Supplementary Information

Glacial Lake Observatory (GLO): Annual dataset of glacial lakes in Nepal and transboundary catchments (2017–2024)

Lauren D. Rawlins[1], C. Scott Watson[1], Rakesh Bhambri[2], Nitesh Khadka[3,4], Mohan B. Chand[5]

[1] School of Geography and water@leeds, University of Leeds, UK
[2] Wadia Institute of Himalayan Geology, Dehradun, Uttarakhand, India
[3] State Key Laboratory of Mountain Hazards and Engineering Resilience, Institute of Mountain Hazards and Environment, Chinese Academry of Sciences, Chengdu 610299, China
[4] University of Chinese Academy of Sciences, Beijing 101408, China
[5] Department of Environmental Science and Engineering, Himalayan Cryosphere, Climate and Disaster Research Centre (HiCCRDC), School of Science, Kathmandu University, Dhulikhel, Nepal

Correspondence to: C. Scott. Watson (C.S.Watson@leeds.ac.uk)

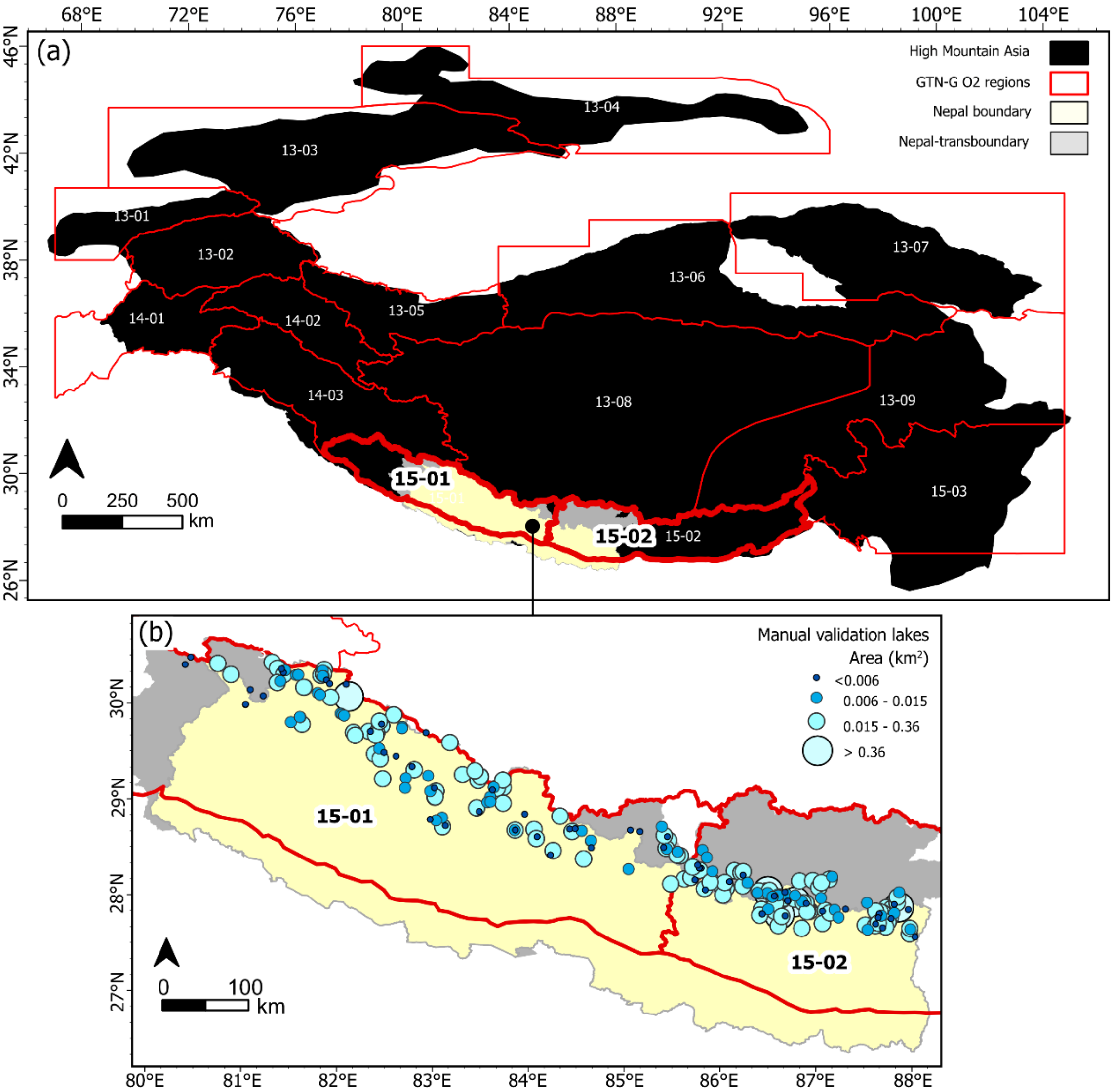

Figure S1: (a) The Global Terrestrial Network for Glaciers (GTN-G) O2-regions for High Mountain Asia (HMA; GTN-G, 2023). The Nepal-transboundary glacial lake dataset spans the O2-regions of 15-01 and 15-02 (Central Himalayas) with a 200 m overlap into 13-08. HMA outline from Bolch et al. (2019); (b) The coverage of the manual digitised dataset used for deep learning validation, showing a spatially variable and representative sample of lake sizes across the Nepal-transboundary region.

Table S1: Number of Sentinel-1 and Sentinel-2 images searched for and used in GEE for producing mosaicked yearly composites of the High Mountain Asia (HMA) region.

Year	Sentinel-1 images (HMA)	Sentinel-2 images (HMA)
2017	1418	3182
2018	1484	9089
2019	1404	6483
2020	1470	6147
2021	1441	18,761
2022	1292	17,913
2023	1232	18,617
2024	1158	18,308
Total images	**10,899**	**98,500**

