Wildfire Severity and Structural Damage Prediction in Wildland-Urban Interface Areas

Chanyoung Park — chp025@ucsd.edu
Jahnavi Naik — janaik@ucsd.edu
Reema Alsaeed — ralsaeed@ucsd.edu
Rushyendra Katabathuni — rkatabathuni@ucsd.edu
Mentor: Ilyes Meftah — imeftah@cotality.com
Mentor: Jonathan Sam — jonsam@cotality.com

cotality™
UC San Diego™
HALICIOĞLU DATA SCIENCE INSTITUTE

Background, Motivation & Objective

Wildfires in California increasingly threaten communities in the **Wildland–Urban Interface (WUI)**, where residential structures intersect with flammable vegetation. Despite growing concern, tools for predicting **structure-level damage** remain limited, yet are critical for **improving risk assessment, resident preparedness, insurance modeling, and mitigation planning**.

Can we predict whether an individual residential structure will be Destroyed (>50%) or experience No Damage during a wildfire event in California?

Using publicly available structural, spatial, and environmental data, we aim to:

- Develop a structure-level damage prediction model
- Identify key spatial and structural features of destruction
- Evaluate model performance across wildfire events

Data

Prediction Target

We predicted whether a structure was **Destroyed** or **Undamaged** based on post-fire inspections by **CAL FIRE Damage Inspection (DINS)** reports.

Data Sources

Damage & Structural Attributes

- **CAL FIRE Damage Inspection (DINS)** — damage classification, construction materials
- **National Structure Inventory (NSI)** — square footage, stories, build year, and site elevation

Environmental Features

- **MTBS** — burn severity class
- **LANDFIRE** — fuels and vegetation characteristics
- **NLCD** — land cover classification
- **PRISM** — weather variables

Spatial Context

- **MTBS wildfire perimeters** — fire boundaries
- **SILVIS Labs WUI dataset** — Wildland–Urban Interface classification perimeters

Methods

Data Processing

- Filtered DINS records to residential structures within WUI census blocks across 30 California wildfires
- Spatially joined structural damage records to 30m raster pixels containing environmental features
- Imputed missing categorical values, one-hot encoded categorical variables, and removed highly sparse features

Modeling & Evaluation

- Feature selection via top 95% cumulative importance on LightGBM
- Wildfire-level cross-validation grouping — test fires completely unseen during training

