

Supplementary A: The explanation of k value selection for KNN

The value of k will affect the final segmentation result after the refining process in our research. There is a similar result when the k value is small, such as 3 (Figure S1b), 5 (Figure S1c), or 10 (Figure S1d) in the plot Y16-2. The k value of 5 has a better visual result, such as the yellow tree in the middle (Figure S1c). On the contrary, when choosing a big k value, such as 50 (Figure S1e) or 150 (Figure S1f), the results of refining are not very correct in some individual trees, such as the red tree in the lower left corner, and the dark green tree in the upper middle (Figure S1e and f).

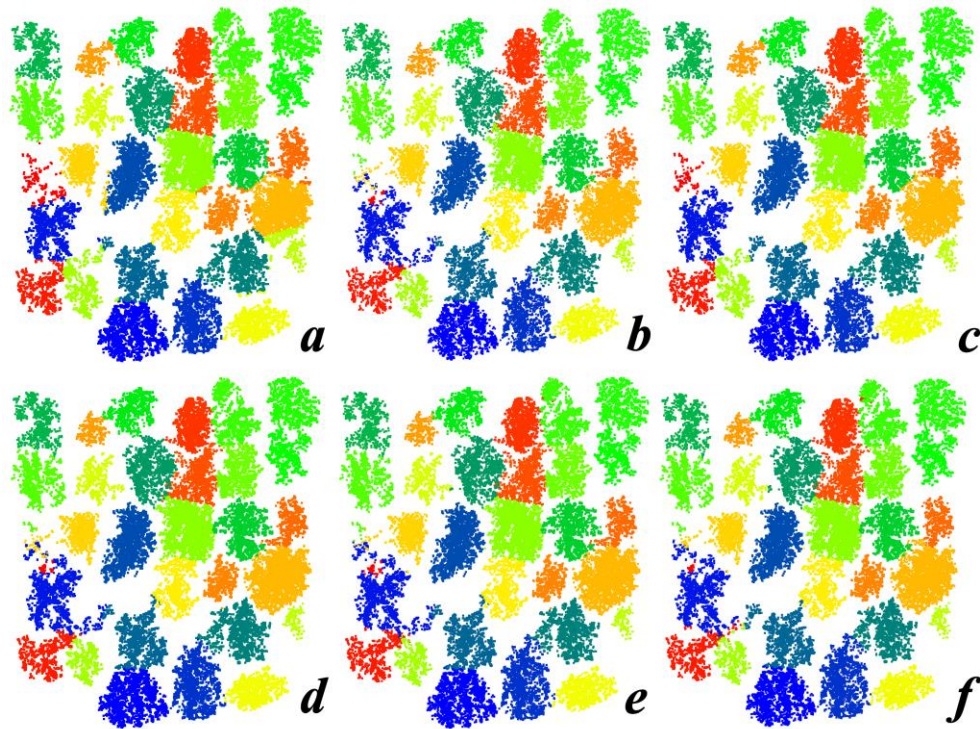


Figure S1. Test of different k values for classification in Y16-2: (a) initial segmentation; final segmentation with $k = 3$ (b), $k = 5$ (c), $k = 10$ (d), $k = 50$ (e), $k = 150$ (f).

The boxplot of two indicators of the process of refining with different k values shows different results (Figure S2). For the result of vertical range, it does not show the difference between different values of k (Figure S2b). While for the result of horizontal area, a larger value of k tends to have a larger horizontal area compared to the results of manual segmentation, which also indicates the incorrectness on the edges, while the mean value of this indicator is very similar when k values are small (Figure S2a).

The reason comes from the point cloud inhomogeneity on vertical structure; in other words, the lower canopy is generally subjected to fewer laser pulses and is therefore sparser (Figure S3). Point clouds with incorrect edge segmentation are often located in the lower sparse point cloud region of the tree, but because the two trees are not the same height, and because of the skew

present in the trees themselves, these point clouds may be close to the canopy of other trees at higher densities nearby. As is shown in Figure S3, the edge point cloud on the left side of the blue tree can be correctly reclassified into labels of the blue tree when taking smaller k values. However, when taking larger values of k (such as 150), this part of the edge point cloud is considered to be closer to the red tree, because the dense upper part of the red tree point cloud is near it. Therefore, a smaller k value (such as between 3 and 10) can help find a more realistic classification result.

