HIGH-ACCURACY READING OF INDUSTRIAL ANALOG GAUGES VIA VISION-LANGUAGE SHADOW FILTERING

J P Upendra Nayanajith Jayalath[1], M P Ayomal Weerasinghe[1], K G Sarathi Heshara[1], D T Chethana Dassanayake[1], J P Upendra Nayanajith Jayalath [1*]

[1]Faculty of Computing, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka

*Corresponding author: It22362926@my.sliit.lk

Abstract

Analog clocks and radial-dial instruments remain common in industrial and everyday settings, yet automating their reading from a single photograph is complicated by two compounding sources of error: perspective distortion when the instrument is not photographed fronto-parallel, and the inherent noise of keypoint-based hand-angle estimation. Most existing pointer-reading pipelines address only one of these problems at a time. We present a four-stage pipeline that treats both explicitly. A YOLOv8-based pose model (125 training images) detects the clock face together with five alignment keypoints (center and the four cardinal hour markers), driving a homography-based rectification of the input image. A second YOLOv8-pose model (2,093 rectified images) extracts the hour- and minute-hand tip coordinates. A ResNet-18 regressor (4,540 hand crops), trained with a sine/cosine parameterization to avoid angular wraparound, refines the raw keypoint-derived hand angles. An exhaustive search over all 720 physically valid clock states resolves a final hour: minute reading and doubles as a confidence signal for adaptive fast-path/expert-path computation. The hand-keypoint model reaches a pose F1 of 0.936 on a held-out split; the angle refiner converges to a mean validation error of 5.46◦ , pooled across both hands and not yet disaggregated by hand type. The detection-and-alignment model reaches approximately 0.95 mean Average Precision at 50% IoU (mAP50), measured on its training set, since this small, 125-image dataset has no separate held-out split. A preliminary measurement shows 147.9ms of detection stage latency on general-purpose CPU hardware. We report these as per-component results and identify a unified end-toend accuracy benchmark, per-hand error disaggregation, cross instrument generalization, and field deployment as the primary remaining work.

Keywords: Analog Clock Reading, Angle Regression, Homography Rectification, Physics-Guided Validation, Pose Estimation