

Deep Learning Method for Height Estimation of Sorghum in the Field Using LiDAR

Matthew Waliman and Avidesh Zakhori; University of California, Berkeley; Berkeley, CA

Abstract

Automatic tools for plant phenotyping have received increased interest in recent years due to the need to understand the relationship between plant genotype and phenotype. Building upon our previous work, we present a robust, deep learning method to accurately estimate the height of biomass sorghum throughout the entirety of its growing season. We mount a vertically oriented LiDAR sensor onboard an agricultural robot to obtain 3D point clouds of the crop fields. From each of these 3D point clouds, we generate a height contour and density map corresponding to a single row of plants in the field. We then train a multiview neural network in order to estimate plant height. Our method is capable of accurately estimating height from emergence through canopy closure. We extensively validate our algorithm by performing several ground truthing campaigns on biomass sorghum. We have shown our proposed approach to achieve an absolute height estimation error of 7.47% using ground truth data obtained via conventional breeder methods on 2715 plots of sorghum with varying genetic strains and treatments.

1. Introduction

Understanding the relationship between genotypes and phenotypes of plants is essential to the optimization of the biofuel production pipeline. By accurately characterizing physiological traits of plants, it is possible to determine connections between plant gene sequences and biomass yield. Specifically, sorghum has been demonstrated to be a suitable source of fuel in practice [1, 2]. To determine the genotype-phenotype map for sorghum, rapid phenotyping methods are necessary for efficient data collection. Currently, plant phenotyping is done by hand, presenting a bottleneck in the growing pipeline [3]. Manual collection of physiological traits is labor intensive, time consuming, error prone and does not provide nearly enough data to supplement the genotypes available. Therefore, it is essential to develop systems that automatically, accurately and efficiently phenotype in situ plants.

Plant height is of particular interest to geneticists and breeders as it is a good indicator of biomass yield [4]. In this paper, we propose a deep learning method that estimates biomass sorghum height at all stages of the growing period. Specifically, we estimate the height of plant stalks using data collected in field conditions measured by a scanning LiDAR mounted onto a mobile robotic platform, which traverses through rows of densely positioned plants.

2. Related Work

Our method is heavily influenced by our previous work and its shortcomings. Specifically, in [5] we develop two methods of height estimation: a peak finding method for LiDAR scans collected before canopy closure and a percentile method for after

canopy closure. The peak based method assumes that peaks in the height contour correspond to individual plant stems in the crop row. While this assumption holds true before canopy closure, after the canopy closes, individual peaks in the height contour are no longer representative of individual stems. In order to be able to estimate height for plants after the canopy closes, we developed a second percentile based method that offers comparable results to the peak method across all plant heights [5].

In this work, we seek to develop a single algorithm that is capable of accurately estimating biomass sorghum height at all stages of the growing period, even at heights it has not yet previously seen. In order to accomplish this, we turn to deep learning methods, which have had recent successes in the domain of plant phenotyping [6, 7, 8].

Many deep learning methods rely on Convolutional Neural Networks (CNNs) which serve as encoders to move 2D images into a latent space. Due to the irregular format of point clouds it is not possible to use 2D CNNs natively. Researchers have offered many solutions such as transforming the point cloud to a regular structure such as a voxel structure [9]. This however renders the data unnecessarily voluminous and with unavoidable computational cost of 3D convolutions. Another option is to develop a special type of network that is capable of directly consuming the point clouds such as in [10]. While a promising solution, point clouds in this formulation are invariant to transformations which is not applicable in the case of our data. A third approach is to project the 3D shape to multiple 2D views. Su et. al find that a classifier trained on multiple 2D views of a 3D shape actually outperform a classifier trained directly from the 3D representation [11].

In this work we propose a multiview approach that reduces 3D data to 2D by projecting along various dimensions in order to be able to use 2D convolutions. In addition to a height contour from our previous work that captures a $y-z$ projection of the point cloud we project the point cloud into the $x-y$ plane to create a density map. The rest of the paper is organized as follows: Section 3 describes our experimental setup, in Section 4, we outline our proposed approach; Section 5 includes experimental results. Finally in Section 6, we discuss implications of our findings and future work.

