

GLOF-Vulnerability Identification based on Automatic Method for Mapping Glacial Lakes in Cordillera Blanca (Peru)

SIBGRAPI Paper ID: 185

Abstract—Glacial lake outburst floods (GLOFs) pose a significant threat to high-mountain communities and infrastructure, particularly in glacierized regions such as the Cordillera Blanca, Peru. This study presents an automated method for mapping glacial lakes using Sentinel-2 satellite imagery with enhanced band combinations and a segmentation-based foundational model (SAM 2.1). Our method enables consistent multitemporal lake extraction, which successfully mapped 80% of 448 manually identified glacial lakes. We performed a comparative analysis of images taken from May 2016 and May 2024, and identified five lakes potentially vulnerable to GLOFs. Notably, Lake Parón and Lake Piticocha experienced significant surface area expansion of 13.79 and 3.72 hectares, respectively. The other three lakes showed significantly expansion despite its smaller size. GLOF simulations incorporating local topography and lake size indicate potential impacts on urban areas (e.g., Caraz city) and agricultural land. These findings highlight the importance of standarized and automated glacial lake monitoring for early risk assessment and disaster preparedness in the face of climate-driven glacial change.

I. INTRODUCTION

Climate change and glacial retreat have a direct influence on glacial lakes bodies [1]. Due to accelerated climate warming and glacier melting, glacial lakes became increasingly sensitive to environmental changes. This highlights the urgent need to monitor their evolution and understand the mechanisms driving their response to climate change. Performing a timely and accurate mapping of glacial lakes is crucial to identifying Potentially Dangerous Glacial Lakes (PDGLs), which pose a significant threat through the potential occurrence of Glacial Lake Outburst Floods (GLOFs) [2]. These flood events can have severe impacts on downstream communities, agricultural land, and infrastructure.

Frequent multitemporal mapping serves as a critical first step in assessing glacial lake evolution and identifying those lakes most vulnerable to outburst events. However, field-based observations are often limited due to the harsh natural environment and poor accessibility in mountain areas. In this context, remote sensing has emerged as a key technology, offering broad spatial coverage and repeated observations over multiple time intervals that enable the extraction of spatio-temporal characteristics of glacial lakes [3], [4].

Accurate delineation of glacial lake boundaries is essential for effective water resource management, climate impact studies, and GLOF risk assessment [5]. However, remote sensing-based mapping faces several challenges; these include image artifacts, low spectral resolution of free satellite images,

and adverse atmospheric conditions such as fog, shadows, sun glint, clouds, and water turbidity. Another challenge is the spectral similarity between water and thin ice that can complicate classification, even for trained experts, such conditions may lead to overestimation or underestimation of lake boundaries, often requiring extensive post-processing [6].

More recently, deep learning methods have demonstrated improved performance in mapping and monitoring the spatio-temporal evolution of glacial lakes [2], [7]–[9]. However these approaches typically require thousands of training labels to perform effectively, often created manually. In this study, we explore the use of segmentation foundational models, specifically the Segment Anything Model 2 (SAM 2) developed by Meta AI Research [10], due to its performance, versatility, and cross-platform adaptability. The robustness of SAM in cryospheric environments, such as glacier termini, supraglacial lakes, icebergs, and crevasses has been demonstrated in [11].

Building on these advances, we propose an automated method for mapping and delineating glacial lake boundaries using semantic segmentation techniques powered by foundational models. Our method enables the extraction of glacial lake outlines across different time periods to assess their evolution. Then, we identified potentially dangerous lakes and analyzed them through GLOF simulations, providing critical insights into the possible spatial extent and impact of such flood events.

After this Introduction, we present the area of interest and the dataset that we have built for our studies in Section 2. In Section 3, we present the proposed approach and in Section 4, we present the experimental results. In Section 5, we present some conclusions and future directions.

