VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7930

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

COMPARATIVE ANALYSIS OF YOLOV5SM, YOLOV8, AND YOLOV11 FOR IMAGE-BASED TEMPEH QUALITY RECOGNITION

Isa Mahfudi[1*]; Mila Kusumawardani[1]; Moechammad Sarosa[1]; Chandrasena Setiadi[1]

Electrical Engineering Department[1]
Politeknik Negeri Malang, Malang, Indonesia[1]
https://www.polinema.ac.id/[1]
isa_mahfudi@polinema.ac.id[*], mila.kusumawardani@polinema.ac.id, msarosa@polinema.ac.id, chandrasenasetiadi@polinema.ac.id

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract— *Tempeh is a traditional Indonesian fermented food whose quality is influenced by fermentation and environmental conditions. Quality assessment is still commonly performed manually, leading to subjectivity and inconsistency. This study compares three modern object detection models—YOLOv5sM, YOLOv8, and YOLOv11—for digital image–based tempeh quality recognition. A dataset of 1,000 images (500 good and 500 defective) was collected using a Logitech C270 camera under controlled lighting conditions. YOLOv5sM was trained with data augmentation (Mosaic, flip, rotation), while YOLOv8 and YOLOv11 were trained without augmentation to isolate architectural differences. All models were trained for 100 epochs using identical hyperparameters and evaluated on a 10% test set. Results show that YOLOv11 achieved the highest accuracy (98%), outperforming YOLOv8 (94%) and YOLOv5sM (88%). Although mAP@0.5 reached 99.5% across models, stricter evaluation using mAP@0.5:0.95 revealed performance differences (96.2%, 96.9%, and 97.0%, respectively). The superior performance of YOLOv11 is attributed to its C3K2 and C2PSA modules, which enhance fine-grained feature extraction and localization precision. These findings indicate that YOLOv11 is the most suitable architecture for automated tempeh quality inspection.*

Keywords*: Augmentation, Computer Vision, Object Detection, Tempeh Quality, YOLO*

Intisari—*Tempe merupakan produk pangan fermentasi tradisional Indonesia yang kualitasnya sangat dipengaruhi oleh kondisi fermentasi dan faktor lingkungan. Penilaian kualitas tempe masih umumnya dilakukan secara manual, sehingga rentan terhadap subjektivitas dan inkonsistensi. Penelitian ini bertujuan membandingkan tiga model deteksi objek modern—YOLOv5sM, YOLOv8, dan YOLOv11—untuk pengenalan kualitas tempe berbasis citra digital. Dataset yang digunakan terdiri dari 1.000 citra (500 tempe baik dan 500 tempe cacat) yang dikumpulkan menggunakan kamera Logitech C270 dengan sistem akuisisi citra dan pencahayaan terkontrol. YOLOv5sM dilatih menggunakan teknik augmentasi data (Mosaic, flip, rotasi), sedangkan YOLOv8 dan YOLOv11 dilatih tanpa augmentasi untuk mengisolasi pengaruh perbedaan arsitektur. Seluruh model dilatih selama 100 epoch dengan hiperparameter yang sama dan dievaluasi menggunakan 10% data uji. Hasil menunjukkan bahwa YOLOv11 memperoleh akurasi tertinggi (98%), diikuti YOLOv8 (94%) dan YOLOv5sM (88%). Meskipun nilai mAP@0.5 seluruh model mencapai 99,5%, evaluasi lebih ketat menggunakan mAP@0.5:0.95 menunjukkan perbedaan performa (96,2%; 96,9%; dan 97,0%). Kinerja superior YOLOv11 dikaitkan dengan modul C3K2 dan C2PSA yang meningkatkan ekstraksi fitur tekstur halus dan presisi lokalisasi. Dengan demikian, YOLOv11 direkomendasikan untuk sistem inspeksi kualitas tempe otomatis.*

Kata Kunci*: Augmentasi, Kualitas Tempe, Penglihatan Komputer, Pendeteksian Objek, YOLO.*

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7930

INTRODUCTION

Tempeh is one of Indonesia's traditional fermented food products with high nutritional value, particularly as a source of plant-based protein, B-complex vitamins, minerals, and various bioactive compounds beneficial to human health [1], [2], [3], [4], [5]. It is widely consumed across many regions and plays an important role in both household-scale and commercial-scale food industries [6]. However, the quality of tempeh is strongly influenced by fermentation conditions such as temperature, humidity, ventilation, and the cleanliness of raw materials [7]. Deviations from these parameters may lead to fermentation failure, producing tempeh with brownish discoloration, black spots, cracked texture, unpleasant odor, and structures that are either excessively soft or overly firm [8]. Such failures not only reduce sensory quality but also cause economic losses, particularly for small producers who rely on daily batch-based production.

