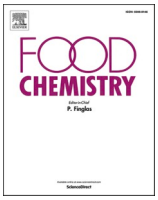
Food Chemistry 495 (2025) 146430

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem



Leveraging pre-trained computer vision models for accurate classification of meat freshness☆

Marcelo M. Hidalgo [a], Robson C. Lima [a,*], Elisabete A. De Nadai Fernandes [a], Márcio A. Bacchi [a], Gabriel A. Sarriés [b]

[a] Nuclear Energy Center for Agriculture, University of São Paulo, Avenida Centenário 303, 13416-000 Piracicaba, SP, Brazil
[b] College of Agriculture Luiz de Queiroz, University of São Paulo, Avenida Pádua Dias 11, 13418-000 Piracicaba, SP, Brazil

ARTICLE INFO

Keywords:
Computer vision
Texture analysis
Pattern recognition
Meat quality assessment

ABSTRACT

Increasing concerns about food quality and safety have led to research into ways to assess meat freshness. Advances in deep learning, particularly image classification, enable up new possibilities for fast and non-destructive methods of evaluating meat properties. This study explored a novel approach for classifying meat freshness based on image data by leveraging features extracted from pre-trained deep convolutional neural networks (DCNNs), followed by random encoding of aggregated deep activation maps (RADAM). The encoded features were subsequently used to train traditional machine learning (ML) classifiers. This approach yielded state-of-the-art results, with classification metrics ranging from 93 to 100 % when classifying beef and chicken meat freshness across three datasets and under multiple hyperparameter settings. Not only does this surpass the performance reported in literature, but it also offers a simpler and more efficient methodology. Findings suggest that the proposed approach could be a practical and effective solution for industry deployment.

1. Introduction

Meat is a globally consumed food product, with production volumes increasing annually to meet rising demand (OECD/FAO, 2022). Meat quality aspects such as freshness and safety are a major concern for the food industry (Zhu et al., 2024) Improper storage or handling can accelerate microbial proliferation and the breakdown of proteins and lipids, which alter key sensory attributes such as odor, color, and texture (Chen et al., 2024; Shao et al., 2021).

Quantification of physicochemical parameters such as spoilage can be done by sensory methods, which rely on organoleptic evaluation, or analytical (Chen et al., 2024). Traditional analytical methods for assessing meat spoilage include measurements such as pH, water holding capacity and drip loss, total volatile basic nitrogen, and total viable count (Bekhit et al., 2021; Cai et al., 2024). More recently, foodomics - an integrative approach encompassing proteomics, metabolomics, lipidomics, and transcriptomics - has greatly advanced our molecular-level understanding of spoilage processes. Utilizing high-throughput technologies such as LC-MS, GC–MS, NMR, and RNA-Seq, foodomics captures complex biochemical signatures that reflect spoilage dynamics and sensory deterioration with high specificity and depth (Bevilacqua et al., 2017; Valdés et al., 2022).

In parallel, sensor-based technologies have been increasingly applied for food quality assessment (Chen et al., 2024). Examples include electronic noses (*E*-noses) and electronic tongues (E-tongues), which mimic human olfactory and gustatory perception through arrays of gas and electrochemical sensors, respectively (Putri et al., 2023; Wojnowski et al., 2017). Beyond E-nose and E-tongue technologies, optical sensor-based approaches have diversified into several advanced techniques, including visible (VIS), near-infrared (NIR), mid-infrared (MIR), ultraviolet-visible (UV-VIS), Raman spectroscopy, and hyperspectral imaging (Schreuders et al., 2021; Wu et al., 2022).

Although analytical and foodmics approaches yield detailed and reliable results, their application in routine meat quality monitoring is constrained by high costs, complex instrumentation, and the need for skilled personnel and dedicated lab infrastructure (Cajka, 2024; Shi et al., 2024). Additionally, time-consuming protocols and extensive sample preparation increase variability and hinder scalability, limiting their suitability for constant, on-site freshness assessment across the supply chain (Shi et al., 2024; Valdés et al., 2022).

