

## Biomass Prediction with 3D Point Clouds from LiDAR

Liyuan Pan<sup>1,2,\*</sup>, Liu Liu<sup>2</sup>, Anthony G. Condon<sup>1</sup>, Gonzalo M. Estavillo<sup>1</sup>, Robert A. Coe<sup>1,3</sup>, Geoff Bull<sup>1,3</sup>  
Eric A. Stone<sup>1,2</sup>, Lars Petersson<sup>4</sup>, Vivien Rolland<sup>1,\*</sup>  
<sup>1</sup> CSIRO A&F, <sup>2</sup> ANU, <sup>3</sup> Australian Plant Phenomics Facility, <sup>4</sup> CSIRO Data61  
liyuan.pan@csiro.au, vivien.rolland@csiro.au

### Abstract

With population growth and a shrinking rural workforce, agricultural technologies have become increasingly important. Above-ground biomass (AGB) is a key trait relevant to breeding, agronomy and crop physiology field experiments. However, measuring the biomass of a cereal plot requires cutting, drying and weighing processes, which are laborious, expensive and destructive tasks. This paper proposes a non-destructive and high-throughput method to predict biomass from field samples based on Light Detection and Ranging (LiDAR). Unlike previous methods that are based on the density of a point cloud or plant height, our biomass prediction network (BioNet) additionally considers plant structure. Our BioNet contains three modules: 1) a completion module to predict missing points due to canopy occlusion; 2) a regularization module to regularize the neural representation of the whole plot; and 3) a projection module to learn the salient structures from a bird's eye view of the point cloud. An attention-based fusion block is used to achieve final biomass predictions. In addition, the complete dataset, including hand-measured biomass and LiDAR data, is made available to the community. Experiments show that our BioNet achieves  $\approx 33\%$  improvement over current state-of-the-art methods.

### 1. Introduction

Based on the United Nations Department of Economic and Social Affairs projections, the world population is expected to reach 10.8 billion in 2100 [9]. Agricultural production will hence need to increase, whilst the pressure on already scarce land resources will continue to rise and the available workforce to shrink [4]. Therefore, new technologies, effective land usage, and efficient management practices are required to sustainably enhance productivity.

Biomass is a crucial variable for assessing the production and performance of agricultural systems. Biomass predic-



Figure 1. Overview of fields. Our method aims to predict the biomass of each plot via the scanned point cloud. The point cloud is obtained by the PhenoMobile-Lite<sup>®</sup> (zoom-in of the red box). (Best viewed in colour on screen.)

tions inform growers and scientists about the amount and timing of fertilizer, pesticides, and water that is required to optimize crop performance and improve harvest production [20]. Moreover, biomass is an important target for genetic improvement to increase crop yield [30].

However, traditionally biomass estimation methods require cutting the culms at ground level for a defined portion of an experimental plot, and weighing the plant material after drying in an oven to constant weight [31]. The applicability and reliability of this method is limited by: 1) variable accuracy, as the sample represents only a small section of the whole experimental plot, which could bias the results for non-uniform field plots; 2) the method being destructive, thereby limiting the number of samples on which it can be applied; and 3) the sampling and subsequent processing requiring transport, drying, and manual handling, being slow, expensive and restrictive in large experiments. In summary, the measurement of biomass is laborious, subject to considerable error and not amenable to high-throughput screens required in modern field experiments or breeding trials.

Therefore, a non-destructive and high-throughput alternative method is required. In recent years, computer vision technologies have been extensively utilized to support dif-

\*corresponding authors

ferent tasks in agriculture, including LiDAR-based methods [7, 54, 3]. Existing LiDAR-based biomass estimation methods can be categorized into two streams: canopy height-based and point density-based approaches. Canopy height-based approaches [43, 46, 11] use the canopy height as a proxy to estimate biomass in maize, wheat, and rice. However, variations in crop height are limited in wheat breeding experiments. Consequently, researchers tend to utilize the 3D nature of a point cloud. Voxel-based, volume-based, and point distribution-based methods are preferred as the density of the point cloud is shown to correlate strongly with the manually measured biomass [19, 45, 10]. Although these methods can be applied without a machine learning mechanism, limited accuracy restricts their implementation. Additionally, in small grain cereals such as wheat, the following limitations currently exist:

- Occlusion in the canopy, especially after the flowering stage and in high-density plots. Hidden areas cannot be detected by LiDAR, hence reducing the accuracy of above methods.
- Lack of relevant datasets. Non-destructive biomass estimation at fine levels is currently not possible due to the lack of accurate and high-throughput phenotypic data and algorithms.

In this paper, we propose an end-to-end network, dubbed BioNet, to predict above-ground biomass. BioNet contains three modules: 1) a completion module (Sec. 3.2); 2) a regularization module (Sec. 3.3); and 3) a projection module (Sec. 3.4). We adopt an attention-based fusion block (Sec. 3.5) to fuse the output features from the three modules and output the final predicted biomass.

Our main contributions are summarized as follows.

- 1) We propose a simple and effective network, named the Biomass prediction Network (BioNet), to predict above-ground-biomass robustly and accurately with point clouds generated by LiDAR.
- 2) We share a dataset with point clouds of 306 plots representing a total of 26 varieties of small grain cereals, captured at two developmental stages and two sowing densities, as well as their associated ground truth biomass.
- 3) Experimental results on the real dataset demonstrate the superiority of our approach over current SOTA methods.

## 2. Related work

This section first summarizes studies that estimate above-ground biomass (AGB) with LiDAR for field crop research. Then, we briefly review commonly used learning methods for point clouds.

**Biomass Prediction with LiDAR.** Canopy height, as a surrogate for biomass, has been widely adopted [27, 15, 43]. Saeys *et al.* [36] estimated crop density using statistical

models for two LiDAR scanning frequencies. Tilly *et al.* [43] added a field spectrometer to their platform and derived bivariate biomass regression models by fusing 3D data, and spectral data [42]. Li *et al.* [23] used airborne LiDAR, and Pearson’s correlation analysis and structural equation modelling (SEM) to estimate plant height and leaf area index (LAI). To estimate canopy height, several studies used stereo reconstruction from aerial imagery [1], or ground platforms [13] as an alternative to LiDAR. However, prediction of biomass from canopy height is unlikely to be useful in the selection for improved biomass since the variation for height is limited in breeding programs.

In addition to canopy height, other variables, such as point volume, LiDAR Projected Volume (LPV), and 3D indices [19, 45, 38, 6], have a strong correlation with measured biomass. Point density-based methods which utilize the 3D nature of point clouds have been applied in cotton [38], arctic shrubs [18] and trees [32], or for wheat grown in a single environment [45]. Jimenez-Berni *et al.* [19] proposed a voxel-based method (3DVI) by dividing the point cloud into voxels of equal size and calculating the ratio between the number of voxels containing points and the number of subdivisions in the horizontal plane. They also proposed a complement to 3DVI, that required splitting point clouds into layers, applying a correction factor to each layer, and summarizing the corrected point fractions present in each layer. Importantly, 3DVI is currently considered the ‘gold-standard’ for measuring AGB. 3DVI is robust and accurate for real-world applications and is used in [45, 41].