II. AREA OF INTEREST AND THE DATASET

Figure 1 shows the Cordillera Blanca - Peru (8.5° – 10.0° S; 77° – 78° W) which is the most glacierized tropical mountain range in the world, containing approximately 25% of all tropical glaciers [12]. It serves as a key region for research on GLOFs and associated risk management, with scientific interest in the area dating back to the 1940s [13].

The study area is defined by the boundaries of Huascarán National Park (HNP), which spans approximately $3,600\text{km}^2$. The park contains all peaks above 6,000 meters above sea level (ASL), including Huascarán (6,768m ASL), the highest mountain in Peru. The region is characterized by significant climate variability and dynamic geomorphological processes,



Fig. 1. Cordillera Blanca - Peru (8.5° – 10.0° S; 77° – 78° W).

surrounding features such as ice, snow, vegetation, and soil:
B11 (Short-Wave Infrared), B8 (Near-Infrared), and B4 (Red).
115
116

III. PROPOSED APPROACH

We present a method for automatically mapping and extracting
118 glacial lakes from satellite imagery, with a specific focus
119 on the Cordillera Blanca region. Based on the analysis of the
120 results, potentially GLOF-vulnerable lakes are identified, and
121 a model is built to assess the associated flood risks.
122

A. Glacial Lakes Mapping and Extraction

Image segmentation is essential for quantifying the size
124 and spatial distribution of glacial lakes. We use the recently
125 released **Segment Anything Model 2 (SAM 2)** developed
126 by Meta AI Research, a foundational model in the field of
127 artificial intelligence that leverages self-supervised learning to
128 generalize across tasks without requiring task-specific training
129 data [10]. **SAM version 2.1** offers improved performance in
130 detecting small objects and handling occlusions caused by
131 features such as mountain shadows in the glacial lake context.
132

The foundation model is used in two stages of the method:
133

- sam2.1_hiera_l is applied for the first image of the
134 collection to localize glacial lakes. The centroids of the
135 detected lakes are saved and later used as point markers
136 for SAM's prompts, which are applied to the subsequent
137 temporal images in the collection.
138
- sam2.1_hiera_base_plus is then applied to seg-
139 ment lakes in the remaining images using the saved
140 prompts.
141

Next, we compute the area of each detected lake and track
142 the temporal changes of their surface area throughout the study
143 period. Figure 3 shows an example of the results obtained by
144 our pipeline for a single ROI.
145

These results enable further analysis to identify glacial
146 lakes that are potentially dangerous and vulnerable to GLOFs.
147 Once these high-risk lakes are identified, GLOF modeling is
148 performed to simulate potential flood scenarios and assess their
149 spatial impacts.
150

B. Modeling GLOFs method

To simulate potential GLOF scenarios, we developed a
152 two-stage method that models water propagation based on
153 Flood Fill Algorithm [22] from a glacial lake over a Digital
154 Elevation Model (DEM) and estimates the spatial impact based
155 on Gaussian functions [23], [24]. The method consists of the
156 following stages:
157

1) *Water propagation:* The simulation begins at the cen-
158 troid of the glacial lake (x_0, y_0) , where the water level is
159 artificially raised by a specified height n , representing a
160 potential overflow trigger. From this initial point, the algorithm
161 iteratively propagates downslope by selecting neighboring
162 DEM cells with the steepest gradient, as long as the slope
163 exceeds a predefined threshold. The process continues until no
164 other valid downward path is available, simulating the likely
165 flow route of a GLOF event based on local topography.
166

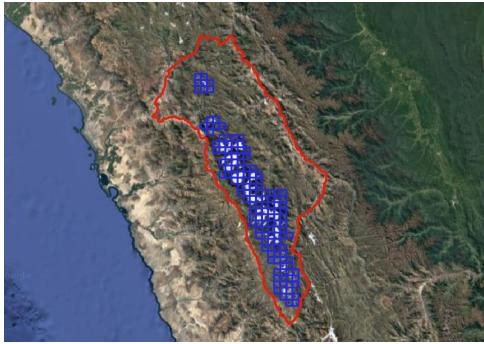


Fig. 2. Grid of $5\text{ km} \times 5\text{ km}$ cells defining the Regions of Interest (ROIs)
covering the Cordillera Blanca study area.

