Toward an Automated Pruning for Apple Trees Based on Computer Vision Techniques

The 16th International Conference on Advances in Computer-Human Interactions

K. Hu and O.D.A. Prima
Graduate School of Software and Information Science, Iwate Prefectural University
xxx@iwate-pu.ac.jp

Introduction

Fuji Apples in Japan: High quality, high price

However, Labor shortage challenges in Japan's agriculture

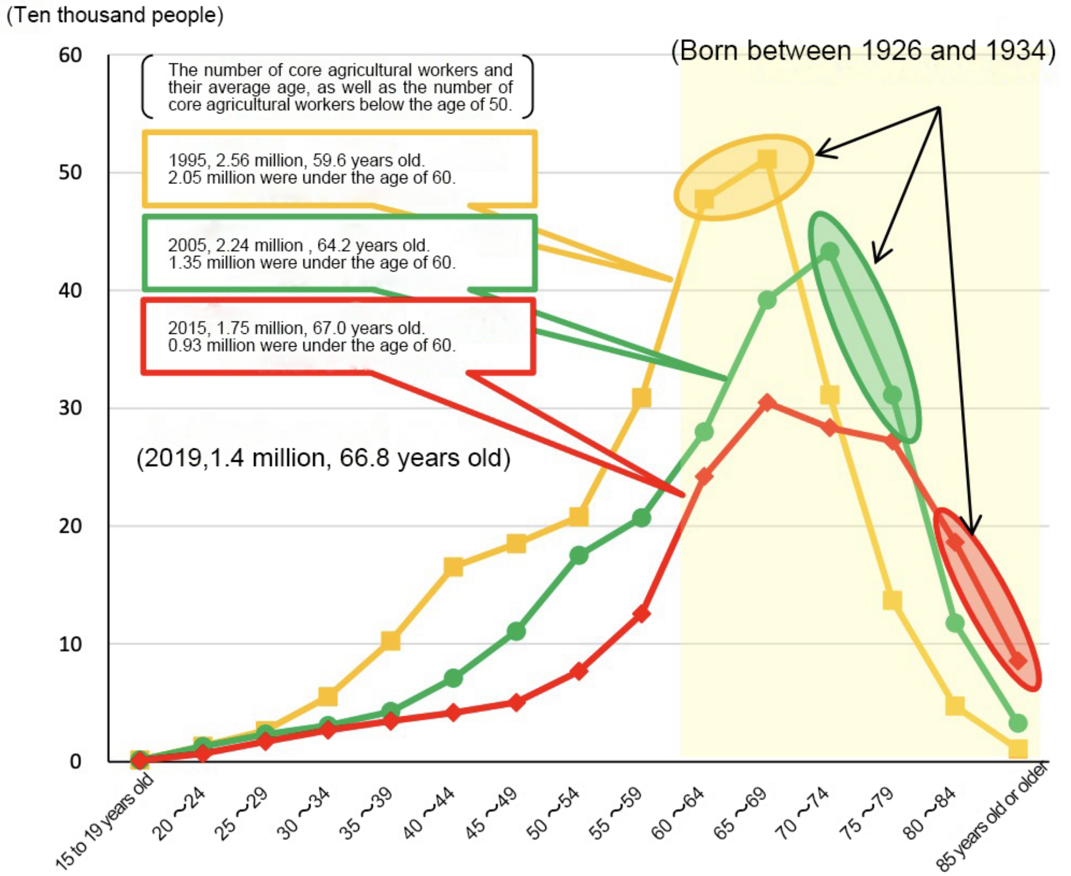
Agricultural automation technology: Solving labor issues

Pruning: Key aspect in agriculture, not yet fully automated

The benefits of proper pruning:

- Stimulate the growth of new buds and shoots
- Facilitate management and harvesting.
- Promote fruit growth
- Increase fruit yield

Fuji apples



Ministry of Agriculture, Forestry and Fisheries "Agriculture and Forestry Census" (Aggregate Compilation), "Agricultural Structural Dynamics Survey" 2022.