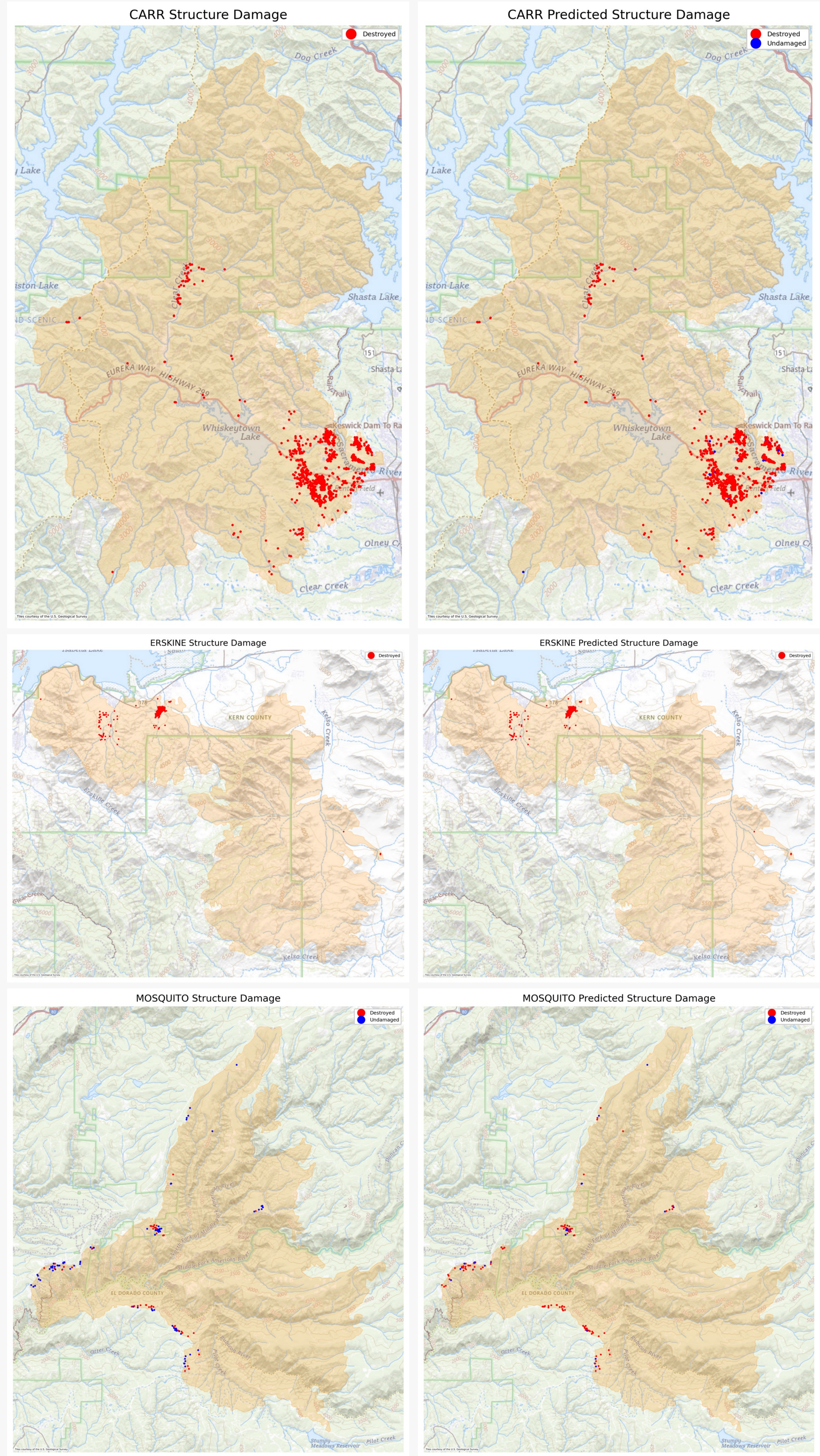
Results

Overall Model Performance

Model	Class	Precision	Recall	F1	Support
	No Damage	0.642	0.888	0.745	303
	Destroyed	0.974	0.893	0.932	1402
Random Forest	Macro avg	0.808	0.890	0.838	1705
	Weighted avg	0.915	0.892	0.898	1705
	Accuracy		*89.2%*		1705
	No Damage	0.824	0.508	0.629	303
	Destroyed	0.902	0.976	0.938	1402
Neural Network	Macro avg	0.863	0.742	0.783	1705
	Weighted avg	0.888	0.893	0.883	1705
	Accuracy		*89.3%*		1705

In terms of real-world applications of our models, the **Random Forest is better at identifying structures likely to survive**, catching more undamaged structures but producing more false alarms for destruction. In contrast, the **Neural Network is better at identifying structures likely to be destroyed**, more reliably detecting high-risk structures but missing more undamaged ones.

Spatial Prediction Visualization



Left: damage outcomes (red = Destroyed, blue = Undamaged). Right: predicted damage class. Predictions from Neural Network on **Carr, Erskine, and Mosquito Fires**.

Structural Loss Estimates by Fire Event

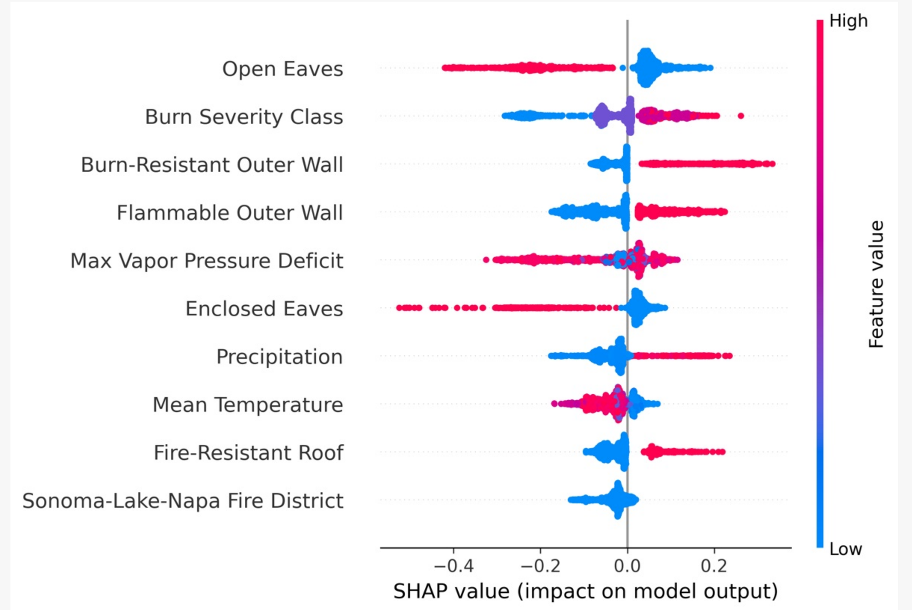
Predicted structural loss is computed by comparing actual structural value lost with property values of model-predicted destroyed structures.

Structural value loss by fire event ($M). Difference = Actual − Predicted; positive = under-prediction.

			Random Forest		Neural Network	
Fire	**Total ($M)**	**Actual ($M)**	**Pred.**	**Diff.**	**Pred.**	**Diff.**
Carr	406.52	406.52	350.50	+56.02	398.62	+7.90
Mosquito	41.20	11.90	13.05	−1.16	31.67	−19.77
Erskine	38.05	38.05	36.89	+1.16	38.05	0.00

Feature Importance (SHAP)

SHAP values reveal the most influential predictors.



SHAP beeswarm plot (Neural Network). Each point is one structure; pink = high feature value, blue = low. Features sorted by mean |SHAP|.

Enclosed eaves are strongly protective, consistently reducing predicted destruction probability. **Open eaves** also appear protective in our model, though this could be due to those structures being concentrated in lower-risk areas. Conversely, structures with **burn-resistant outer walls** and **fire-resistant roofs** show increased destruction risk, which could likely be attributed to those structures being located in high-risk areas. We can observe that **burn severity class** is an influential environmental predictor, with higher severity strongly amplifying destruction risk.

Conclusion

- **Structure-level wildfire damage** can be predicted with strong performance using publicly available structural, environmental, and spatial data
- **Structural characteristics** (eaves configuration, siding type, roof construction, etc.) and **burn severity** consistently emerge as influential predictors
- These results demonstrate the value of **structure-level modeling** for understanding structure damage in the Wildland–Urban Interface
- By identifying high-risk structures and influential factors, this approach provides **actionable insights** to support risk assessment, preparedness planning, insurance evaluation, and mitigation strategies

Project Website QR Code