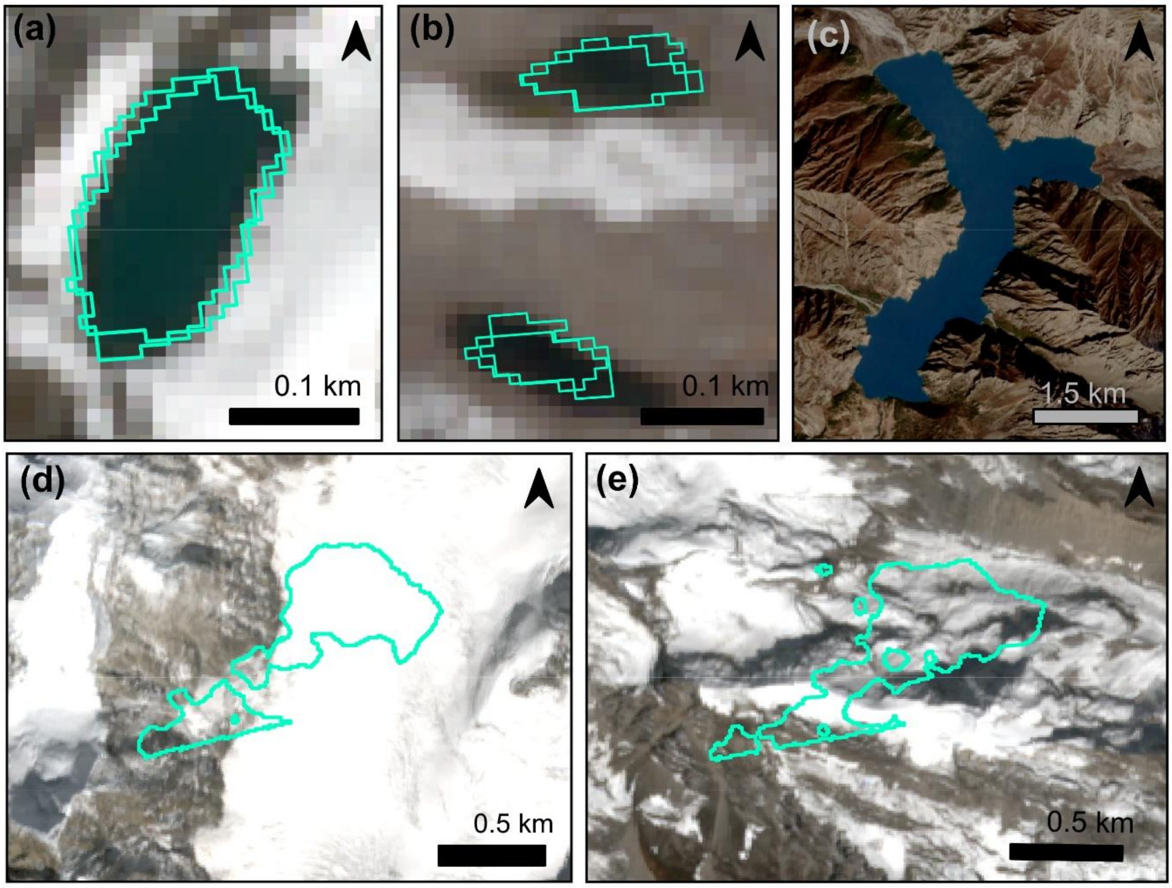

Figure S2: Examples of lake detections removed from the glacial lake dataset. Both (a) and (b) show glacial lakes which were detected twice. These were identified and removed using their duplicated GLO_IDs. (c) The Shey Phoksundo landslide lake which was removed from the dataset; (d) and (e) erroneous polygons captured on mountain terrain, removed using elevation standard deviation values ≥ 50 m.

Table S2: Validation statistics for comparing deep learning lakes from Sentinel-1 and Sentinel-2 data (from 2020 only) to the published glacial lake inventories of Zhang et al. (2024a) only and then supplemented with Kumar et al. (2025). The deep learning Sentinel-1 glacial lake dataset is also compared against the Sentinel-2 glacial lake dataset. F1 scores are given in bold.

Validation (Reference) Data	S1 Deep Learning (2020) vs Validation Data						
	TP (area)	**FP (area)**	**FN (area)**	**Prec.**	**Recall**	**F1**	**IoU**
TZhang	1381567	172219	430941	0.89	0.76	**0.82**	0.70
TZhang+Kumar	1445514	108200	653220	0.93	0.69	**0.79**	0.66
