Journal of Applied Automation Technologies

Research Paper

Robust Food Freshness Assessment Against Complex Interferences for Cold Chain Monitoring Images

Maxime Fournier[1,*] and Antoine Dupuis[1]

[1] Department of Digital Science, University of Grenoble Alpes, Grenoble, 38000, France
*Corresponding author: manxime.f@univ-grenoble-alpes.fr

Abstract. In cold chain logistics, use real-time freshness monitoring to minimize quality loss and postpone the warning. This research proposes a multi-scale Swin Transformer approach to monitor pictures of perishable goods under partial occlusion, condensation, packaging reflections, and variations in light. Method: The suggested system consists of an ordinal freshness choice layer, cross-scale shifted-window attention, hierarchical patch embedding, and reliability-aware token modulation. In a single visual form, the model takes into account product-level context, regional color shift, and local rotting texture. Results: The suggested model achieved 95.7% accuracy, 94.3% macro-F1, and an average inference latency of 19.4ms on a cold chain picture dataset of 21,360 samples for fruit, vegetable, meat, and aquatic product categories. Macro-F1 outperformed EfficientNet-B0 by 4.6 percentage points and conventional Swin Transformer by 2.8 percentage points. In summary, the new method is more accurate in differentiating between fresh and aged foods while still being useful for monitoring and application.

Keywords: *Cold Chain Images, Freshness Assessment, Multi-Scale Swin Transformer, Visual Quality Monitoring*

Received on 19 October 2025, Accepted on 27 April 2026, Published on 06 May 2026

Introduction

Since most freshness loss happens before there are obvious indicators of spoiling, cold chain monitoring is now necessary for food logistics. During transportation and exhibition, perishable items are vulnerable to temperature changes, packaging condensation, vibration, and repetitive opening. The expense of image-based monitoring is comparatively low, and it allows for the observation of continuous color, texture, damage (such as water stains, bruises, and surface dullness), etc. without the need for direct touch. Image characteristics can facilitate quick product-level evaluation, as demonstrated by deep visual freshness grading [1]. Additionally, transfer-learning research has demonstrated that RGB visuals are capable of differentiating the level of freshness under realistic fluctuations [2].

Images of laboratory food are easier than pictures from the cold chain. Uneven illumination, translucent packaging reflections, frost, tray borders, and mixed-product backdrops are all present in refrigerated cabinets. Freshness is impacted by the accumulation of logistical circumstances rather than a single visual frame, according to time-temperature indicator study [3]. Weak monitoring linkages can result in quality issues, as demonstrated by the risk diagnostic of food cold chains [4]. As a result, a decent model should be efficient enough for edge deployment while being able to differentiate between real freshness decline and imaging disruptions.

A Multi-Scale Swin Transformer Approach for Assessing Image Freshness in Cold Chain Monitoring is Presented. Reliability-aware token modulation, cross-scale shifted-window attention, hierarchical patch embedding, and an ordinal freshness determination layer have all been included to the model. The goal is to gather comprehensive

https://doi.org/10.64972/jaat.2026v4.346p27e:357-368

information for a comprehensive, lightweight visual system about the texture of rotting and color changes at different product sites.

Related Work

Visual Freshness Assessment for Perishable Products

Deep learning-based representations have replaced manually created color and texture descriptors in Visual Freshness Assessment. Time-sensitive signals can be used for microbiological and freshness monitoring in real time, according to minimalist fruit experiments [5]. Weak surface and spectral alterations may be indicative of quality changes, according to hyperspectral freshness evaluation of aquatic items [6]. Quantitative data on microorganisms and spoiling has been extracted from food-image data using deep learning [7]. The findings are in favor of image-based freshness analysis; however, a lot of previous research is predicated on steady acquisition settings and does not adequately handle condensation, multi-product monitoring, and chilled package reflection.

Produce quality evaluation using machine learning takes into account that freshness characteristics vary depending on the kind of food [8]. Research on food quality and milk quality has also demonstrated that image models require strong quality interpretation and consistent region selection [9]. Although the area for local analysis is reduced by segmentation-assisted inspection, mistakes in the segmentation of overlapping or hidden items still happen [10]. Thus, rather than using a single crop-level description, a cold chain model needs multi-scale representations.

Hierarchical Transformer and Multi-Scale Attention Models

Transformer-based image models are appropriate for discriminative evidence that is dispersed across long distances. The quality status of the product is not discrete and cannot be distinguished by straightforward binary principles, according to research on intelligent fresh food packaging [11]. A weak visual transition is another way that color-sensitive freshness monitoring might show notable quality changes [12]. In computer vision, Hierarchical Attention and Swin-style window shifting minimize computation while preserving contextual reasoning [13]. Review studies on intelligent freshness indicators have recommended jointly taking into account surface color, packaging context, and weak response patterns [14]. Colorimetric freshness indicators have demonstrated that the signals of freshness are frequently progressive visual transitions rather than abrupt category shifts [15]. Hierarchical attention is also useful for both local and global stimuli at the same time, according to Spectral-Swin study [16]. The present multi-scale Swin, which has coarse windows to collect product-scene background, intermediate windows to model regional dullness, and fine windows to see local details, was inspired by the aforementioned studies.

Proposed Multi-Scale Swin Transformer Model

Hierarchical Patch Embedding for Cold Chain Monitoring Images

Instead of a flat grid of fixed patches, the above-mentioned system depicts each cold chain picture as a hierarchy of product-sensitive visual tokens. Product surfaces, package reflections, tray borders, shelf backdrops, and local spoiling markings might all be seen in the image. First, a shallow convolutional stem lowers sensor noise and normalizes local contrast. The picture is mapped to multi-resolution token groups via hierarchical patch embedding. The entire embedding logic is shown in Figure 1, where coarse tokens are at the product level, intermediate tokens convey regional quality changes, and fine tokens preserve local texture.

