Disaster Scene Description and Indexing (DSDI)

Asad Anwar Butt, National Institute of Standards and Technology; Johns Hopkins University
George Awad, National Institute of Standards and Technology; Georgetown University
Jeffrey Liu, MIT Lincoln Laboratory
William Drew, Office of Homeland Security and Preparedness

National Institute of Standards and Technology
U.S. Department of Commerce

Introduction - DSDI

- Video and imagery data can be extremely helpful for public safety operations.
- Natural Disasters, e.g.,
  - Wildfire
  - Hurricanes
  - Earthquakes
  - Floods
- Man-made Disasters, e.g.,
  - Hazardous material spills
  - Mining accidents
  - Explosions

WEATHER

Weather Fury: The 18 billion-dollar weather disasters of 2021

NOAA recently determined that the United States experienced a record 18 weather disasters during the first nine months of 2021.

Author: **Christopher P Laudicina**
Published: **5:02 PM EDT October 15, 2021**
Updated: **5:02 PM EDT October 15, 2021**

WASHINGTON — The National Oceanic and Atmospheric Administration (NOAA) recently determined that the United States experienced 18 weather disasters during the first nine months of 2021, each of which caused at least one billion dollars in damages.

The total cost of these disasters was in excess of $104 billion in damages. Hurricane Ida was the costliest natural disaster of 2021, with an estimated $60 billion price tag, making it one of the five costliest hurricanes in U.S. history.

https://www.wusa9.com/article/weather/most-expensive-weather-events-2021/65-86e708cf-666e-4494-a2ef-b4b11b942463

NEWS

US experiences record-breaking natural disasters in 2021

A car flooded on a local street as a result of the remnants of Hurricane Ida is seen in Somerville, N.J. Thursday, Sept. 2, 2021. (AP Photo/Eduardo Munoz Alvarez)

Posted: Oct 24, 2021 / 05:47 PM EDT / Updated: Oct 24, 2021 / 05:48 PM EDT

(WSYR)- Another year, another potentially record-breaking number of natural billion-dollar disasters in the United States.

https://www.news10.com/news/us-experiences-record-breaking-natural-disasters-in-2021/

Introduction - DSDI

- Prior knowledge about affected areas can be very useful for the first responders.
- Oftentimes, the communication systems go down in major disasters, which makes it very difficult to get any information regarding the damage.
- Automated systems to gather information before rescue workers enter the area can be very helpful.

Introduction - DSDI

- Computer vision capabilities have rapidly advanced recently with the popularity of deep learning.
  - Research groups have access to large image and video datasets for various tasks.
- However, the capabilities do not meet public safety needs.
  - Lack of relevant training data.
- Most current image and video datasets have no public safety hazard labels.
  - State-of-the-art systems trained on such datasets fail to provide helpful labels.

Training Dataset

- In response, the MIT Lincoln Laboratory developed a dataset of images collected by the Civil Air Patrol of various natural disasters.
- The Low Altitude Disaster Imagery (LADI) dataset was developed as part of a larger NIST Public Safety Innovator Accelerator Program (PSIAP) grant.
- Two key properties of the dataset are:
  - Low altitude
  - Oblique perspective of the imagery and disaster-related features.
- The DSDI test data and ground truth from 2020 is also available for teams to use as training data.

Training Dataset

- LADI Dataset*:
  - Hosted as part of the AWS Public Dataset program.
  - Consists of 20,000+ human annotated images with ~500,000 machine annotated images
  - The images are from locations with FEMA major disaster declaration for a hurricane, earthquake, or flooding.
  - Lower altitude criteria distinguishes the LADI dataset from satellite datasets to support development of computer vision capabilities with small drones operating at low altitudes.
  - A minimum image size was selected to maximize the efficiency of the crowd source workers; lower resolution images are harder to annotate.
- 2020 DSDI Test Set:
  - 5 hours of video.
  - Segmented into small video clips (shots) of maximum 20 sec.
  - Videos are from earthquake, hurricane, and flood affected areas.
  - Total number of shots: 1825

*LADI dataset documentation and basic tutorials are hosted on GitHub: https://github.com/ladi-dataset

Test Dataset

- A test dataset of about 6.7 hours of video was distributed for this task.
- The test dataset was segmented into small video clips (shots) of a maximum of 20 sec, with a median length of 8.34 sec.
- The videos are from earthquake, flooding, fire, and erosion affected areas.
- Consisted of a mix from international and domestic sources.
- Total number of shots: 2801

Testing Data: Example Videos

Testing Dataset - Categories

- Hierarchical labeling scheme: 5 coarse categories, each with 4-9 more specific annotations.

Damage	Environment	Infrastructure	Vehicles	Water
Misc. Damage	Dirt	Bridge	Aircraft	Flooding
Flooding/Water Damage	Grass	Building	Boat	Lake/Pond
Landslide	Lava	Dam/Levee	Car	Ocean
Road Washout	Rocks	Pipes	Truck	Puddle
Rubble/Debris	Sand	Utility Or Power Lines/Electric Towers		River/Stream
Smoke/Fire	Shrubs	Railway		
	Snow/Ice	Wireless/Radio Communication Towers		
	Trees	Water Tower		
		Road		

J. Liu, D. Strohschein, S. Samsi and A. Weinert, "Large Scale Organization and Inference of an Imagery Dataset for Public Safety," 2019 IEEE High Performance Extreme Computing Conference (HPEC), 2019, pp. 1-6, doi: 10.1109/HPEC.2019.8916437

Annotation

- We used full time annotators instead of crowdsourcing.
- For each category, a practice page was created.
- This page included multiple examples for each label.
- The annotators were also given 2 videos as a test to mark the labels visible in them.
- This allowed the annotators to become familiarized with the task and labels before starting a category.

