Datasets

- Region Of Interest is within the RED Square
- Labels are normalized bounding box coordinates w.r.t original image dimensions given in .JSON format

Wheat Dataset

Barley Dataset

Canola Dataset

Dataset Preparation

- Cropping the area of image within the RED Square
- Adjusting the bounding box coordinates to cropped image dimensions

ROI Image of Wheat Dataset

ROI Image of Barley Dataset

ROI Image of Canola Dataset

Object detection using classical computer vision techniques

Techniques explored:

- HOG + SVM

HOG+SVM

- Sample kernel images of Grain Class (Positive example) and Not belonging to Grain Class (Negative example)
- Extract HOG features
- Train SVM Classifier on the HOG features extracted
- Resize the full image at multiple scales.
- Use Sliding window to classify the presence of Grain kernel.
- Apply non-maximum suppression to remove redundant and overlapping bounding boxes.

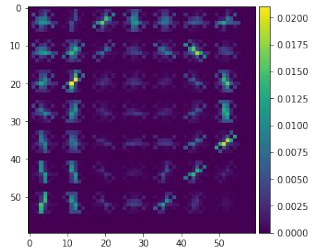
STEP 1: Dataset created for Wheat and Not Wheat Images

Wheat

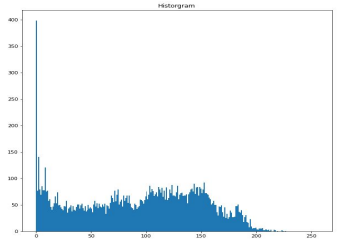
Not Wheat

STEP 2: HOG Features of Wheat Kernel

Wheat Kernel (71 x 54)

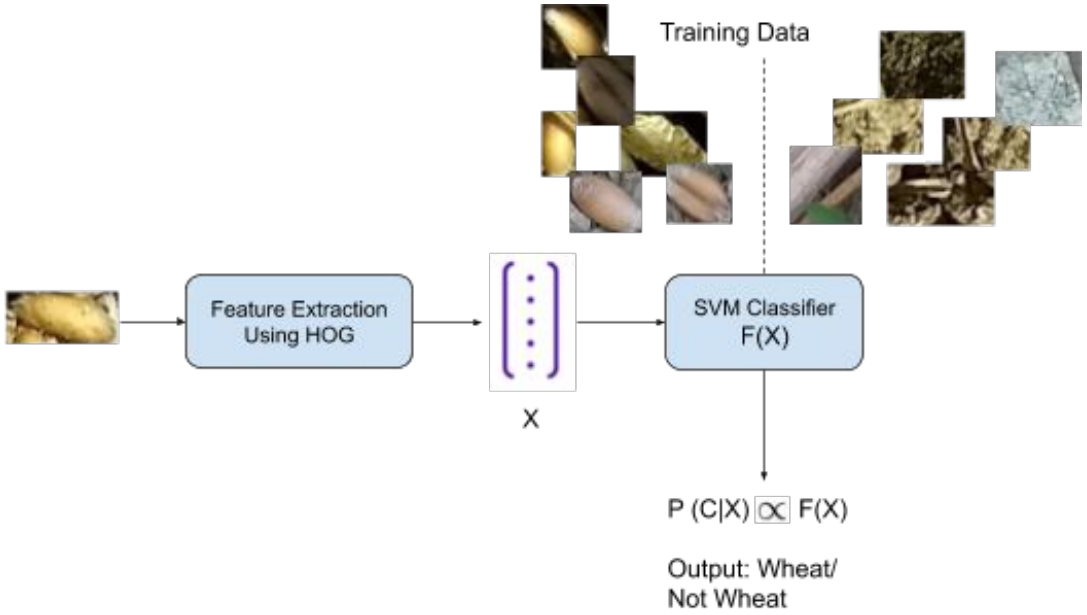


Hog Features of Wheat Kernel rescaled to 60 x 60



Histogram of Gradients in the Wheat Kernel

STEP 3: HOG + SVM Model Classifier



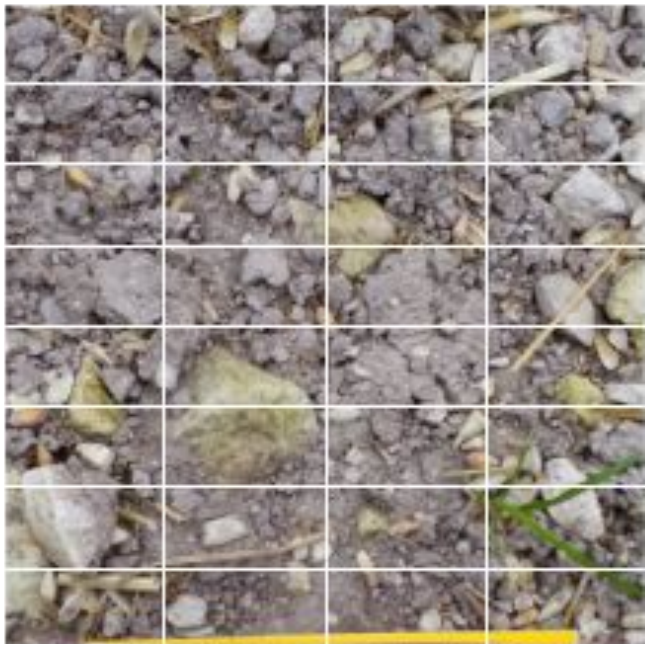
STEP 4: Test Score (scaled from 0-1) using SVM on Hog Dataset of Wheat and Not Wheat Kernel Images

	precision	recall	f1-score	support
not wheat	0.97	0.97	0.97	495
wheat	0.97	0.97	0.97	495
accuracy			0.97	990
macro avg	0.97	0.97	0.97	990
weighted avg	0.97	0.97	0.97	990

