Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Deep Learning-Based Non-Destructive Seafood Freshness Assessment: A Systematic Review of Sensing Technologies, Validation, and Deployment Readiness

Muhammad Munsarif

m.munsarif@unimus.ac.id

Universitas Muhammadiyah Semarang

Richmond Ampah-Mensah

Universitas Muhammadiyah Semarang

Research Article

Keywords: seafood freshness, deep learning, non-destructive assessment, computer vision, hyperspectral imaging, deployment readiness

Posted Date: October 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11263422/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Seafood freshness assessment is essential for quality control, shelf-life management, and supply-chain decisions, yet conventional sensory, chemical, and microbiological measurements are often destructive or laboratory-dependent. This systematic review synthesizes evidence on deep learning (DL)-based non-destructive seafood freshness assessment, focusing on sensing modality, model architecture, predictive performance, reference-standard validity, methodological robustness, and deployment readiness. Scopus and Web of Science records published in English from 2021 to 2026 were screened. Of 1,057 identified records, 306 remained after filtering and deduplication; 32 full texts were assessed and 28 studies were included. Four technology families emerged: red-green-blue (RGB) computer vision; hyperspectral, near-infrared, Raman, and fluorescence spectroscopy; colorimetric or fluorescent intelligent sensors; and multimodal or olfactory-sensor fusion. Reported performance was often high, with classification accuracies above 95% and regression R^2 values above 0.90 in several studies, but comparisons were limited by heterogeneous species, freshness labels, datasets, and validation designs. Studies anchored to total volatile basic nitrogen (TVB-N), total viable count/aerobic plate count (TVC/APC), K-value, thiobarbituric acid reactive substances (TBARS), or biogenic amines provided stronger freshness validation than storage-time or visual-class labels alone. Methodological appraisal classified 22 studies as some concern and six as high concern, with no study classified as low concern. Mobile, edge, portable-sensor, and robotic implementations show translational promise, but external and industrial validation remain limited. Technological maturity therefore depends on biological validity, independent validation, computational feasibility, and real-world deployment, not predictive accuracy alone.

1. Introduction

Seafood freshness is a multidimensional quality state shaped by sensory deterioration, biochemical reactions, and microbial growth. Recent non-destructive studies therefore use reference indicators such as total volatile basic nitrogen (TVB-N), total viable count/aerobic plate count (TVC/APC), K-value, thiobarbituric acid reactive substances (TBARS), and biogenic amines to connect sensor signals with measurable deterioration (Fan et al. 2024; Han et al. 2025b).

Conventional freshness determination can provide biologically meaningful measurements, but destructive sampling, laboratory requirements, analysis time, and expert dependence constrain continuous or point-of-use monitoring. This has stimulated interest in optical and sensor-based approaches that infer freshness without destroying the product. The recent evidence base spans RGB imaging, hyperspectral and near-infrared spectroscopy, fluorescence, Raman spectroscopy, colorimetric sensing, fluorescent tags, and electronic-nose systems (Moosavi-Nasab et al. 2021; Yu et al. 2026).

Visible appearance provides an accessible signal for freshness assessment. Fish eyes, gills, skin, and whole-product surfaces undergo color and texture changes that can be learned by convolutional neural networks (CNNs), while smartphone acquisition offers a low-cost path toward practical inspection (Akgül et al. 2023; Zheng et al. 2025). Spectral sensing extends this principle by detecting wavelength-dependent molecular or biochemical changes that may precede obvious visual deterioration (Guo et al. 2025; Han et al. 2025b).

The modeling landscape has evolved from transfer-learning CNNs toward lightweight YOLO variants, CNN-LSTM hybrids, Transformers, and multimodal attention-based fusion. Transfer learning is particularly relevant when biological datasets are limited, whereas lightweight networks explicitly address memory and latency constraints

(Gao et al. 2024; Hamidy et al. 2026). Transformer models have also emerged for hyperspectral freshness prediction, where self-attention can model long-range spectral dependencies (Jiang et al. 2026).

However, predictive performance alone is an incomplete measure of maturity. Very high classification accuracy may reflect storage-time labels, visually separable classes, or tightly controlled acquisition rather than direct prediction of sample-specific biochemical freshness (Genç et al. 2025a; Zheng et al. 2025). Studies directly anchored to TVB-N, TVC, K-value, TBARS, or spoilage metabolites provide a stronger bridge to established freshness science (Han et al. 2025a; Hao et al. 2025).

Generalizability is another unresolved issue. Many systems remain species-, device-, laboratory-, or storage-condition specific, and external validation is less common than internal train/test evaluation (Jiang et al. 2026; Nandi et al. 2026). Similarly, claims of real-time or industrial potential exceed the number of studies that actually benchmark models on smartphones, embedded hardware, portable platforms, or automated sorting systems (He and Wei 2026; Yu et al. 2026).

