Integrating Crowd-sourced Annotations of Tree Crowns using Markov Random Field and Multispectral Information

Qipeng Mei, Janik Steier, Dorota Iwaszczuk*

Remote Sensing and Image Analysis, Technical University of Darmstadt, Germany
(qipeng.mei, janik.steier, dorota.iwaszczuk)@tu-darmstadt.de

Keywords: Crowdsourcing, Markov Random Field, Labels, Crown delineation, Forest mapping.

Abstract

Benefiting from advancements in algorithms and computing capabilities, supervised deep learning models offer significant advantages in accurately mapping individual tree canopy cover, which is a fundamental component of forestry management. In contrast to traditional field measurement methods, deep learning models leveraging remote sensing data circumvent access limitations and are more cost-effective. However, the efficiency of models depends on the accuracy of the tree crown annotations, which are often obtained through manual labeling. The intricate features of the tree crown, characterized by irregular contours, overlapping foliage, and frequent shadowing, pose a challenge for annotators. Therefore, this study explores a novel approach that integrates the annotations of multiple annotators for the same region of interest. It further refines the labels by leveraging information extracted from multi-spectral aerial images. This approach aims to reduce annotation inaccuracies caused by personal preference and bias and obtain a more balanced integrated annotation.

1. Introduction

In the field of forestry research, it is fundamental to accurately map individual tree crowns, i.e., the projected area of a tree crown on a horizontal plane. It plays a significant role as input variable in more accurate analysis, modelling and management, such as the carbon storage estimation, biodiversity assessment, urban forest management, forest fire simulation and forest health description (Zhao et al., 2023; Zhu et al., 2021).

A common way of the individual tree crown delineation is field measurement. Accessed by professional measurement devices, it is able to obtain the common forest mensuration variables, such as the crown length, crown base height, crown diameter, crown radius, crown projected area and crown shape (Zhu et al., 2021). Although we can obtain precise data through field measurements, there are also some challenges associated with it. Large-scale field measurement is a time-consuming process and limited by access issues, such as privately-owned or otherwise inaccessible area (Zhao et al., 2023).

In contrast, with the advent and development of remote sensing technology, and in particular the popularity of unmanned aerial vehicles (UAVs) and multimodal data sources (Röder et al., 2018; Bulatov et al., 2016b; Freudenberg et al., 2022), depicting tree crowns on remotely sensed data is more economical and not constrained by access issues. Benefiting from advancements in algorithms and computing capabilities, in the past decade, supervised deep learning models offer significant advantages in accurately mapping individual tree crown as well as to replace traditional labour-intensive visual interpretation. Taking the widely used convolutional neural network (CNN) model as an example, its unique convolutional layers can extract both surface and abstract features of the image. Then, through a series of sequentially trained hidden layers constructed with a large number of interconnected neurons, it is able to understand images in a manner similar to human cognitive processes and produce reliable results.

For the mapping of the individual tree crowns, deep learning-based instance segmentation (G. Braga et al., 2020) is an suitable approach, which can not only locate the position of the individual tree in an image, but also to depict its tree crown. However, the model's capability stems from both its excellent internal structure and the reliable dataset used for training, which is equally important and cannot be overlooked. Besides quantity, researchers are increasingly emphasizing the importance of quality in training data. Even the most powerful models cannot compensate for the deficiencies of low-quality training data (Oksuz et al., 2020; Whang et al., 2023).

When it comes to the quality of datasets, the quality of annotations cannot be bypassed. Accurately annotating individual tree crowns is fraught with several obstacles that the annotator must face:

1. Characteristics of the tree crown: Tree crowns have irregular contours and overlapping foliage (Stewart et al., 2021).
2. Tree arrangement: Trees are for the most part in the ground without a specific distribution pattern. Especially in dense forests, it is extremely challenging to distinguish a single tree with the naked eye (Freudenberg et al., 2022; Ball et al., 2023).
3. Other ground features: Some features resembling trees in appearance (green belts and lawns,etc.) and shadows bring trouble to the annotator's judgment.
4. Image quality: The limited spatial resolution of the images and varying lighting conditions impede the distinguishability of tree crowns.
5. Issues of the annotator: The patience, fatigue and attitude of the annotator also subjectively affect the quality of the labeling.