	S2 Deep Learning (2020) vs Validation Data						
TZhang	1551198	188459	261350	0.89	0.85	**0.87**	0.78
TZhang+Kumar	1636017	103422	462657	0.94	0.78	**0.85**	0.74
	S1 vs S2 (Unique Lakes)						
2020	1424323	132517	318083	0.92	0.82	**0.87**	0.76

Table S3: Validation statistics for comparing a 10% sample of deep learning lakes from Sentinel-1 and Sentinel-2 datasets to manually-digitised lake outlines produced independently by two authors. The F1 statistic is given in bold.

Year	Author (1) **vs** Author (2)						
	TP (area)	**FP (area)**	**FN (area)**	**Prec.**	**Recall**	**F1**	**IoU**
2020	9912200	365000	730100	0.96	0.93	**0.95**	0.90
	Author (1) vs S2 sample						
2017	8916400	938200	791500	0.90	0.92	**0.91**	0.84
2020	9465399	1120700	818600	0.89	0.92	**0.91**	0.83
2024	9865400	1059500	780400	0.90	0.93	**0.91**	0.84
	Author (2) **vs** S2 sample						
2020	9715800	869000	931900	0.92	0.91	**0.92**	0.84
	Author (1) vs S1 sample*						
2020	8147300	2129900	1437800	0.79	0.86	**0.82**	0.71
	S1 vs S2 *(Whole datasets)*						
2020	121569400.00	27952400.00	13959900.00	0.81	0.90	**0.85**	0.74

