Camera Align and Folder Drag are All You Need: Rapid Crop Lodging Aerial Assessment without Segmentation

2025/08/05

Haozhou WANG [1]
王 浩舟
Zheng Gong [2] Huaiyang Liu [3]
Huibin Li [3] Wei Guo [1]
Baofeng Su [2] Yun Shi [3]

[1] Graduate School of Agricultural and Life Sciences The University of Tokyo

2 College of Mechanical and Electronic Engineering, Northwest A&F University, China

3 State Key Laboratory of Efficient Utilization of Arid and Semi-Arid Arable Land in Northern China, Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, China

1. Introduction

Lodging after extreme weather

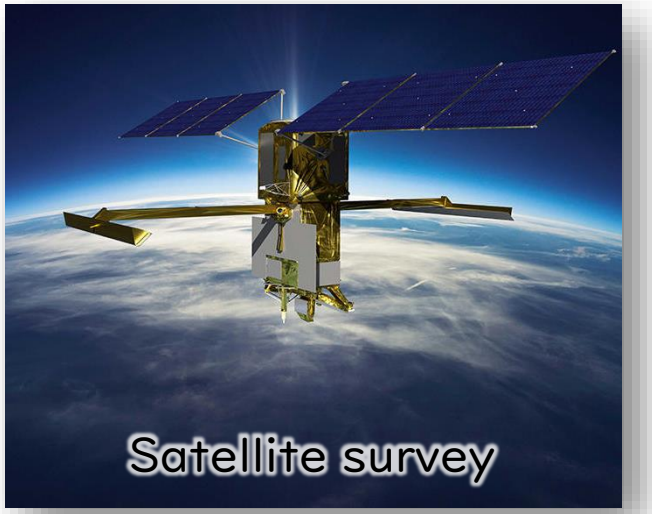
Lodging caused over 20% world annual grain loss

Post-disaster compensation and filed management decision relies on the quick and accurate estimate of lodging area

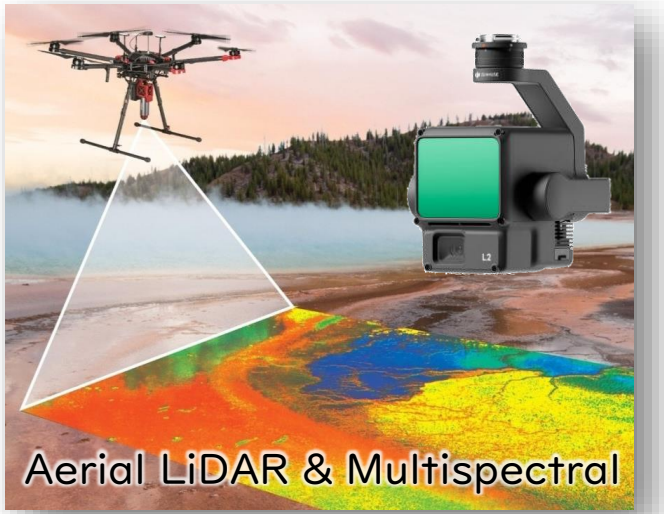
1. Introduction

Current methods and limitations

Low efficiency
High labor cost
Human errors



Low sparsal resolution
Cloud interference



Good for lodging estimate

Higher device cost

RGB UAV

1. Introduction

Challenges of using RGB UAV for lodging survey

1. High computation time to do photogrammetry (SfM-MVS) to produce geo-maps (DOM + DSM)
2. Digital surface model (DSM) from RGB images not always accurate for heights -> img texture
3. Complex and variate natural & crop condition makes segmentation difficult -> deep learning
4. High segmentation annotation time for training deep learning model

Time is critical for insurance company and government to make in-time decision

Objective: Can we decrease the time and labor cost in this processing?

