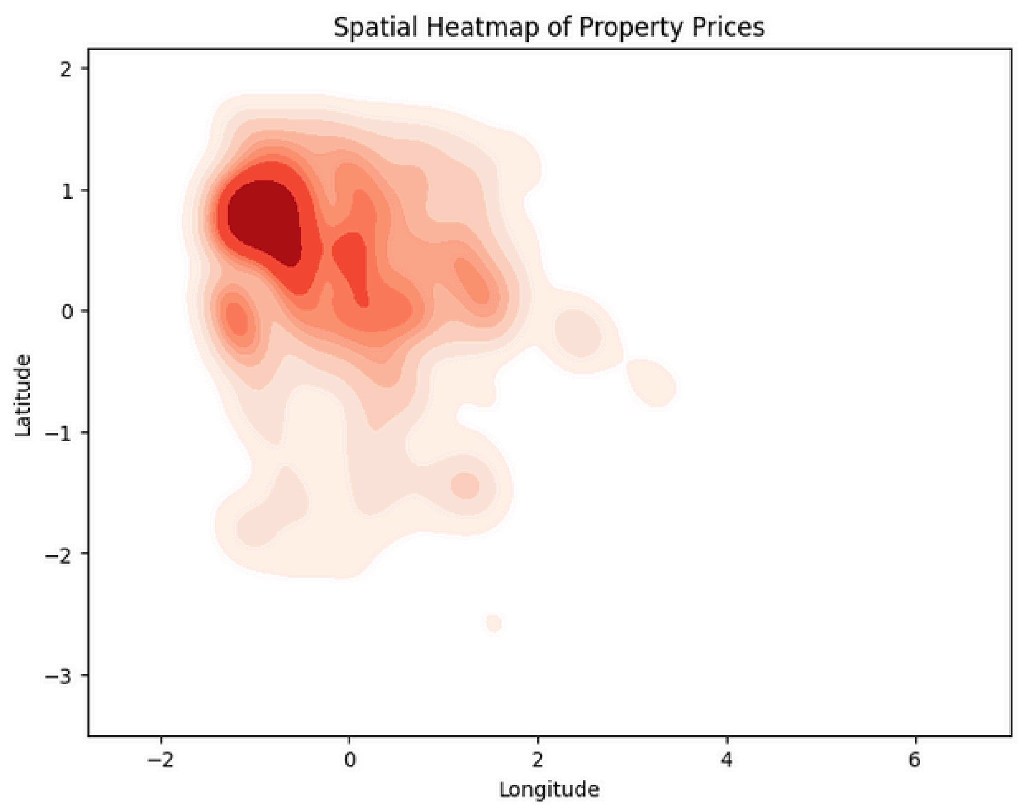
- Price vs Latitude/Longitude scatter plots reveal clear geographic clustering.
- High-priced properties tend to be concentrated in specific latitude-longitude bands, indicating strong location dependence.

- 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

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

High-Value Property Indicators

- Dense but organized road networks
- Planned urban layouts
- Presence of green spaces near residential clusters
- Open surroundings with low congestion

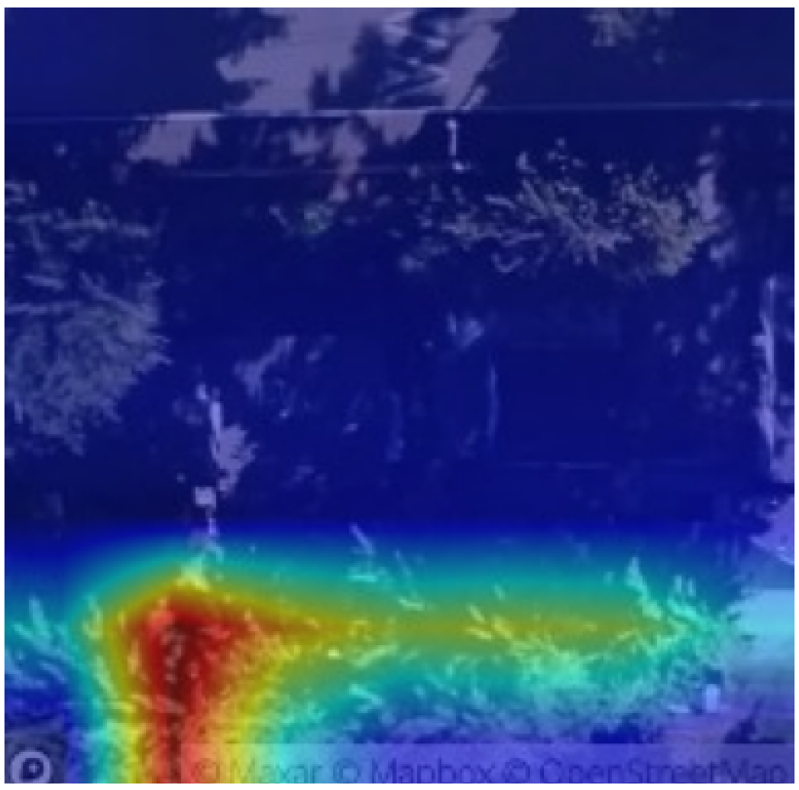
Lower-Value Property Indicators

- Sparse or irregular road connectivity
- Highly congested or industrial-looking zones
- Minimal vegetation
- Poorly structured neighborhood layouts