Despite the importance of maintaining tempeh quality in the food supply chain, quality evaluation is still predominantly conducted through manual inspection based on visual observation and producer experience. This traditional method is subjective, non-standardized, requires specific expertise, and is prone to inconsistent judgments [9], [10], [11]. Differences among evaluators in assessing color, mycelium patterns, and fermentation level often lead to misclassification of marketable products. These limitations create an urgent need for automated detection technology [9] based on computer vision and deep learning [11], [12], [13], [14], which can assess tempeh quality objectively, rapidly, and accurately [15].

In recent years, rapid advancements in computer vision and deep learning have enabled the development of digital image–based food quality inspection systems [11], [15]. One-stage object detectors such as MobileNet-SSD, EfficientDet-Lite, and the You Only Look Once (YOLO) family have been widely adopted across various applications due to their balance of accuracy, speed, and computational efficiency. Since its introduction in 2016, YOLO has evolved through multiple generations. YOLOv2 introduced architectural refinements and more stable anchor boxes, while YOLOv3 improved performance using the Darknet-53 backbone. YOLOv4 leveraged augmentation techniques such as Mosaic and the CSPDarknet backbone to achieve higher accuracy [16], and YOLOv5 became a widely used PyTorch implementation due to its ease of deployment [17], [18]. Subsequent iterations include YOLOv6, which focuses on speed, YOLOv7 with superior real-time performance [19], YOLOv8 featuring an anchor-free design [20], and the more recent YOLOv9 [21], YOLOv10 [22], and YOLOv11 [23], [24] all offering improved accuracy and computational efficiency for diverse application scenarios.

Several studies have compared the performance of YOLO models across different domains, demonstrating that each version offers unique advantages in terms of accuracy and computational efficiency. YOLOv5, for instance, has been widely adopted in automotive and autonomous vehicle applications due to its stable and fast performance [25], [26]. YOLOv8, as a newer generation with anchor-free architecture, has been reported to achieve improved mAP and inference speed in tasks such as weed classification, plant disease detection [27], and food quality inspection. YOLOv11, the latest generation, provides enhanced architectural efficiency and improved detection accuracy, making it increasingly relevant for real-world deployment on resource-constrained devices [28].

Recent work has undertaken systematic evaluations of modern YOLO architectures to determine their relative performance across various detection tasks. For example, Casas et al. compared YOLOv5 and YOLOv8 on corrosion segmentation benchmarks, reporting notable differences in mean average precision (mAP) and F1-score across diverse settings [29]. Similarly, Wijaya et al. evaluated YOLOv5 and YOLOv8 variants for road damage detection, highlighting performance trade-offs between accuracy and inference speed [30]. More recent studies that include YOLOv11 in comparative assessments demonstrate that architectural enhancements such as improved feature aggregation and lightweight attention modules can lead to higher detection accuracy and computational efficiency compared to earlier versions [31]. A comprehensive review of the YOLO framework traces its development from YOLOv5 through YOLOv11, underscoring systematic advances in backbone design, feature fusion strategies, and deployment efficiency [32]. Recent benchmarking studies explicitly comparing YOLOv5, YOLOv8, and YOLOv11 across defect-detection domains demonstrate consistent performance improvements in newer architectures, particularly in mAP@0.5:0.95 and inference efficiency [17], [28], [31]. However, no study has evaluated these three architectures specifically for tempeh quality recognition.

Despite these advances in YOLO benchmarking, research on automated tempeh quality assessment remains very limited. To date, no study has systematically compared multiple modern YOLO generations—particularly YOLOv5sM, YOLOv8, and YOLOv11—for distinguishing between high-quality and defective tempeh using digital images. This gap raises an important question: which YOLO architecture provides the optimal balance between detection accuracy, computational efficiency, and deployment feasibility for real-world tempeh inspection systems?

Based on this gap, the present study aims to conduct a comparative analysis of three modern YOLO models—YOLOv5sM, YOLOv8, and YOLOv11—for digital image–based tempeh quality recognition. These architectures were selected because they represent major evolutionary stages of the YOLO family: YOLOv5sM as a stable PyTorch-based baseline, YOLOv8 as a more efficient anchor-free design, and YOLOv11 as the newest model with enhanced feature representation and computational efficiency. Each model was trained on a dataset of tempeh images consisting of two main classes, "good tempeh" and "defective tempeh," under controlled experimental conditions. Performance evaluation was conducted using accuracy metrics (precision, recall, mAP@0.5, mAP@0.5:0.95) and computational parameters (model size, inference speed, and resource requirements). This study tests the hypothesis that newer architectures—particularly YOLOv11—will outperform earlier generations in detection accuracy and computational efficiency for tempeh quality recognition.