02 Methods and Materials

2.1 Study area & devices

Experiment Rice field (46.69 ha) at Jiangsu, China

Lodged after Bebinca typhoon

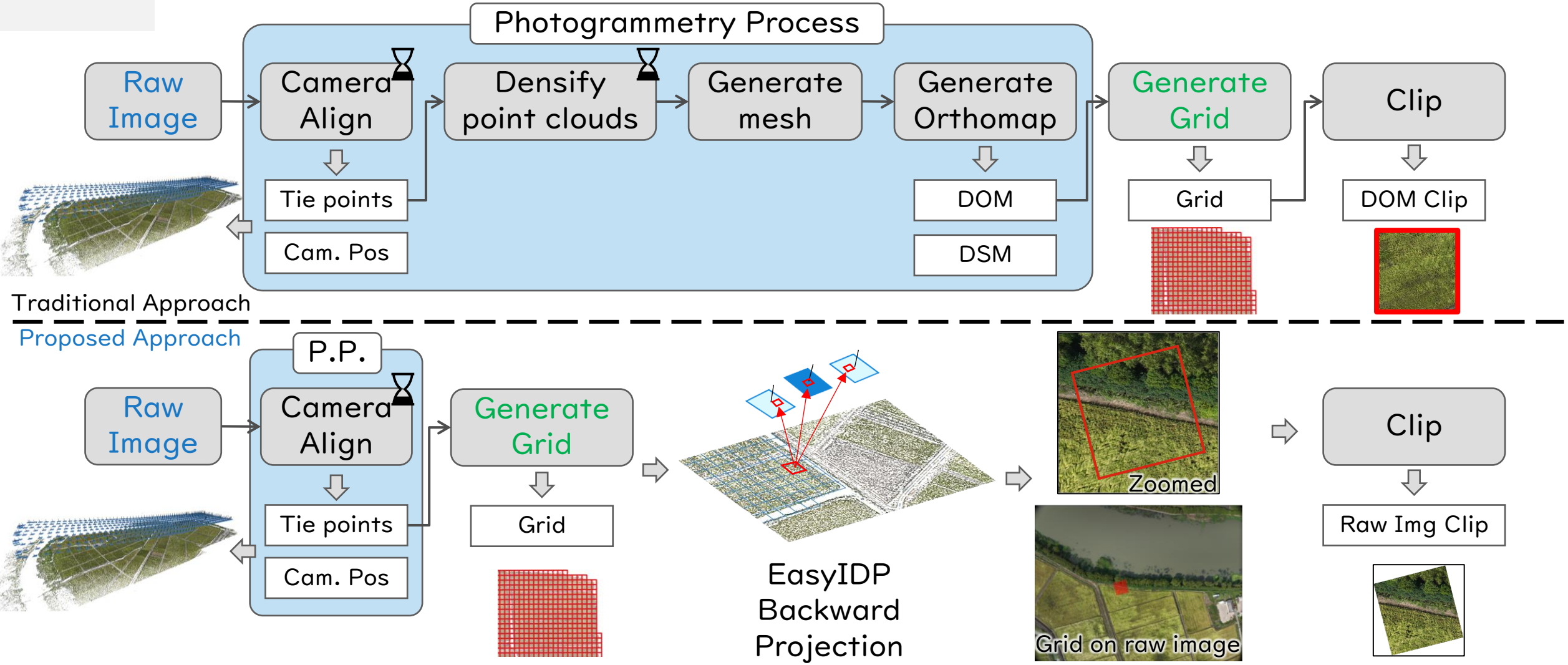
Data collection devices

DJI Mavic 3

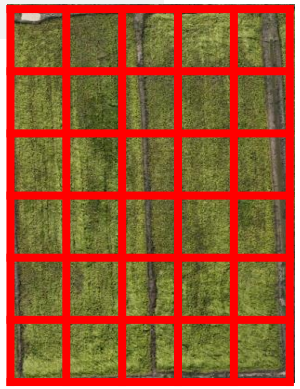
Only use RGB images

Survey time:	41 mins
Flight Height:	120m
Lateral overlap:	75%
Side overlap:	85%

2.2 Fast Preprocessing Approach



2.3 Fast lodging area distinguish approach



Full DOM

Traditional Approach

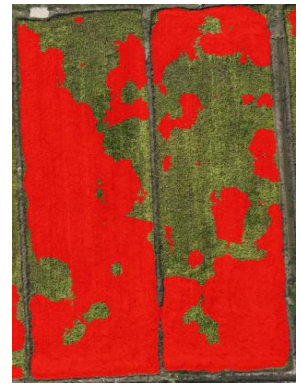
DOM clips

Segmentation Annotation

Seg. model training

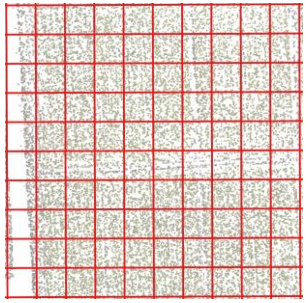
UNet

Seg. model inferencing

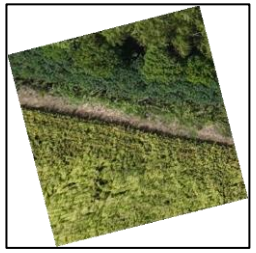


Lodging Results

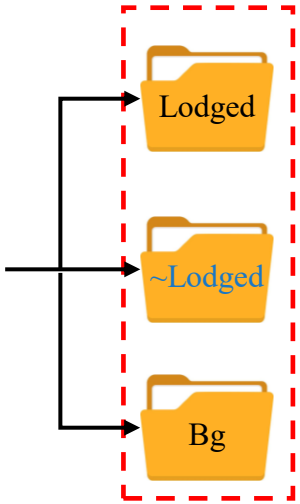
Proposed Approach



Grid on tie points



Raw image clips