Grad-CAM Observations

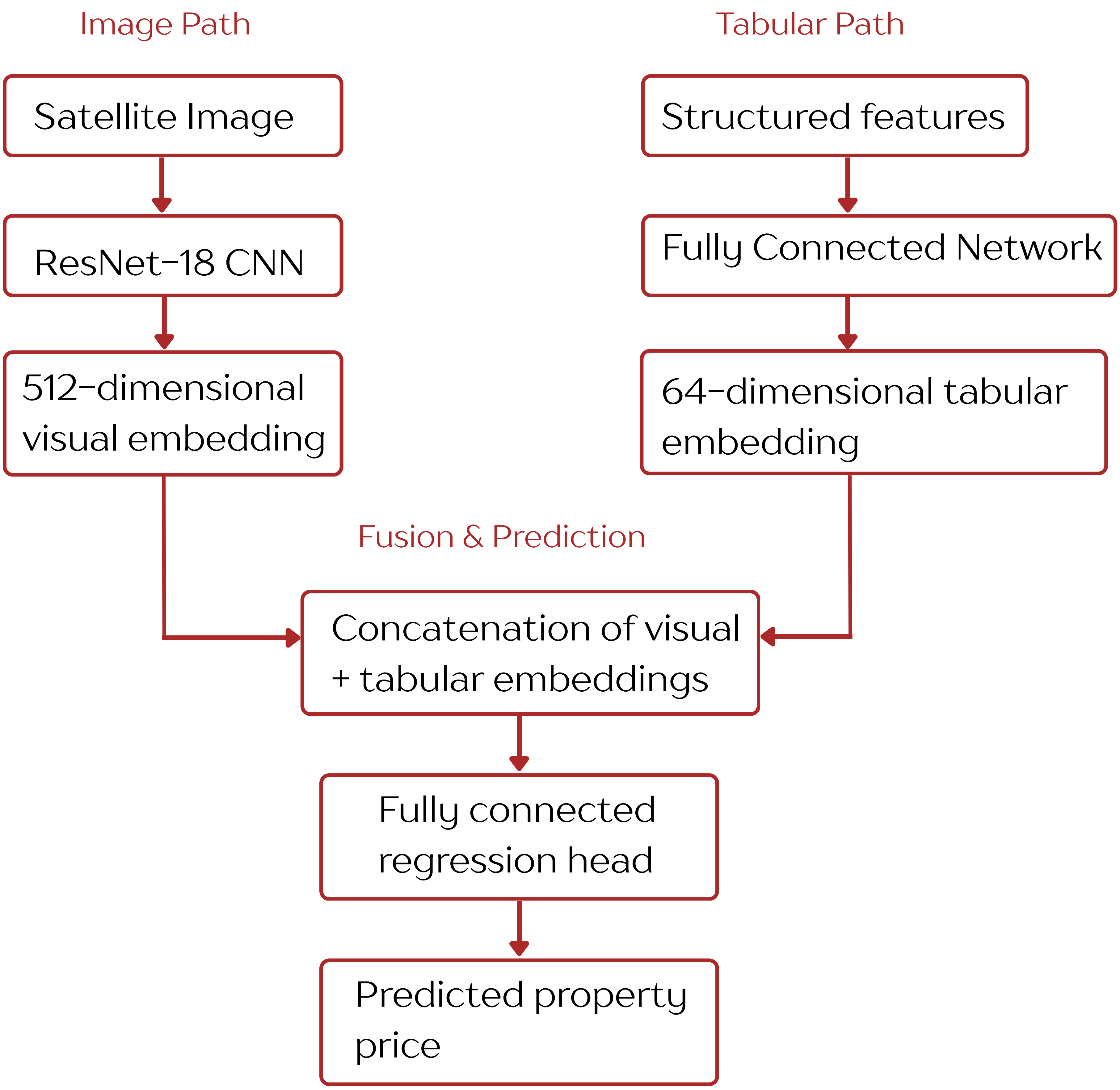
- Grad-CAM heatmaps show that the CNN focuses on:
  Road structures
  Built-up density patterns
  Green/open regions
- This confirms that the model learns environmental context, not just building footprints.

Grad-CAM: Influential Regions in Satellite Image



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

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-Only	~231k	~0.57
Multimodal (Tabular + Images)	~228k	~0.59

Baseline Model

- The mean baseline predicts the average house price for all samples.
- As expected, it shows no explanatory power ($R^2 = 0.00$) and a high RMSE.
- This serves as a lower bound for performance comparison.

Tabular Only Model

- The tabular model uses structured housing features such as:
  * bedrooms
  * bathrooms
  * living area
  * latitude and longitude
- It achieves a substantial improvement over the baseline, with:
  * RMSE ≈ 231k
  * $R^2 \approx 0.57$
- This confirms that size and location are strong predictors of property value.

Multimodal Model

- The multimodal model integrates:
  * Tabular features (structural and locational)
  * Visual features extracted from satellite imagery using a CNN
- The model achieves the best overall performance, with:
  * Lowest RMSE (≈ 228k)
  * Highest R^2 (≈ 0.59)
- This indicates that satellite imagery provides additional contextual information beyond structured data alone.

Key Observations

- While the improvement over the tabular model is modest, it is consistent and meaningful.
- Satellite imagery contributes environmental signals such as:
  road connectivity
  neighborhood density
  presence of green/open spaces
- These signals help refine price predictions, particularly for properties with similar structural attributes.

Conclusion

The multimodal model achieves the best overall performance by combining strong tabular predictors with contextual visual information from satellite imagery.