Supplementary Material: Functional Map of the World

Gordon Christie[1] Neil Fendley[1] James Wilson[2] Ryan Mukherjee[1]
[1]The Johns Hopkins University Applied Physics Laboratory [2]DigitalGlobe
{gordon.christie,neil.fendley,ryan.mukherjee}@jhuapl.edu
james.wilson@digitalglobe.com

Abstract

In this document, we provide:

1: Full acknowledgments.
2: Descriptions and distributions of metadata features.
3: Additional collection details.
4: Additional results.
5: Examples from our dataset.

1. Acknowledgments

This work would not have been possible without the help of everyone on the fMoW Challenge team, who we thank for their contributions. A special thanks to: Kyle Ellis, Todd Bacastow, Alex Dunmire, and Derick Greyling from DigitalGlobe; Rebecca Allegar, Jillian Brennan, Dan Reitz, and Ian Snyder from Booz Allen Hamilton; Kyle Bowerman and Gődény Balázs from Topcoder; and, finally, Myron Brown, Philippe Burlina, Alfred Mayalu, and Nicolas Norena Acosta from JHU/APL. We also thank the professors, graduate students, and researchers in industry and from the CV community for their suggestions and participation in discussions that helped shape the direction of this work.

The material in this paper is based upon work supported by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via Contract 2017-17032700004. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation therein.

2. Metadata Features and Statistics

1. **ISO Country Code** ISO Alpha-3 country code (String). There are a total of 247 possible country codes, 207 of which are present in fMoW.
2. **UTM Zone** Universal Transverse Mercator. There are 60 UTM zones, which are 6° in width. We provide a number for the UTM zone (1-60), along with a letter representing the latitude band. There are a total of 20 latitude bands, which range from "C" to "X" ("I" and "O" are not included).
3. **Timestamp** UTC timestamp. Datetime format (Python): "%Y-%m-%dT%H:%M:%SZ" (String).
4. **Cloud Cover** Fraction of the image strip, not image chip, that is completely obscured by clouds on a scale of 0-100 (Integer).
5. **Scan Direction** The direction the sensor is pointed when collecting an image strip. Either "Forward", when the image is collected ahead of the orbital path or "Reverse" when the image is taken behind the orbital path (String).
6. **Pan Resolution** Ground sample distance of panchromatic band (pan-GSD) in the image strip, measured in meters (Double). `start`, `end`, `min`, and `max` values are also included. `start` and `end` represent the pan-GSD for the first and last scan lines, respectively. `min` and `max` represent the minimum and maximum pan-GSD for all scan lines, respectively.
7. **Multi Resolution** Ground sample distance of multispectral bands (multi-GSD) in the image strip, measured in meters (Double). `start`, `end`, `min`, and `max` values are also included. `start` and `end` represent the multi-GSD for the first and last scan lines, respectively. `min` and `max` represent the minimum and maximum multi-GSD for all scan lines, respectively.
8. **Target Azimuth** Azimuth angle of the sensor with respect to the center of the image strip, measured in degrees (Double). `start`, `end`, `min`, and `max` values are also included. `start` and `end` represent the target azimuth for the first and last scan lines, respectively. `min` and `max` represent the minimum and maximum target azimuth for all scan lines, respectively.
9. **Sun Azimuth** Azimuth angle of the sun measured from north, clockwise in degrees, to the center of the

image strip, measured in degrees (Double). `min` and `max` values are also included. `min` and `max` represent the minimum and maximum sun azimuth for all scan lines, respectively.

10. **Sun Elevation** Elevation angle of the sun measured from the horizontal, measured in degrees (Double). `min` and `max` values are also included. `min` and `max` represent the minimum and maximum sun elevation for all scan lines, respectively.
11. **Off-Nadir Angle** The off nadir angle of the satellite with respect to the center of the image strip, measured in degrees (Double). `start`, `end`, `min`, and `max` values are also included. `start` and `end` represent the off-nadir angle for the first and last scan lines, respectively. `min` and `max` represent the minimum and maximum off-nadir angle for all scan lines, respectively.

Country Codes Here we show the counts for each unique country code in fMoW. Counts are incremented once for each sequence instead of once per metadata file.