**** Sentinel-1 lakes are chosen based on their intersection with the manually digitised lakes from Sentinel-2.***

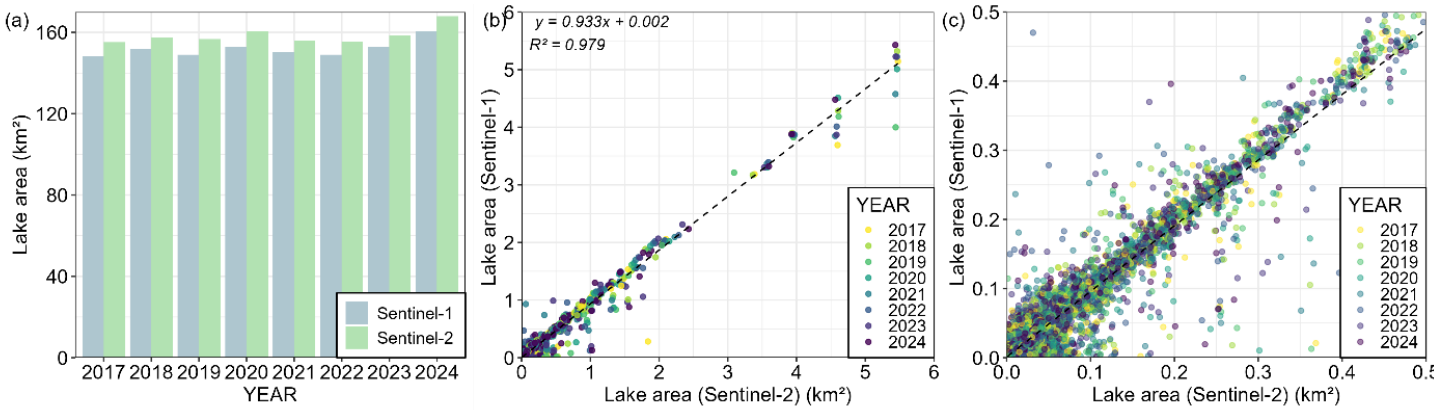

Figure S3: Comparison of Sentinel-1 and Sentinel-2 derived lake areas where a lake outline was mapped/ intersecting in both datasets (n=15,365). (a) Total mapped lake area for each year. (b) Comparison of Sentinel-2 and Sentinel-1 derived lake areas. (c) Plot (b) clipped to 0–0.5 km^2 to visualise the distribution of smaller lakes.