Therefore, it is difficult for even the most specialized experts to get the best labels independently.

To overcome these challenges, we introduce a novel approach associated with the information extracted from multi-spectral

---

* Corresponding author

aerial images, to integrate the annotations from multiple annotators, aim to reduce annotation inaccuracies caused by personal preference and bias and obtain a more balanced integrated annotation.

1.1 Previous Work

In instance segmentation, vector data typically serve as carriers in annotation datasets. They store label as geometric shapes (like points, lines, and polygons) and solely the endpoint coordinates rather than the values of individual pixels, which remain constant regardless of scaling. Converting vector annotations to the raster domain for specific numerical operations is one approach for annotation integration. Walter (2018) proposed a method to apply a majority vote in raster domain based on the "Wisdom of the Crowd", assuming that "If many individuals measure the same object, the average geometry should closely approximate the real geometry". Collmar et al. (2023) expanded on this concept to integrate tree crown labels obtained from crowdsourcing through a two-step process. These methods primarily focus on polygon integration for a single object.

Inspired by their work, we have taken a step further to implement tree crown annotation integration in intricate scenarios (Mei et al., 2024). While focusing on individual trees, it also attempted to match pixels to suitable trees in densely forested areas. In contrast to our previous study, this work, we not only consider the mutual constraints of different annotators' perspectives, but also introduce the information provided by remotely sensed data.

More specifically, we use Normalized Difference Vegetation Index (NDVI), a widely used indicator of vegetation health and which can be calculated using commonly available multi-spectral sensors, to support the integration approach. Its inclusion enhances the match between individual pixels and their respective trees. Furthermore, it facilitates clearer differentiation between tree crown boundaries and other surface features, such as man-made structures, which typically exhibit much lower values than vegetation.
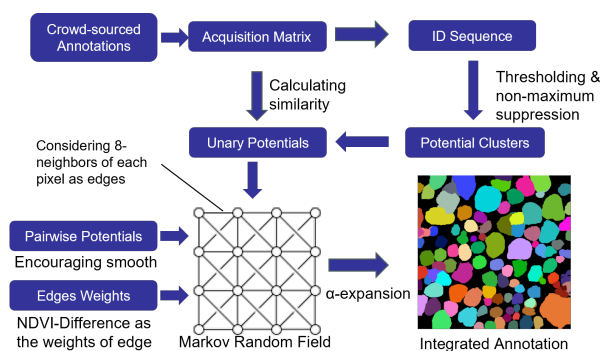
2. Methodology

Figure 1. Workflow of the approach.

Figure 1 depicts the workflow of our approach. In general, it revolves around constructing a Markov Random Field (MRF) and minimizing the associated energy. Therefore, it can be roughly divided into three steps: pre-processing, MRF construction, and energy minimization.

2.1 Pre-processing

2.1.1 Acquisition Matrix In order to take into account the annotations of the same ROI by multiple annotators, we construct a matrix, named the acquisition matrix, for aligning annotations from different annotators. Its shape is $(n + 1, H, W)$, where n represents the number of manual annotations associated with the same ROI, and 1 is dedicated to serve as an expansion layer. The height and width of the image denote as H and W. In this matrix, each of the n layers records the Identifier (ID) of the tree crown label at the corresponding pixel position based on the original annotation. On the expansion layer, we record the frequency f of labelling a pixel as a tree crown. Therefore, we are able to obtain the ID-sequence (IDS) for each pixel, which contains the perspectives of all the annotators.

2.1.2 Potential Clusters Each annotator has its own understanding of the number of trees and the outline of the tree crown in the same ROI. To find a consensus, we define two type of pixel, in addition to the background pixel, for subsequent process, central and marginal ID-sequence (cIDS and mIDS). For them, we have the following assumptions:

1. Pixels in the region around the tree center, where most annotators agree on the tree crown, share the same IDS, referred to cIDS.
2. Pixels at the border of the tree crown may exhibit different IDS due to varying perspectives of one or more annotators, referred to mIDS.
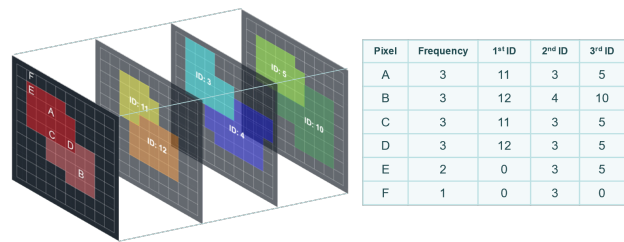

Figure 2. Example of cIDS and mIDS.