[("USA", 18750), ("FRA", 7470), ("ITA", 6985), ("RUS", 6913), ("CHN", 6597), ("DEU", 4686), ("GBR", 4496), ("BRA", 3820), ("CAN", 3128), ("TUR", 2837), ("JPN", 2542), ("IDN", 2448), ("ESP", 2402), ("AUS", 2105), ("DZA", 1849), ("IND", 1804), ("UKR", 1735), ("CZE", 1713), ("POL", 1386), ("MEX", 1274), ("ARG", 1248), ("NLD", 1236), ("SYR", 1224), ("BEL", 1190), ("PHL", 1179), ("IRQ", 1129), ("EGY", 1041), ("ZAF", 924), ("CHL", 888), ("LTU", 871), ("LBY", 863), ("KOR", 809), ("CHE", 788), ("LVA", 772), ("PRT", 722), ("YEM", 701), ("BLR", 601), ("GRC", 592), ("AUT", 572), ("SVN", 570), ("ARE", 566), ("IRN", 540), ("COL", 509), ("TWN", 509), ("TZA", 475), ("NZL", 465), ("PER", 459), ("HTI", 417), ("KEN", 405), ("NGA", 383), ("VEN", 378), ("PRK", 371), ("ECU", 351), ("IRL", 335), ("MYS", 328), ("BOL", 313), ("FIN", 288), ("KAZ", 268), ("MAR", 266), ("TUN", 257), ("CUB", 256), ("EST", 247), ("SAU", 246), ("HUN", 222), ("THA", 219), ("NPL", 196), ("HRV", 187), ("NOR", 183), ("SVK", 175), ("SEN", 172), ("BGD", 171), ("HND", 167), ("SWE", 166), ("BGR", 165), ("HKG", 154), ("DNK", 153), ("MDA", 147), ("ROU", 142), ("ZWE", 141), ("SRB", 140), ("GTM", 140), ("DOM", 134), ("LUX", 133), ("SDN", 132), ("VNM", 126), ("URY", 120), ("CRI", 119), ("SOM", 112), ("ISL", 110), ("LKA", 110), ("QAT", 108), ("PRY", 107), ("SGP", 106), ("OMN", 105), ("PRI", 95), ("NIC", 87), ("NER", 85), ("SSD", 82), ("UGA", 79), ("SLV", 79), ("JOR", 78), ("CMR", 77), ("PAN", 74), ("PAK", 72), ("UZB", 70), ("CYP", 67), ("KWT", 67), ("ALB", 66), ("CIV", 65), ("BHR", 65), ("GIN", 64), ("MLT", 63), ("JAM", 62), ("AZE", 62), ("GEO", 60), ("SLE", 59), ("ETH", 58), ("LBN", 57), ("ZMB", 55), ("TTO", 54), ("LBR", 52), ("BWA", 51), ("ANT", 50), ("BHS", 50), ("MNG", 46), ("MKD", 45), ("GLP", 45), ("COD", 45), ("KO-", 42), ("BEN", 42), ("GHA", 41), ("MDG", 36), ("MLI", 35), ("AFG", 35), ("ARM", 33), ("MRT", 33), ("KHM", 32), ("CPV", 31), ("TKM", 31), ("MMR", 31), ("BFA", 29), ("BLZ", 29), ("NCL", 28), ("AGO", 27), ("FJI", 26), ("TCD", 25), ("MTQ", 25), ("GMB", 23), ("SWZ", 23), ("BIH", 21), ("CAF", 19), ("GUF", 19), ("PSE", 19), ("MOZ", 18), ("NAM", 18), ("SUR", 17), ("GAB", 17), ("LSO", 16), ("ERI", 15), ("BRN", 14), ("REU", 14), ("GUY", 14), ("MAC", 13), ("TON", 13), ("ABW", 12), ("PYF", 12), ("TGO", 12), ("BRB", 12), ("VIR", 11), ("CA-", 11), ("DJI", 11), ("FLK", 11), ("MNE", 11), ("KGZ", 11), ("ESH", 10), ("LCA", 10), ("BMU", 10), ("COG", 9), ("ATG", 9), ("BDI", 9), ("GIB", 8), ("LAO", 8), ("GNB", 8), ("DMA", 8), ("KNA", 8), ("GNQ", 7), ("RWA", 7), ("BTN", 7), ("TJK", 6), ("TCA", 5), ("VCT", 4), ("WSM", 3), ("IOT", 3), ("AND", 3), ("ISR", 3), ("AIA", 3), ("MDV", 2), ("SHN", 2), ("VGB", 2), ("MSR", 2), ("PNG", 1), ("MHL", 1), ("VUT", 1), ("GRD", 1), ("VAT", 1), ("MCO", 1)]

UTM Zones Here we show the counts for each unique UTM zone in fMoW. Counts are incremented once for each sequence instead of once per metadata file.

