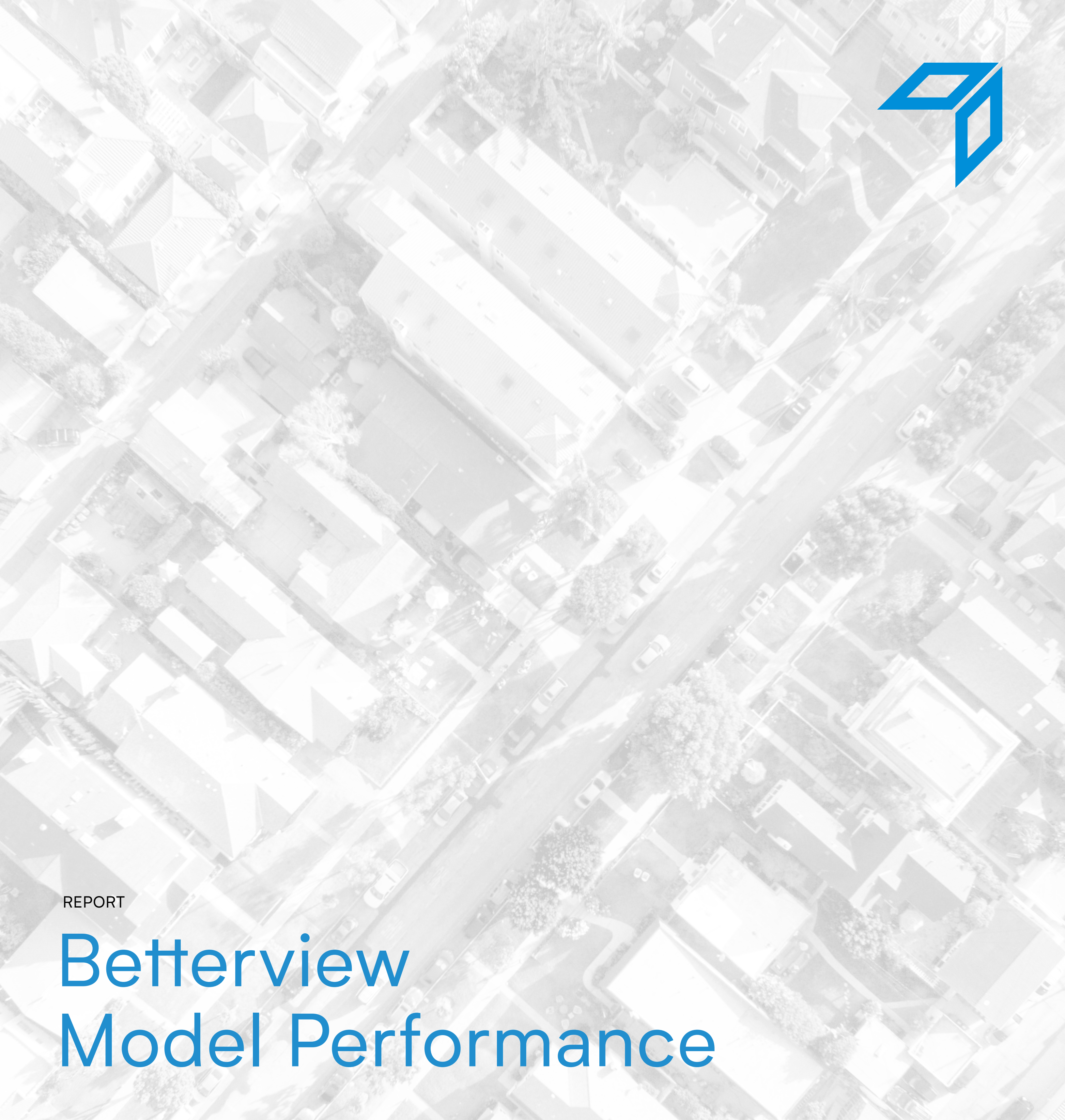
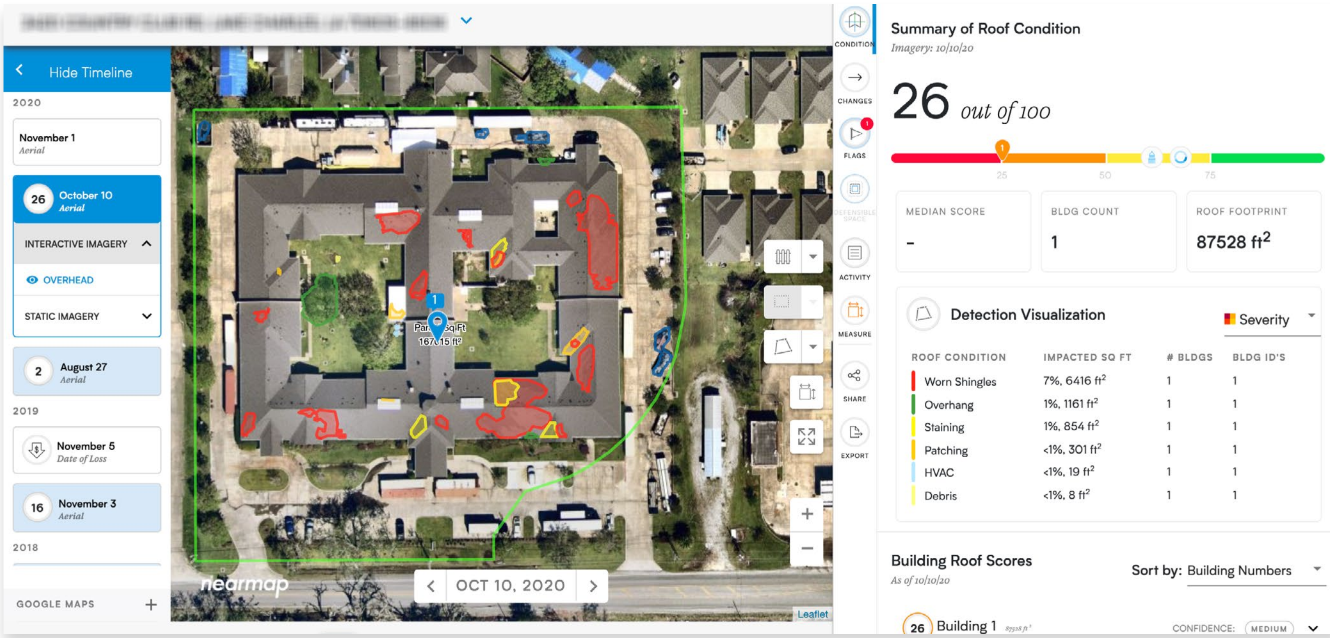


