Weed Detection using Semi-Supervised Learning

Team 57

I. Understanding the Problem Statement

This project aims at developing a weed detection system using Semi-Supervised Learning mainly for agricultural applications. The core task is essentially an object-detection problem where the model needs to accurately identify and localize weeds in images of sesame plants in a field.

Key aspects of the problem are:

- **Data Constraints:** The available dataset is partitioned into three subsets:
  - **Labeled Data:** 200 images with corresponding YOLOv8-format annotations.
  - **Unlabeled Data:** 1000 images without annotations.
  - **Test Data:** 50 images with annotations.
- **Semi-Supervised Learning:**
  - **Why:** Annotating large datasets for object detection is expensive and labour-consuming. Since the labeled data is less (200 images), a purely supervised approach would risk overfitting and likely fail to generalize other images.
  - **Benefits:** Integrating unlabeled data into training :
    - Improves the model's ability to generalize by feeding a wider variety of data and thus enhancing performance.
    - Reducing the dependency on large annotated datasets.
- **Evaluation Metrics:** The model is evaluated using a composite metric that balances both classification and localization performance, combining F1-Score with the mean Average Precision (mAP).
  Evaluation metric = 0.5*(F1-Score)+0.5*(mAP@[.5:.95])

This section establishes the challenges and motivates the use of advanced techniques to effectively utilize the large unlabeled dataset alongside the less number of labeled examples.

II. Base Model

Firstly, we have trained our base model by supervised learning solely on the 200 labeled images as a pre-trained model using YOLOv8. The evaluated parameters are :

- **F1-Score:** 0.85
- **mAP@[.5:.95] :** 0.59
- **Evaluation metric:** 0.72

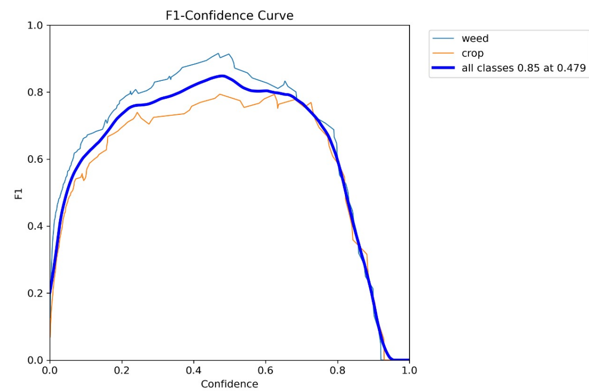

Fig. 1: F1 Confidence Curve for Base Model

Based on this base model, we have taken different methods in order to determine the better solution.

III. Approaches

We have tried preparing **four** approaches in training the model while working on this problem statement :

- **Pseudo-Labeling Approach 1 -** In this approach, we batched every 100 unlabeled images for pseudo-labeling in a stepwise manner. The images were chosen depending on the confidence threshold value. If the prediction of the model is greater than the confidence threshold, those images were combined with the original dataset and then we re-trained the model. Then this re-trained model was used to predict the pseudo-labels for the next 100 unlabeled images along with the already combined dataset and this was done ten times cumulatively in a loop.
  For our final code for this, we have set epochs = 25.For the first 10 epochs new images are fetched in batches of 100. We began with the threshold = 0.85, then gradually decreased the threshold by 0.05 until 0.65. The results were satisfactory for this approach. The obtained values were :
  F1 score : 0.885
  mAP@[.5:.95]: 0.602
  Evaluation metric : 0.743
  Why this value of epoch and confidence threshold? We initially tried choosing epoch = 50. At a higher value like this, in the first 10 epochs, the model used images for retraining; however, after that, it began skipping the images. This could cause a significant bias towards the images used for training and reduces the effectiveness of utilization of the entire dataset. We began gradually decreasing the confidence threshold so that the number of images getting skipped gets reduced and hence improving the utilization of the dataset.

- **Two-Stage Augmented Pseudo-Labeling -** Here after training the the base model, firstly we weakly augment the unlabeled data to increase the unlabeled dataset and to create variety in the images. Then we used the trained base model to predict pseudo labels on this weakly augmented dataset and consider only those data images whose prediction are greater than the confidence threshold set. Then these images were combined with the original labeled data and this entire set was used to re train the previously trained model.
  After this the weakly augmented images whose prediction was greater than threshold were made to undergo strong augmentation. Then the retrained model was used to predicted labels on the strong augment images and similarly only those were selected which had pseudo- labels greater than the threshold. Then finally these images were combined with the original 200 labeled images and the model was retrained to get the final scores.
  F1 score : 0.912
  mAP@[.5:.95]: 0.578
  Evaluation metric : 0.745
  From this approach we achieved a better value of f1 score but the mAP value decrease.