[("31U", 5802), ("32T", 4524), ("33T", 4403), ("30U", 4186), ("32U", 3864), ("33U", 3315), ("31T", 3150), ("18T", 2672), ("17T", 2339), ("34U", 2049), ("37S", 1718), ("30T", 1686), ("37U", 1672), ("23K", 1627), ("18S", 1481), ("11S", 1388), ("16T", 1283), ("54S", 1244), ("38S", 1229), ("31S", 1227), ("35U", 1137), ("35V", 1116), ("52S", 1115), ("16S", 1110), ("51P", 1086), ("51R", 1069), ("36S", 1046), ("35T", 1038), ("36R", 1037), ("49M", 1026), ("48M", 1021), ("10T", 1010), ("53S", 1001), ("10S", 955), ("14R", 935), ("19T", 928), ("30S", 912), ("17S", 875), ("17R", 874), ("43P", 854), ("50S", 796), ("36U", 767), ("50R", 751), ("33S", 751), ("32S", 746), ("14S", 730), ("34T", 728), ("12S", 716), ("37M", 705), ("13S", 676), ("37T", 667), ("36T", 653), ("15S", 629), ("55H", 618), ("34S", 604), ("29S", 600), ("38P", 598), ("15T", 586), ("22J", 585), ("18Q", 549), ("15R", 539), ("35S", 511), ("10U", 497), ("21H", 492), ("36V", 491), ("19H", 482), ("48R", 476), ("49S", 459), ("48S", 446), ("49Q", 444), ("29T", 438), ("16P", 429), ("56H", 425), ("14Q", 422), ("40R", 420), ("39R", 413), ("39U", 406), ("18N", 385), ("35J", 383), ("37V", 380), ("50T", 379), ("56J", 355), ("34V", 351), ("43V", 347), ("29U", 346), ("38U", 345), ("17M", 328), ("38T", 323), ("19P", 323), ("51S", 317), ("54H", 311), ("49R", 295), ("34H", 293), ("22K", 293), ("48N", 276), ("20H", 273), ("50Q", 268), ("28P", 262), ("18L", 260), ("24M", 258), ("24L", 256), ("21J", 255), ("41V", 254), ("13T", 254), ("47N", 253), ("40U", 253), ("45R", 251), ("43Q", 245), ("51Q", 243), ("51T", 240), ("39S", 239), ("19K", 238), ("19Q", 237), ("59G", 236), ("43R", 234), ("12T", 230), ("49T", 227), ("41U", 223), ("32V", 219), ("30V",

