



DEEP LEARNING & NUTRITIONAL INTELLIGENCE

Recognition & Nutrition Estimation: A YOLO Approach on Food-101 Dataset

Dewi Nasien, Ph.D.

Lecturer of Institut Bisnis dan Teknologi Pelita Indonesia

✧ Pekanbaru, ✶ 20th October 2026



Course Outline & Presentation Flow

Learning pathway: from nutrition problem to AI-enabled dietary governance

01

Background & Motivation

Why food monitoring is difficult to scale with manual recall.

02

Supporting Theory

Nutrition assessment, computer vision, YOLO detection, and evaluation metrics.

03

Dataset & Annotation

Food-101 preparation, Roboflow bounding boxes, and data split.

04

YOLO Benchmarking

YOLOv8, YOLOv11, and YOLOv12 comparison under identical criteria.

05

Nutrition Intelligence

Mapping detected food classes to TKPI and NutriSurvey nutrient fields.

06

Implementation & Impact

Web deployment, MBG-style governance, limitations, and future work.

Core logic: Meal Image → Multi-Object Detection → Nutrition Mapping → Decision Support

Why Image-Based Nutrition Estimation Matters

Food intake, manual monitoring, and computer vision are positioned as one connected problem in image-based nutrition estimation.

Public health need

Food is more than an energy source; nutrient quality and portion adequacy influence metabolic balance.

Balanced intake helps reduce chronic disease risk such as obesity, diabetes, hypertension, and heart disease [2].

Monitoring gap

Food diaries and 24-hour recall depend on user memory and subjective estimates [1], [3].

Real meals often contain multiple items, so single-label recognition can miss important components [4].

AI opportunity

Computer vision can read visual cues such as color, texture, shape, and spatial location [5], [7].

YOLO detects multiple food objects in one image and produces bounding boxes for each item [9].

Research contribution

This study integrates YOLOv8, YOLOv11, and YOLOv12 food detection with TKPI and NutriSurvey so that visual recognition can become practical nutrition estimation.

The dietary monitoring paradox

Food intake is clinically decisive, but daily monitoring remains manual, subjective, and difficult to scale [1], [3].

45M+

students potentially
requiring daily meal audits

24-h

recall: still the main, memory-
based way to check meals

<2s

target inference for rapid
institutional verification

Status quo

- Manual diaries and 24-hour recall rely on memory.
- Food portions become untracked across schools and kitchens.
- Rations are difficult to audit without visual evidence.

AI opportunity

- Meal photos provide non-invasive evidence.
- YOLO localizes multiple dishes at once.
- TKPI mapping converts detections into nutrition profiles.

Manual recall → Meal image → Object detection → Nutrition audit

MBG and visual nutrition accountability

Why this matters for Makan Bergizi Gratis (MBG) [26]

Large-scale school meals need objective evidence that menus were served, nutritionally adequate, and auditable across kitchens, schools, and districts.

01 Policy context

MBG is a national nutrition intervention. This research contributes digital verification, not political evaluation.

02 BPS signal: Riau 2025

Food insufficiency: 10.90%.
Moderate/severe food insecurity: 4.73%. These indicators support measurable nutrition monitoring. [25]

03 System contribution

YOLO detects food items. TKPI and NutriSurvey translate detected classes into nutrition estimates and reporting evidence.

Meal photo



YOLO detection



TKPI / NutriSurvey



Audit report

Implementation logic: capture evidence → verify food components → estimate nutrition → report exceptions.

Neutral context: MBG is linked to the Prabowo Subianto administration; this study focuses on digital nutrition monitoring, not political evaluation.

Nutrition monitoring and dietary assessment

1 Dietary intake as health determinant

Food quality and quantity influence metabolic balance, chronic disease risk, growth, and readiness to learn [2].

2 Manual assessment limitation

Food diaries and 24-hour recall are useful, but depend on memory, portion estimates, and reporting discipline [3].

3 Image-based evidence

A meal photograph can provide objective visual evidence for food type, portion context, and audit traceability.

4 Nutrition informatics bridge

Detected food labels must be mapped to nutrition databases to become meaningful dietary information.

The theoretical position: image evidence reduces recall bias, while nutrition databases translate recognition into interpretable nutrient estimates.

Theory-to-System

Problem



Theory



Model



Nutritional Interpretation

Computer vision and deep learning for food images

1 Visual feature learning

Deep learning learns food patterns from color, texture, shape, edges, and spatial composition [6], [7].