☆ This article is part of a Special issue entitled: 'AI for Food Chemistry' published in Food Chemistry.

* Corresponding author.
E-mail address: robsoncamposdelima@usp.br (R.C. Lima).

https://doi.org/10.1016/j.foodchem.2025.146430
Received 1 November 2024; Received in revised form 9 September 2025; Accepted 15 September 2025
Available online 25 September 2025

Another approach, computer vision using RGB images, presents versatility, robustness, and wide applicability in industrial settings, including food inspection. Leveraging advanced image processing and machine learning techniques, computer vision systems analyze visual cues related to meat freshness (Sánchez et al., 2023), such as color (Wu & Sun, 2013), marbling (Liao et al., 2025), and surface texture (Modzelewska-Kapituła & Jun, 2022). This approach is scalable and non-destructive for real-time meat quality assessment (Kaushal et al., 2024; Saha et al., 2025).

1.1. Lightweighting computer vision

Deep convolutional neural networks (DCNNs) have achieved remarkable success in computer vision tasks, including object recognition, segmentation, and visual quality assessment (Kaushal et al., 2024; Zhao et al., 2025). However, as network architectures become deeper and more complex, the associated demands for training data and computational resources increase, making full model optimization costlier and time-consuming (Lee et al., 2022). While advances in hardware have partially mitigated performance issues, they still rely on large, high-quality labeled datasets (Luo et al., 2024; Rex et al., 2022; Saha et al., 2025). This requirement poses challenges in applications such as food quality monitoring, where collecting and labeling data is often labor-intensive, time-consuming, and costly (Luo et al., 2024; Rex et al., 2022).

An alternative to training DCNNs involves extracting and leveraging high-level representations from pre-trained models without requiring additional training or fine-tuning (Filus & Domańska, 2023). One such method is the Random encoding of Aggregated Deep Activation Maps (RADAM), which leverages the rich texture information embedded in intermediate network layers and transforms it into a compact numerical vector suitable for traditional machine learning classifiers (Scabini et al., 2023). By avoiding large-scale training of DCNNs, RADAM enables the development of accurate and robust models that require significantly less computational power and can operate effectively with smaller datasets and simpler hardware configurations (Huang et al., 2006; Kasun et al., 2013).

The raw texture features in RADAM originate from the intermediate activation maps of a pre-trained DCNN, which capture spatial patterns at multiple semantic levels, ranging from fine-grained textures to coarser, global structures. The RADAM aggregates these activation maps and encodes them using randomized autoencoders (RAEs) to produce compact, texture-rich feature vectors.

Because the activation maps have different spatial resolutions, each map is resized via bilinear interpolation to match the spatial dimensions of the selected intermediate stage of the DCNN, enabling their aggregation into a single structure, referred to as the aggregated activation map. Since deeper layers typically retain coarser semantic features with reduced resolution, upscaling such maps tends to preserve more structural information than downscaling early-layer maps (Scabini et al., 2023). Therefore, the encoding process implicitly prioritizes activation maps from later stages, leveraging the hierarchical nature of convolutional architectures.

After being resized to a common resolution, all activation maps are concatenated along the channel axis, forming a single tensor that brings together texture information from different layers of the network. This tensor is then reshaped into a 2D matrix (the aggregated activation map) in which each row corresponds to a pixel location, and each column holds features learned at that position.

The aggregated activation map is then processed by a randomized autoencoder (RAE) specific to each image. Instead of traditional training, the RAE applies a fixed, randomly generated linear projection followed by a nonlinear activation to encode the high-dimensional data into a lower-dimensional latent space. Then, a second linear projection is analytically computed to reconstruct the original data from this latent representation, avoiding iterative optimization. This process effectively compresses the deep features into a compact vector that preserves important texture information. The resulting low-dimensional feature vector can be used directly for classification or retrieval tasks.

