Assessing Fine Root Production in Terrestrial Forests: A Comparative Analysis of AI and Human Annotation Using Minirhizotron Images

Imogen Carter*[1,2], Grace Handy*[1,2], Rob Mackenzie[1,2], Marie Arnaud[1,2,3], Gael Denny[1,2], Abraham Smith[4], and Adriane Esquivel-Muelbert[1,2]

1: Birmingham Institute of Forest Research, University of Birmingham, Birmingham, United Kingdom of Great Britain – England, Scotland, Wales
2: School of Geography Earth and Environmental Sciences, University of Birmingham, Birmingham, United Kingdom of Great Britain – England, Scotland, Wales
3: Institute of Ecology and Environmental Sciences Paris (iEES-PARIS), Sorbonne University, Paris, France
4: Department of Computer Science, University of Copenhagen, Copenhagen, Denmark

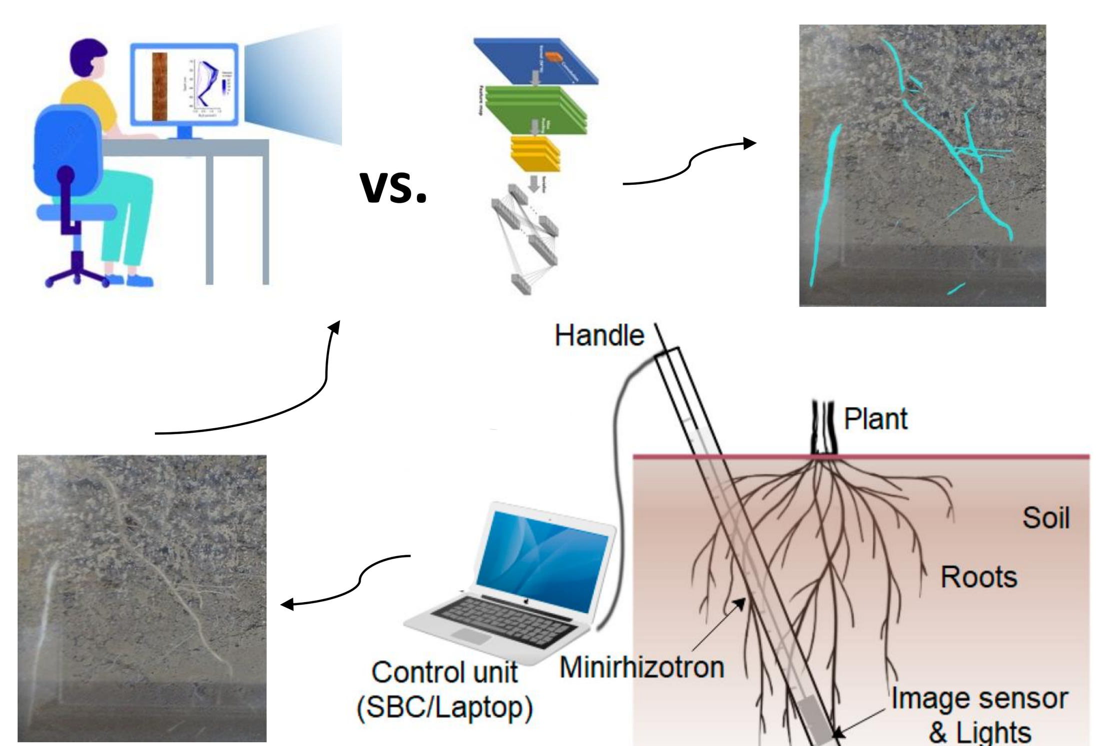
gih849@student.bham.ac.uk
i.carter@bham.ac.uk

Sharing is encouraged

Introduction

Fine roots are a major source of stabilised carbon in soils. However, the response of fine root dynamics to increased atmospheric CO_2 in temperate forests are poorly understood. Minirhizotrons can help to quantify fine root production and associated carbon dynamics in long-term, *in-situ* experiments such as Free Air CO_2 Enrichment (FACE) experiments. However, the manual annotation of hundreds of thousands of minirhizotron images is highly time-consuming and entails observer bias. Artificial Intelligence (AI) technology for image processing is fast developing, and Deep Neural Networks can already automate image analysis in simple crop systems. Here, we quantify how automated root analysis with AI (using RootPainter) compares with manual annotations by multiple root scientists, and the implications for root dynamics in a temperate forest (BIFoR FACE).