Each ROI defines a bounding box to extract Sentinel-2
109 image collections from Google Earth Engine, selecting only
110 scenes with less than 5% cloud cover. The corresponding
111 image collections are publicly available on our GitHub re-
112 pository². For glacial lake mapping, three spectral bands were
113 used due to their effectiveness in distinguishing water from
114

¹COPERNICUS/S2_HARMONIZED

²<https://github.com/NuryStud/automated-glof-mapping>

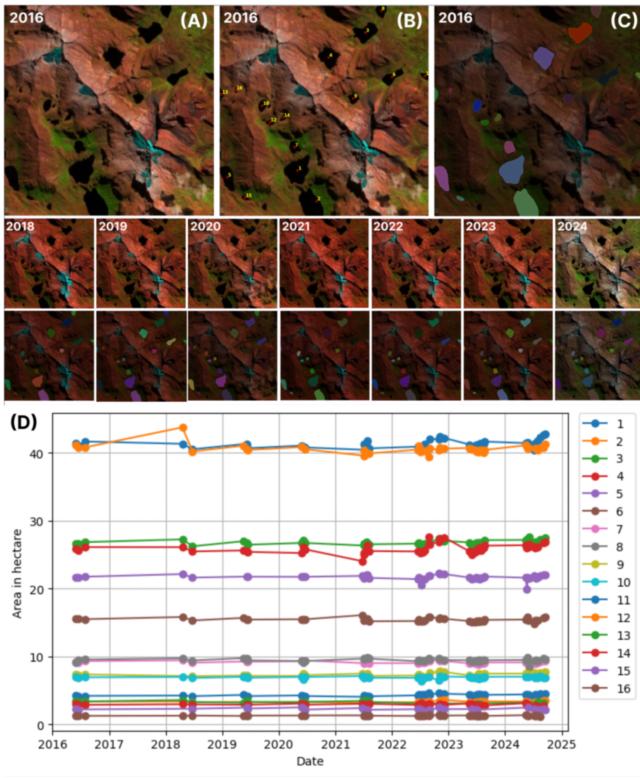


Fig. 3. Glacial lakes mapping and monitoring for an image collection (A) First image of image collection (B) A total of 16 glacial lakes are detected and their centroid are used as point prompts for subsequently processing (C) Glacial lakes are segmented across the remaining image collection (2016 to 2025) using the predefined prompts (D) Temporal evolution of the surface area of each mapped lake throughout the study period.

2) *Spatial Impact Estimation*: Following the path of water propagation, an *influence map* is generated to estimate the spatial extent of the potential flood impact. For each cell affected by the simulated flow, a circular buffer is applied with a radius proportional to the originating lake's surface area. Within each buffer zone, the influence is spatially distributed using a Gaussian decay function, providing a probabilistic estimate of flood impact intensity across the landscape.

IV. EXPERIMENTAL RESULTS

In this section, we present some experimental results of the proposed method on the Huascarán National Park dataset.

A. Glacial lakes mapping and tracking

A total of 360 glacial lakes were automatically mapped using the proposed method in Sec. III-A, out of 448 glacial lakes in the Cordillera Blanca. The remaining unmapped lakes were manually inspected across the entire region, these missed lakes are typically small, with areas of less than 100 pixels ($10000m^2$). Due to the Sentinel-2 spatial resolution ($100 m^2$ per pixel), such small lakes have less defined boundaries, making their segmentation more prone to errors and potentially leading to a misleading perception of lake expansion or contraction, to mitigate this issue, only lakes an

area greater than $10000m^2$ were considered in this study.

