SpaceNet 8 - The Detection of Flooded Roads and Buildings

Ronny Hänsch
German Aerospace Center (DLR)
ronny.haensch@dlr.de

Jacob Arndt, Dalton Lunga
Oak Ridge National Laboratory
{arndtjw, lungadd}@ornl.gov

Matthew Gibb, Tyler Pedelose, Arnold Boedihardjo, Desiree Petrie, Todd M. Bacastow
Maxar Technologies
firstname.lastname@maxar.com

Abstract

The frequency and intensity of natural disasters (i.e. wildfires, storms, floods) has increased over recent decades. Extreme weather can often be linked to climate change, and human population expansion and urbanization have led to a growing risk. In particular floods due to large amounts of rainfall are of rising severity and are causing loss of life, destruction of buildings and infrastructure, erosion of arable land, and environmental hazards around the world. Expanding urbanization along rivers and creeks often includes opening flood plains for building construction and river straightening and dredging speeding up the flow of water. In a flood event, rapid response is essential which requires knowledge which buildings are susceptible to flooding and which roads are still accessible. To this aim, SpaceNet 8 is the first remote sensing machine learning training dataset combining building footprint detection, road network extraction, and flood detection covering 850km^2, including 32k buildings and 1,300km roads of which 13% and 15% are flooded, respectively.

1. Introduction

Floods are one of the major types of natural disasters responsible for loss of life, damage to buildings and infrastructure, immense costs for destroyed property, cleanup and rebuilding, healthcare, the erosion of arable land, and a host of other environmental problems (e.g. the 2016 and 2017 floods in Houston, Texas, USA, carried human waste into reefs more than 100km away [23]). A total of 2.23 million km^2 were flooded and 255-290 million people directly affected by overflowing rivers in the last 15 years alone [25]. The frequency and severity, and thus the damage caused by floods are likely to increase given the ever growing human population, the urbanization of flood plains, and the influence of climate change [10]. Flood detection and moni-

toring also connects to the Sustainable Development Goals defined by the United Nations, e.g. to assess the vulnerability to flooding, mapping degraded land, and disaster risk reduction [15].

Being able to send first aid quickly during, or shortly after, a flood event can mean saving lives. A fast response requires an assessment of the damage, i.e. most importantly to buildings and infrastructure, since this information is essential in determining where help is needed and which roads are blocked by flood waters. Satellite imagery analysis is one of the tools that is able to provide a fast initial mapping of the flooded region [5, 14] (sometimes combined with other information sources e.g. from social networks [21]) and to help monitoring the situation after the disaster [16, 19]. Remote sensing is also used to map areas with high flood risk [3, 24], to forecast floods [2, 13] and their evolution [12], and assess post-flood effects [1].

Several public benchmark datasets exist addressing flood detection and related topics. The *Flood Extent Detection* dataset [7] published in 2021 by NASA Impact provides more than 30k 256×256 dual-polarimetric Sentinel 1 image patches. The goal is to detect open areas of flood water. The *FloodNet* dataset [17] provides more than 2k optical images acquired by an unmanned aircraft system platform after Hurricane Harvey of which 20% are annotated. This dataset also contains buildings and road annotations. However, it is from a single location and event only with a small amount of labeled data. The *Sen12-FLOOD* dataset [18] is based on Sentinel 1&2 images and consists of 412 time series with on average 9 optical and 14 SAR images per sequence that cover a flood event in $\sim 45\%$ of the cases. The image resolution is with $10 \times 10m$ rather low and the flood label is provided only on the image level. Natural disasters are broadly covered by the *xBD dataset* [8] with a particular focus on building damage assessment from optical satellite imagery. The complete dataset contains more than 850k building annotations in 16 countries for several
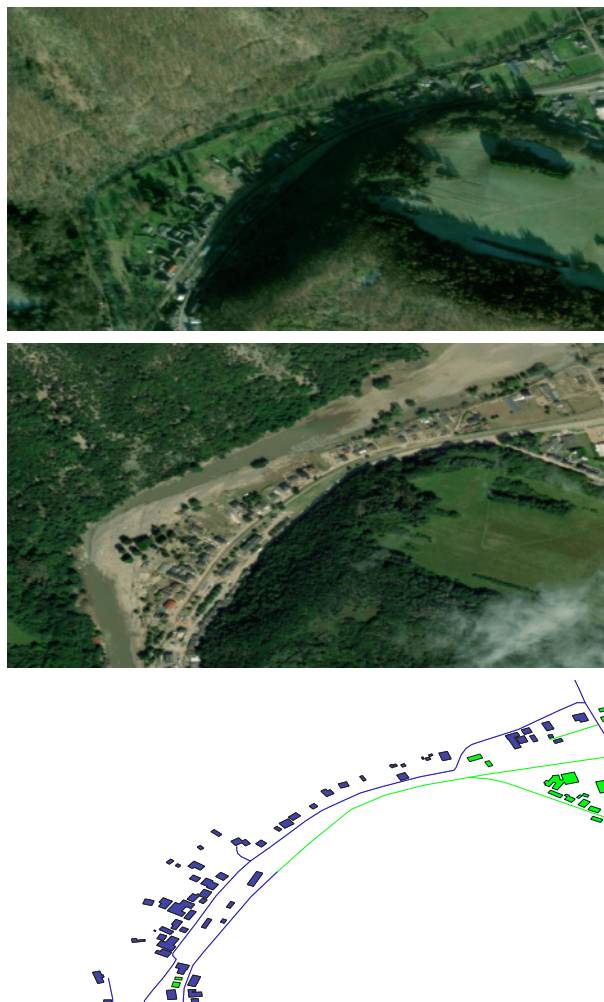
Figure 1. Example Maxar satellite imagery from the AOI in Germany of SpaceNet 8. Note the significant changes between the pre- (top) and post-event (middle) images due to different look angles, shadows, and of course the flood event itself. The bottom row shows part of the provided reference data, i.e. buildings and roads (marked in blue if "flooded"). Buildings are marked as "flooded" if there is flood water in close spatial proximity; roads if they are visually covered in flood water or rubble. Satellite imagery © 2022 Maxar Technologies.