**Learning in Point Cloud.** PointNet [33] is a pioneering effort that directly processes point sets. It uses multi-layer perceptron (MLP) networks [39] to extract features from point sets. PointNet++ [34] is an extension of PointNet with an added hierarchical structure. In this approach, the set of points is partitioned into overlapping local regions by the distance metric of the underlying space. Several other methods [22, 48, 49, 24, 25] were proposed which utilize local features or a pooling operation that is more advanced than max-pooling. For example, Wang *et al.* [48] proposed a new neural network module dubbed EdgeConv suitable for CNN-based high-level tasks on point clouds, including classification and segmentation. EdgeConv is differentiable and can be plugged into existing architectures. Compared with previous 3D point CNNs, GS-Net [51] exploits EigenGraph to group distant points with similar and relevant geometric information, and aggregates features from nearest neighbours in both Euclidean space and Eigenvalue space.

3D shape completion plays an important role in robotics and perception. Yuan *et al.* [53] proposed the pioneering point cloud completion (PCN) work using a simple encoder-decoder network. TopNet [40] proposed a hierarchical tree-structure network. GAN-base network RL-GAN-Net [37] was invented for real-time point cloud completion.

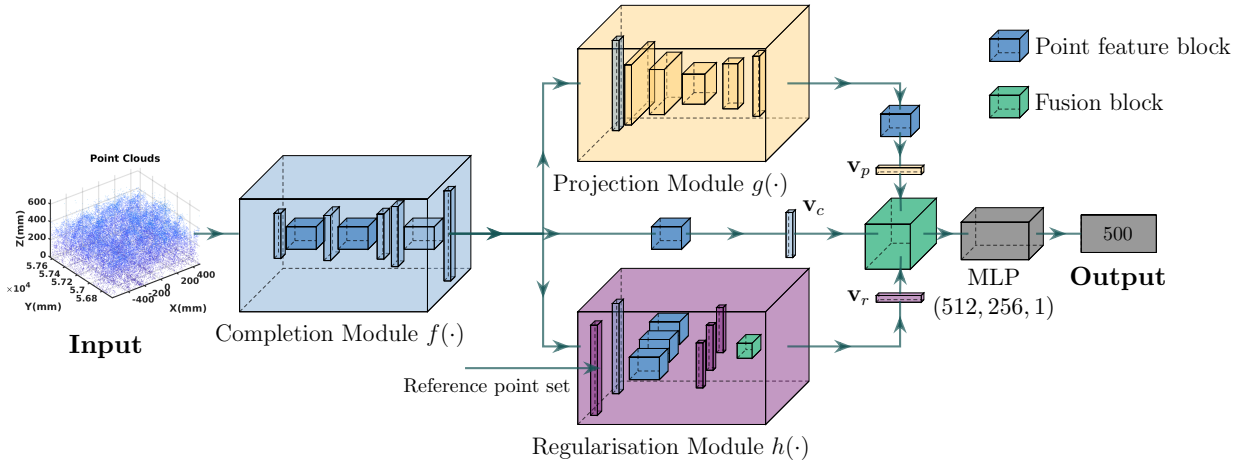


Figure 2. Architecture of BioNet. We first feed the point cloud  $\mathcal{X}$  to the completion module to get a completed point cloud  $\mathcal{Y} = f(\mathcal{X}; \vec{\mathcal{F}})$ . A feature vector  $\mathbf{v}_c$  is extracted from  $\mathcal{Y}$  with a MLP block. Meanwhile, the point cloud  $\mathcal{Y}$  is fed to the regularization and projection module to extract a feature vector  $\mathbf{v}_r$  and  $\mathbf{v}_p$ , respectively. Modulated by the feature vector of a reference point cloud, the regularized planting practice, the projection module first projects  $\mathcal{Y}$  as a bird-eye view image, and then extracts a salient structural feature vector  $\mathbf{v}_p = g(\mathcal{Y}; \vec{\mathcal{G}})$ . With  $\mathbf{v}_c$ ,  $\mathbf{v}_r$  and  $\mathbf{v}_p$ , we adopt an attention mechanism to fuse the three feature vectors. A MLP block outputs our final biomass prediction.

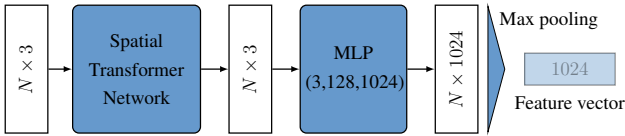


Figure 3. Architecture of the point feature block (small dark-blue cuboids in Fig. 2). The block aims to extract a feature vector from the input point set. (Best viewed in colour on screen.)

### 3. Biomass prediction network

In this section, we first present an overview of our network and then introduce the three modules which contribute to accurate and robust biomass predictions. Finally, we present an attention fusion method to aggregate outputs of the three modules together.

#### 3.1. Overview

A point cloud of a plot is represented as a set of 3D points  $\mathcal{X} = \{\mathbf{X}_i | i = 1, \dots, N\}$ , where each point  $\mathbf{X}_i$  is a coordinate vector  $(x_i, y_i, z_i)$  and  $N$  is the number of 3D points in the plot. With the input point cloud, our BioNet is able to predict the AGB  $\hat{w}$ . To train our BioNet, we use ground-truth biomass  $w$ .

The pipeline of BioNet is given in Fig. 2. BioNet contains three key modules: the completion module  $f(\cdot)$  with parameters  $\vec{\mathcal{F}}$ , the regularization module  $h(\cdot)$  with parameters  $\vec{\mathcal{H}}$ , and the projection module  $g(\cdot)$  with parameters  $\vec{\mathcal{G}}$ . The completion module (Sec. 3.2) predicts the miss-

ing points due to canopy occlusion caused by neighbouring plants. The regularization module (Sec. 3.3) is to regularize the neural representation of the whole cereal plot with a reference wheat plant. The projection module (see Sec. 3.4) first projects the whole cereal plot as a bird's eye view, and then extracts salient structural features. Finally, output features from the three modules are fused to predict the above-ground biomass  $\hat{w}$ .

To train BioNet, we use both a prediction loss  $\mathcal{L}_{\text{pre}}$  and a completion loss  $\mathcal{L}_{\text{comp}}$ . The final loss is the sum of these two losses

$$\mathcal{L} = \mathcal{L}_{\text{pre}} + \lambda \mathcal{L}_{\text{comp}}, \quad (1)$$

where  $\lambda$  is the weight parameter to balance the two losses.

