Determining Pavement PCIs Using a Stacking Ensemble Learning Approach

Pablo Raigoza, Devin X. Cheng, and Nubia J. Camacho-Reynaga
California State University, Chico

1 Background

According to ASTM International, ASTM D6433, the Pavement Condition Index (PCI) is a numerical indicator of the pavement condition that ranges from 0 to 100 with 0 being the worst possible condition and 100 being the best non distressed condition. Traditionally, the PCI values are calculated based on the pavement condition surveys. The more surface distresses found in the pavement, the higher the deducts from 100; therefore, the lower the PCI. With the rapid development of AI, it is possible to rate the PCI of a pavement section through computer vision and novel machine learning algorithms.

The objective is to develop an effective supervised training model utilizing top-down views of pavement image data and corresponding labeled PCIs. Then, the trained model can be used to predict the PCIs of any new pavement images.

2 Modeling Approach

As shown in Figure 1, we used two models, instance segmentation and oriented object detection, to help identify distress types such as alligator cracking, longitudinal and transverse cracking, and bricks. Our classification model was not directly trained on distress types because it could identify key features on its own. Our experiments consisted of optimizing the size of the model, the learning rate, and the image size, using different optimizers, and varying the number of epochs, etc. Our experiment results were better when using YOLOv8s-seg for segmentation, YOLOv8l-obb for oriented object detection, and YOLOv8s-cls for classification.

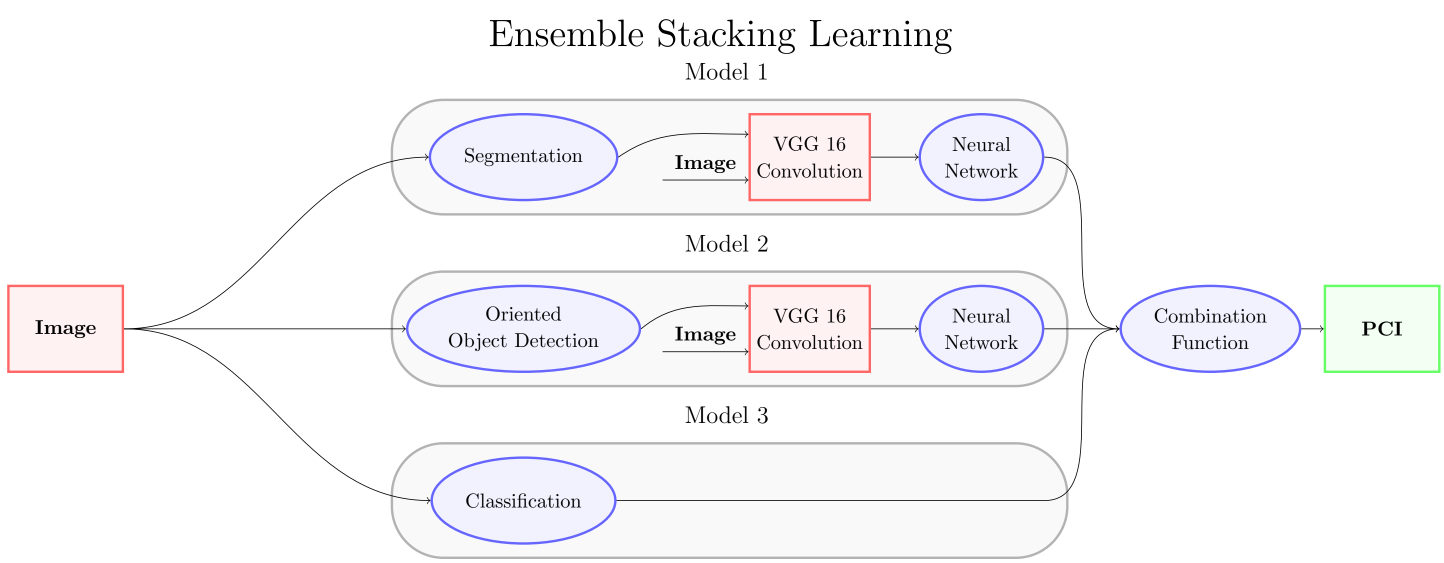

Figure 1: Red squares are unlearned inputs, blue ovals are models that we train/improve, and green is output PCI

2.1 YOLOv8 Segmentation Model and VGG 16 Convolution Model

Instance segmentation defines a mask around labeled distresses. We used the online platform Roboflow to label these distresses. We re-introduced and combined the convolved original image and the convolved annotated image provided by our instance segmentation model into a neural network which makes predictions on PCI.