The mapped (segmented) lakes by our proposed method, with its location determined by the centroid coordinate derived from the center of the segmented lake, are saved in a GeoJson file and can be visualized using the geojson.io platform [25], as shown in Fig. 4. Each blue-marked circle represents a lake, illustrating the spatial distribution of the 360 glacial lakes mapped for the proposed method throughout the Cordillera Blanca region in Peru.

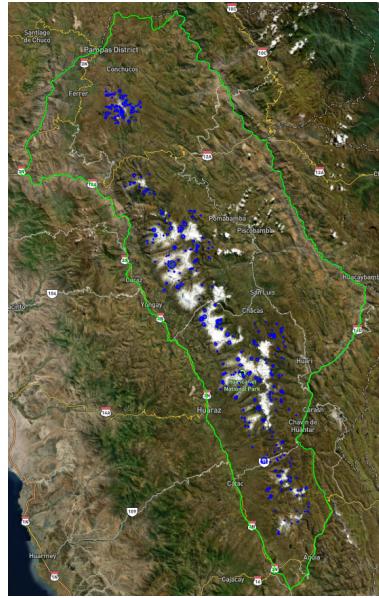


Fig. 4. Mapped glacial lakes in Cordillera Blanca in May 2024.

Figure 5 shows a boxplot distribution of glacial lake areas in May 2024 when compared to May 2016 (same seasonal period, exactly the end of the rainy season and the beginning of the dry season). The horizontal axis shows the boxplot categorized by the type of change (no change, negative change, or positive change.) The analysis reveals that 194 lakes increased in area, 157 decreased, and nine showed no change, up to the precision of the resolution. The summary statistics of the percentage change in lake area between the two-time points are provided in Table I.

- Among the growing lakes, the mean percentage increase in area is 2.89%. It is also noted that 25% (Q1) of the lakes expanded by more than 4.08%. This fact can be an indication that these lakes may be undergoing significant hydrological expansion due to glacier melt and climate variability.
- For the shrinking lakes, the mean percentage decrease is 2.60%. One-fourth of the lakes lost more than 3.69% of their area, possibly due to severe retreat, rapid drying or sedimentation.

B. GLOF-Vulnerable Glacial Lakes

The potentially dangerous glacial lakes identified in this study are presented in Figure 6. These 5 lakes exhibit the

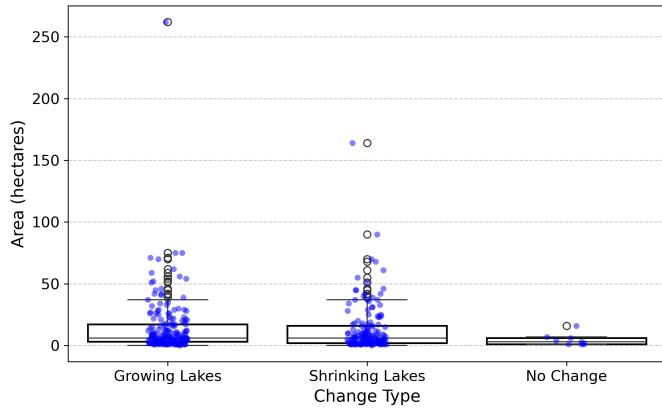


Fig. 5. Glacial lake areas in May 2024 grouped by type of change (Growing Lakes; Shrinking Lakes; and No Change) compared to May 2016.

TABLE I
SUMMARY STATISTICS OF PERCENTAGE AREA CHANGES FOR GLACIAL LAKES BETWEEN MAY 2024 AND MAY 2016.

	count	mean	std	min	Q1	Q2	Q3	max
Growing lakes	194	2.89	2.42	0.02	0.10	2.03	4.08	9.90
Shrinking Lakes	157	2.60	2.24	0.05	0.90	1.88	3.69	9.59
No Change	9	0	0	0	0	0	0	0

222 highest percentage increase in area between May 2016 and
223 May 2024 and lakes with significantly more increase of
224 water.