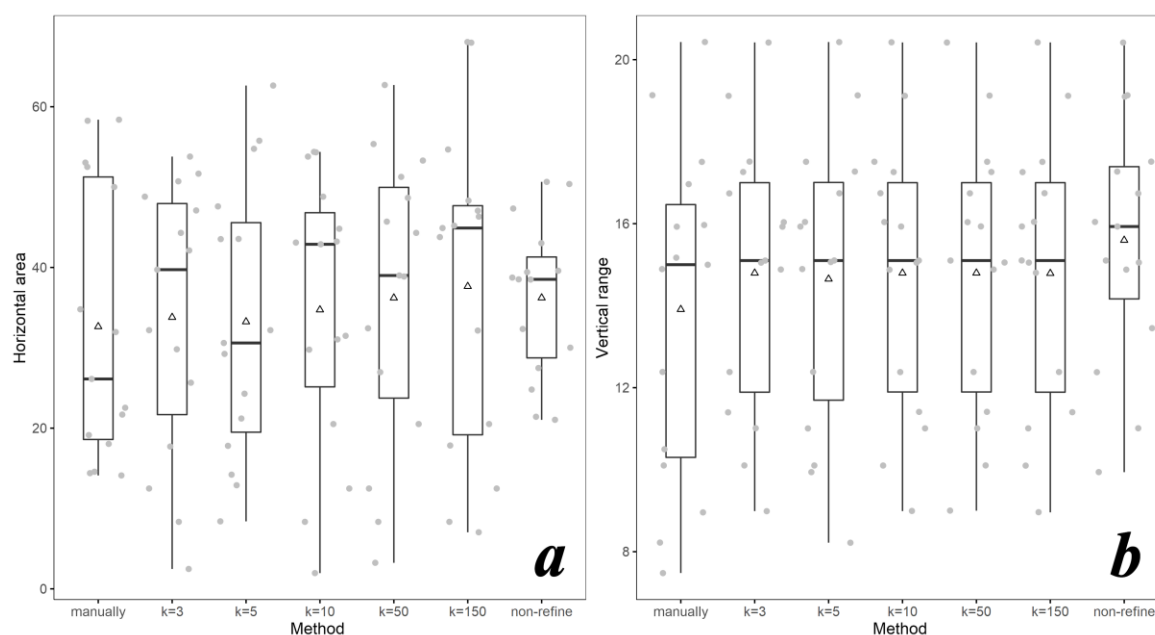


Figure S2. Boxplot of two indicators (horizontal area (a) and Vertical range (b) of segmented individual trees) of the different processing methods (manually, segmentation with the refining process ($k = 3, 5, 10, 50, 150$), and segmentation without the refining process).

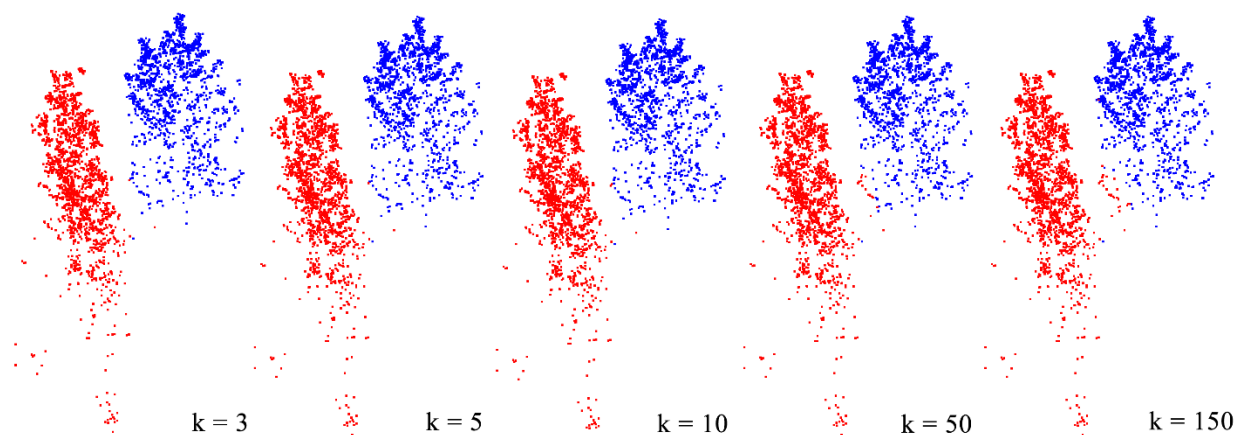


Figure S3. Test of different k values for classification of two individual trees in the process of refining.