disaster types, but flood data comes from two areas only (Midwestern US and Nepal) and contains 57k buildings of which 20% are damaged.

With the requirements of fast response in flood scenarios and the limitations of existing flood datasets in mind, SpaceNet[1] introduces the *SpaceNet 8 Challenge* that provides building and road annotations based on pre-event high-resolution optical imagery from Maxar with corresponding flood attributes from post-event imagery. Three different AOIs cover 850km^2 in total, including ∼32k build-

---

[1]https://spacenet.ai/

ings and ∼1,300km roads of which ∼13% and ∼15% are flooded, respectively.

SpaceNet is a collaborative initiative of companies, government agencies, research institutions, and professional societies that aims to foster the development of open source computer vision and machine learning methods for remote sensing data. Launched in August 2016 by IQT Labs' CosmiQ Works and Maxar as a non-profit project, it offers a repository of freely available overhead imagery with co-registered annotations. Today, SpaceNet consists of Maxar, Amazon Web Services (AWS), IEEE GRSS, and Topcoder with collaboration from Oak Ridge National Laboratory and releases openly licensed, precision-labeled satellite imagery, sponsors prize challenges to solve difficult problems, and releases the winning models.

In the last six years, SpaceNet has published eight challenges focused on building detection (SpaceNet 1, 2, 4, 6, 7) and road network extraction (SpaceNet 3, 5 [26, 28]) using very high-resolution optical data including non-standard characteristics such as off-nadir acquisitions (SpaceNet 4 [29]), and also exploring Synthetic Aperture Radar (SAR) in SpaceNet 6 [22]) and multi-temporal data (SpaceNet 7 [6, 27]).

For SpaceNet 8, we aimed for establishing a clear connection to previous challenges as well as going beyond what currently exists. The former allows building on the previous accomplishments, i.e. leveraging the massive amount of labeled data as well as the open source tools and top performing models of previous challenges published by SpaceNet. The latter is achieved by including a realistic level of complexity so that challenge outcomes are better aligned with the needs associated with the United Nations' Sustainable Development Goals.

As previous SpaceNet challenges, SpaceNet 8 addresses automated mapping of building footprints and road network with travel time estimation. A first novel aspect is that both tasks have to be solved simultaneously instead of independently by introducing a multi-class segmentation problem. In addition, the main novelty is that SpaceNet 8 does not only address foundation mapping but combines it with flood detection. The overall task is detecting buildings and roads from pre-event satellite imagery while determining for each object instance whether it is affected by a recent flood event in post-event imagery (see Fig. 1).