225 Among them, lake Patarcocha shows the greatest relative
226 increase, with a 9.90% expansion. However, this corresponds
227 to a modest absolute increase of 480 m². Other notable cases
228 include two smaller lakes shown in the second and third rows
229 of Fig. 6, which experienced increases of 9.38% and 7.63%
230 respectively, equivalent to absolute area changes of 850 m²
231 and 660 m². Laguna Paron, one of the largest lakes in the
232 Cordillera Blanca, recorded an increase of 5.54% increase in
233 area, amounting to a substantial gain of 137.900 m². Similarly,
234 Piticocha expanded by 5.22%, equivalent to an increase of
235 37.200 m², which is a significant increase in its size. These
236 changes may reflect growing risks related to glacial lake
237 outburst floods (GLOFs) and highlight the need for ongoing
238 monitoring.

239 C. Modeling Potential GLOFs

240 Assessment of GLOF scenarios was conducted for the lakes
241 identified in Subsection IV-B. Elevation data were obtained
242 from the OpenTopography Global Digital Elevation Model
243 (DEM) API [26] and used to simulate water flow propagation
244 and estimate the spatial extent of potential flood impacts,
245 following the method described in Subsection III-B. For each
246 simulation, is indicated in the figures the value of n used to
247 model the flood, representing the artificially raised water level.

248 The objective of this modeling is to identify downstream
249 areas at risk in the event of a sudden lake outburst. The sim-
250 ulation results allow for visualizing the extent of inundation,
251 assessing potential damage to infrastructure and settlements,

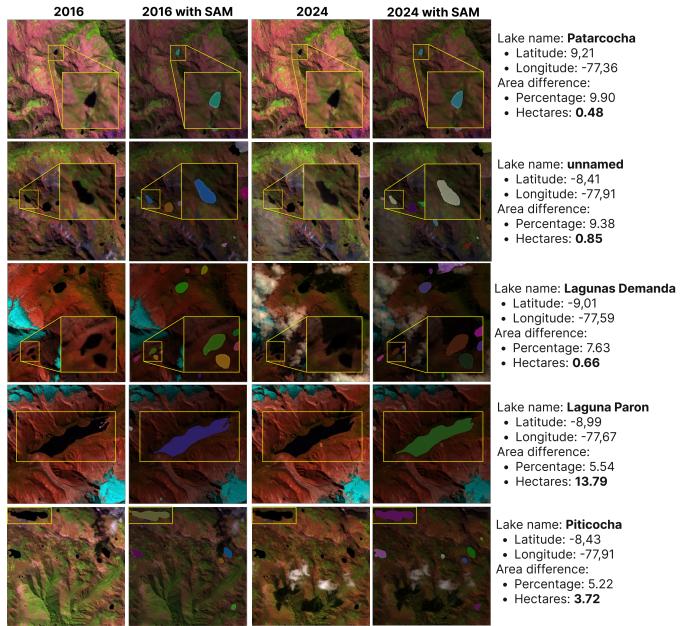


Fig. 6. Glacial lakes with the highest percentage area increase from May 2016 to May 2024, highlighting growing GLOF risks. Each row shows a lake (circled in yellow): columns 1 and 3 present satellite images; columns 2 and 4 show segmentation results from our method.

and supporting early warning and risk mitigation strategies. In
253 the following subsections, individual simulations for selected
254 high-risk lakes are presented, with particular attention to
255 populated areas, infrastructure and agricultural area in the
256 potential flood paths.

257 1) *Simulation of Laguna Paron:* As shown in Fig. 7,
258 the simulated GLOF from Laguna Paron would significantly
259 affect the city of Caraz, particularly built-in areas, posing
considerable risk to the local population.

