SOCAL HOUSE PRICE ESTIMATOR

HYBRID MACHINE LEARNING MODEL
TABULAR DATA + IMAGES

Javier García Esteve

PROJECT MOTIVATION

Why Tabular Data Isn't Enough

- Price depends on numeric factors (bedrooms, bathrooms, sqft...)
- Location
- But also on visual factors, such as:
  - Condition & maintenance level
  - Architectural style
  - Landscaping & curb appeal
  - Overall quality

Limitation of Tabular Models

- Tabular models capture structural attributes
- But they are blind to how the property actually looks

Goal

- Combine both worlds:
  - Structured tabular data
  - Automatically extracted visual information from house images

→ A richer, more realistic home-price prediction model

DATASET OVERVIEW

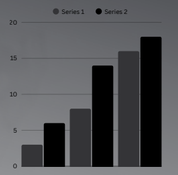
15,474 Southern California listings, each with:

- Tabular metadata:
- Bedrooms · Bathrooms · Sqft · City · Address · Price
- Target variable: price
- Property image

EDA insights:

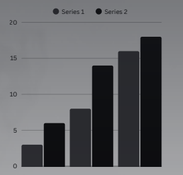
- Price highly skewed → log transform
- City dominates price variation → need good encoding

TABULAR MODELING (BASELINES)



Model	MAE	RMSE	R^2
Linear Regression	0.3097	0.3981	0.4
Random Forest	0.3073	0.3905	0.4227
XGBoost	**0.2007**	**0.286**	**0.6904**

FEATURE ENGINEERING & HYPERPARAMETER TUNING



Created features that reflect real estate logic:

- sqft_per_bed
- sqft_per_bath
- log_sqft
- total_rooms
- city_target_enc (mean log-price per city)

Feature importance shows:

- City encoding = strongest factor
- Size metrics = next strongest
- Spaciousness features add extra lift

Hyperparameter Tuning

- Gridsearch

Result:
XGBoost + FE + tuning improved to $R^2 = 0.78$

Model	MAE_test	RMSE_test	R2_test
Linear Regression (baseline)	0.309733	0.398098	0.400037
Linear Regression (feat-eng)	0.209571	0.286105	0.69012
Random Forest (baseline)	0.307297	0.390504	0.42271
Random Forest (feat-eng)	0.201842	0.27573	0.712187
XGBoost (baseline)	0.200721	0.285987	0.690375
XGBoost (feat-eng)	0.171125	0.24494	0.772876
XGBoost (feat-eng, tuned)	**0.166249**	**0.242512**	**0.777357**

WHY INCORPORATE IMAGES?

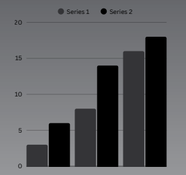
Tabular data misses key visual factors:

- Renovation quality
- Architectural style
- Condition
- Landscaping
- Curb appeal
- Luxury finishes
- Neighborhood visual cues

Users often can’t provide this information manually.
Even homeowners may not know or can't objectively describe these attributes.

But everyone can take a picture!
An image can capture many of these visual signals instantly.

Solution:
let the model learn these visual cues automatically from images.



CNN MODEL ARCHITECTURE

Transfer Learning: EfficientNetB0 tuned

- EfficientNet backbone
- Unfreeze last 10 layers
- Lower learning rate
- Strong augmentations (flip, zoom, rotation, contrast)
- Added head:
  - GlobalAveragePooling
  - Dense layer(128) → embedding
  - Dropout 0.3
  - Dense layer(1) → log-price output

CNN models performed poorly for standalone price prediction.

The 128-dim embedding is the valuable output — used as visual features inside the hybrid model.

MODELS TESTED

Model Texto	MAE_log_price	RMSE_log_price	R2_log_price
EfficientNetB0 (baseline)	0.712699	0.941124	-2.353031
EfficientNetB3	0.73529	0.985476	-2.676506
EfficientNetB0 Tuned (last 10 unfrozen)	**0.604645**	**0.781624**	**-1.312809**