1.1 Biological basis of freshness deterioration

Seafood deterioration involves microbial proliferation, protein and nucleotide degradation, volatile-base formation, lipid oxidation, metabolite accumulation, and sensory changes. TVB-N and TVC/APC capture chemical and microbiological deterioration, K-value reflects nucleotide degradation, TBARS reflects lipid oxidation, and biogenic amines provide metabolite-level evidence of spoilage (Fan et al. 2024; Guo et al. 2025). These targets are complementary rather than interchangeable.

1.2 Non-destructive sensing pathways

RGB methods learn visible changes in eyes, gills, skin, texture, and surface color, whereas NIR and hyperspectral systems capture spectral interactions associated with tissue composition and biochemical deterioration (Akgül et al. 2023; Han et al. 2025b). Fluorescence and Raman measurements provide additional molecular information, while colorimetric or fluorescent sensor arrays convert volatile spoilage chemistry into camera-readable optical patterns (Ma et al. 2021; Zhang et al. 2026). Electronic-nose systems instead represent freshness through volatile profiles (He and Wei 2026).

1.3 Deep-learning architectures

CNNs remain foundational for spatial, color, texture, and sensor-pattern representation. YOLO-based architectures integrate localization and classification and can be optimized for efficient inspection (Gao et al. 2024; Hou et al. 2025). CNN-LSTM hybrids combine local feature extraction with sequential dependency modeling, while Transformers use self-attention to model long-range relationships in spectral inputs (Fan et al. 2024; Jiang et al. 2026). MobileNet, GhostNet, and related designs target resource-constrained deployment, and cross-attention or gated fusion architectures integrate complementary modalities (Yu et al. 2026; Zou et al. 2026).

This review therefore evaluates DL-based non-destructive seafood freshness assessment through an integrated framework linking biological target and sensing modality to model architecture, predictive performance, freshness-reference validity, methodological robustness, generalizability, and deployment readiness. The central argument is that a mature system must demonstrate more than high accuracy: it must show that the predicted signal is biologically meaningful, independently validated, computationally feasible, and transferable to realistic operating conditions.

1.4 Research questions

RQ1. How have deep learning-based non-destructive methods been applied for seafood freshness assessment, and how can these approaches be categorized according to seafood species, anatomical targets, sensing modalities, and deep-learning architectures?

RQ2. How do the identified approaches compare in predictive performance and in validation against established freshness indicators?

RQ3. What are the principal methodological limitations, generalizability constraints, and indicators of deployment readiness?

2. Materials and methods

2.1 Review design and information sources

A systematic literature review was conducted using PRISMA 2020 as the reporting framework (Page et al. 2021). Scopus and Web of Science were used as the information sources. Eligibility was restricted to peer-reviewed English-language articles published from 2021 through 2026. The review used predefined research questions, eligibility criteria, structured data extraction, methodological appraisal, and narrative/comparative synthesis. No review protocol was prospectively registered; the procedures used for screening, extraction, appraisal, and synthesis are reported below.

2.2 Search strategy

The search combined three concept blocks: seafood products; deep learning/computer vision/imaging; and freshness/quality/non-destructive assessment. The core Boolean expression was: ("seafood" OR "fish" OR "fishes" OR "shellfish" OR "shrimp" OR "prawn" OR "squid" OR "aquatic product*" OR "marine product*") AND ("deep learning" OR "convolutional neural network*" OR "CNN" OR "vision transformer*" OR "ViT" OR "YOLO" OR "computer vision" OR "image analysis" OR "hyperspectral imag*") AND ("freshness" OR "freshness assessment" OR "freshness evaluation" OR "freshness prediction" OR "quality assessment" OR "quality inspection" OR "spoilage" OR "TVB-N" OR "total volatile basic nitrogen" OR "non-destructive" OR "nondestructive").

2.3 Eligibility criteria

Studies were included when they investigated seafood or aquatic food products, addressed freshness or spoilage, used a non-destructive sensing or imaging component, incorporated deep learning or directly relevant DL-centered computer vision, reported quantitative model evaluation, were original peer-reviewed articles, were written in English, and were published during 2021–2026. Exclusions comprised non-seafood products, studies unrelated to freshness deterioration, destructive-only laboratory assessment, conventional machine-learning-only studies outside the DL scope, reviews and other non-original publications, and studies without sufficient quantitative performance information.

2.4 Study selection and deduplication

Database records were filtered by publication year, article type, and language. Deduplication followed a DOI-first procedure, supplemented by normalized-title matching and author/year checks. Titles and abstracts of unique records were screened against the eligibility criteria, followed by full-text assessment. Full-text reports were

assessed against the eligibility criteria. Four reports were excluded: two because they did not use deep learning, one because the assessment was destructive-only, and one because it did not report sufficient quantitative performance results.

Study selection was conducted by two authors using a sequential two-reviewer process. M.M. performed the initial title-and-abstract screening against the eligibility criteria, and R.A.-M. reviewed the records retained for full-text eligibility assessment. Because screening was sequential rather than independent parallel screening, inter-rater agreement was not calculated. Borderline cases and disagreements were rechecked by both authors against the eligibility criteria and resolved by consensus.