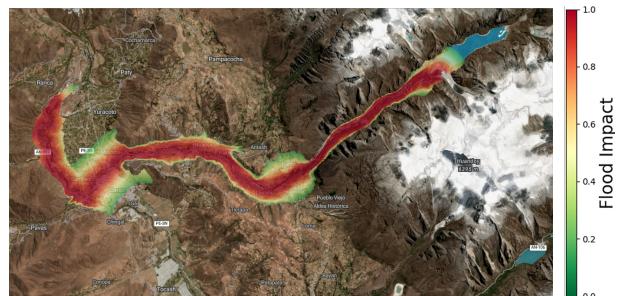


Fig. 7. Scenario of GLOF simulation ($n = 38$) for Paron lake. The color scale on the right of the figures shows the expected amount of impact of a outburst phenomenon.

260 2) *Simulation of Piticocha:* The simulation results shown
261 in Fig. 8 indicate that areas along the river between Aija
262 Caserio and Piedra Colorado would be affected by potential
263 GLOF from lake Piticocha. The primary impact zone includes
264 agricultural lands, which could suffer significant damage due
265 to flooding.

266 3) *Simulation of Unnamed Lake:* The simulation results
267 shown in Fig. 9 indicate that a potential GLOF from the



Fig. 8. Scenario of GLOF simulation for Piticocha lake ($n = 10$).

unnamed lake would follow the river channel without impacting nearby communities or agricultural areas. The flood path remains confined to uninhabited and non-productive zones.



Fig. 9. Scenario of GLOF simulation for the Unnamed lake ($n = 8$).

4) *Simulation of Patarcocha:* The simulation results shown in Fig. 10 indicate that a potential GLOF from Lake Patarcocha could affect nearby agricultural regions, particularly the area around Chucpin.

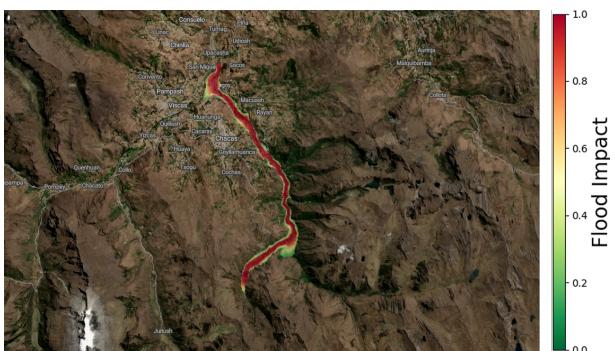


Fig. 10. Scenario of GLOF simulation for the Patarcocha lake ($n = 11$).

5) *Simulation of Lagunas Demanda:* The simulation results shown in Fig. 11 indicate that a potential GLOF from Lagunas Demanda would follow the river channel without affecting any nearby communities or agricultural areas.

V. CONCLUSION

In this work, we proposed an automatic method for mapping and extracting glacial lakes using a segmentation-based foundational model applied to Sentinel-2 imagery, with enhanced

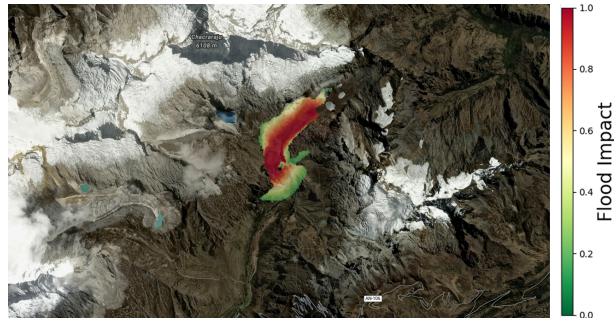


Fig. 11. Scenario of GLOF simulation for the Lagunas Demanda lake ($n = 10$).

band combinations to highlight lakes features. The method enables consistent monitoring of glacial lakes in the Cordillera Blanca from 2016 to 2025. To ensure temporal consistency, glacial lake surface areas from May 2016 and May 2024 were compared.

