Selvyn Perez
Xinyu Yao
https://github.com/xinyuyao22/Global-Wheat-Detection

Data Science Capstone Group Proposal

Problem Description

For the Capstone project, we plan to join an ongoing Kaggle competition: Global Wheat Detection[1], where we are attempting to predict bounding boxes around each wheat head in images. These images can be used to estimate the density and size of wheat heads in different varieties. Farmers can use the data to assess health and maturity when making management decisions in their fields.

Accurate wheat head detection in outdoor field images can be visually challenging. There is often overlap of dense wheat plants, and the wind can blur the photographs. Both make it difficult to identify single heads. Additionally, appearances vary due to maturity, color, genotype, and head orientation. Finally, because wheat is grown worldwide, different varieties, planting densities, patterns, and field conditions must be considered. Models developed for wheat phenotyping need to generalize between different growing environments. Current detection methods involve one- and two-stage detectors (Yolo-V3 and Faster-RCNN), but even when trained with a large dataset, a bias to the training region remains. The baseline method[2] trained using Faster-RCNN, with a ResNet34 as the backbone yielded a mAP@0.5 of 0.68 and a mean RMSE of 13.75 wheat heads per image which corresponds to rRMSE=34%.

Dataset Description

Global Wheat Head Detection (GWHD)[3] dataset is a large and diverse dataset of high resolution RGB labelled images to develop and benchmark wheat head detection methods. The data is images of wheat fields, with bounding boxes for each identified wheat head, though not all images include wheat heads / bounding boxes. The training dataset corresponds to 3422 images from Europe and North America representing 73% of the whole GWHD dataset images. To evaluate model performance, including robustness against unseen images, the test data set includes all the images from Australia, Japan, and China, representing 1276 images. All images share a common format of 1024×1024 px with a resolution of 0.1-0.3mm per pixel.

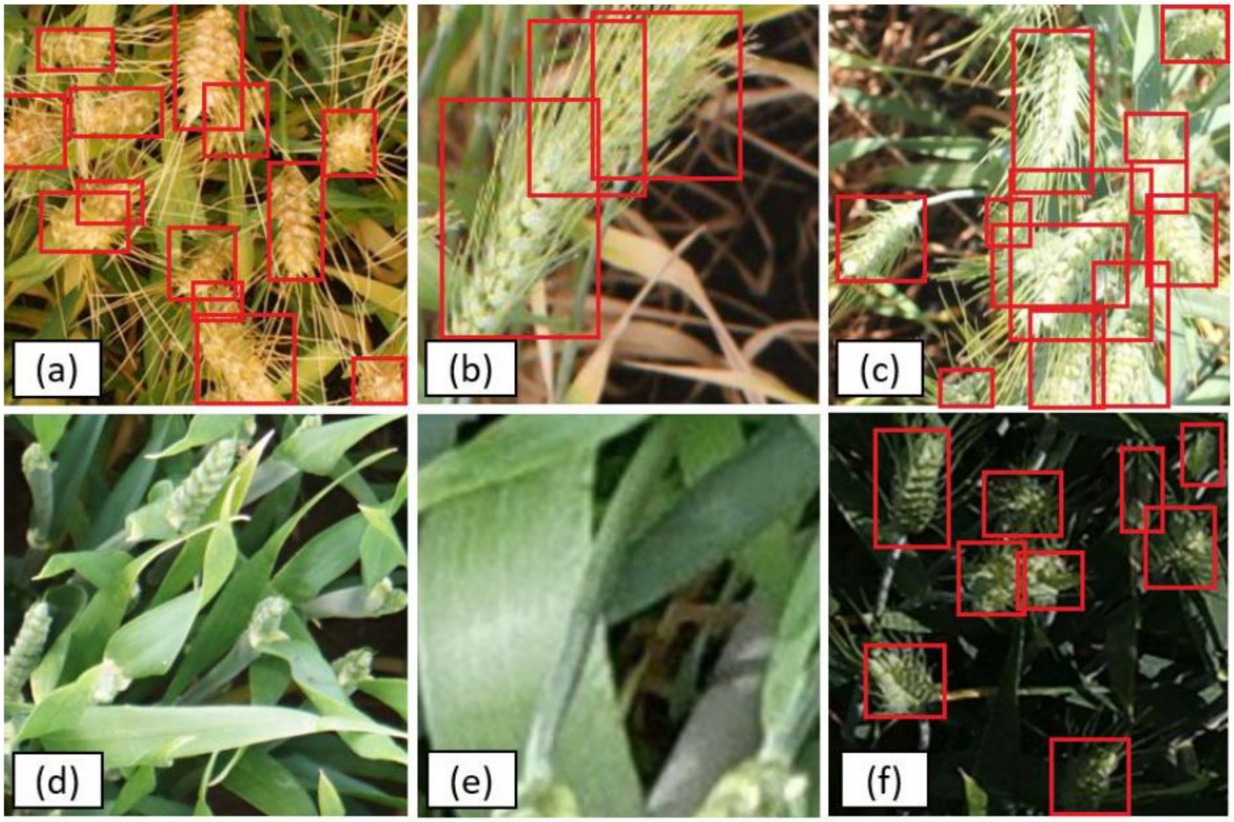


Figure 1: Examples of wheat heads difficult to label. It includes overlapping heads (a-c), heads at emergence (d), heads that are partly cut at the border of the image (e), and images with a low illumination (f). These examples are zoomed-in views from images contained in the dataset, with different zoom factors. Note that the image (d) was removed from the dataset because of the ambiguity of heads at emergence.