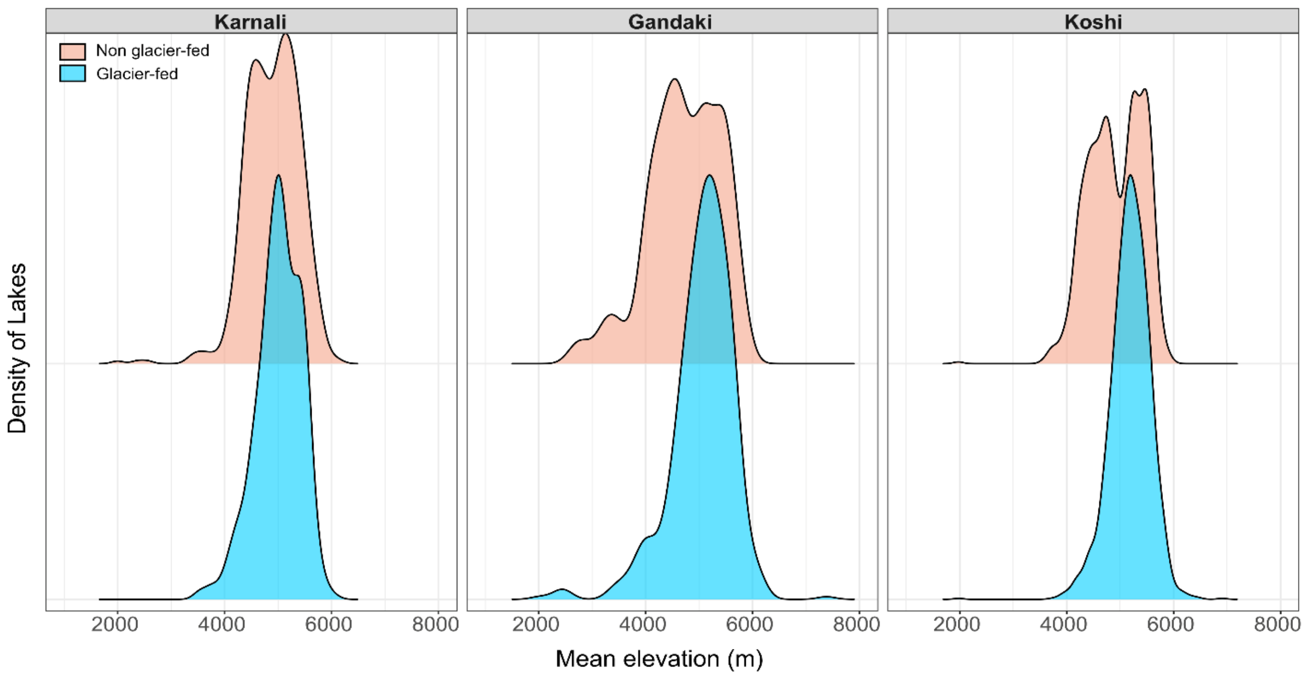

Figure S4: Density distribution of glacier-fed and non-glacier-fed lakes from the combined Sentinel-1/-2 delineations across mean elevation (m) within the three major Nepal-transboundary river basins: Karnali, Gandaki, and Koshi. Shaded areas represent the relative density of lakes by elevation, highlighting the elevational preferences of glacier-fed (blue) and non–glacier-fed (orange) lakes in each basin.

Table S4: The total number of lakes and their total cumulative annual area (km^2) per transboundary country (Nepal, China, India) per year from 2017 to 2024. Sentinel-1 and Sentinel-2 (bold) values are given. Sentinel-1 and -2 datasets are available at Rawlins et al. (2025).

Year	Nepal		China		India	
	Number of Lakes	Total area (km^2)	Number of Lakes	Total area (km^2)	Number of Lakes	Total area (km^2)
2017	1252 | **1576**	61.3 | **71.2**	1077 | **1155**	91.1 | **94.8**	25 | **31**	1.1 | **1.0**
2018	1251 | **1429**	63.8 | **69.1**	1059 | **1181**	91.2 | **95.4**	26 | **21**	1.0 | **0.8**
2019	1163 | **1455**	62.2 | **69.4**	1025 | **1243**	88.8 | **96.4**	16 | **25**	0.8 | **0.9**
2020	1173 | **1631**	63.6 | **73.9**	1019 | **1295**	90.8 | **99.0**	19 | **25**	0.9 | **1.0**
2021	1170 | **1421**	62.8 | **70.4**	1055 | **1322**	91.2 | **97.4**	17 | **10**	0.9 | **0.3**
2022	1235 | **1277**	64.2 | **66.2**	999 | **1295**	91.7 | **97.9**	20 | **14**	1.1 | **0.5**
2023	1208 | **1470**	64.2 | **71.4**	1013 | **1245**	91.9 | **97.8**	17 | **21**	0.9 | **1.0**
2024	1203 | **1786**	65.6 | **78.1**	1062 | **1325**	95.9 | **101**	25 | **41**	1.2 | **1.2**