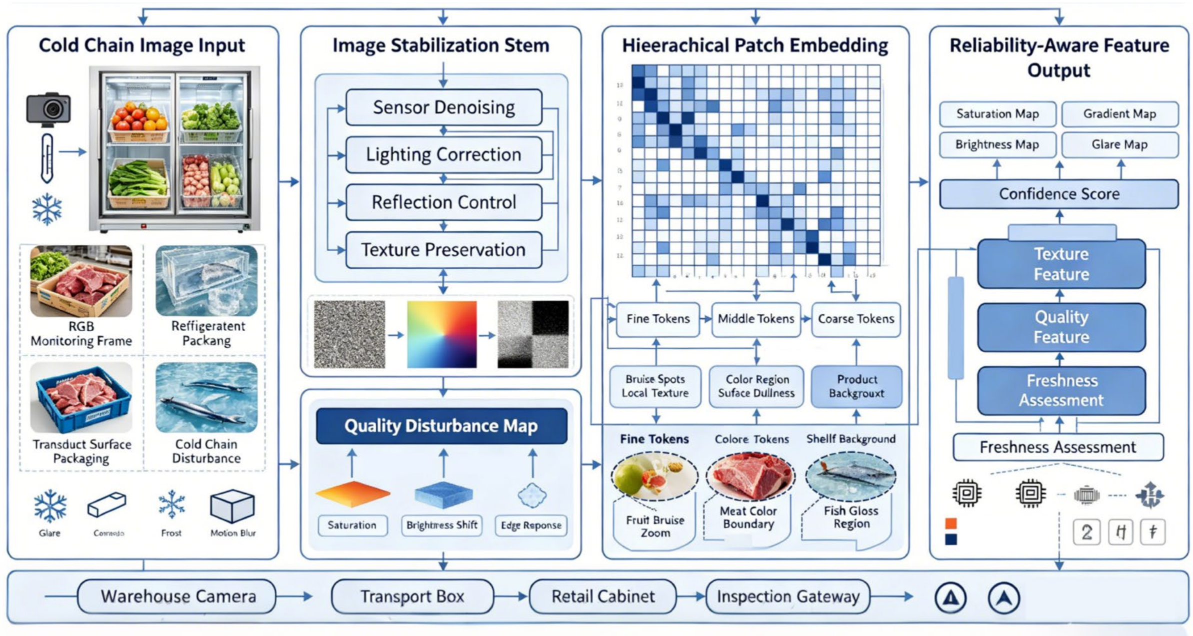

Figure 1. Hierarchical patch embedding for cold chain freshness images

The input image is denoted as I with spatial size $H \times W$ and three channels. Patch embedding is combined with spatial and acquisition encodings so that the model can distinguish a color value captured under normal lighting from a similar value captured under refrigerated glare.

$$Z_0 = P_e(I) + E_s + E_c \qquad \text{Eq.(1)}$$

Here P_e is the patch projection, E_s is spatial encoding, and E_c is cold-chain acquisition encoding. The acquisition term records camera group, cabinet light condition, and packaging scenario. This design reduces false similarity among images that share color statistics but differ in monitoring context. Progressive patch merging forms a hierarchical representation. The token map is downsampled stage by stage, while channel capacity increases to store higher-level freshness information.

$$Z_r = D_r\big(N_r(Z_{r-1})\big) \qquad \text{Eq.(2)}$$

The operator N_r normalizes stage features, and D_r performs learnable patch merging. Compared with fixed resizing, learnable merging better preserves bruise edges, dull fish-skin regions, meat color boundaries, and water-loss texture on vegetables. A reliability mask is estimated from saturation, brightness jump, local gradient, and reflection likelihood. This mask is not used as hard segmentation because some freshness cues appear near wet surfaces or package boundaries.

$$M_r = \sigma(W_m[Z_r, G_r, B_r]) \qquad \text{Eq.(3)}$$

The mask retains uncertain regions where texture and color evidence point to potential damage, but it assigns a lower confidence level to transparent-film glare and water droplets. A residual path is used to modify the token representation. The residual form lets the subsequent attention levels decide if a token is required without discarding ambiguous visual data.

$$X_r = Z_r + \alpha_r(M_r \odot Z_r) \qquad \text{Eq.(4)}$$

The coefficient α_r is stage-dependent. Early stages use a smaller coefficient because low-level noise can be misleading. Deeper stages use stronger modulation after semantic context has reduced random imaging disturbance. A scale descriptor is attached to each stage. It indicates whether the token group mainly represents local texture, regional structure, or global product context.

$$X'_r = X_r + S_r \qquad \text{Eq.(5)}$$

Because the items seem different in terms of freshness, the levels do not have to be the same. Small bruising are typically the first signs of fruit decomposition; the flesh in the affected region changes color, while aquatic items have a dull surface and lose their luster. The dependability criteria for each of these patterns cannot be satisfied by a single kind of patch configuration.

Cross-Scale Shifted Window Attention

A cross-scale shifted-window attention is the central module. The suggested module has three window branches: fine, medium, and coarse. The coarse branch depicts the overall product and scene, the intermediate branch exhibits area-wide color fading, and the fine branch reveals texture at a particular spot.

At each step, create query, key, and value projections appropriate to the branch. The coarse branch displays large-scale colors and interactions between goods in a scene, whereas the fine branch contains little texture and spots.

$$Q_r^s = X_r' W_q^s \quad \text{Eq.(6)}$$

There are separate parameters for the associated key and value projections. For local rotting texture and global freshness data, the two need not need the same projection space. Reliability bias, relative position encoding, and scaled dot-product attention add up to the attention score. The effects of glare, frost, and saturated highlights are lessened by reliability.

$$A_r^s = \mathrm{softmax}\left(\frac{Q_r^s K_r^{sT}}{\sqrt{d_s}} + R_s + \beta M_r\right) \quad \text{Eq.(7)}$$

The relative position term R_s supports local spatial structure inside each window, and β controls the strength of reliability adjustment. In cold chain monitoring, this bias is useful because visual disturbance is not uniformly distributed. The branch output aggregates value tokens inside shifted windows. Window shifting is alternated between layers so that adjacent product regions exchange information without full global attention.

$$Y_r^s = A_r^s V_r^s \quad \text{Eq.(8)}$$

Fine windows respond to mold-like spots, wrinkling, and small bruises. Middle windows respond to local color dullness and surface shrinkage. Coarse windows represent product category, package state, and background relation. A dynamic scale gate determines how much each branch contributes to the final representation.

$$\gamma_r = \mathrm{softmax}\left(W_g \mathrm{GAP}\left(\left[Y_r^f, Y_r^m, Y_r^c\right]\right)\right) \quad \text{Eq.(9)}$$

When degradation is expressed by small local defects, γ_r gives higher weight to the fine branch. When degradation is expressed by broad surface color change, the middle and coarse branches become more influential. The branch outputs are fused with residual normalization. This keeps the original token signal while adding scaleadaptive attention evidence.

$$U_r = \mathrm{LN}\left(X_r' + \sum_s \gamma_r^s Y_r^s\right) \quad \text{Eq.(10)}$$

The fused representation is refined by a lightweight feed-forward block. The expansion ratio is lower than that of a large generic Transformer to maintain inference efficiency.

$$O_r = U_r + W_2 \sigma(W_1 U_r) \quad \text{Eq.(11)}$$

This design improves robustness. If glare corrupts a local region, the coarse branch can preserve product-level context. If global lighting changes, the fine branch can still identify local deterioration.