2. Data

For many major disaster events, the Maxar Open Data Program has published and continues to release high resolution satellite imagery collected from Maxar's Earth observation satellites. The imagery is released under a Creative Commons 4.0 Non-Commercial License, allowing for rapid access for humanitarian and response organizations. Over the past five years, there have been numerous events related
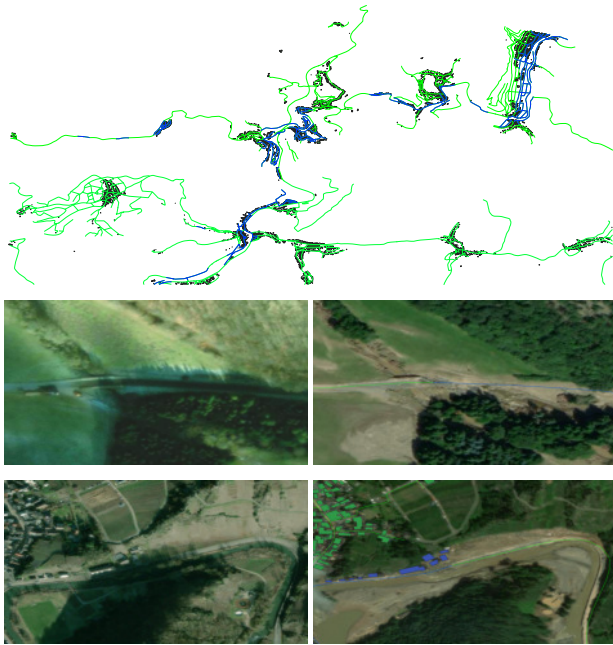
Figure 2. This AOI contains images acquired near Bonn, Germany, before and after a flood in 2021. The top shows the annotations: Not-flooded (flooded) roads and buildings are marked in green (blue). The middle and bottom rows show a road partially destroyed by the flood event (reference data overlaid with the post event image on the right). Satellite imagery © 2022 Maxar Technologies.
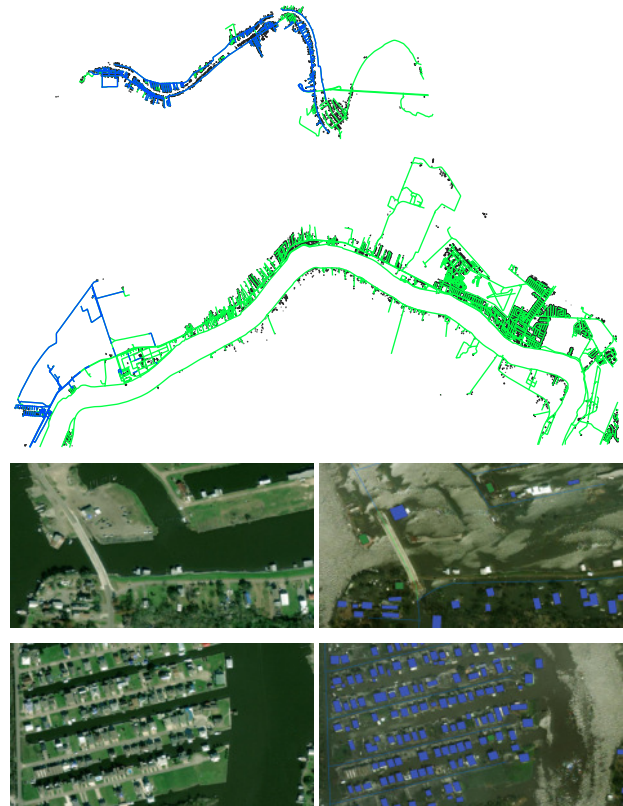

Figure 3. This AOI contains Maxar satellite imagery acquired over Louisiana, USA, before and after a flood in 2021. The top shows a part of the annotations: Not-flooded (flooded) roads and buildings are marked in green (blue). The middle and bottom rows show flooded building areas (reference data overlaid with the post event image on the right). Satellite imagery © 2022 Maxar Technologies.

to flooding. The Maxar Open Data Program, releases both pre- and post-event imagery, when available, enabling the ability to compare and analyze the situation on the ground.

For the SpaceNet 8 Challenge, three distinct areas of interest (AOIs) were selected: Germany after heavy rains in early July 2021, Louisiana after Hurricane Ida end of August 2021, and a third location to be used for blind testing that will be revealed after the challenge is concluded. A total of twelve images are used for training data collection and scoring, consisting of both pre- and post-flooding event imagery. While the Maxar satellites do provide panchromatic and multispectral imagery, we choose to use the pansharpened RGB (0.3-0.8m resolution) data routinely published through the Open Data Program since this is what would be available during such an event to make the task more realistic.

Using the Maxar Open Data Program imagery, data annotations are created by hand, considering both the pre- and post-event images. Starting with the pre-event imagery, analysts labeled all building footprints and highway centerlines. Highways are attributed with a classification of the road type (e.g. residential), the surface type (paved or unpaved), and the number of lanes. After the pre-event baseline dataset was created, analysts then used the post-event

images to attribute buildings and segments of roads that appeared flooded in the imagery. A road segment is considered flooded if it is visibly covered with water or rubble in the post-event imagery. If a building is assigned the flooded attribute, it means that there is flood water in its immediate proximity which is a proxy for probable flooding (since it is not possible from satellite imagery to determine whether a building is actually flooded inside). In the event that cloud cover obscured features in the post-event image, as is often the case in flooding disasters, analysts were instructed to only attribute what was visible, and not make any assumptions where features could not be seen.

Table 1 shows the statistics of the created reference data. Figure 2 and 3 show the annotations and two example patches from the AOIs in Germany and Louisiana. Note the significant differences between the pre- and post-event images such as look angle and light conditions making automatic analysis extremely challenging. While the Louisiana AOI contains a densely populated area along a large river

in mostly flat terrain, the German AOI shows small villages and towns along a small river in hilly terrain. This diversity in both, properties of the geographic location and image acquisition factors, is what any image analysis tool would need to be able to handle in a realistic scenario.

The data is tiled by creating a regular geographic grid for the pre-event image with non-overlapping, 1300×1300 pixel patches. The same grid is then used to tile the reference data as well as the post-event images. Since pre- and post-event images have different resolutions, the post-event patches cover the same area on the ground but have different pixel sizes. We leave it to the participants to address these differences in resolution, e.g. by up-/downsampling the smaller/larger patch.