HYBRID MODEL

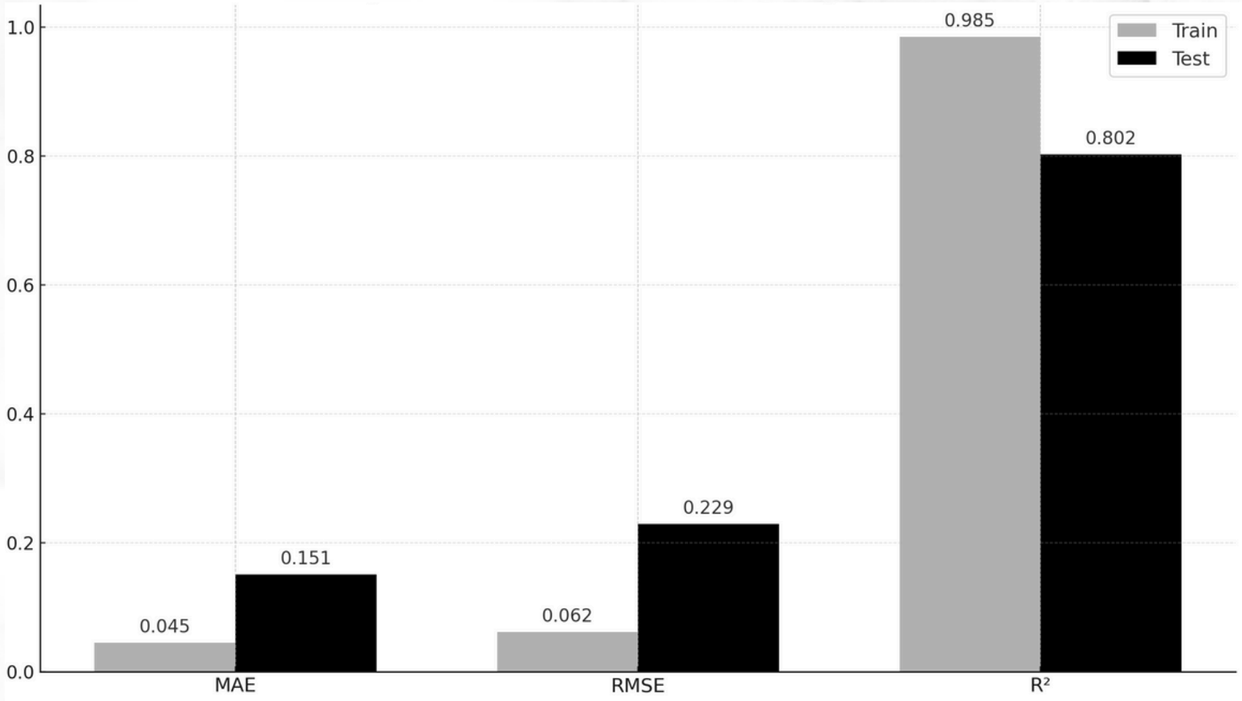
Hybrid = Tabular Features + Image Embedding → XGBoost

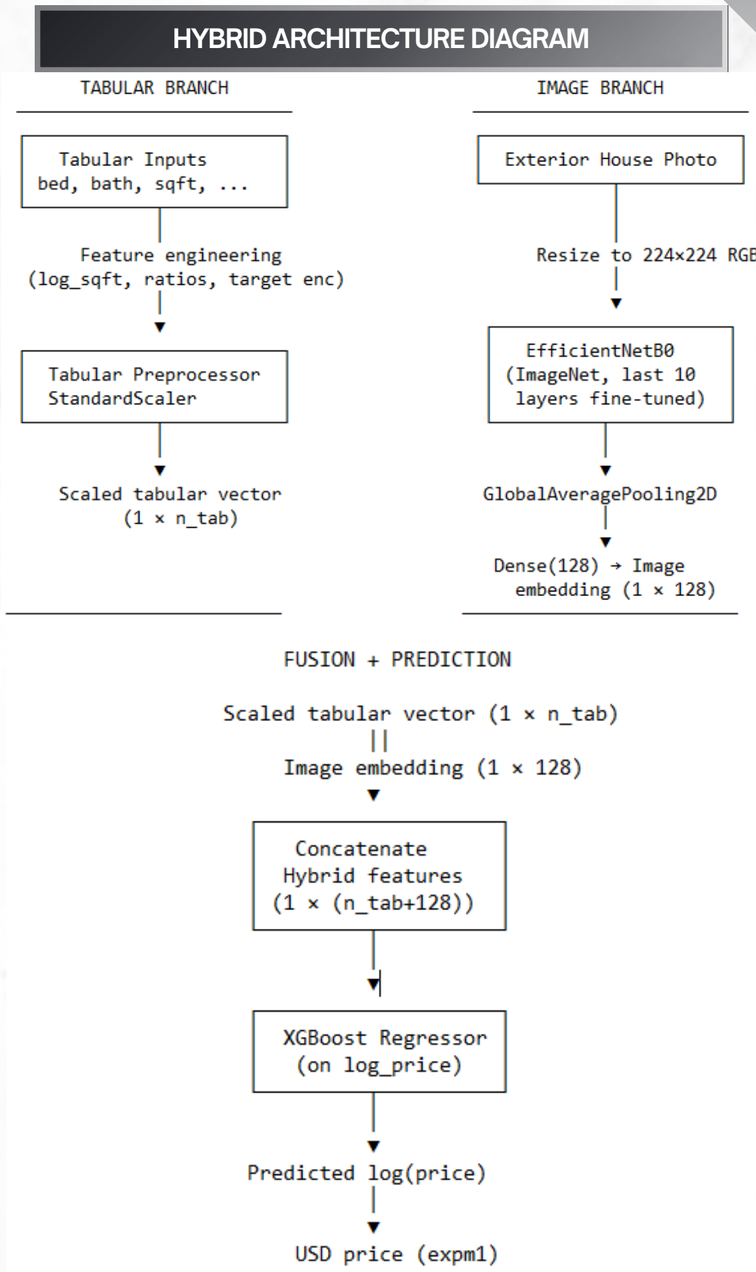
Steps:

1. Scale tabular features
2. Extract 128-dim embedding from EfficientNetB0 (fine-tuned)
3. Combine tabular features + image embedding into one vector
4. Feed this hybrid vector into XGBoost regressor
5. Predict log(price) → convert back to USD

Performance:

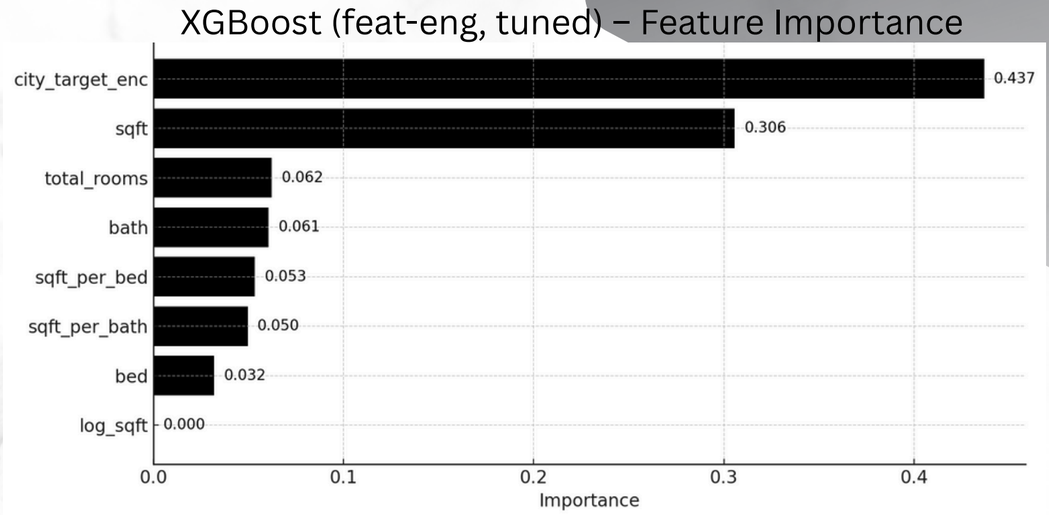
- Hybrid Model: $R^2 = 0.80$

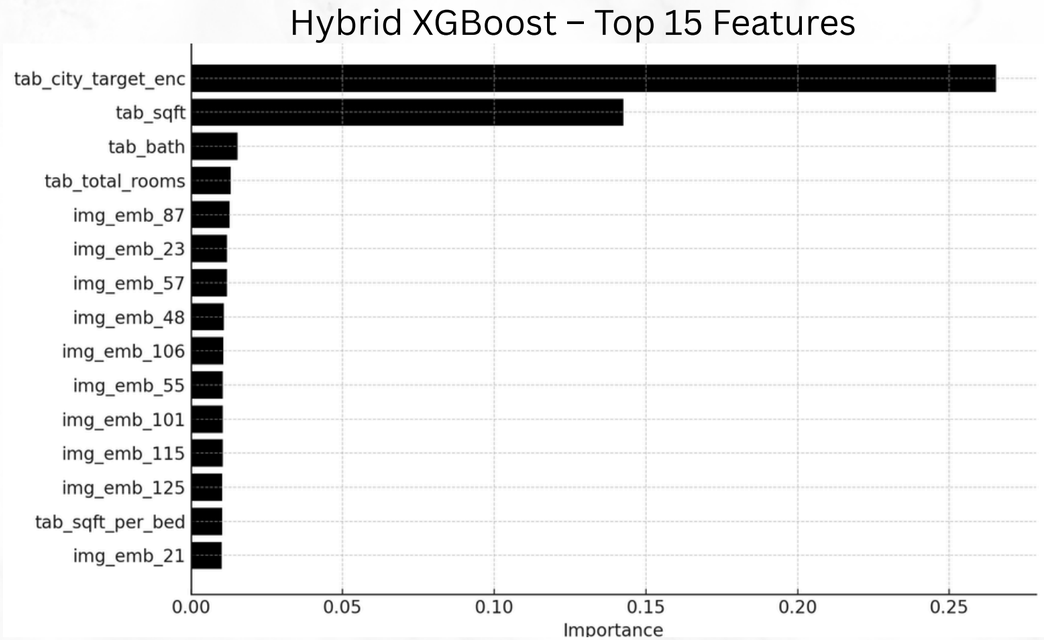




TABULAR-ONLY VS HYBRID MODEL

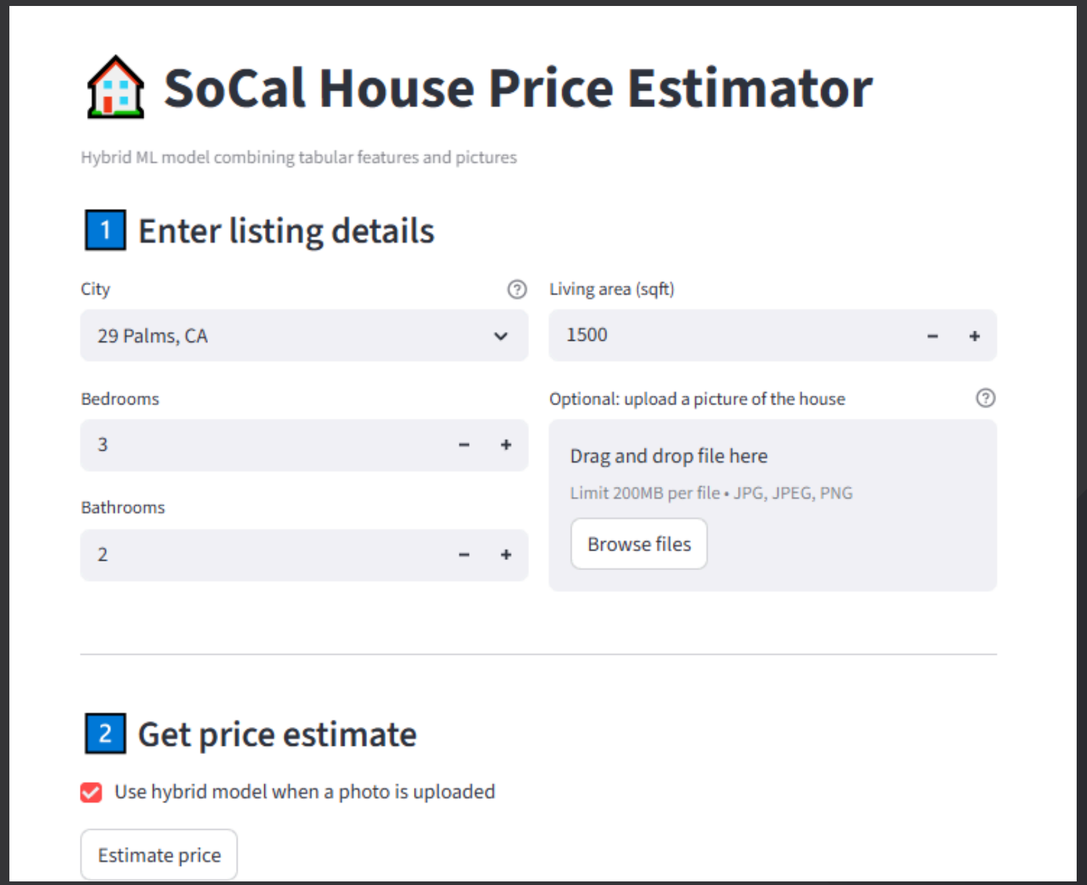
Model	MAE_test	RMSE_test	R2_test
XGBoost (feat-eng, tuned – tabular only)	0.166249	0.242512	0.777357
Hybrid (tabular FE + image embeddings)	0.150762	0.228766	0.80188





DEMO

“To make this accessible, I built a Streamlit app where the user inputs the property details and optionally uploads an image. If an image is provided you can choose to either run the hybrid model or just the tabular one. If you don’t provide an image, just the tabular-only model runs.”



CONCLUSION & FUTURE WORK

Key Conclusions

- Feature engineering significantly boosts performance
- CNN embeddings add visual intelligence
- Hybrid model is the global best performer
- Real-time app demonstrates practical application

Future Ideas

- Add **Interior pictures**
- **Multiple images per listing**
- **Satellite imagery**
- **Multi-task models (price + condition scoring)**

THANK YOU!