Figure 2 is the example of cIDS and mIDS. For a clear explanation, we have selected 6 key pixels. A and B are representatives of the cIDS, surrounded by pixels that share the same IDS. C, D, and E are located at the border of the tree crowns. C has the same IDS as A, indicating that its identity is recognized by all annotators. The IDS of D and E may be slightly different from cIDS due to differing opinions of several annotators and are considered to be mIDS. Moreover, it is certain that their number is much smaller than cIDS. The low f of F suggests that it may stem from bias from a single annotator.

In order to refine the cIDS that most annotators agree on as the anchors for potential clusters, we use the combination of thresholding and non-maximum suppression. The process is as follow:

1. Determine the pixels that most annotators have agreed are part of the tree crowns by retaining pixels with f greater than a certain threshold.
2. Apply non-maximum suppression based on the number of IDS to eliminate redundancies that may point to the same tree crown, which attend to preserve cIDS while eliminate mIDS.

The final refined potential clusters obtained represent the agreed-upon trees within the corresponding ROI.

2.1.3 Normalized Difference Vegetation Index

In addition to subjective human annotation, we introduce objective spectral information to support the integration. The Normalized Vegetation Index (NDVI) is a widely used indicator of vegetation health that ranges from -1 to 1 and is calculated using the following formula:

$$NDVI = \frac{NIR - RED}{NIR + RED} \qquad (1)$$

where
$NDVI$ = Normalized Difference Vegetation Index
NIR = spectral radiance in near-infrared
RED = spectral radiance in red (visible)

Overall, it is negative for water bodies, close to zero for rocks, sands, or concrete surfaces, and positive for vegetation as well as positively correlated with the vitality of the vegetation(Jones and Vaughan, 2010).
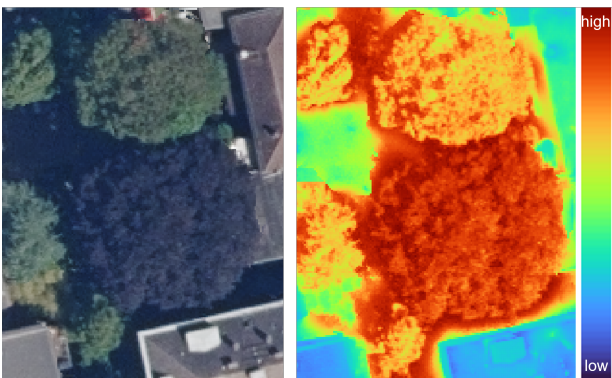

Figure 3. Example of Normalized Vegetation Index.

Figure 3 is the example of the NDVI among various surface objects. The difference in NDVI between vegetation and man-made structures is significant, with boundaries being more visually distinct than in regular RGB images. Additionally, the variations in characteristics between trees are evident in the NDVI. However, within individual trees, the fluctuations in NDVI are relatively minor.

Therefore, we assume that the introduction of inter-pixel NDVI differences enables the following refinement of the labeling.

1. The NDVI difference between vegetated and non-vegetated pixels makes the separation more pronounced.
2. The NDVI might have variations from tree to tree due to the health status or characteristics.

2.2 Markov Random Field

The MRF is an undirected probabilistic graphical model used to model the relationship between random variables. It compose of a set of nodes and edges connecting these nodes, where nodes represent random variables, and edges represent the dependency between variables (Geman and Graffigne, 1986). Its structure is suitable for characterizing images and implementing specific tasks such as depth map generation (Bulatov et al., 2016a). Qiu et al. (2022) employed MRF as the post-processing the deep learning model for car detection and utilized elevation as a constraint, which demonstrates the potential of MRF in incorporating support data to optimize labels.

Here, we establish a MRF to represent the annotation in a ROI. Each pixel (i) in the annotation is considered as a node and connected by edges with its eight neighbors' nodes (j). Their set is denoted by $\mathcal{N}$ resp. $j \in \mathcal{N}(i)$.