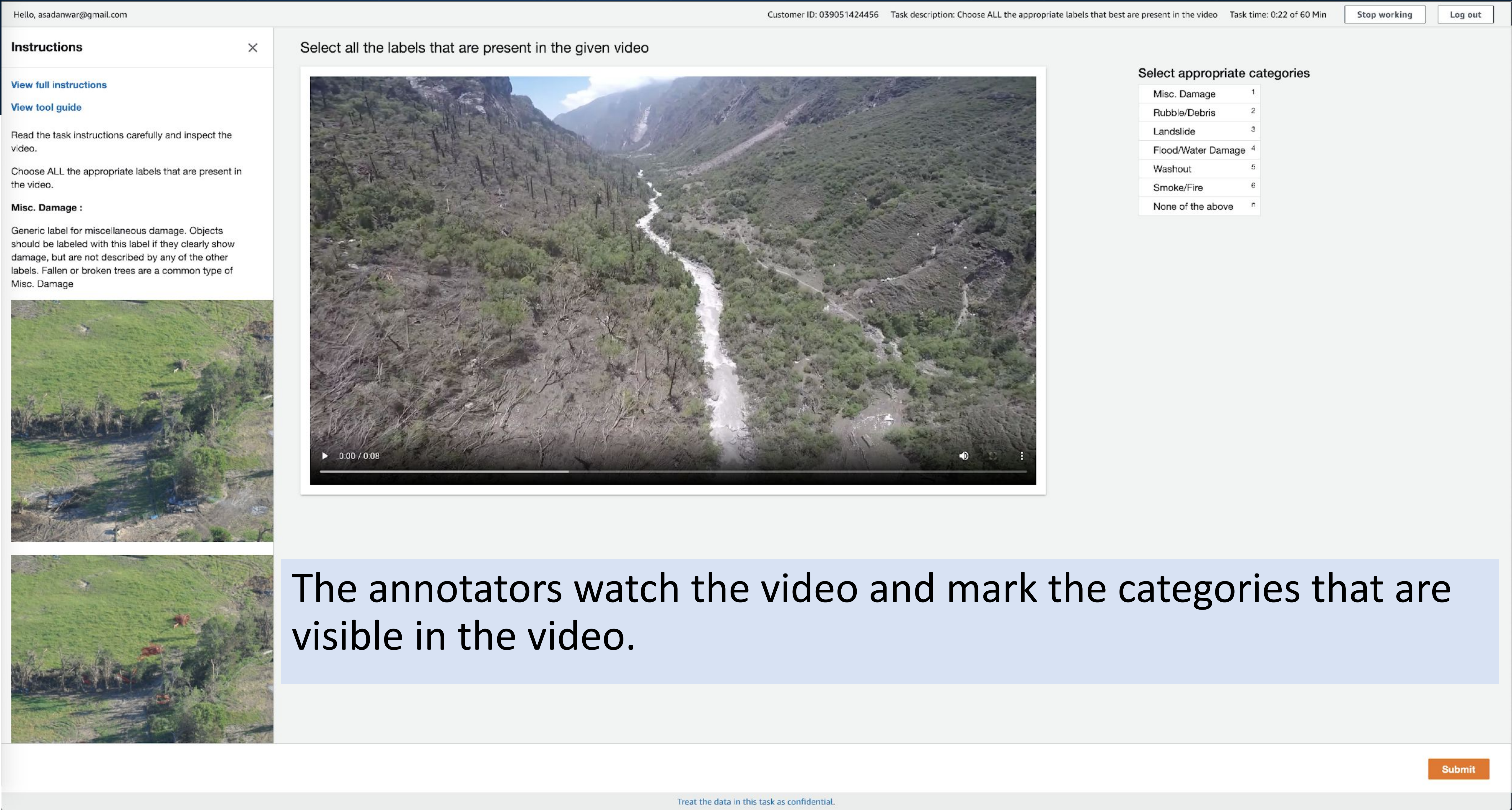
Annotation

NIST

- We had 2 full time annotators to annotate the testing dataset.
- Both annotators had also worked on the pilot task in 2020.
- We used the Amazon Augmented AI (Amazon A2I) tool.

NIST Disclaimer: Any mention of commercial products within this presentation is for information only; it does not imply recommendation or endorsement by NIST.

- The annotators worked independently on each category.
- For each coarse category, they marked all the specific labels that were present in the video.
- To create the final ground truth, for each shot, the union of labels were used.



The annotators watch the video and mark the categories that are visible in the video.

DSDI Task

- Systems are required to return a ranked list of up to 1000 shots for each of the 32 features.
- Each submitted run specified its training type:
  - LADI-based (L): The run only used the supplied LADI dataset for development of its system.
  - Non-LADI (N): The run did not use the LADI dataset, but only trained using other dataset(s).
  - LADI + Others (O): The run used the LADI dataset in addition to any other dataset(s) for training purposes.