**Prediction loss  $\mathcal{L}_{\text{pre}}$ .** We adopt the smooth  $\ell_1$  loss. Smooth  $\ell_1$  loss  $\mathcal{S}(\cdot)$  is widely used in regression tasks, for its robustness and low sensitivity to outliers [16]. This loss enforces the predicted biomass to be similar to the ground truth biomass, and is given by

$$\mathcal{L}_{\text{pre}}(\hat{w}, w) = \frac{1}{M} \sum_{m=1}^M \mathcal{S}(w_m - \hat{w}_m), \quad (2)$$

where  $M$  is the number of plots and the subscript  $m \in \{1, \dots, M\}$  is the plot index.

**Completion loss  $\mathcal{L}_{\text{comp}}$ .** Taking the original point cloud of a plot as input, the completion module outputs a high-fidelity, dense point cloud. To measure the difference between the completed point cloud  $\mathcal{Y}$  and the ground truth

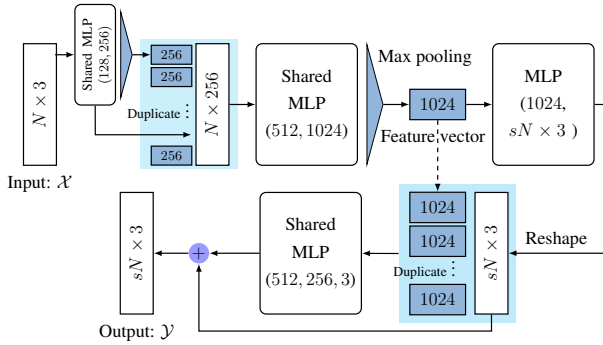


Figure 4. Architecture of the completion module. The module aims to complete the low-density point cloud  $\mathcal{X}$  with a scale factor  $s = 2$ .

point cloud  $\mathcal{G}$ , we adopt the Earth Mover’s Distance (EMD) [12, 53, 26]. The  $\mathcal{L}_{\text{comp}}$  is defined as

$$\mathcal{L}_{\text{comp}}(\mathcal{Y}, \mathcal{G}) = \min_{\phi: \mathcal{Y} \rightarrow \mathcal{G}} \frac{1}{|\mathcal{Y}|} \sum_{\mathbf{Y} \in \mathcal{Y}} \|\mathbf{Y} - \phi(\mathbf{Y})\|. \quad (3)$$

Here,  $\phi$  is a bijection that minimizes the average distance between corresponding points [5].

Next, we describe each key component of BioNet.

### 3.2. Completion module

The depth of laser pulse penetration varies with the canopy structure. In early developmental stage (vegetative), small grain cereal plants such as wheat have few leaves which are all close to the stem. At later developmental stages (flowering and later), the same plants widen to form a canopy of varying density, depending on the species, variety and sowing density. It is difficult for laser pulses to penetrate the canopy due to the occlusion of leaves and panicles, leading to partially-scanned 3D point clouds. This poses a challenge for existing methods which estimate the volume of scans, as only points collected from the top of the canopy are used to compute biomass. As discussed above, the inaccurate prediction due to occlusion should be noted as a limitation of LiDAR-based biomass measurements.

Therefore, we add a completion module to BioNet, to predict missing points due to occlusion. This module directly maps a partial point cloud  $\mathcal{X}$  to a dense, completed point cloud  $\mathcal{Y}$  without any voxelization. Inspired by [26, 53, 47], we generate  $\mathcal{Y}$  in a coarse-to-fine fashion.

The completion module is an encoder-decoder network, and is given in Fig. 4. The encoder network takes the input point cloud  $\mathcal{X}$  as input, and extracts a  $k$ -dimensional feature vector  $\mathbf{v} \in \mathbb{R}^k$  with  $k = 1024$ . The decoder network takes the feature vector  $\mathbf{v}$  as input, and outputs the completed point cloud  $\mathcal{Y}$ , by combining the advantages of a fully connected decoder and a folding-based decoder [2, 52]. In Fig. 5, we show an example of the input point cloud  $\mathcal{X}$  and the

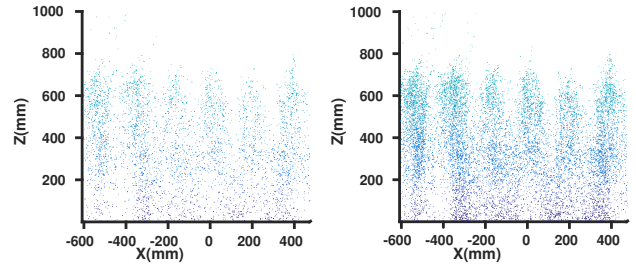


Figure 5. We give an example of the input point cloud  $\mathcal{X}$  (Left) and the completed point cloud  $\mathcal{Y}$  (Right).

completed point cloud  $\mathcal{Y}$ . The result shows that our completion module is able to predict missing points caused by occlusion.

The completed point cloud  $\mathcal{Y}$  is then fed to the regularization and projection modules.

### 3.3. Regularization module

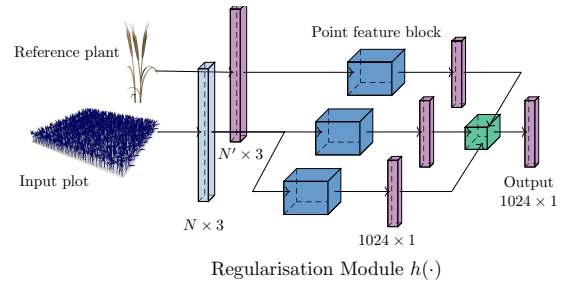


Figure 6. Architecture of the regularization module. A reference wheat plant is introduced to regularize the neural representation of the whole cereal plot. (Best viewed in colour on screen.)

As discussed earlier, it is common that the point set shows a non-uniform density across different areas of a given plot. Such non-uniformity introduces a significant challenge for point set feature learning and accurate biomass predictions. Features learned in dense data may not generalize to sparsely sampled regions. Conversely, models trained with sparse point clouds may not recognize fine-grained local structures.

Besides non-uniform density, the canopy structure of small grain cereals varies depending on the species, variety, developmental stage and sowing density. Consequently, features learnt from a given variety/stage/density may not generalize to another.