A total of 448 glacial lakes were mapped, with 80% automatically detected and extracted without post-processing. Five lakes were identified as potentially GLOF-vulnerable. Lake Parón and Lake Piticocha showed surface area increases of 13.79 and 3.72 hectares, respectively. Lake Patarcocha, Lagunas Demanda, and an unnamed lake ($8^{\circ} 24' 23.47''$ S, $77^{\circ} 54' 41.71''$ W) recorded area increases of 9.9%, 9.38%, and 0.66%.

GLOF simulations based on local topography indicated that Lake Parón poses a risk to built-up areas, including Caraz city. Piticocha and Patarcocha may affect nearby agricultural zones in a GLOF scenario.

This approach underscores the value of automated, standardized glacial lake monitoring for timely risk assessment and mitigation of GLOF hazards in the Cordillera Blanca and other glacierized regions.

REFERENCES

- [1] R. Hugonet, R. McNabb, E. Berthier, B. Menounos, C. Nuth, L. Girod, D. Farinotti, M. Huss, I. Dussaillant, F. Brun *et al.*, “Accelerated global glacier mass loss in the early twenty-first century,” *Nature*, vol. 592, no. 7856, pp. 726–731, 2021.
- [2] H. Zhao, X. Liu, and S. Wang, “Using landsat-8 imagery and generative adversarial network for glacial lakes mapping in high mountain regions,” in *2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, vol. 2. IEEE, 2021, pp. 877–881.
- [3] C. Song, B. Huang, L. Ke, and K. S. Richards, “Remote sensing of alpine lake water environment changes on the tibetan plateau and surroundings: A review,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 92, pp. 26–37, 2014.
- [4] R. Ahmed, S. T. Ahmad, G. F. Wani, R. A. Mir, P. Ahmed, and S. K. Jain, “High resolution inventory and hazard assessment of potentially dangerous glacial lakes in upper jhelum basin, kashmir himalaya, india,” *Geocarto International*, vol. 37, no. 25, pp. 10 681–10 712, 2022.
- [5] F. Chen, “Comparing methods for segmenting supra-glacial lakes and surface features in the mount everest region of the himalayas using chinese gaofen-3 sar images,” *Remote Sensing*, vol. 13, no. 13, p. 2429, 2021.
- [6] Y. Zhang, J. Zhao, X. Yao, H. Duan, J. Yang, and W. Pang, “Inventory of glacial lake in the southeastern qinghai-tibet plateau derived from sentinel-1 sar image and sentinel-2 msi image,” *Remote Sensing*, vol. 15, no. 21, p. 5142, 2023.