Submissions

NIST

1	Number of Teams

3

2	Number of Submissions

12

3	LADI-Based Training

8

4	LADI + Others

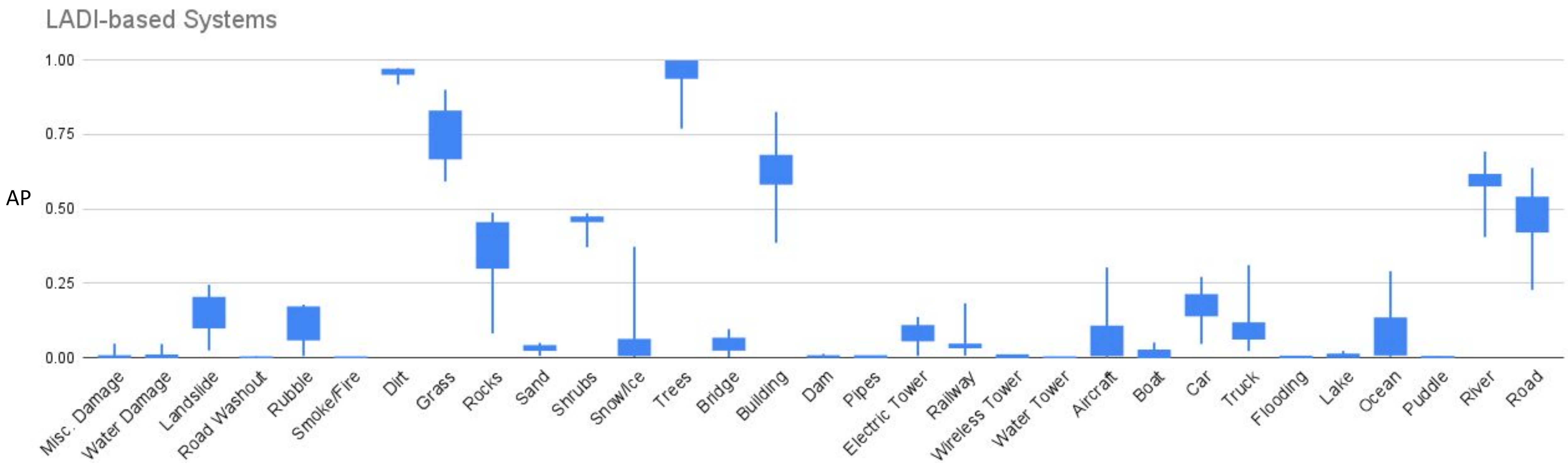
4

Evaluation Metrics

- The following evaluation metrics were used to compare the submissions:

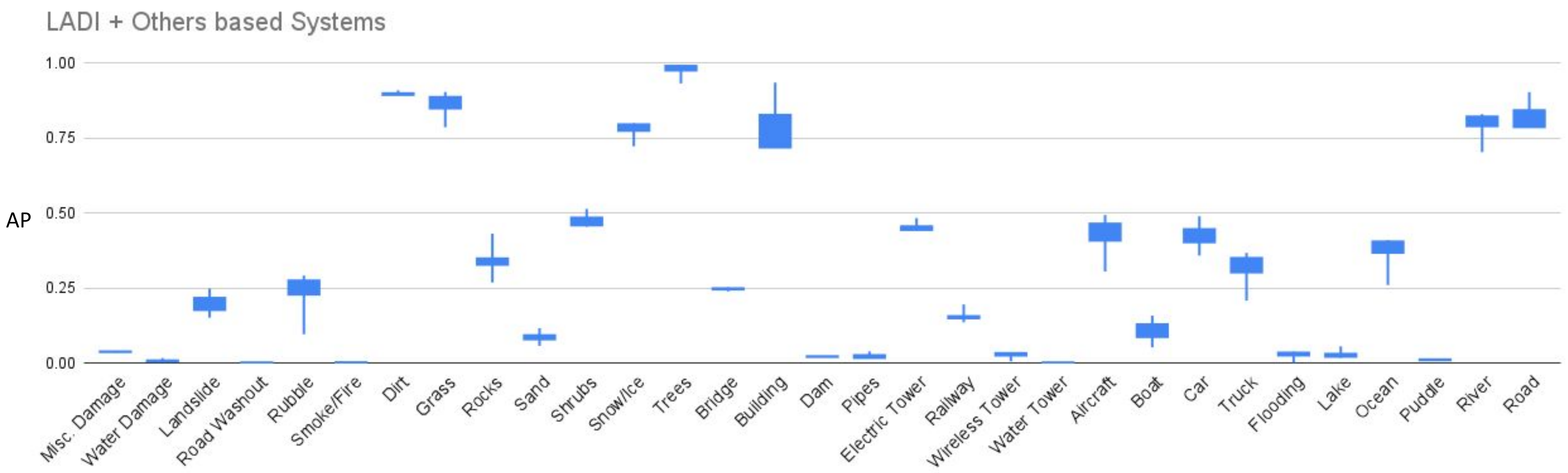
Metric	Description
Speed	Clock time per inference (reported by participants).
Mean Average Precision (MAP)	Average precision is calculated for each feature, and the mean average precision reported for a submission.
Recall	True positive, true negative, false positive, and false negative rates.

Results by Features (LADI based)



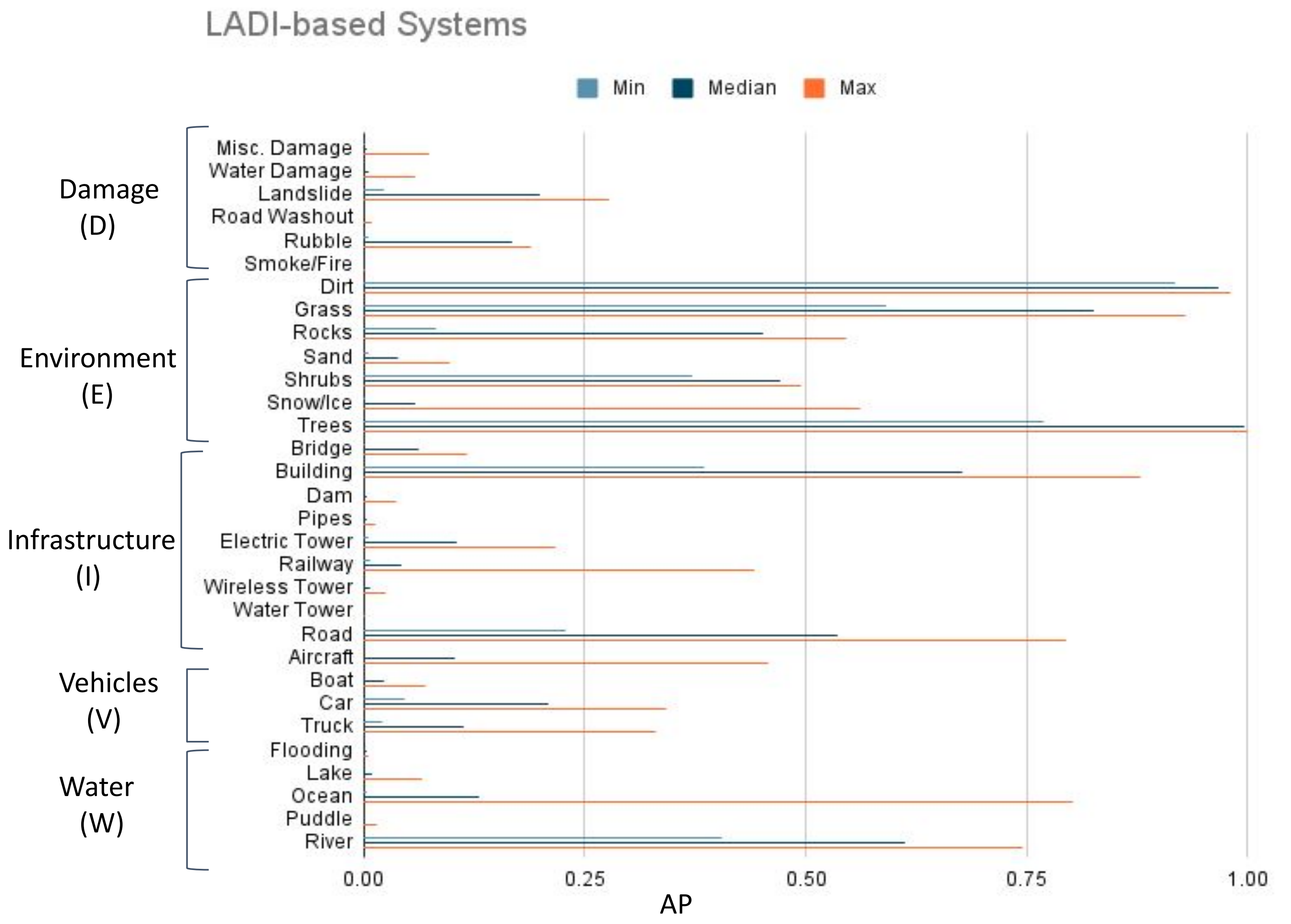
- Average precision values for each feature categorized by training type.
- 8 LADI-based runs.