Therefore, we use the attention mechanism [44, 14, 8, 29] to selectively aggregate the feature by a weighted summarization, which helps to tackle the non-uniform density in the point cloud. Moreover, to ‘regularize’ the point cloud of a plot that contains multiple cereal plants, we use a single standard wheat plant as a reference. Features extracted from the point cloud of the reference plant are used to rec-

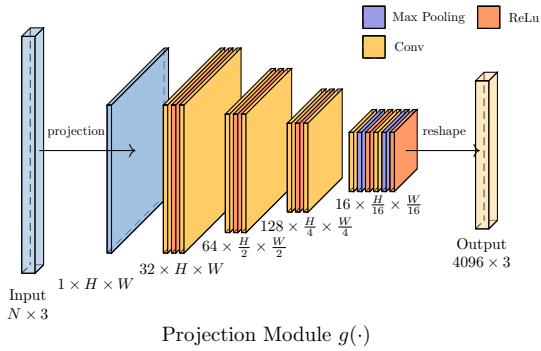


Figure 7. Architecture of the projection module. The module aims to extract salient structural features from a bird's eye view projection of the point cloud. (Best viewed in colour on screen.)

tify features learnt from the point cloud of the whole plot. The above strategy allows our model to learn features from irregular plant structure when predicting the biomass compared to existing density-based methods. The architecture of our regularization module is given in Fig. 6. Specifically,

First, a query feature  $\mathbf{Q} \in \mathbb{R}^k$  is extracted from the reference plant  $\mathcal{X}_{\text{ref}}$  with a MLP block. Next, a key feature  $\mathbf{K} \in \mathbb{R}^k$  is extracted from the completed point cloud  $\mathcal{Y}$  with a MLP block. After performing a matrix multiplication between  $\mathbf{Q}$  and the transpose of  $\mathbf{K}$ , we apply a softmax layer to calculate an attention map  $\mathbf{A} \in \mathbb{R}^{k \times k}$ .

$$\mathbf{A} = \text{Softmax}_i(\mathbf{Q} \cdot \mathbf{K}^T), \quad (4)$$

where  $i \in \{1, \dots, k\}$  is the row index and  $\text{Softmax}_i(\cdot)$  is the (row-wise) softmax function [17].

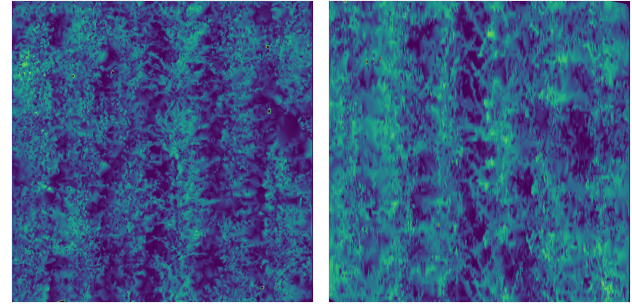
Then, a value feature vector  $\mathbf{V} \in \mathbb{R}^k$  is extracted from the complete point cloud  $\mathcal{Y}$  with a MLP block. Finally, the output of our regularization module is computed as a weighted sum of value features by performing a matrix multiplication between  $\mathbf{A}$  and  $\mathbf{V}$ .

In summary, the query feature  $\mathbf{Q}$  from the reference plant is used to draw an attention map  $\mathbf{A}$  on the completed point cloud  $\mathcal{Y}$ , and  $\mathbf{A}$  is further used to regularize the value feature vector  $\mathbf{V}$  of  $\mathcal{Y}$ .

### 3.4. Projection module

In an actual application, a point cloud is sparsely sampled from 3D surfaces in an irregular and off-order way. In addition, plants are sown in rows in a plot, separated by even gaps. The gaps allow plants to grow but also contribute to the non-uniform point distribution in each cloud. To feed the non-uniform and sparse point cloud to a network in training, previous attempts included iterative farthest point sampling, volumetric grids and view-based projections [34, 50, 28].

In this study, we represent the 3D point cloud as a 2D depth image to achieve robustness to transformations and



(a) Early stage (b) Flowering stage

Figure 8. Bird's eye view of the same plot at two different growing stages. We generate a greyscale image based on points  $(x_i, y_i, z_i)$  in a point cloud. The image coordinates correspond to  $(x_i, y_i)$  and pixel intensity values correspond to  $z_i$ .

permutations. A strong structure can be found when looking down at the plot, due to the standardised planting practices which ensure seeds are planted in evenly-spaced rows. In Fig. 8, we show the representation of a plot with small grain cereals at two growing stages (one at the early stage and one at the flowering stage), where it is possible to recognize six plant rows (shown as columns) in either stages.

In our representation, each point  $(x_i, y_i, z_i)$  in a cloud is mapped to discrete coordinates  $(x_i, y_i)$ , where its intensity value is  $z_i$ . Compared with point sampling, the representation helps to feed all points to the network in training. In addition, the representation is able to filter the soil point which can be treated as noise in the point set. After the represented image, we adopt the convolutional neural network (CNN) to extract a  $k$ -dimensional feature vector from the 2D depth image. The architecture of the projection module is given in Fig. 7. Then, with a point feature block we extract the feature vector  $\mathbf{v}_p$ .

The above three modules form the basis of our BioNet.

### 3.5. The fusion block

As explained above, our network starts with a point cloud  $\mathcal{X}$  being fed to the completion module to get a completed point cloud  $\mathcal{Y}$ . A feature vector  $\mathbf{v}_c$  is extracted from  $\mathcal{Y}$  with a MLP block. Meanwhile, the point cloud  $\mathcal{Y}$  is separately fed to the regularization and projection module to extract a feature vector  $\mathbf{v}_r$  and  $\mathbf{v}_p$ .

With  $\mathbf{v}_c$ ,  $\mathbf{v}_r$  and  $\mathbf{v}_p$ , we feed the three feature vectors to a feature fusion block. The fusion block adopts the attention mechanism [44], and is given by

$$\text{Fuse}(\mathbf{v}_r, \mathbf{v}_p, \mathbf{v}_c) = \text{Softmax}\left(\frac{\mathbf{v}_r \mathbf{v}_p^T}{\sqrt{k}}\right) \mathbf{v}_c, \quad (5)$$

where  $\sqrt{k}$  is a scaling factor and  $k$  is the dimension of feature vectors. The output of the fusion block is fed to a MLP block to extract our final biomass prediction  $\hat{w}$ .

## 4. Experiments

### 4.1. Dataset

We collected data from a field experiment conducted in 2019 at Yanco, NSW Australia. In this experiment, 26 varieties of small grain cereals (wheat and triticale) were sown at two different densities to generate a range of AGB values and plant architectures. Seeding rates were 250 ('high', standard Australian yield trials) and 50 ('low') seeds/m<sup>2</sup>. Both sub-trials were sown in May, and each crop variety was randomly replicated for each sowing density, giving 78 plots per sub-trial, 156 plots in total.

Each plot was scanned at two different developmental stages, early (August, vegetative stage) and late in the growing season (October, flowering stage). Point clouds were generated with LiDAR mounted on a PhenoMobile-Lite<sup>®</sup> driven above all plots at 2 m height. The experimental imaging set-up is shown in Fig. 9. The PhenoMobile-Lite<sup>®</sup>, conceived as a manually operated buggy, is designed to be lightweight, cost-effective, and transportable across multiple field sites. It provides reliable field phenotyping amenable to deployment in multi-site managed environmental facilities for targeted trait and germplasm evaluation [35]. To measure the associated ground truth above-ground biomass, ground-level cuts were then taken from each plot immediately after the LiDAR scanning. The cuts, taken from 1 m<sup>2</sup> (early stage) and 0.5 m<sup>2</sup> (flowering stage) areas of each plot, were oven-dried at 70°C for 7 days until a constant dry weight was reached. They were then weighed to calculate AGB on area basis (g/m<sup>2</sup>) for each cut.