2.2.1 Unary Potentials

The unary potentials refer to the costs of assigning nodes to potential clusters, mentioned in Section 2.1.2, provided to the MRF. It focuses solely on the current node without considering the state of other nodes. In order to compute the unitary potential of each node over potential clusters c, we first compute the similarity of their IDS to cIDS in response to the possibility of them being assigned to each of the potential clusters $P(c(i))$.

$$P(c(i)) = \frac{n_s}{n_t} \qquad (2)$$

Here,
n_s = number of items in IDS identical to cIDS
n_t = number of the total items in IDS

Then, we convert this possibility into the unary potentials for each node with the standard negative logarithm trick to ensure that a low possibility results in a high unary cost:

$$E_u(c(i)) = -\log_2 P(c(i)), \qquad (3)$$

where $E_u(c(i))$) denote the unary potentials.

2.2.2 Pairwise Potentials

The pairwise potentials take into account the cost caused by the adjacency relationship, which make the assignment of nodes subject to neighboring nodes. The NDVI difference $\Delta_N(i,j)$ between each node and its neighboring nodes is converted into edge weight W_{ij} by the following formula:

$$W_{ij} = \lambda \cdot \exp\left(-\frac{\Delta_N(i,j)}{\omega}\right), \qquad (4)$$

where
ω = the sensitivity to $\Delta_N(i,j)$
λ = the gain of the weights

Hence, the pairwise potentials are defined as:

$$E_p(c(i), c(j)) = W_{ij} \cdot \begin{cases} 0, \text{ if } c(i) = c(j) \\ 1, \text{ if } c(i) \neq c(j) \end{cases} \qquad (5)$$

In general, pairwise potential encourages the smoothing. When neighboring nodes are classified under the same cluster, it does not incur any cost for the MRF. However, unary pontetials enable to constrain excessive smoothing and incur substantial costs with sensible parameter configurations.

Furthermore, the $\Delta_N(i,j)$ are able to optimize the assignment. Due to the W_{ij}, in the case of $c(i) \neq c(j)$, lower $\Delta_N(i,j)$ result in higher W_{ij}. Conversely, the pairwise potentials caused by high $\Delta_N(i,j)$ are acceptable. Therefore, our MRF tends to keep neighboring pixels with similar NDVI in the same cluster, while allowing pixels with significant NDVI differences to belong to different clusters.

2.3 Energy Minimization

The total cost of the MRF caused by unary and pairwise potentials is expressed as the following energy equation.

$$E(c) = \sum_i \left[E_u(c(i)) + \sum_{j \in N(i)} E_p(c(i), c(j)) \right] \quad (6)$$

The clusters distribution that minimizes this energy equation is the most balanced integrated annotation. It is generated by the mutual constraints of the different annotators and controlled by objective spectral information. However, directly solving for the minimum energy of the MRF is an NP-hard problem. We use a graph cuts algorithm, alpha-expansion moves, to achieve the approximate energy minimization (Boykov et al., 2001).

2.4 Quantitative Metrics

Quantitative evaluation of annotations is a challenging issue. Manual annotation is regarded as the ground truth in the training of deep learning models. Hence, appropriate reference data of annotation is crucial.

Ball et al. (2023) and Collmar et al. (2023) used the labels delineated by domain-specific experts as the ground truth for evaluation. Freudenberg et al. (2022) introduced LiDAR data to assist in generating reference labels.

In our study, we use the tree cadaster[1] as the reference data, which maintained by the parks department of the City of Frankfurt am Main and collected in the field, has been accessible online to the public since 2014, with the latest version updated in August 2023. It contains almost all trees in public areas of Frankfurt am Main, stored as vector data. The tree center of each tree is recorded, as well as the tree cover stored as a circular vector.

We have defined the following metrics for quantifying the information we are interested in:

1. Overall Intersection over Union (IoU): The IoU between annotation and tree's coverage from the cadastre.
2. Tree's center coverage (TCC): The proportion of recorded tree's centers that are covered in the annotation.
3. Single tree label (STL): The ratio of the number of labels that cover only a single tree center to the total number of labels in the annotation.

It is worth noting that for each ROI, we finally performed maximum absolute scaling on the metrics, aiming to evaluate the relative gap between annotations (manual and integrated) to the optimal values.