Image data as tiled GeoTIFFs, reference annotations as GeoJSON, as well as an example baseline implementation (see Section 3) are freely available via the SpaceNet webpage.

AOI	L	G	M
Size (km^2)	741.1	65.5	43.4
Buildings	18,892	3,723	9,000
Flooded Buildings	2,736	1,175	344
Percent Flooded Buildings (%)	14	32	4
Highways (km)	885	163.6	287.7
Flooded Highways (km)	148.4	34.6	21.2
Percent Flooded Highways (%)	17	21	7

Table 1. Data statistics of the different AOIs of SpaceNet 8, i.e. Germany (G), Louisiana (L), and a third mystery AOI (M) that will be undisclosed after the contest.

3. Baseline

The building labels are prepared by converting the Geo-JSON annotations into single channel binary masks (0 for no-building and 1 for building). Following [26], we prepare multi-class segmentation labels for road speed by aggregating the speed labels into bins of 10 miles per hour (mph). The total range of values for speed range from 0 mph to 70 mph. We include an eighth channel to the road speed labels which is a binary road segmentation channel. Flood labels are created as four channel masks. Channel 1 corresponds to non-flooded buildings, channel 2 for flooded buildings, channel 3 for non-flooded roads, and channel 4 for flooded roads.

The proposed baseline consists of two independently trained convolutional neural networks and post-processing steps to convert rasterized predictions into vector data suitable for submission. One convolutional neural network, which we call the foundation network, is responsible for segmenting buildings and roads from the pre-event imagery without any flood attribution and for producing road speed estimates. The second network is responsible for predicting the flood status of roads and buildings by using both pre- and post-event imagery as input. Post-processing steps are included to convert predictions from the foundation network to vector data. Following this conversion, a final post-processing step merges the flood predictions into the vector representation of the foundation estimates, resulting in a final dataset suitable for submission to the evaluation server of the challenge. Figure 4 gives an overview of these three primary components of the proposed baseline method for SpaceNet 8.

3.1. Foundation Features Network

Building on previous SpaceNet baselines and proposed solutions, we use a U-Net [20] style segmentation model with ResNet34 encoder [9] as the foundation network. This is the same architecture that was used in [26] for city-scale road segmentation. We find that this architecture is suitable for segmenting foundation features (both road speed and buildings) from the pre-event imagery provided by the SpaceNet 8 dataset. In order to produce both road speed predictions and building predictions, we make one small modification to the architecture and include two convolutional output layers. One is responsible for predicting the eight classes of road speed (channels 1-7 for speed estimation, channel 8 for binary road prediction), while the other is responsible for binary segmentation of buildings.

We train the foundation network following a similar strategy outlined in [26] in terms of loss function and weights for each term. Our objective function consists of a focal loss and soft-Dice loss for road speed segmentation and binary cross-entropy loss for building segmentation. We weight the building and road losses equally with $\alpha_b = 0.5$. Within the road loss, we set $\alpha_{m_c} = 0.75$. The custom loss function is as follows:

$$\mathcal{L} = \alpha_b \mathcal{L}_{BCE} + (1 - \alpha_b)(\alpha_{m_c} \mathcal{L}_{focal} + (1 - \alpha_{m_c}) \mathcal{L}_{Dice}) \tag{1}$$