REPORT

Betterview Model Performance

THE STATISTICAL PERFORMANCE OF BETTERVIEW'S PROPRIETARY COMPUTER VISION MODELS



This document presents the statistical performance of Betterview's proprietary computer vision models used in PropertyInsight. The numbers are calculated at the property level, meaning that they measure each model's ability to detect the presence or absence of a condition or feature across an entire property.

The test samples used to compile these metrics were compiled using the following principles:

- Balanced distribution – each dataset contains an even number of positive examples (images that contain the condition) and negative examples (images that do not contain the condition). This ensures that the performance metrics are comparable across models and not distorted by the relative frequencies of these different conditions in the real world. For example, structural damage is much less common "in the wild" than overhang; if we were to measure the performance on datasets that reflect this real-world distribution, the structural damage model would have much worse precision than overhang due solely to how rare it is.

- Representative images – within each positive and negative set, examples are representative of the types of occurrences to occur in the real world. This is distinctly different than measuring performance on the dataset used to train the model, which typically contains a higher percentage of "challenging" examples (e.g., blue roofs for the tarp model) in order to effectively teach the model about these edge cases.

- Sample size – each dataset contains a minimum of 1,000 images to ensure statistical significance of the metrics and minimize uncertainty.

We measure the performance of each model by comparing the model predictions to a ground truth dataset created by a team of expert labelers. The labelers have been trained using detailed technical documentation of each condition and thorough review of tens of thousands of images. Once trained, they identify these conditions through detailed visual inspection of the imagery.

We present the performance of each model using two complementary metrics: Precision and Recall:

- Precision represents the percent of the time that the model prediction is correct when it says that a condition is present

- Recall represents the percent of real instances of the condition that are picked up by the model

A higher precision means fewer false positives. A higher recall means fewer false negatives.

The uncertainty in each metric is also provided to convey the variability in results based on the sample used. A model with precision of 96% and uncertainty of ±2% can be interpreted as: we have 90% degree of confidence that the precision of the model is between 94% – 98%.

The following table reflects the current model suite underlying Betterview's PropertyInsight product. We routinely improve our models as we grow our training sets and advance our machine learning algorithms, and as we do we update this table to reflect the new versions.

Group	*Model*	*Precision*	*Precision Uncertainty*	*Recall*	*Recall Uncertainty*
Roof/Property Maladies	Tarp*	99%	±1%	87%	±5%
	Structural Damage	98%	±<1%	92%	±2%
	Ponding*	92%	±1%	93%	±2%
	Rust	92%	±<1%	86%	±3%
	Missing Shingles	97%	±<1%	87%	±2%
	Patching	88%	±1%	67%	±2%
	Overhang	86%	±1%	92%	±2%
	Staining	85%	±1%	98%	±1%
	Worn Shingles*	85%	±<1%	97%	±1%
	Roof Debris	79%	±2%	68%	±3%
Roof/Property Features	Trampoline	99%	±<1%	91%	±2%
	Solar Panels	98%	±<1%	94%	±1%
	Swimming Pool*	98%	±<1%	97%	±1%
	HVAC*	90%	±1%	90%	±2%
	Yard Debris	52%	±1%	78%	±2%
	Water Hazard	98%	±<1%	85%	±2%
Roof Material	Asphalt Shingles	95%	±1%	93%	±1%
	Ballasted	94%	±1%	94%	±1%
	Metal Panel	93%	±1%	96%	±1%
	Clay Tile	92%	±2%	92%	±3%
	EPDM	92%	±1%	96%	±2%
	PVC/TPO	92%	±2%	92%	±2%
	Modified Bitumen	88%	±4%	80%	±3%
Roof Shape	Hip	96%	±1%	94%	±1%
	Flat	95%	±1%	94%	±1%
	Gable	89%	±1%	93%	±1%

* Performance metrics for these models are based on a dataset containing less than 1,000 images.

AUTHOR

Madeleine Lopeman

LEAD DATA SCIENTIST

Madeleine leads Betterview's Data Science team, which crafts datasets used to train Betterview's models, evaluates models to measure and improve performance, and conducts studies to quantify the value of Betterview's data for insurance carriers. Madeleine received her PhD from Columbia University and has held positions at Guy Carpenter and Blackboard Insurance.