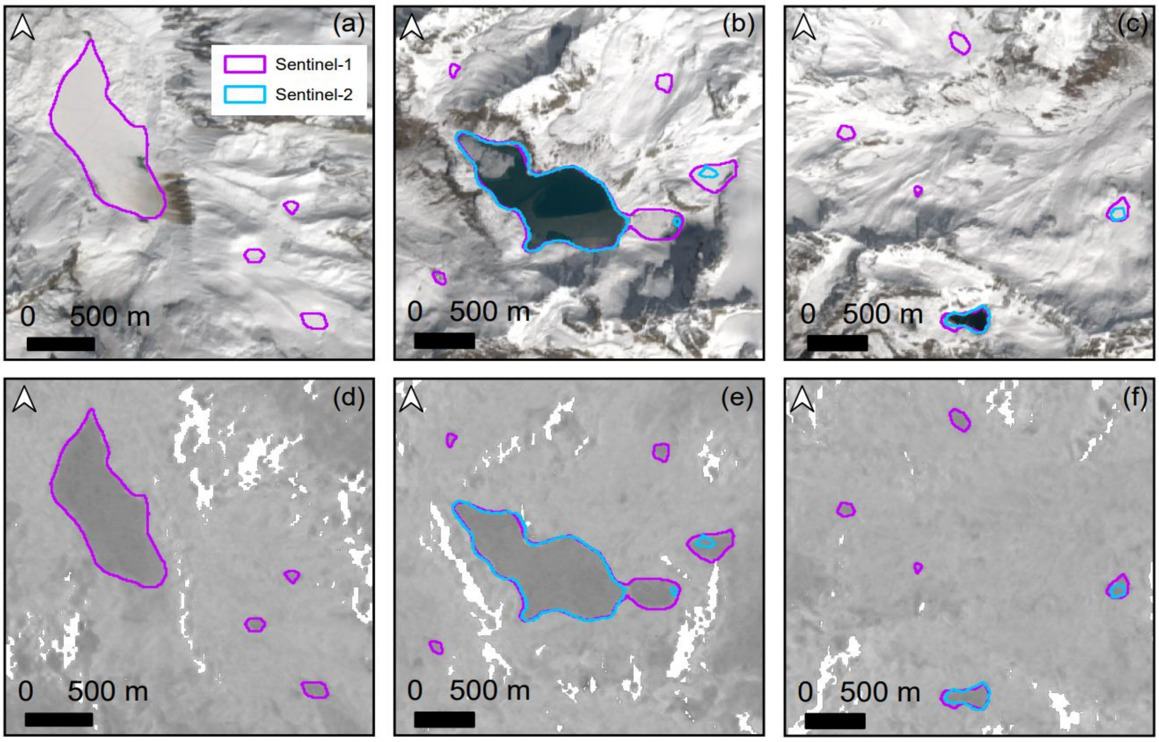

Figure S5: Lake outlines from Sentinel-1 (purple) and Sentinel-2 (blue) for the year 2022, when lake numbers classified via Sentinel-2 were lower. (a-c) Lake outlines overlaid on Sentinel-2 RGB composite images used for classification, showing a high presence of snow and ice; (d-f) Lake outlines overlaid on Sentinel-1, showing the ability of Sentinel-1 to delineate full lake extents when surface lake ice is present.

