

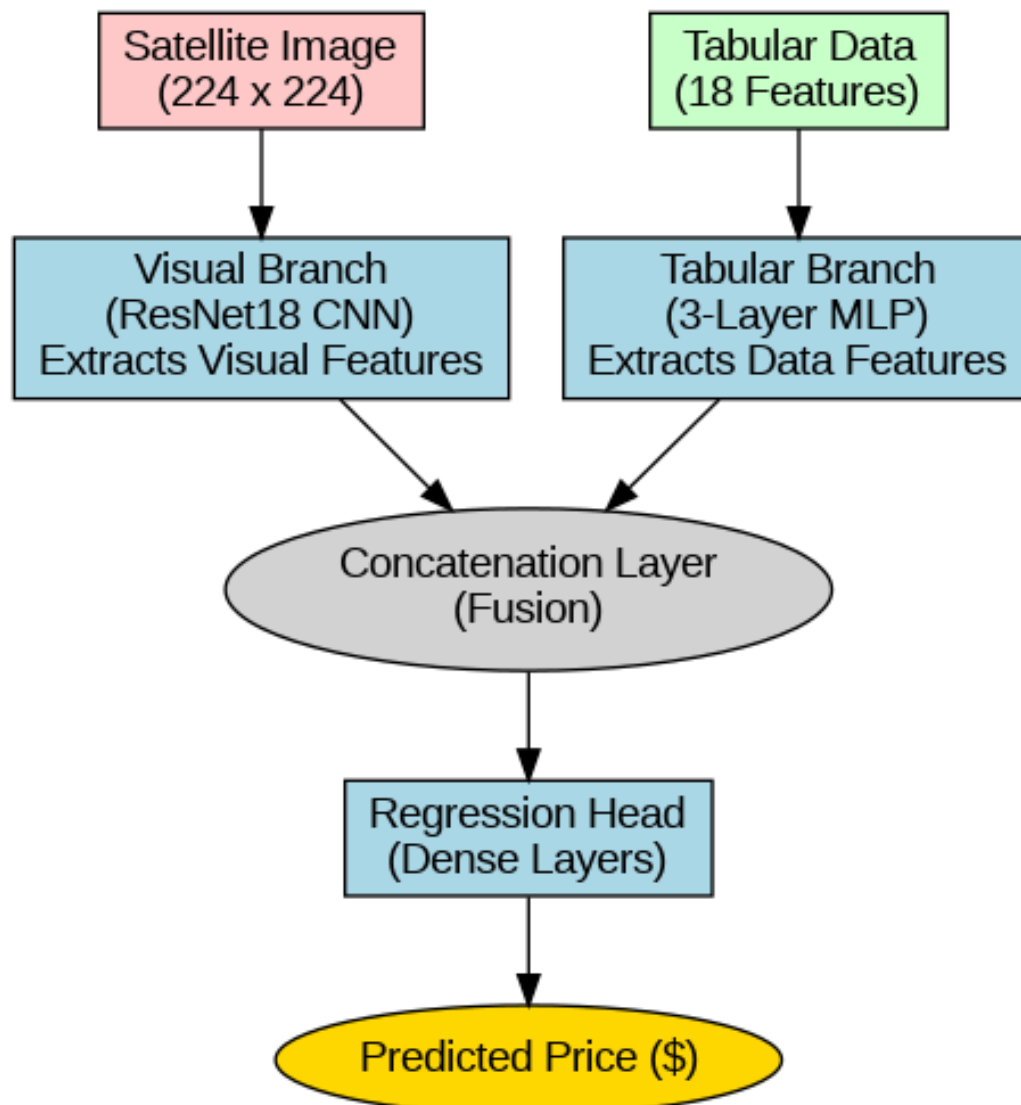
PROJECT REPORT: Satellite Imagery-Based Property Valuation