- 331 [7] R. Wu, G. Liu, R. Zhang, X. Wang, Y. Li, B. Zhang, J. Cai, and
332 W. Xiang, "A deep learning method for mapping glacial lakes from the
333 combined use of synthetic-aperture radar and optical satellite images,"
334 *Remote Sensing*, vol. 12, no. 24, p. 4020, 2020.
- 335 [8] N. Qayyum, S. Ghuffar, H. M. Ahmad, A. Yousaf, and I. Shahid,
336 "Glacial lakes mapping using multi satellite planetscope imagery and
337 deep learning," *ISPRS International Journal of Geo-Information*, vol. 9,
338 no. 10, p. 560, 2020.
- 339 [9] S. Wang, M. V. Peppa, W. Xiao, S. B. Maharjan, S. P. Joshi, and J. P.
340 Mills, "A second-order attention network for glacial lake segmentation
341 from remotely sensed imagery," *ISPRS Journal of Photogrammetry and*
342 *Remote Sensing*, vol. 189, pp. 289–301, 2022.
- 343 [10] N. Ravi, V. Gabeur, Y.-T. Hu, R. Hu, C. Ryali, T. Ma, H. Khedr,
344 R. Rädle, C. Rolland, L. Gustafson *et al.*, "Sam 2: Segment anything in
345 images and videos," *arXiv preprint arXiv:2408.00714*, 2024.
- 346 [11] S. Shankar, L. A. Stearns, and C. J. van der Veen, "Segment anything in
347 glaciology: An initial study implementing the segment anything model
348 (sam)," 2023.
- 349 [12] G. Kaser and H. Osmaston, *Tropical Glaciers*. Cambridge: Cambridge
350 University Press, 2002.
- 351 [13] A. Emmer, "Glofs in the wos: bibliometrics, geographies and
352 global trends of research on glacial lake outburst floods (web of
353 science, 1979–2016)," *Natural Hazards and Earth System Sciences*,
354 vol. 18, no. 3, pp. 813–827, 2018. [Online]. Available: <https://nhess.copernicus.org/articles/18/813/2018/>
- 355 [14] L. G. Thompson, E. Mosley-Thompson, and K. A. Henderson, "Ice-core
356 palaeoclimate records in tropical south america since the last glacial
357 maximum," *Journal of Quaternary Science*, vol. 15, no. 4, pp. 377–394,
358 2000.
- 359 [15] O. Solomina *et al.*, "Lichenometry in the cordillera blanca, peru: "little
360 ice age" moraine chronology," *Global and Planetary Change*, vol. 59,
361 no. 1-4, pp. 225–235, 2007.
- 362 [16] L. Barcenas *et al.*, "Current state of glaciers in the tropical andes:
363 a multicentury perspective on glacier evolution and climate change,"
364 *Cryosphere*, vol. 7, no. 1, pp. 81–102, 2013.
- 365 [17] E. Mayr *et al.*, "Climate trends and glacier retreat in the cordillera
366 blanca, peru, revisited," *Global and Planetary Change*, vol. 84-85, pp.
367 39–44, 2012.
- 368 [18] M. Clark *et al.*, "Glaciological problems set by the control of dangerous
369 lakes in cordillera blanca, peru. i. historical failures of moranic dams,
370 their causes and prevention," *Journal of Glaciology*, vol. 18, no. 79, pp.
371 239–254, 1977.
- 372 [19] M. L. Zapata, "La dinámica glaciar en lagunas de la cordillera blanca,"
373 *Acta Montana*, vol. 123, pp. 37–60, 2001.
- 374 [20] M. Carey, "Living and dying with glaciers: people's historical vulnerability
375 to avalanches and outburst floods in peru," *Global and Planetary
376 Change*, vol. 47, no. 2-4, pp. 122–134, 2005.
- 377 [21] A. Emmer, "Geomorphologically effective floods from moraine-dammed
378 lakes in the cordillera blanca, peru," *Quaternary Science Reviews*, vol.
379 177, pp. 220–234, 2017.
- 380 [22] T. Pavlidis, *Algorithms for graphics and image processing*. Springer
381 Science & Business Media, 2012.
- 382 [23] Á. V. Jóhannesson, S. Siegert, R. Huser, H. Bakka, and B. Hrafinkelsson,
383 "Approximate bayesian inference for analysis of spatiotemporal flood
384 frequency data," *The Annals of Applied Statistics*, vol. 16, no. 2, pp.
385 905–935, 2022.
- 386 [24] F. J. Gomez, K. Jafarzadegan, H. Moftakhari, and H. Moradkhani,
387 "Probabilistic flood inundation mapping through copula bayesian multi-
388 modeling of precipitation products," *Natural Hazards and Earth System
389 Sciences*, vol. 24, no. 8, pp. 2647–2665, 2024.
- 390 [25] Mapbox contributors, "geojson.io," <https://geojson.io>, 2025, accessed:
391 2025-05-24.
- 392 [26] OpenTopography, "Global digital elevation model (dem) api," 2025,
393 [Online database]. Retrieved on May 25, 2025, from <https://portal.opentopography.org/API/globaldem>.