Supplement of

Automated snow cover detection on mountain glaciers using space-borne imagery

Aberle et al.

Correspondence to: Rainey Aberle (raineyaberle@u.boisestate.edu)

S1 PlanetScope radiometric adjustments

We observed substantial variations in the dynamic range of PlanetScope 4-band surface reflectance images, even after the prior harmonization with Sentinel-2 imagery (Fig. S1a). To improve the consistency and comparability of these images for classification purposes, we implemented adjustments to the surface reflectance values. Our approach involved aligning the mean band values of pixels within the top 20% of elevations within the glacier area with the characteristics of snow, while setting the lowest band values to 0.0. First, we constructed a polygon outlining the uppermost 20% of elevations within the glacier area using the respective glacier boundary and digital elevation model (DEM; Fig. S1b–c). Next, for each image, we calculated the median surface reflectance value for all pixels contained within this polygon. The surface reflectances for all bands were then adjusted according to the following relationship,

$$\rho_{adj} = \rho * x - y$$

where ρ is the surface reflectance of a given image pixel and ρ_{adj} is the adjusted surface reflectance value. The constant x is the ratio between the range of the original surface reflectance values and the adjusted surface reflectance values:

$$x = \frac{\rho_{adj,max}}{\rho_{max} - \rho_{min}}$$

where $\rho_{adj,max}$ is the maximum surface reflectance of the adjusted image and ρ_{max} and ρ_{min} are the maximum and minimum values of the original image, respectively. The constant y is the minimum surface reflectance of each band, such that subtracting y results in minimum band values of 0.0. $\rho_{adj,max}$ is equal to 0.94 for the blue band, 0.95 for the green band, 0.94 for the red band, and 0.78 for the near infrared band, based on snow reflectance at each band's center wavelength from Painter et al. (2009).

Figure S1 depicts an original and adjusted PlanetScope 4-band surface reflectance image at Lemon Creek Glacier, with the polygon outlining the top 20% of elevations within the glacier area shown. This image adjustment process produces not only a more consistent PlanetScope image collection, but also typically results in greater contrast between snow, ice, and firn within individual images, which we found to improve classification performance metrics.