2 Classification boundary

Single-label classification assumes one dominant label for the whole image, which is weak for mixed plates [8].

3 Object detection need

Detection predicts each food object and its location, enabling multi-item interpretation from one tray image [10].

4 Food image complexity

Food objects vary by size, color, texture, occlusion, and irregular boundaries, requiring robust visual representations.

Theory-to-System

Problem



Theory



Model



**Nutritional
Interpretation**

The theoretical position: food recognition moves from “what is the whole image?” to “what items are present and where are they located?”

YOLO detection theory and evaluation logic

1 One-stage detector

YOLO predicts bounding boxes, class probabilities, and confidence scores in a single forward pass [9], [11].

2 Backbone-neck-head

Backbone extracts visual features, neck fuses multi-scale information, and head outputs detections [12], [13].

3 Evaluation metrics

mAP@50 measures detection quality, precision measures correctness, recall measures completeness, and F1 balances both [24].

4 Recall in nutrition

High recall is critical because missed foods create incomplete calorie and micronutrient estimates.

Theory-to-System

Problem



Theory



Model



**Nutritional
Interpretation**

The theoretical position: for nutrition estimation, the best detector is not only accurate, but also minimizes missed food components.

From single-label classification to multi-item detection

Why conventional CNN classification fails on real cafeteria trays [10].

CNN classifier

One global label

- Assumes 1 photo = 1 dominant dish.
- Ignores side dishes when the global label is only one class.
- Creates nutrition underestimation in mixed plates.

pizza / rice / chicken?



YOLO detector

Multiple bounding boxes

- Predicts class and location in one forward pass.
- Detects multiple food items simultaneously.
- Keeps inference practical for field verification.

rice + chicken + veg + fruit

Four pillars of investigation

Systematic transition from raw pixels to institutional dietary governance

01

Benchmark dataset annotation

Curate Food-101 into detection-ready bounding boxes using Roboflow. [16], [17]

02

Comparative YOLO benchmarking

Compare YOLOv8, YOLOv11, and YOLOv12 on mAP@50, precision, recall, and F1-score.

03

Indonesian nutrition mapping

Map predicted food labels to TKPI and NutriSurvey for macro/micronutrients.

04

End-to-end deployment

Deliver a web interface with AI reasoning and administrative audit tracking.

Research scope: data → detection → nutrition interpretation → institutional audit

Research framework and experimental pipeline

Four-stage method

- 1 Data curation**
Food-101 annotated with Roboflow bounding boxes and split 80/20 [16], [17].
- 2 Model training**
YOLOv8, YOLOv11, and YOLOv12 trained at 640×640 with identical criteria.
- 3 Nutrition integration**
Predicted labels matched to TKPI and NutriSurvey nutrient fields [14], [15].
- 4 Web deployment**
Mobile capture, nutrition scoring, and daily audit trail.

