

Date of Submission (month day, year) : July 6, 2026

Department of Mechanical Engineering		Student ID Number D239103	Supervisors Naoki Uchiyama Junji Takahashi
Applicant's name	MUHAMMAD HAZIQ BIN NOOR AKASHAH		

Abstract (Doctor)

Title of Thesis	Image Processing and K-Nearest Neighbor Based Object-Size Estimation for Agricultural and Fishing Industrial Applications
-----------------	---

Approx. 800 words

Quality inspection in agriculture and manufacturing has traditionally relied on manual observation, contact-based measurement, or operator-dependent judgment. Although these approaches remain practical, they are often labour-intensive, subjective, and difficult to maintain consistently as production volume increases or target objects vary greatly. These limitations are critical for fragile, deformable, repetitive, or flexible products that are difficult to measure non-contact. This thesis investigates a vision-based measurement framework for converting image-derived information into physically meaningful inspection outputs. The study focuses on strawberry inspection and knotted-net measurement in net manufacturing. Although these domains differ in material properties, appearance, and operating conditions, both require visual features to be transformed into physical quantities. Therefore, this thesis treats computer vision not only as an object-recognition tool, but as the front end of a measurement pipeline combining feature extraction, calibration, and regression-based mapping. Chapter 3 establishes strawberry segmentation as the visual front end for agricultural inspection. Accurate segmentation is essential because downstream measurement depends on the extracted fruit region. Incomplete masks, merged fruit regions, or background contamination can reduce the reliability of later size- or weight-related estimates. A custom dataset of 551 strawberry images was used to compare Detectron2 and YOLOv11 as deep-learning segmentation frameworks. YOLOv11 outperformed Detectron2 in segmentation performance and efficiency, achieving 96.0% mAP@50:95, 99.4% mAP@50, 97.9% accuracy, 96.2% recall, and 98.4% F1 score, with an inference time of 0.0692 s per image. These results show that YOLO-family architectures can balance segmentation quality and speed while providing reliable fruit-region features. Chapter 4 extends the agricultural application from visual recognition to physical estimation by developing a strawberry detection and weight-estimation system using YOLOv7 and K-nearest neighbors regression. Detected strawberry regions are converted into pixel-area features and mapped to fruit weight. The YOLOv7 detector achieved mAP@0.5 of 0.991, demonstrating reliable fruit localization. Several regression models were evaluated, and KNN achieved the best test performance, with MSE = 0.898, MAE = 0.700, and R² = 0.974. For individual strawberries in the range of 7–27 g, prediction error ranged from

0.71% to 10.15%. For grouped strawberries, error ranged from 0.13% to 8.53% in organized arrangements and from -10.06% to 13.89% in closely packed arrangements. These results show that image-derived pixel area can support non-contact weight estimation when fruit is visible. However, the wider error in packed arrangements highlights the limitation of two-dimensional estimation when overlap and occlusion reduce visible area as a proxy for mass. Chapter 5 shifts the application domain from agriculture to net manufacturing. A vision-based mesh-size measurement framework was developed using lightweight round-shape-based knot localization, camera correction, and optimized KNN regression. The objective was to determine whether knot centers could serve as measurable visual features and whether inter-knot pixel distances could be converted into mesh-size estimates under structured imaging conditions. The KNN parameter study showed that $k = 2$ achieved the best calibration performance, with $MAE = 1.21\%$, $RMSE = 1.55\%$, and $R^2 = 0.995$. In practical evaluation, the optimized KNN model achieved $MAE = 1.02\%$ and $RMSE = 1.68\%$ under static conditions, and $MAE = 0.98\%$ and $RMSE = 1.66\%$ under dynamic conditions. These results demonstrate that a simple visual front end can provide accurate and stable mesh-size estimates when the target structure is repetitive and imaging is controlled. This study emphasizes temporal stability when measurement output is intended for monitoring or control. Chapter 6 extends the net-manufacturing application toward a more robust dimensional inspection framework. While the lightweight method established mesh-size measurement feasibility, its dependence on handcrafted assumptions about knot roundness, contrast, and visual separation can reduce robustness under realistic production-line conditions. To address this limitation, YOLOv11-based knot detection was integrated with centroid extraction, row-wise directional linking, distortion correction, and KNN-based real-world distance estimation. A dataset of 1,054 industrial net images was collected using a 150° ultra-wide-angle camera. YOLOv11m achieved mAP_{50} of 0.99 and mAP_{50-95} of 0.86. In production-line evaluation, the system achieved a mean detection accuracy of 94.87% and a mean link success rate of 89.80%. KNN regression with $k = 4$ achieved $MAE = 2.31\%$, $RMSE = 2.45\%$, and $R^2 = 0.995$, with prediction errors within $\pm 3.53\%$. The system visualized mesh-size deviations using color-coded links, enabling interpretation of acceptable, undersized, and oversized regions. Chapter 7 presents the generalized framework for vision-based measurement, concluding remarks of this thesis and prospective future works.

Chapter 7 presents the conclusions and future work of this thesis. The thesis proposes a unified KNN-calibrated vision-based measurement framework for natural and flexible products in primary-industry-related applications. Image processing and deep learning are used to extract measurement-related visual features, including strawberry regions, fruit pixel area, knot positions, and inter-knot pixel distances. These image-derived measurements are then calibrated or mapped to physical quantities using K-nearest neighbors regression. By applying the same measurement logic to strawberry weight estimation and knotted-net mesh-size measurement, the thesis demonstrates that low-cost monocular vision systems can provide meaningful, repeatable, and interpretable non-contact inspection outputs when the visual front end, calibration method, and regression model are designed as a complete sensing pipeline.