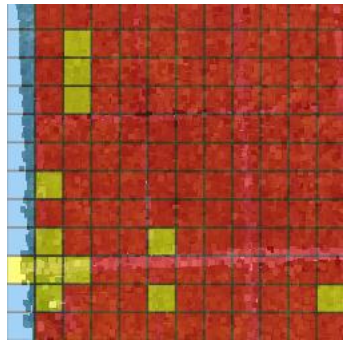


Classification Annotation

Cls. model training

CNN

Cls. model inferencing



Lodging Results

03 Results and Discussion

3.1 Processing time consuming comparison

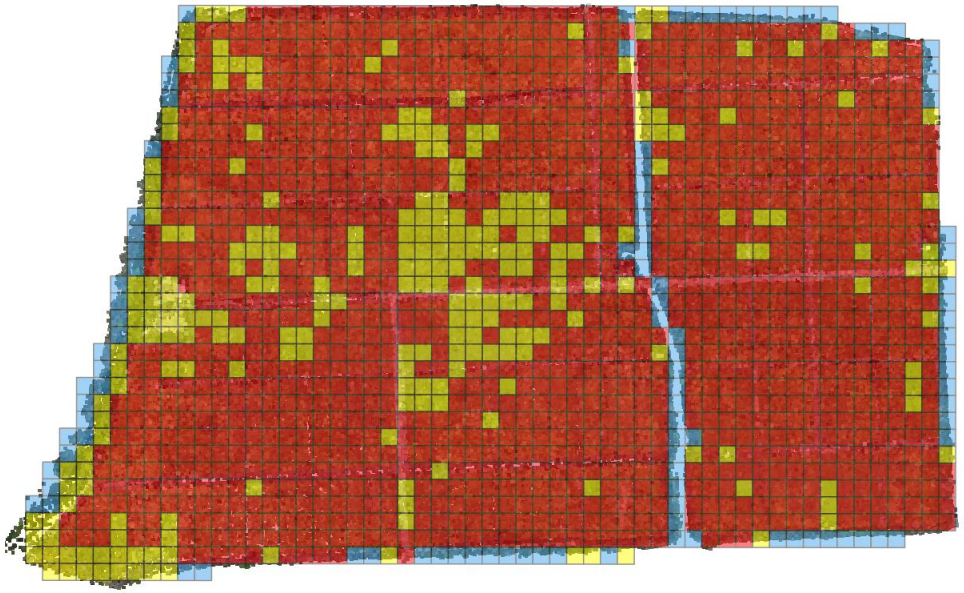
Traditional Steps	Time	Our Proposed steps	Time
Align photos	18 min	Align Photos	23 min
Densify point cloud	111 min	-	
Generate mesh	2 min	-	
Generate DEM	2 min	-	
Generate DOM	9 min	EasyIDP backward projection	9 min
Crop grids	10 min	Crop grids	7 min
Segmentation annotation	53 min	Classification annotation	25 min
Training seg. model	23 min	Train cls. model	10 min
Inference seg. model	4 min	Inference cls. model	2 min
Total time	**232 min**	**Total time**	**76 min**

3.2 Processing accuracy comparison

UNet segmentation approach

Lodged

Proposed classification approach (20m grid)