3. Setup

In the 2018 growing season, four field campaigns were carried out on sorghum fields at the University of Illinois Urbana Champaign and University of California Davis. The surveys were spaced at regular intervals between June 8th and October 10th, offering a widespread sampling of sorghum height at various stages of its growing season.

Each field is subdivided into plots measuring 3 m \times 3 m

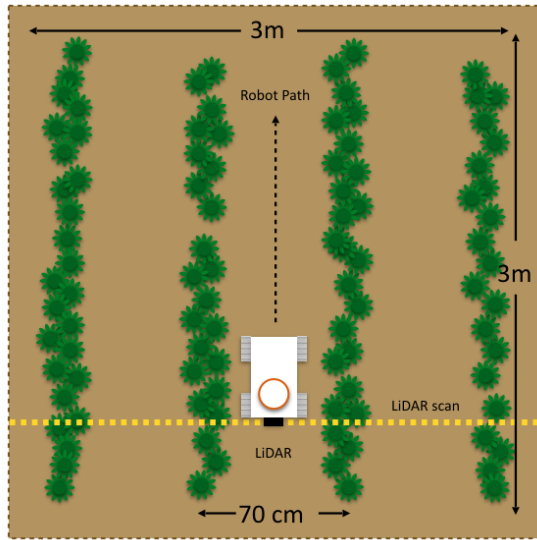


Figure 1: A top down view of a single plot.

shown in Figure 1. The Illinois and Davis fields contained 864 plots and 250 plots respectively. Each sorghum plot contains either a unique genetic strain or unique microbial treatment. Since geneticists are primarily interested in comparing statistical parameters of phenotypes across plots, our goal is to estimate the average stem height for each plot.

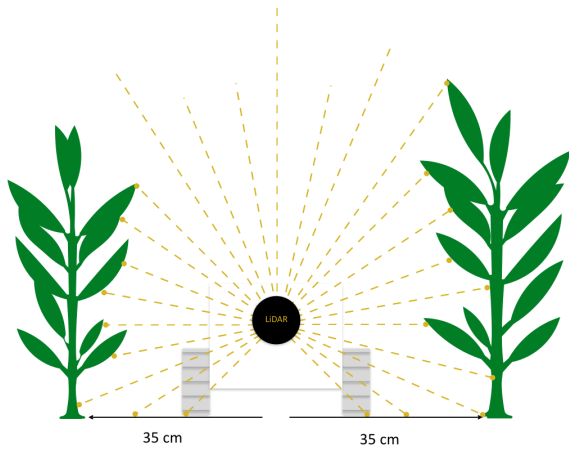


Figure 2: Back view of the sensor set up on the agricultural robot.

As seen in Figure 1, each plot consists of four rows of sorghum, planted by precision planter 70 cm apart from each other. To collect data, a vertical facing Hokuyo UST-10LX is mounted onto mobile robotic platform as shown in Figure 2, traversing the center row of each plot.

To develop an accurate and robust method for obtaining ground truth for all field conditions and varying plant growth stages, we conducted several campaigns to validate our algorithm. To mimic the existing method of phenotype data collection, our team ground truthed 800 plots with varying genetic strains over five dates employing the ‘by-eye’ method currently employed by

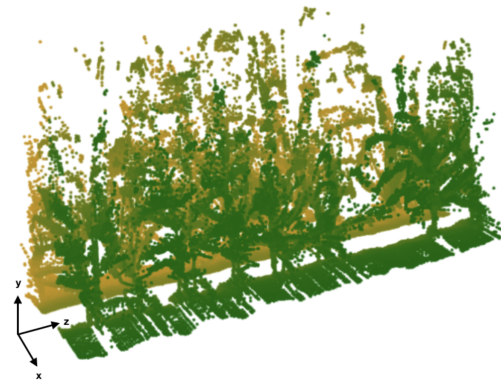


Figure 3: 3D point cloud for a single crop row shown with axes.

breeders. This entails placing one’s eye to the level of one of the rows and visualizing line that encompasses the majority of the tops of the plants. The measured distance from the ground was assigned to be the ground truth height estimate for the plot. The plants ranged from 5 cm to 400 cm in our experiments.