Introduction

Traditional pruning rules, such as:

- Ventilation, light penetration, and transparency of the branches.
- The branches should be smaller than the trunk.
- The lowest branch should be 2 to 3 feet above the ground, etc.

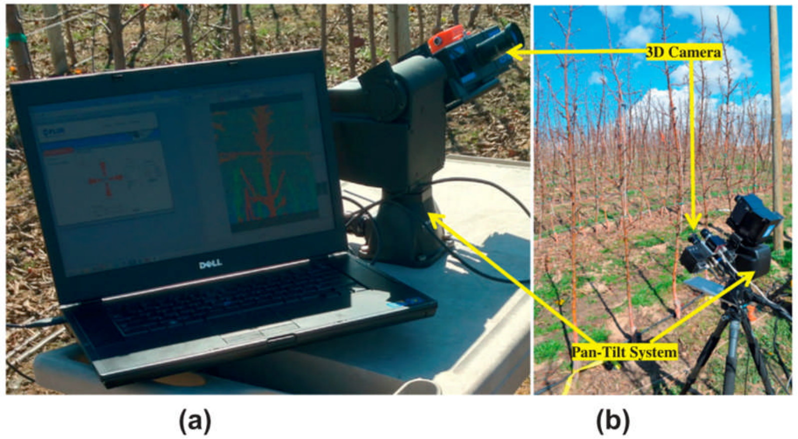
Therefore, many studies are focused on how to obtain the structure of apple trees in order to achieve pruning automation.

Iwate Prefecture Hanamaki City Apple Orchard

Related work

- M. Karkee et al. (2014) proposed a method for constructing 3D models of apple trees using Time-of-Flight (ToF) 3D cameras and implementing automated pruning of tall spindle apple trees through optimized algorithms.

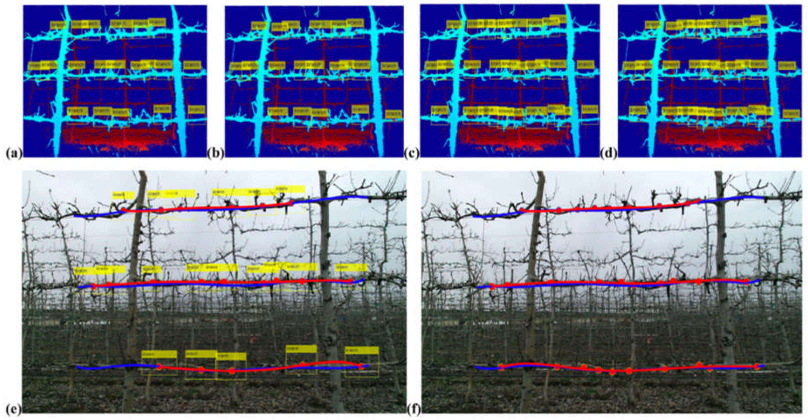
Extracting fine tree branches with ToF 3D cameras is complex.



2014 - Identification of pruning branches in tall spindle apple trees for automated pruning.

- J. Zhang et al. (2018) proposed a branch detection method for apple trees with fruiting wall architecture using depth features and Regions-Convolutional Neural Network (R-CNN) based on pseudo-color images.

Specific environmental limitations.



2018 - Branch detection for apple trees trained in fruiting wall architecture using depth features and Regions-Convolutional Neural Network (R-CNN).

Objective

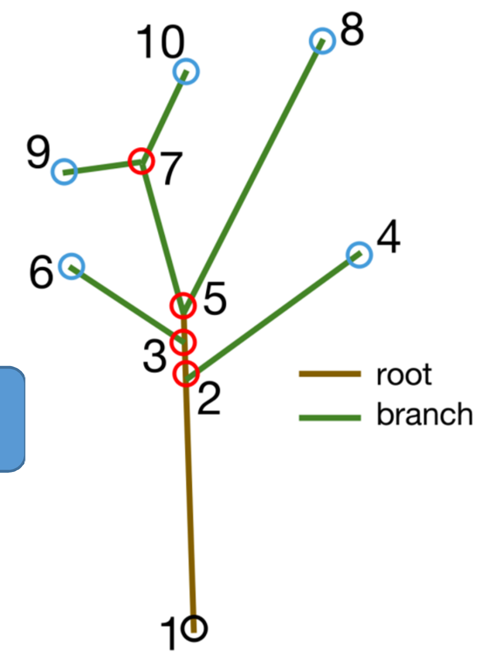
Our research proposes a new computer vision method based solely on RGB cameras, aiming to address two critical issues:

- Difficulty in detecting slender branches.
- Specific environmental limitations.

Materials and Methods

Extracting tree structure from an RGB image

Extract → Skeletonize → Structure



- Extracting the branches and trunk sections.
- Skeletonized extraction of images (feature extraction).
- Exploring the topological structure of apple trees.

Dataset

Selected 9 Trees from Hanamaki Orchard Japan

Creating annotations using Labelme tool.

Dataset information:

- Image size : 1920 x 1080
- Train images: 143
- Validate images: 20

Original image

Annotation image

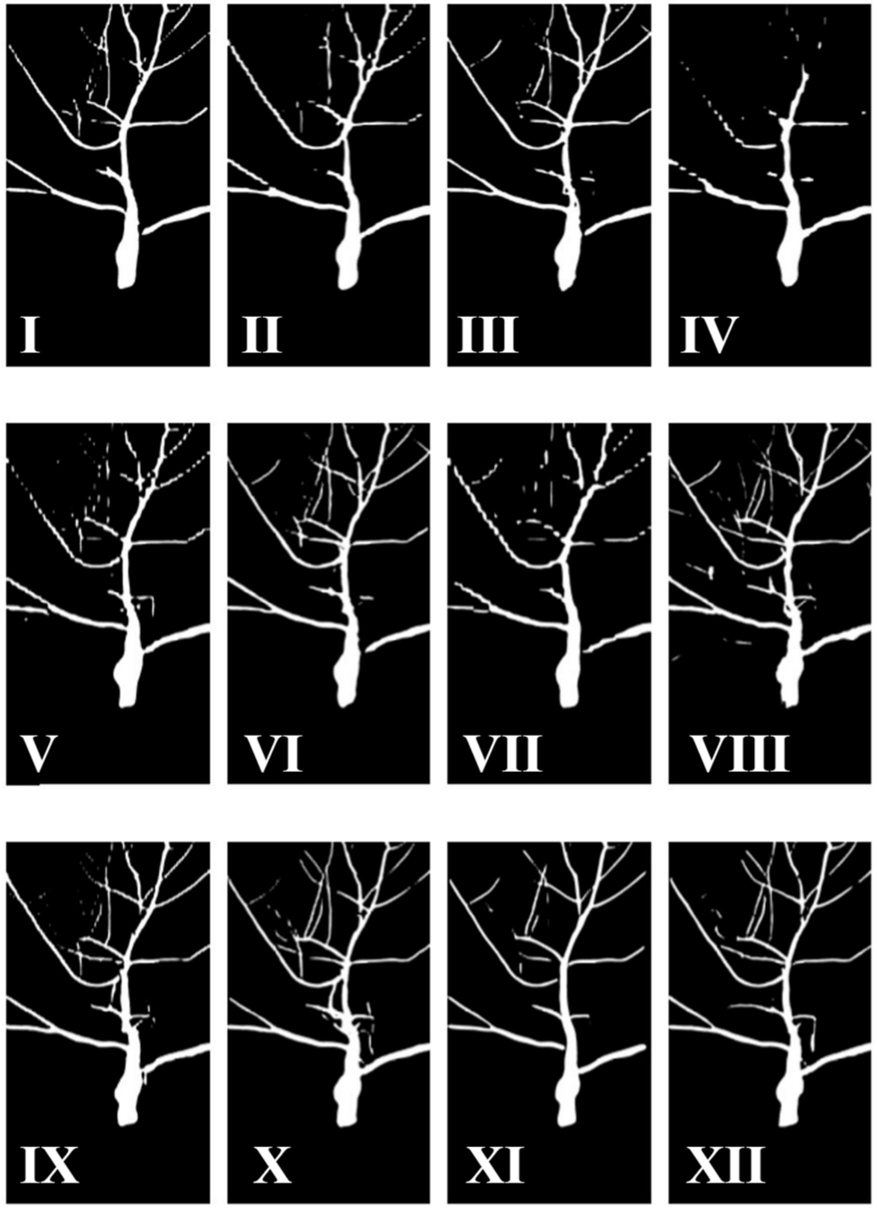
Experiment

- Tree extraction from RGB images using semantic segmentation models

ID	Name	Backbone	Input Size	Batch Size	Iterations	Acc	IoU
I	DeepLabV3+	R50	480x480	6	10000	65.26	54.48
II	APCNet	R50	512x512	6	10000	67.09	58.39
III	Semantic+FPN	R50	512x512	16	5000	72.1	62.35
IV	PSPNet	R50	480x480	8	3500	53.29	46.44
V	CCNet	R50-D8	512x512	6	10000	68.98	59.19
VI	UPerNet	R50	512x512	8	5000	69.99	60.44
VII	FCN	R101	512x512	3	10000	64.48	55.27
VIII	UNet+FCN	UNet-S5-D16	64x64	128	1000	73.5	62.29
IX	CGNet	M3N21	680x680	12	10000	74.87	65.13
