

2020 IEEE GRSS Data Fusion Contest: Global Land Cover Mapping With Weak Supervision

The 2020 Data Fusion Contest, organized by the IEEE Geoscience and Remote Sensing Society (GRSS) Image Analysis and Data Fusion Technical Committee (IADF TC) and the Technical University of Munich, aims to promote research in automatic large-scale land cover mapping from globally available multimodal satellite data with weak supervision (Figure 1). As in previous years [1]–[3], the contest will be set up as a benchmark competition. To promote practical innovation, the 2020 Data Fusion Contest will have two sequentially organized challenge tracks:

- ▶ *Track 1*: land cover classification with low-resolution labels
- ▶ *Track 2*: land cover classification with low- and high-resolution labels.

To train the land cover prediction models, we use the SEN12MS data set [4], which includes more than 180,000 triplets of corresponding *Sentinel-1* synthetic aperture radar data, *Sentinel-2* multispectral imagery, and moderate-resolution imaging spectroradiometer (MODIS)-derived low-resolution land cover maps [5] sampled across the globe (Figure 2). While all data are provided at a ground sampling distance of 10 m, the *Sentinel* images have a native resolution of approximately 10–20 m per pixel, whereas the MODIS-derived land cover has a native resolution of 500 m per pixel. For the contest, we use a simplified version of the international geosphere–biosphere program (IGBP) classification scheme [6], which has 10 land cover classes such as forest, shrubland, savanna, grassland, wetlands, crop-



land, urban/built-up areas, snow/ice, barren terrain, and water. The main challenge is to train classification models for high-resolution land cover prediction (with a target resolution of 10 m) from low-resolution and noisy annotations. The validation and testing employ semi-manually derived high-resolution land cover maps of scenes with undisclosed geolocations. It should be noted that the validation and testing samples are not contained in the SEN12MS data set.

In Track 1, the semi-manually derived high-resolution land cover maps for the validation set are undisclosed. The objective is to predict land cover labels using only low-resolution labels for training and hyperparameter tuning. In Track 2, we provide high-resolution labels for the validation set, and the goal is to train models for land cover mapping using low-resolution labels and a limited number of high-resolution labels. For both tracks, performance is assessed through the average accuracy of all classes.

The 2020 Data Fusion Contest focuses on a fundamental issue related to weakly supervised learning: global generalization, which is a major challenge in a range of fields, from Earth observation to computer vision and machine learning. In conventional land cover classification, it is common to use human annotations for training at the same resolution as the input image. However, this approach is expensive and difficult to apply on a global scale. The contest serves as a benchmark to evaluate the best approaches in global land cover classification with weak supervision. The large-scale training data set facilitates the development of deep-learning algorithms that are generalized worldwide, and the globally sampled test data set enables the generalization to be evaluated.

GET THE DATA, AND ENTER THE CONTEST

The SEN12MS training data set is available at <https://mediatum.ub.tum.de/1474000>. The validation data set without any corresponding high-resolution labels was posted to the IEEE DataPort (<https://iee-dataport.org/competitions/2020-ieee-grss-data-fusion-contest>) on 13 December 2019. The evaluation server with a public leaderboard opened on 13 January 2020, enabling participants to submit their validation-set prediction results to the CodaLab competition website (<https://competitions.codalab.org/competitions/22289>) for feedback about their approaches.

To enter the test phase, participants submitted a short description of their approach to Track 1 by 28 February, received the test data set on 1 March, and had five days to submit their land cover maps. On 6 March, the evaluation server for Track 1 closed, high-resolution labels for the validation data set were released, and the evaluation server for Track 2 opened. Participants have until 20 March to complete Track 2 and must submit a short description of their approach to Track 2 by 25 March to be eligible to win. The test phase is kept short to ensure an objective comparison of the methods. After the submitted classification maps are compared with the undisclosed ground truth for testing, the winners will be announced on 27 March, and the evaluation server will reopen to support development in the field.

More information regarding the data download and evaluation-server registration can be found at the IADF TC website (<http://www.grss-ieee.org/community/technical-committees/data-fusion/>). Questions and comments about the data and the contest can be submitted to the LinkedIn group of the IADF TC (<https://www.linkedin.com/groups/3678437/>).

AWARDEES, AWARDS, AND PRIZES

The following seven teams will be declared winners:

- ▶ the first-, second-, third-, and fourth-ranked teams in Track 1
- ▶ the first-, second-, and third-ranked teams in Track 2.

The winning teams will be invited to submit papers to the 2020 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) in Waikoloa, Hawaii. Each manuscript must describe the method used to achieve the winning results. The teams will present their papers during an oral invited session dedicated to the contest, and the papers will be included in the IGARSS 2020 proceedings.

The seven winning teams will be awarded an IEEE certificate of recognition during a ceremony at the IGARSS 2020 awards banquet. The first-ranked team in both tracks will be awarded a special prize and coauthor an article (with a limit of three coauthors per team) summarizing the outcomes of the contest, which will be submitted to *IEEE*



FIGURE 1. The 2020 Data Fusion Contest banner.