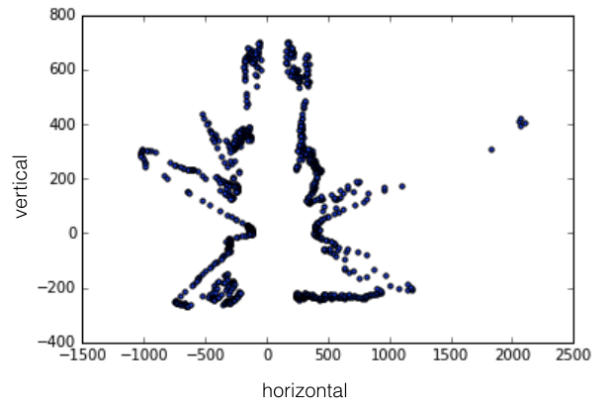


Figure 4: A single LiDAR scan.

4. Proposed Approach

In order to take advantage of the 3D information offered by a pointcloud while minimizing computational cost, we reproject the 3D pointcloud to two different 2D views. Our first projection to a height contour was proposed in [12] and was later used for height estimation in [5]. In order deal with challenging situations such as canopy closure or occluding objects, we additionally reproject the pointcloud into the $x - y$ dimension.

Our proposed method takes in two sets of inputs derived from a single 3D point cloud, shown in Figure 3. By convention, we choose the z axis to denote the direction of motion of the robot. For the $x - y$ projection, we choose a density map representation rather than a contour so that the density map encoder can learn complex image features that can help disambiguate noise arising from the height contour.

We refer to a single rotation of the LiDAR as a single scan, shown in Figure 4. As the robot traverses a row of plants, it collects scans at regular intervals, shown in Figure 5. We range limit

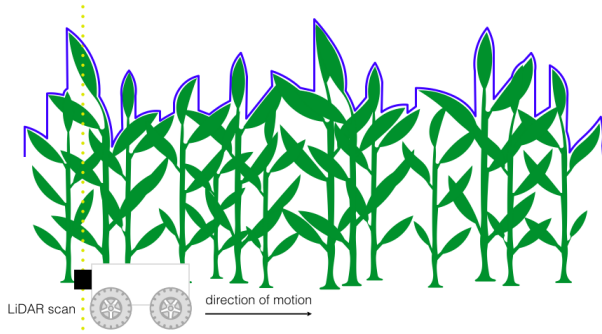


Figure 5: Agricultural robot travels across a row, collecting vertical LiDAR scans. The blue line represents the height contour generated by concatenating these scans in the direction of motion of the robot.

each scan to 70 cm on either side of the robot in order to limit the data to the first row of plots. To generate a height contour, we find the maximum y -value for each scan and concatenate these maximum heights together over the entire run of the data collection to create a height map as shown in Figure 6.

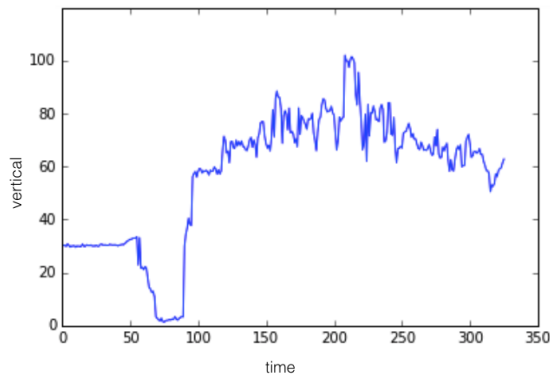


Figure 6: Height contour for a single row.

From the 3D pointcloud, we also project all the points to the $x-y$ plane. We then rasterize these points in the $x-y$ plane into a 1 cm by 1 cm grid, where the brightness of each pixel corresponds to the number of total points found in each 1x1 cm pixel as seen in Figure 7. We refer to this as a "density map". As a final post-processing step, it is necessary to standardize the input shapes so that they can be fed into the network. While the length of the height contour vector may vary depending on scan length, an empirical investigation reveals that there are no scans that exceed 512 time-points. Thus we crop the height contour to 512 timepoints resulting in a 1×512 vector. Similarly, we find that it is uncommon for sorghum plants to exceed 4 meters in height. Combined with the knowledge that crop rows are planted 70 cm apart, we standardize the density map to a $400\text{cm} \times 140\text{cm}$ area around the center of the LiDAR sensor.