3.2. Flood Attribution Network

We use a Siamese convolutional neural network [11] to generate flood predictions. Several existing works show that a Siamese convolutional neural network is a suitable method for detecting change in high resolution satellite imagery [4, 30]. This network consists of two identical branches with shared weights and enables feature extraction from pre- and post-event image pairs. We use the same type of architecture that we use for the foundation features network. As a Siamese network, one branch receives the pre-event image as input, while the other branch receives the post-event image as input. Each branch's output features are concatenated, followed by two more convolutional layers with the last layer producing the flood prediction mask.
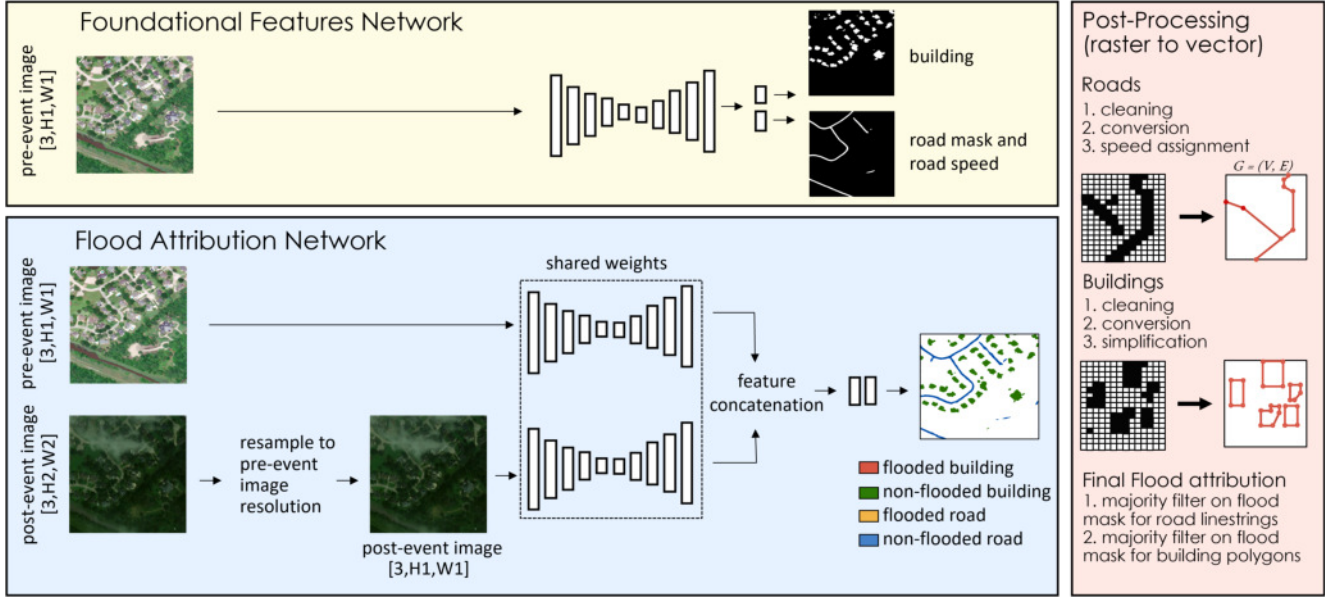
Figure 4. Schematic of the baseline method. The foundation network is a U-Net style segmentation network with a ResNet34 encoder. The flood attribution network consists of the same U-Net style segmentation network and ResNet34 encoder, but as siamese networks whereby a pre-event and post-event image are fed through the network with output features concatenated before the final convolutional layers.

3.3. Implementation Details

To train the foundation features network, we use an Adam optimizer, an initial learning rate of $1e - 4$, a batch size of 2, and train for 50 epochs. We decay the learning rate by a factor of 0.5 every 20 epochs.

We train the Siamese network using cross entropy loss along with similar implementation as the foundation features. We again use an Adam optimizer, an initial learning rate of $1e - 4$, a batch size of two, and train for 50 epochs. We decay the learning rate by a factor of 0.5 every 20 epochs.

We train both networks using a single NVIDIA Tesla V100 32GB GPU on a NVIDIA DGX-2 system.

3.4. Post-Processing

Model predictions are post-processed to produce the final flood estimate suitable for submission to the evaluation server. Road predictions from the foundation network are post-processed to produce a road network using the procedure outlined in [26] for graph extraction and speed estimation. Building predictions from the foundation features are post-processed by first applying morphological opening with a structuring element of size 5. The cleaned mask is then polygonized, followed by topology preserving polygon simplification with a distance tolerance of 0.75 meters. A final filtering step is used to remove polygons with an area smaller than 5 meters.

We generate the final flood prediction submission by assigning a flood or non-flood attribute to each vector feature.

The flood or non-flood assignment for each vector feature is determined by a majority vote using all pixels from the flood prediction mask that intersect the vector feature. If no flood prediction values are present in the flood prediction mask, but a foundation feature exists, we set its flood status to non-flooded.

4. Evaluation

We designed an evaluation approach that accounts for the real-world parameters of emergent events. Specifically, this entails a scoring scheme that enables different weighting for various model prediction failure vectors, for example, failing to detect a building should receive a larger penalty than failing to correctly attribute a true building detection. Other scenarios that motivate this scoring design include but are not limited to marginalizing trivial solutions, unbalanced classes, and reducing error propagation (e.g., false negative results for buildings should not be counted towards class label correctness).

Leveraging our prior SpaceNet challenges, we utilize two key base metrics to evaluate the quality of submitted solutions. For building flood assessments, the Intersection over Union (IoU) is used to quantify the spatial intersection of the predictions and the reference data. For road network prediction, the Average Path Length Similarity (APLS) [28] is used as the base score to measure how close the estimated road network is to the reference data with respect to travel time or distance between pairs of nodes.

For the building damage assessment, there are three

states associated with a building object: flooded building, non-flooded building, and no building. In order to more fully account for the general quality of the submitted predictions, we examine each of the possible building states against all other states with respect to their IoU scores. This is modeled as a 3×3 matrix where the columns (prediction) and rows (reference data) are the three states described above i.e., flooded building, non-flooded building, and no building, and each cell is the average IoU. For the "no building" state, we consider a fixed area that is the union of the test spatial regions to allow standard comparison across the set of submitted solutions. Based on the above matrix, we develop various aggregate functions to compute an overall quality score which is used towards comparing the different solutions.