Schematic of the Use of AI vs Humans in Minirhizotron Image Analysis



Images adapted from Rahman et al., 2020, *Sensors*; Smith et al., 2022, *NewPhyt*; Zhou, 2023 in *RPubs*

Results

Outputs varied significantly between annotators, with **less experienced** annotators consistently identifying **more root length**

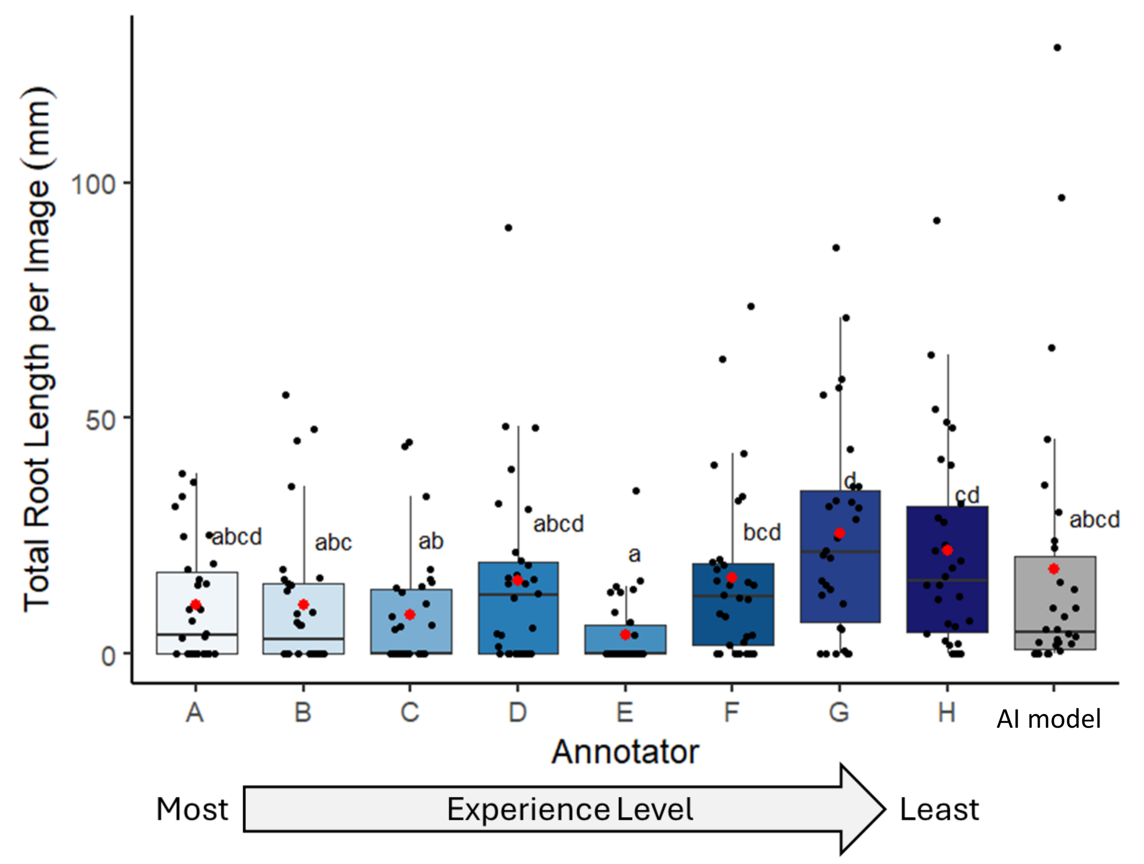


Figure 1: *Total root length output per image from 8 annotators (A-H) with varying levels of root experience, that manually traced minirhizotron images using Rootfly, and from our AI model. Points represent each individual image (n=30). Red points represent mean values, and compact letter display represents pairwise comparisons using Dunn's tests (p<0.005).*