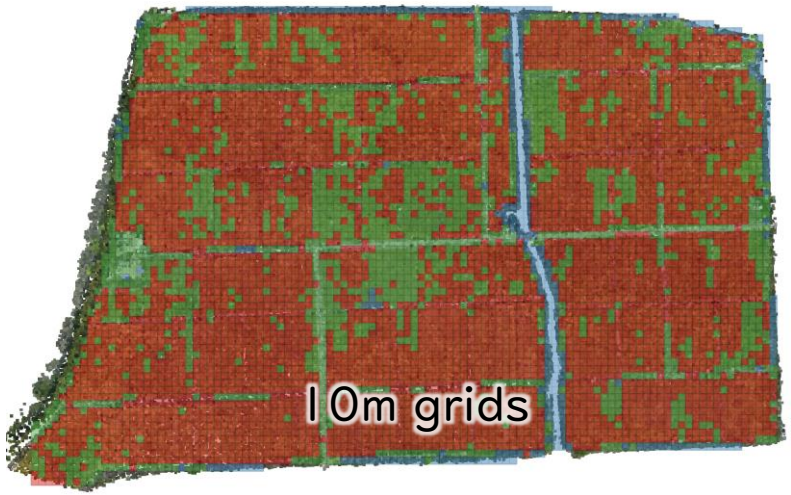


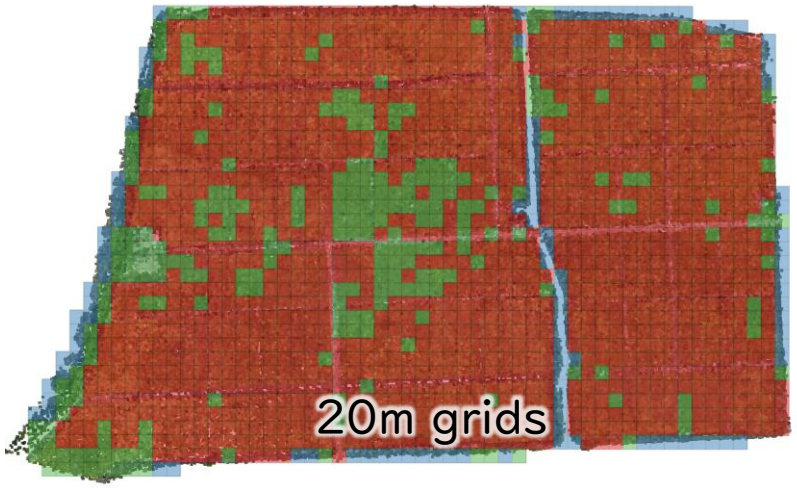
Lodged Non-lodged Background

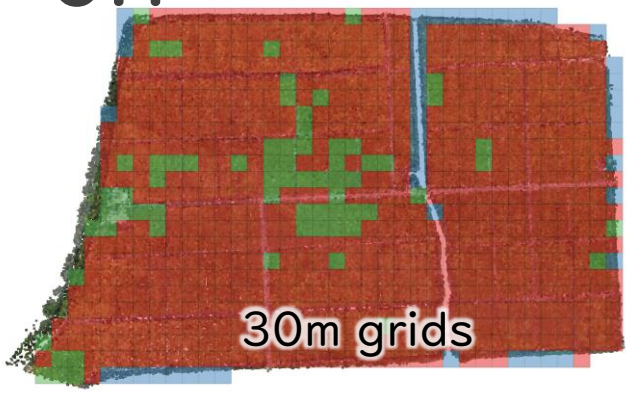
Method	Item	Accuracy	Precision	Recall	F1	Clip num
Unet Segmentation	**Total**	0.86	0.86	0.94	0.90	580
Our proposed	**Total**	0.83	0.77	0.83	0.78	647

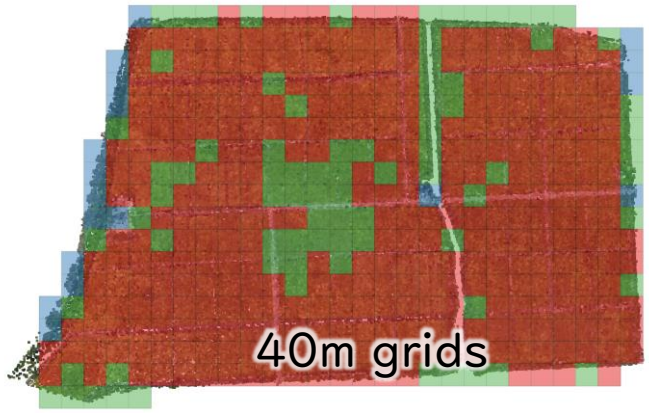
3.2 Processing time and grid size trade-off

