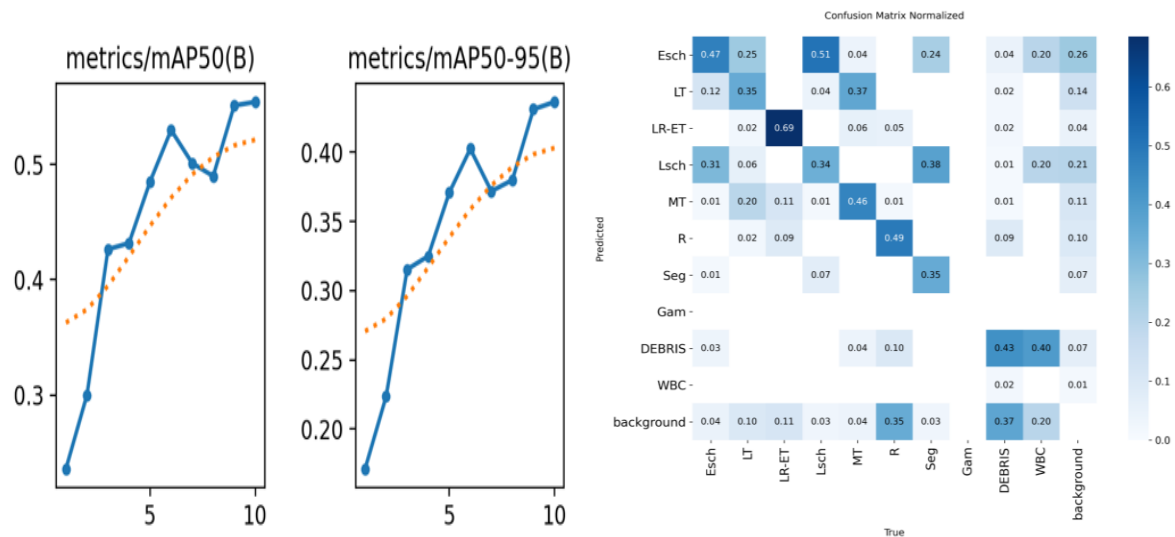# RESULTS

Results from YOLO model:

FIGURE 1: metrics/mAP50 and metrics/mAP50-95 plots, as well as a confusion matrix from running a YOLO model on the dataset. YOLOv8s achieved strong localization performance, with mAP50 = 0.554 and recall = 0.609, indicating that the model was generally able to detect parasite-containing regions across most slides. The confusion matrix shows that YOLO reliably distinguishes between the different annotated parasite categories, but some overlap remains between visually similar stages.
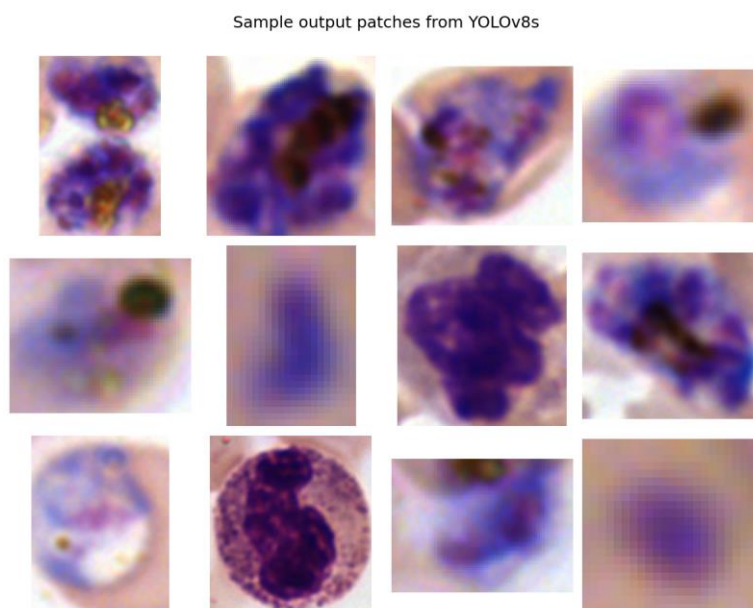

FIGURE 2: Sample output patches from YOLOv8s

TABLE 1: Output metrics for YOLOv8s model (used alone)