The novelty of this research lies in the in-depth comparative analysis of multiple YOLO architectures for tempeh quality detection, a domain that has not been explored previously. The main contributions of this study include:

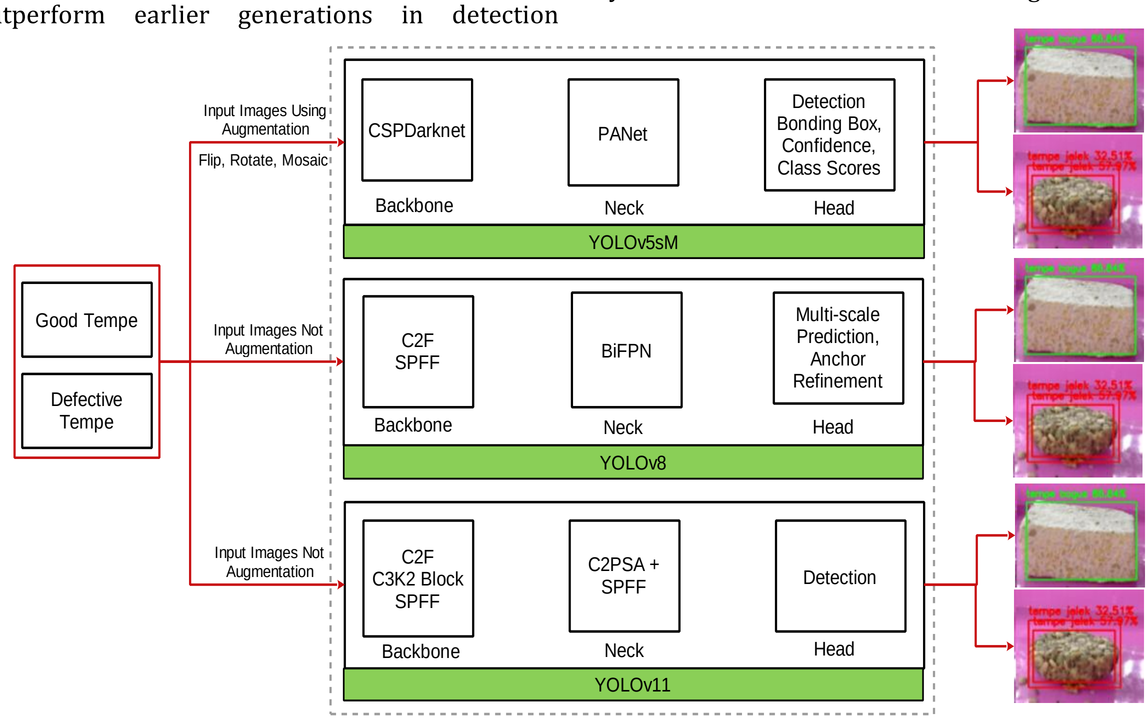
1. Implementing several modern YOLO architectures specifically for tempeh quality classification;
2. Analyzing detection performance based on accuracy and computational efficiency metrics; and

identifying the most optimal YOLO model for automated tempeh inspection systems, including its feasibility for deployment on resource-constrained devices.

MATERIALS AND METHODS

Research Design

This study employed a comparative experimental design to evaluate the performance of three YOLO architectures—YOLOv5sM, YOLOv8, and YOLOv11—using the same tempeh image dataset. The workflow consisted of dataset acquisition, preprocessing, model training, and comparative evaluation based on detection accuracy and computational efficiency. The overall system architecture is illustrated in Figure 1.



Source : (Research Results, 2025)

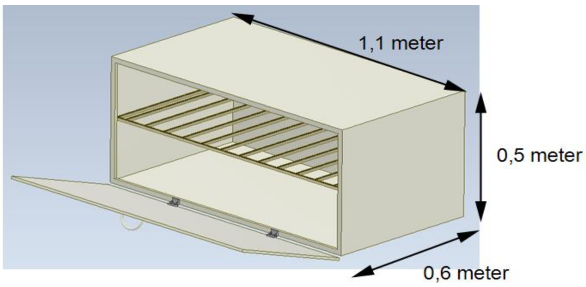
Figure 1. System Architecture for the Comparison of Three YOLO Models

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7930

Dataset and Preprocessing

A total of 1,000 tempeh images were collected, consisting of 500 good tempeh samples and 500 defective tempeh samples. The images were captured using a Logitech C270 camera with a native resolution of 1280 × 720 pixels, positioned 15 cm above the sample tray in a perpendicular (90° top-down) orientation to ensure consistent image geometry. The acquisition process was conducted inside a closed enclosure to eliminate external light interference, using controlled white LED illumination with a color temperature of 6500K and no ambient light exposure during capture. Prior to training, all images were resized to 640 × 640 pixels. The dataset was divided into 70% training, 20% validation, and 10% testing sets using stratified sampling to maintain balanced class distribution across all subsets. For YOLOv5sM, data augmentation techniques including Mosaic, horizontal flip, and rotation were applied during training, whereas YOLOv8 and YOLOv11 were trained using non-augmented images to isolate the influence of architectural differences on model performance. Figure 2 presents the design and dimensions of the fermentation box, while Figure 3 illustrates the actual image acquisition setup.