X	UPerNet	Swin-S	512x512	8	30000	76.98	64.24
XI	UPerNet	ViT-B+LN+MLN	512x512	4	60000	56.89	43.27
XII	SegFormer [19]	MIT-B5	512x512	4	60000	76.72	64.29

*) Each model was trained using the OpenMMlab framework.

The UPerNet and SegFormer models were found to be superior in terms of accuracy and IoU.



Experiment

❑ Tree Branch Skeletonization

- STEP 1: Apply morphology dilation to the image resulting from segmentation to connect non-contiguous branches.
- STEP 2: Apply thinning to the image using skeletonization method.

Segmentation result After STEP 1 After STEP 2

Experiment

Structuring Tree Branches

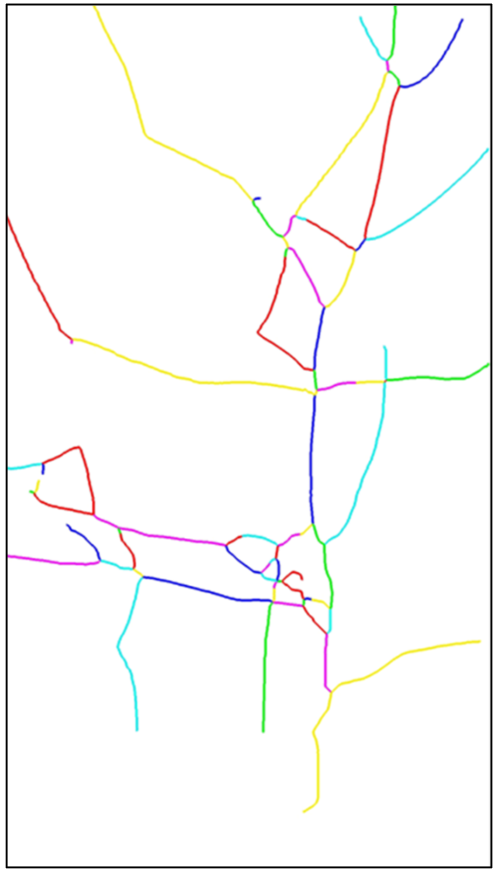
Breadth-first Search for searching the tree data structure

```
Input: Masked image (M)
Output: The topological structure of an apple tree
Function findConnection(M):
    process queue ← [root point];
    node link ← [root point];
    while process queue  lenght > 0 do
        now point ← process  queue.pop(0);
        for each in now point  arround(M) do
            node link[now].nextAdd(each);
            process queue.add(each);
        end
    end
    return node link;
```