METRIC	VALUE
mAP50	0.554
mAP50-95	0.436
Precision	0.518
Recall	0.609

Results from EfficientNet model:


```
--- Classification Report ---
              precision    recall   f1-score    support

      DEBRIS      0.660     0.680      0.670        100
        Esch      0.473     0.265      0.340         98
         Gam      0.000     0.000      0.000          0
       LR-ET      0.391     0.200      0.265         45
          LT      0.442     0.339      0.384         56
        Lsch      0.476     0.588      0.526         68
          MT      0.550     0.458      0.500         48
           R      0.663     0.793      0.722         82
         Seg      0.351     0.812      0.491         32
         WBC      0.471     1.000      0.640          8

    accuracy                           0.527        537
   macro avg      0.448     0.514      0.454        537
weighted avg      0.527     0.527      0.509        537
```
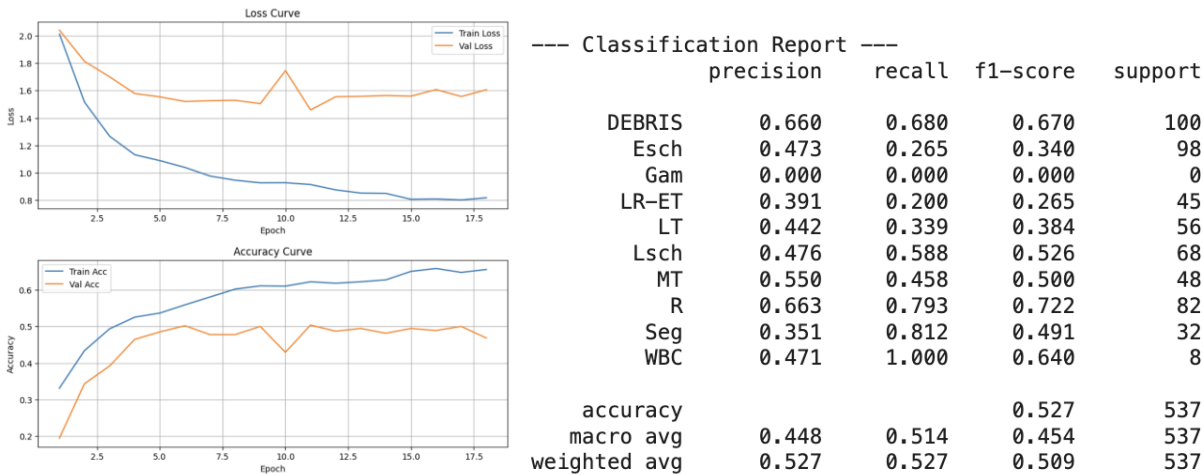
FIGURE 3: Accuracy and loss curves, as well as classification report, for EfficientNet. EfficientNet-B0 showed moderate learning, with training accuracy improving but validation accuracy plateauing early on, suggesting limited generalization. The classification report shows uneven performance across classes, with higher precision in common classes and low recall in underrepresented life stages.
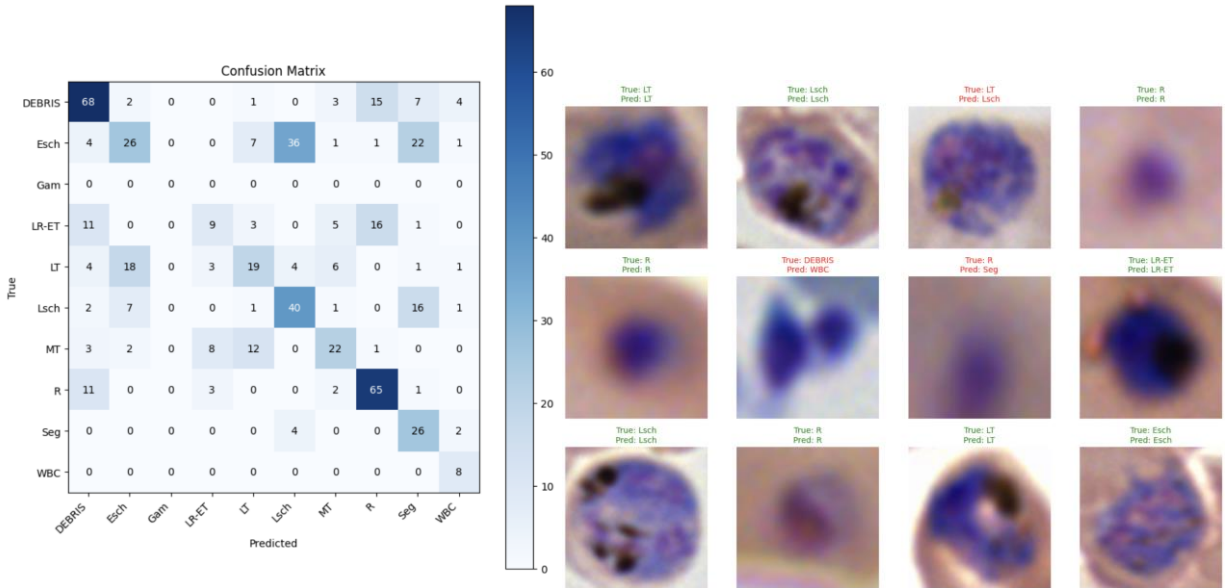
FIGURE 4: Confusion matrix and sample output patches from EfficientNet. EfficientNet frequently confused several life stages, especially among transitional trophozoite and schizont classes, reflecting difficulty in distinguishing fine-grained morphological differences. The sample outputs show that predictions were more reliable for visually distinct classes but inconsistent for subtle or rare stages.

