Inferknow

Better Response to Wildfires

Bryan Morgan - Don Moon - Francis Leung - Ray Xiao

Mission

Provide California's Emergency Coordinators with the Best Information to Quickly Save Lives and Property

Problem

1. Wildfires are highly destructive and disruptive
    a. 2018 California wildfires caused $400 billion in damages

2. Becoming more frequent and severe, to date in 2020:
    a. ~35% more than the 5-year average, 7,786 reported fires
    b. 5x more severe than 5-year average, with ~2.2m acres burned

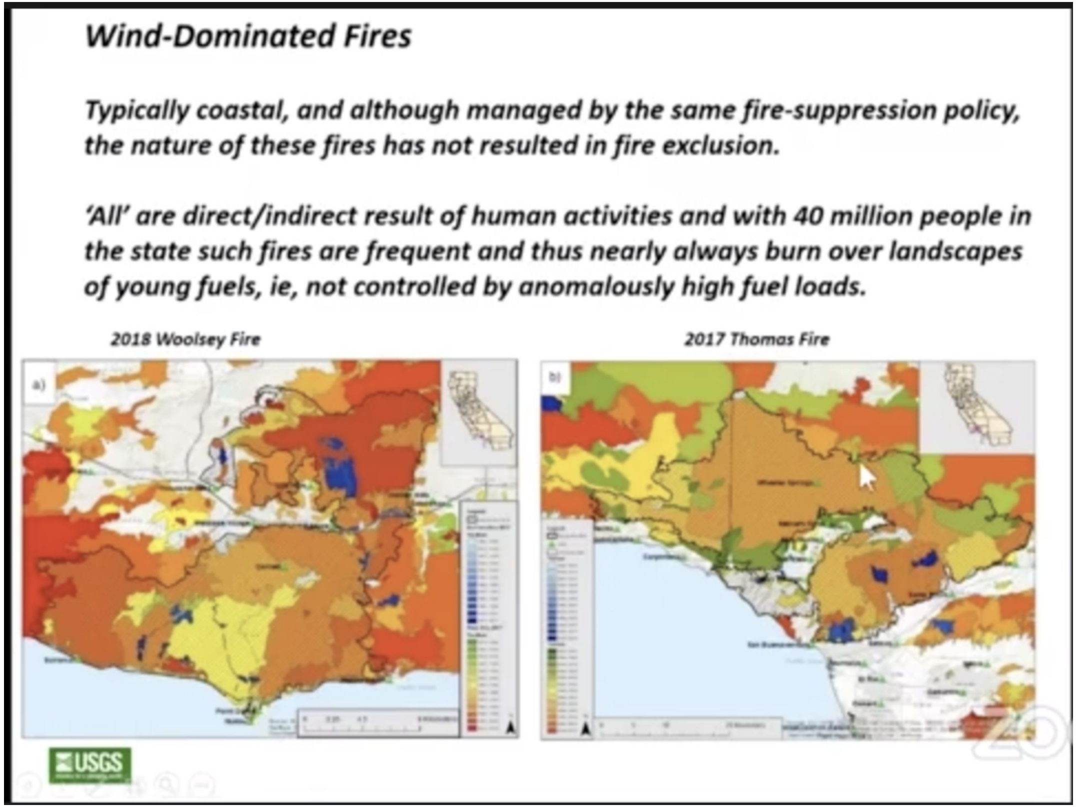
3. Emergency Response Coordinators have to be well informed to combat these natural disasters

Target Customer and Proposed Solution

- **Customer:** *California Geographic Area Coordination Center (GACC)* whose principal mission is to provide cost-effective and timely coordination for the response for wildfires.

- **Our Product:**
    - An integrated solution that will conduct the following after a reported fire:
        - Model fire propensity (size)
        - Scan satellite imagery for Areas of Interest (Buildings, important infrastructure etc.)
        - Indicate areas of most urgent action (most populated or most critical infrastructure etc.)

Focus on SoCal Fires



SoCal has the unique aspect of a drought season, wind-dominant fires, and always burn over landscapes.