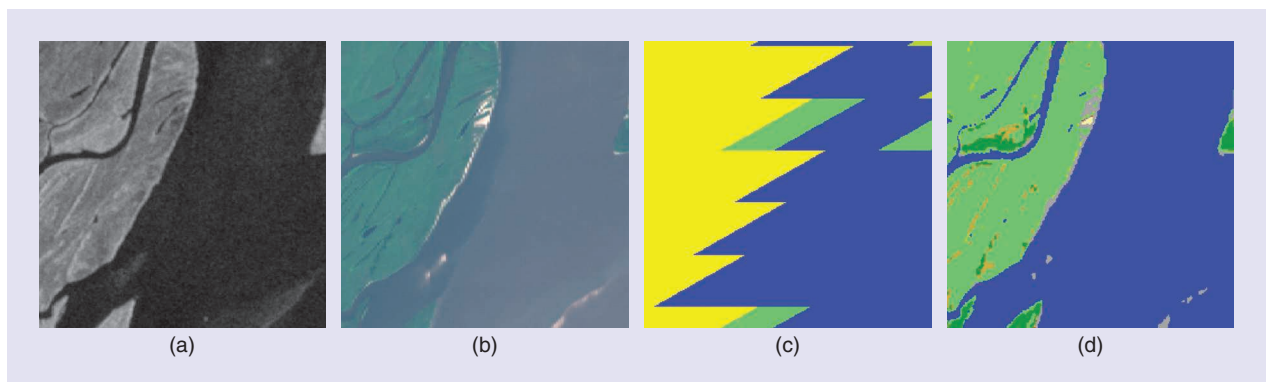


FIGURE 2. The SEN12MS training data, including the (a) *Sentinel-1* image (single polarization), (b) *Sentinel-2* image (red-green-blue), and (c) MODIS low-resolution land cover map (simplified IGBP scheme). (d) The Data Fusion Contest high-resolution land cover map that will be used for validation and testing.

Contest Data

The data of the 2020 GRSS Data Fusion Contest will remain available to the GRSS community for benchmarking algorithms and publishing research works. The data are usable free of charge for scientific purposes, but the contest terms and conditions on the contest website remain applicable (<http://www.grss-ieee.org/community/technical-committees/data-fusion/2020-ieee-grss-data-fusion-contest/>). Please read them carefully.

Join the GRSS IADF TC

You can contact the GRSS Image Analysis and Data Fusion Technical Committee (IADF TC) chairs at iadf_chairs@grss-ieee.org. If you are interested in joining the IADF TC, please complete the form on our website (<http://www.grss-ieee.org/community/technical-committees/data-fusion/>), or send an email to the chairs including your

- ▶ first and last name
- ▶ institution/company
- ▶ country
- ▶ IEEE membership number (if available)
- ▶ email address.

Members receive information regarding research and applications on image analysis and data fusion topics as well as updates concerning the annual Data Fusion Contest and all other IADF TC activities. Membership in the IADF TC is free! You may also join the LinkedIn GRSS data fusion discussion forum: <http://www.linkedin.com/groups/IEEE-GRSS-Data-Fusion-Discussion-3678437>.

Journal of Selected Topics in Applied Earth Observations and Remote Sensing. To maximize impact and promote research in global land cover mapping with weak supervision, the open access option will be used for these submissions. The GRSS will cover the costs related to the open access fees as well as the winning teams' participation in the awards banquet and will also sponsor the prizes for the winning teams.

ACKNOWLEDGMENTS

The contest is organized in collaboration with the Research Group for Signal Processing in Earth Observation (SiPEO) at the Technical University of Munich (TUM). The IADF TC chairs would like to thank TUM-SiPEO for providing the competition data and the GRSS for continuously supporting the Data Fusion Contest through funding and resources. The Copernicus Sentinel 2019 data are available from the European Space Agency (<https://sentinel.esa.int/>). The land cover maps derived from the moderate-resolution imaging spectroradiometer are available from the Land Processes Distributed Active Archive Center (<https://lpdaac.usgs.gov/>).

AUTHOR INFORMATION

Naoto Yokoya (naoto.yokoya@riken.jp) received his M.Eng. and Ph.D. degrees in aerospace engineering from the University of Tokyo in 2010 and 2013, respectively. He is the leader of the Geoinformatics Unit at the Riken Center for Advanced Intelligence Project, Tokyo. Previously, he was an assistant professor with the University of Tokyo and an Alexander von Humboldt research fellow with the German Aerospace Center, Oberpfaffenhofen, and the Technical

University of Munich, Germany. His research interests include the development of image processing, data fusion, and machine-learning algorithms for understanding remote sensing images, with applications to disaster management. He is currently chair of the IEEE Geoscience and Remote Sensing Society (GRSS) Image Analysis and Data Fusion Technical Committee and also secretary of the IEEE GRSS All Japan Joint Chapter. He is an associate editor of *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* and a Member of the IEEE.