3. Experimentation

We selected eight locations within Frankfurt am Main as ROIs, which cover cemeteries, streets, squares, and backyards within the city area (see Fiegure 4). The trees within these areas display a variety of distributions, ranging from haphazard arrangements akin to those found in cemeteries to organized layouts resembling those along streets and in squares. Each aerial images
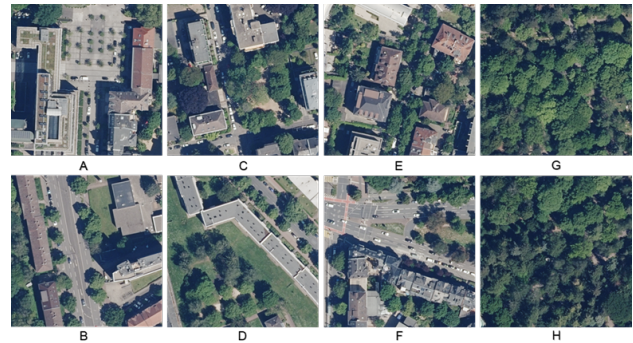
---

[1] https://www.offenedaten.frankfurt.de/dataset/baumkataster-frankfurt-am-main

Figure 4. Experimental areas.

is in size of 512×512 pixels with a with a Ground Sampling Distance (GSD) of 20 $cm/pixel$ and consist of R, G, B, NIR channels. For each ROI, four experts independently annotated the tree crowns on the true color image.

	A	B	C	D	E	F	G	H
Annotator 1	48	19	45	30	50	34	139	127
Annotator 2	37	18	45	32	76	39	107	99
Annotator 3	44	18	69	26	65	40	97	96
Annotator 4	43	19	39	33	61	34	75	119

Table 1. Number of labels in the region of interest.

Table 1 shows the number of labels made by each annotator for the corresponding ROI. In uniform tree distribution areas, label counts are almost consistent, but in irregularly distributed areas, label counts vary significantly.
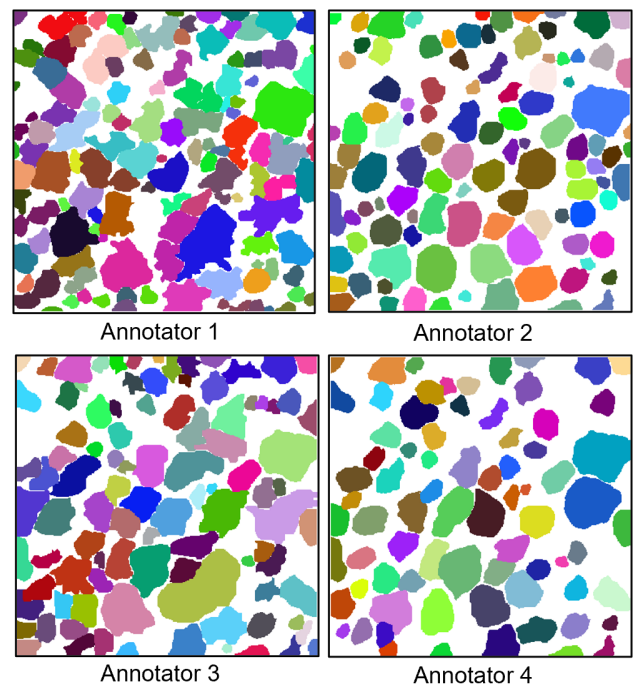

Figure 5. Annotation example of ROI G.

In addition, in complex areas such as forests, differences in labelling between different annotators are evident (see Figure 5). It is reflected in the depiction of outlines and the identification of tree locations. Therefore, to account for the possibility of trees being overlooked due to the negligence of a particular annotator, we set the threshold at 75%. That means, in our ex-

periment, pixels identified as representing trees by at least 3 individuals will be considered for further processing.

In non-maximum suppression, we retain the IDS with higher counts as potential clusters and remove those with fewer counts, which have at least 50% of the items in common with the potential clusters.

4. Results

For manual and integrated annotation, we conduct quantitative and qualitative evaluations to evaluate the quality of annotation from different perspectives.