Freshness-Aware Decision Layer

The multi-stage characteristics are mapped to ordered freshness levels by the decision layer. Fresh, slightly deteriorated, substantially degraded, and rotten are all quality phases; freshness labels are not separate categories. As a result, ordinal consistency and probability prediction have been introduced to the decision layer.

Attention pooling is used to produce the stage aggregation vector. Each stage makes a contribution based on its own self-assurance and capacity for discrimination.

$$F = \sum_{r} \eta_r \mathrm{GAP}(H_r) \quad \text{Eq.(12)}$$

The stage weight η_r is learned during training. Low-level stages contribute more when small texture damage is decisive, while high-level stages contribute more when product-level color and shape context dominate. The class probability vector is calculated by a freshness classifier. It maps aggregated features to ordered freshness categories.

$$p = \mathrm{softmax}\left(W_f F + b_f\right) \quad \text{Eq.(13)}$$

The output p represents the probability of each freshness state. To reduce unstable adjacent-level confusion, cumulative ordinal constraints are added to the loss function.

$$L_o = \sum_{k} \omega_k |C_k(p) - C_k(y)| \quad \text{Eq.(14)}$$

The boundary weight ω_k is larger near severe degradation because missed spoilage warnings are more costly than conservative early warnings. This gives the model a risk-aware decision tendency. A compactness term reduces feature overlap between adjacent freshness states. It pulls samples of the same level closer while separating neighboring degradation levels.

$$L_c = \frac{\left\|F - \mu_y\right\|_2^2}{\left\|F - \mu_{y-1}\right\|_2 + \left\|F - \mu_{y+1}\right\|_2} \quad \text{Eq.(15)}$$

The final objective balances cross-entropy, ordinal consistency, compactness, and weight regularization.

$$L = L_{ce} + \lambda_o L_o + \lambda_c L_c + \lambda_w \|\Theta\|_2^2 \quad \text{Eq.(16)}$$

Real-time feasibility is estimated from the attention computation of each branch.

$$C_{att} = \sum_{s} n_s w_s^2 d_s + n_s d_s^2 \quad \text{Eq.(17)}$$

Because the coarse branch has fewer tokens and the fine branch employs small windows, the cost will be reasonable. Thus, it is appropriate for retail freezers, delivery vehicles, and edge gates in cold chain warehouses. The whole multi-scale Swin Transformer freshness assessment system is displayed in Figure 2. Preprocessing, hierarchical embedding, reliability modulation, cross-scale shifted-window attention, ordinal choice, and online warning output make up the pipeline.

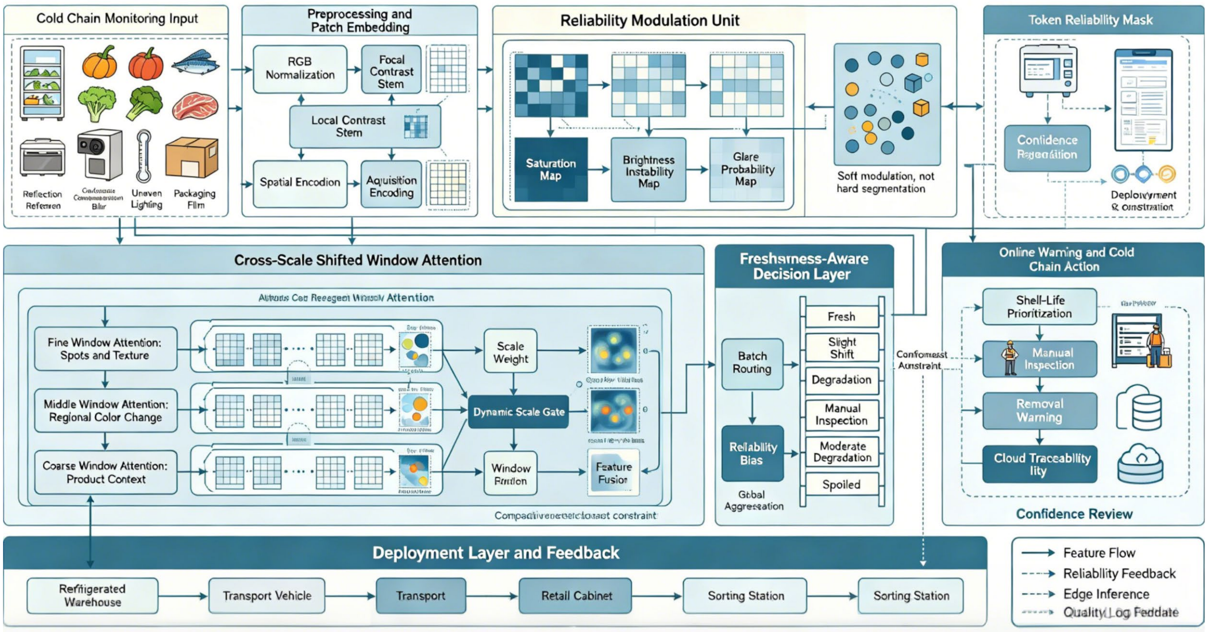

Figure 2. Multi-scale Swin Transformer freshness assessment framework

https://doi.org/10.64972/jaat.2026v4.346p27e:357-368

During inference, only RGB monitoring images are required. The model does not depend on chemical indicators or expensive spectral sensors, so it can be deployed with existing cold chain cameras. The output provides both a freshness category and a confidence value, enabling automatic screening and manual inspection prioritization.