For the road network estimates, we measure the average pairwise travel time/distance discrepancy (APLS) to the reference data. We model the flooded segments as non-traversable paths. Similar to the building damage assessment above, we consider path matching for two states: networks composed solely of traversable segments (i.e., flooded roads removed) and networks composed solely of non-traversable segments (i.e., non-flooded roads removed). This state-oriented comparison allows us to account for how well a model estimates matches against reference traversal paths concurrently with its prediction of non-traversable roads. This reduces ambiguities that may occur for instances where the APLS on the traversable network is high and the APLS on the non-traversable network is low.

A submitted solution is assigned a single score that is composed of the building damage and road network estimates quality scores. This final score can take on two forms: the first considers a normalized metric between the two evaluation scores above (e.g., mean of standardized building and network scores); and the second proposes to utilize the rank order of each evaluation component (building and road network) and compute a final rank based on an aggregation of the individual component ranks.

5. Experiments

We performed extensive preliminary experiments to decide for a train/test split best suited for the challenge. On the one hand, the problem should be challenging, i.e. we aimed for a setup that is not already perfectly solved by our simple baseline model. On the other hand, the challenge needed to be solvable avoiding that most if not all solutions fail to produce acceptable results.

We decided to split the available data on the AOI-level as this is the most realistic application scenario. A model trained over a given set of data needs to be able to generalize to new geographic locations and image acquisition factors if it is supposed to be applicable during a real flood event. We include the AOIs in Germany and Louisiana into the training set, while the third AOI is dedicated as a private test and not directly available to the participants during the challenge.

For this split, the baseline achieves an average APLS score of 0.12 and a mean IoU of 0.42 (averaged over flooded and not-flooded object instances). The performance for not-flooded roads and buildings is significantly better than for flooded objects.

Interestingly, if the third blind testing AOI is split into two halves where one part is included into the training set and the model is evaluated over the second part, performance improves to an APLS score of 0.45 for roads and a mean IoU of 0.66 for buildings. On the other hand, performance breaks down if the model is applied on the German AOI and trained on the other two regions. This clearly illustrates the challenges of real world scenarios where the model is applied to different geographic locations than it was trained on.

Figure 5 shows three example patches from the validation set of training the baseline (i.e. Louisiana). Buildings are mostly well recognized with the typical shortcomings of the U-Net architecture such as rounded corners. The road network itself is well extracted, too. However, road intersections are often missed, i.e. road segments end before the intersection and fail to connect to other road segments. While this would barely influence pixel-wise scores such as IoU, it causes the connectivity of the network to decrease which explains the low APLS scores of the baseline.

In realistic application scenarios it is not possible to wait for the next cloud-free image but one has to work with what is available. As described above, if an area is not visible in the post-event image due to cloud cover, the flood attribute is omitted. The last row of Figure 5 shows that this annotation rule has been learned by the network.

With the performance level of the baseline, it has only a limited utilization potential for real disaster scenarios. However, given this baseline method, similar to past SpaceNet Challenges, we set the stage for challenge participants to extend and improve upon our implementation. We are particularly interested in solutions which combine the flood attribution, road speed, and building feature learning into a single end-to-end framework, rather than separate networks. Such a solution would be a potential step towards removing the last post-processing step that merges flood attribution and foundation features.

6. Conclusion

Floods are one of the most common natural disasters. Their frequency and severity are only going to increase in the future due to climate change and urbanization of flood plains among other things. Being able to perform rapid damage assessment during a flood event is of utmost importance for first responders. In particular, we think the ability
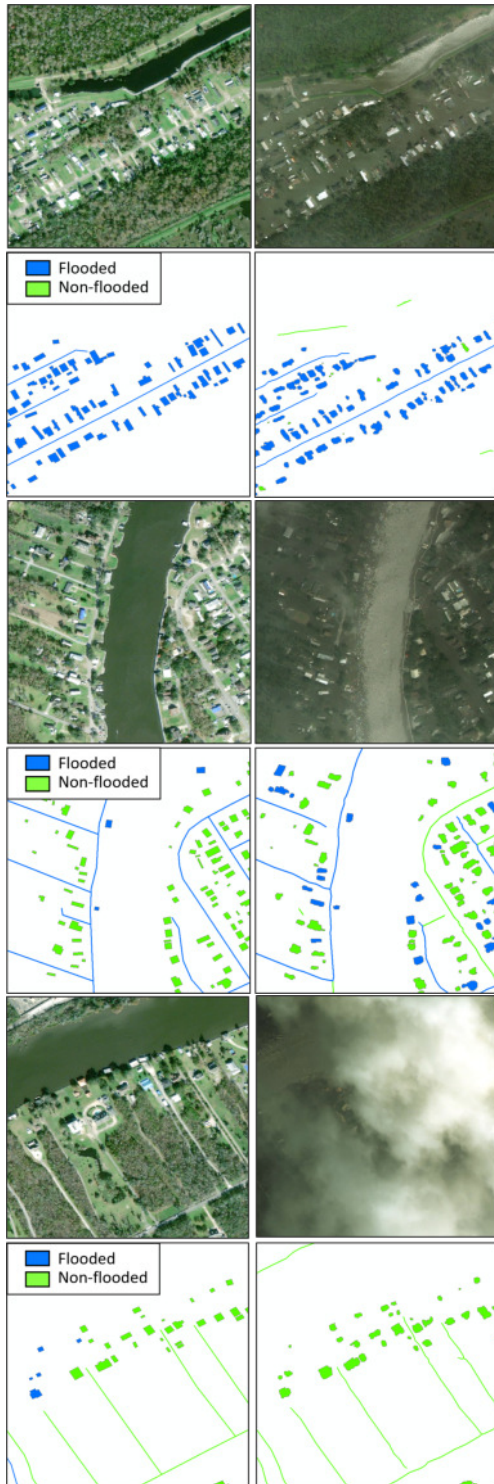
Figure 5. Three patches from the validation set (Louisiana): Pre- and post-event images, reference data, and prediction results for flooded roads and buildings using the baseline method. Satellite imagery © 2022 Maxar Technologies.