Result

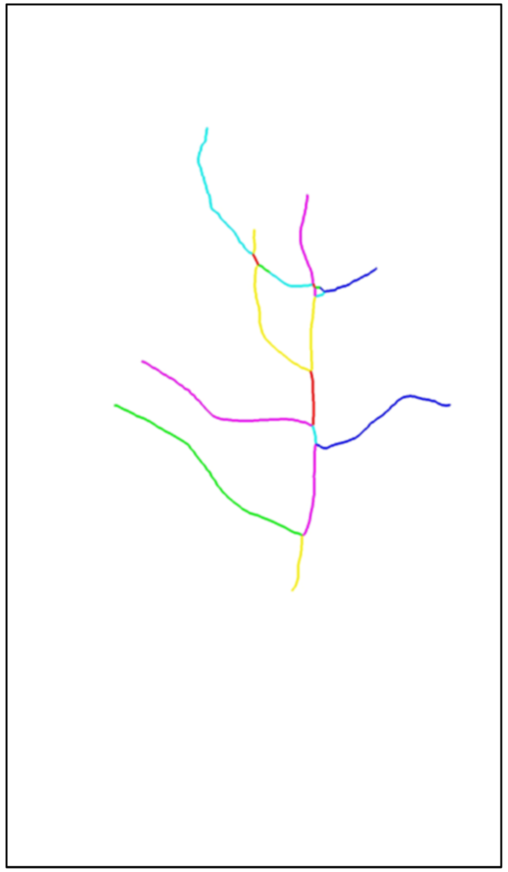
- Automated apple tree topological structure extraction results

Input (RGB)



Topological Structure

Input (RGB)

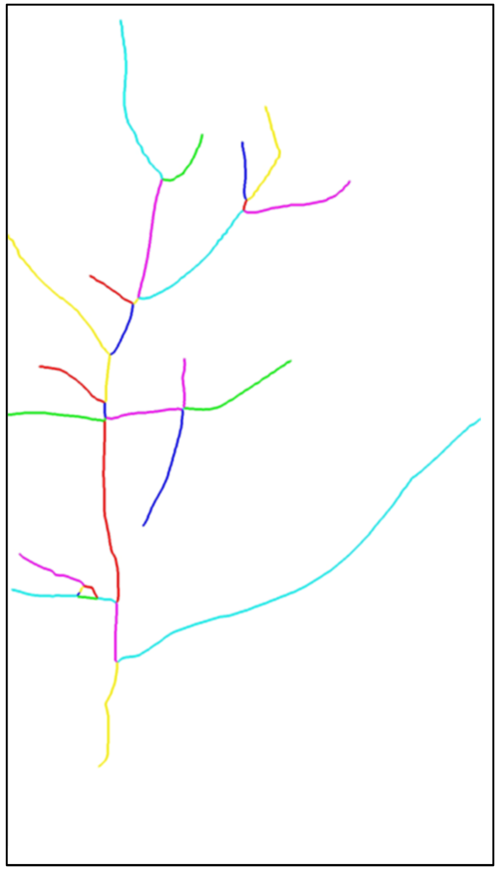


Topological Structure

Result

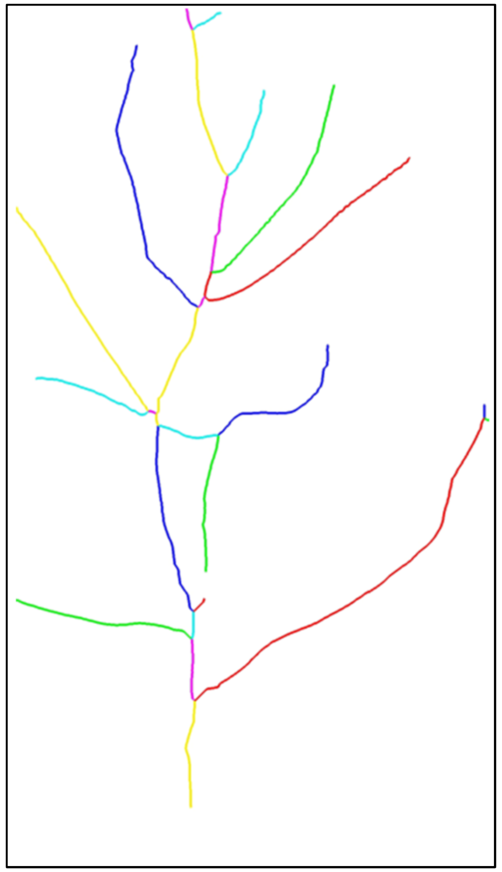
❑ Automated apple tree topological structure extraction results