212), (“13Q”, 212), (“40V”, 210), (“16R”, 210), (“20T”, 210), (“38R”, 204), (“36J”, 203), (“46T”, 200), (“45T”, 197), (“44U”, 196), (“15Q”, 190), (“50L”, 190), (“32P”, 184), (“60H”, 182), (“47P”, 182), (“20P”, 181), (“24K”, 178), (“17Q”, 178), (“35K”, 169), (“20J”, 168), (“11U”, 165), (“18H”, 164), (“52T”, 163), (“11T”, 161), (“36N”, 158), (“39V”, 157), (“20K”, 157), (“39Q”, 155), (“12U”, 149), (“38V”, 147), (“18P”, 147), (“23L”, 147), (“18G”, 146), (“31N”, 146), (“19J”, 142), (“33P”, 141), (“40Q”, 136), (“13R”, 136), (“47T”, 132), (“47R”, 126), (“48U”, 124), (“32R”, 123), (“15P”, 121), (“39P”, 117), (“48P”, 117), (“33R”, 116), (“45U”, 113), (“43S”, 111), (“44N”, 109), (“54T”, 109), (“32N”, 109), (“36W”, 108), (“17P”, 108), (“36P”, 105), (“31R”, 104), (“56K”, 101), (“20Q”, 101), (“39T”, 97), (“16Q”, 96), (“29R”, 95), (“25L”, 92), (“45Q”, 91), (“46Q”, 91), (“48T”, 90), (“44Q”, 89), (“42V”, 87), (“29N”, 87), (“43U”, 86), (“4Q”, 86), (“47Q”, 85), (“48Q”, 84), (“30N”, 83), (“19G”, 82), (“25M”, 81), (“42Q”, 80), (“44P”, 80), (“20L”, 77), (“50J”, 77), (“53U”, 76), (“38N”, 75), (“27W”, 75), (“44R”, 75), (“33V”, 74), (“34R”, 72), (“49L”, 70), (“36M”, 69), (“40S”, 69), (“12R”, 68), (“37P”, 68), (“52R”, 65), (“14T”, 64), (“50U”, 62), (“35H”, 62), (“50H”, 61), (“28R”, 60), (“54U”, 59), (“46V”, 58), (“44T”, 56), (“21K”, 56), (“55G”, 56), (“22L”, 56), (“35P”, 55), (“31P”, 54), (“29P”, 54), (“35R”, 52), (“30R”, 51), (“19U”, 50), (“53T”, 49), (“46U”, 49), (“50N”, 48), (“47S”, 48), (“42R”, 48), (“37Q”, 47), (“19L”, 47), (“14U”, 47), (“28Q”, 46), (“37N”, 45), (“19F”, 45), (“42U”, 44), (“36K”, 42), (“37R”, 40), (“37W”, 40), (“41S”, 38), (“42S”, 38), (“38Q”, 37), (“30P”, 37), (“42T”, 36), (“35L”, 36), (“46R”, 36), (“52U”, 35), (“60G”, 35), (“27V”, 34), (“45V”, 34), (“35W”, 34), (“13U”, 34), (“35M”, 34), (“18M”, 32), (“17L”, 32), (“41W”, 32), (“17N”, 31), (“21N”, 31), (“23M”, 30), (“21L”, 29), (“28S”, 28), (“58K”, 28), (“22M”, 28), (“41R”, 27), (“18R”, 27), (“10V”, 26), (“57U”, 26), (“34K”, 26), (“49U”, 25), (“6V”, 25), (“38L”, 25), (“20G”, 25), (“33L”, 24), (“60K”, 24), (“55K”, 23), (“51N”, 23), (“22H”, 22), (“22N”, 22), (“47V”, 22), (“41T”, 21), (“44V”, 21), (“36Q”, 21), (“46S”, 20), (“22T”, 20), (“34N”, 19), (“20U”, 19), (“12Q”, 19), (“12V”, 19), (“19N”, 18), (“31Q”, 18), (“21M”, 18), (“52L”, 18), (“56V”, 18), (“52V”, 18), (“23J”, 16), (“45W”, 16), (“9U”, 16), (“34J”, 16), (“27P”, 16), (“43W”, 15), (“1K”, 14), (“33M”, 14), (“40W”, 14), (“40K”, 14), (“43T”, 14), (“55T”, 14), (“51U”, 13), (“53K”, 13), (“34M”, 13), (“32M”, 13), (“37L”, 13), (“21P”, 12), (“50P”, 12), (“35N”, 12), (“6K”, 11), (“59H”, 11), (“33K”, 11), (“20M”, 11), (“49N”, 11), (“5Q”, 10), (“6W”, 10), (“26Q”, 10), (“39L”, 10), (“47U”, 10), (“34W”, 10), (“50K”, 10), (“8V”, 10), (“20S”, 10), (“40T”, 9), (“51V”, 9), (“42W”, 8), (“60W”, 8), (“53H”, 8), (“50V”, 8), (“20F”, 8), (“53L”, 7), (“18F”, 7), (“35Q”, 7), (“30Q”, 7), (“44S”, 7), (“15M”, 7), (“5V”, 7), (“54J”, 7), (“39W”, 6), (“49P”, 6), (“50M”, 6), (“19V”, 6), (“21F”, 6), (“20N”, 5), (“14P”, 5), (“34P”, 5), (“53J”, 5), (“38M”, 5), (“51K”, 5), (“29Q”, 4), (“11R”, 4), (“49V”, 4), (“48V”, 4), (“51M”, 4), (“38W”, 4), (“33N”, 4), (“45S”, 4), (“27Q”, 4), (“55J”, 3), (“19M”, 3), (“53V”, 3), (“2W”, 3), (“32Q”, 3), (“2L”, 3), (“16M”, 3), (“57W”, 3), (“43M”, 3), (“53W”, 2), (“43N”, 2), (“52J”, 2), (“28M”, 2), (“56T”, 2), (“33H”, 2), (“21T”, 2), (“44W”, 2), (“15V”, 1), (“33W”, 1), (“60V”, 1), (“18K”, 1), (“31M”, 1), (“54M”, 1), (“58P”, 1), (“58W”, 1), (“40X”, 1), (“58G”, 1), (“57V”, 1), (“16U”, 1), (“59K”, 1), (“52N”, 1), (“2K”, 1), (“33Q”, 1), (“34Q”, 1), (“11V”, 1), (“56W”, 1), (“26P”, 1), (“28W”, 1), (“59W”, 1), (“38K”, 1), (“26S”, 1), (“7L”, 1), (“56U”, 1), (“55V”, 1)]