Results by Features (LADI + Others based)

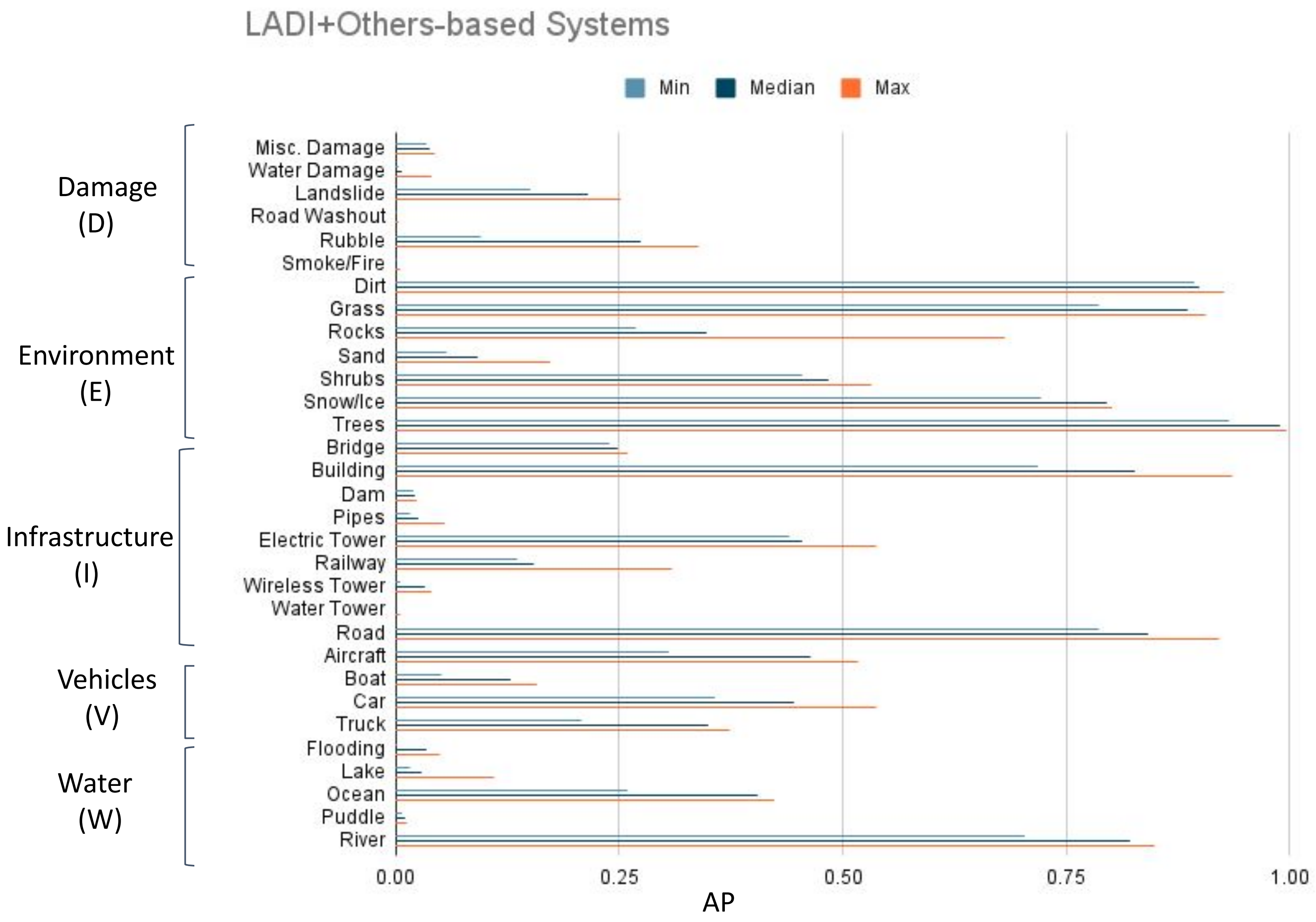


- Average precision values for each feature categorized by training type.
- 4 LADI+Others-based runs.