Agronomic and bioinformatics experts worked together for five months to collect the 306 point clouds and corresponding manually-measured ground-truth AGB. The dataset was then split into 204 training plots and 102 testing plots. We make our small grain cereals biomass prediction (SGCBP) dataset<sup>1</sup> publicly available.

### 4.2. Experimental Setup

**Baselines.** The state-of-the-art method is 3DVI [19], which is a traditional density-based, non-deep method. It is demonstrated to be effective and accurate in real-world applications and is used in [45, 41]. Given that our method is the pioneer of learning-based biomass prediction methods, there is no open-source baseline available at this stage. Thus, to prove the effectiveness of our BioNet, we build several baselines based on the state-of-the-art point cloud learning methods. The selected methods include PointNet [33], PointNet++ [34], DGCNN [48] and GS-Net [51]. All networks are retrained on the proposed new dataset and are optimized with the same prediction loss function.

**Evaluation metrics.** We use standard metrics to evalu-

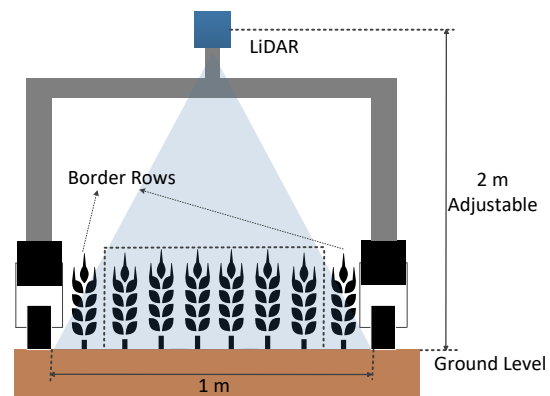


Figure 9. Schematic of our imaging set-up, using a PhenoMobile-Lite<sup>®</sup> to collect point clouds. PhenoMobile-Lite comprising LiDAR laser scanner and touch-screen computer mounted on an aluminium frame with adjustable wheel spacing to accommodate different plot widths (1.7–2.2 m) and ground-clearance for canopy heights up to 1.5 m. The height-adjustable sensor boom (2.0–2.5 m) enables data capture from crop emergence to maturity. In the standardised planting practices, the border rows are cut manually. This is because the plant at border rows has a larger growing space and more sunlight than the inner rows, making the canopy structure of border rows different from the inner rows.

ate the quality of the AGB prediction using BioNet from the LiDAR point cloud. Specifically, these metrics include the mean absolute error 'MAE', the absolute relative error 'MAE<sub>rel</sub>', and the root mean square error 'RMSE'. Here, the 'MAE<sub>Rel</sub>' is defined as

$$\text{MAE}_{\text{rel}} = \frac{1}{M} \sum_{m=1}^M \frac{|\hat{w}_m - w_m|}{w_m}, \quad (6)$$

where  $M$  is the number of point clouds and the subscript  $m \in \{1, \dots, M\}$  is the plot index. Here,  $\hat{w}_m$  is the  $m^{\text{th}}$  predicted above-ground biomass and  $w_m$  is the  $m^{\text{th}}$  manually measured ground-truth biomass.

**Implementation details.** BioNet is implemented in PyTorch and is trained from scratch using the Adam optimizer [21] with a learning rate of  $10^{-3}$  and a batch size of 24. Our model is trained on four NVIDIA Tesla P100 GPUs. During the training procedure, we augment the data by randomly sampling and perturbing the point locations.

### 4.3. Experimental Results

We compare our results with baselines on the collected real datasets (note: no other datasets are available). BioNet predictions achieve competitive results compared with the state-of-the-art methods [33, 34, 48, 51, 19] (Table 1). Here, 'Ours<sub>c</sub>', 'Ours<sub>r</sub>', and 'Ours<sub>p</sub>' denote our model with only the completion module, the regularization module, and the projection module, respectively. 'Ours' refers to our model with its three modules and an attention mechanism-based

<sup>1</sup><https://doi.org/10.25919/xv6v-6h56>

Table 1. *Quantitative comparisons of biomass prediction methods on our real dataset. The best results are shown in bold. The model of each learning-based method is trained on the whole training set. We report the result on the whole testing set ('Overall'), the 'Early stage' set, and the 'Flowering stage' set separately. We achieve competitive results compared to state-of-the-art methods in all three testing sets. To demonstrate the effectiveness of each module in our BioNet, we also report the result of only using the completion module ('Our<sub>c</sub>'), the regularization module ('Our<sub>r</sub>'), and the projection module ('Our<sub>p</sub>') separately. In the last column, we show the relative improvement ('RI', (%)) compared to the state-of-the-art method 3DVI. Our BioNet achieves  $\approx 33\%$  improvement on average.*

Method \ Metric	PointNet [33]	PointNet++ [34]	DGCNN [48]	GS-Net [51]	3DVI [19]	Our <sub>c</sub>	Our <sub>r</sub>	Our <sub>p</sub>	Ours	RI
Overall										
MAE ↓	139.61	142.80	129.66	145.63	115.15	92.46	112.85	108.36	<b>71.23</b>	38.14
MAE <sub>rel</sub> ↓	27.40	27.52	25.35	26.13	19.02	18.79	18.49	18.68	<b>12.08</b>	36.49
RMSE ↓	189.80	188.85	161.61	180.85	151.79	119.69	146.87	140.93	<b>99.33</b>	34.56
Early stage										
MAE ↓	92.59	120.62	83.07	83.68	60.45	69.83	72.06	79.56	<b>43.11</b>	28.68
MAE <sub>rel</sub> ↓	32.61	41.07	33.40	33.06	21.14	23.89	21.87	23.23	<b>13.54</b>	35.95
RMSE ↓	108.50	159.17	98.67	98.10	69.83	78.62	87.99	91.72	<b>57.71</b>	17.36
Flowering stage										
MAE ↓	188.52	181.69	170.10	210.05	172.04	120.22	155.27	109.41	<b>100.47</b>	41.60
MAE <sub>rel</sub> ↓	21.98	17.96	16.39	19.44	16.82	13.49	14.98	10.94	<b>10.57</b>	35.51
RMSE ↓	247.48	230.96	207.74	238.15	204.77	150.98	189.61	135.87	<b>129.09</b>	36.96

fusion block. The relative improvement is on average 33% but it can reach up 41% in some of the cases. Although there are some gains using the individual modules, the results demonstrate combination of the three modules and the fusion block can significantly improve the predictions.

Here, all learning-based models are trained on the whole training set (mix of early and flowering stage data). We report the result on the whole testing set ('Overall'), as well as on the 'Early stage' set, and the 'Flowering stage' set, separately. Here, '3DVI' and 'DGCNN' achieve the second and third-best performance. '3DVI' is more accurate at the early stage, while both '3DVI' and 'DGCNN' work similarly at the flowering stage. Notably, 'Ours' is superior at both the early and the flowering stage.