2.5 Data extraction

Extraction covered bibliographic characteristics; seafood species and anatomical target; storage design; sensing modality; preprocessing and feature strategy; DL architecture; prediction task; freshness reference standard; classification or regression performance; validation design; model size or inference information where available; limitations; and deployment evidence. Classification outcomes included accuracy, precision, recall, F1-score, AUC, and mAP, whereas regression outcomes included R^2, RMSE/RMSEP, MAE, and RPD.

M.M. initially extracted and compiled the study information, and R.A.-M. checked the extracted data against the corresponding included articles. Disagreements were reviewed jointly and resolved by consensus. Information not reported in the original studies was recorded as not reported and was not inferred.

2.6 Methodological appraisal and deployment-readiness assessment

An author-developed methodological appraisal framework was applied across six domains: sample representativeness and biological independence (D1), freshness reference-standard validity (D2), data-split independence and leakage control (D3), external/generalizability validation (D4), reporting transparency (D5), and validity of deployment claims (D6). The framework was developed for transparent comparative synthesis and was not treated as a validated risk-of-bias instrument. Each domain was judged as low concern, some concern, or high concern. Low concern indicated that the relevant methodological feature was adequately reported and no material threat to the central study claim was identified; some concern indicated incomplete reporting or a plausible limitation whose effect on the central claim was uncertain; and high concern indicated a clear limitation likely to materially affect interpretation. The overall judgment reflected the highest-concern domain in relation to its likely effect on the central study claim. Appraisal was performed sequentially, with initial evidence characterization by M.M. and second-author checking by R.A.-M.; borderline or discrepant judgments were rechecked against the source article and domain definitions and resolved by consensus. Domain judgments considered biological-unit independence and representativeness (D1), the strength of objective freshness anchoring (D2), specimen-level split independence and leakage control (D3), the independence and breadth of validation (D4), sufficiency of methodological reporting (D5), and whether deployment claims were demonstrated on real hardware or practical prototypes rather than only proposed (D6). Deployment evidence was considered stronger when a study implemented or benchmarked a system on a smartphone, edge platform, portable sensor, or robotic prototype rather than only proposing future use.

2.7 Evidence synthesis

Because the included studies differed substantially in species, sensing technologies, freshness definitions, architectures, datasets, and metrics, statistical pooling was not considered appropriate. Evidence was therefore

synthesized narratively and comparatively across the linked dimensions of modality, architecture, validation, and deployment.

Deployment readiness was evaluated as a combination of reference-standard validity, generalizability, computational feasibility, and practical implementation evidence rather than predictive performance alone. This framing guided the comparative synthesis of sensing, validation, and deployment evidence.

2.8 Use of AI-assisted tools

During manuscript preparation, OpenAI ChatGPT was used to assist with manuscript organization and language/readability refinement. The tool did not make autonomous decisions regarding study inclusion or exclusion, data extraction, methodological appraisal, source verification, or scientific interpretation. All such decisions were performed or confirmed by the human authors, who reviewed and edited the final text and take full responsibility for the manuscript.

3. Results

3.1 Study selection

The database search identified 1,057 records: 568 from Scopus and 489 from Web of Science. Restricting the publication period to 2021–2026 removed 353 records, leaving 704. Article-type and English-language filtering left 484 records. Deduplication removed 178 records, yielding 306 unique records for title/abstract screening. Of these, 274 were excluded. Thirty-two full-text reports were assessed; four were excluded, comprising two studies that did not use deep learning, one destructive-only assessment, and one study without sufficient quantitative performance results. Twenty-eight studies were included in the qualitative synthesis and methodological appraisal. The study-selection process is summarized in Fig. 1.

3.2 Characteristics of included studies

The 28 studies covered finfish, shrimp/prawn, crayfish, crab, and oyster. Experimental designs ranged from conventional refrigerated storage trials to non-isothermal or dynamic-temperature studies, smart-packaging systems, sensor arrays, and multimodal sensing. RGB imaging frequently targeted whole products, eyes, or gills, while spectral studies examined muscle or fillet chemistry. Sample reporting was heterogeneous: some studies described hundreds of independent specimens, whereas others generated many images, views, patches, or augmented samples from comparatively few biological units. Table 1 summarizes the included studies by seafood target, sensing modality, freshness target, and DOI.

3.3 Sensing modalities and model architectures

Four broad technology families emerged. First, RGB/computer-vision systems used whole-product, eye, gill, or surface images with transfer-learning CNNs, YOLO, Mask R-CNN, modified ResNeXt, or lightweight backbones. Akgül et al. combined YOLOv5 with Inception-ResNet-v2 and reported 97.67% accuracy for anchovy and 96.0% for horse mackerel, whereas Gao et al. reported 98.958% accuracy with a compact DWG-YOLOv8 model (Akgül et al. 2023; Gao et al. 2024).