We propose a dual pronged network able to leverage information from both the $y-z$ and $x-y$ planes to estimate height as seen in 8. Our network takes two inputs, a height contour (1×512 vector) and a density map (400×140 array) which are derived

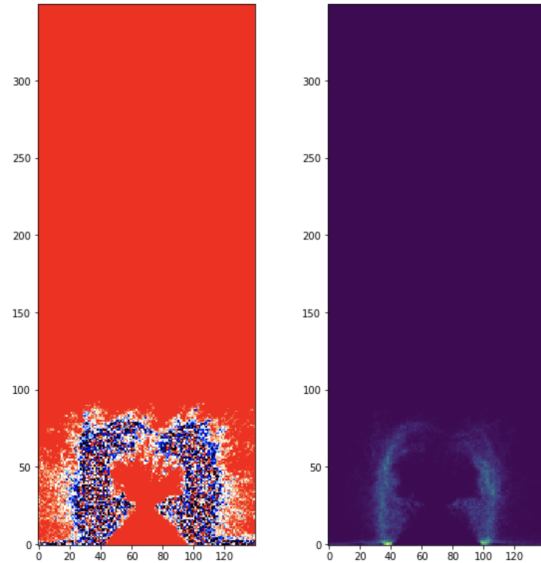


Figure 7: Density Map for a single crop row shown in two different colormaps.

from reprojections of a pointcloud of a crop row to the yz plane and the xy plane respectively. The height contour encoder is composed a series of diminishing fully connected layers outputting a single height contour feature vector. The density map encoder is comprised of 3 convolutional modules. Each convolutional module is made up of two 3×3 convolutional layers, succeeded by a max-pooling layer which reduces the size of the output by half in each spatial dimension. The filter depth doubles in size for each successive module. We additionally apply dropout after the last convolutional module. The convolutional output is flattened, a second dropout layer is applied, and finally passes through single dense layer to create a density map feature vector. The density map and height contour feature vectors are concatenated and passed to the regression network which regresses the height estimate for the sample.

Our network is implemented in Tensorflow and was trained on a single GeForce GTX1080Ti. We trained on a data set of 2715 data points, with a random 1715-1000 train-test split. The network was trained in an end-to-end manner for 100 epochs using the Adam optimizer. Mean average percent error (MAPE) was used as our loss function:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{H_i - Y_i}{Y_i} \right| \quad (1)$$

where H_i is the predicted height and Y_i is the ground truth height.

5. Results

Figure 9 shows a tight correlation between the estimated height using our proposed method, and 'by eye' ground truth method for a wide range of crop heights. Averaged over 2715 measurement, our deep learning method achieves an absolute average error of 7.47% for plants between 5 cm and 400 cm. As seen in Table 1, the corresponding error rates for our previously approaches, the Peak and Percentile Method [5] are 8.46% and 8.51%, respectively. Specifically, table 1 shows that our deep