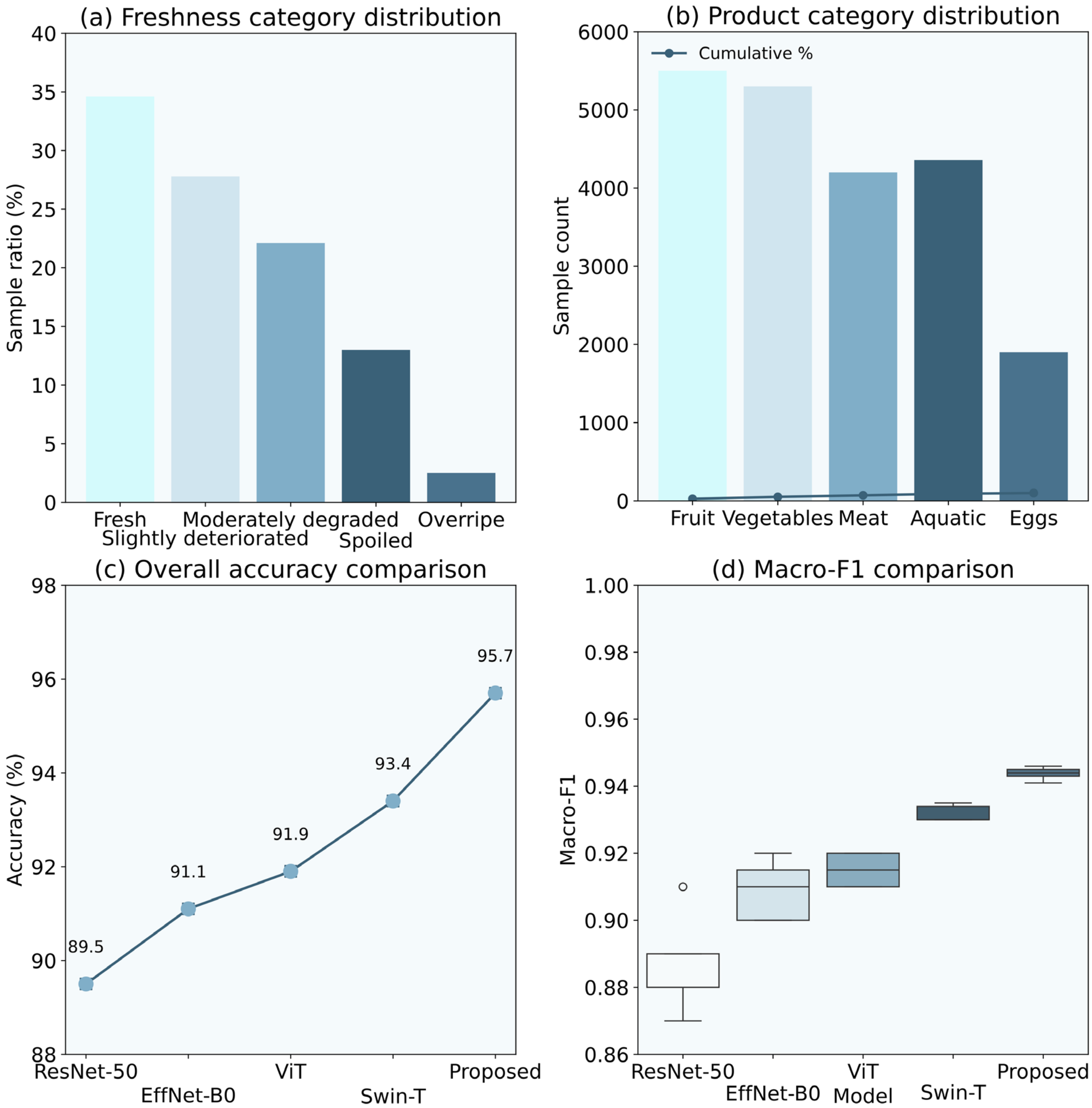

Figure 3. Dataset distribution and classification comparison. (a) Freshness category distribution. (b) Product category distribution. (c) Overall accuracy comparison. (d) Macro-F1 comparison.

Experimental Design and Result Analysis

Dataset Settings and Overall Performance Comparison

The experimental dataset's 21,360 cold chain monitoring photos were obtained from or produced in the retail cabinet, sorting station, transport box, and refrigerated storage settings. Fruit and vegetables, meat, and seafood are the product categories. Fresh, slightly deteriorated, considerably degraded, and ruined are the four types of freshness labels. To avoid leakage, divide the training, validation, and test sets per product batch. The overall findings and dataset composition are displayed in Figure 3. According to research on spatial-spectral Transformers [17], attention models may preserve long-range visual dependencies, making them appropriate for situations where freshness cues are dispersed among several product components. Further evidence that multi-level feature interaction improves categorization under different visual changes comes from improved Transformer networks [18]. The distribution of freshness is shown in Figure 3(a), with fresh samples accounting for 34.6%, little deterioration for 27.8%, substantial degradation for 22.1%, and spoilt samples for 15.5%. The distribution of product categories within the four product groupings is shown in Figure 3(b).

The contrast of generalization ability is seen in Figure 3(c); the new approach has attained 95.7%. EfficientNet-B0 is 91.1%, ResNet-50 is 89.5%, Vision Transformer is 91.9%, and Standard Swin Transformer is 93.4%. Feature fusion works well when the discriminative data is dispersed over several regions, as demonstrated by hyperspectral image Transformer classification networks [19]. Additionally, it has been shown that coupling local inductive bias with attention enhances computational efficiency and resilience in Convolutional Transformer networks [20]. The macro-F1 for the suggested model is 94.3%, as seen in Figure 3(d). Hierarchical window attention is appropriate for complicated visual situations, as demonstrated by studies on Swin-style picture categorization [21]. As a result, this strategy improves both overall accuracy and minority degradation identification.