to rapidly identify which buildings are damaged and which roads are still traversable is valuable. Such information can be derived from remote sensing data, specifically, from images of optical satellites that cover wide areas in high resolution. Machine learning models such as deep neural networks allow to automatically analyze such data and guide human operators on the ground if properly trained on representative training data.

In this paper we introduce the SpaceNet 8 dataset which aims exactly at these needs. We build upon previous SpaceNet challenges and combine building detection and road network extraction to a multi-class segmentation problem and extend it further with flood detection. Thus, corresponding methods have to detect objects and road segments, assign every object instance with a flood attribute, and estimate the driving speed for each road.

We provide annotated high-resolution optical satellite imagery where the flood attribute indicates for a road segment that it is covered by water and for a building that there is flood water in its intermediate proximity. The imagery is provided through Maxar's Open Data Program, including realistic challenges such as cloud cover and differing image acquisition factors (e.g. light conditions) between pre- and post-event images. Covering 850km^2, 32k buildings, 1,300km roads distributed over three different AOIs and flood events, the SpaceNet 8 Challenge provides the currently largest and most diverse dataset for flood detection.

We believe that this dataset will foster the development of new approaches that are able to accurately and robustly detect flooded roads and buildings. These will be helpful guides to human operators on the ground during flood events and help to perform damage assessment, monitor the evolution of the flood event, and organize a response to provide help where needed in the fastest way possible.

Acknowledgements

References

[1] C.V. Angelino, L. Cicala, N. Fiscante, and M. Focareta. Estimated post-flood effects through free, full and open satellite remote sensing data to support civil protection. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 4215–4218, 2016. 1

[2] Nasreddine Belabid, Feng Zhao, Luca Brocca, Yanbo Huang, and Yumin Tan. Near-real-time flood forecasting

based on satellite precipitation products. *Remote Sensing*, 11(3), 2019. 1

[3] G. Robert Brakenridge. *Flood Risk Mapping From Orbital Remote Sensing*, chapter 3, pages 43–54. American Geophysical Union (AGU), 2018. 1

[4] R. Caye Daudt, B. Le Saux, and A. Boulch. Fully convolutional siamese networks for change detection. In *IEEE International Conference on Image Processing (ICIP)*, October 2018. 4

[5] J.-F. Cretaux, M. Berge-Nguyen, M. Leblanc, R. Abarca Del Rio, F. Delclaux, N. Mognard, C. Lion, R.-K. Pandey, S. Tweed, S. Calmant, and Philippe Maisongrande. Flood mapping inferred from remote sensing data. *International Water Technology Journal*, 1(1):48–62, 2011. 1

[6] A. Van Etten, D. Hogan, J. Manso, J. Shermeyer, N. Weir, and R. Lewis. The multi-temporal urban development spacenet dataset. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6394–6403, Los Alamitos, CA, USA, jun 2021. IEEE Computer Society. 2

[7] Shubhankar Gahlot, Muthukumaran Ramasubramanian, Iksha Gurung, Ronny Hansch, Andrew Molthan, and Manil Maskey. Curating flood extent data and leveraging citizen science for benchmarking machine learning solutions. *Earth and Space Science Open Archive*, page 9, 2022. 1

[8] Ritwik Gupta, Richard Hosfelt, Sandra Sajeev, Nirav Patel, Bryce Goodman, Jigar Doshi, Eric Heim, Howie Choset, and Matthew Gaston. xbd: A dataset for assessing building damage from satellite imagery, 2019. 1

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016. 4

[10] Abdullah Kahraman, Elizabeth J. Kendon, Steven C. Chan, and Hayley J. Fowler. Quasi-stationary intense rainstorms spread across europe under climate change. *Geophysical Research Letters*, 48(13):e2020GL092361, 2021. e2020GL092361 2020GL092361. 1

[11] Gregory Koch, Richard Zemel, and Ruslan Salakhutdinov. Siamese neural networks for one-shot image recognition. 2015. 4

[12] Roderick Lammers, Alan Li, Sreeja Nag, and Vinay Ravindra. Prediction models for urban flood evolution for satellite remote sensing. *Journal of Hydrology*, 603:127175, 2021. 1