Second, spectroscopy and hyperspectral approaches predicted continuous biochemical or microbial targets. Han et al. reported 1D-CNN prediction of TVB-N with $R^2p=0.9397$ and TVC with $R^2p=0.9318$ in crayfish, while Jiang et al.

reported Transformer R^2 values of 0.9716 for TVB-N and 0.9721 for TVC in Atlantic salmon (Han et al. 2025b; Jiang et al. 2026).

Third, intelligent colorimetric and fluorescent sensors transformed volatile-spoilage chemistry into optical patterns for DL classification. Ma et al. reported up to 99.94% classification accuracy with Wide-ResNet-50 on a portable MOF colorimetric platform, and Zhang et al. reported 98.36% overall CNN accuracy using an ammonia-responsive fluorescent tag (Ma et al. 2021; Zhang et al. 2026).

Fourth, multimodal systems combined complementary signals. Guo et al. fused NIR and Raman spectra for salmon TBARS prediction, whereas Zou et al. used visual-olfactory cross-attention fusion and reported that multimodal fusion outperformed single-modality configurations (Guo et al. 2025; Zou et al. 2026).

3.4 Predictive performance and freshness validation

High classification accuracy was common but not uniform. Genç et al. reported 99.07% accuracy with CNN-LSTM for frozen-thawed shrimp, while Zheng et al. obtained 84.0% for 24-hour freshness classes and 72.0% for the more difficult 12-hour classes in large yellow croaker (Genç et al. 2025a; Zheng et al. 2025). The latter finding illustrates that apparent performance depends strongly on task granularity.

Regression studies provided a different form of evidence by predicting continuous freshness indicators. Guo et al. reported CNN R^2p=0.866 and RPD=4.136 for TBARS, while Moosavi-Nasab et al. reported LDNN R^2p=0.853, RMSEP=3.159, and RPD=3.001 for TVB-N (Guo et al. 2025; Moosavi-Nasab et al. 2021). These outcomes cannot be meaningfully ranked against classification accuracy because the target and metric structures differ.

Reference-standard strength also varied. Several studies directly used TVB-N, TVC/APC, K-value, TBARS, biogenic amines, or structured sensory/QIM measures, whereas others relied primarily on storage time or visual classes. This distinction is critical because a model can separate acquisition days accurately without necessarily measuring the biochemical construct of freshness.

3.5 Validation, methodological appraisal, and deployment readiness

Overall appraisal judgments were low concern in zero studies, some concern in 22, and high concern in six. The most recurrent weakness was limited external/generalizability validation. High-concern judgments were concentrated where weak freshness ground truth, potential non-independence of repeated images, limited leakage control, or no reported external validation materially affected the central claim.

Deployment evidence was more convincing in a smaller subset. He and Wei benchmarked their oyster system on Raspberry Pi 4B and Jetson TX2 NX hardware, while Yu et al. reported a mobile implementation with approximately 12 ms inference and 97.20% MobileNetV1 accuracy (He and Wei 2026; Yu et al. 2026). Hasan et al. demonstrated robotic sorting, and Zhang et al. demonstrated a smartphone application for fluorescent-tag interpretation (Hasan et al. 2024; Zhang et al. 2026).

Other studies primarily established proof-of-concept potential. Even where model size, latency, or portable sensing was favorable, industrial validation across independent batches, sites, devices, or supply-chain conditions remained uncommon.

4. Discussion

4.1 Principal findings

This review shows that DL-based non-destructive seafood freshness assessment has progressed beyond a single computer-vision paradigm. The field now includes RGB imaging, spectral sensing, intelligent optical sensors, electronic noses, and multimodal fusion. The strongest evidence does not come from a single architecture; rather, it arises when an appropriate sensing mechanism is paired with biologically meaningful ground truth, independent validation, and feasible deployment.

4.2 Sensing modality: information versus practicality

RGB and smartphone imaging provide the lowest barrier to acquisition and can support automated assessment of visible deterioration. Their main limitation is that external appearance may lag or incompletely represent biochemical spoilage. Spectral approaches can access molecular information and have achieved strong continuous prediction of TVB-N and TVC, but instrument complexity and calibration remain translational constraints (Han et al. 2025b; Jiang et al. 2026).

Colorimetric and fluorescent sensor arrays offer an intermediate pathway: they convert volatile chemistry into camera-readable patterns, potentially combining chemical sensitivity with portable optical readout. Smartphone-oriented sensor systems and intelligent packaging are therefore especially relevant to decentralized freshness monitoring (Pan et al. 2025; Zhang et al. 2026).

4.3 Architecture and predictive performance

Architecture should be interpreted jointly with task and modality. CNNs remain effective for image and sensor-pattern classification, YOLO variants add efficient localization/detection, CNN-LSTM models address sequential structure, and Transformers offer long-range spectral modeling. Lightweight architectures are particularly important when inference must occur on mobile or edge hardware (Gao et al. 2024; Yu et al. 2026).