3. Dataset Collection

The location selection phase was used to identify potential locations that map to our categories while also ensuring geographic diversity. Potential locations were drawn from several Volunteered Geographic Information (VGI) datasets, which were conflated and curated to remove duplicates and ensure geographic diversity. The remaining locations were then processed using DigitalGlobe’s GeoHIVE crowdsourcing platform. Members of the GeoHIVE crowd were asked to validate the presence of categories in satellite images, as shown in Figure 1. The interface uses centerpoint location information to draw a circle around a possible object of interest. The interface then asks users to rapidly verify the existence of a particular label, as extracted from the VGI datasets, using the ‘1’, ‘2’, and ‘3’ keys to represent existence, non-existence, and cloud cover.

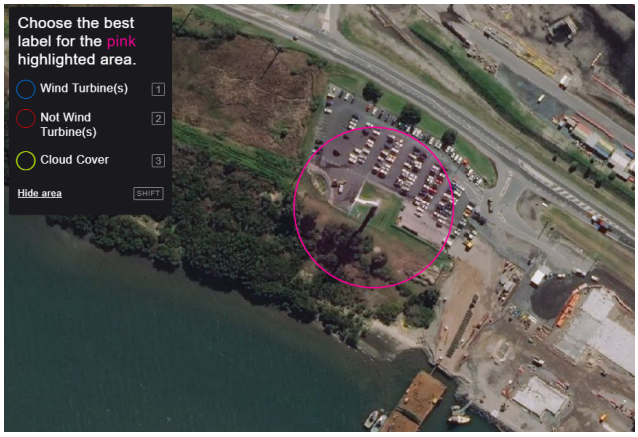


Figure 1: Sample image (“wind farm”) of what a GeoHIVE user might see while validating potential fMoW features. Instructions can be seen in the top-left corner that inform users to press the ‘1’, ‘2’, or ‘3’ keys to validate existence, non-existence, or cloud obscuration of a particular object.

For validation of object localization, a different interface is used that asks users to draw a bounding box around the

object of interest after being given an initial seed point. The visualization for this is shown in Figure 2, and the seed point can be seen as a green dot located on the object of interest. Users are additionally provided some instructions regarding how large of a box to draw, which may vary by object class. This interface is more complex than the location selection interface, which is why it is performed after object existence can be confirmed and non-cloudy high-quality imagery is obtained. A smaller and more experienced group of users is also used for this task to help ensure the quality of the annotations.

To help illustrate why full image annotation of fMoW categories is difficult, we show an example from the dataset in Figure 3. The primary category, which is located near the center of the image, is "gas station". As shown, it is difficult to identify the functional purpose of the surrounding buildings, and if map data is not available, it would be easy for humans to make mistakes when annotating. It is also possible to see how object detectors may correctly detect other categories from fMoW. By providing bounding boxes as input, we can avoid the issue of scoring results for which annotations do not exist.

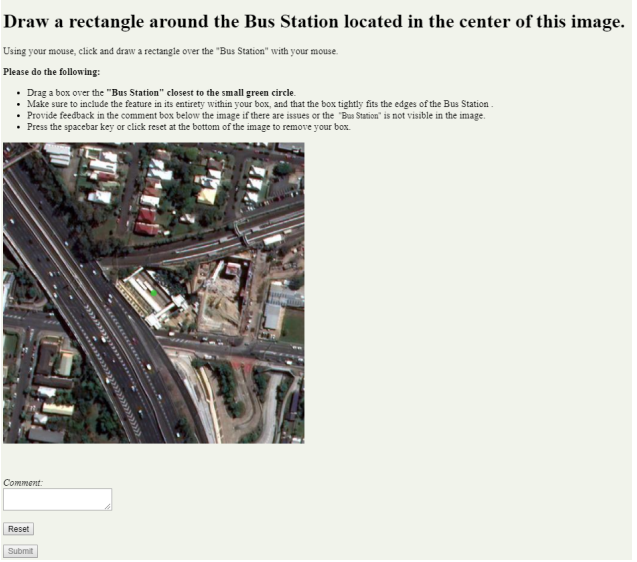
Another difficult example is shown in Figure 4, which shows an instance of the "educational institution" category in Japan. While an initial box is provided to the annotators, it is difficult to determine which buildings should be grouped as part of the same category when placing a bounding box.