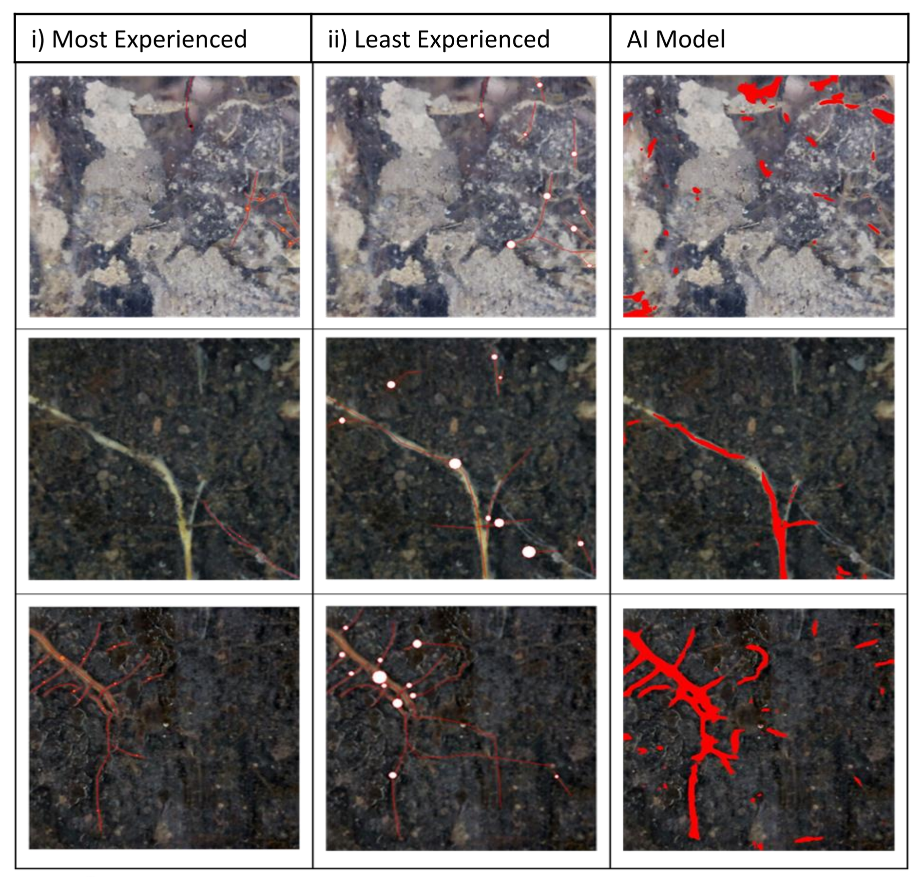


Figure 2: *3 annotated test images (randomly selected for exemplification). For i and ii, red lines represent traced length, and circles represent traced maximum diameter. In For AI model, red represents all areas where the model deemed root to be present.*

Results

The **AI model over-annotated** root length compared to **expert** manual annotation