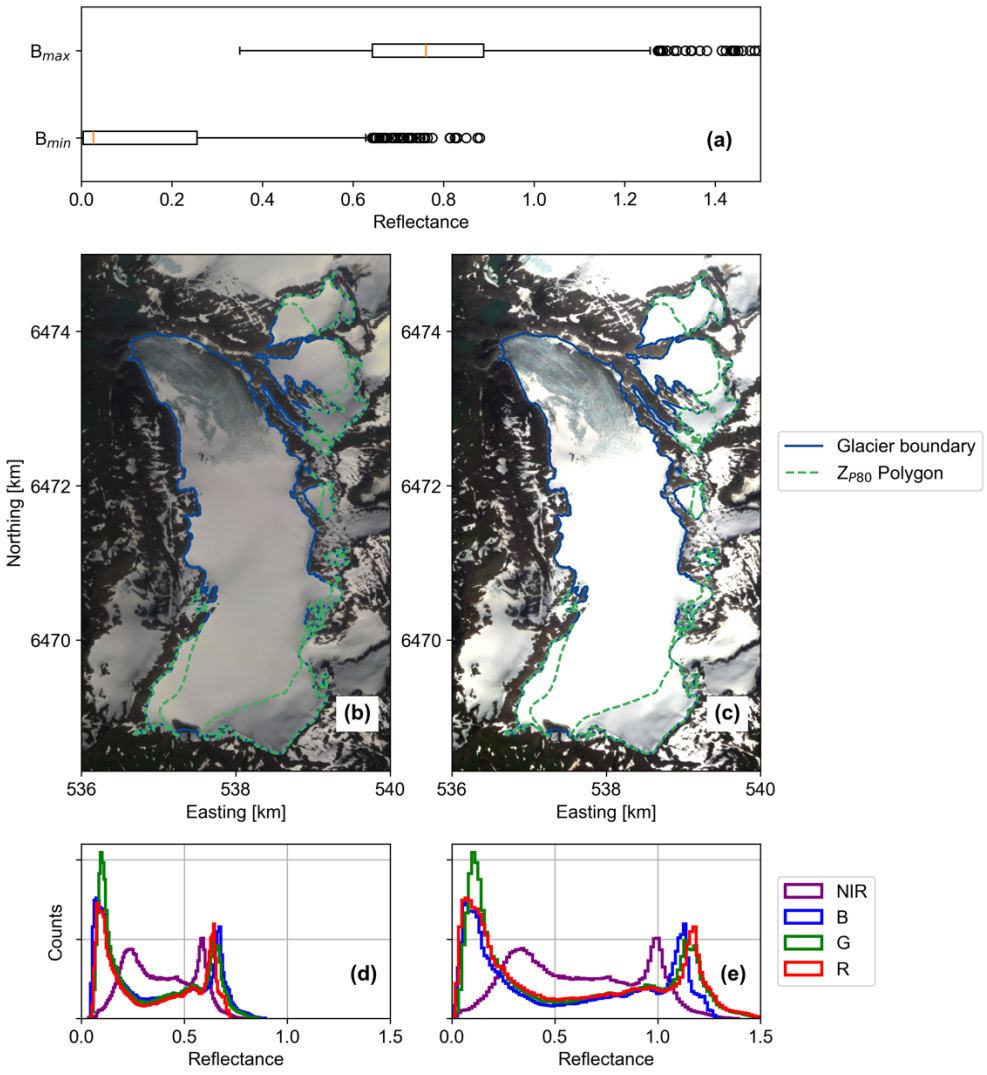

Figure S1: (a) Distributions of Blue band ("B") minimum and maximum surface reflectance values for all PlanetScope image mosaics collected at Lemon Creek Glacier for 2013–2022, demonstrating the relatively large variations in dynamic ranges between images. (b) Original image and (c) adjusted image captured 29 July 2020 at Lemon Creek Glacier. The glacier boundary and the polygon outlining the top 20% of elevations within the glacier area (or 80th percentile elevation, "Z_{P80} Polygon") are shown in each image. Histograms of surface reflectance values for each band are shown for (d) the original image and (e) the adjusted image.

S2 Performance assessment: all machine learning algorithms

To assess the performance of the classification models reported in the main body of the paper, we calculated the overall accuracy and learning curves for all nine machine learning models tested. Table S1 shows the overall accuracy for all models tested. The optimal classification models that were used to construct timeseries of snow-covered area (SCA) include the

Nearest Neighbors for Landsat 8/9 and PlanetScope imagery and the Support Vector Machine model for Sentinel-2 SR and TOA imagery (Table S1, highlighted in gray).

Table S1: Mean overall accuracies for all tested machine learning models resulting from K-folds cross-validation. Bolded and highlighted values indicate the most accurate classifier for each satellite image product. Note that the optimal models may have comparable accuracies to other models, but had higher accuracies at greater precision than shown.

Model name	Landsat 8/9 SR	PlanetScope 4-band SR	Sentinel-2 SR	Sentinel-2 TOA
AdaBoost	74 %	87 %	82 %	53 %
Decision Tree	92 %	96 %	96 %	93 %
Naïve Bayes	82 %	93 %	91 %	83 %
Nearest Neighbors	**94 %**	**97 %**	93 %	92 %
Neural Network	90 %	93 %	94 %	89 %
Random Forest	92 %	96 %	95 %	93 %
Support Vector Machine	93 %	97 %	**98 %**	**93 %**
Quadratic Discriminant Analysis	93 %	97 %	97 %	90 %
Logistic Regression	21 %	95 %	96 %	93 %