2.2 YOLOv8 Object Oriented Detection Model and VGG 16 Convolution Model

Oriented object detection defines boxes around the labeled distresses, with the additional feature of identifying an angle for the distress box. The dataset we used for training the detection model was the instance segmentation dataset, which was converted to oriented object detection. Using a Residual Neural Network technique, we re-introduced the original image and the annotated image from the object detection model into a VGG16 convolution layer, whose output was flattened as the input of a standard neural network to predict PCI.

The motivation behind this model is to increase recall rates beyond the segmentation model. We did this by combining all distress types into one class to simplify the problem.

2.3 YOLOv8 Classification Model

As input into the YOLOv8 classification model we identified and filtered some images that were deemed outliers in the dataset. These images were labeled as either high PCI with high distresses, or low PCI with low distress. We also found images in similar road sections were labeled similarly; therefore, we augmented images that represented the section. We didn't augment images that misrepresent PCIs in the road section.

2.4 Combined Model with PCI Output

We took the predicted PCIs from the three models and created a function that combined these PCIs into one prediction. Through experimenting, we used a heuristic approach to combine these PCI predictions.

---

3 Modeling Results

3.1 YOLOv8 Segmentation Model

The first training dataset was labeled with common distresses found in the images (done manually with Roboflow). After data augmentation (rotation every 90 degrees and flipping horizontally and vertically), models were trained with different pretrained YOLOv8 segmentation models. Figure 2 shows an example of predicting segmented distresses with the left side test image showing the label masks, and the right side image showing predicted distresses. The overall model results are shown in Figure 3.

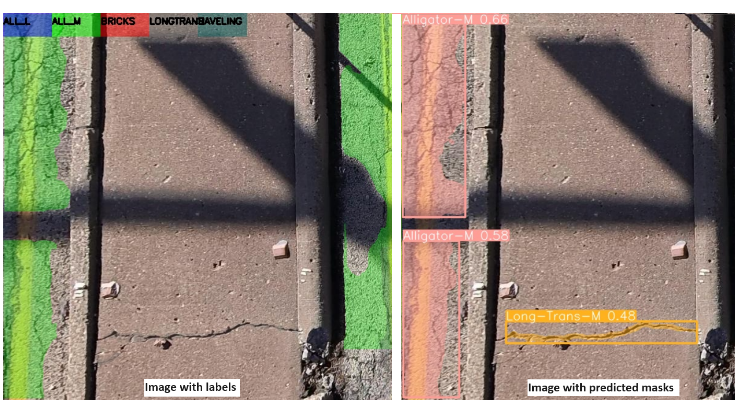

Figure 2: Segmentation prediction example

3.2 YOLOv8 Object Oriented Detection Model

Results were slightly better than most metrics predicted from the segmentation model. During training, we found that the largest model was able to predict the most accurate masks. The mAP of this model achieved on the validation set was 0.602 with precision of 0.6428 and recall of 0.5454.

