Counting Pollen Viability via Deep Learning

Simon Castle-Green, Michael P. Pound, and Tony P. Pridmore

University of Nottingham, Nottingham NG8 1BB UK {simon.castle-green, michael.pound, tony.pridmore}@nottingham.ac.uk

Pollen counts are an important measure for biologists working in plant phenotyping and assist in understanding a plant's ability to tolerate stresses resulting from different growing conditions. Accurately assessing pollen can prove challenging within high-throughput experiments [2]. Current pollen counting methods include: manual counting, image analysis software [1], & impedance flow cytometry [4]. However, these methods are cumbersome, don't provide viability (e.g. fertile vs. sterile pollen grains) counts, and are costly to implement, respectively.

In this paper we propose a deep neural network capable of generating viability counts of pollen grains in microscopic images. The principal challenge in producing accurate viability counts lies in the variability of the dataset, which comprises pollen grain images of cultivars of multiple species subjected to differing heat stress treatments and stained using different techniques. In addition, the quality of images is variable, with differing levels of lighting, magnification, resolution, compression, background noise, foreign objects, partial occlusion and clutter. Figure 1 shows training set samples that illustrates the kinds of variability encountered.

Building on previous work analysing wheat spikes and spikelets [6], a stacked hour-glass architecture [5] was implemented, allowing segmentation of input images into 3 channels and providing



Fig. 1: Image samples from training set.

a mechanism to generate fertile, sterile and total pollen grain counts.

The model was trained on 130 annotated microscopic images, including samples from both Rice (Figs. 1a-1f) and Arabidopsis plants (Figs. 1g-1i) cultivated under a variety of growing conditions, and stained with either Potassium Iodide (Figs. 1a-1h) or Fluorescein Diacetate (FDA) (Figure 1i). Data samples typically contain both fertile and sterile grains and annotation files provide co-ordinate sets for each type.

Due to hardware limitations, it is not feasible to process whole images at full size. Downsampling the image results in loss of detail due to the small size of pollen grains. Therefore, a random cropping approach was used, processing a portion of each image per sample. This maintains context within the

2 S. Castle-Green et al.

image and allows images to be reused multiple times within the training process. Data augmentation techniques such as randomised jitter, scaling and rotation were also used. Training was run over 100 epochs with each image sampled 10 times per epoch. RMSProp was used for network optimisation and MSE for the loss function.

Individual F1 scores per channel were used for evaluating performance. At the 100th epoch, channel accuracy scores of 0.894, 0.737 and 0.881 were achieved for Fertile, Sterile and All, respectively. Three primary modes of failure were identified resulting in a reduction in accuracy. Grain types were sometimes misclassified where visible differences between sterile and fertile grains appeared to be extremely subtle. Incorrect identification of similarly shaped foreign objects triggered some false positives. Finally, in heavily clustered areas some grain predictions were lost, possibly due to the threshold value and neighbouring grains. The lower score in the sterile channel is believed to be due to unbalanced class distribution between grain types across image samples. Future work could implement selective cropping to create balanced distributions. Whilst the third channel containing all pollen grains didn't provide improvement upon the fertile and sterile channels, it was seen to assist with training, marginally increasing accuracy and reducing volatility compared to experiments without.

As with training, hardware limitations prevent inferring complete images in a single pass. Therefore, the inference method used in [6] was extended to introduce a sliding window approach, cropping tiles from the image, processing them individually and then stitching them back into a full image. To en-



Fig. 2: Inference predicted 1195 yellow and 71 red masks representing fertile and sterile pollen grains respectively.

sure overlapping pollen grains residing at the edges of adjoining tiles are not double-counted, an R-tree structure [3] was used to store detected grain locations. New grains are checked for existing representation using a pre-defined distance threshold prior to being added to the tree.

The inference process produced counts in the region of 8-90 seconds (dependant on input image size) running on a 2.3 GHz 8-Core Intel Core i9 CPU. This is much improved on manual counts (5–68min [2]), and although marginally slower than [1], our method requires no pre-processing, provides viability counts and generates output images (Figure 2) aiding human verification.

While the model does not perform perfectly in all conditions, we believe it is sufficiently accurate to be useful to biologists, allowing rapid approximation of viability counts without requiring specialist equipment or complex pipelines. Manual annotation of pollen grain images to create training data is laborious and prone to error and it is anticipated that future work will use active learning to improve performance.

Acknowledgments - We thank Behzad Talle & Zoe Wilson, School of Biosciences, University of Nottingham, for the dataset used in this research.

References

- Costa, C.M., Yang, S.: Counting pollen grains using readily available, free image processing and analysis software. Annals of Botany **104**(5), 1005–1010 (2009)
- Dreccer, M.F., Molero, G., Rivera-Amado, C., John-Bejai, C., Wilson, Z.: Yielding to the image: How phenotyping reproductive growth can assist crop improvement and production. Plant science 282, 73–82 (2019)
- Guttman, A.: R-trees: A dynamic index structure for spatial searching. In: Proceedings of the 1984 ACM SIG-MOD international conference on Management of data. pp. 47–57 (1984)
- Heidmann, I., Schade-Kampmann, G., Lambalk, J., Ottiger, M., Di Berardino, M.: Impedance flow cytometry: a novel technique in pollen analysis. PloS one 11(11) (2016)
- Newell, A., Yang, K., Deng, J.: Stacked hourglass networks for human pose estimation. In: European conference on computer vision. pp. 483–499. Springer (2016)
- Pound, M.P., Atkinson, J.A., Wells, D.M., Pridmore, T.P., French, A.P.: Deep learning for multi-task plant phenotyping. In: Proceedings of the IEEE International Conference on Computer Vision Workshops. pp. 2055– 2063 (2017)