Source : (Research Results, 2025)

Figure 2. The design of the fermentation box along with its dimensions

Source : (Research Results, 2025)

Figure 3. The actual condition of the fermentation box, LED illumination, and camera position during image acquisition

Model Training

All three models were trained using identical hyperparameters to ensure fair comparison. Training was conducted for 100 epochs, with a batch size of 16 and a learning rate of 0.01 using the SGD optimizer. The selected hyperparameters follow commonly recommended configurations for medium-scale object detection tasks, balancing training stability and computational efficiency on limited GPU resources. Early stopping was not applied to maintain consistent training duration across models. For each architecture, the checkpoint achieving the highest validation mAP was selected for final evaluation.

Model evaluation

Model performance was evaluated exclusively on the test set, which was not used during training or validation. The evaluation metrics included:

a. Precision
b. Recall
c. mAP@0.5
d. mAP@0.5:0.95

Inference speed and model size were also recorded to assess computational efficiency. Comparative analysis was conducted to determine differences in detection accuracy and generalization capability across architectures.

Training Environment

All experiments were performed on Google Colab using an NVIDIA Tesla T4 GPU (16 GB memory). The models were implemented in Python using the PyTorch framework and the Ultralytics YOLO library. Each model was trained separately under identical hardware and software conditions to ensure consistency.

RESULTS AND DISCUSSION

Training Performance

The training phase evaluates how effectively each YOLO model learns visual patterns from the tempeh image dataset. Model convergence is reflected by decreasing training loss and increasing mAP values across epochs. Table 1 summarizes the training loss, mAP metrics, and computation time per epoch for YOLOv5sM, YOLOv8, and YOLOv11 over 100 epochs.

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7930

Table 1. Training Performance Across Epochs for YOLOv5sM, YOLOv8, and YOLOv11

Epoch	Model	Train Loss	Train mAP@0.5 (%)	Train mAP@0.5:0.95 (%)	Time/Epoch (s)
10	YOLOv5sM	2.684	63.2	39.8	38
	YOLOv8	2.103	72.4	48.6	28
	YOLOv11	1.842	78.9	56.1	22
20	YOLOv5sM	1.812	76.5	58.9	38
	YOLOv8	1.312	84.6	66.4	28
	YOLOv11	0.982	89.7	72.5	22
30	YOLOv5sM	1.214	84.1	66.2	38
	YOLOv8	0.874	91.4	74.3	28
	YOLOv11	0.621	94.2	81.1	22
40	YOLOv5sM	0.923	88.9	71.6	38
	YOLOv8	0.642	94.7	78.9	28
	YOLOv11	0.452	96.8	85.4	22
50	YOLOv5sM	0.754	92.1	75.4	38
	YOLOv8	0.512	96.5	82.3	28
	YOLOv11	0.345	97.9	87.6	22
60	YOLOv5sM	0.612	94.3	78.2	38
	YOLOv8	0.402	97.7	85.9	28
	YOLOv11	0.221	98.8	90.4	22
70	YOLOv5sM	0.511	96.2	81.1	38
	YOLOv8	0.318	98.4	88.3	28
	YOLOv11	0.165	99.3	92.8	22
80	YOLOv5sM	0.436	97.1	83.9	38
	YOLOv8	0.251	98.9	90.2	28
	YOLOv11	0.121	99.5	94.7	22
90	YOLOv5sM	0.382	97.9	85.2	38
	YOLOv8	0.196	99.1	91.0	28
	YOLOv11	0.092	99.7	95.6	22
100	YOLOv5sM	0.331	98.4	86.4	38
	YOLOv8	0.158	99.3	92.2	28
	YOLOv11	0.061	99.8	96.1	22

Source : (Research Results, 2025)

Table 1 shows that all models achieved stable convergence, reflected by decreasing training loss and increasing mAP values. At epoch 100, YOLOv11 obtained the lowest loss (0.061) and highest mAP@0.5 (99.8%), outperforming YOLOv8 and YOLOv5sM. In addition, YOLOv11 demonstrated the fastest training time (22 s/epoch), compared to YOLOv8 (28 s) and YOLOv5sM (36 s), indicating more efficient learning and improved architectural design.

