

Microsoft Rice Disease Solution

To reproduce solution:

1. Upload the provided notebook to colab
2. Choose GPU runtime - Ensure you use the Tesla V100-SXM2 16 GB VRAM GPU to get a similar score, using a different GPU might result in different scores.
3. Run all to get the submission file
4. Upload the submission file to Zindi for scoring

Solution summary

- Used both the RGB and RGN images for training. A balanced mix of two rgn images and two rgb images were used for training. This together with mixup greatly reduced overfitting of the model.
- The model was trained on a 10 Stratified cross validation strategy to ensure class imbalance was maintained across the folds. The model was trained on the first five folds.
- Early stopping was used during experimentation to get the cut off epochs where there was no further performance improvement past certain epochs
- To further curb overfitting, extensive image augmentation was utilised
- To accommodate the huge model, Automatic Mixed Precision was used to fit the model and data in memory
- Pseudo labelling with predictions over or equal to a probability of 0.999
- Wandb logging tool was used to track experimentation results

Parameters used

1. Model - swinv2_large_window12to24_192to384_22kft1k
2. Training batch size - 4
3. Evaluation batch size - 4
4. Dropout - 0.5
5. Learning rate - 1e-5
6. Image size - 384
7. Scheduler - CosineAnnealingWarmRestarts
8. Optimizer - AdamW