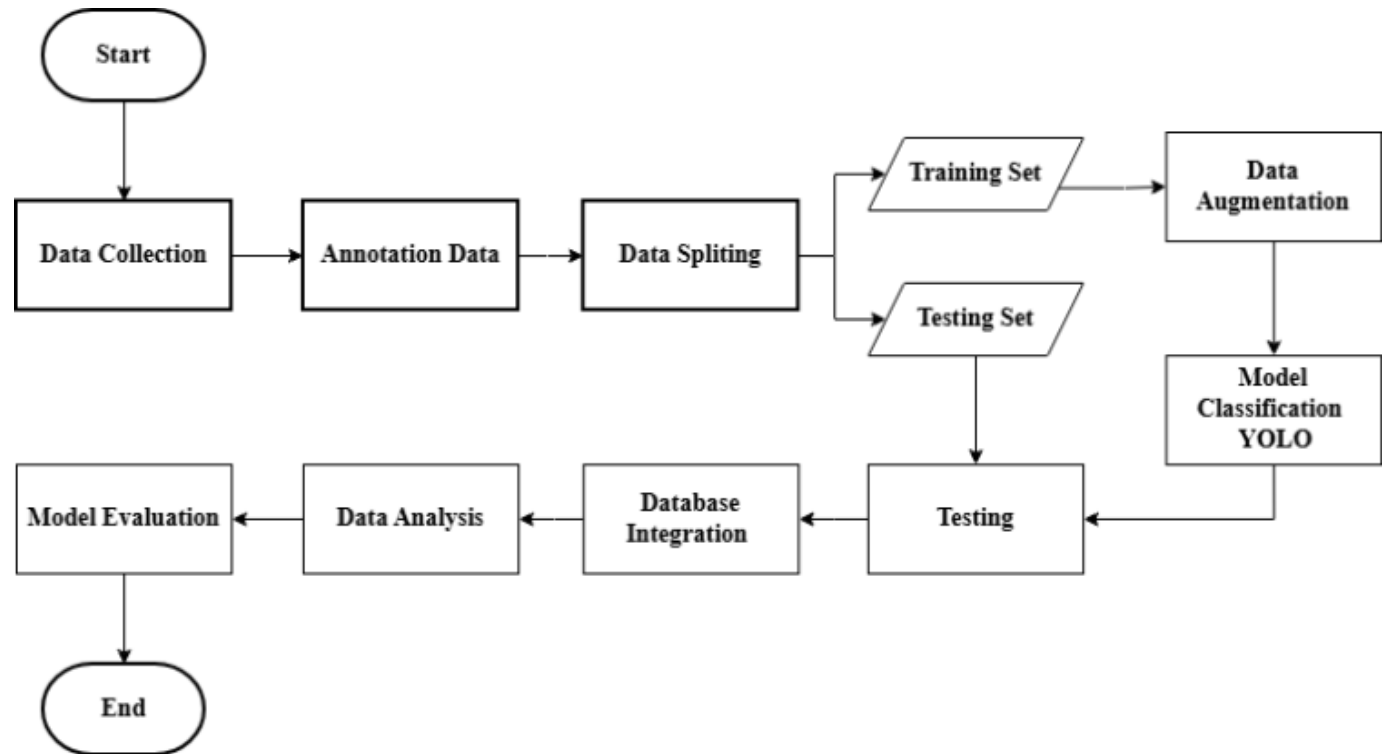
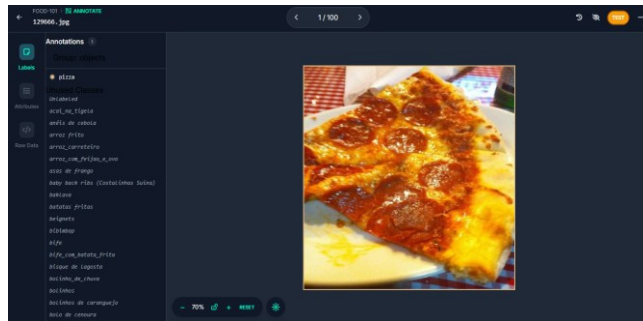


Figure: End-to-end experimental workflow from data collection to model evaluation.

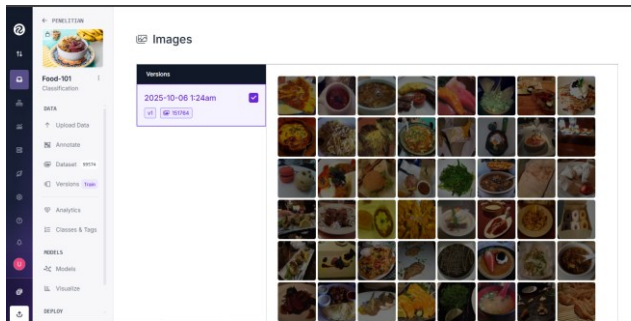
Dataset curation with Roboflow

Food-101 benchmark: 101 categories, 101,000 images, spatial annotation workflow [16]



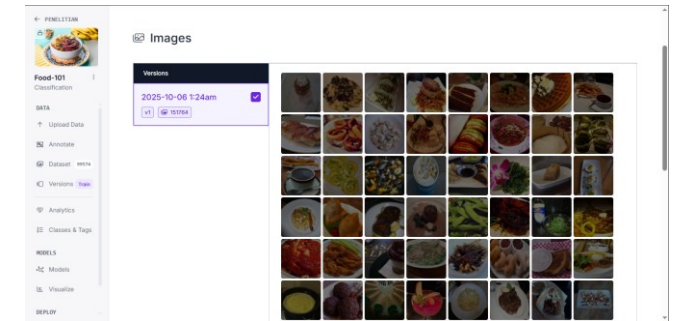
Annotation

Tight box coordinates for each dish.