Validation Phase

The validation phase evaluates the generalization capability of each model using data not seen during training. This stage assesses performance stability and potential overfitting under more realistic conditions. Table 2 presents the validation loss and mAP metrics for YOLOv5sM, YOLOv8, and YOLOv11 across different IoU thresholds.

Table 2. Validation Performance Across Epochs for YOLOv5sM, YOLOv8, and YOLOv11

Epoch	Model	Val Loss	Val mAP@0.5 (%)	Val mAP@0.5:0.95 (%)
10	YOLOv5sM	3.124	59.6	41.5
	YOLOv8	2.284	69.8	52.7
	YOLOv11	1.914	78.2	62.1
20	YOLOv5sM	2.015	76.2	58.8
	YOLOv8	1.482	83.7	70.4
	YOLOv11	1.133	88.9	77.3
30	YOLOv5sM	1.427	84.3	67.5
	YOLOv8	0.964	92.1	81.2
	YOLOv11	0.674	94.8	85.6
40	YOLOv5sM	1.032	89.7	74.1
	YOLOv8	0.712	94.9	83.9
	YOLOv11	0.458	96.7	88.3
50	YOLOv5sM	0.823	92.8	77.6
	YOLOv8	0.524	96.4	86.7
	YOLOv11	0.346	97.9	90.5
60	YOLOv5sM	0.653	94.6	80.2
	YOLOv8	0.398	97.8	88.9
	YOLOv11	0.244	98.7	92.8
70	YOLOv5sM	0.518	95.9	82.7
	YOLOv8	0.317	98.3	90.1
	YOLOv11	0.168	99.2	94.3

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7930

Epoch	Model	Val Loss	Val mAP@0.5 (%)	Val mAP@0.5:0.95 (%)
80	YOLOv5sM	0.432	96.8	84.4
	YOLOv8	0.256	98.7	91.8
	YOLOv11	0.127	99.4	95.2
90	YOLOv5sM	0.378	97.5	85.9
	YOLOv8	0.212	99.1	92.7
	YOLOv11	0.095	99.6	95.9
100	YOLOv5sM	0.334	98.1	86.5
	YOLOv8	0.176	99.3	93.4
	YOLOv11	0.071	99.8	96.4

Source : (Research Results, 2025)

Table 2 confirms that YOLOv11 demonstrates the strongest generalization performance. At epoch 100, it achieved the lowest validation loss (0.071), outperforming YOLOv8 (0.176) and YOLOv5sM (0.334). Similarly, YOLOv11 obtained the highest mAP@0.5:0.95 (96.4%), compared with YOLOv8 (93.4%) and YOLOv5sM (86.5%), indicating more precise localization under stricter IoU thresholds. This performance advantage is attributed to architectural enhancements, particularly the C3K2 block for efficient feature extraction and the C2PSA attention mechanism, which improves sensitivity to fine-grained defects such as subtle texture irregularities. The use of GIoU loss further enhances bounding-box regression stability, contributing to improved detection precision and reduced misclassification.

Training and Validation Performance

Overall, consistent performance trends were observed across training and validation phases. YOLOv11 maintained the lowest loss and highest mAP values in both stages, indicating stable convergence and strong generalization capability. The relatively small gap between training and validation performance suggests that no significant overfitting occurred, and that architectural improvements in YOLOv11 contribute to more efficient feature representation and localization precision.

Confusion Matrix Analysis

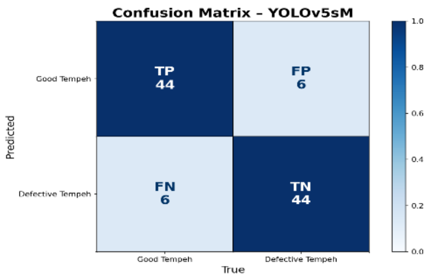
The confusion matrix was used to evaluate classification performance on the 100-image test set (50 good and 50 defective samples). Table 3 summarizes the results in terms of True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN).

Table 3. Summary of Confusion Matrix Results for the Three YOLO Models

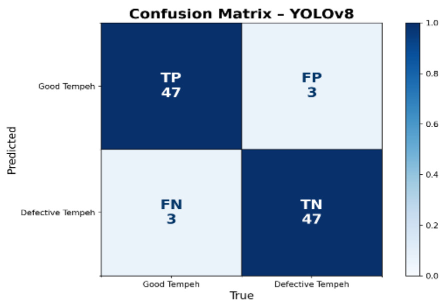
Model	TP	FP	FN	TN	Accuracy
YOLOv5sM	44	6	6	44	88%
YOLOv8	47	3	3	47	94%
YOLOv11	49	1	1	49	98%

Source : (Research Results, 2025)