To assess the robustness of the optimal models selection, we calculated learning curves to illustrate variations in accuracy scores for different training dataset sizes and subsets across all models (Viering & Loog, 2023). The training dataset sizes ranged from 100 and 6,500 points, tested at various increments. Figure S2 shows the learning curves for all models, focusing on training dataset sizes between 500 and 6,500 points in 1,000-point increments. For each training dataset size n, a random sample of n points was used for model training, and $n*0.2$ points were allocated for testing and accuracy score calculation. Notably, the cross-validated accuracy of the optimal models exhibited minimal variation (±5%) in response to changes in the training sample or size (Fig. S2: Nearest Neighbors in dark orange, Support Vector Machine in dark purple). The accuracy scores for the optimal models are nearly constant for greater than or equal to 2,500 points for the Sentinel-2 TOA model and greater than or equal to 1,500 for all other image product models. This consistent performance instills confidence in the optimal models, indicating relative insensitivity to variations in the input training dataset size and subset.

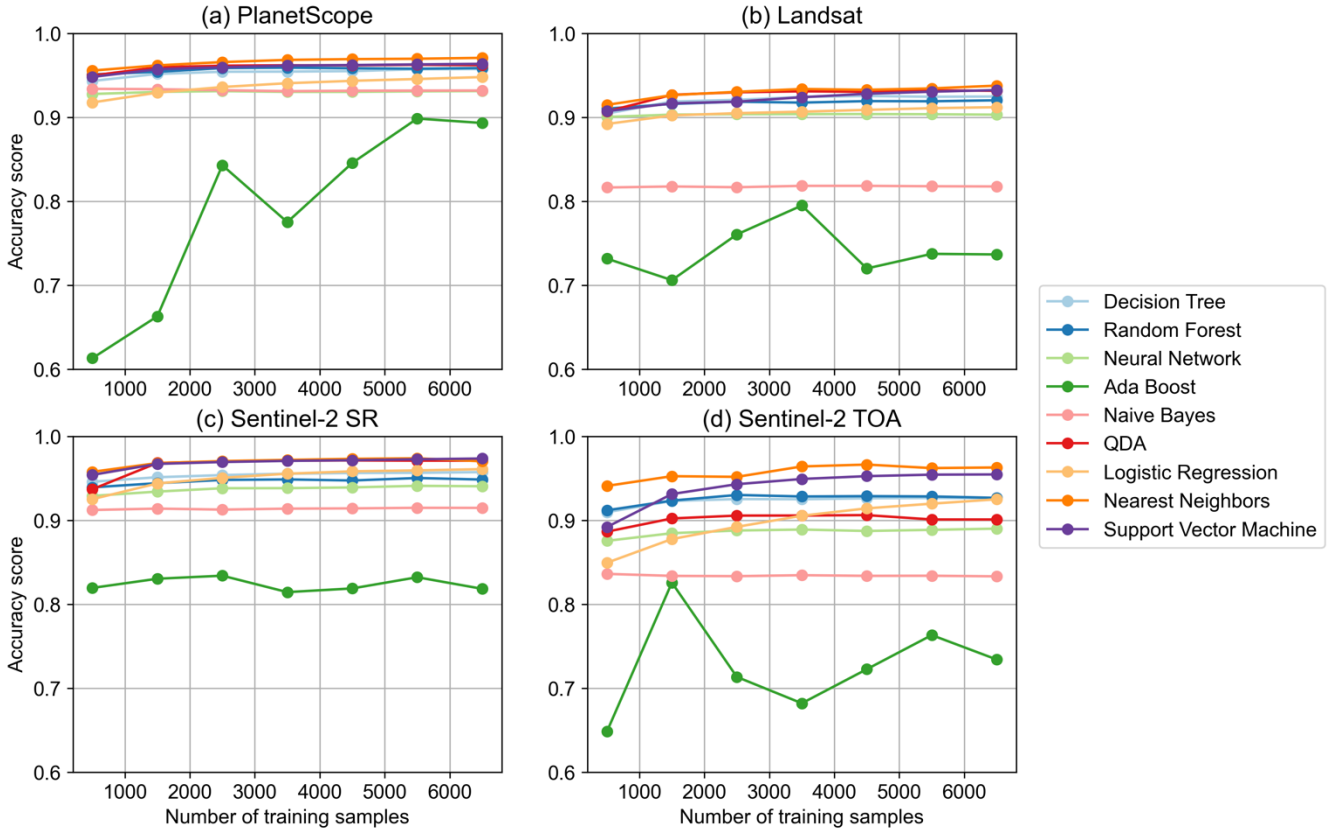

Figure S2: Learning curves for each machine learning model tested for each image product. Accuracy scores are shown for the testing set.

S3 Full time series of snow-covered area

Figure S4 shows the time series of SCA for the USGS Benchmark Glaciers using Landsat, PlanetScope, and Sentinel-2 imagery. The timeseries broadly demonstrates the density of observations in time and increasing coverage upon the launch of additional satellites from Sentinel-2 and PlanetScope starting in ~2016. Notably, the PlanetScope-derived time series (Fig. S4, green circles) tends to be much noisier compared to the other image products (i.e., the SCA varies greatly on an intra-weekly basis), as discussed in the main text (Sect. 4.2). Nonetheless, these frequent observations allow for improved understanding of temporal patterns in snow cover across different satellite platforms, providing additional insights into the challenges and nuances associated with each dataset.