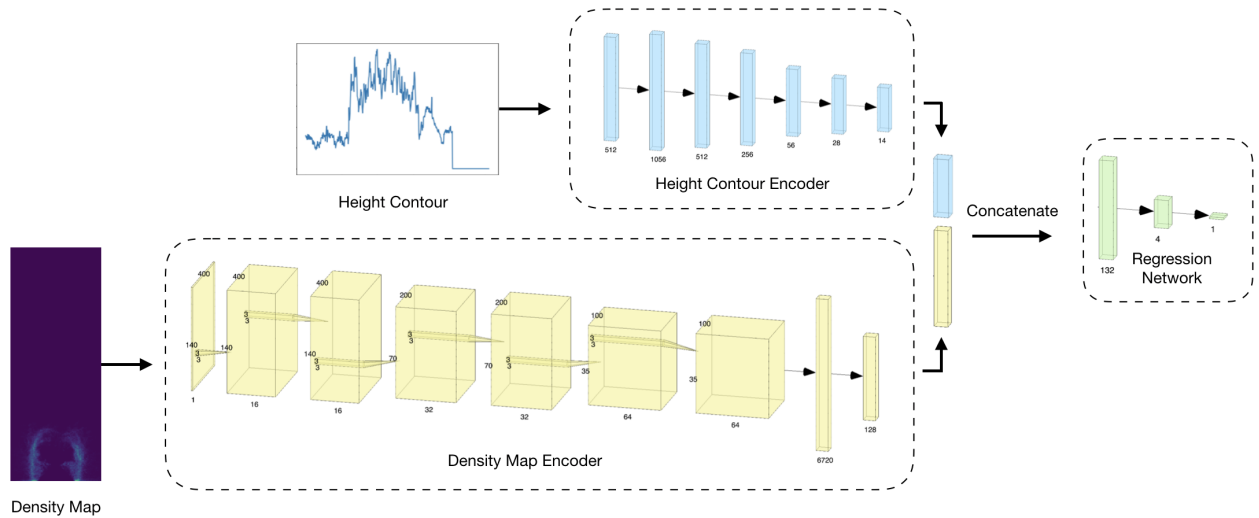


Figure 8: Diagram of proposed pipeline.

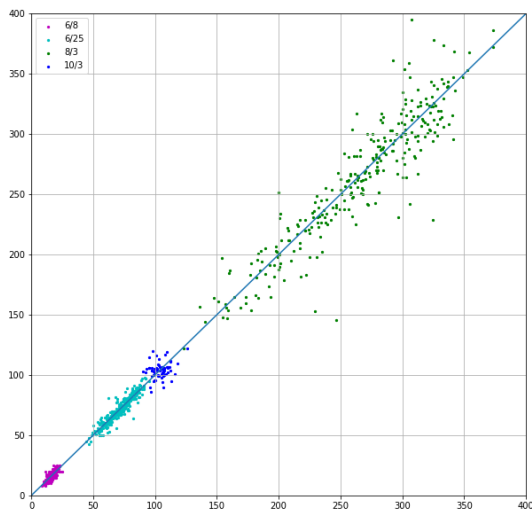


Figure 9: Measured height is on the x-axis and the estimated height is on the y-axis. The data shown here comes from four ground truthing campaigns conducted at sorghum fields at the University of Illinois Urbana Champaign and University of California Davis. The campaigns were spaced evenly throughout the growing period to capture the greatest height range of the sorghum plants.

learning method significantly outperforms both of our previous approaches, for heights ranging between 0-30 cm and 50-100 cm. Table 2 compares the accuracy of our method versus previous approaches for different data sets captured at different times. As seen, for all dates except for 10/3, our method achieves a smaller absolute error than peak and percentile methods.

The variance of the deep learning method is far smaller than both the Peak and Percentile methods across all datasets as shown

in Table 3. We show the error distribution for each dataset in Figure 10. Each of the plots show the spread of error for a specific height range. The deep learning method results in smaller bias and lower variance than both of our previous methods.

	0-30 cm	50-100 cm	100-400 cm	Average
Peak	12.6	5.1	8.3	8.46
Percentile	12.9	5.7	6.6	8.51
Deep Learning	10.9	5.0	7.1	7.47

Table 1: Mean Absolute Percent Error For Each Height Range.

	6/8	6/25	8/3	10/3	Average
Peak	12.6	5.2	11.6	8.3	8.46
Percentile	12.9	5.6	6.2	6.6	8.51
Deep Learning	10.9	5.0	5.5	7.2	7.47

Table 2: Mean Absolute Percent Error For Each Collection Date.

	6/8	6/25	8/3	10/3
Peak	0.013	0.022	0.006	0.003
Percentile	0.135	0.056	0.062	0.100
Deep Learning	0.008	0.001	0.005	0.002

Table 3: Error Variance For Each Collection Date.