To further demonstrate that our BioNet learns structure features and has generalization ability, we perform a cross-validation experiment using our two developmental time points. To do this, the dataset was split into two early and flowering stages categories. We used 2/3 of the point clouds of each category as training and 1/3 as testing data. We train our model on either early or flowering stage data, and test our model on both temporal stages separately (Table 2). In this sub-experiment, baseline methods are second- ('3DVI') and third-best ('DGCNN'). The two learning-based methods ('Ours' and 'DGCNN') are trained on the early or flowering stage only and tested on the two stages separately. Because '3DVI' does not need to be trained, its performance is only driven by the test set. For models trained and tested on the same growing stage, both 'Ours' and 'DGCNN' achieve competitive results compared to the traditional density-based method '3DVI'. However, for the model trained on the early stage and tested on the flowering stage, or the model trained on the flowering stage

and tested on the early stage, 'DGCNN' failed to achieve good results. In contrast, our BioNet showed results superior to both baseline methods to predict biomass at the flowering stage, regardless of the training set. The experiment demonstrates that our BioNet learns structure features and has useful generalization abilities.

Note that, in Table 2, the model trained on the early stage and tested on the flowering stage achieves the best performance. However, the model trained on the flowering stage and tested on the early stage only achieves the second-best performance. We suspect that the point cloud for a plot at the early stage has a stronger and clearer structure than at the flowering stage. Feature extracted from the early stage can contribute to the prediction at the flowering stage. However, we hypothesize that at the flowering stage, canopy occlusion limits distinguishable features, reducing the accuracy of our model when predicting another stage.

#### 4.4. Ablation study

The main body of our BioNet is based on a multi-layer perceptron (MLP) network [39] with three proposed modules. When combined, the completion ('Our<sub>c</sub>'), regularization ('Our<sub>r</sub>'), and projection ('Our<sub>p</sub>') modules achieve superior performance on our real-world dataset. To study the contribution of each one of these modules to the final performance of BioNet, we investigate their individual impact on prediction accuracy (Table 1). First, adding any one of our three modules significantly improves the predicted result. Second, by fusing the feature vectors with an attention-based fusion block, the accuracy of the predicted biomass was improved by  $\geq 35\%$  in terms of MAE<sub>rel</sub>, compared to only using one individual module. This is likely because BioNet benefits from the combination of representa-

Table 2. Quantitative comparisons of biomass prediction methods on our real dataset. The best results are shown in **bold**. We test the generalization ability of the top-3 methods Ours, 3DVI, and the DGCNN. We first split the dataset into two categories, the early stage and the flowering stage. We use 2/3 of point clouds of each category as the training set and 1/3 as the testing set. The two learning-based methods are trained on the early or flowering stage only, and tested on the two stages separately.

Training	Testing	Early stage			Flowering stage		
		DGCNN [48]	3DVI [19]	Ours	DGCNN [48]	3DVI [19]	Ours
Early stage	MAE ↓	101.75	62.14	<b>50.53</b>	592.91	147.64	<b>121.86</b>
	MAE <sub>rel</sub> ↓	24.71	16.33	<b>13.37</b>	65.77	17.09	<b>13.70</b>
	RMSE ↓	116.05	74.75	<b>62.18</b>	613.57	181.82	<b>226.45</b>
Flowering stage	MAE ↓	563.11	<b>62.14</b>	114.46	133.28	147.64	<b>111.13</b>
	MAE <sub>rel</sub> ↓	147.63	<b>16.33</b>	29.50	11.48	17.09	<b>9.87</b>
	RMSE ↓	566.32	<b>74.75</b>	124.46	150.15	181.82	<b>145.65</b>

Table 3. Ablation on the modules of our BioNet. The best results are shown in **bold**. To demonstrate the effectiveness of each module in our BioNet, we test in the following cases: 1) main body (a point feature block in Fig. 3 + a MLP block in Fig. 2) without the three modules; 2) completion + regularization module; 3) completion + projection module; and 4) regularization + projection module. In case 2/3/4, we only have two modules that extract feature vectors feeding to the fusion block. Since our fusion block needs three feature vectors, we concatenate feature vectors output by the chosen two modules directly. Each model is tested on the whole testing set (Overall).

Settings	Case 1	Case 2	Case 3	Case 4	Ours
Completion	×	✓	✓	×	✓
Regularisation	×	✓	×	✓	✓
Projection	×	×	✓	✓	✓
MAE ↓	146.42	102.55	94.67	98.27	<b>71.23</b>
MAE <sub>rel</sub> ↓	35.11	20.65	17.77	20.66	<b>12.08</b>
RMSE ↓	198.65	128.13	119.86	124.65	<b>99.33</b>

Table 4. Ablation on the fusion block of our BioNet. To demonstrate the effectiveness of the fusion block, we use  $\mathbf{v}_r, \mathbf{v}_p$ , or  $\mathbf{v}_c$  as Query, Key, or Value in Eq. (5) respectively. We adopt Fuse( $\mathbf{v}_r, \mathbf{v}_p, \mathbf{v}_c$ ) in our BioNet for its best performance.

Settings	MAE ↓	MAE <sub>rel</sub> ↓	RMSE ↓
concatenate	96.15	19.36	125.17
Fuse( $\mathbf{v}_r, \mathbf{v}_p, \mathbf{v}_c$ )	<b>71.23</b>	<b>12.08</b>	<b>99.33</b>
Fuse( $\mathbf{v}_p, \mathbf{v}_r, \mathbf{v}_c$ )	77.82	12.96	110.72
Fuse( $\mathbf{v}_r, \mathbf{v}_c, \mathbf{v}_p$ )	71.27	12.35	104.48
Fuse( $\mathbf{v}_c, \mathbf{v}_r, \mathbf{v}_p$ )	74.28	13.78	100.11
Fuse( $\mathbf{v}_c, \mathbf{v}_p, \mathbf{v}_r$ )	76.28	13.74	110.77
Fuse( $\mathbf{v}_p, \mathbf{v}_c, \mathbf{v}_r$ )	83.66	14.35	125.23

tions learnt by these three modules. The improved performance of our full model shows that all three modules and the fusion block contribute to final biomass predictions.

To further demonstrate the advantage of our three modules, we test four module combinations: 1) main body (MB) alone; 2) MB + completion module + regularization module; 3) MB + completion module + projection module; and 4) MB + regularization module + projection module. The

results of this experiment are shown in Table 3. Since our fusion block needs three feature vectors, we directly concatenate feature vectors output by the two chosen modules. Each model is tested on the whole testing set (Overall).