NorCal you can explain ~ 50% of the fire size variation with just temperature, while prior year precipitation and high winds have more of an effect in SoCal

Solution Deep Dive

Data Sources

Fire Propensity:

- 1998 - 2015 California Wildfire Occurrences (Kaggle)
- 1998 - 2015 Hourly weather data (NOAA)

Computer vision:

- Satellite images, labelled for buildings and roads (Spacenet)
- Raw satellite footage (Planet Labs)

Baseline Fire Propensity Model Features

Hourly Weather Data

- Temperatures
- Relative Humidity
- Wind Speed

Rolling Average/Sum of Weather Data

- Rolling 7, 30, 60, 90 Day Average Temperatures
- Rolling 7, 30, 60, 90 Day Average Relative Humidity
- Rolling 7, 30, 60, 90 Day Sum Average Relative Humidity
- Average 2 Day Wind Speed Average

Fire Cause Description

Discovery Month/ Hour

Propensity Model Initial Findings

Total Records: 36K
Split: 80% Train, 20% Test

7-Class (A-G) Fire Propensity Model Performance

Model	F1-Score Macro	F1-Score Micro (Accuracy)	F1-Score Weighted
GBT Classifier	.13	.5937	.4949
Random Forest Classifier	.107	.606	.457
XGBoost	.1707	.59315	.5507

Takeaways:

- Data is imbalanced with small fires (ABC) accounting for 95%
- Multi-class classifiers shows poor performance

Additional Features to Propensity Model

Count the number of acres burned nearby a location of interest for the past 3 and 12 months

- 0.01, 0.07, 0.22, 0.38, 0.7, 1.6, 3 miles.
- These miles match the the size of the each fire class (A,B,C,D,E,F,G).
- **Rationale**: more fires occured in the past nearby, less fuels to burn and less likely to have large fires
- Theses features are proxies for how much fuels left to burn nearby

Fire Propensity Iterations (2-Class)

Category	Metrics	Baseline Model	Baseline Model (Oversampling)	Latest Model
ABC	Precision	.97	.98	.98
	Recall	1	.56	.71
	F1 Score	.99	.71	.83
DEFG	Precision	.10	.05	.06
	Recall	.01	.69	.63
	F1 Score	.01	.09	.12

Computer Vision Modeling

- Retrained a pre-trained ensemble model from SpaceNet
- Scoring metric IOU
- SP2 models (1-3) did not extend well to Planet Data

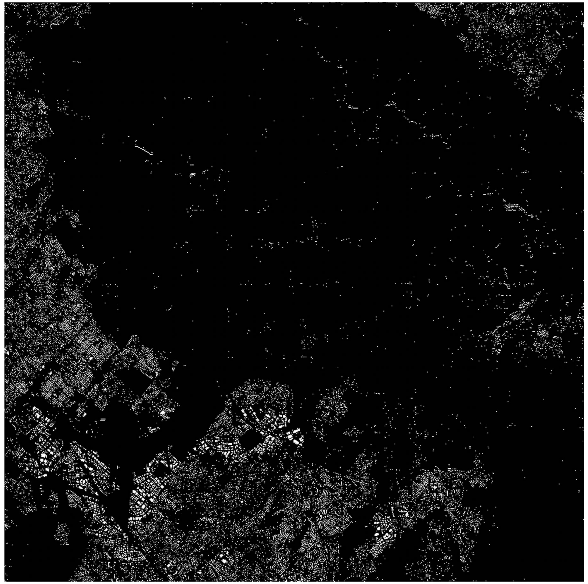
	Cloud Training		Testing			Extending to Planet Data			
Model name	Dataset (# images)	Training Time	Precision	Recall	F-1 score	True Positives	Precision	Recall	F-1 score
inferknow_v1	SP2 Vegas (3,100)	10 epochs 3 hours	0.90	0.82	0.86	0	0.00	0.00	0.00
inferknow_v2	SP2 Vegas (3,100)	20 epochs 6 hours	0.90	0.85	0.87	0	0.00	0.00	0.00
inferknow_v3	SP2 All (7,500)	25 epochs 16 hours	0.86	0.75	0.80	1	0.33	0.00	0.00
inferknow_v4	SP7 (1,100)	30 epochs 4 hours	0.11	0.05	0.07	253	0.11	0.02	0.00
inferknow_v5	SP7 (1,200)	50 epochs 7 hours	0.12	0.04	0.08	225	0.12	0.02	0.03