Deep learning was not uniformly superior to conventional analytical models. In fluorescence and hyperspectral studies, RBFNN, PLSR, LS-SVM, or other chemometric approaches remained competitive under some conditions (Fan et al. 2024; Moosavi-Nasab et al. 2021). Accordingly, the scientific question should not be whether DL is universally best, but when its representation-learning capacity adds value relative to task complexity, dataset scale, and deployment constraints.

4.4 Reference-standard validity

The distinction between predicting freshness and predicting a proxy for freshness is central. Direct prediction of TVB-N, TVC/APC, K-value, TBARS, or biogenic amines provides a clearer mechanistic link than storage-day classification alone (Fan et al. 2025; Han et al. 2025a). Storage time remains operationally useful, but its relationship with deterioration depends on temperature, handling, species, packaging, and initial microbial status.

This issue also affects interpretation of very high accuracy. Controlled image classes can be separable because of consistent lighting, background, or temporal progression. Without specimen-level biochemical validation and independent testing, headline accuracy may overstate the degree to which a model measures transferable freshness.

4.5 Methodological robustness and generalizability

The most consistent weakness across the evidence base was the scarcity of external validation. Many studies used a single species, batch, laboratory, sensor, or storage regime. Cross-species transfer learning and robustness testing under varying lighting or angles are useful advances, but they remain less stringent than independent multi-site validation (Fan et al. 2025; Pan et al. 2025).

Biological independence is another important concern. Multiple angles, repeated storage observations, patches, and augmentation can generate large image counts from few independent specimens. Hamidy et al., for example, used extensive multi-angle imaging from ten fish, illustrating why image count should not be conflated with biological sample size (Hamidy et al. 2026). Future work should report specimen-level partitioning explicitly and prevent related observations from crossing training and test sets.

4.6 The performance–deployment gap

Laboratory accuracy is not equivalent to deployment readiness. A deployable system must tolerate device variability, illumination, temperature, packaging, sensor drift, sanitation requirements, throughput constraints, and limited compute. Actual edge and mobile benchmarking therefore provides qualitatively different evidence from a statement that a model could be deployed in the future (He and Wei 2026; Yu et al. 2026).

The evidence suggests a translational continuum: proof-of-concept, internally validated model, independent validation, portable prototype, and real-world or industrial validation. Most included studies occupy the first three stages; fewer reach demonstrated portable or embedded implementation, and robust industrial validation remains rare.

4.7 Implications and future research

For food-control practice, the critical transition is from model-centric benchmarking to decision-centric validation: systems should be evaluated against the inspection, release, shelf-life, or sorting decision they are intended to support, under the environmental and hardware conditions in which that decision will be made.

Future research should prioritize larger and biologically diverse datasets; specimen-level independent splitting; external and cross-device validation; standardized freshness reference measurements; multimodal fusion where modalities provide complementary information; and explicit reporting of model size, latency, hardware, and energy constraints. Explainability may also be useful where model outputs need to be audited by quality-control personnel.

For supply-chain use, the most promising systems are likely to be those co-designed around a specific operational decision. Smartphone vision may be appropriate for rapid retail or consumer screening, portable spectroscopy for higher-value analytical checks, smart packaging for continuous monitoring, and edge-enabled multisensor systems for processing-line automation. The sensing technology should therefore be selected from the operational context rather than from model accuracy alone.

4.8 Strengths and limitations of the review

This review integrates modality, architecture, freshness-reference validity, validation rigor, methodological appraisal, and deployment evidence in a single analytical framework. It also distinguishes biological sample independence from nominal image count and separates demonstrated deployment from proposed future use.

Several limitations should be considered. The search was restricted to Scopus and Web of Science, English-language peer-reviewed articles, and the 2021–2026 period. The search terminology may not capture every relevant

sensor or architecture synonym. Substantial heterogeneity in species, sample size, freshness definitions, metrics, and validation designs prevented meaningful meta-analysis. Some included studies did not report sufficient detail on biological independence, external validation, or deployment conditions, which was reflected in the appraisal judgments.

5. Conclusion

Deep learning has enabled a diverse set of non-destructive approaches for seafood freshness assessment, ranging from RGB computer vision and hyperspectral spectroscopy to intelligent colorimetric sensors, fluorescent tags, electronic noses, and multimodal fusion. Several studies report high classification accuracy or strong regression performance, but these values are not directly comparable across different targets and validation designs.

The central conclusion is that predictive accuracy alone is insufficient to establish technological maturity. Stronger evidence is obtained when a model is anchored to objective freshness indicators, evaluated with biologically independent data, tested across relevant conditions, and demonstrated on realistic hardware or sensing platforms. The current literature shows substantial proof-of-concept capability and emerging mobile, edge, packaging, and robotic implementations, yet external and industrial validation remain the principal gaps. Future progress therefore depends on integrating sensing quality, model efficiency, reference-standard validity, generalizability, and real-world deployment into a single evaluation pathway.