4.1 Quantitative Evaluation
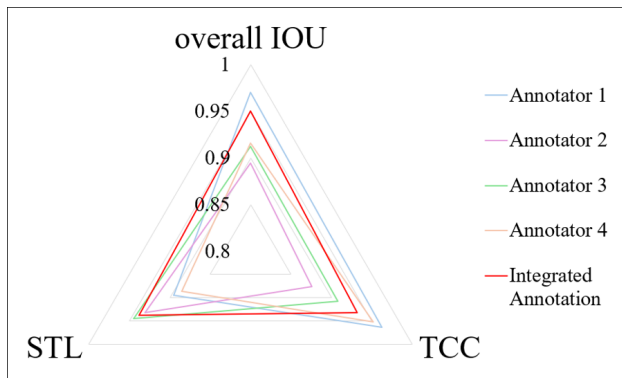
Figure 6. Annotation preference.

Figure 6 shows the metrics distribution over difference annotations. It is evident that manual annotations often exhibit certain biases. For example, annotator 1 tends to delineate more accurate tree crown boundaries and strives to cover all trees, but it's prone to lumping multiple trees into one label. Conversely, a strict definition of a single tree may result in their inability to cover all trees and to delineate more matching boundaries. Such as the results of annotator 2 and 3.

Although the integrated annotation is not optimal in all metrics, it is the most balanced. It is able to approach the optimal value in each metric under the mutual constraints of different annotators' perspectives and the support of additional remote sensing information.

Compared with the average results of manual annotation, the integrated annotation is significantly better (see Figure 7). This further illustrates that the contributions of different annotators can yield a relatively reliable annotation while suppressing bias and personal preference. Furthermore, compared to our previous approach (Mei et al., 2024), due to the additional information provided by NDVI, the improvement in overall IoU and STL is noteworthy, which proves our assumptions in Section 2.1.3. In TCC, due to its primary dependence on the acquisition matrix, attention to the ROI is more global, hence, there is no significant enhancement. However, the TCC of integrated (with and without NDVI) annotations generally outperform the of manual annotations, once again demonstrating the effectiveness of our approach.

4.2 Qualitative Evaluation

In addition to quantitatively analyzing the quality of annotation, we perform subjective observations on the annotations.
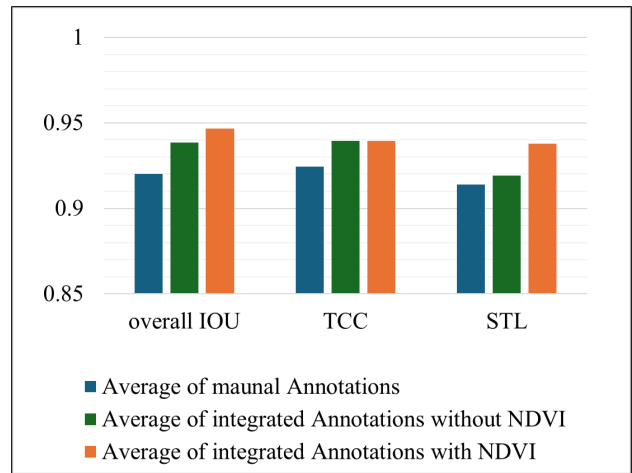

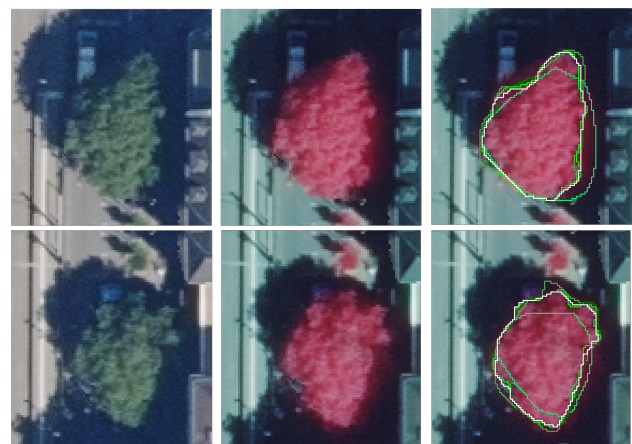Figure 7. Comparing integrated with average manual annotation.

Figure 8. Integrated (white) and manual (green) annotations on the single tree (left:true-color image, middle: near-infrared pseudo color image, right: annotations).