Technique to use

Image Augmentation: Albumentations[4], a fast and flexible open source library for image augmentation with many various image transform operations available.

Network Architecture for Object Detection: ResNeST[5], Cascade Mask R-CNN[6], and EfficientDet[7].

Ensemble Method: Weighted Box Fusion (WBF)[8], an ensemble algorithm that boosts the performance by ensembling predictions from different object detection models.

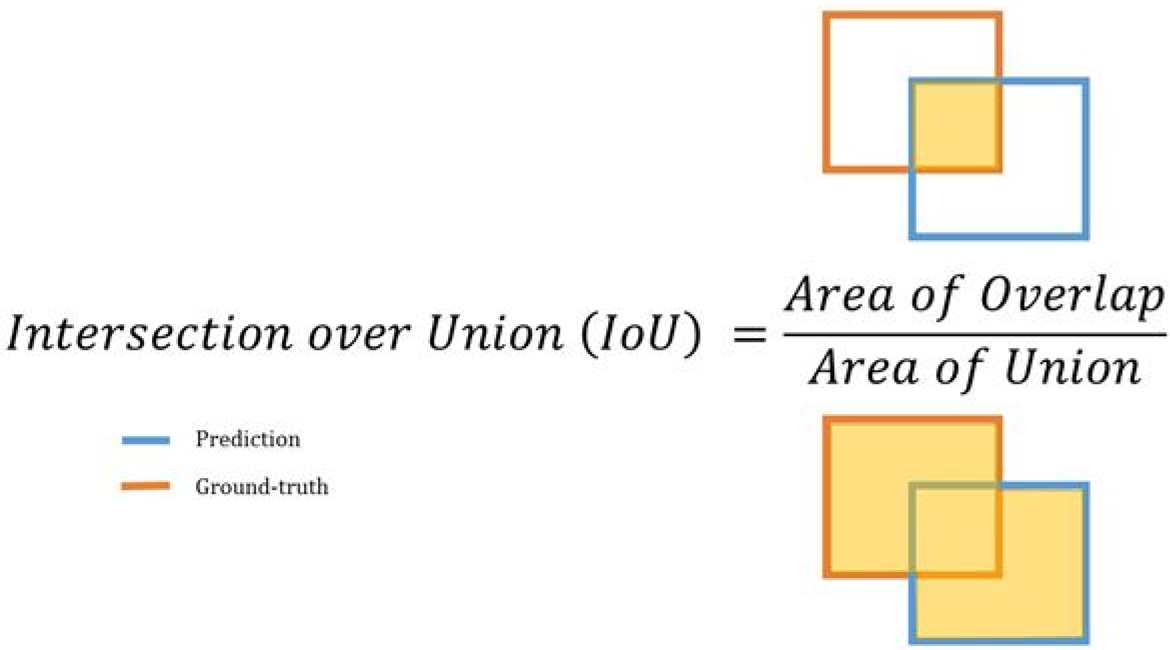
Framework to use

Pytorch, an open source machine learning framework.

Evaluation

This competition is evaluated on the mean average precision at different intersections over union (IoU) thresholds. IoU is a measure of the magnitude of overlap between two bounding boxes (or, in the more general case, two objects). It calculates the size of the overlap between two objects, divided by the total area of the two objects combined.

It can be visualized as the following:



The two boxes in the visualization overlap, but the area of the overlap is insubstantial compared with the area taken up by both objects together. IoU would be low - and would likely not count as a "hit" at higher IoU thresholds.

Schedule (Provide a rough schedule for completing the project)

Jun 6 - Jul 31: Trying different network architecture

Aug 1 – 4: Get final submission of competition.

Aug 3 – 6: Work on final report draft

Aug 7 – 13: Get the final report done and prepare for final presentation

Aug 14 – 20: Prepare for final project presentation and journal submission

Aug 20 – 24: Get final submission to a journal

[1] https://www.kaggle.com/c/global-wheat-detection

[2] David, E., Madec, S., Sadeghi-Tehran, P., Aasen, H., Zheng, B., Liu, S., Kirchgessner, N., Ishikawa, G., Nagasawa, K., Badhon, M.A. and Pozniak, C., 2020. Global Wheat Head Detection (GWHD) dataset: a large and diverse dataset of high resolution RGB labelled images to develop and benchmark wheat head detection methods. *arXiv preprint arXiv:2005.02162*.

[3] http://www.global-wheat.com/

[4] Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M. and Kalinin, A.A., 2020. Albumentations: fast and flexible image augmentations. *Information*, *11*(2), p.125.

[5] Zhang, H., Wu, C., Zhang, Z., Zhu, Y., Zhang, Z., Lin, H., Sun, Y., He, T., Mueller, J., Manmatha, R. and Li, M., 2020. ResNeSt: Split-Attention Networks. *arXiv preprint arXiv:2004.08955*.

[6] Liu, Y., Wang, Y., Wang, S., Liang, T., Zhao, Q., Tang, Z. and Ling, H., 2019. Cbnet: A novel composite backbone network architecture for object detection. *arXiv preprint arXiv:1909.03625*.

[7] Tan, M., Pang, R. and Le, Q.V., 2019. Efficientdet: Scalable and efficient object detection. *arXiv preprint arXiv:1911.09070*.

[8] Solovyev, R. and Wang, W., 2019. Weighted Boxes Fusion: ensembling boxes for object detection models. *arXiv preprint arXiv:1910.13302*.