To further demonstrate the advantage of the fusion block, we concatenate the three feature vector  $\mathbf{v}_c, \mathbf{v}_r$  and  $\mathbf{v}_p$  directly and then feed that to the MLP block. Without the fusion block, the model achieves 96.15/19.36%/125.17 in terms of 'MAE/MAE<sub>rel</sub>/RMSE'. In addition, we use  $\mathbf{v}_r, \mathbf{v}_p$ , or  $\mathbf{v}_c$  as Query, Key, or Value in Eq. (5) respectively. We test the following cases: 1) Fuse( $\mathbf{v}_r, \mathbf{v}_p, \mathbf{v}_c$ ) (the setting of our fusion block); 2) Fuse( $\mathbf{v}_p, \mathbf{v}_r, \mathbf{v}_c$ ); 3) Fuse( $\mathbf{v}_r, \mathbf{v}_c, \mathbf{v}_p$ ); 4) Fuse( $\mathbf{v}_c, \mathbf{v}_r, \mathbf{v}_p$ ); 5) Fuse( $\mathbf{v}_c, \mathbf{v}_p, \mathbf{v}_r$ ); and 6) Fuse( $\mathbf{v}_p, \mathbf{v}_c, \mathbf{v}_r$ ). The results shown in Table 4 show that our attention fusion method in any order outperforms the direct concatenation strategy.

## 5. Conclusion

This paper aims at estimating the above-ground biomass (AGB) using LiDAR. The agricultural dataset collected for this study is made available to the community. Unlike previous methods which predicted biomass based on the density of the point cloud or plant height, we propose a biomass prediction network (BioNet) that also considers plant structure via a double attention-based fusion block, a completion module and a point cloud representation module. BioNet achieved higher AGB predictions than the previous state-of-the-art method (3DVI) and is therefore proposed as the new reference. In the future, we aim to introduce more sensors in our system to improve prediction accuracy.

## 6. Acknowledgement

This work was partially funded by CSIRO's Machine Learning and Artificial Intelligence Future Science Platform. The dataset was collected as part of Grains Research and Development Corporation (GRDC) project #CSP00168. This research was enabled by the Australian Plant Phenomics Facility (APPF), which itself is funded by the Australian Government under the National Collaborative Research Infrastructure Strategy (NCRIS).



## References

- [1] Helge Aasen, Andreas Burkart, Andreas Bolten, and Georg Bareth. Generating 3d hyperspectral information with lightweight uav snapshot cameras for vegetation monitoring: From camera calibration to quality assurance. *ISPRS Journal of Photogrammetry and Remote Sensing*, 108:245–259, 2015.
- [2] Panos Achlioptas, Olga Diamanti, Ioannis Mitliagkas, and Leonidas Guibas. Learning representations and generative models for 3d point clouds. In *International conference on machine learning*, pages 40–49. PMLR, 2018.
- [3] Michael Alonzo, Roman J Dial, Bethany K Schulz, Hans-Erik Andersen, Eric Lewis-Clark, Bruce D Cook, and Douglas C Morton. Mapping tall shrub biomass in alaska at landscape scale using structure-from-motion photogrammetry and lidar. *Remote Sensing of Environment*, 245:111841, 2020.
- [4] Asher Bender, Brett Whelan, and Salah Sukkarieh. A high-resolution, multimodal data set for agricultural robotics: A ladybird’s-eye view of brassica. *Journal of Field Robotics*, 37(1):73–96, 2020.
- [5] Dimitri P Bertsekas. A distributed asynchronous relaxation algorithm for the assignment problem. In *1985 24th IEEE Conference on Decision and Control*, pages 1703–1704. IEEE, 1985.
- [6] Jaime Caballer Revenga, Katerina Trepekli, Stefan Oehmcke, Fabian Gieseke, Christian Igel, Rasmus Jensen, and Thomas Friborg. Prediction of above ground biomass and c-stocks based on uav-lidar, multispectral imagery and machine learning methods. In *EGU General Assembly Conference Abstracts*, pages EGU21–15708, 2021.
- [7] Tatiana Fernanda Canata, Maurício Martello, Leonardo Felipe Maldaner, Jadir de Souza Moreira, and José Paulo Molin. 3d data processing to characterize the spatial variability of sugarcane fields. *Sugar Tech*, pages 1–11, 2021.
- [8] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. End-to-end object detection with transformers. In *European Conference on Computer Vision*, pages 213–229. Springer, 2020.
- [9] U. DESA. Population division (2019). how certain are the united nations global population projections? In *Population Facts*, number 6. United Nations, Department of Economic and Social Affairs, 2019.
- [10] Jan UH Eitel, Troy S Magney, Lee A Vierling, Tabitha T Brown, and David R Huggins. Lidar based biomass and crop nitrogen estimates for rapid, non-destructive assessment of wheat nitrogen status. *Field Crops Research*, 159:21–32, 2014.
- [11] Jan UH Eitel, Troy S Magney, Lee A Vierling, Heather E Greaves, and Guang Zheng. An automated method to quantify crop height and calibrate satellite-derived biomass using hypertemporal lidar. *Remote Sensing of Environment*, 187:414–422, 2016.
- [12] Haoqiang Fan, Hao Su, and Leonidas J Guibas. A point set generation network for 3d object reconstruction from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 605–613, 2017.
- [13] Maria G Salas Fernandez, Yin Bao, Lie Tang, and Patrick S Schnable. A high-throughput, field-based phenotyping technology for tall biomass crops. *Plant physiology*, 174(4):2008–2022, 2017.
- [14] Jun Fu, Jing Liu, Haijie Tian, Yong Li, Yongjun Bao, Zhiwei Fang, and Hanqing Lu. Dual attention network for scene segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.
- [15] Robin Gebbers, Detlef Ehlert, and Rolf Adamek. Rapid mapping of the leaf area index in agricultural crops. *Agronomy Journal*, 103(5):1532–1541, 2011.
- [16] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015.
- [17] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.
- [18] Heather E Greaves, Lee A Vierling, Jan UH Eitel, Natalie T Boelman, Troy S Magney, Case M Prager, and Kevin L Griffin. Estimating aboveground biomass and leaf area of low-stature arctic shrubs with terrestrial lidar. *Remote Sensing of Environment*, 164:26–35, 2015.
- [19] Jose A Jimenez-Berni, David M Deery, Pablo Rozas-Larraondo, Anthony Tony G Condon, Greg J Rebetzke, Richard A James, William D Bovill, Robert T Furbank, and Xavier RR Sirault. High throughput determination of plant height, ground cover, and above-ground biomass in wheat with lidar. *Frontiers in plant science*, 9:237, 2018.
- [20] Kasper Johansen, Mitchell JL Morton, Yoann Malbeteau, Bruno Aragon, Samer Al-Mashharawi, Matteo G Ziliani, Yoseline Angel, Gabriele Fiene, Sonia Negrao, Magdi AA Mousa, et al. Predicting biomass and yield in a tomato phenotyping experiment using uav imagery and random forest. *Frontiers Artif. Intell.*, 3:28, 2020.
- [21] Diederick P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.
- [22] Jiaxin Li, Ben M Chen, and Gim Hee Lee. So-net: Self-organizing network for point cloud analysis. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9397–9406, 2018.
- [23] Wang Li, Zheng Niu, Ni Huang, Cheng Wang, Shuai Gao, and Chaoyang Wu. Airborne lidar technique for estimating biomass components of maize: A case study in zhangye city, northwest china. *Ecological indicators*, 57:486–496, 2015.
- [24] Liu Liu, Dylan Campbell, Hongdong Li, Dingfu Zhou, Xibin Song, and Ruigang Yang. Learning 2d-3d correspondences to solve the blind perspective-n-point problem. *arXiv preprint arXiv:2003.06752*, 2020.
- [25] Liu Liu, Hongdong Li, Haodong Yao, and Ruyi Zha. Pluckernet: Learn to register 3d line reconstructions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1842–1852, June 2021.
- [26] Minghua Liu, Lu Sheng, Sheng Yang, Jing Shao, and Shi-Min Hu. Morphing and sampling network for dense point cloud completion. In *Proceedings of the AAAI Conference*