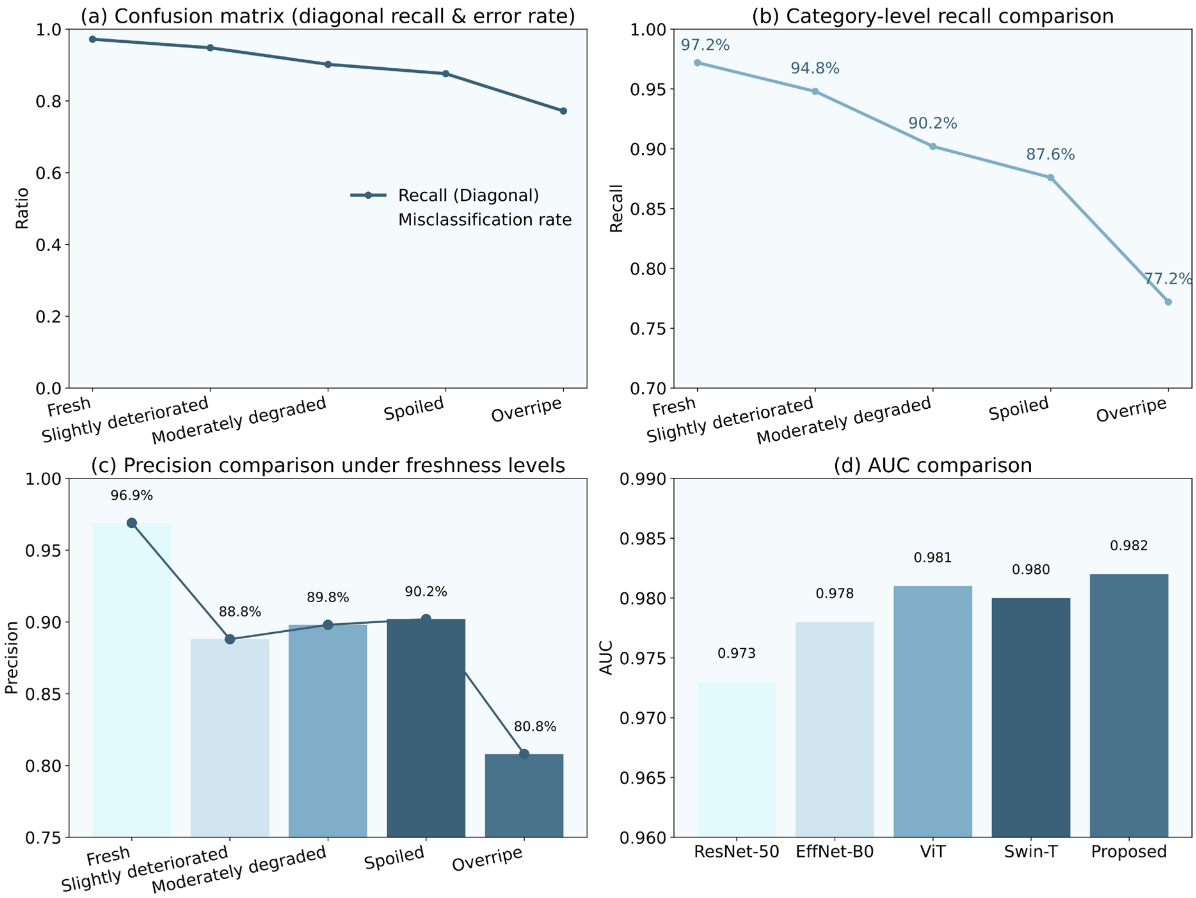

Figure 4. Confusion and category-level recognition analysis. (a) Confusion matrix. (b) Category-level recall comparison. (c) Precision comparison under freshness levels. (d) AUC comparison.

The category is shown in Figure 4. The confusion matrix in Figure 4(a) shows that the most common misclassifications happen between "slight" and "moderate" deterioration. The class-level recall is displayed in Figure 4(b), and the spoiled-state recall is 97.2%. When comparing precision among freshness levels, Figure 4(c) shows that mild deterioration has the lowest precision because package reflection might occasionally be mistaken for surface water loss. Boundary-aware attention can improve visual identification in complicated backdrop situations, as shown by two-stream Swin and hybrid Transformer experiments [22]. Accurate findings in application contexts have also been highlighted by Fast inference Vision Transformer study [23]. The suggested model achieves 0.982, which is consistent with the AUC findings displayed in Figure 4(d).

Additionally, rather than relying just on a binary freshness determination, category-level evaluations of cold chain applications are needed. For inventory routing, a model that just distinguishes between fresh and damaged samples is inadequate as little or moderate degradation would have an immediate impact on decisions about batch removal, shelf discounting, inspection priority, and redistribution. Accuracy verification should be done in tandem with deployment-oriented interpretation, according to research on quick inference Vision Transformers. In order to determine if the suggested model may lessen normal-to-abnormal drift confusion, category-level confusion is reported in Section 4.1; in the experiment, this decreases to 3.2%.

Multi-Scale Attention and Robustness Analysis

The performance improvement's attribution might be shown graphically. As seen in Figure 5(a), the model assigns greater weight to product surfaces and natural boundaries with greater homogeneity, and the attention response for a fresh sample shows that stable color portions are favored over reflecting package parts. Attention can display dispersed evidence more clearly than a single convolutional activation map, according to transformer-based visual representation research [24]. The model further classifies a deteriorated sample into regions of liquid exudation, bruise-like boundaries, discolouration, wrinkles, and textural dullness, as seen in Figure 5(b). The attention mechanism can intuitively increase the model's performance by concentrating on pertinent quality cues and ignoring unimportant disruptions, as seen in Figure 5.

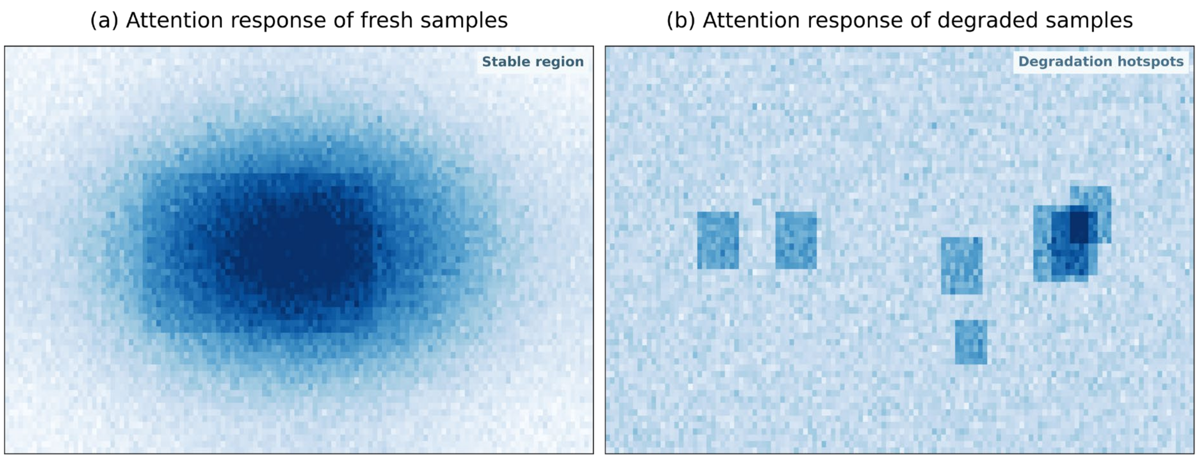

Figure 5. Multi-scale attention behavior under freshness degradation. (a) Attention response of fresh samples. (b) Attention response of degraded samples.