6. Discussion

We presented a deep multiview fusion method for estimating the average height of a plot of sorghum from LiDAR scans obtained from an agricultural robot. Evaluation on our datasets verify that our method significantly outperform current methods and is capable of generalizing to plants at all stages of the growing season ranging in height from 5 cm to 400 cm.

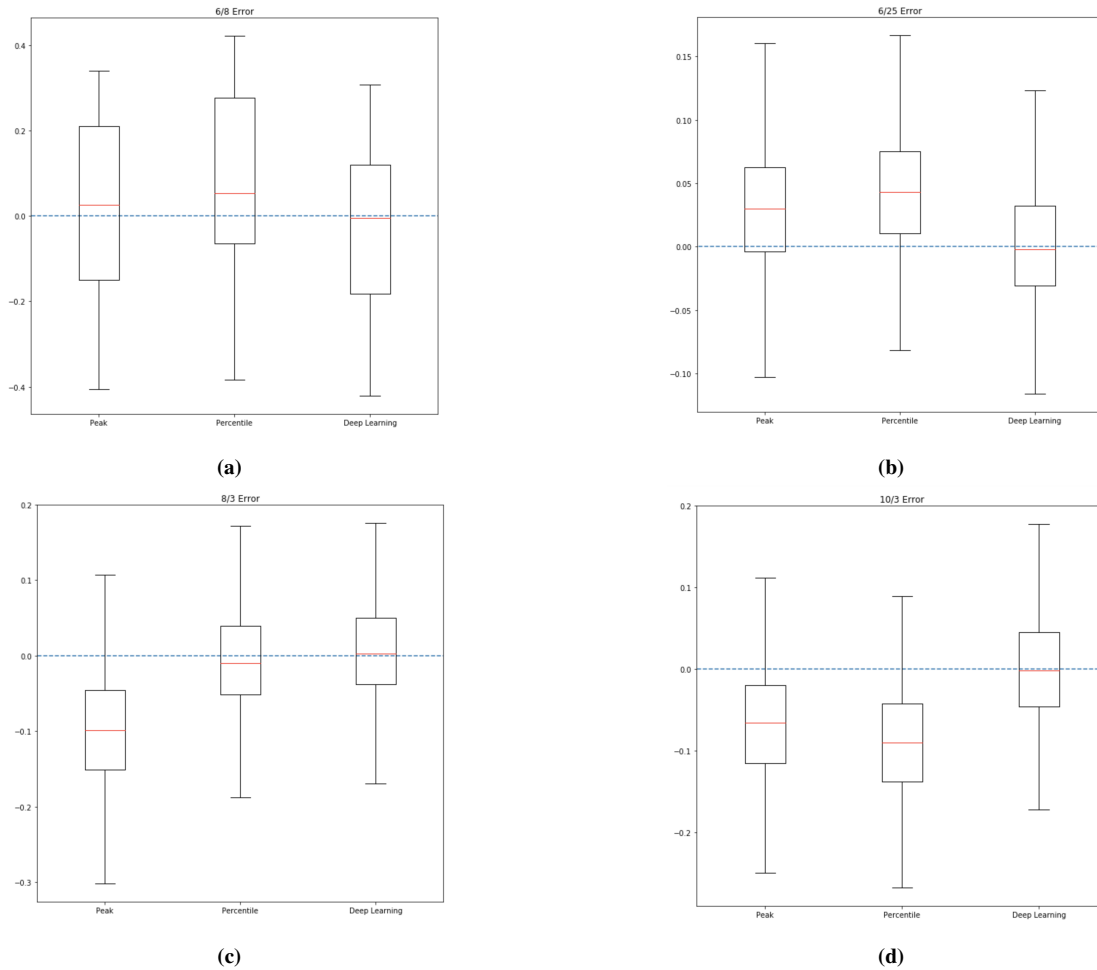


Figure 10: Error distributions comparison of Peak, Percentile and Deep Learning methods across each of the collected datasets. The red line shows the median error for each method and the box shows the upper and lower quartile ranges. The dashed blue line indicates the zero error axis. (a) 6/8 data set for plant heights in the 5-30 cm range. (b) 6/25 data set for plant heights in the 30-100 cm range. (c) 8/3 data set for plant heights in the 130-400 cm. (d) 10/3 data set for plant heights in the 80-130 cm.