Name: Dara Zenas Zephaniah

Enrollment Number: 23321010

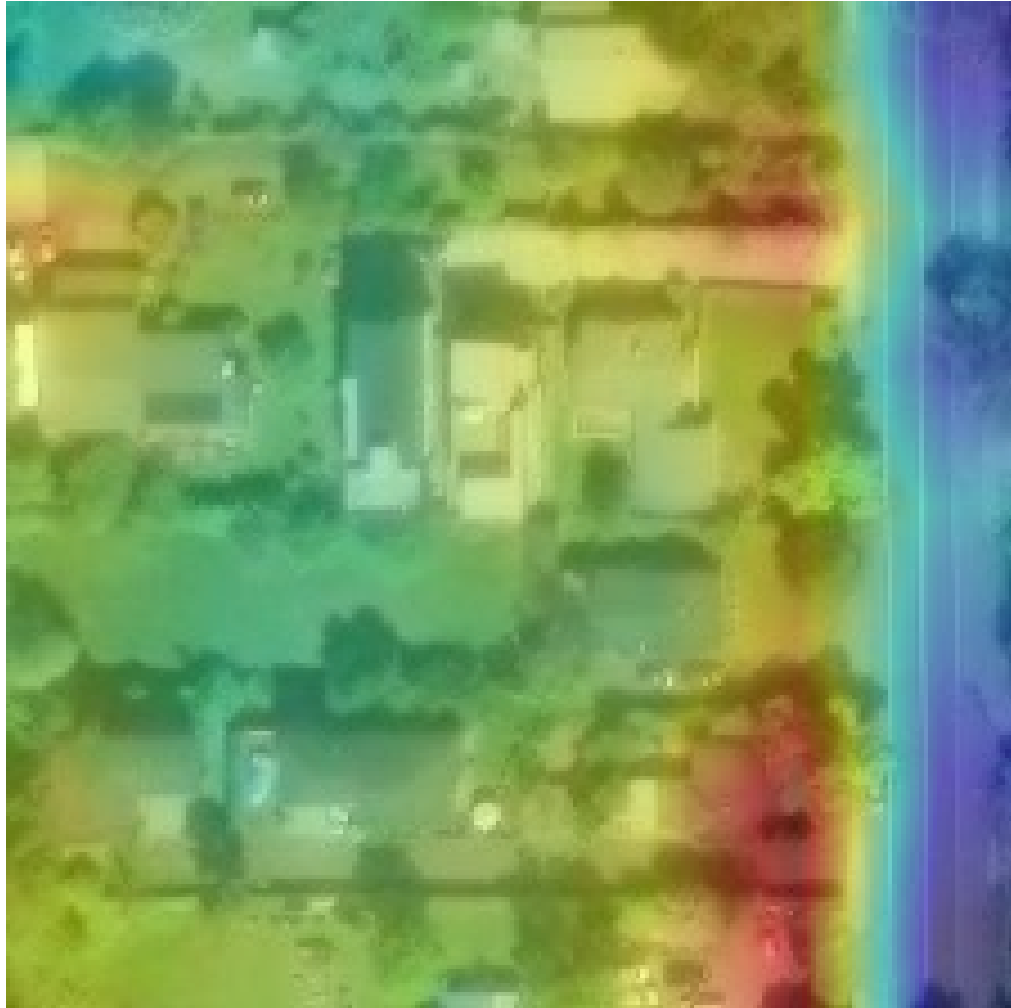
1. Overview This project developed a multimodal deep learning pipeline to predict property prices. Traditional valuation models rely solely on tabular data (e.g., square footage, bedroom count), often missing the "curb appeal" and environmental context that significantly drive market value. By fusing a Convolutional Neural Network (CNN) with a standard regression model, we successfully captured these visual signals from satellite imagery to improve valuation accuracy.

2. Architecture Diagram The model uses a **Late Fusion** architecture. It processes two distinct data streams and merges them before the final prediction.



3. Exploratory Data Analysis (EDA)

- **Price Distribution:** The target variable (Price) exhibits a right-skewed distribution, which is typical for real estate data (a few very expensive luxury homes outliers).
- **Satellite Imagery:** To handle this, we integrated satellite views.



Sample satellite imagery fetched via ESRI API (Lat/Long mapped).