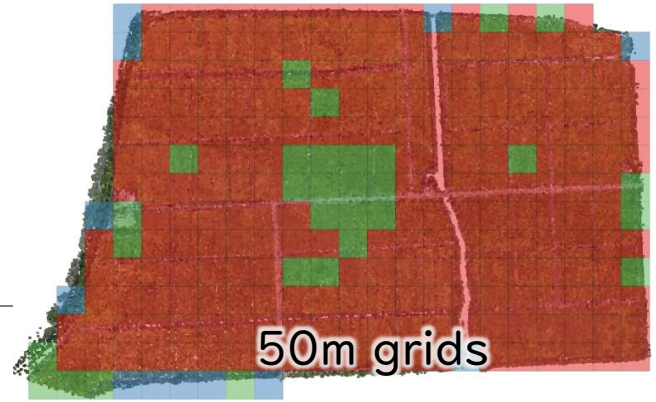
Methods	Total Processing Time
10meters Grids (6130images)	106 min
20meters Grids (1620images)	76 min
30meters Grids (722images)	65 min
40meters Grids (428images)	66 min
50meters Grids (264images)	61 min









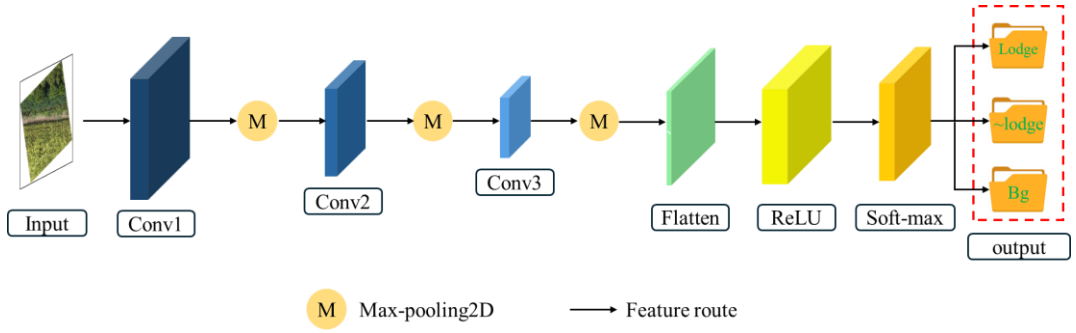


3.4 Future works

Expand to full research area

Expand to more crops
(e.g. cotton, wheat, maize)

3.4 Future works



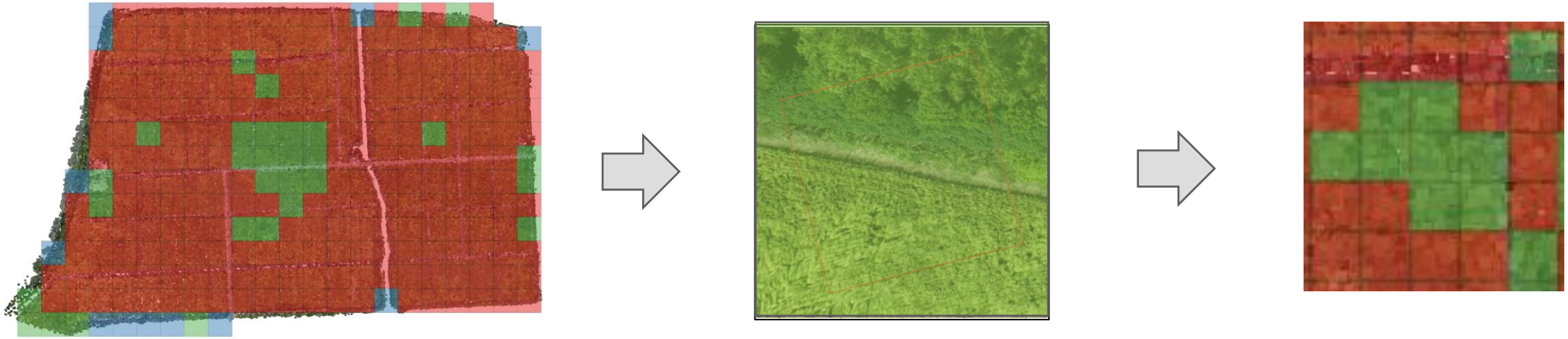
Just using simple CNN classification network to quick test our idea

Compared UNet segmentation is also not the latest technology

SOTA classification & segmentation networks comparison

Diffusion, SAM2, YOLO-CLS, etc…

3.4 Future works



50m grid fast classify

Mixed grid

10m detailed classify

Apply the pyramid to accelerate grid annotation for classification

04 Conclusion

4. Conclusion

By using only camera align and using classification to replace segmentation:

- Decrease lodging processing time from 232 min to 76 min (-67.24%)
- Classification preformed reasonable and acceptable worse than segmentation
- Grid size, efficiency and performance trade-off needs to be further explored

Thank You !

haozhou-wang@outlook.com

https://www.haozhou.wang

https://lab.fieldphenomics.com