Declarations

Author Contributions: Muhammad Munsarif: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Validation, Project administration, Writing – original draft; Richmond Ampah-Mensah: Conceptualization, Methodology, Investigation, Data curation, Validation, Visualization, Writing – review & editing.

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

Data Availability: No new primary datasets were generated or analyzed for this review. Study-level evidence was derived from the published articles included in the systematic review.

References

1. Akgül I, Kaya V, Zencir Tanır Ö (2023) A novel hybrid system for automatic detection of fish quality from eye and gill color characteristics using transfer learning technique. PLOS ONE. https://doi.org/10.1371/journal.pone.0284804
2. Fan Y et al (2024) Deep learning models with optimized fluorescence spectroscopy to advance freshness of rainbow trout predicting under nonisothermal storage conditions. Food Chemistry. https://doi.org/10.1016/j.foodchem.2024.139774
3. Fan Y et al (2025) Integrating EEM fluorescence spectroscopy with machine and deep learning models for multi-output prediction of fish freshness. Applied Food Research. https://doi.org/10.1016/j.afres.2025.101374
4. Gao S et al (2024) Intelligent classification and identification method for *Conger myriaster* freshness based on DWG-YOLOv8 network model. Food Bioengineering 3(3):269–279. https://doi.org/10.1002/fbe2.12097

5. Genç İY et al (2025b) Quality prediction of seabream *Sparus aurata* by deep learning algorithms and explainable artificial intelligence. Food Chemistry. https://doi.org/10.1016/j.foodchem.2025.143150
6. Genç İY, Gürfidan R, Açikgözoğlu E (2025a) Quality determination of frozen-thawed shrimp using machine learning algorithms powered by explainable artificial intelligence. Food Analytical Methods 18:935–945. https://doi.org/10.1007/s12161-025-02768-0
7. Guo M et al (2025) Deep learning algorithm-assisted non-destructive detection of TBARS values of salmon flesh using multi-modal molecular spectra fusion. Food Chemistry. https://doi.org/10.1016/j.foodchem.2025.145649
8. Hamidy SM et al (2026) Deep learning-based prediction of fish freshness and purchasability using multi-angle image data. Foods. https://doi.org/10.3390/foods15010068
9. Han Q et al (2025a) Nondestructive detection of biogenic amines in muscle of Chinese mitten crab: A basis for quality assessment using infrared spectroscopy and deep learning. Food Chemistry. https://doi.org/10.1016/j.foodchem.2025.146354
10. Han Q et al (2025b) Non-destructive detection of freshness in crayfish (*Procambarus clarkii*) based on near-infrared spectroscopy combined with deep learning. Food Control. https://doi.org/10.1016/j.foodcont.2024.110858
11. Hao D et al (2025) A non-destructive deep learning-based method for shrimp freshness assessment in food processing. Processes. https://doi.org/10.3390/pr13092895
12. Hasan M et al (2024) Framework for fish freshness detection and rotten fish removal in Bangladesh using Mask R-CNN method with robotic arm and fisheye analysis. Journal of Agriculture and Food Research. https://doi.org/10.1016/j.jafr.2024.101139
13. He A, Wei G (2026) A novel oyster freshness prediction model based on hierarchical feature fusion with XGBoost and CNN-LSTM for edge deployment. IEEE Sensors Journal. https://doi.org/10.1109/JSEN.2025.3648226
14. Hou M et al (2025) Innovations in seafood freshness quality: Non-destructive detection of freshness in *Litopenaeus vannamei* using the YOLO-shrimp model. Food Chemistry 463(Pt 1):141192. https://doi.org/10.1016/j.foodchem.2024.141192
15. Jiang C et al (2024) Deep-learning-assisted portable chemo-responsive dye-based odor sensor array for rapid sensing of shrimp and fish freshness. Chemical Engineering Journal. https://doi.org/10.1016/j.cej.2024.149283
16. Jiang Z et al (2026) Non-destructive freshness assessment of Atlantic salmon via hyperspectral imaging and an SPA-enhanced Transformer framework. Foods. https://doi.org/10.3390/foods15040725
17. Li J et al (2026) A smart colorimetric film coupling deep eutectic solvents and deep learning for real-time salmon freshness monitoring. Food Chemistry 512:148904. https://doi.org/10.1016/j.foodchem.2026.148904
18. Ma P et al (2021) Integrated portable shrimp-freshness prediction platform based on ice-templated metal-organic framework colorimetric combinatorics and deep convolutional neural networks. ACS Sustainable Chemistry & Engineering. https://doi.org/10.1021/acssuschemeng.1c04704
19. Moosavi-Nasab M et al (2021) Evaluation of the total volatile basic nitrogen (TVB-N) content in fish fillets using hyperspectral imaging coupled with deep learning neural network and meta-analysis. Scientific Reports. https://doi.org/10.1038/s41598-021-84659-y
20. Murali S et al (2025) Indian mackerel freshness evaluation using image processing and convolutional neural networks approach. Journal of Food Measurement and Characterization 19(9):6968–6978.