Results from ResNet-50 model:


```
...    Test Accuracy: 0.5490909090909091

       Classification report:
                 precision    recall  f1-score   support

       DEBRIS        0.74      0.70      0.72       104
         Esch        0.51      0.18      0.26       107
        LR-ET        0.51      0.70      0.59        46
           LT        0.38      0.48      0.43        56
         Lsch        0.40      0.52      0.45        67
           MT        0.55      0.74      0.63        50
            R        0.85      0.63      0.73        82
          Seg        0.39      0.68      0.49        31
          WBC        0.55      0.86      0.67         7

     accuracy                            0.55       550
    macro avg        0.54      0.61      0.55       550
 weighted avg        0.58      0.55      0.54       550
```
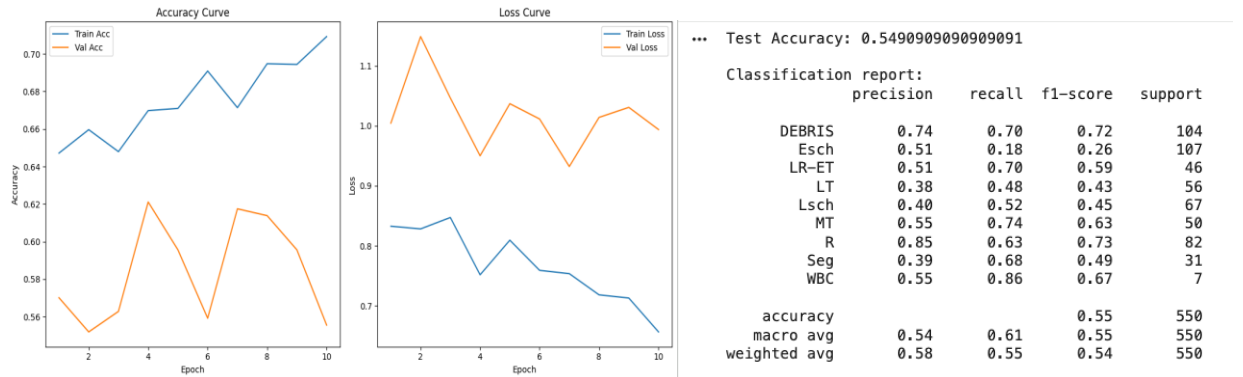
FIGURE 5: Accuracy and loss curves, as well as classification report, for ResNet. ResNet-50 demonstrated clearer learning trends than EfficientNet, with better separation between training and validation curves and accuracy near 55–59%. The classification report shows improved precision and recall in several classes; therefore, ResNet extracted more discriminative features than EfficientNet.
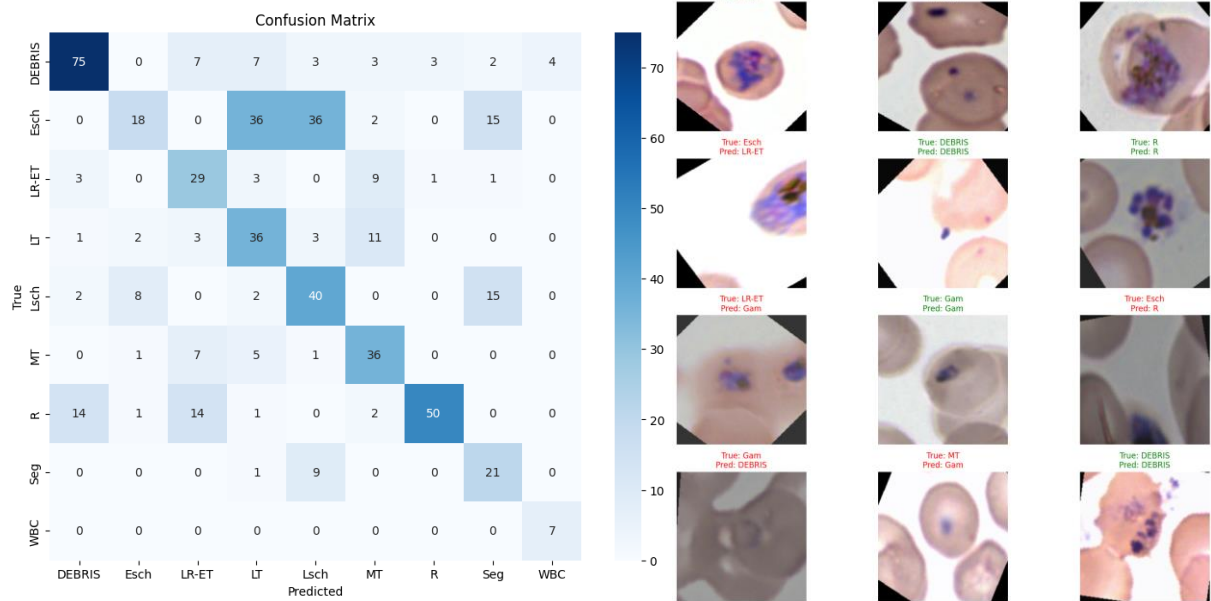
FIGURE 6: Confusion matrix and sample output patches from EfficientNet. The ResNet confusion matrix shows reduced misclassification compared to EfficientNet, but the model still struggled with rare or visually overlapping stages. Sample predictions show that ResNet handles common patterns well while misclassifying ambiguous cases.