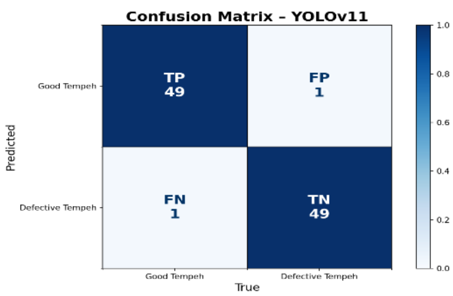
YOLOv11 achieved the highest test accuracy (98%), with only one false positive and one false negative. In comparison, YOLOv8 achieved 94%, while YOLOv5sM reached 88%. This difference indicates that newer architectures provide improved robustness in distinguishing subtle texture variations between good and defective tempeh samples. The visualizations of the confusion matrices for each model can be seen in Figures 4, 5, and 6.



Source : (Research Results, 2025)
Figure 4. Confusion Matrix of YOLOv5sM Model



Source : (Research Results, 2025)
Figure 5. Confusion Matrix of YOLOv8 Model



Source : (Research Results, 2025)
Figure 6. Confusion Matrix of YOLOv11 Model.

Classification Report Evaluation

Table 4 presents the classification performance of each model on the test dataset using precision, recall, and mAP metrics. Overall, all three models demonstrate high detection performance for digital image–based tempeh quality recognition. The precision and recall values are consistently high across architectures, reflecting strong classification capability in distinguishing good and defective tempeh samples.

Table 4. Classification Report Evaluation for YOLOv5sM, YOLOv8, and YOLOv11

Model	Precision (%)	Recall (%)	mAP@0.5 (%)	mAP@0.5:0.95 (%)	Time (s)
YOLOv5sM	88	88	99.5	96.2	13585
YOLOv8	94	94	99.5	96.9	3272
YOLOv11	98	98	99.5	97.0	3841

Source : (Research Results, 2025)

Table 4 presents the classification performance on the test dataset. Precision and recall values are consistent with the confusion matrix results, where YOLOv11 achieved 98%, YOLOv8 reached 94%, and YOLOv5sM obtained 88%. These results confirm that newer YOLO architectures provide improved classification robustness in distinguishing subtle texture variations between good and defective tempeh samples.

Practical Implications for Tempeh Inspection

The findings of this study have direct implications for automated tempeh quality control systems. The superior localization performance of YOLOv11, particularly under stricter IoU thresholds, indicates its ability to detect subtle surface defects such as uneven mycelium growth, small cracks, and localized discoloration. In practical production environments, such defects may not always be consistently identified through manual inspection, which is often subjective and dependent on operator experience. The reduced false detection rate observed in YOLOv11 enhances reliability in real-world applications, minimizing the risk of misclassifying defective products as acceptable. This is particularly important for small- and medium-scale tempeh producers, where consistent quality assurance directly influences consumer trust and economic sustainability. Furthermore, the demonstrated computational efficiency suggests that the model can potentially be deployed on edge devices for real-time inspection systems, supporting automated sorting or quality grading during production.

CONCLUSION

This study conducted a comparative evaluation of three modern YOLO architectures—YOLOv5sM, YOLOv8, and YOLOv11—for automated tempeh quality recognition using digital images. The results demonstrate that architectural advancements significantly influence detection accuracy and localization precision. Among the evaluated models, YOLOv11 consistently achieved the highest performance across training, validation, and test phases, particularly in mAP@0.5:0.95 and overall classification accuracy. The findings confirm that newer YOLO architectures provide improved robustness in detecting subtle surface defects, such as uneven mycelium growth and localized discoloration. The relatively small generalization gap between training and validation results indicates stable learning behavior without significant overfitting. Therefore, YOLOv11 is identified as the most suitable architecture for automated tempeh quality inspection systems, especially in environments requiring reliable and consistent quality control. From a practical perspective, the superior detection performance and computational efficiency of YOLOv11 support its potential deployment in real-time inspection systems, including implementation on edge-based devices for small- and medium-scale tempeh production. However, this study was conducted using a controlled dataset with limited environmental variability. Future research should explore larger and more diverse datasets, including varying lighting conditions, fermentation stages, and different tempeh packaging types. Additionally, investigating model optimization techniques such as quantization or pruning may further enhance deployment feasibility in low-resource industrial environments.

REFERENCE

[1] O. H. Kristiadi and A. T. Lunggani, "Tempe Kacang Kedelai Sebagai Pangan Fermentasi Unggulan Khas Indonesia: Literature Review," *J. Andaliman-Jurnal Gizi Pangan, Klin. Dan Masy.*, vol. 2, no. 2, pp. 48–56, 2022, doi: 10.24114/jgpkm.v2i2.40334.

