Title : Rice grain counting and analysis

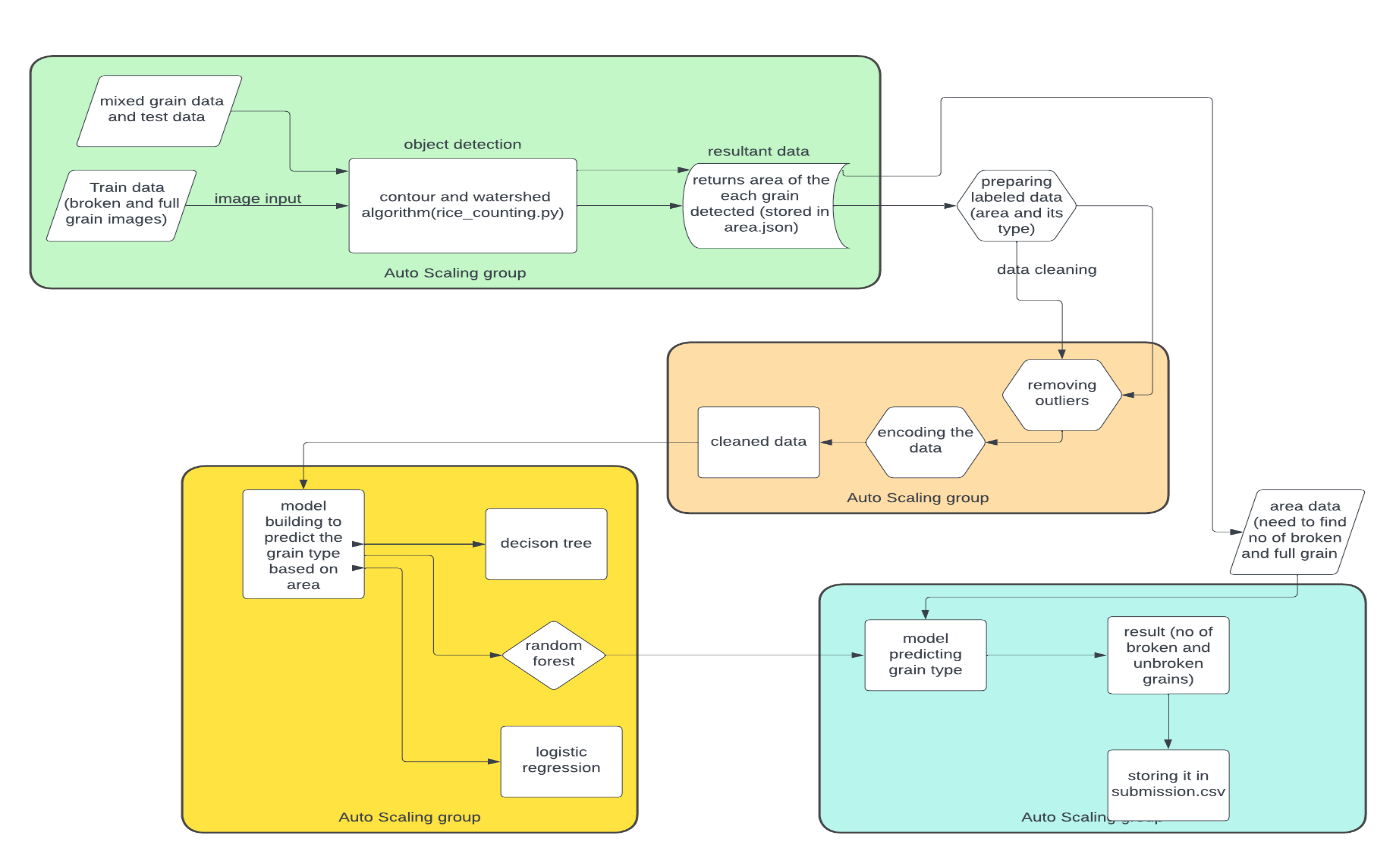
Problem statement:

1. To count the total no. of rice grains from the given image
2. To differentiate between broken and full grain

Methodology:

1. To count the total no. of rice grains from the given image
   1. I have considered various segmentation and edge detection algorithm such as traditional contour detection , canny edge algorithm , water shed algorithm ,yolo , somel etc.,
   2. Given the time constraint only tried contour detection and watershed algorithm has been tired
   3. Watershed algorithm showed best result out of 2 algorithms considered
   4. created a module(rice\_counting.py) which returns area of each grain along with some statistics(total count of rice , average area , median area)
2. To differentiate between broken and full grain
   1. Created a data file containing area and its label (broken or full)
   2. Used this data to train a machine learning model to predict the grain type based on the area
   3. After fine tuning the random forest model is selected and non label data is passed to the model
   4. The result no of broken and full is stored in submission.csv