Results when output from YOLO is used as input for ResNet:

FIGURE 7: Train loss/val loss/val accuracy plot and classification report for YOLO+ResNet model. Using YOLO-segmented patches as input significantly improved the ResNet classifier's performance, with validation accuracy stabilizing around 59% and more consistent classwise F1-scores. This confirms that providing cleaner, parasite-centered crops helped the classifier learn more stage-specific morphological features.
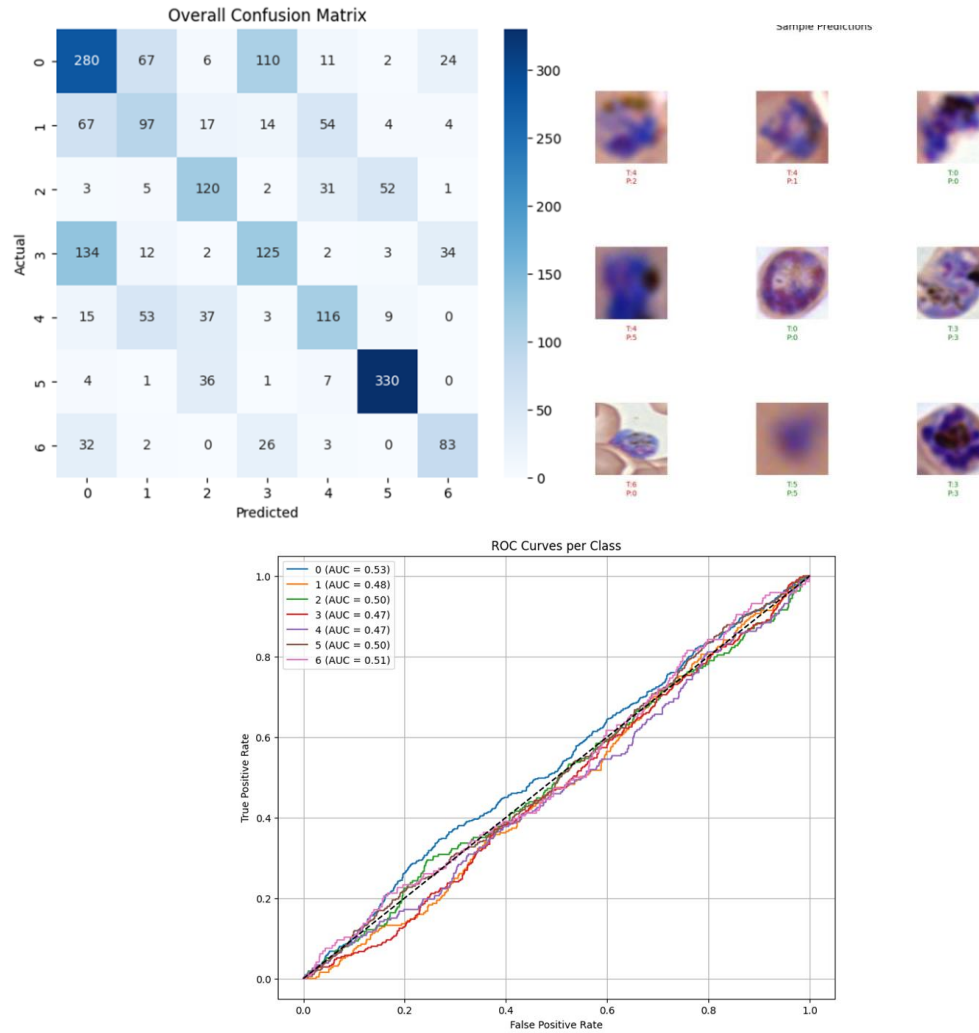
FIGURE 8: Confusion matrix, sample patches, and ROC curves per class for YOLO+ResNet model. The YOLO+ResNet confusion matrix shows a clearer diagonal structure compared to previous models, indicating improved stage differentiation. ROC curves further suggest that several classes achieved above-random discrimination, confirming that the combined pipeline strengthens classification due to more precise parasite inputs.