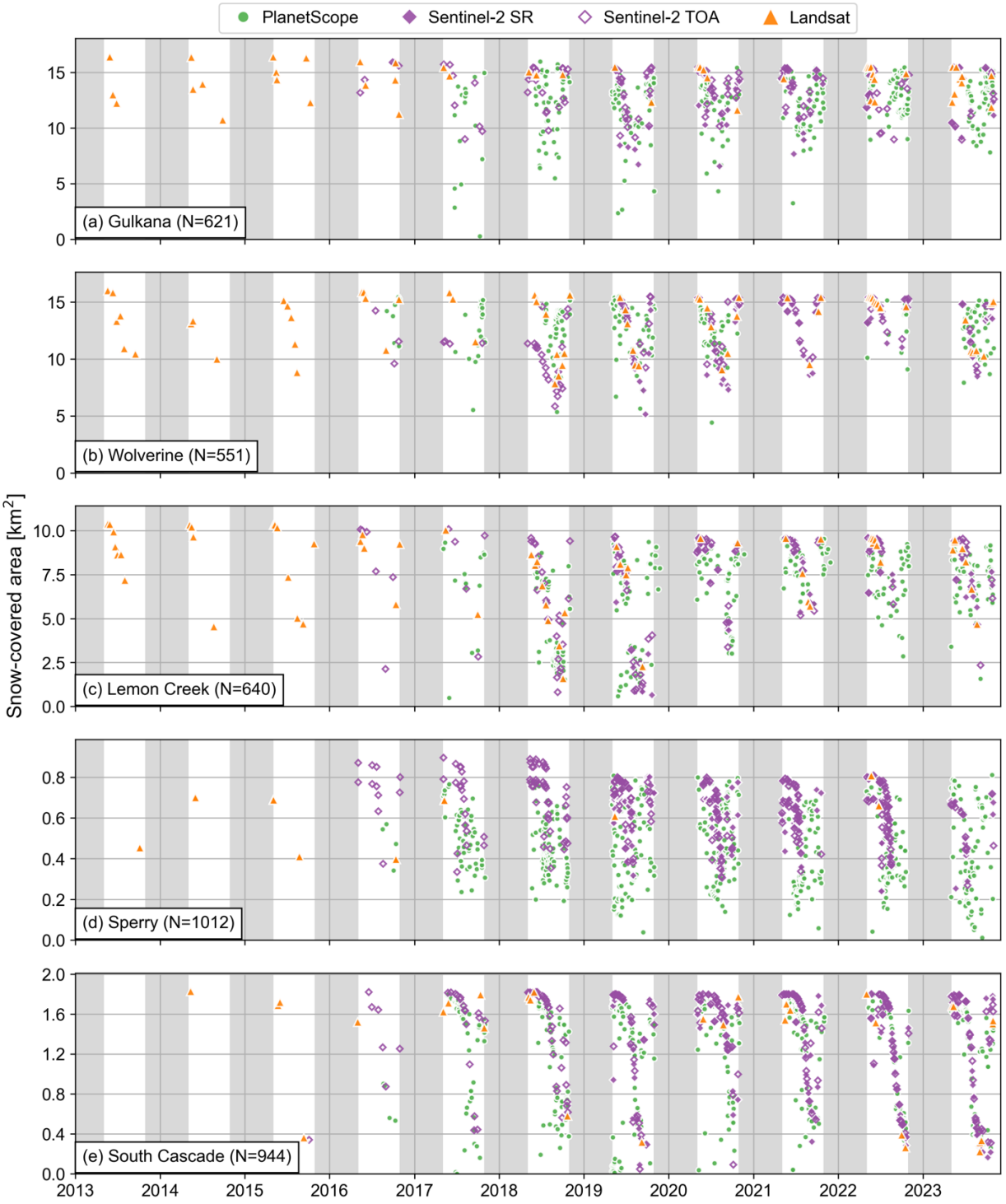

Figure S3: (Left panels) Time series of snow-covered area for the USGS Benchmark Glaciers with marker types distinguishing each image product. Gray shaded regions indicate dates outside the observation period (1 Nov–30 April).

To increase the reliability of accumulation area ratio (AAR) and snow-covered elevation estimates amid changing glacier areas and surface elevations through time, glacier boundaries and DEMs glacier boundaries were updated over the 2013–2023 study period using all available data products from the USGS Benchmark Glacier data release version 8 (McNeil et al., 2022). For each summer melt season, we used the closest glacier boundary and DEM in time to construct time series of the SCA, AAR, and snowline elevations. The timestamps corresponding to the glacier boundaries and DEMs used for each snow detection year are detailed in Table S2. This approach improves temporal consistency and accuracy in capturing the dynamic nature of glacier surface areas and elevations throughout the study period.

