

Model	@IoU=0.50			@IoU=0.75		
	Precision	Recall	F-Score	Precision	Recall	F-Score
Mask R-CNN (Ours) [7]	99.36 ± 0.36	99.86 ± 0.22	99.61 ± 0.28	94.30 ± 2.65	96.00 ± 1.74	95.14 ± 2.21
F. R-CNN w/ MobileNet [6]	99.46 ± 0.34	99.80 ± 0.21	99.63 ± 0.28	96.58 ± 0.94	97.38 ± 0.57	96.98 ± 0.75
RetinaNet w/ ResNet-50 [18]	99.30 ± 0.26	99.90 ± 0.22	99.60 ± 0.23	88.94 ± 3.68	92.48 ± 2.27	90.67 ± 2.97
SSD300 w/ VGG16 [20]	98.88 ± 0.48	99.88 ± 0.16	99.38 ± 0.31	88.72 ± 1.92	92.22 ± 1.02	90.43 ± 1.47
FCOS w/ ResNet-50 [29]	99.38 ± 0.31	99.84 ± 0.23	99.61 ± 0.26	95.62 ± 1.13	97.30 ± 0.86	96.45 ± 0.98

Table 1. Results based on bounding box borders for our Digit Area Detection (DAD) module using various models. We calculate the precision, recall, and F-score values with Intersection over Union (IoU) thresholds of 0.50 and 0.75 for 5 trials. Afterwards, we average these values and calculate their standard deviation. The dataset is split into 5 random training and testing subsets for the 5 trials. The five compared models are trained and tested on the same subsets in every trial.

Model	@IoU=0.50			@IoU=0.75		
	Precision	Recall	F-Score	Precision	Recall	F-Score
Mask R-CNN (Ours) [7]	97.94 ± 0.46	99.48 ± 0.39	98.70 ± 0.41	93.94 ± 1.75	96.22 ± 1.22	95.07 ± 1.49
F. R-CNN w/ MobileNet [6]	98.42 ± 0.95	99.26 ± 0.62	98.84 ± 0.76	94.00 ± 2.40	95.90 ± 1.49	94.94 ± 1.94
RetinaNet w/ ResNet-50 [18]	85.32 ± 2.06	99.60 ± 0.22	91.90 ± 1.27	77.38 ± 2.58	89.60 ± 2.28	83.04 ± 2.39
SSD300 w/ VGG16 [20]	96.40 ± 1.01	98.84 ± 0.70	97.60 ± 0.82	80.30 ± 2.10	84.20 ± 2.66	82.20 ± 2.32
FCOS w/ ResNet-50 [29]	94.66 ± 1.00	99.38 ± 0.27	96.96 ± 0.62	86.08 ± 3.61	92.62 ± 1.74	89.22 ± 2.73

Table 2. Results based on bounding box borders for our Digit Area Orientation Alignment (DAOA) module using various models. We calculate the precision, recall, and F-score values with Intersection over Union (IoU) thresholds of 0.50 and 0.75 for 5 trials. Afterwards, we average these values and calculate their standard deviation. The dataset is split into 5 random training and testing subsets for the 5 trials. The five compared models are trained and tested on the same subsets in every trial.

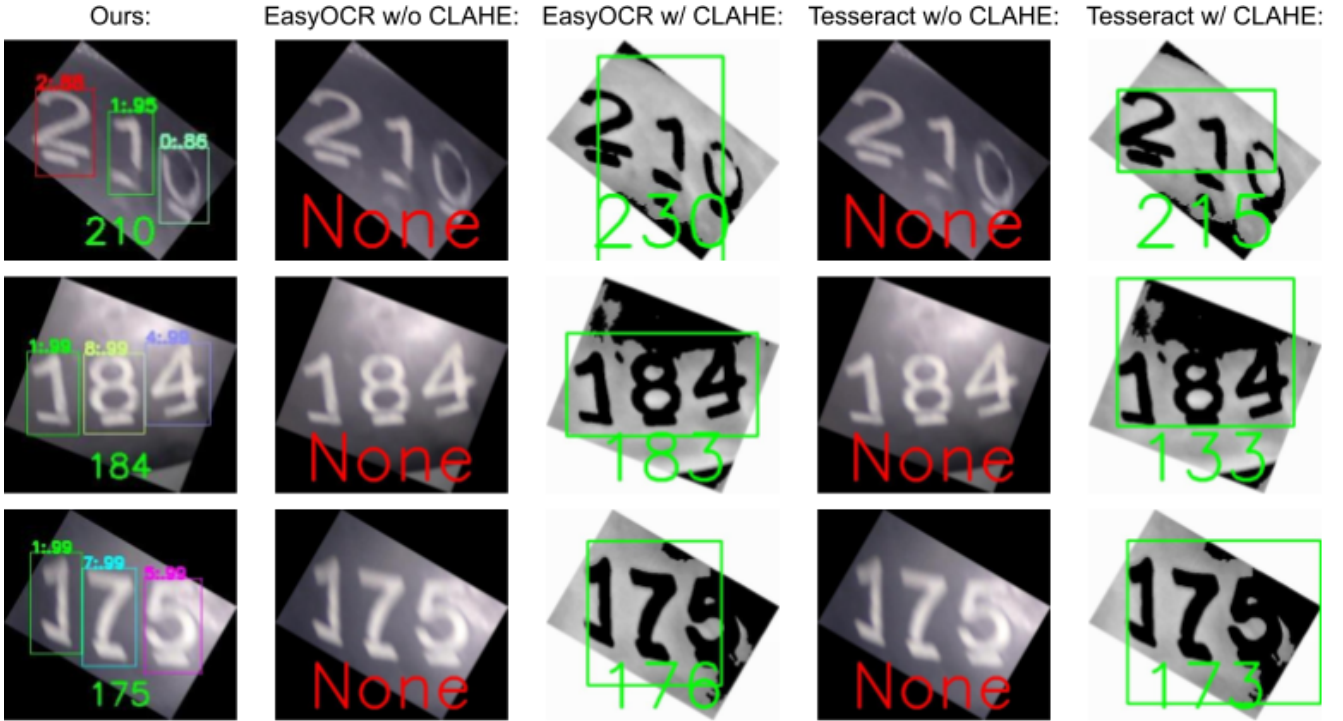


Figure 1. Column 1: Our method recognizes digits with high confidence. Column 2 and 4: EasyOCR/Tesseract recognizes nothing on the pebble digits. Column 3 and 5: Using some image preprocessing, EasyOCR/Tesseract can recognize some digits on pebbles.