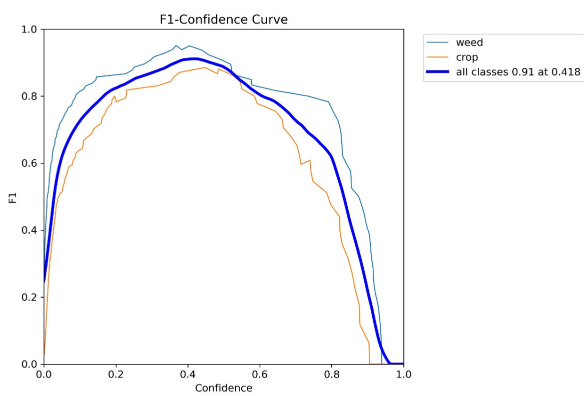

Fig. 2: F1 Confidence Curve for Two-Stage Augmented

- **Pseudo-labeling Approach 2 -** In this approach, the base model is applied to all the 1000 unlabeled images at once, and its predictions are used as pseudo ground truth. Only those predictions that meet a specified confidence threshold are retained. We varied the confidence level in order to determine the better one.This effectively enlarges the training dataset without additional manual annotation.

Confidence Threshold	F1-score	mAP[.5:.95]	Evaluation Metric
0.70	0.89	0.57	0.73
0.75	0.90	0.581	0.7431
0.80	0.879	0.608	0.7438
0.85	0.8549	0.596	0.7254

  Note that the value of the metric increases as we increase our confidence threshold from 0.70 and then after reaching a maximum, it decreases as we increase further towards 0.85. From the values, we determine that the best performance occurs at Confidence threshold = 0.80.
- **Augmentation of labeled data -** The base model is trained with the 200 labeled images augmented by Horizontal flip, small rotations, brightness variation, colour jittering and mild blurring. However, the results were not satisfactory. Possible reasons could be overfitting to augmented data, degradation in quality of augmented data due to Gaussian blurring, imbalance of the data, lack of variety in the labeled data.
- **Mean Teacher -** The mean teacher approach uses a student–teacher framework where the teacher model, computed as an exponential moving average of the student, provides consistent predictions. This method enforces stability under input perturbations.
  F1 score : 0.766
  mAP@[.5:.95]: 0.457
  Evaluation metric : 0.616
  The mean teacher approach incurred significant computational overhead and slower convergence. Its sensitivity to hyperparameter tuning, difficulty stabilizing predictions for subtle weeds, errors in object localization due to label normalization could have made it less effective compared to simpler alternatives.

IV. Insights

When we tried to train the base model by augmenting the labeled images by some small rotations and horizontal flips, we found the that the results quality deteriorated. this suggests that the test data images are very similar to labeled data images because of which when we augment the labeled data and combine with the original data, the quality of the training data decreases.

V. Discussion and Conclusion

*The most recommended approach is **Pseudo-Labeling Approach 1***. In **Pseudo-Labeling Approach 2** with confidence threshold = 0.80, a large portion of the unlabeled data gets skipped; as a result, the former model, which utilizes more data is preferable. In comparison with **Two-Stage Augmented Pseudo-labeling** in which the augmented images hampers the quality of the data,the Pseudo-Labeling Approach 1 is logically more consistent and has more chances to turn out to be more useful in wider datasets and in practice. Therefore, even though three of these approaches fetch a similar or close value of the evaluation metric, the **Pseudo-Labeling Approach 1** is more recommended.

This project illustrates how farmers could use our method as a tool to better manage their fields by lowering the requirement for labour-intensive weed control. Additionally, it may be useful when combined with other precision farming tools, which would make it an affordable option for farmers. We can manage the identified weeds by applying accurate, targeted pesticide spraying, for example. All things considered, the suggested weed-plant identification model offered the agricultural sector a viable answer that might increase crop yields and lessen its negative effects on the environment.