[2] A. Romulo and R. Surya, "International Journal of Gastronomy and Food Science Tempe : A traditional fermented food of Indonesia and its health benefits," *Int. J.*

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7930

Gastron. Food Sci., vol. 26, no. May, 2021, doi: 10.1016/j.ijgfs.2021.100413.

[3] R. Ratnaningsih, N. Kusumawaty, A. C. Iwansyah, E. R. N. Herawati, and D. Kristanti, "History, manufacture, nutritional content , bioactive compounds , and health benefits of tempeh and tofu as alternative protein in Indonesia : a review," *Aust. J. Crop Sci.*, vol. 19, no. 07, pp. 839–852, 2025, doi: 10.21475/ajcs.25.19.07.p381.

[4] J. Suliburska, I. A. Harahap, and E. Capanoglu, "Fermented soy products : A review of bioactives for health from fermentation to functionality," *Compr. Rev.*, vol. 24, no. September, pp. 1–24, 2025, doi: 10.1111/1541-4337.70080

[5] S. Q. Teoh, N. L. Chin, C. W. Chong, A. M. Ripen, S. How, and J. J. L. Lim, "A Review on Health Benefits and Processing of Tempeh with Outlines on Its Functional Microbes," *Futur. Foods*, vol. 9, p. 100330, 2024, doi: 10.1016/j.fufo.2024.100330.

[6] K. Górska, E. Pejcz, and J. Haasym, "Tempeh and Fermentation — Innovative Substrates in a Classical Microbial Process," *Appl. Sci.*, vol. 15, pp. 1–24, 2025, doi: 10.3390/app15042171.

[7] A. M. Jacob, C. Gaby, H. Rachman, J. Sulistyo, and O. Krisbianto, "Determination of optimum fermentation time through microscopic, sensory, and eating quality comparison of hyacinth beans, sword beans , and soybean tempeh," *AGROINTEK*, vol. 19, no. 4, pp. 937–949, 2025, doi: 10.21107/agrointek.v19i4.32918.

[8] M. Astawan, A. P. G. Prayudani, P. Hadiningtias, and T. Wresdiyati, "Effect of tea extract (Camellia sinensis) on shelf life and intrinsic quality of tempeh," *Canrea J. Food Technol. Nutr. Culin.*, vol. 7, no. 1, pp. 1–14, 2024, doi: 10.20956/canrea.v7i1.1349.

[9] K. Shehzad, U. Ali, and A. Munir, "Computer Vision for Food Quality Assessment : Advances and Challenges," *Glob. J. Mach. Learn. Comput.*, vol. 1, no. 1, pp. 76–92, 2025, doi: 10.70445/gjmlc.1.1.2025.76-92.

[10] A. M. Mhaskar, P. P. Mangrulkar, R. E. Rane, and R. M. Sairise, "Computer Vision in Food Quality Control: Applications and Innovations," *IJFANS Int. J. Food Nutr. Sci.*, vol. 11, no. 8, pp. 3830–3846, 2022, doi: 10.12345/IJFANS.2022.11.08.

[11] K. G. Liakos, V. Athanasiadis, E. Bozinou, and S. I. Lalas, "Machine Learning for Quality Control in the Food Industry : A Review," *Foods*, vol. 14, pp. 1–35, 2025, doi: 10.3390/foods14060934.

[12] Y. Hendrawan *et al.*, "Classification of soybean tempe quality using deep learning," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 924, no. 1, 2021, doi: 10.1088/1755-1315/924/1/012022.

[13] C. H. S. A, R. Farah, and Z. Nizam, "Machine Vision for Quality Control in Halal Food Production : A Deep Learning Approach," *J. Moeslim Res. Tech.*, vol. 2, no. 3, pp. 108–117, 2025, doi: 10.70177/technik.v2i3.2351.

[14] D. Liu, E. Zuo, D. Wang, L. He, L. Dong, and X. Lu, "Deep Learning in Food Image Recognition: A Comprehensive Review," *Appl. Sci.*, vol. 15, no. 14, pp. 1–18, 2025, doi: 10.3390/app15147924.

[15] L. Zhu, P. Spachos, E. Pensini, and K. N. Plataniotis, "Current Research in Food Science Deep learning and machine vision for food processing : A survey," *Curr. Res. Food Sci.*, vol. 4, no. March, pp. 233–249, 2021, doi: 10.1016/j.crfs.2021.03.009.

[16] J. Woo, J. Baek, S. Jo, and S. Y. Kim, "A Study on Object Detection Performance of YOLOv4 for Autonomous Driving of Tram," *Sensors*, vol. 22, no. 22, pp. 9026–9037, 2022, doi: 10.3390/s22229026.