Computer Vision Model Output

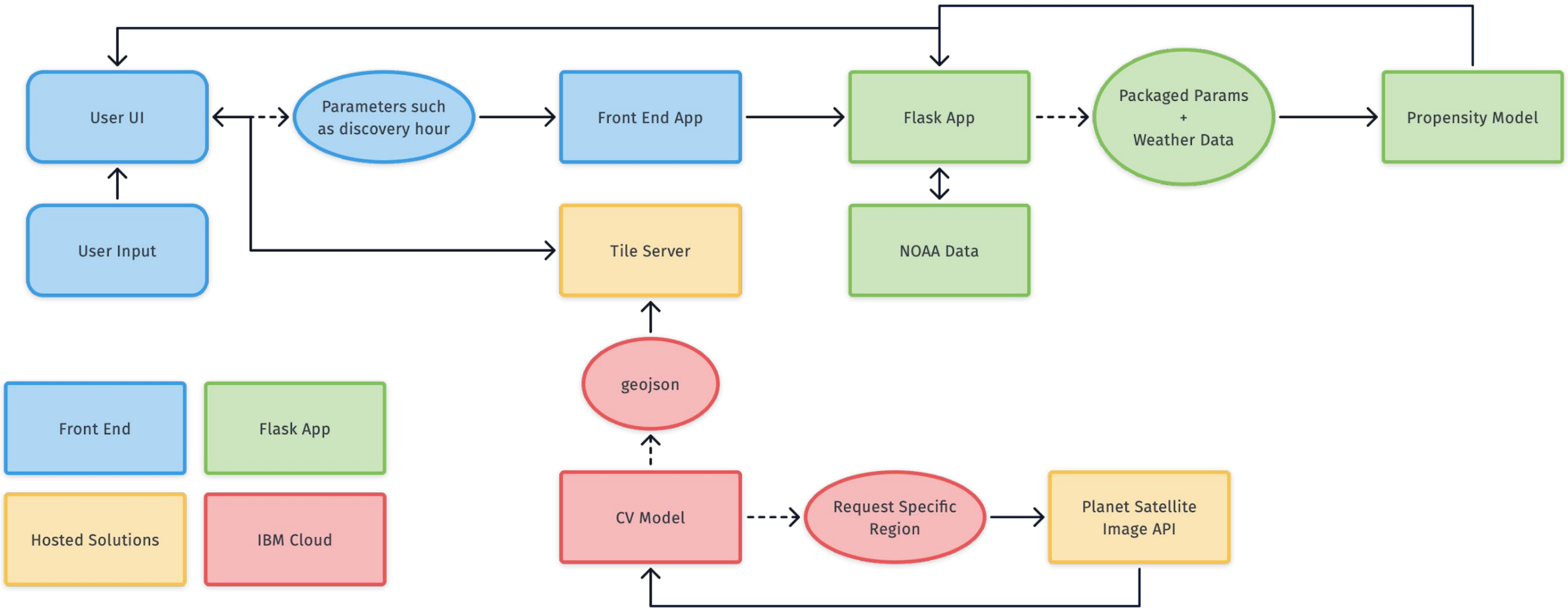
Original Image

Inferknow_v3

Inferknow_v5 (final)



System Design and Infrastructure



Product Demo

References

- CALFIRE (2020). Stats and Events. California
  https://www.fire.ca.gov/stats-events/

- NOAA (2020). Land-Based Station Data. Asheville, NC.
  https://www.ncdc.noaa.gov/data-access/land-based-station-data

- Planet Team (2017). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. https://api.planet.com

- Short, Karen C. 2017. Spatial wildfire occurrence data for the United States, 1992-2015 [FPA*FOD*20170508]. 4th Edition. Fort Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2013-0009.4

Appendix: Team Responsibilities

Bryan Morgan ● System design ● CV datasets + framework research ● Planet API EDA ● Infrastructure development ● Front-end development	**Don Moon** ● Fire propensity infrastructure ● Fire propensity data collection/engineering ● Fire propensity modeling
Francis Leung ● EDA: Computer vision (Spacenet Data) ● Literature review ● Integration with Planet API to pull new satellite images ● Developed computer vision model (acquiring training data, set up cloud training, testing and iterations, creating final geojson outputs)	**Ray Xiao** ● Wildfire research ● Fire propensity EDA ● Fire propensity modeling

Future Steps

- Propensity:
  - Improve model performance with more historical and topographical features

- Computer Vision:
  - Additional AOIs
  - Better performance (accuracy and time)

- Infrastructure
  - We would like to automate triggers
  - We would also like to pre-process building GeoJSON data to supply to tile server

Limitations of Existing Research

- Over 300 academic papers on fire prediction have been published
- Fire size prediction research is limited by sample data size and lack of California-specific focus
- No integrated solution