4. Additional Results

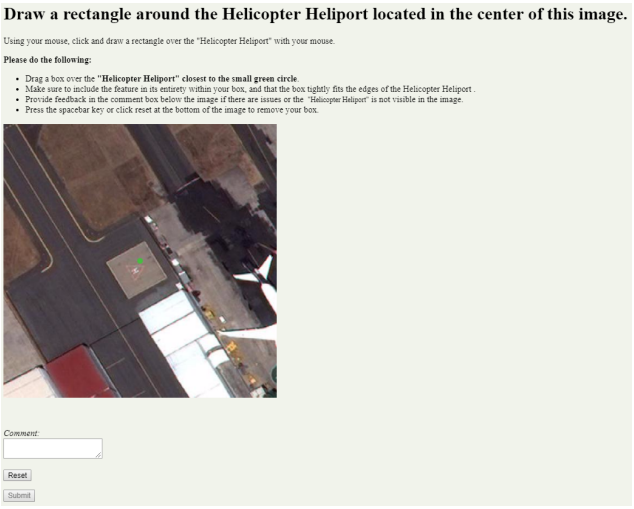
Introduced in the main paper, CNN-I-1 and CNN-IM-1 make predictions for each individual view. All other methods repeat their prediction over the full sequence. Again, we note that these tests are clearly not fair to some categories, such as "construction site", where some views may not even contain the category. However, we show results of these tests for completeness. Only the average values, which do not include "false detection" results, are shown in the main paper. We show per-category results in Table 1.

5. Dataset Examples

Figure 5 shows one example for each category in our dataset. For viewing purposes, regions within the full image chip were extracted using the scaled bounding box coordinates for the categories. For the baseline approaches presented in the main paper, smaller boxes were given more context than larger boxes. Therefore, for smaller-sized categories (*e.g.*, smoke stacks) it may appear that there is a lot more context than expected. It is important to keep in mind that the images for each category in the full dataset vary in quality, recognizability, were taken under different weather conditions (*e.g.*, snow cover) and seasons, contain drastically different context (*e.g.*, desert vs. urban), and other variations.



(a) ground transportation station



(b) helipad

Figure 2: Sample images of the interface used to more precisely localize objects within an image. In each example, a green dot is placed near the center of the pertinent object. Users are able to draw a bounding box by clicking and dragging. Instructions at the top of each example inform the user how to use the interface and also provide any category-specific instructions that may be relevant. Comments regarding issues such as clouds or object misclassification can be entered near the bottom of the page before submitting an annotation.

Figure 3: This image shows an example from fMoW with an instance of the "gas station" category, which is located near the center of the image. This shows how it is difficult to identify the functional purpose of the surrounding buildings. If map data is not available, it is very easy for humans to make mistakes when annotating.

Figure 4: This image shows an example from fMoW with an instance of the "educational institution" category. This example is located in Japan. This shows the difficulty of determining which buildings/areas should be included within the bounding box as part of the category.