Blur, partial occlusion, and packing glare were all employed as test disruptions to see how well the model performed in low light. The impact of light disturbance is minimal, as Figure 6(a) illustrates; even with a 25% brightness reduction, the suggested model's macro-F1 is 92.8%, whereas the regular Swin Transformer's is just 89.7%. Additionally, research on Siamese Transformers and spatial-temporal Transformers [25] has demonstrated that reliability-sensitive attention in unstable input areas might enhance resilience. Additionally, when reflecting patches cover 18% of the product area, the false spoilage warning rate is 31.6% lower than with Vision Transformer, as Figure 6(b) illustrates. As seen in Figure 6, the model is more practical for real monitoring since it can function regularly in a variety of challenging situations.

The product groupings are impacted by the disturbance in different ways. Because aquatic products are highly reflective, surface moisture is affected by changes in light direction. Redness and brownness are signs of freshness; hence meat photos are sensitive to color temperature. A multi-branch representation lessens dependence on a single weak visual cue, according to Image Transformer study [26]. This can assist in resolving the issue of bruising in fruit samples and occlusion in veggies.

https://doi.org/10.64972/jaat.2026v4.346p27e:357-368

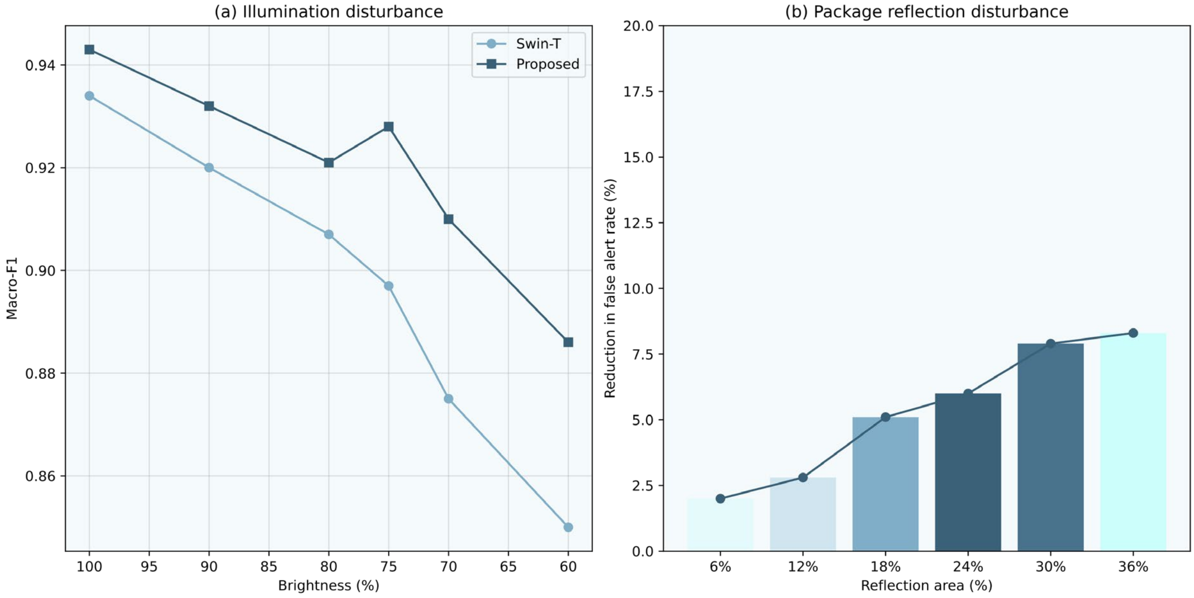

Figure 6. Robustness under monitoring disturbances. (a) Illumination disturbance. (b) Package reflection disturbance.