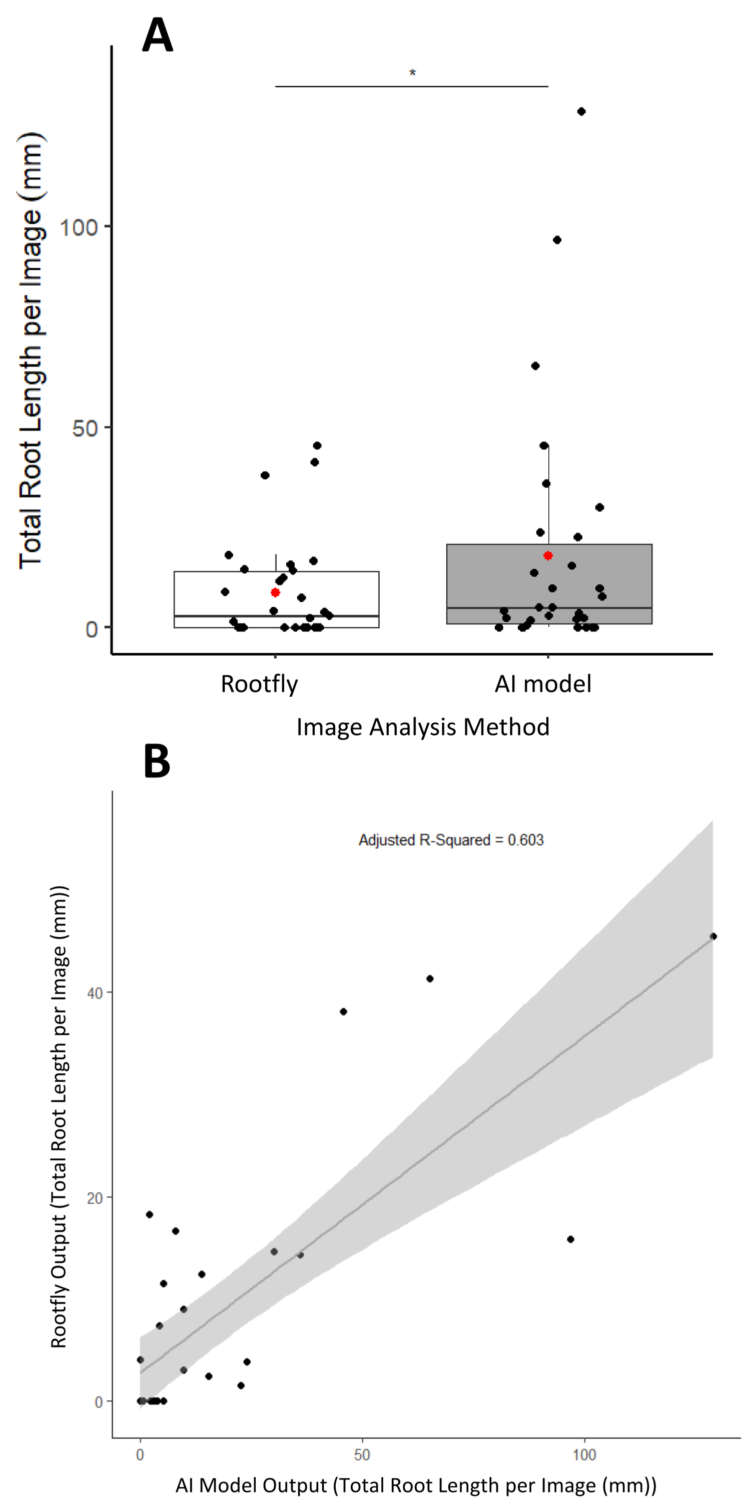


Figure 3: ***A*** *Total root length output per image from an expert consensus manual annotation using Rootfly, and an AI model. Points represent each individual image (n=30). Red points represent mean values. Wilcoxon signed-rank test (v=63, p=0.01).* ***B*** *Correlation between total root length output per image from expert consensus manual annotation, and AI model. Points represent each individual image (n=30). Spearmans rank correlation; rs=0.73.*

No significant net root length change difference in a **time** series between the **AI model** and **expert** manual annotation

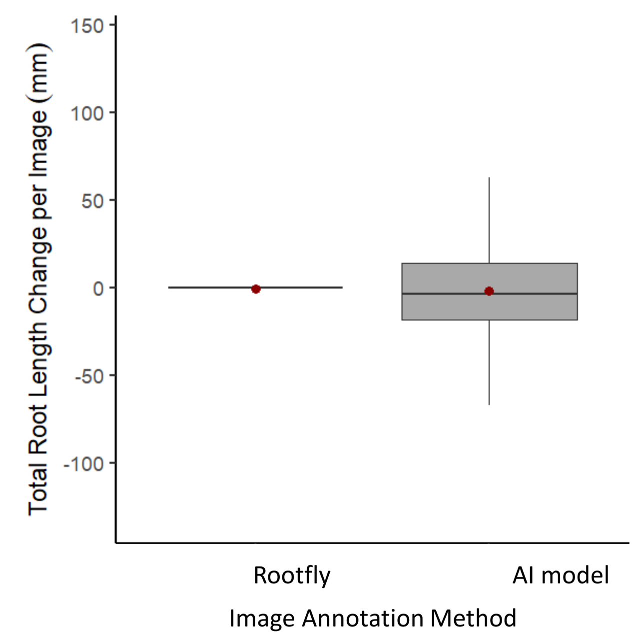


Figure 4: *Total root length net change over a 4-month period per minirhizotron image (n=529) from an expert manual annotation using Rootfly and an AI model. Red points represent mean values. Wilcoxon signed rank test (v=63714, p=0.08).*

Methods

- 8 root annotators of varying expertise manually annotated images *(n=30)* from BIFoR FACE minirhizotron dataset using Rootfly.
- AI trained with 966 images and resulting models used to annotate the same 30 images.
- Root length output per image compared between users and AI annotation
- Using a separate 4-month subset of images *(n=529),* root length net change compared between outputs from expert manual annotation and AI annotation

Conclusions

•Manual segmentation is contingent on individual annotators and their experience level.
•The AI model consistently over-estimated root length and exhibited unsatisfactory segmentations by eye. However, overall net root length changes were still meaningful, indicating that relevant effects of change over time could still be quantified with a seemingly inaccurate model.
•This has significant implications for analysis of long-term image datasets.

Acknowledgements

With special thanks to the wider BIFoR team, the Woodland Trust for financially supporting this research, and the Midlands Innovation Talent Commission for funding my attendance to EGU24.