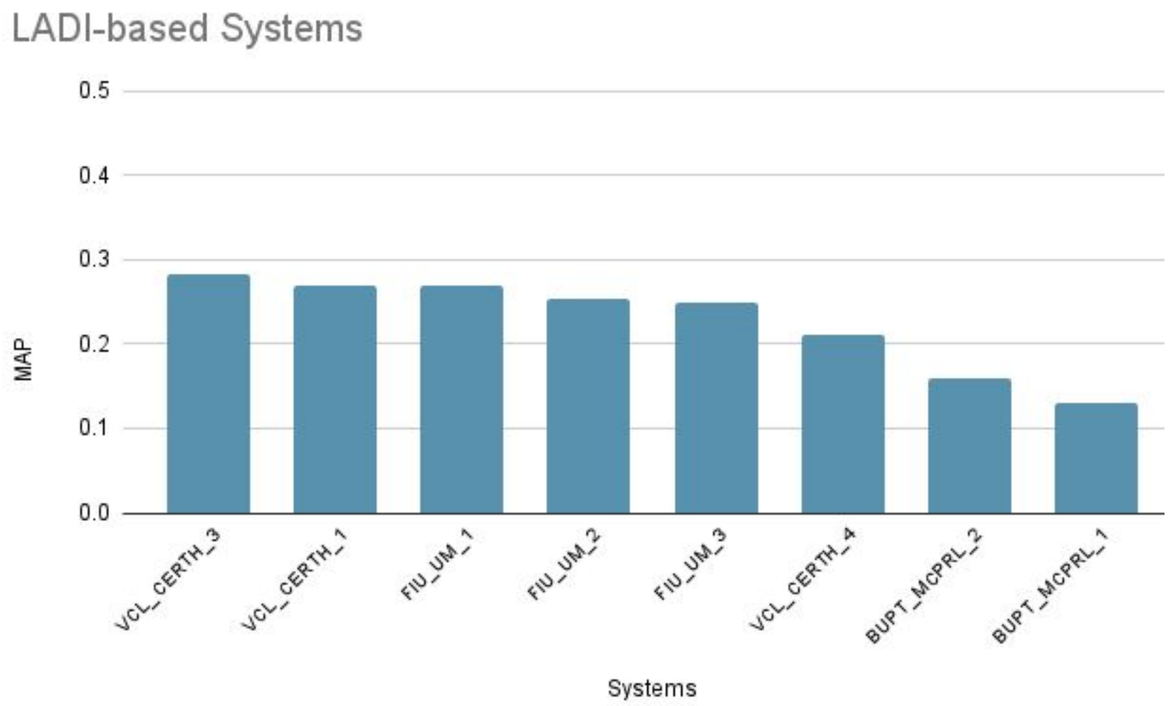
Results by Categories (LADI-Based)

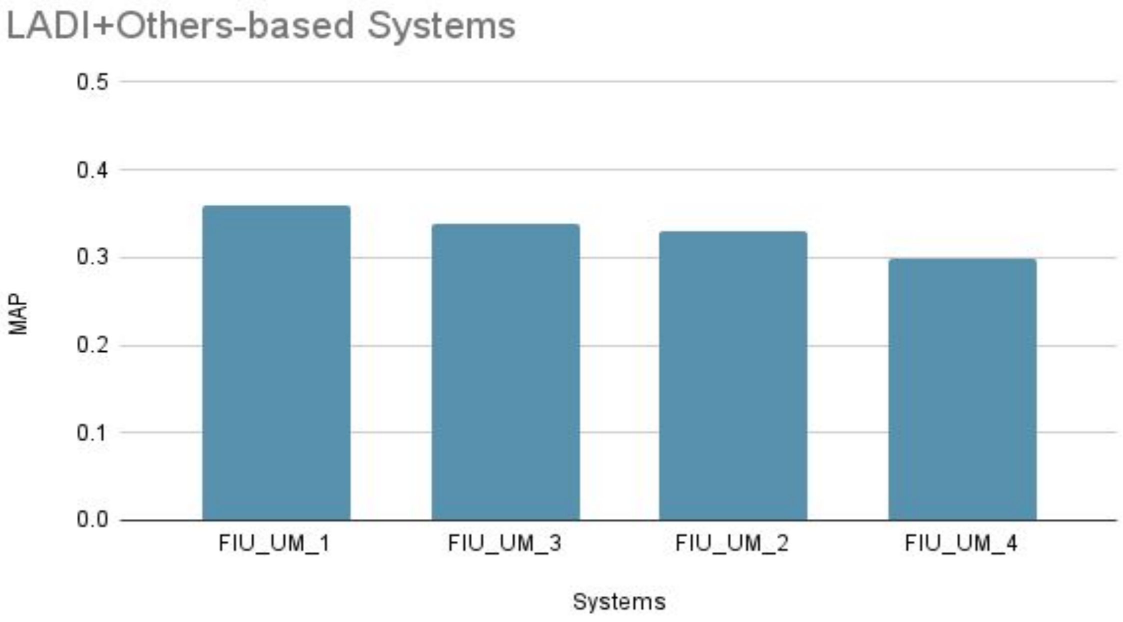


Results by Categories (LADI+Others-based)