[17] R. Khanam, T. Asghar, and M. Hussain, "Comparative Performance Evaluation of YOLOv5 , YOLOv8 , and YOLOv11 for Solar Panel Defect Detection," *Solar*, vol. 5, no. 6, pp. 1–25, 2025, doi: 10.3390/solar5060076.

[18] Y. Zhang, Z. Guo, J. Wu, Y. Tian, H. Tang, and X. Guo, "Real-Time Vehicle Detection Based on Improved YOLO v5," *Sustainability*, vol. 14, no. 19, pp. 12274–12293, 2022, doi: 10.3390/su141912274.

[19] C. Wang, A. Bochkovskiy, and H. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023, pp. 1–15, doi: 10.1109/CVPR52729.2023.00921.

[20] J. Ryu, D. Kwak, and S. Choi, "YOLOv8 with Post-Processing for Small Object Detection Enhancement," *Appl. Sci.*, vol. 15, pp. 1–13, 2025, doi: 10.3390/app15116091.

[21] V. Ponduri, L. Mohan, S. Basheera, V. R. V, U. Chandan, and K. Praveena, "Highly efficient YOLOV9 model for detecting extremely small-scale objects," *Eng. Res. Express*, vol. 7, no. 1, pp. 15287–15302, 2025, doi: 10.1088/2631-8695/adc223.

[22] Q. Wu and X. Nie, "Improved YOLOv10 : A Real-Time Object Detection Approach in Complex Environments," *Sensors*, vol. 25, pp.

1–17, 2025, doi: 10.3390/s25010169.

[23] L. He, Y. Zhou, L. Liu, and J. Ma, "Research and Application of YOLOv11-Based Object Segmentation in Intelligent Recognition at Construction Sites," *Buildings*, vol. 14, no. 12, pp. 3777–3812, 2024, doi: 10.3390/buildings14123777.

[24] I. Mahfudi, B. Helmiananda, D. Marya, M. Kusumawardani, and C. Setiadi, "Rancang Bangun Sistem Pengenalan Ekspresi Bayi Berbasis YOLOv11 Untuk Pemantauan Jarak Jauh," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 6, pp. 10533–10540, 2025, doi: 10.36040/jati.v9i6.16509.

[25] S. Swain and A. K. Tripathy, "Real-time lane detection for autonomous vehicles using YOLOV5 Segmentation Model," *Int. J. Sustain. Eng.*, vol. 17, no. 1, pp. 718–728, 2024, doi: 10.1080/19397038.2024.2311905.

[26] Z. Ren, H. Zhang, and Z. Li, "Improved YOLOv5 Network for Real-Time Object Detection in Vehicle-Mounted Camera Capture Scenarios," *Sensors*, vol. 23, no. 10, pp. 4589–4605, 2023, doi: 10.3390/app14188368.

[27] L. Gao *et al.*, "A Lightweight YOLOv8 Model for Apple Leaf Disease Detection," *Appl. Sci.*, vol. 14, no. 15, pp. 6710–6731, 2024, doi: 10.3390/app14146190.

[28] R. Reveles-Martínez, H. Gamboa-Rosales, E. Sánchez-Femat, J. Saldívar-Pérez, T. Ibarra-Pérez, L. C. Reveles-Gómez, O. A. Guirette-Barbosa, J. I. Galván-Tejada, C. E. Galván-Tejada, H. Luna-García, and J. M. Celaya-Padilla, "Benchmarking YOLOv8 to YOLOv11 Architectures for Real-Time Traffic Sign Recognition in Embedded 1:10 Scale Autonomous Vehicles," Technologies, vol. 13, no. 11, p. 531, 2025, doi: 10.3390/technologies13110531.

[29] E. Casas, L. Ramos, C. Romero, and F. Rivas-echeverría, "A comparative study of YOLOv5 and YOLOv8 for corrosion segmentation tasks in metal surfaces," *Array*, vol. 22, p. 100351, 2024, doi: 10.1016/j.csite.2024.104894.

[30] Y. Indrihapsari, D. Wijaya, S. A. Ardy, and I. I. Siswanto, "A Performance Comparison of YOLOv5 and YOLOv8 for Road Damage Object Detection on a Mixed GRDDC – PUPR Dataset," *Elinvo (Electronics, Informatics, Vocat. Educ.*, vol. 10, no. 2, pp. 147–168, 2025.

[31] F. Aldi, I. Nozomi, M. Hafizh, and T. Novita, "Comparative Analysis of YOLOv11 with Previous YOLO in the Detection of Human Bone Fractures," *J. Comput. Networks, Archit. High Perform. Comput.*, vol. 7, no. 3, pp. 777–790, 2025, doi: 10.54254/2755-2721/136/2025.21668.

[32] 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, pp. 336–373, 2024, doi: 10.3390/app15052669.