Considering the potential of RADAM to reduce the need for extensive model retraining, this study investigates its effectiveness and suitability for classifying meat freshness from image data. We aim to evaluate whether RADAM, when combined with different pre-trained DCNN backbones and traditional machine learning classifiers, can achieve competitive classification performance while offering improved computational efficiency. The work is framed as an external visual inspection task, with the intended application of computer vision models in non-destructive and rapid inspection scenarios. While this approach does not capture all aspects of freshness, it represents an important step toward assessing whether outer visual features alone can serve as reliable indicators.

2. Methodology

The RADAM method was evaluated using three meat image datasets, testing different implementation settings, including four pre-trained DCNNs for feature extraction and three machine learning classifiers for freshness classification. In all cases, freshness assessment was limited to the visual appearance of the meat samples, consistent with the objective of exploring outer visual cues as discriminative features. The goal was to investigate the trade-off between feature compactness and discriminative power in RADAM-based pipelines. The optimized configurations were compared with the results of other methods for the same datasets. An overview of the entire pipeline is illustrated in Fig. 1.

2.1. Datasets

Three image datasets focused on meat freshness was used: two containing images of beef and one of chicken meat. Fig. 2 shows some examples found in each dataset. All dataset images were resized to 224 × 224 pixels before being processed by RADAM for feature extraction, as this resolution provided the best accuracy-computational cost trade-off according to the evaluations of Scabini et al. (2023).

2.1.1. Meat quality assessment dataset

The first bovine meat dataset, hereafter mentioned as ‘Meat Quality Assessment Dataset’, was originally presented by Ulucan et al. (2019). These authors monitored the spoilage progression in meat stored in trays. Periodic images of sliced meat were acquired using a high-resolution IP camera installed in the butcher section of a processing plant. A total of 1896 images with a resolution of 1280 × 720 pixels, taken from a superior view (Fig. 2a). Those images were equally divided into two categories, *fresh* and *spoiled*, each comprising 948 images. These categories were assigned by a processing unit specialist based on visual and sensory parameters. Fresh samples were characterized by uniform red color, firm texture, and absence of off-odors, whereas spoiled samples showed color darkening, softening, and perceptible spoilage odor.

2.1.2. Meat freshness image dataset

The second bovine meat dataset, named ‘Meat Freshness Image Dataset’, was analyzed by Büyükarıkan (2024). It holds 2266 RGB images categorized as *fresh* ($n = 853$), *half-fresh* ($n = 789$) and *spoiled* ($n = 624$). These freshness stages were determined through visual inspection at the collection site, reflecting progressive deterioration of the meat.

All images were captured at a close range, underwent automatic EXIF-orientation stripping and stretched to 416 × 416 pixels to emphasize meat regions and minimize background noise (Fig. 2b). Fresh images depicted bright red meat with firm texture, half-fresh images represented an intermediate stage observed between fresh and spoiled, showing early signs of color darkening and slight softening, and spoiled images presented extensive discoloration and often noticeable spoilage