https://doi.org/10.1007/s11694-025-03429-w
21. Nandi B et al (2026) Spectra-MobileNet: A frequency-aware ordinal framework for fine-grained prawn freshness assessment. WSEAS Transactions on Signal Processing 22:193–205. https://doi.org/10.37394/232014.2026.22.17
22. Page MJ et al (2021) The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ 372:n71. https://doi.org/10.1136/bmj.n71
23. Pan Y et al (2025) A smartphone-based non-destructive multimodal deep learning approach using pH-sensitive pitaya peel films for real-time fish freshness detection. Foods. https://doi.org/10.3390/foods14101805
24. Yu Q et al (2026) A dual-response colorimetric sensor array integrated with deep learning for mobile intelligent freshness assessment of aquatic products. Engineering Applications of Artificial Intelligence 177(Pt 2):114952. https://doi.org/10.1016/j.engappai.2026.114952
25. Zhang L et al (2024) Deep learning used with a colorimetric sensor array to detect indole for nondestructive monitoring of shrimp freshness. ACS Applied Materials & Interfaces. https://doi.org/10.1021/acsami.4c04223
26. Zhang W, Tian Y, Chen Y, Wei M (2026) CNN-integrated ratiometric fluorescent tag for rapid monitoring of seafood freshness at refrigerated temperatures. Current Research in Food Science. https://doi.org/10.1016/j.crfs.2026.101560
27. Zhang Y, Wei C, Zhong Y, Wang H, Luo H, Weng Z (2022) Deep learning detection of shrimp freshness via smartphone pictures. Journal of Food Measurement and Characterization 16(5):3868–3876. https://doi.org/10.1007/s11694-022-01473-4
28. Zheng Y, Zhang Q, Wang X, Guo Q (2025) Classifying the freshness of large yellow croaker at 12- and 24-hour intervals using computer vision technique and convolutional neural network. Smart Agricultural Technology 10:100767. https://doi.org/10.1016/j.atech.2025.100767
29. Zou et al (2026) ST-CAFNet: A visual-olfactory multimodal cross-attention fusion network for freshness prediction of cold plasma-treated salmon. Computers and Electronics in Agriculture. https://doi.org/10.1016/j.compag.2026.112222

Table

Table 1 Characteristics of the included studies

Study	Seafood/target	Modality	Freshness target	DOI
Akgül et al. (2023)	Anchovy; horse mackerel	RGB eye + gill	Visual/storage class	10.1371/journal.pone.0284804
Fan et al. (2024)	Rainbow trout	EEM fluorescence	K-value; TVC; VBN	10.1016/j.foodchem.2024.139774
Fan et al. (2025)	Bighead carp; trout	EEM fluorescence	TVB-N; K-value; TVC	10.1016/j.afres.2025.101374
Gao et al. (2024)	Conger myriaster	RGB	Freshness class	10.1002/fbe2.12097
Genç et al. (2025a)	Shrimp	RGB	Visual/storage class	10.1007/s12161-025-02768-0
Genç et al. (2025b)	Sea bream	RGB eye + gill	Storage freshness class	10.1016/j.foodchem.2025.143150
Guo et al. (2025)	Atlantic salmon	NIR + Raman	TBARS/MDA	10.1016/j.foodchem.2025.145649
Hamidy et al. (2026)	Sea bass	Multi-angle RGB	QIM/color/purchasability	10.3390/foods15010068
Han et al. (2025a)	Chinese mitten crab	NIR	Biogenic amines	10.1016/j.foodchem.2025.146354
Han et al. (2025b)	Crayfish	NIR	Whiteness; TVB-N; TVC	10.1016/j.foodcont.2024.110858
Hao et al. (2025)	Marine shrimp	RGB	TVB-N thresholds	10.3390/pr13092895
Hasan et al. (2024)	Fish	Fish-eye RGB	Visual fresh/rotten	10.1016/j.jafr.2024.101139
He & Wei (2026)	Oyster	Electronic nose	Sensory; APC; TVB-N	10.1109/JSEN.2025.3648226
Hou et al. (2025)	Pacific white shrimp	RGB	Sensory; TVB-N; TVC	10.1016/j.foodchem.2024.141192
Jiang et al. (2024)	Shrimp and fish	Colorimetric odor array	Volatile spoilage	10.1016/j.cej.2024.149283
Jiang et al. (2026)	Atlantic salmon	HSI	TVB-N; TVC	10.3390/foods15040725