	CNN-I-1	CNN-I	LSTM-I	CNN-IM-1	CNN-IM	LSTM-IM
false_detection	0.669	0.737	0.732	0.834	0.840	0.821
airport	0.782	0.864	0.819	0.900	0.905	0.835
airport hangar	0.642	0.746	0.685	0.659	0.696	0.726
airport terminal	0.642	0.726	0.757	0.655	0.758	0.782
amusement park	0.701	0.751	0.736	0.854	0.901	0.846
aquaculture	0.606	0.743	0.767	0.720	0.798	0.790
archaeological site	0.451	0.532	0.507	0.518	0.624	0.622
barn	0.606	0.678	0.675	0.642	0.697	0.682
border checkpoint	0.241	0.268	0.311	0.307	0.465	0.497
burial site	0.732	0.788	0.802	0.781	0.821	0.830
car dealership	0.654	0.712	0.771	0.672	0.716	0.748
construction site	0.298	0.436	0.423	0.306	0.347	0.407
crop field	0.856	0.879	0.871	0.908	0.933	0.929
dam	0.776	0.805	0.778	0.776	0.839	0.861
debris or rubble	0.263	0.330	0.536	0.228	0.365	0.439
educational institution	0.477	0.517	0.482	0.517	0.585	0.601
electric substation	0.761	0.852	0.865	0.806	0.847	0.859
factory or powerplant	0.410	0.461	0.461	0.496	0.534	0.542
fire station	0.405	0.382	0.450	0.351	0.471	0.516
flooded road	0.214	0.254	0.240	0.585	0.634	0.809
fountain	0.661	0.744	0.720	0.729	0.811	0.857
gas station	0.717	0.779	0.806	0.703	0.767	0.785
golf course	0.866	0.906	0.926	0.901	0.932	0.898
ground transportation station	0.653	0.691	0.733	0.661	0.734	0.764
helipad	0.700	0.814	0.866	0.734	0.834	0.804
hospital	0.319	0.385	0.395	0.349	0.447	0.468
impoverished settlement	0.396	0.484	0.546	0.763	0.764	0.691
interchange	0.758	0.852	0.691	0.898	0.912	0.927
lake or pond	0.549	0.700	0.625	0.616	0.661	0.676
lighthouse	0.645	0.727	0.751	0.761	0.805	0.854
military facility	0.517	0.564	0.627	0.588	0.630	0.685
multi-unit residential	0.401	0.433	0.472	0.406	0.517	0.523
nuclear powerplant	0.548	0.575	0.759	0.598	0.650	0.494
office building	0.213	0.229	0.245	0.195	0.225	0.213
oil or gas facility	0.721	0.757	0.767	0.763	0.824	0.859
park	0.561	0.624	0.653	0.629	0.658	0.685
parking lot or garage	0.731	0.778	0.791	0.665	0.694	0.704
place of worship	0.576	0.637	0.642	0.640	0.703	0.729
police station	0.219	0.216	0.225	0.243	0.199	0.317
port	0.632	0.646	0.621	0.680	0.710	0.642
prison	0.530	0.614	0.657	0.577	0.656	0.729
race track	0.832	0.893	0.880	0.889	0.936	0.924
railway bridge	0.667	0.704	0.759	0.708	0.762	0.794
recreational facility	0.871	0.908	0.925	0.865	0.911	0.909
road bridge	0.659	0.712	0.728	0.728	0.742	0.758
runway	0.767	0.847	0.806	0.857	0.899	0.900
shipyard	0.405	0.416	0.326	0.398	0.390	0.411
shopping mall	0.554	0.617	0.622	0.625	0.676	0.675
single-unit residential	0.633	0.700	0.705	0.649	0.711	0.658
smokestack	0.700	0.756	0.762	0.684	0.792	0.782
solar farm	0.791	0.862	0.884	0.832	0.852	0.882
space facility	0.878	0.878	0.788	0.917	0.885	0.971
stadium	0.853	0.866	0.903	0.844	0.871	0.879
storage tank	0.913	0.933	0.920	0.874	0.930	0.921
surface mine	0.757	0.789	0.754	0.795	0.837	0.848
swimming pool	0.878	0.916	0.903	0.865	0.894	0.881
toll booth	0.840	0.874	0.878	0.904	0.949	0.947
tower	0.644	0.741	0.765	0.667	0.749	0.777
tunnel opening	0.783	0.852	0.880	0.911	0.943	0.932
waste disposal	0.531	0.562	0.516	0.470	0.583	0.632
water treatment facility	0.782	0.842	0.786	0.757	0.841	0.864
wind farm	0.881	0.932	0.934	0.931	0.950	0.972
zoo	0.523	0.531	0.563	0.502	0.606	0.637
Average	0.622	0.678	0.684	0.669	0.722	**0.735**

Table 1: F1 scores for different approaches on an individual image basis. Color formatting was applied to each column independently. The average values shown at the bottom of the table are calculated without the false detection scores. CNN-I-1 and CNN-IM-1 make predictions for each individual view. All other methods repeat their prediction over the full sequence.

Figure 5: One example per category in fMoW.