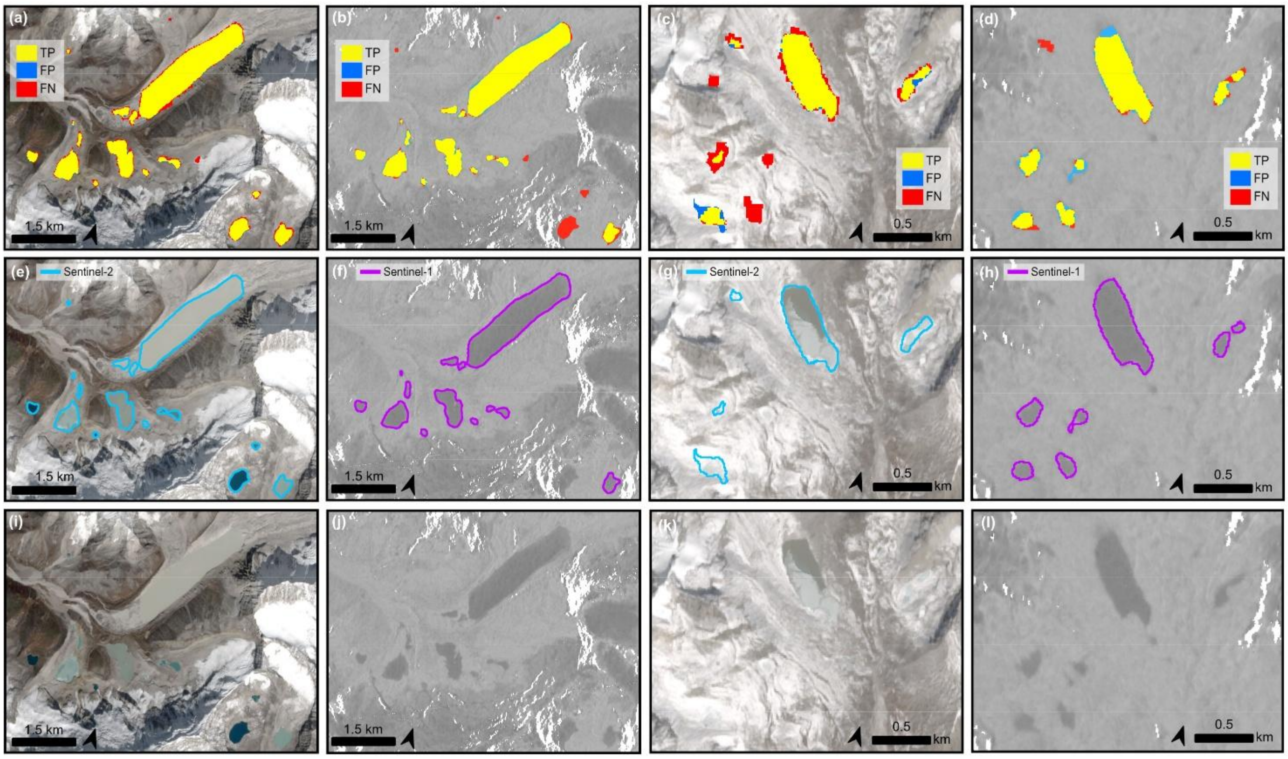

Figure S6: Examples of deep learning classification performance relative to reference inventories for Sentinel-2 and Sentinel-1 (2020) and subsequent lake delineations overlain on Sentinel-2 RGB and Sentinel-1 SAR training imagery. (a-d) True positives (TP), false positives (FP), and false negatives (FN), with left-hand pair (a,b) showing better performance in Sentinel-2, whereas right-hand pair (c,d) showing better performance in Sentinel-1 where snow and ice are more prevalent; (e-h) Corresponding deep learning lake outlines overlain on 2020 training imagery; (i-l) corresponding imagery only.

References

Bolch, T., Kulkarni, A., Kääb, A., Huggel, C., Paul, F., Cogley, J. G., Frey, H., Kargel, J. S., Fujita, K., Scheel, M., Bajracharya, S., and Stoffel, M.: The State and Fate of Himalayan Glaciers, Science, 336, 310–314, https://doi.org/10.1126/science.1215828, 2012.

GTN-G: GTN-G Glacier Regions. Global Terrestrial Network for Glaciers., https://dx.doi.org/10.5904/gtng-glacreg-2023-07, 2023.

Kumar, A., Mal, S., Schickhoff, U., and Dimri, A. P.: Basin-scale spatio-temporal development of glacial lakes in the Hindukush-Karakoram-Himalayas, Glob. Planet. Change, 245, 104656, https://doi.org/10.1016/j.gloplacha.2024.104656, 2025.

Rawlins, L. D., Watson, C. S., Bhambri, R., Khadka, N., and Chand, M. B.: Glacial Lake Observatory (GLO): A dataset of glacial lakes in Nepal and transboundary catchments (2017–2024) [Dataset], https://doi.org/10.5281/zenodo.17802334, 2025

Zhang, T., Wang, W., and An, B.: Heterogeneous changes in global glacial lakes under coupled climate warming and glacier thinning, Commun. Earth Environ., 5, 374, https://doi.org/10.1038/s43247-024-01544-y, 2024a.