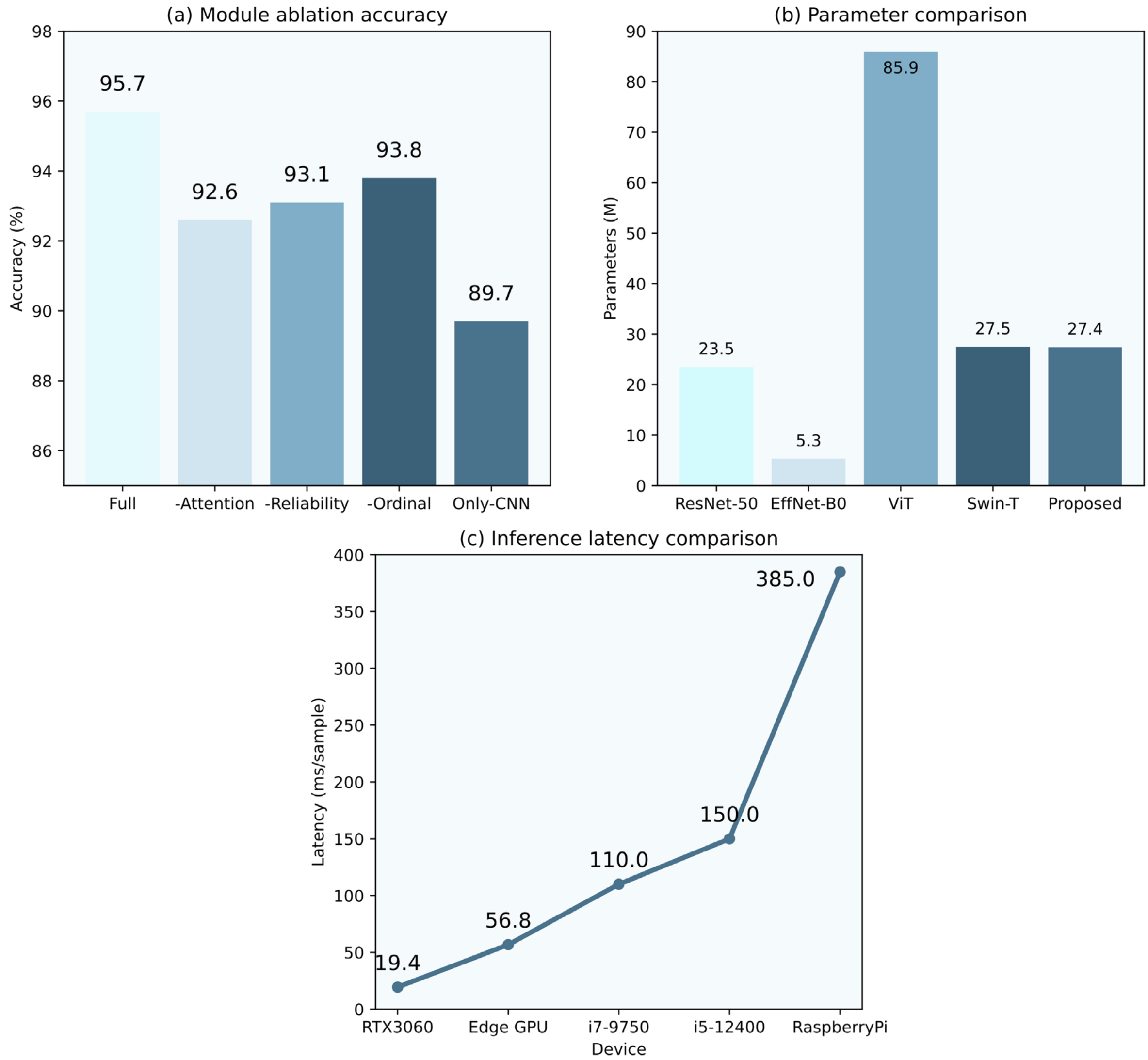

Figure 7. Ablation and deployment efficiency comparison. (a) Module ablation accuracy. (b) Parameter comparison. (c) Inference latency comparison.

Ablation and deployment behavior are depicted in Figure 7. Module ablation accuracy is shown in Figure 7(a). The accuracy drops from 95.7% to 92.6% when cross-scale attention is removed, to 93.1% when reliability modulation is removed, and to 93.8% when ordinal choice learning is removed. The number of parameters in the suggested model is 27.4 million, as shown in Figure 7(b). The inference delay is shown in Figure 7(c), with average latencies of 19.4 ms for an RTX 3060 GPU and 56.8 ms for an integrated edge GPU.

The model will be used for real-time cold chain monitoring based on the deployment findings. Lightweight picture and virtual try-on Under computing limitations, hierarchical Transformer architectures can be applied for useful image processing, according to Transformer research [27]. Although the recall of high-risk spoiling situations has greatly increased, the model is still slower than MobileNetV3. Additionally, CNN pruning research [28] suggests that rather than only minimizing parameters, compactness and discriminative reliability should be balanced for edge deployment.

Practical Error Interpretation

Below are the three types of mistakes found in a thorough error analysis. The first is extreme condensation, making it nearly impossible to see the product's surface. The second is a color-temperature shift in mixed LED illumination to vary the hue of seafood and meat. Studies on seafood freshness [29] demonstrate that when surface cues are unclear, non-visual signals can support image-based categorization. Similar to the third error scenario, natural product textures like vegetable veins or fruit freckles can occasionally be misinterpreted for small damage.

As a result, the above-mentioned model is only appropriate for automatic screening and not for making the ultimate safety choice. According to studies on fish freshness detection [30], domain verification is still necessary for safety-critical applications, while lightweight machine-learning judgments are appropriate for quick screening. Batch routing, shelf-life prioritizing, and inspection scheduling in the output are all guided by high confidence. The product batch will be sent for manual or physicochemical testing if the confidence level is low.

Conclusion

This research introduced a multi-scale Swin Transformer approach for freshness evaluation in cold chain monitoring photos. The model consists of cross-scale shifted-window attention, ordinal freshness choice learning, reliability-aware token modulation, and hierarchical patch embedding. The design separates real-quality deterioration from glare, condensation, light variations, and background interference, making it appropriate for cold picture settings.

According to experimental data, the suggested approach outperforms convolutional networks, Vision Transformer, and the conventional Swin Transformer baseline in terms of overall accuracy, macro-F1 score, spoiled-state recall, resilience, and deployment efficiency. The greatest benefit is seen in identifying mild to moderate deterioration; freshness cues are weak at this point and can be confused for other image noise. The ablation findings show that ordinal choice learning, cross-scale attention, and reliability modulation all provide supplementary improvements.

The system may be expanded in future work to integrate picture streams with data from temperature, humidity, gas sensors, and time-temperature indicators. Additionally, cross-site validation is needed in real refrigerated warehouses, moving cars, and retail display displays. Quantization, edge-cloud collaborative inference, and model compression may all be leveraged to increase the usefulness of large-scale cold chain quality monitoring. Freshness evaluation is a challenge of both visual identification and operational quality control; a large-scale real-world application should also take into account the variety of items, clear package information, camera maintenance, warning responsibility, and so forth.

Author Contributions

Maxime Fournier contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision. Antoine Dupuis contributes to draft preparation, methodology, software, validation, analysis, investigation. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized ChatGPT-4o for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the formulation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

References

[1] Fu, Y., Nguyen, M., & Yan, W. Q. (2022). Grading methods for fruit freshness based on deep learning. SN Computer Science, 3(4), 264. https://doi.org/10.1007/s42979-022-01152-7

[2] Kazi, A., & Panda, S. P. (2022). Determining the freshness of fruits in the food industry by image classification using transfer learning. Multimedia Tools and Applications, 81(6), 7611-7624. https://doi.org/10.1007/s11042-022-12150-5

[3] Gao, T., Sun, D., Tian, Y., & Zhu, Z. (2021). Gold-silver core-shell nanorods based time-temperature indicator for quality monitoring of pasteurized milk in the cold chain. Journal of Food Engineering, 306, 110624. https://doi.org/10.1016/j.jfoodeng.2021.110624

[4] Wu, J., & Hsiao, H. (2021). Food quality and safety risk diagnosis in the food cold chain through failure mode and effect analysis. Food Control, 120, 107501. https://doi.org/10.1016/j.foodcont.2020.107501

[5] Adiani, V., Gupta, S., & Variyar, P. S. (2021). A simple time temperature indicator for real time microbial assessment in minimally processed fruits. Journal of Food Engineering, 311, 110731. https://doi.org/10.1016/j.jfoodeng.2021.110731

[6] Zhang, W., Cao, A., Shi, P., & Cai, L. (2021). Rapid evaluation of freshness of largemouth bass under different thawing methods using hyperspectral imaging. Food Control, 125, 108023. https://doi.org/10.1016/j.foodcont.2021.108023

[7] Soni, A., Al-Sarayreh, M., Reis, M. M., & Brightwell, G. (2021). Hyperspectral imaging and deep learning for quantification of Clostridium sporogenes spores in food products using 1D-convolutional neural networks and random forest model. Food Research International, 147, 110577. https://doi.org/10.1016/j.foodres.2021.110577