[13] Hafiz Suliman Munawar, Ahmed W. A. Hammad, and S. Travis Waller. Remote sensing methods for flood prediction: A review. *Sensors*, 22(3), 2022. 1

[14] Davide Notti, Daniele Giordan, Fabiana Caló, Antonio Pepe, Francesco Zucca, and Jorge Pedro Galve. Potential and limitations of open satellite data for flood mapping. *Remote Sensing*, 10(11), 2018. 1

[15] Claudio Persello, Jan Dirk Wegner, Ronny Hänsch, Devis Tuia, Pedram Ghamisi, Mila Koeva, and Gustau Camps-Valls. Deep learning and earth observation to support the sustainable development goals, 2021. 1

[16] T.J. Pultz and R.A. Scofield. Applications of remotely sensed data in flood prediction and monitoring: report of the ceos disaster management support group flood team. In *IEEE International Geoscience and Remote Sensing Symposium*, volume 2, pages 768–770 vol.2, 2002. 1

[17] Maryam Rahnemoonfar, Tashnim Chowdhury, Argho Sarkar, Debvrat Varshney, Masoud Yari, and Robin Roberson Murphy. Floodnet: A high resolution aerial imagery dataset for post flood scene understanding. *IEEE Access*, 9:89644–89654, 2021. 1

[18] C. Rambour, N. Audebert, E. Koeniguer, B. Le Saux, M. Crucianu, and M. Datcu. Flood detection in time series of optical and sar images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2020:1343–1346, 2020. 1

[19] Alberto Refice, Annarita D'Addabbo, and Domenico Capolongo. *Methods, Techniques and Sensors for Precision Flood Monitoring Through Remote Sensing*, pages 1–25. Springer International Publishing, Cham, 2018. 1

[20] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing. 4

[21] Rizwan Sadiq, Zainab Akhtar, Muhammad Imran, and Ferda Ofli. Integrating remote sensing and social sensing for flood mapping. *Remote Sensing Applications: Society and Environment*, 25:100697, 2022. 1

[22] Jacob Shermeyer, Daniel Hogan, Jason Brown, Adam Van Etten, Nicholas Weir, Fabio Pacifici, Ronny Hänsch, Alexei Bastidas, Scott Soenen, Todd Bacastow, and Ryan Lewis. Spacenet 6: Multi-sensor all weather mapping dataset. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 768–777, 2020. 2

[23] Amanda Shore, Jordan A. Sims, Michael Grimes, Lauren I. Howe-Kerr, Carsten G. B. Grupstra, Shawn M. Doyle, Lauren Stadler, Jason B. Sylvan, Kathryn E. F. Shamberger, Sarah W. Davies, Lory Z. Santiago-Vázquez, and Adrienne M. S. Correa. On a reef far, far away: Anthropogenic impacts following extreme storms affect sponge health and bacterial communities. *Frontiers in Marine Science*, 8:305, 2021. 1

[24] Hannes Taubenböck, Michael Wurm, Maik Netzband, Hendrik Zwenzner, Achim Roth, A. Rahman, and Stefan W. Dech. Flood risks in urbanized areas – multi-sensoral approaches using remotely sensed data for risk assessment. *Natural Hazards and Earth System Sciences*, 11:431–444, 2011. 1

[25] B. Tellman, J. A. Sullivan, C. Kuhn, A. J. Kettner, C. S. Doyle, G. R. Brakenridge, T. A. Erickson, and D. A. Slayback. Satellite imaging reveals increased proportion of population exposed to floods. *Nature*, 596(7870):80–86, 2021. 1

[26] Adam Van Etten. City-scale road extraction from satellite imagery v2: Road speeds and travel times. In *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1775–1784, 2020. 2, 4, 5

[27] Adam Van Etten and Daniel Hogan. The spacenet multitemporal urban development challenge. In Hugo Jair Es-

calante and Katja Hofmann, editors, *Proceedings of the NeurIPS 2020 Competition and Demonstration Track*, volume 133 of *Proceedings of Machine Learning Research*, pages 216–232. PMLR, 06–12 Dec 2021. 2

[28] Adam Van Etten, Jacob Shermeyer, Daniel Hogan, Nicholas Weir, and Ryan Lewis. Road network and travel time extraction from multiple look angles with spacenet data. In *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, pages 3920–3923, 2020. 2, 5

[29] N. Weir, D. Lindenbaum, A. Bastidas, A. Etten, V. Kumar, S. Mcpherson, J. Shermeyer, and H. Tang. Spacenet mvoi: A multi-view overhead imagery dataset. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 992–1001, Los Alamitos, CA, USA, nov 2019. IEEE Computer Society. 2

[30] Yang Zhan, Kun Fu, Menglong Yan, Xian Sun, Hongqi Wang, and Xiaosong Qiu. Change detection based on deep siamese convolutional network for optical aerial images. *IEEE Geoscience and Remote Sensing Letters*, 14(10):1845–1849, 2017. 4