Input (RGB)



Topological Structure

Input (RGB)



Topological Structure

Conclusion and Future work

Conclusion

- Successfully extracted apple branches and topological structure.
- The SegFormer model was found to be the most effective segmentation model.

Future work :)

- Develop automatic pruning based on the topological structure.
- Compare the apple tree images before and after pruning.
- Collect more image datasets.
- Modify the hyperparameters of the neural network.

Thank you for your attention

Reference

[1] M. Karkee et al., “Identification of pruning branches in tall spindle apple trees for automated pruning,” Computers and Electronics in Agriculture, 103, pp. 127–135, 2014.

[2] Y. Majeed et al., “Apple Tree Trunk and Branch Segmentation for Automatic Trellis Training Using Convolutional Neural Network Based Semantic Segmentation,” IFAC PapersOnLine, 51-17, pp. 75–80, 2018.

[3] J. Zhang et al., “Branch detection for apple trees trained in fruiting wall architecture using depth features and Regions-Convolutional Neural Network (R-CNN),” Computers and Electronics in Agriculture, 155, pp. 386–393, 2018.

[4] D. Fumey, P. Lauri, Y. Guedon, C. Codin, and E. Costes, “Effects of ´ Pruning on the Apple Tree: from Tree Architecture to Modeling,” Proc. IXth IS on Orchard Systems Ed.: T.L. Robinson Acta Hort. 903, ISHS, 2011.

[5] Labelme, https://github.com/wkentaro/labelme [retrieved at April 4, 2023]

[6] OpenMMLab, https://github.com/open-mmlab [retrieved at April 4, 2023]

[7] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2016.

[8] A. Kirillov, R. Girshick, K. He, and P. Dollar, “Panoptic Feature Pyramid ´ Networks,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp. 6392-6401, 2019.

[9] J. He, Z. Deng, L. Zhou, Y. Wang, and Y. Qiao, “Adaptive Pyramid Context Network for Semantic Segmentation,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp. 7511-7520, 2019.

[10] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid Scene Parsing Network,” IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 6230-6239, 2017.

[11] T. Xiao, Y. Liu, B. Zhou, Y. Jiang, and J. Sun, “Unified Perceptual Parsing for Scene Understanding,” European Conference on Computer Vision, pp. 418-434, 2018.

[12] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, pp. 3431-3440, 2015.

[13] Z. Huang et al., “CCnet: Criss-cross attention for semantic segmentation,” Proceedings of the IEEE/CVF international conference on computer vision, 2019.

[14] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convo- lutional networks for biomedical image segmentation,” Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015, 2015.

[15] T. Wu, S. Tang, R. Zhang, J. Chao, and Y. Zhang, “CGNet: A LightWeight Context Guided Network for Semantic Segmentation,” in IEEE Transactions on Image Processing, 30, pp. 1169-1179, 2020.

[16] Z. Liu et al., “Swin transformer: Hierarchical vision transformer using shifted windows,” Proceedings of the IEEE/CVF international conference on computer vision, 2021.

[17] D. Alexey et al., “An image is worth 16x16 words: Transformers for image recognition at scale,” ILCR Conference, pp. 1-21, 2021.

[18] E. Xie et al., “SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers,” Neural Information Processing Systems, 2021.