# Furrow Mapping of Sugarcane Billet Density Using Deep Learning and Object Detection

Jordan Scott
School of Engineering
and Built Environment
Griffith University
Nathan, QLD 4111, Australia
Email: jordan.scott@griffithuni.edu.au

Andrew Busch
School of Engineering
and Built Environment
Griffith University
Nathan, QLD 4111, Australia
Email: a.busch@griffith.edu.au

***Abstract*—Australia's sugar industry is currently undergoing significant hardships, due to global market contractions from COVID-19, increased crop forecasts from larger global producers, and falling oil prices. Current planting practices utilize inefficient mass-flow planting techniques, and no attempt to map the seed using machine vision has been made, to date, in order to understand the underlying problems. This paper investigates the plausibility of creating a labeled sugarcane billet dataset using a readily-available camera positioned beneath a planter and analysing this using a YOLOv3 network. This network resulted in a high mean average precision at intersect over union of 0.5 (mAP50) of 0.852 on test images, and was used to provide planting metrics by generating a furrow map.**

## I. Introduction

Sugarcane is predominantly planted in Australia using mass flow billet planting systems which can cause significant damage to the billets and uneven distribution of seed across the field [1]–[6]. To compensate, growers plant over five times the amount of seed compared to hand-planted whole stick planting methods [2], [7]. This is economically wasteful and can reduce annual yield due to weaker plants [4], [8], [9].

A study by MIT found that image augmentation techniques were useful for classifying billet damage [1] and in conjunction with deep learning and object detection, this can provide operators with useful, real-time metrics. This paper begins by investigating capturing data in the furrow using an off the shelf camera and then proposes a method for detecting and mapping sugarcane billets in the furrow.

The main contributions of this paper follow:

- An annotated billet dataset of images captured in the furrow.
- An accurate You Only Look Once (YOLO) model for the detection of billets.
- The development of a furrow map showing billet density.

## II. Background

### A. Machine Learning

Machine learning uses training data to build a mathematics model, by calculating weights, biases and loss, in order to make a prediction on unseen data [10]. Deep learning uses a multi-layer architecture to essentially stack non-linear models together, with each layer transforming the input representation into a higher-level representation at its output. The application of back-propagation in deep learning allows for the model to train by adjusting its internal parameters at each layer, based on the performance at the previous layer, starting with a loss function at the final output layer then working inward back through the layers [11], [12].

### B. Convolutional Neural Networks (CNNs)

CNNs are an implementation of the deep learning architecture, which use stacked convolutional layers in the learnable layers [10]. CNNs are very powerful for computer vision, as they are designed to cope very well with data in 3-dimensional arrays, such as an image consisting of RGB channels, using discrete convolution to apply feature maps [12], [13]. CNNs outperform other methods of image classification and object detection, when applied using supervised learning methods [14], [15].

### C. You Only Look Once (YOLO)

The YOLO model is a regression based CNN which uses 24 convolutional layers in its network architecture. A YOLO model takes an entire image into account when training and testing, which allows the CNN to encode contextual information regarding classes, as well as their features [16]. This allows the network to generalise better, when compared with other models. The YOLO network divides the image into a grid and predicts multiple bounding boxes and corresponding confidence for each class. This grid structure and YOLO's ability to use a single network evaluation allows for extremely fast inference at test time [16].

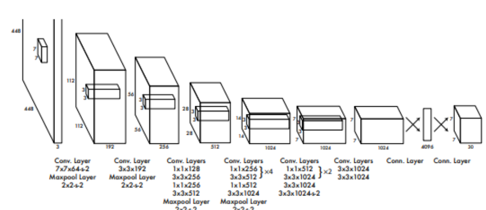

Fig. 1. YOLO detection network architecture. [17]

## III. METHODOLOGY

### A. Data Collection

A machine vision system had been created for the purpose of data capture but was not implemented due to travel restrictions. A GoPro®Hero8 was attached beneath the planter looking into the open furrow in two orientations, forward- and rear-facing. Data was captured over two days of planting using 60 frames per second (fps) and 240 fps video. In total, 9 videos were collected with around 100 minutes of footage in varying natural lighting conditions.

### B. Constructing the Dataset

The data collected was of poor quality due to dust and lighting issues caused by the planter moving at an average speed of 12km/hr through freshly formed beds. The best 240fps video was split into frames and every 60th frame used in order to increase the variability of the dataset. The images were then cropped to a Region of Interest (ROI). Each billet in the image was labelled using a bounding box technique as a single class. High quality annotated lab data was added to provide finer grained detail of a billet. In total, 1500 annotated images were used in the final iteration.

### C. Training the YOLOv3 Model

A YOLOv3 model was trained initially with 500 images and transfer learning using 2 different models and weights, YOLOv3-spp and YOLOv3-tiny. 500 high quality lab images were then added and transfer learning again used to create models. Each model used 80% training data from one source and 20% test data from another. A batch size of 4 increased the throughput and variability in each training iteration and allowed training on a NVIDIA®1050 GPU. Mosaic tiling and HSV augmentation were also applied [18]. Another 500 images were added to the final training set and a final YOLOv3-spp model trained. Each model was trained for 300 epochs and the epoch with the best results was used for comparison.