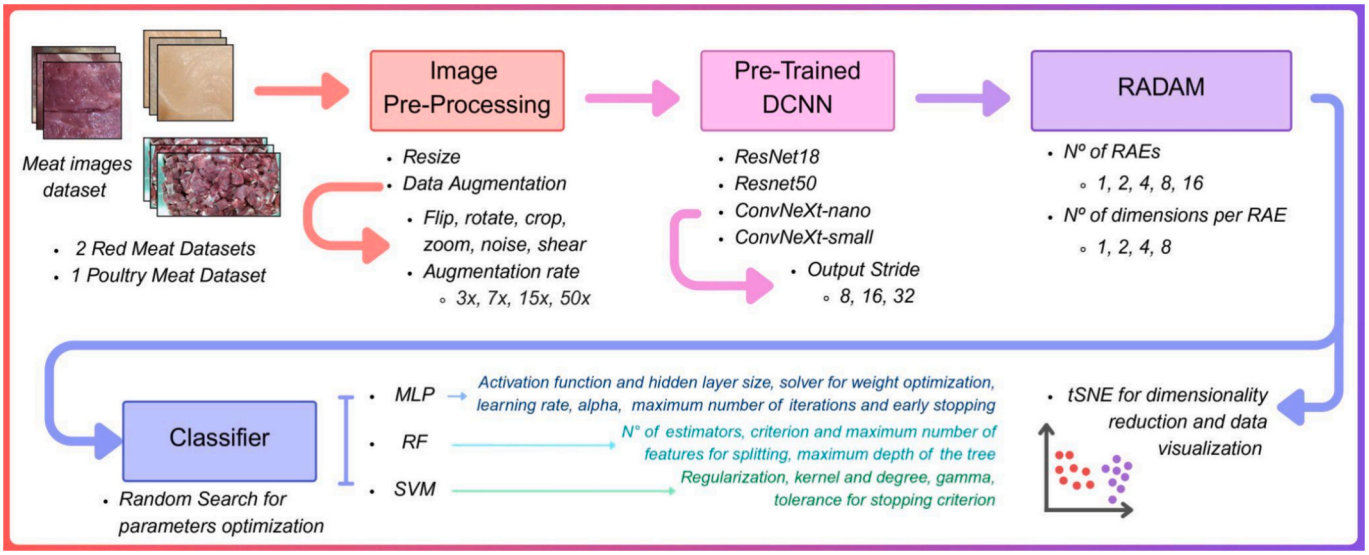


Fig. 1. Workflow overview of the methodology used in this study.

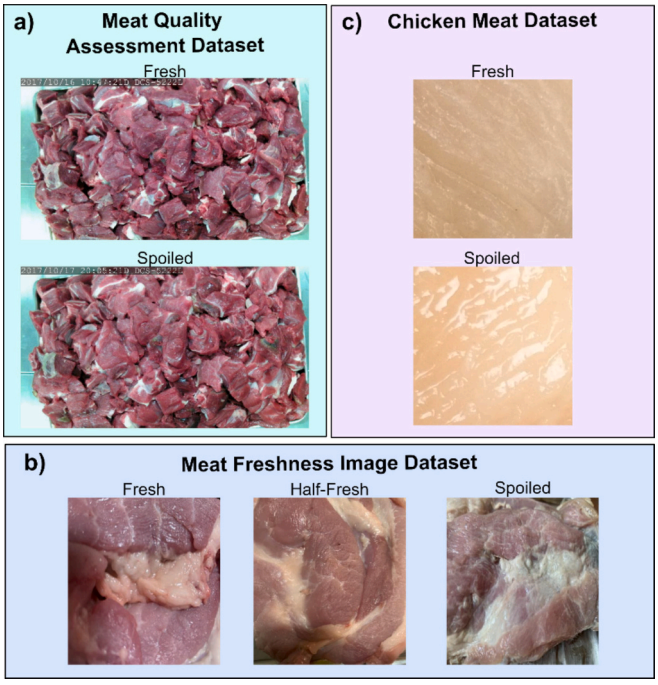


Fig. 2. Examples of each class within each dataset analyzed in this study.

indicators.

2.1.3. Chicken meat dataset

The 'Chicken Meat Dataset', analyzed by Calvin and Prakasa (2020), comprises 433 images of malay chicken breast meat in two categories: *spoiled* ($n = 218$) and *fresh* ($n = 215$). Fresh samples were excised 6–8 h post-mortem and trimmed to a thickness of 0.5 cm. Initial images were captured using an 8MP smartphone camera positioned approximately 10 cm from the sample. These samples were then held at ambient temperature for approximately 22 h before being re-photographed and classified as spoiled. Finally, all images were resized to 400 × 400 pixels to eliminate background noise.

2.2. RADAM settings

To optimize classification performance and computational efficiency, four hyperparameters of the RADAM framework were systematically varied: (1) the DCNN architecture used for feature map extraction; (2) the output stride configuration, which determines the spatial resolution of the activation maps; (3) the number of RAEs and; (4) the number of projections per RAE. These parameters directly modulate the quantity and quality of the information captured and subsequently encoded.