Pedram Ghamisi (p.ghamisi@gmail.com) received his M.Sc. degree (Hons.) in remote sensing from the K.N. Toosi University of Technology, Tehran, Iran, in 2012 and his Ph.D. degree in electrical and computer engineering from the University of Iceland, Reykjavik, in 2015. In addition, in 2015, he was an Alexander von Humboldt research fellow with the Technical University of Munich, Germany, and Heidelberg University, Germany. From 2015 to 2018, he was a research scientist with the German Aerospace Center, Oberpfaffenhofen. Since 2018, he has been head of the machine-learning group, Helmholtz-Zentrum Dresden-Rossendorf. He is also chief technology officer and co-founder of VasoGnosis Inc., Milwaukee, Wisconsin, where he is involved in the development of advanced diagnostic and analysis tools for brain diseases using cloud computing and deep learning algorithms. He is also vice-chair of the IEEE Image Analysis and Data Fusion Technical Committee. His research interests include interdisciplinary research on remote sensing and machine (deep) learning, image and signal processing, and multisensor data fusion. He is a Senior Member of the IEEE and an associate editor of *MDPI-Remote Sensing* and *IEEE Geoscience and Remote Sensing Letters*.

Ronny Hänsch (rww.haensch@gmail.com) received his undergraduate degree in computer science and his Ph.D. degree from the Technische Universität Berlin, Germany, in 2007 and 2014, respectively. His research interests include computer vision, machine learning, object detection, neural networks, and ensemble theory. He works in the field of image-based object detection and classification with a focus on remote sensing data and, in particular, polarimetric synthetic aperture radar images. His current research interests focus on ensemble methods for image analysis. He is cochair of the IEEE Geoscience and Remote Sensing Society Image Analysis and Data Fusion Technical Committee and the International Society for Photogrammetry and Remote Sensing Working Group II/1 (Image Orientation). He is an associate editor of *IEEE Geoscience and Remote Sensing Letters*.

Michael Schmitt (m.schmitt@tum.de) received his Dipl.-Ing. degree in geodesy and geoinformation, his Dr.-Ing. degree in remote sensing, and his habilitation in data fusion from the Technical University of Munich (TUM), Germany, in 2009, 2014, and 2018, respectively. Since 2015, he has been a senior researcher and deputy head at the Professorship for Signal Processing in Earth Observation at TUM; in 2019 he was additionally appointed adjunct

teaching professor with the Department of Aerospace and Geodesy of TUM. In 2016, he was a guest scientist at the University of Massachusetts, Amherst. His research focuses on image analysis and machine learning applied to the extraction of information from remote sensing observations. In particular, he is interested in remote sensing data fusion with a focus on synthetic aperture radar and optical data. He is a cochair of the International Society for Photogrammetry and Remote Sensing Working Group 1/3 (SAR and Microwave Sensing). He frequently serves as a reviewer for a number of renowned international journals and conferences and has received several best reviewer awards. He is a Senior Member of the IEEE and an associate editor of *IEEE Geoscience and Remote Sensing Letters*.

REFERENCES

- [1] N. Yokoya et al., "Open data for global multimodal land use classification: Outcome of the 2017 IEEE GRSS Data Fusion Contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1363–1377, May 2018. doi: 10.1109/JSTARS.2018.2799698.
- [2] Y. Xu et al., "Advanced multi-sensor optical remote sensing for urban land use and land cover classification: Outcome of the 2018 IEEE GRSS Data Fusion Contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 6, pp. 1709–1724, June 2019. doi: 10.1109/JSTARS.2019.2911113.
- [3] B. L. Saux, N. Yokoya, R. Hansch, M. Brown, and G. Hager, "2019 Data Fusion Contest [Technical Committees]," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 1, pp. 103–105, Mar. 2019. doi: 10.1109/MGRS.2019.2893783.
- [4] M. Schmitt, L. H. Hughes, C. Qiu, and X. X. Zhu, "SEN12MS: A curated dataset of georeferenced multi-spectral sentinel-1/2 imagery for deep learning and data fusion," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. IV-2/W7, pp. 153–160, Sept. 2019. doi: 10.5194/isprs-annals-IV-2-W7-153-2019.
- [5] D. Sulla-Menashe and M. Friedl, "MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006," 2015. [Online]. Available: <https://lpdaac.usgs.gov/products/mcd12q1v006/>
- [6] T. R. Loveland and A. Belward, "The international geosphere biosphere programme data and information system global land cover data set (discover)," *Acta Astronaut.*, vol. 41, nos. 4–10, pp. 681–689, 1997. doi: 10.1016/S0094-5765(98)00050-2.

GRS

IEEE Membership Can Help You Reach Your Personal and Professional Goals



Gain access to the latest IEEE news, publications and digital library. Give and receive personal mentoring. Network locally and globally with IEEE members. And that's only the beginning. Discover how IEEE can help jumpstart your career.

"Participating in IEEE has developed me as a well-rounded engineer and helped me shine during networking events."

-Likhitha Patha

Electrical Engineering Student, IEEE Brand President, Virginia Polytechnic Institute and State University

Visit www.ieee.org/join today.