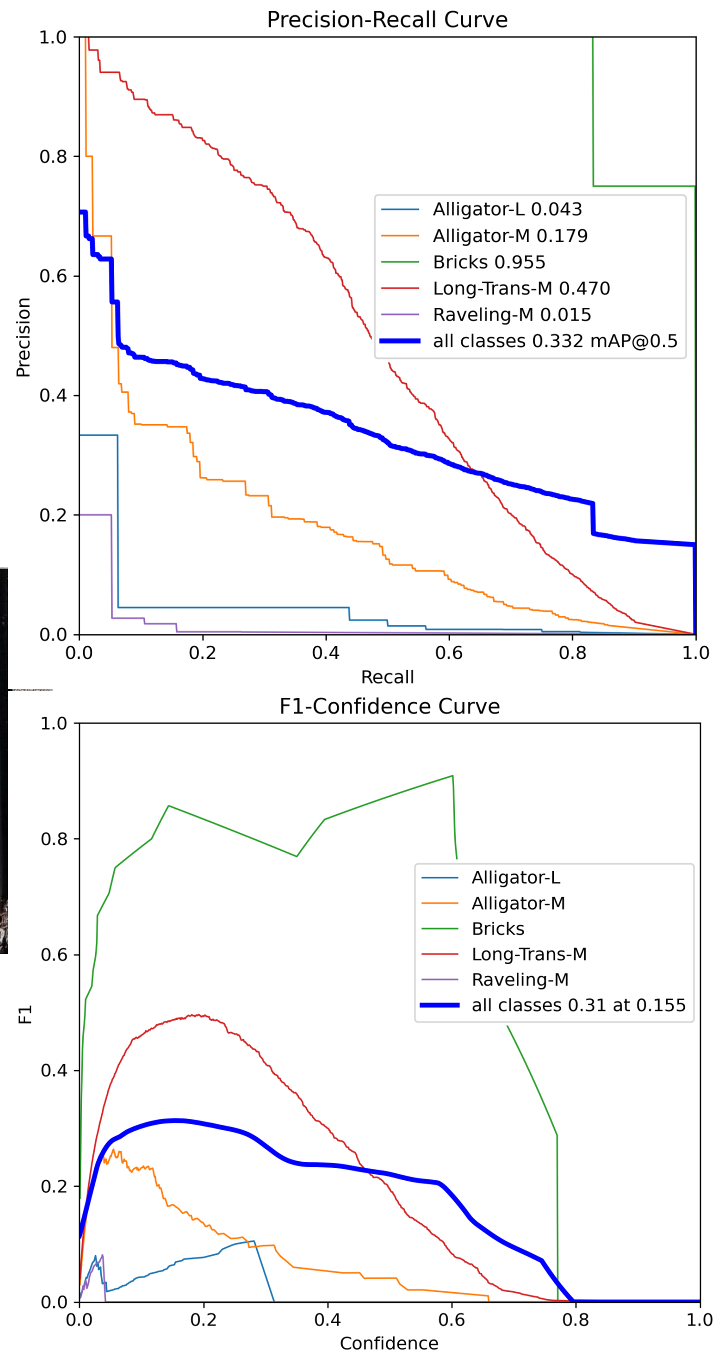

Figure 3: Precision curve and F-1 Confidence Curve

3.3 YOLOv8 Classification Model

Using the given training dataset, we conducted a series of training experiements with different sizes of YOLOv8 classification models (nano, small, medium, large, x-large), plus different Optimizers (Adam, AdamW, SGD), and plus different learning rates, moments, and epochs. Figure 4 shows the classification results. The number of epochs should not be set too high; otherwise, the model could be overtrained.