Glacier name	Snow detection year	Glacier boundaries date	DEM date
Gulkana	2013	2014-10-07	2014-10-07
	2014		
	2015		
	2016	2016-08-30	2016-08-30
	2017		
	2018	2021-09-19	2021-09-19
	2019		
	2020		
	2021		
	2022		
	2023		
Wolverine	2013	2012-08-22	2012-08-22
	2014	2015-08-13	2015-08-13
	2015		
	2016	2016-09-01	
	2017		2018-09-12
	2018	2018-09-12	
	2019	2019-08-19	2020-05-02
	2020	2020-10-19	
	2021		2020-10-19
	2022		
	2023		
Lemon Creek	2013	2013-09-04	2021-10-05
	2014	2014-09-08	
	2015		
	2016	2016-08-28	
	2017		
	2018	2018-09-02	
	2019		
	2020	2021-10-05	
	2021		
	2022		
	2023		
Sperry	2013	2014-09-07	2014-09-07
	2014		
	2015		

	2016		
	2017		
	2018		
	2019		
	2020		
	2021		
	2022		
	2023		
South Cascade	2013	2015-10-14	2015-10-14
	2014		
	2015		
	2016		
	2017		
	2018		
	2019	2021-08-13	2021-08-13
	2020		
	2021		
	2022		
	2023		

Table S2. Timestamps of glacier boundaries and digital elevation models used for each snow detection year for each study site.