An exciting direction for future work would be to adapt Pointnet [10] to process pointsets that are not transformation invariant. Such a network could be trained for segmentation of individual plants and height and width measurement on a per plant basis.

7. Acknowledgements

The information, data, or work presented herein was funded in part by the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy, under Award Number DE-AR0000598. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

References

- [1] S Prasad, Anoop Singh, N Jain, and HC Joshi, "Ethanol production from sweet sorghum syrup for utilization as automotive fuel in india," *Energy & Fuels*, vol. 21, no. 4, pp. 2415 - 2420, 2007.
- [2] Mustafa Balat and Havva Balat, "Recent trends in global production and utilization of bio-ethanol fuel," *Applied energy*, vol. 86, no. 11, pp. 2273–2282, 2009.
- [3] M. Minervini, H. Scharr, and S. A. Tsafaris, "Image analysis: The new bottleneck in plant phenotyping [applications corner]," *IEEE Signal Processing Magazine*, vol. 32, no. 4, pp. 126–131, July 2015.
- [4] Reynolds, M. P., and Borlaug, N. E. Impacts of breeding on international collaborative wheat improvement. *J. Agric. Sci.* 144, 3. 2006
- [5] Waliman, Matthew, and Avidah Zakhor. "Height estimation of biomass sorghum in the field using LiDAR." *Electronic Imaging 2019*, no. 13 (2019): 137-1.
- [6] Aich, Shubhra, Anique Josuttes, Ilya Ovsyannikov, Keegan Strueby, Imran Ahmed, Hema Sudhakar Duddu, Curtis Pozniak, Steve Shirliff, and Ian Stavness. "Deepwheat: Estimating phenotypic traits from crop images with deep learning." In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 323-332. IEEE, 2018.
- [7] Ubbens, Jordan R., and Ian Stavness. "Deep plant phenomics: a deep learning platform for complex plant phenotyping tasks." *Frontiers in plant science* 8 (2017): 1190.

- [8] Namin, Sarah Taghavi, Mohammad Esmaeilzadeh, Mohammad Najafi, Tim B. Brown, and Justin O. Borevitz. "Deep phenotyping: deep learning for temporal phenotype/genotype classification." *Plant methods* 14, no. 1 (2018): 66.
- [9] Qi, Charles R., Hao Su, Matthias Nießner, Angela Dai, Mengyuan Yan, and Leonidas J. Guibas. "Volumetric and multi-view cnns for object classification on 3d data." In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 5648-5656. 2016.
- [10] Qi, Charles R., 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 (CVPR)*, pp. 652-660. 2017.
- [11] Su, Hang, Subhransu Maji, Evangelos Kalogerakis, and Erik Learned-Miller. "Multi-view convolutional neural networks for 3d shape recognition." In *Proceedings of the IEEE international conference on computer vision (ICCV)*, pp. 945-953. 2015
- [12] Jin, Jihui, Gefen Kohavi, Zhi Ji and Avidesh Zakhor. "Top Down Approach to Height Histogram Estimation of Biomass Sorghum in the Field." *Computational Imaging*, vol. 17 2018.

Author Biography

Matthew Waliman received his B.S. in Electrical Engineering and Computer Science from the University of California Berkeley in 2018. He is currently a research engineer at UCSF building deep learning models that segment brain and spinal cord volumes from MRI. His areas of interest include: 3D Computer Vision, Graphics, Optics, HCI and Augmented Reality.

Avidesh Zakhor is Qualcomm Chair and professor in EECS at U.C. Berkeley. Her research interests are in 3D computer vision, machine learning, signal and image processing. She has won a number of best paper awards from IEEE signal processing, circuits and systems, and solid state, and semiconductor manufacturing societies. She was a Hertz fellow, General Motors scholar, IEEE fellow, and received the Presidential Young Investigators (PYI) award in 1990. She was selected as electronic Imaging scientist of the year by SPIE in 2018.

JOIN US AT THE NEXT EI!

IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!



- **SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •**
- **INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •**

www.electronicimaging.org