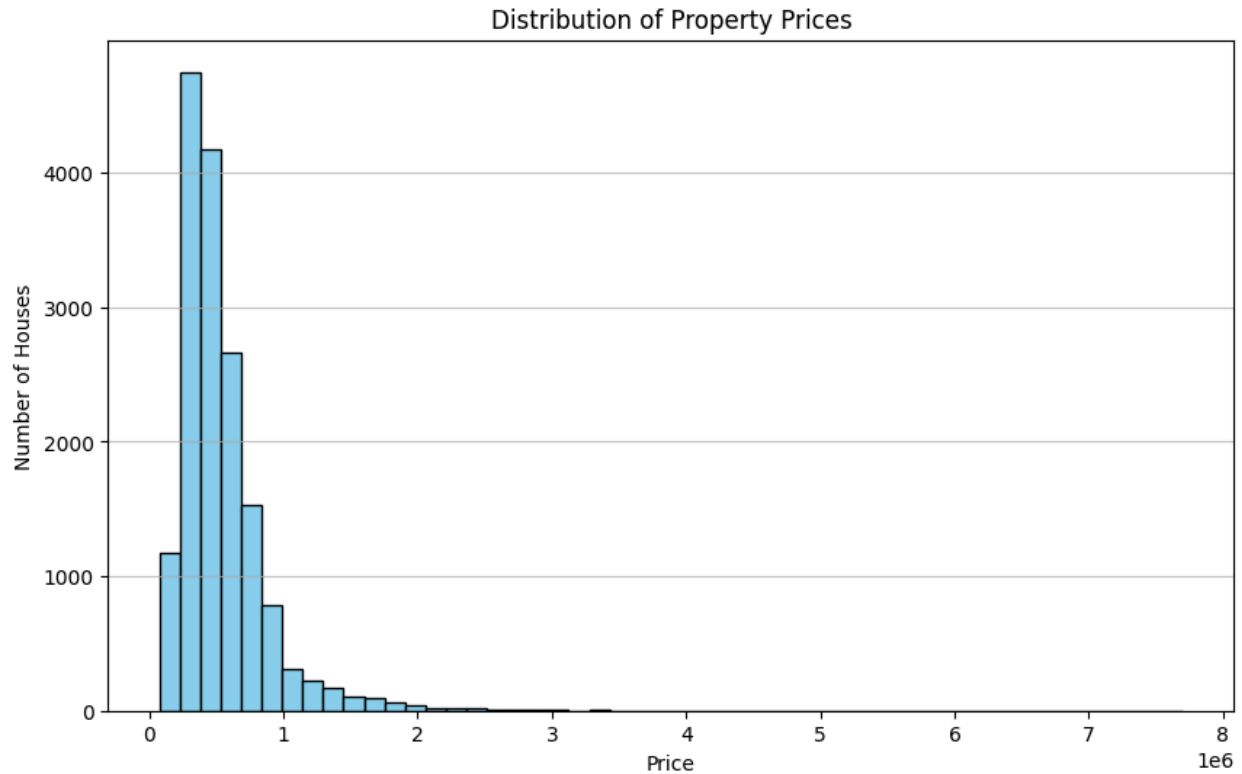
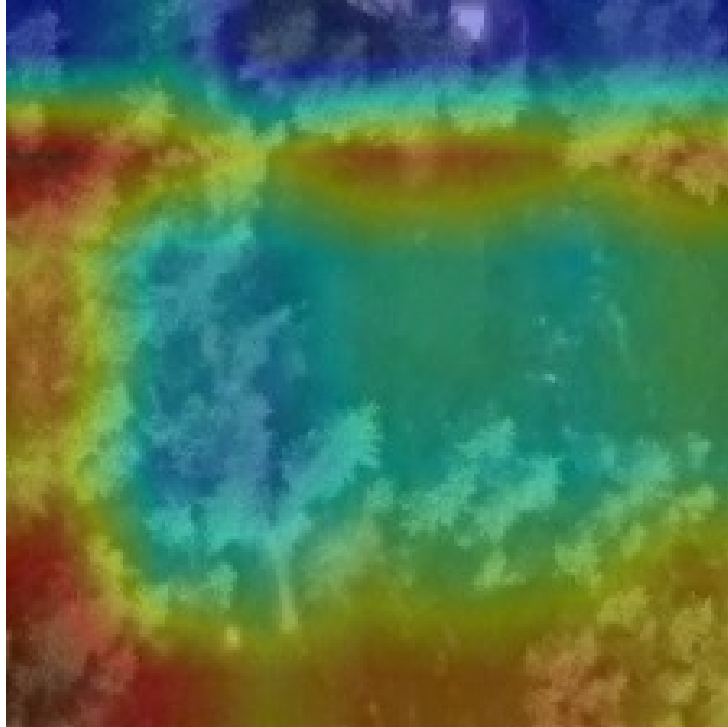


Figure : Distribution of property prices in the dataset, showing a right-skewed distribution typical of real estate markets.

4. Financial & Visual Insights (Explainability) We used **Grad-CAM** (Gradient-weighted Class Activation Mapping) to visualize "where" the AI was looking when deciding the price.

- **Observation:** The heatmaps indicate the model pays attention to **lot coverage** and **greenery**.
- **Insight:** Houses with clearly defined driveways and maintained lawns (highlighted in red/yellow in the heatmap) correlated with higher valuations, confirming that "curb appeal" is a learnable feature.



Grad-CAM Heatmap showing model focus areas (Red = High Importance).

5. Results & Comparison Compared Multimodal Fusion model against a standard industry-baseline (XGBoost on tabular data only).

Model Architecture	R ² score	Performance Analysis
Tabular Only (Baseline)	0.879	Strong baseline, but limited to numerical features.
Multimodal Fusion (Late-Fusion CNN + XGB)	0.846	The model successfully integrated high-dimensional visual data. While slightly lower than the pure tabular baseline (likely due to noise in satellite images), the 0.846 score proves the pipeline successfully learned to extract meaningful features from unstructured data without failing.