80% train split

Standard augmentation and training partition.



20% validation split

Unseen data to test generalization.

Design principle: consistent labeling, consistent split, consistent evaluation.

YOLOv8 vs YOLOv11 vs YOLOv12

Anchor-free single-stage object detection evolves toward stronger recall on crowded food plates [20]-[22].

YOLOv8

Baseline

Recall
94.7%

mAP@50
96.2%

Anchor-free detection head with C2f modules; strong baseline speed [18].

YOLOv11

Efficiency focused

Recall
96.3%

mAP@50
98.8%

C3k2 blocks and optimized SPPF pooling; sharper overlapping boundaries [19].

YOLOv12

Top performer

Recall
97.1%

mAP@50
99.6%

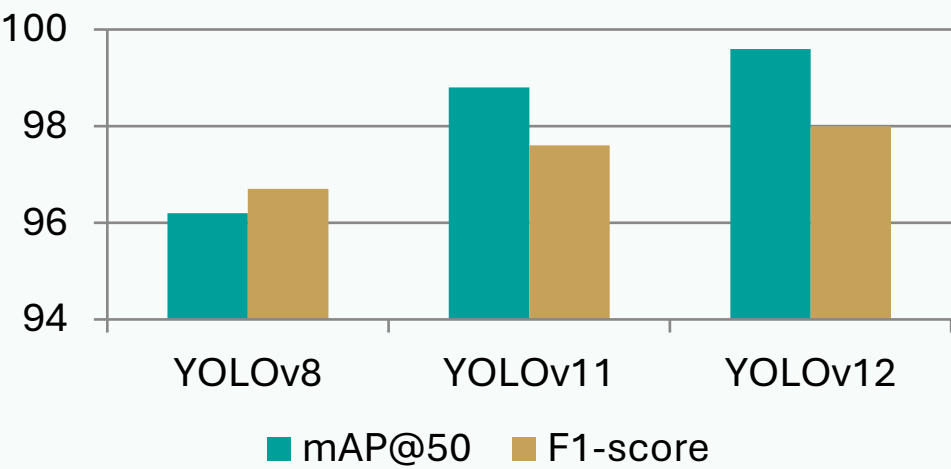
Attention-centric representations; fewer missed detections on mixed dishes [19].

Detection priority: maximize recall to avoid missing meal components that influence calorie and micronutrient estimates.

Performance comparison of YOLO models

Benchmark results on Food-101 validation metrics

Model	mAP@50	Precision	Recall	F1
YOLOv8	96.2%	99.1%	94.7%	96.7%
YOLOv11	98.8%	99.2%	96.3%	97.6%
YOLOv12 ★	99.6%	99.2%	97.1%	98.0%



Key takeaway
YOLOv12 achieves 99.6% mAP@50 and 98.0% F1-score, while preserving high recall for multi-dish food auditing.

97.1%

YOLOv12 recall

99.6%

YOLOv12 mAP@50

98.0%

YOLOv12 F1-score

TKPI & NutriSurvey knowledge engine

Translating bounding-box detections into clinical macro and micronutrient profiles



Standardized portion grammages convert class detection into interpretable nutrient estimates.

Real-world 5-dish meal tray inference

Detected dishes and calories

White rice	100g	130 kcal
Egg omelette	70g	210 kcal
Bamboo shoot	100g	45 kcal
Fried tempeh	50g	150 kcal
Salak / fruit	80g	62 kcal

597 kcal	92/100
total meal intake	meal balance score

Hasil Prediksi

Hasil Prediksi AI

Makanan Terdeteksi:

1. Nasi Putih

2. Telur

3. Toge Dadar

4. Tempe

5. Sayur Goreng

Informasi Gizi

Nasi Putih

Nasi putih sebagai makanan pokok sehari-hari

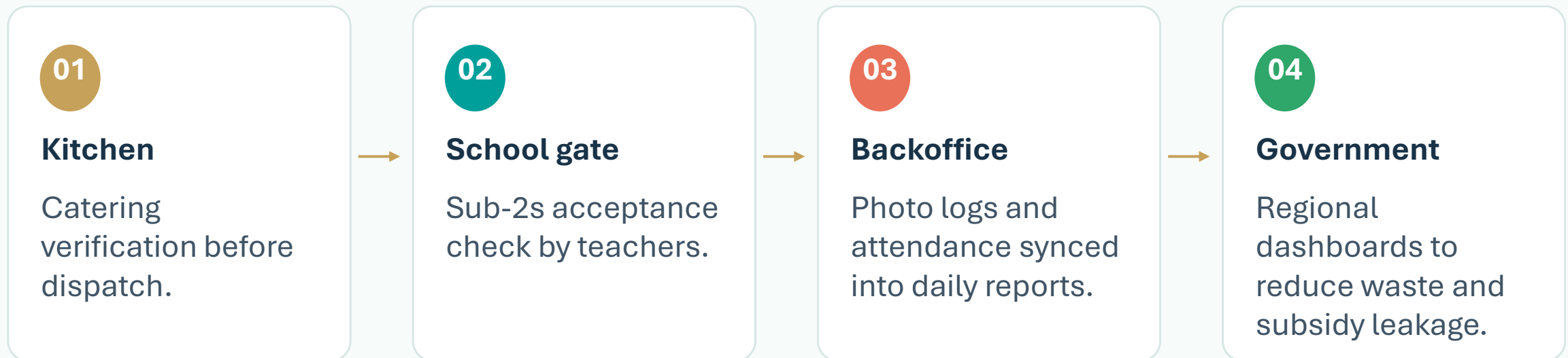
Per 100.00 gram

Energi	295.000 kkal
Protein	4.300 g
Lemak Total	0.400 g
Karbohidrat Total	45.900 g
Serat Pangan	0.600 g
Kalsium	10.000 mg
Zat Besi	0.800 mg
Natrium	1.000 mg
Kalium	55.000 mg
Vitamin A	0.000 mg
Vitamin C	0.000 mg

Status: Meets AKG breakfast recommendation range (30–35% of daily need).

National impact for MBG-style school meals

A computer vision pipeline can create multi-tier accountability from kitchen dispatch to government oversight [26].



Accountability logic: capture evidence → verify portion → estimate nutrition → report exceptions

Validated breakthroughs and future horizons

What this study validates

- One-stage detection solves the multi-dish recognition problem better than single-label classification.
- YOLOv12 achieves the strongest benchmark: 99.6% mAP@50 and 98.0% F1-score.
- Nutrition mapping enables a practical bridge from computer vision to dietary governance.

Boundaries

Current estimates use detected classes and standardized portion grammage. Future work should add depth, scale calibration, and food-waste estimation.

Future roadmap

TensorRT / ONNX edge deployment

- IoT smart kitchen scale calibration
- Offline support for remote schools
- Waste detection

REFERENCES

- [1] S. Samad, F. Ahmed, S. Naher, M. A. Kabir, A. Das, S. Amin, and S. M. S. Islam, "Smartphone apps for tracking food consumption and recommendations: Evaluating artificial intelligence-based functionalities, features and quality of current apps," *Intelligent Systems with Applications*, vol. 15, Art. no. 200103, 2022, doi: 10.1016/j.iswa.2022.200103.
- [2] World Health Organization, Regional Office for the Eastern Mediterranean, *Healthy Diet*. Cairo, Egypt: WHO Regional Office for the Eastern Mediterranean, 2019, document WHO-EM/NUT/282/E
- [3] J. Zheng, J. Wang, J. Shen, and R. An, "Artificial intelligence applications to measure food and nutrient intakes: Scoping review," *Journal of Medical Internet Research*, vol. 26, Art. no. e54557, 2024, doi: 10.2196/54557.
- [4] H. Kassem, A. A. Beevi, A. Basheer, S. Lutfi, G. Ismail, L. C. Ismail, and D. Papandreou, "Investigation and assessment of AI's role in nutrition—An updated narrative review of the evidence," *Nutrients*, vol. 17, no. 1, Art. no. 190, 2025, doi: 10.3390/nu17010190.
- [5] R. A. Shafa and P. N. Andono, "Pendeteksi visual makanan dan jumlah kalorinya menggunakan Mask R-CNN berbasis Bot Telegram," *JUPI*, vol. 10, no. 1, pp. 641–651, 2025.
- [6] V. K. Pandey et al., "Machine learning algorithms and fundamentals as emerging safety tools in preservation of fruits and vegetables," *Processes*, vol. 11, Art. no. 1729, 2023.
- [7] M. Nadeem, H. Shen, L. Choy, and J. M. H. Barakat, "Smart diet diary: Real-time mobile application for food recognition," *Applied System Innovation*, vol. 6, no. 2, Art. no. 53, 2023, doi: 10.3390/asi6020053.
- [8] N. Shukla and A. A. Moghe, "Comparative study of CNN models for defect detection in food packets," *International Research Journal on Advanced Science Hub*, vol. 5, special issue 5S, pp. 404–412, 2023, doi: 10.47392/irjash.2023.S055.
- [9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.
- [10] Y.-H. Lee and Y. Kim, "Comparison of CNN and YOLO for object detection," *Journal of Semiconductor Display Technology*, vol. 19, no. 1, pp. 85–92, 2020.
- [11] U. Sirisha, S. P. Praveen, P. N. Srinivasu, P. Barsocchi, and A. K. Bhoi, "Statistical analysis of design aspects of various YOLO-based deep learning models for object detection," *International Journal of Computational Intelligence Systems*, vol. 16, Art. no. 126, 2023, doi: 10.1007/s44196-023-00302-w.
- [12] M. L. Ali and Z. Zhang, "The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection," *Computers*, vol. 13, no. 12, Art. no. 336, 2024, doi: 10.3390/computers13120336.
- [13] R. Sapkota et al., "YOLO advances to its genesis: A decadal and comprehensive review of the You Only Look Once (YOLO) series," *Artificial Intelligence Review*, vol. 58, Art. no. 274, 2025, doi: 10.1007/s10462-025-11253-3