Firstly, when it comes to individual tree labels, the integrated annotations are satisfying and visually more accurate (see Figure 8). Influenced by multiple factors such as image quality, surface similarity (such as shadows resembling tree colors), annotator's patience and fatigue, etc., manual annotation may include pixels that do not belong to the tree crown or overlook some pixels.

Our integrated annotation benefits from the mutual constraints of different annotators and the simple yet effective support of NDVI differences, enabling the production of a label that closely fits the outline of the tree crown.

However, when the tree crown is too small and all annotators attempt to depict it with labels larger than the tree crown, the suppressive effect of NDVI is limited. As shown in Figure 9, even the shadows are still included in the labels.

In the case of trees intersecting with each other, our approach also can find a suitable boundary for the tree crowns (see Figure 10). Moreover, pixels located at the margins are able to be more inclined to the corresponding trees as the result of the heath status or characteristics of the tree.

In conclusion, the combination of the mutual constraints of different annotators and the NDVI Difference enables the gener-
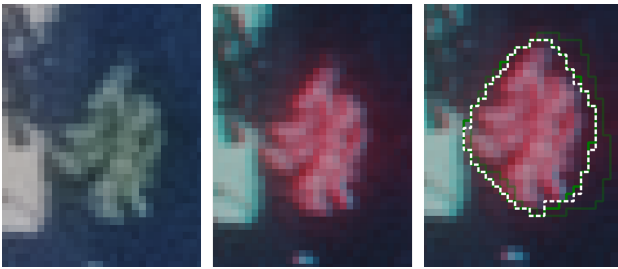
Figure 9. Integrated (white) and manual (green) annotations on the single tree, limitation (left:true-color image, middle: near-infrared pseudo color image, right: annotations).
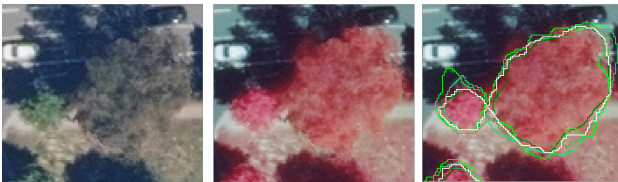

Figure 10. Integrated (white) and manual (green) annotations on the intersecting trees (left:true-color image, middle: near-infrared pseudo color image, right: annotations).

ation of an integrated annotation that is more representative of the actual situation. And favorable outcomes are also observed in regions where multiple trees intersect with each other.

5. Conclusion

In this study, we introduce a novel integrated approach of crowd-sourced tree crown annotations. It takes into account the subjective cognition of the annotator and is supported by objective multi-spectral aerial images to obtain a balanced integrated annotation.

More specifically, the annotations from difference annotators are aligned by an acquisition matrix we propose, and undergo specific numerical processing (non-maximal suppression and thresholding) to obtain the basic consensus by the tree crown annotation of different annotators on the same region of interest. Building upon this foundation, we employed a probabilistic graphical model, namely a Markov random field, supplemented by objective indicator, the Normalized Vegetation Index, to achieve mutual constraint among different annotators' annotations by energy minimization.

Experimentation in various regions of Frankfurt am Main demonstrates its potential. In the quantitative evaluation, our approach is shown to balance the preferences and biases of different annotators and outperform the average annotator performance on each metric. In the qualitative evaluation, the integrated annotations are visually and subjectively more realistic. The introduction of objective data enable to partially rectify human judgment errors, yet its efficacy remains constrained in the face of inaccuracies upheld by the majority of annotators. It is worth noting that our approach performs suitable in scenarios with trees arranged individually as well as in densely. Therefore, it can be applied to the integration of annotations in large-scale scenarios. Overall, the integrated annotations obtained through the mutual constraints of multiple annotators are satisfactory.

In addition, our approach has strong generality. The acquisition matrix is a general component that can be adapt to the annota-

tions of any object with complex outline, such as human, water body. Pairwise potentials and edge weights of the Markov random field can be adjusted according to specific tasks. For example, thermal images for the human as well as Normalized Difference Water Index for water bodies.

In our future work, we will assess the efficacy and extent of enhancement that this approach brings to deep learning models. Meanwhile, it also has the potential to become a post-processing part for deep learning model. It has the possibility of integrating multi-modal prediction results as well as combining it with a general segmentation large model,such as Segment Anything Model (Kirillov et al., 2023), to improve specific segmentation tasks.