S4 Potential impacts of slope on exposed firn

The area of exposed firn on the surface is sensitive to a number of factors, including short-term climate perturbations and surface slope gradient. In a year with higher-than-average air temperatures, the decrease in accumulation area and increase in exposed firn area will typically be greater for glaciers with shallower surface slopes than those with steeper slopes (Geyman et al., 2022). For example, if the zero-degree isotherm is raised by 0.5 ℃ at two different glaciers with the same widths, the equilibrium line altitude (ELA) will increase by about 50 m, assuming a dry adiabatic lapse rate of 9.8 ℃/km (Lutgens & Tarbuck, 1998). As shown in Fig. S4, the ELA increase can then lead to a greater decrease in accumulation area at the lower-sloped glacier, potentially exposing a greater area of firn on the surface.

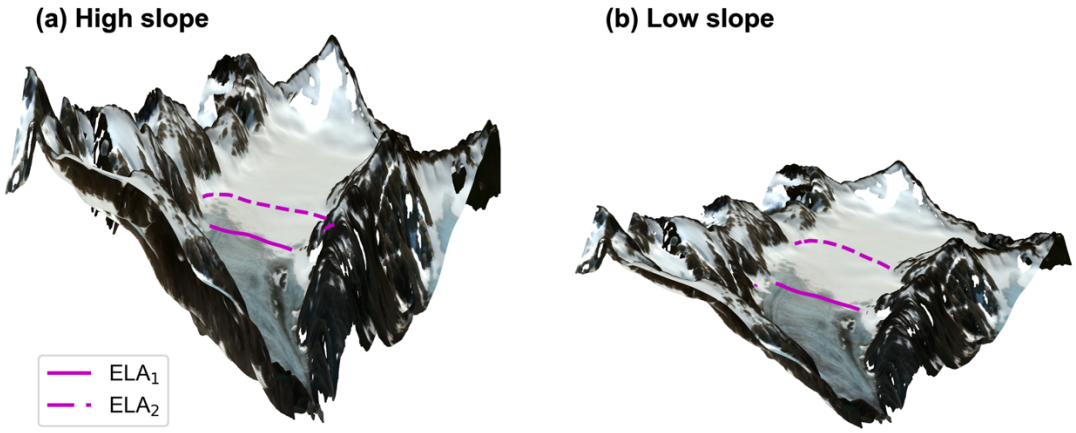

Figure S4: Schematics demonstrating the relative change in equilibrium line altitude (ELA) and its impact on accumulation area for two glaciers with comparatively a) high slope and b) low slope. ELA_1 is set to the observed snowline and ELA_2 is set to 51 m higher. When the ELA (i.e., the zero mass balance contour) increases by 51 m at both glaciers (ELA_2), the accumulation area shrinks more at the low slope glacier. On glaciers that have extensive firn, this ELA increase can also lead to more exposed firn on the surface. Schematics are based on imagery for Lemon Creek Glacier, Alaska, captured by Sentinel-2 on 31 August 2021 and projected onto the most recent digital elevation model from the USGS (McNeil et al., 2022). Elevations have been scaled to demonstrate differing slope gradients.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Geyman, E. C., J. J. van Pelt, W., Maloof, A. C., Aas, H. F., & Kohler, J. (2022). Historical glacier change on Svalbard predicts doubling of mass loss by 2100. *Nature*, *601*(7893), 374–379. https://doi.org/10.1038/s41586-021-04314-4

Lutgens, F. K., & Tarbuck, E. J. (1998). *The Atmosphere: An Introduction to Meteorology*. Prentice Hall.

McNeil, C. J., Sass, L., Florentine, C. E., Baker, E. H., Peitzsch, E. H., Whorton, E. N., et al. (2022). Glacier-Wide Mass Balance and Compiled Data Inputs: USGS Benchmark Glaciers. *Alaska Science Center*. https://doi.org/10.5066/F7HD7SRF

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., & Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment*, *113*(4), 868–879. https://doi.org/10.1016/j.rse.2009.01.001

Viering, T., & Loog, M. (2023). The Shape of Learning Curves: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(6), 7799–7819. https://doi.org/10.1109/TPAMI.2022.3220744

110