- on *Artificial Intelligence*, volume 34, pages 11596–11603, 2020.
- [27] E Louise Loudermilk, J Kevin Hiers, Joseph J O’Brien, Robert J Mitchell, Abhinav Singhania, Juan C Fernandez, Wendell P Cropper, and K Clint Slatton. Ground-based lidar: a novel approach to quantify fine-scale fuelbed characteristics. *International Journal of Wildland Fire*, 18(6):676–685, 2009.
- [28] Daniel Maturana and Sebastian Scherer. Voxnet: A 3d convolutional neural network for real-time object recognition. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 922–928. IEEE, 2015.
- [29] Liyuan Pan, Shah Chowdhury, Richard Hartley, Miaomiao Liu, Hongguang Zhang, and Hongdong Li. Dual pixel exploration: Simultaneous depth estimation and image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4340–4349, June 2021.
- [30] Martin AJ Parry, Matthew Reynolds, Michael E Salvucci, Christine Raines, P John Andralojc, Xin-Guang Zhu, G Dean Price, Anthony G Condon, and Robert T Furbank. Raising yield potential of wheat. ii. increasing photosynthetic capacity and efficiency. *Journal of experimental botany*, 62(2):453–467, 2011.
- [31] AJD Pask, Julian Pietragalla, DM Mullan, and MP Reynolds. *Physiological breeding II: a field guide to wheat phenotyping*. Cimmyt, 2012.
- [32] Joan Ramon Rosell Polo, Ricardo Sanz, Jordi Llorens, Jaume Arnó, Alexandre Escolà, Manel Ribes-Dasi, Joan Masip, Ferran Camp, Felip Gràcia, Francesc Solanelles, et al. A tractor-mounted scanning lidar for the non-destructive measurement of vegetative volume and surface area of tree-row plantations: A comparison with conventional destructive measurements. *Biosystems Engineering*, 102(2):128–134, 2009.
- [33] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 652–660, 2017.
- [34] Charles R Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. *arXiv preprint arXiv:1706.02413*, 2017.
- [35] Greg J Rebetzke, Karine Chenu, Ben Biddulph, Carina Moeller, Dave M Deery, Allan R Rattey, Dion Bennett, Ed G Barrett-Lennard, and Jorge E Mayer. A multisite managed environment facility for targeted trait and germplasm phenotyping. *Functional Plant Biology*, 40(1):1–13, 2012.
- [36] Wouter Saeys, Bart Lenaerts, Geert Craessaerts, and Josse De Baerdemaeker. Estimation of the crop density of small grains using lidar sensors. *Biosystems Engineering*, 102(1):22–30, 2009.
- [37] Muhammad Sarmad, Hyunjoo Jenny Lee, and Young Min Kim. RL-gan-net: A reinforcement learning agent controlled gan network for real-time point cloud shape completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5898–5907, 2019.
- [38] Shangpeng Sun, Changying Li, Andrew H Paterson, Yu Jiang, Rui Xu, Jon S Robertson, John L Snider, and Peng W Chee. In-field high throughput phenotyping and cotton plant growth analysis using lidar. *Frontiers in Plant Science*, 9:16, 2018.
- [39] H Taud and JF Mas. Multilayer perceptron (mlp). In *Geomatic Approaches for Modeling Land Change Scenarios*, pages 451–455. Springer, 2018.
- [40] Lyne P Tchapmi, Vineet Kosaraju, Hamid Rezatofghi, Ian Reid, and Silvio Savarese. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 383–392, 2019.
- [41] Jelle ten Harkel, Harm Bartholomeus, and Lammert Kooistra. Biomass and crop height estimation of different crops using uav-based lidar. *Remote Sensing*, 12(1):17, 2020.
- [42] Nora Tilly, Helge Aasen, and George Bareth. Fusion of plant height and vegetation indices for the estimation of barley biomass. *Remote Sensing*, 7(9):11449–11480, 2015.
- [43] Nora Tilly, Dirk Hoffmeister, Qiang Cao, Shanyu Huang, Victoria Lenz-Wiedemann, Yuxin Miao, and Georg Bareth. Multitemporal crop surface models: accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. *Journal of Applied Remote Sensing*, 8(1):083671, 2014.
- [44] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *arXiv preprint arXiv:1706.03762*, 2017.
- [45] James DC Walter, James Edwards, Glenn McDonald, and Haydn Kuchel. Estimating biomass and canopy height with lidar for field crop breeding. *Frontiers in plant science*, 10:1145, 2019.
- [46] Cheng Wang, Sheng Nie, Xiaohuan Xi, Shezhou Luo, and Xiaofeng Sun. Estimating the biomass of maize with hyperspectral and lidar data. *Remote Sensing*, 9(1):11, 2017.
- [47] Xiaogang Wang, Marcelo H Ang Jr, and Gim Hee Lee. Cascaded refinement network for point cloud completion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 790–799, 2020.
- [48] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *Acm Transactions On Graphics (tog)*, 38(5):1–12, 2019.
- [49] Wenxuan Wu, Zhongang Qi, and Li Fuxin. Pointconv: Deep convolutional networks on 3d point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9621–9630, 2019.
- [50] Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1912–1920, 2015.
- [51] Mingye Xu, Zhipeng Zhou, and Yu Qiao. Geometry sharing network for 3d point cloud classification and segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 12500–12507, 2020.

- [52] Yaoqing Yang, Chen Feng, Yiru Shen, and Dong Tian. Foldingnet: Interpretable unsupervised learning on 3d point clouds. *arXiv preprint arXiv:1712.07262*, 2(3):5, 2017.
- [53] Wentao Yuan, Tejas Khot, David Held, Christoph Mertz, and Martial Hebert. Pcn: Point completion network. In *3D Vision (3DV), 2018 International Conference on*, 2018.
- [54] Yujin Zhao, Xiaoliang Liu, Yang Wang, Zhaoju Zheng, Shuxia Zheng, Dan Zhao, and Yongfei Bai. Uav-based individual shrub aboveground biomass estimation calibrated against terrestrial lidar in a shrub-encroached grassland. *International Journal of Applied Earth Observation and Geoinformation*, 101:102358, 2021.