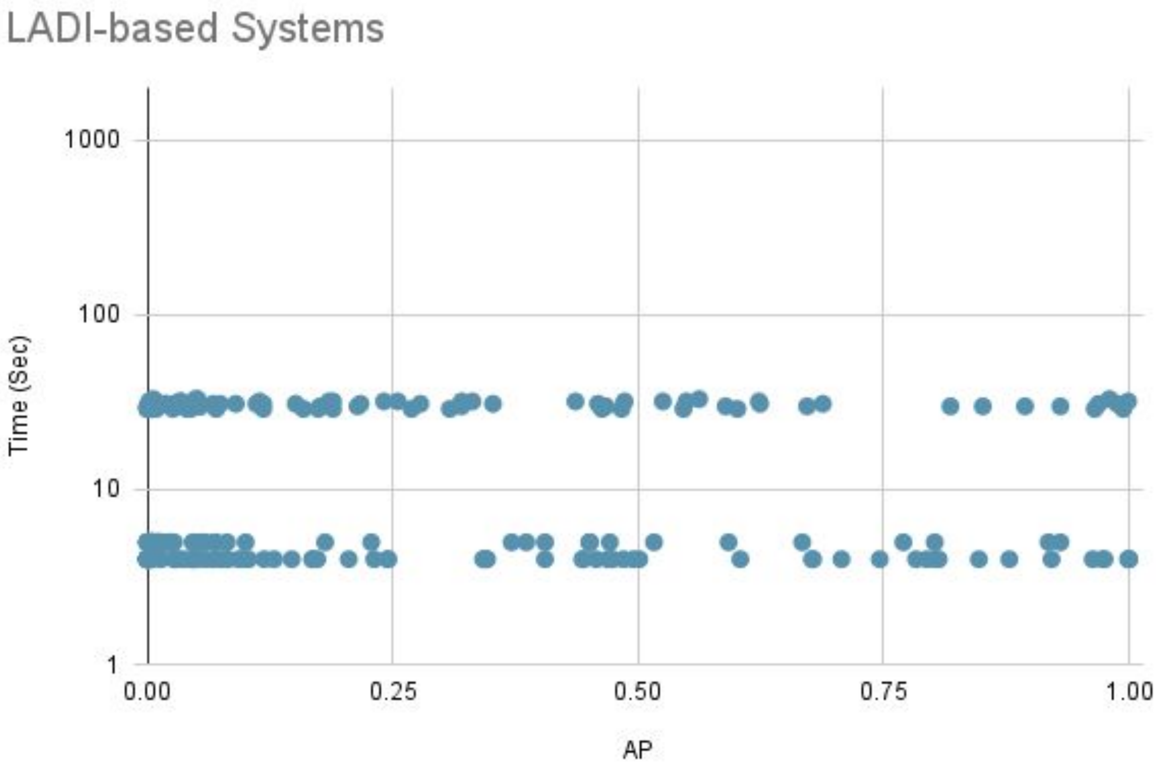


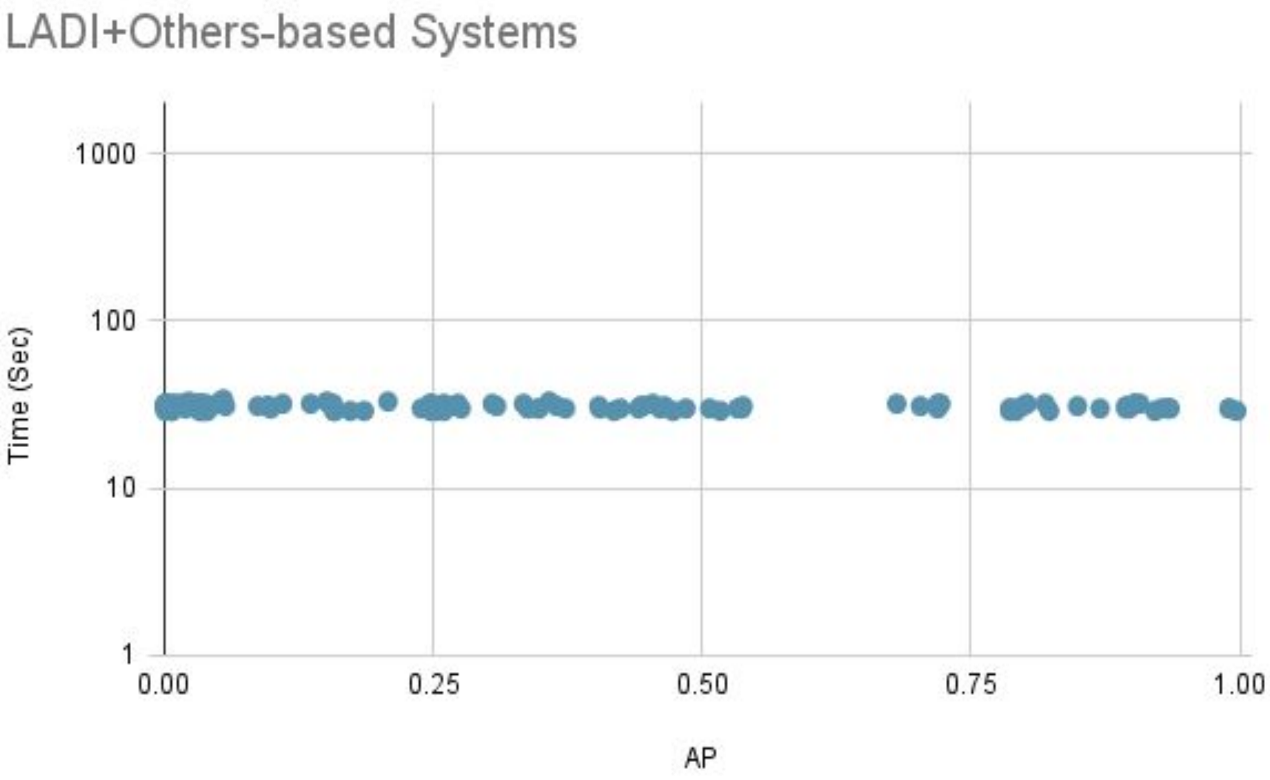
Results by Teams



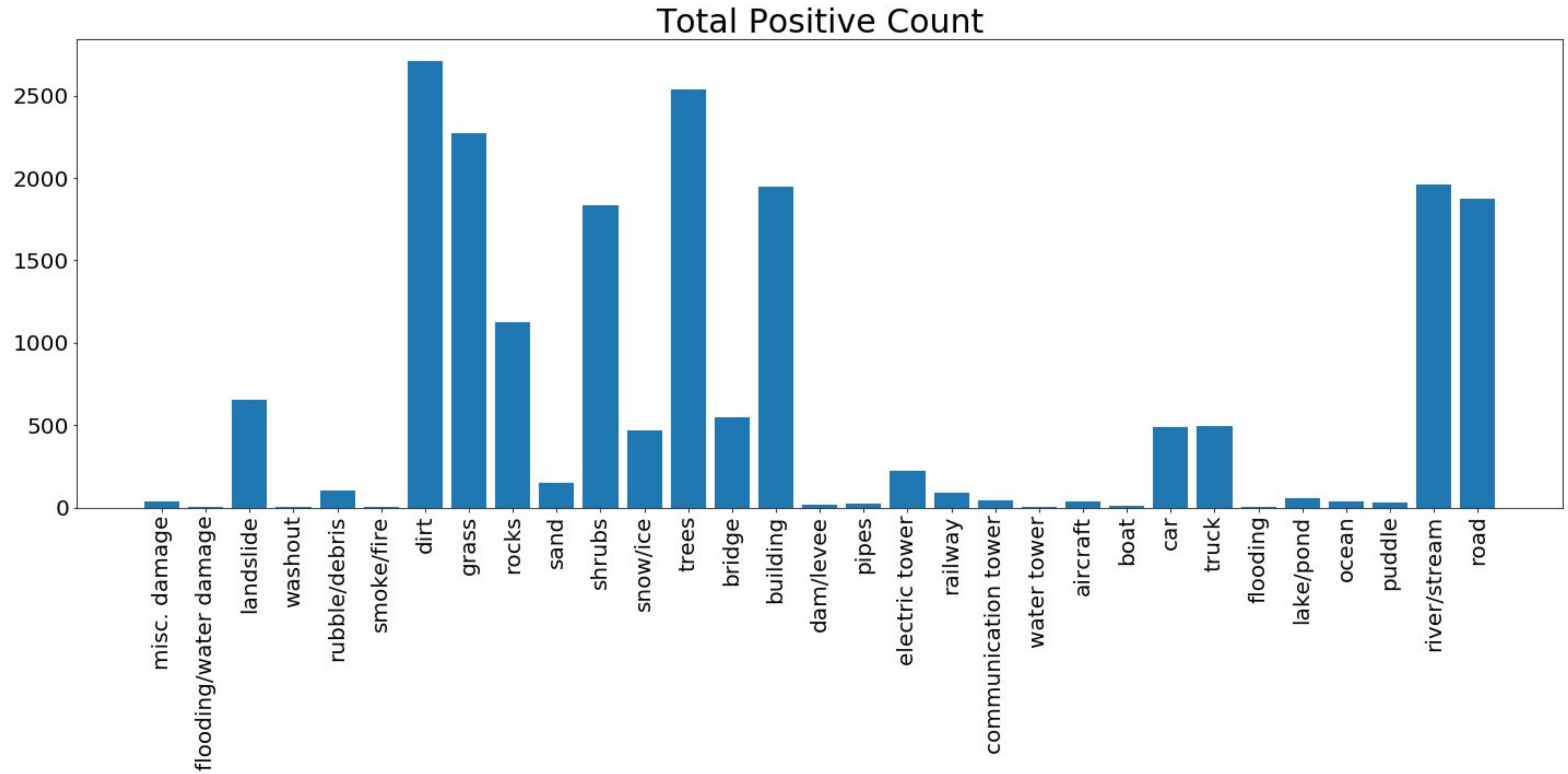


Efficiency


LADI-based Systems
1000
100
10
1
Time (Sec)
0.00
0.25
0.50
0.75
1.00
AP

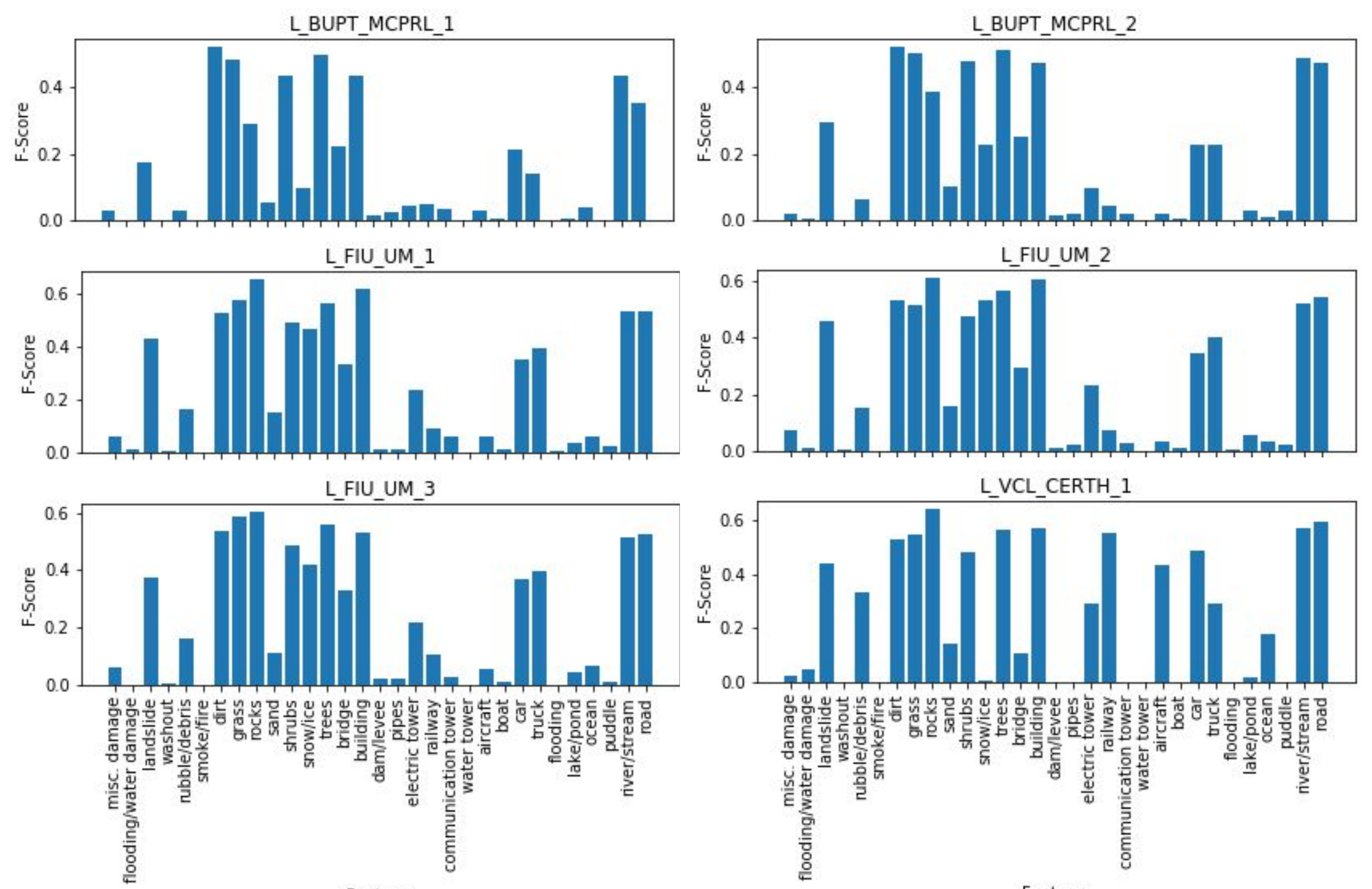

LADI+Others-based Systems
1000
100
10
1
Time (Sec)
0.00
0.25
0.50
0.75
1.00
AP

Total Positives for Features

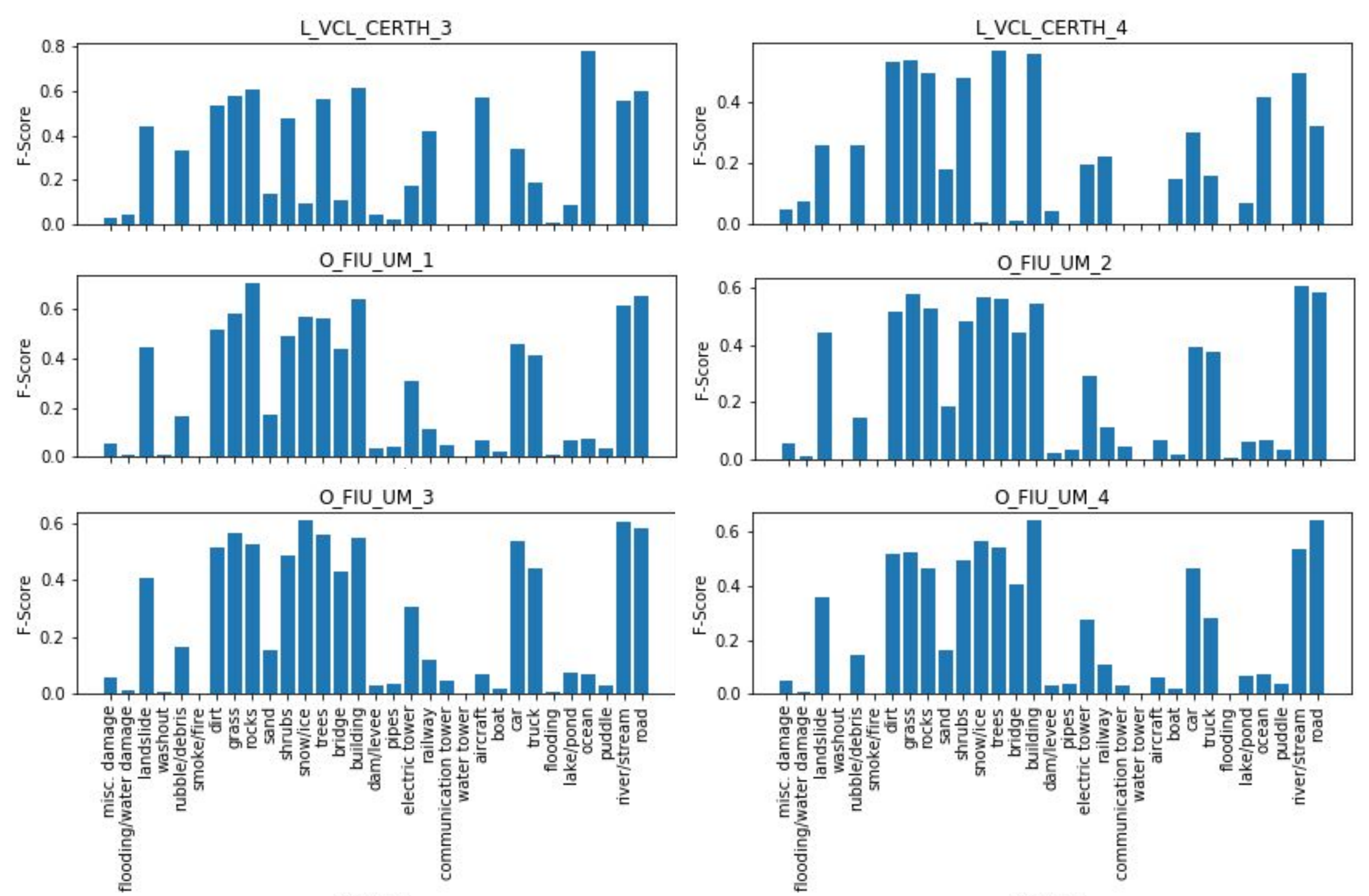


- Graph shows number of shots containing each feature.
- Some features (e.g. dirt, grass, trees, buildings, roads, etc.) occur much more frequently than others.
- Under-representation of some features and annotations in the training datasets (e.g. flooding, misc. damage) affected the distribution of detected features in the testing dataset

F-Measure For Submissions



F-Measure For Submissions



Conclusion and Future Work

- Three teams submitted to the task out of seven that signed up.
- Second year of the task continued with same disaster scenario as used in 2020 (The Nepal earthquake). Additional videos from other disasters added.
- Challenges include:
  - Small dataset and limited resources for annotation.
  - Training dataset distribution of features/annotations caused some bias.
  - Training and testing dataset should be from the same distribution. Hard to do with different nature of calamities.
- We plan to continue with the task for 2022. Accumulating training dataset by adding 2021 testing dataset.
- We are looking for a different test dataset, which focuses on other types of disasters.

Thank you!