[8] Logan, R. D., Scherrer, B., Senecal, J., Walton, N. S., Peerlinck, A., Sheppard, J. W., et al. (2021). Assessing produce freshness using hyperspectral imaging and machine learning. Journal of Applied Remote Sensing, 15(3), 034505. https://doi.org/10.1117/1.jrs.15.034505

[9] Asefa, B. G., Hagos, L., Kore, T., & Emire, S. A. (2022). Feasibility of image analysis coupled with machine learning for detection and quantification of extraneous water in milk. Food Analytical Methods, 15(11), 3092-3103. https://doi.org/10.1007/s12161-022-02352-w

[10] Hemamalini, V., Rajarajeswari, S., Nachiyappan, S., Sambath, M., Devi, T., Singh, B. K., et al. (2022). Food quality inspection and grading using efficient image segmentation and machine learning-based system. Journal of Food Quality, 2022, 5262294. https://doi.org/10.1155/2022/5262294

[11] Almasi, H., Forghani, S., & Moradi, M. (2022). Recent advances on intelligent food freshness indicators; an update on natural colorants and methods of preparation. Food Packaging and Shelf Life, 32, 100839. https://doi.org/10.1016/j.fpsl.2022.100839

[12] Liu, L., Wu, W., Zheng, L., Yu, J., Sun, P., & Shao, P. (2022). Intelligent packaging films incorporated with anthocyanins-loaded ovalbumin-carboxymethyl cellulose nanocomplexes for food freshness monitoring. Food Chemistry, 387, 132908. https://doi.org/10.1016/j.foodchem.2022.132908

[13] Wang, J., Sun, X., Zhang, H., Dong, M., Li, L., Zhangsun, H., et al. (2022). Dual-functional intelligent gelatin-based packaging film for maintaining and monitoring the shrimp freshness. Food Hydrocolloids, 124, 107258. https://doi.org/10.1016/j.foodhyd.2021.107258

[14] Kwon, S., & Ko, S. (2022). Colorimetric freshness indicator based on cellulose nanocrystal-silver nanoparticle composite for intelligent food packaging. Polymers, 14(17), 3695. https://doi.org/10.3390/polym14173695

[15] Luo, X., Zaitoon, A., & Lim, L. T. (2022). A review on colorimetric indicators for monitoring product freshness in intelligent food packaging: Indicator dyes, preparation methods, and applications. Comprehensive Reviews in Food Science and Food Safety, 21(3), 2489-2519. https://doi.org/10.1111/1541-4337.12942

[16] Ayas, S., & Tunc-Gormus, E. (2022). SpectralSWIN: A spectral-Swin Transformer network for hyperspectral image classification. International Journal of Remote Sensing, 43(11), 4025-4044. https://doi.org/10.1080/01431161.2022.2105668

[17] He, X., Chen, Y., & Lin, Z. (2021). Spatial-spectral Transformer for hyperspectral image classification. Remote Sensing, 13(3), 498. https://doi.org/10.3390/rs13030498

[18] Qing, Y., Liu, W., Feng, L., & Gao, W. (2021). Improved Transformer net for hyperspectral image classification. Remote Sensing, 13(11), 2216. https://doi.org/10.3390/rs13112216

[19] Yang, X., Cao, W., Lu, Y., & Zhou, Y. (2022). Hyperspectral image Transformer classification networks. IEEE Transactions on Geoscience and Remote Sensing, 60, 1-15. https://doi.org/10.1109/TGRS.2022.3171551

[20] Zhao, Z., Hu, D., Wang, H., & Yu, X. (2022). Convolutional Transformer network for hyperspectral image classification. IEEE Geoscience and Remote Sensing Letters, 19, 1-5. https://doi.org/10.1109/LGRS.2022.3169815

[21] Hao, S., Wu, B., Zhao, K., Ye, Y., & Wang, W. (2022). Two-stream Swin Transformer with differentiable Sobel operator for remote sensing image classification. Remote Sensing, 14(6), 1507. https://doi.org/10.3390/rs14061507

[22] Ma, C., Jiang, J., Li, H., Mei, X., & Bai, C. (2022). Hyperspectral image classification via spectral pooling and hybrid Transformer. Remote Sensing, 14(19), 4732. https://doi.org/10.3390/rs14194732

[23] Chen, Y., Gu, X., Liu, Z., & Liang, J. (2022). A fast inference Vision Transformer for automatic pavement image classification and its visual interpretation method. Remote Sensing, 14(8), 1877. https://doi.org/10.3390/rs14081877

[24] Wang, X., Yang, S., Zhang, J., Wang, M., Zhang, J., Yang, W., et al. (2022). Transformer-based unsupervised contrastive learning for histopathological image classification. Medical Image Analysis, 81, 102559. https://doi.org/10.1016/j.media.2022.102559

[25] Li, G., Fang, Q., Zha, L., Gao, X., & Zheng, N. (2022). HAM: Hybrid attention module in deep convolutional neural networks for image classification. Pattern Recognition, 129, 108785. https://doi.org/10.1016/j.patcog.2022.108785

[26] Zhang, Z., Li, T., Tang, X., Hu, X., & Peng, Y. (2022). CAEVT: Convolutional autoencoder meets lightweight Vision Transformer for hyperspectral image classification. Sensors, 22(10), 3902. https://doi.org/10.3390/s22103902

[27] Zheng, Y., & Jiang, W. (2022). Evaluation of Vision Transformers for traffic sign classification. Wireless Communications and Mobile Computing, 2022, 3041117. https://doi.org/10.1155/2022/3041117

[28] Zhang, S., Wu, G., Gu, J., & Han, J. (2020). Pruning convolutional neural networks with an attention mechanism for remote sensing image classification. Electronics, 9(8), 1209. https://doi.org/10.3390/electronics9081209

[29] Grassi, S., Benedetti, S., Magnani, L., Pianezzola, A., & Buratti, S. (2022). Seafood freshness: E-nose data for classification purposes. Food Control, 138, 108994. https://doi.org/10.1016/j.foodcont.2022.108994

[30] Yudhana, A., Umar, R., & Saputra, S. (2022). Fish freshness identification using machine learning: Performance comparison of k-NN and Naïve Bayes classifier. Journal of Computing Science and Engineering, 16(3), 153-164. https://doi.org/10.5626/jcse.2022.16.3.153