Li et al. (2026)	Salmon	Colorimetric film	TVB-N; TVC; amines	10.1016/j.foodchem.2026.148904
Ma et al. (2021)	Shrimp	MOF colorimetric array	Volatile amines	10.1021/acssuschemeng.1c04704
Moosavi-Nasab et al. (2021)	Rainbow trout fillet	HSI	TVB-N	10.1038/s41598-021-84659-y
Murali et al. (2025)	Indian mackerel	RGB	Conventional indicators	10.1007/s11694-025-03429-w
Nandi et al. (2026)	Prawn	Frequency-aware RGB	Hedonic grades	10.37394/232014.2026.22.17
Pan et al. (2025)	Fish	Smartphone + pH film	pH; TVB-N; TVC	10.3390/foods14101805
Zhang et al. (2024)	Shrimp	Colorimetric strip	Indole/spoilage	10.1021/acsami.4c04223
Zhang et al. (2022)	Shrimp	Smartphone RGB	Freshness classes	10.1007/s11694-022-01473-4
Zheng et al. (2025)	Large yellow croaker	Smartphone RGB	Storage time	10.1016/j.atech.2025.100767
Zou et al. (2026)	Salmon	Visual + olfactory	Freshness indicators	10.1016/j.compag.2026.112222
Yu et al. (2026)	Salmon; shrimp	Dual-response array	TVB-N/indole responses	10.1016/j.engappai.2026.114952
Zhang et al. (2026)	Shrimp/seafood	Fluorescent tag	Ammonia/volatile amines	10.1016/j.crfs.2026.101560

Figures

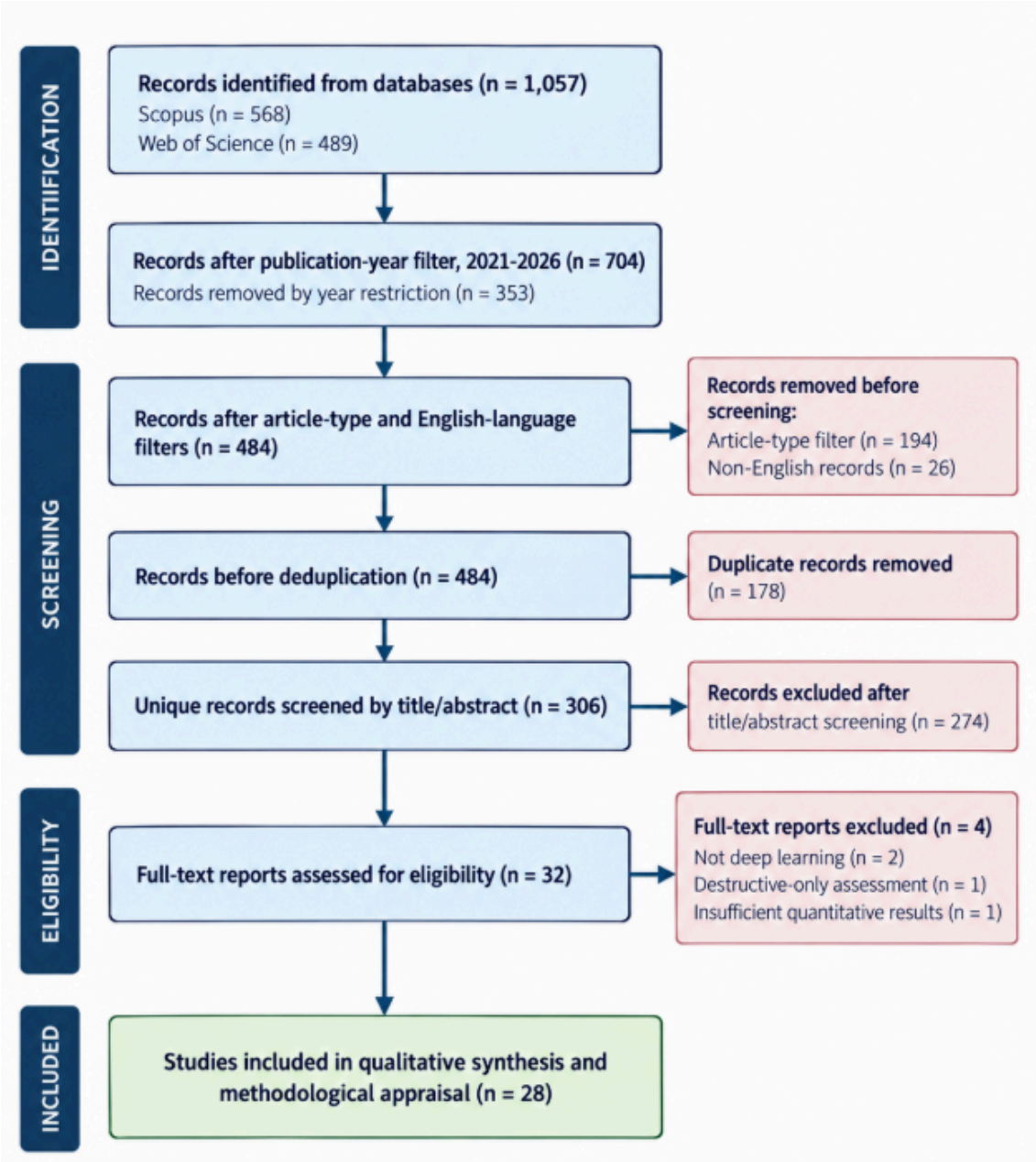


Figure 1

PRISMA 2020 flow diagram for study selection