References

Ball, J. G., Hickman, S. H., Jackson, T. D., Koay, X. J., Hirst, J., Jay, W., Archer, M., Aubry-Kientz, M., Vincent, G., Coomes, D. A., 2023. Accurate delineation of individual tree crowns in tropical forests from aerial RGB imagery using Mask R-CNN. *Remote Sensing in Ecology and Conservation*, 9(5), 641–655.

Boykov, Y., Veksler, O., Zabih, R., 2001. Fast approximate energy minimization via graph cuts. *IEEE Transactions on pattern analysis and machine intelligence*, 23(11), 1222–1239.

Bulatov, D., Kottler, B., Rottensteiner, F., 2016a. Energy minimization of discrete functions with higher-order potentials for depth map generation. *2016 23rd International Conference on Pattern Recognition (ICPR)*, IEEE, 2344–2349.

Bulatov, D., Wayand, I., Schilling, H., 2016b. Automatic tree-crown detection in challenging scenarios. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B3, 575–582.

Collmar, D., Walter, V., Kölle, M., Sörgel, U., 2023. From Multiple Polygons to Single Geometry: Optimization of Polygon Integration for Crowdsourced Data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 10, 159–166.

Freudenberg, M., Magdon, P., Nölke, N., 2022. Individual tree crown delineation in high-resolution remote sensing images based on U-Net. *Neural Computing and Applications*, 34(24), 22197–22207.

G. Braga, J. R., Peripato, V., Dalagnol, R., P. Ferreira, M., Tarabalka, Y., OC Aragão, L. E., F. de Campos Velho, H., Shiguemori, E. H., Wagner, F. H., 2020. Tree crown delineation algorithm based on a convolutional neural network. *Remote Sensing*, 12(8), 1288.

Geman, S., Graffigne, C., 1986. Markov random field image models and their applications to computer vision. *Proceedings of the international congress of mathematicians*, 1, Berkeley, CA, 2.

Jones, H. G., Vaughan, R. A., 2010. *Remote sensing of vegetation: principles, techniques, and applications*. Oxford University Press, USA.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment Anything. *arXiv:2304.02643*.

Mei, Q., Steier, J., Iwaszczuk, D., 2024. Integration Approach for Manual Generated Single Tree Crown Annotations. *Band 32: 44. Wissenschaftlich-Technische Jahrestagung in Remagen 2024*, 141–150.

Oksuz, K., Cam, B. C., Kalkan, S., Akbas, E., 2020. Imbalance problems in object detection: A review. *IEEE transactions on pattern analysis and machine intelligence*, 43(10), 3388–3415.

Qiu, K., Bulatov, D., Lucks, L., 2022. Improving car detection from aerial footage with elevation information and markov random fields. *SIGMAP*, 112–119.

Röder, M., Latifi, H., Hill, S., Wild, J., Svoboda, M., Brůna, J., Macek, M., Nováková, M. H., Gülch, E., Heurich, M., 2018. Application of optical unmanned aerial vehicle-based imagery for the inventory of natural regeneration and standing deadwood in post-disturbed spruce forests. *International Journal of Remote Sensing*, 39(15-16), 5288–5309.

Stewart, D., Zare, A., Marconi, S., Weinstein, B. G., White, E. P., Graves, S. J., Bohlman, S. A., Singh, A., 2021. Randcrowns: A quantitative metric for imprecisely labeled tree crown delineation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 11229–11239.

Walter, V., 2018. Integration of multiple collected polygons with a raster-based approach. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 679–685.

Whang, S. E., Roh, Y., Song, H., Lee, J.-G., 2023. Data collection and quality challenges in deep learning: A data-centric ai perspective. *The VLDB Journal*, 32(4), 791–813.

Zhao, H., Morgenroth, J., Pearse, G., Schindler, J., 2023. A systematic review of individual tree crown detection and delineation with convolutional neural networks (CNN). *Current Forestry Reports*, 9(3), 149–170.

Zhu, Z., Kleinn, C., Nölke, N., 2021. Assessing tree crown volume—A review. *Forestry: An International Journal of Forest Research*, 94(1), 18–35.