### D. Performance evaluation

In order to determine the performance of each model, the precision, recall, mean Average Precision (mAP) and F1 scores were used and compared to determine the best performance, Table I.

TABLE I
MODEL PERFORMANCE INDICATORS

| Model | Precision | Recall | mAP | F1 |
|---|---|---|---|---|
| YOLOv3-spp | 0.686 | 0.851 | 0.793 | 0.76 |
| YOLOv3-tiny | 0.647 | 0.728 | 0.714 | 0.685 |
| YOLOv3-spp | 0.700 | 0.741 | 0.743 | 0.720 |
| YOLOv3-tiny | 0.686 | 0.733 | 0.726 | 0.709 |
| Final model | 0.745 | 0.888 | 0.852 | 0.810 |

## IV. RESULTS AND DISCUSSION

The YOLOv3-spp model trained using transfer learning out-performed the others and was used to detect billets in an unseen test video. This model detected billets with a reasonably high accuracy as seen in Figure 2a in some images but experience false negatives and false positives in others, Figure 2b. These results were used in the following sections.

(a) Test results with 100% precision

(b) Test results with false negatives

Fig. 2. Sugarcane Billet detection using YOLOv3 detector

### A. Seed Density Mapping in the Furrow

The results of each image was output to a text file with the class and bounding box co-ordinates for each object. A python script was written to open each output and count the number of billets per image. This was stored in an array and used to plot a furrow map of seed density as seen Figure 3. A map across the field of seed density will allow for improvements to be made to existing planters and results recorded. This will add the most value to the sugarcane industry as it should determine where underlying issues exist with current planting processes. Each detection run on the NVIDIA®1050 took around 0.06 seconds, which would provide an accurate indication of billets per meter in real-time to a planter operator.

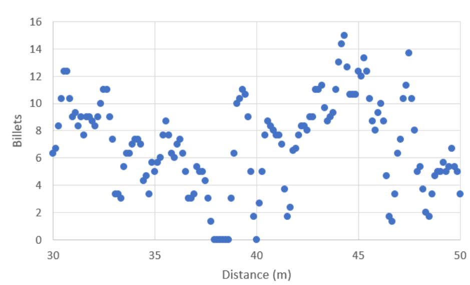


Fig. 3. Furrow seed density mapped over a 20m section

## V. CONCLUSION

The GoPro®data demonstrates the feasibility of using machine learning and object detection to provide furrow mapping of sugarcane billet and useful real-time metrics. Data quality is key to improving the existing model and this experiment proves the need for a higher quality machine vision system to be implemented. A system will be implemented using controlled lighting, a high quality machine vision camera and positive pressure to remove dust from the ROI for future tests.

## References


[1] M. Alencastre-Miranda, J. R. Davidson, R. M. Johnson, H. Waguespack, and H. I. Krebs, "Robotics for sugarcane cultivation: Analysis of billet quality using computer vision," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3828–3835, 2018.

[2] B. Croft, "Srdc project bss208: Literature review of methods of improving the germination of sugarcane," Sugar Research Australia, Report, 2000.

[3] R. Davis, C. Whiteing, C. Norris *et al.*, "A review of opportunities to improve the design and performance of sugarcane harvesters," Sugar Research and Development Corporation, Report, 2010.

[4] C. Norris, "Engineering technology and crop establishment," *The Journal of the Australian Institute of Agricultural Science*, vol. 51, 1985.

[5] L. Thompson, M. Young, M. Foster, A. Bury, S. Cartwright, and P. Please, "A report on the impacts of the sugar industry reform program (sirp): 2004 to 2008," Australian Bureau of Agricultural and Resource Economics and Sciences, Report, 2013.

[6] C. H. W. d. Souza, R. A. C. Lamparelli, J. V. Rocha, and P. S. G. Magalhães, "Mapping skips in sugarcane fields using object-based analysis of unmanned aerial vehicle (uav) images," *Computers and Electronics in Agriculture*, vol. 143, pp. 49–56, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168169917312243

[7] S. Nalawade, A. Mehta, and A. Sharma, "Sugarcane planting techniques: a review," in *Contemporary research in India: National Seminar Recent Trends in Plant Sciences and Agricultural Research (PSAR-2018)*, 2018, pp. 98–104.

[8] N. Ayele, G. Abiy, and N. Tadesse, "Influence of intra-row row setts spacing on yield and yield components of some sugarcane varieties at finchaa sugar estate," *ARPN Journal of Science and Technology*, vol. 4, no. 1, pp. 39–45, 2014.

[9] J. Leslie and B. Leslie, "Yield response surfaces for spacing of sugarcane plants," *Proc. Aust. Soc. Sugar Cane Technol.*, vol. 27, 2005.

[10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[11] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.

[12] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[13] Y. Bengio, "Learning deep architectures for ai," *Foundations*, vol. 2, pp. 1–55, 01 2009.

[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, Conference Proceedings, pp. 1097–1105.

[15] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," *arXiv preprint arXiv:1312.6229*, 2013.

[16] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *ArXiv*, vol. abs/1804.02767, 2018.

[17] J. Redmon, S. K. Divvala, R. B. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *CoRR*, vol. abs/1506.02640, 2015. [Online]. Available: http://arxiv.org/abs/1506.02640

[18] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv preprint arXiv:1712.04621*, 2017.