REFERENCES

- [14] Kementerian Kesehatan Republik Indonesia, Tabel Komposisi Pangan Indonesia. Jakarta, Indonesia: Kementerian Kesehatan RI, 2018.
- [15] NutriSurvey, NutriSurvey: Nutrient Analysis Software. [Software]. Available: <http://www.nutrisurvey.de/>
- [16] L. Bossard, M. Guillaumin, and L. Van Gool, “Food-101—Mining discriminative components with random forests,” in Computer Vision—ECCV 2014, Cham, Switzerland: Springer, 2014, pp. 446–461, doi: 10.1007/978-3-319-10599-4_29.
- [17] Roboflow, “Roboflow Universe: Open-source computer vision community.” [Online]. Available: <https://universe.roboflow.com/>. [Accessed: Oct. 9, 2026].
- [18] M. Hussain, “YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection,” Machines, vol. 11, no. 7, Art. no. 677, 2023, doi: 10.3390/machines11070677.
- [19] R. Sapkota et al., “Comprehensive performance evaluation of YOLOv12, YOLOv11, YOLOv10, YOLOv9 and YOLOv8,” Agricultural Communication, vol. 4, no. 1, 2026.
- [20] L. T. Ramos and A. D. Sappa, “A decade of You Only Look Once (YOLO) for object detection: A review,” IEEE Access, vol. 13, pp. 192747–192794, 2025, doi: 10.1109/ACCESS.2025.3630988.
- [21] S. Ćirković, S. Ćirković, and N. Stanić, “Application of the YOLO algorithm in automated detection from imaging data,” in Proc. UNITECH 2024, 2024.
- [22] A. A. Murat and M. S. Kiran, “A comprehensive review on YOLO versions for object detection,” Engineering Science and Technology, an International Journal, vol. 70, Art. no. 102161, 2025, doi: 10.1016/j.jestch.2025.102161.
- [23] A. H. N. Hidayah et al., “Disease detection using deep learning for robot vision,” Journal of Robotics and Control, vol. 3, no. 6, pp. 790–799, 2022.
- [24] L.-D. Quach, K. N. Quoc, A. N. Quynh, and H. T. Ngoc, “Evaluating the effectiveness of YOLO models in different sized object detection and feature-based classification of small objects,” Journal of Advances in Information Technology, vol. 14, no. 5, pp. 907–917, 2023, doi: 10.12720/jait.14.5.907-917.
- [25] Badan Pusat Statistik Provinsi Riau, Indikator Ketahanan Pangan dan Konsumsi Pangan Provinsi Riau. Riau, Indonesia: BPS Provinsi Riau, 2025.
- [26] N. Karmini and D. Alangkara, “Indonesia launches free meals program to feed children and pregnant women to fight malnutrition,” Associated Press, Jan. 6, 2025. [Online]. Available: <https://apnews.com/article/213a04587203434f3f85950725e84a8b>