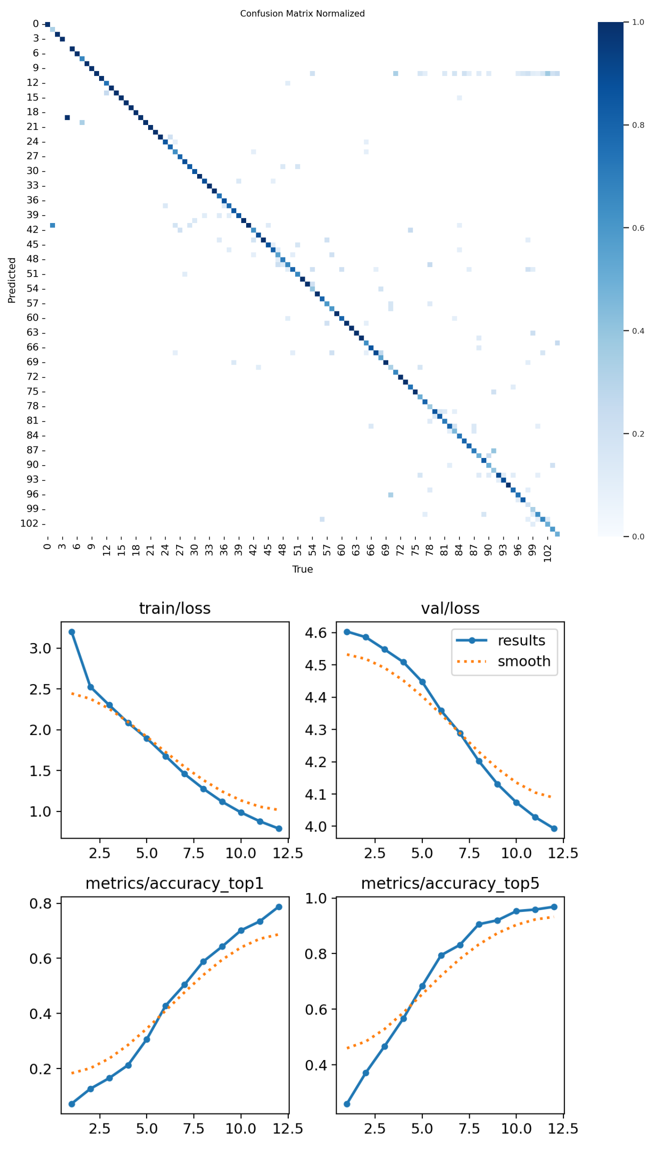

Figure 4: Confusion matrix, loss, top1, top5 accuracy of a trained YOLOv8 classification model

3.4 Combined Model with PCI Output

We found in practice that YOLOv8s-cls had the best prediction, however, it is possible to improve the PCI prediction by combining the other PCI estimates. We executed multiple trials to find a good heuristic function to calculate a combined PCI. The following equation was found to be experimentally best on the testing dataset:

$$\boxed{\text{Combined PCI} = \min(C, O) \cdot 0.95 + \min(C, S) \cdot 0.05} \quad (1)$$

where C is the classification PCI prediction, S is the segmentation PCI prediction, and O is the oriented object detection PCI prediction. Figure 5 demonstrates that YOLOv8s-cls classification model's output is 100, while the actual picture's PCI is 64. By using the combined Equation 1, it generates PCI of 65.8. Therefore, the combined PCI prediction can be considered better because combining several models increases the chances of finding distress.

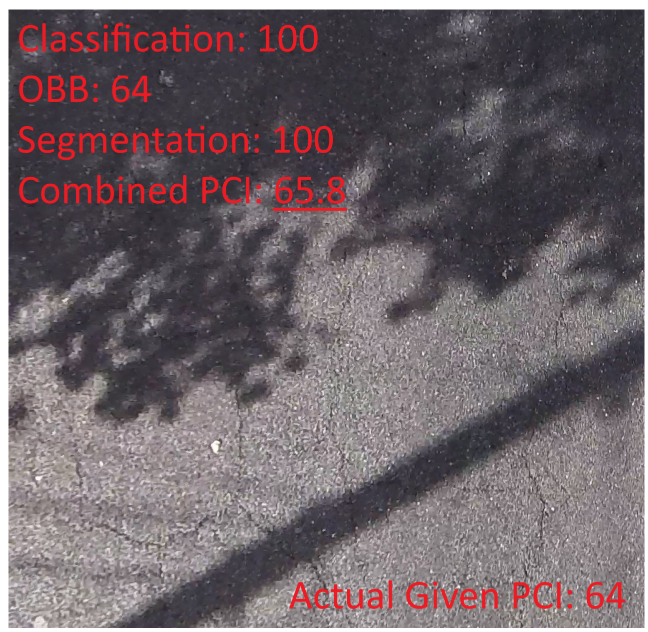

Figure 5: Example of combined model prediction for better PCI

4 Conclusions and Recommendations

Utilizing recent developments in AI, machine learning models can be developed to predict PCIs of pavement accurately and effectively. Another advantage to stacking ensemble learning is the ability for it to scale as we add more models to identify key features. To generate advanced models for PCI predictions, we need high-resolution pavement images, accurate labels, advanced machine learning architectures, and well-designed algorithms. The images provided are two dimensions without depth information. It would be beneficial to develop a PCI prediction model that considers rutting depression in the future.