THE PIPELINE

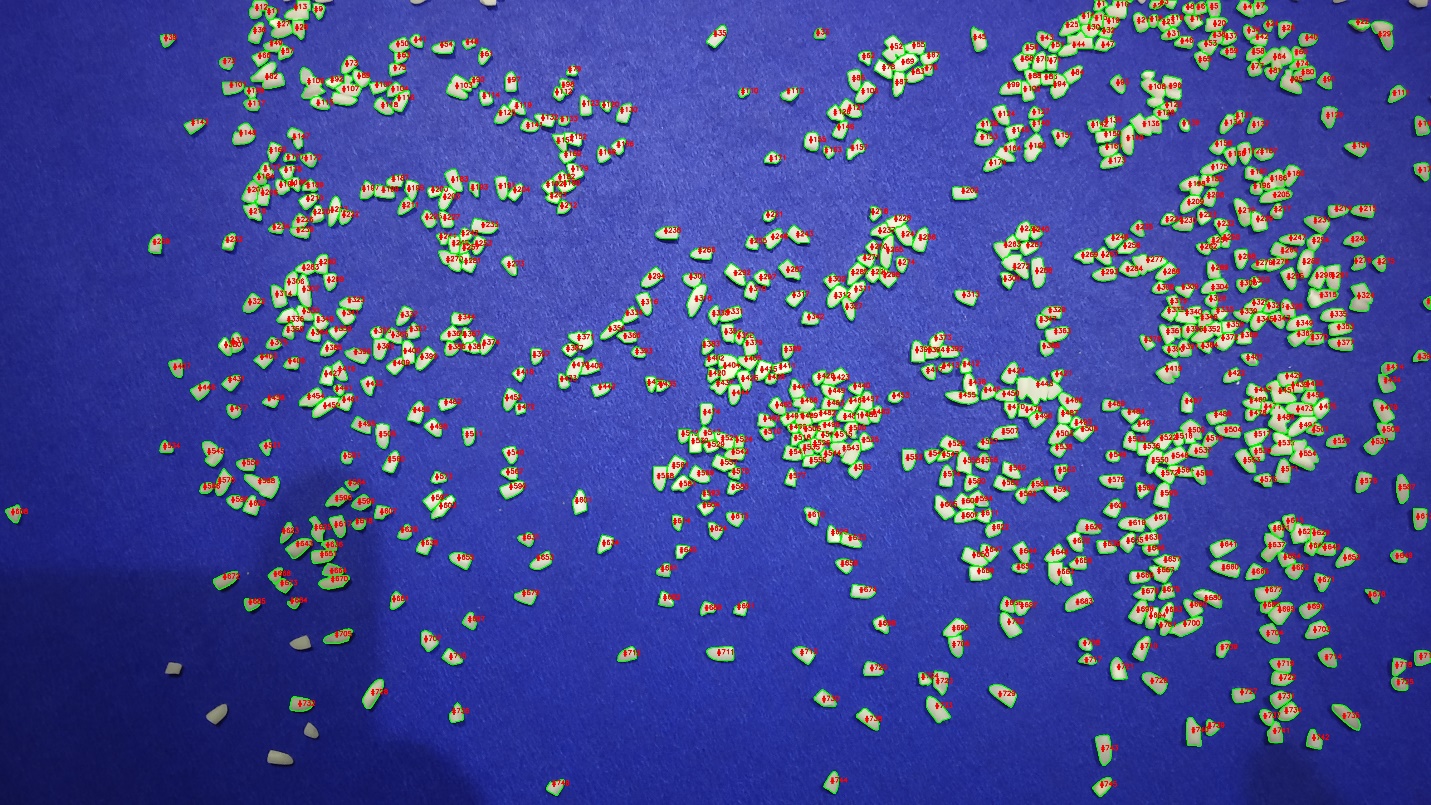


TRAIN DATA:

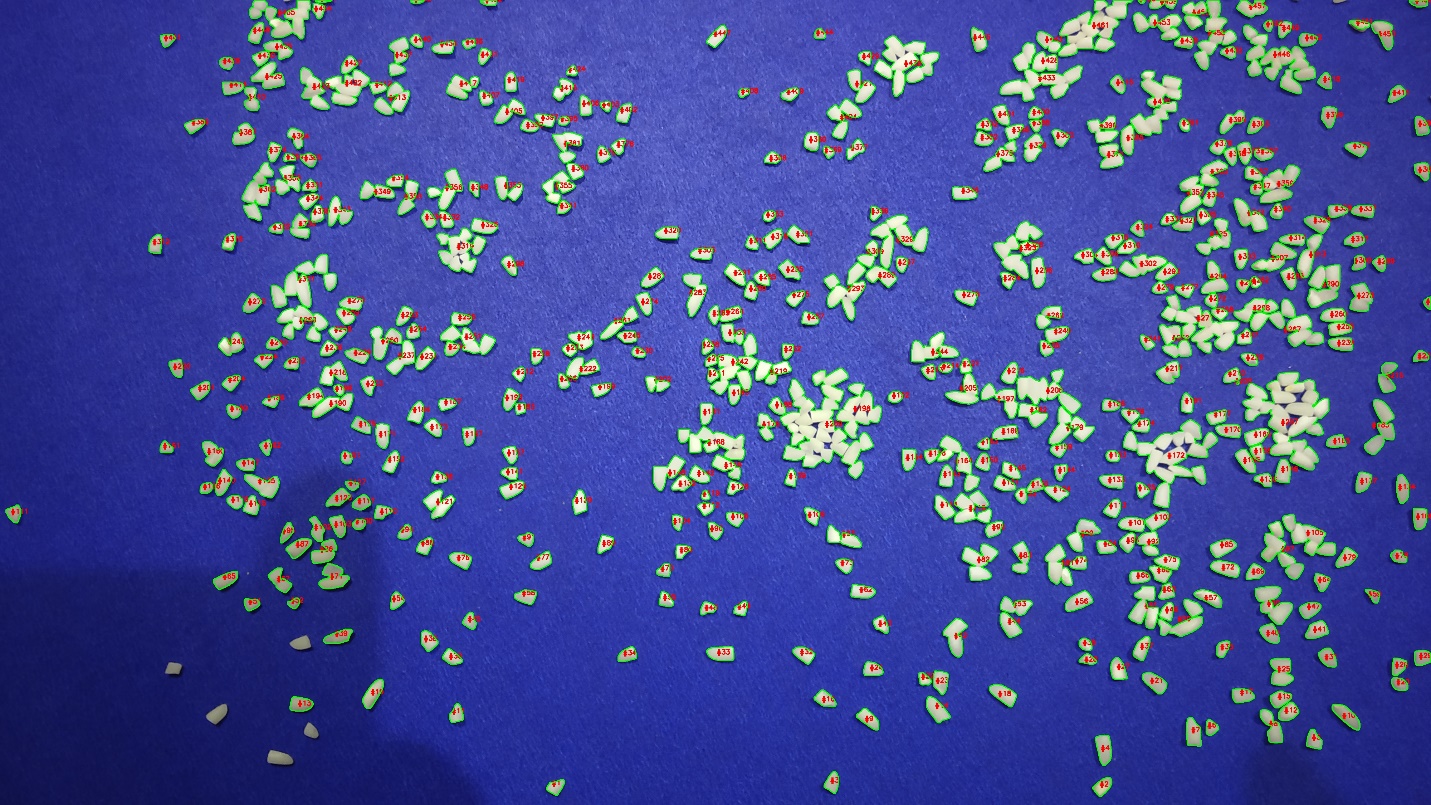
SAMPLE INPUT IMAGE:



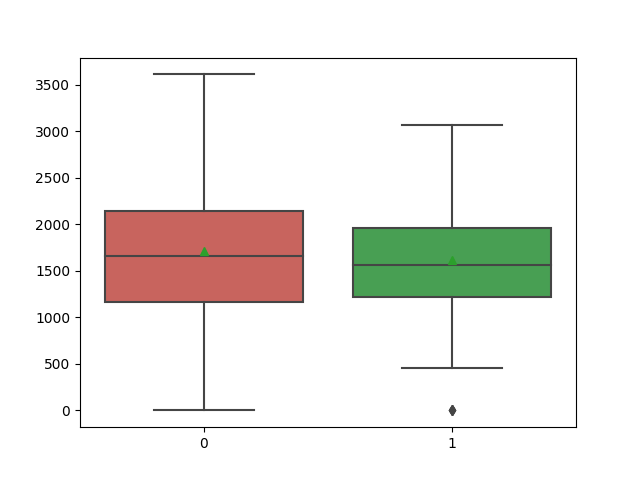
OUPUT IMAGE BY WATERSHED ALGORITHM:



OUTPUT BY CONTOUR ALGORITHM:



Data distribution of broken and full grain



As from the above visualization we can see the we can’t differentiate broken and full grain

Random forest algorithm gives a Auc-roc score of .54 ± 2 which is not different than random guessing

LIMITATION:

1. Can’t differentiate between full and broken grain
2. Grain touching each other is counted as single object by contour method and even though watershed algorithm solves it the segmentation at every place is not perfect

POSSIBLE FUTURE IMPROVEMENT

1. We can try different deep learning algorithms like yolo to do image segmentation