HOG and SVM Model tested on full images using sliding windows

STEP 5: Image Pyramid constructed

```
(500, 500, 3)
(250, 250, 3)
(125, 125, 3)
(63, 63, 3)
(32, 32, 3)
(16, 16, 3)
(8, 8, 3)
(4, 4, 3)
(2, 2, 3)
(1, 1, 3)
```



STEP 6: Each Image from pyramid divided into several windows of size 60X60

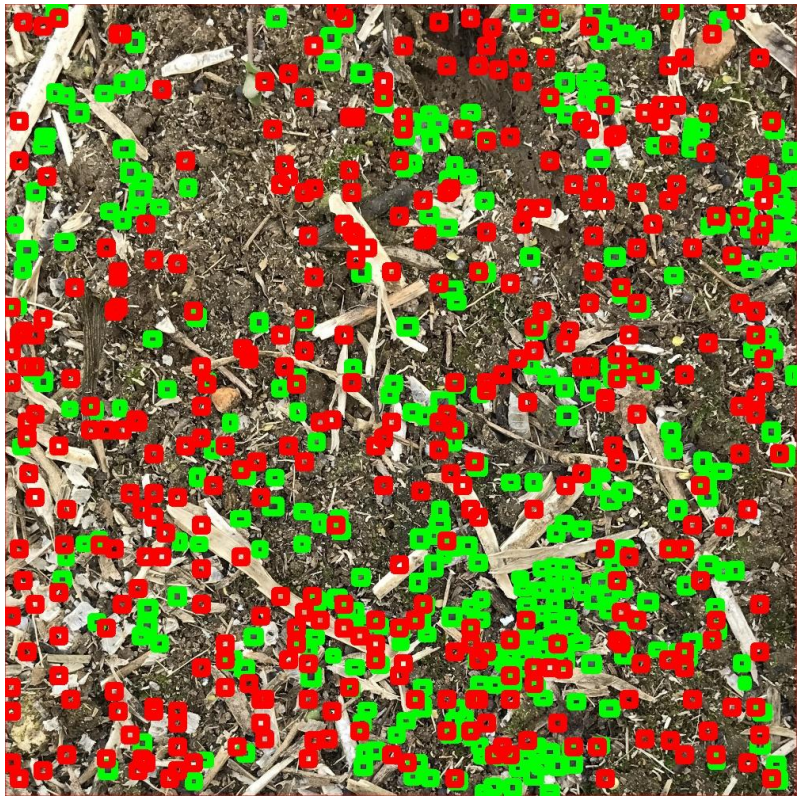
STEP 7: Histogram of Gradients computed and classified for each sliding window on each image from the pyramid

Detections using HOG+SVM

Ground truth boxes marked in **GREEN** and detections marked in **RED** colors

Wheat

Barley



Canola

Results and Evaluation

Evaluation metrics to be used is Average Precision

Average precision is calculated using precision, recall and IOU

- Precision=TP/(TP+FP). Precision is the ratio of correctly predicted positive observations to the total predicted positive observations
- Recall=TP/(TP+FN)). Recall is the ratio of correctly predicted positive observations to the all observations in actual class
- IOU gives a similarity score between the ground truth Bgt and predicted bounding boxes Bp

$$IOU(B_p, B_{gt}) = \frac{area(B_p \cap B_{gt})}{area(B_p \cup B_{gt})}$$

- 11 Point Interpolated precision is calculated by taking the mean precision at a set of eleven equally spaced recall levels [0,0.1,...,1]

$$AP_{11} = \frac{1}{11} \sum_{R \in \{0,0.1,,,,1\}} P_{interp}(R) \qquad P_{interp}(R) = \max_{\tilde{R}:\tilde{R} \geq R} P(\tilde{R})$$

- The precision at each recall level r is interpolated by taking the maximum precision measured for a method for which the corresponding recall exceeds r

$$A_{all} = \sum_{n} (R_{n+1} - R_n) P_{Interp}(R_{n+1}) \qquad P_{interp}(R_{n+1}) = \max_{\tilde{R}:\tilde{R} \geq R_{n+1}} P(\tilde{R})$$

Results of classical computer vision techniques

Grain Type	HOG + SVM
Wheat	14.58 %
Barley	10.75 %
Canola	21.85 %

Disadvantages

HOG + SVM model would then learn the curvature of the grains and base its classification similar to that. However, they are:

- Less scalable as compared to Deep Neural Networks
- Cannot extract own features on full images like texture, colors, shape of grains
- Faster to train (0.200 seconds) but slower to test (225.04 seconds) as compared to CNNs
- Crowding objects, illumination and orientation invariant, different aspect ratios of object lead to false positives.

Future Scope:

- Hard-negative mining for HOG+SVM
  - Retrain classifier in multiple stages using false positives generated in earlier stage.