2.2.1. DCNN

To evaluate RADAM's ability to extract discriminative texture features from deep networks, four pre-trained DCNNs commonly used in image recognition were employed. Two of them were residual networks (ResNet-18 and ResNet-50) and two were ConvNeXt variants (Nano and Small), selected to represent distinct architectural paradigms and depths. For each DCNN, activation maps were extracted from all main stages, corresponding to the natural structural divisions of each architecture.

All DCNNs were implemented in Python using the Timm library (PyTorch Image Models), which provides access to state-of-the-art vision models. All the selected architectures were used with their official pre-trained weights, obtained from training on the ImageNet-1 k dataset, which contains approximately 1.2 million natural images across a thousand object categories.

2.2.2. Output stride

In DCNNs, the output stride (OS) refers to the ratio between the input image resolution and the spatial resolution of the final feature maps before global pooling or classification layers. For example, an OS of 32 implies that each spatial dimension of the output feature map is 1/32 while OS values of 16 and 8 yield progressively denser maps (e.g., 14 × 14 and 28 × 28, respectively, for a 224 × 224 input). Denser feature maps better preserve fine-grained texture and local spatial information, albeit at increased computational cost. Values of 8, 16, and 32 were considered to assess their influence on texture preservation and computational cost.

2.2.3. Number of RAEs and projections per RAE

RADAM encodes texture features using ensembles of random autoencoders (RAEs). Two hyperparameters govern this process: (a) the number of RAEs in the ensemble and (b) the number of projections

produced by each RAE. These parameters were adjusted to modulate the amount of texture information extracted and encoded, enabling the assessment of how information density affects both classification accuracy and computational efficiency. In the original implementation proposed by Scabini et al. (2023), it is proposed the use of four RAEs, each producing a single projection.

The number of RAEs defines the size of the ensemble of random coders whose decoders (i.e., learned weights) are aggregated to form the final descriptor. Increasing the number of RAEs allows for a greater number of projections from the aggregated activation map, which in turn enables the model to capture a wider variety of textural representations. This increases the likelihood of identifying more complex and discriminative patterns.

In the original study presenting RADAM, Scabini et al. (2023) showed that using multiple RAEs improves representation quality by averaging over different random projections. However, performance shows a decreasing marginal growth. Based on this, and seeking to balance performance with computational cost, three values for the number of RAEs were tested: 2, 4 and 8.

The number of projections per RAE corresponds to the size of the spatial feature vector input at each spatial position of the aggregated feature map. In RADAM, each spatial position is represented by a d-dimensional vector, where d is the number of concatenated channels.

Lower dimensions (e.g., 1 or 2) significantly reduce these costs, but may limit representational expressiveness by ignoring inter-channel dependencies. Higher input dimensions (e.g., 8) allow each RAE to model complex multichannel texture interactions, potentially enriching the representation. However, this also leads to costlier linear systems (e.g., in solving least squares problems) and larger output descriptors. In this study, we tested four input dimensionalities per RAE: 1, 2, 4, and 8.

2.3. Classification and fine-tuning

The extracted features by RADAM were then provided to three classification algorithms, Random Forest (RF), Support Vector Machine (SVM) and Multilayer Perceptron (MLP). They were chosen because both are well-suited for handling high-dimensional data, which makes them adequate choices for classification using the texture features extracted with RADAM. They also present some robustness when dealing with noise and other interferences that are common in complex datasets.

To fit the models, a randomized search strategy was employed for hyperparameter optimization, using stratified 5-fold cross-validation and the weighted F1-score as the evaluation metric. Each model was embedded in a pipeline that included feature standardization. The number of randomized iterations was set to 300.

For SVM, the search spanned multiple kernel types (linear, polynomial, RBF), regularization strengths, gamma values, polynomial degrees, and class balancing options. RF was tested with varied the number of trees, splitting criteria (Gini, entropy, log-loss), maximum tree depths and feature selection strategies. The MLP was optimized across different hidden layer architectures, activation functions, solvers, learning rate strategies, regularization strengths, and stopping criteria.

After hyperparameter optimization, the classification was conducted following two different approaches. First, the data were classified using 5-fold cross-validation, which allocates 20 % of the data for testing the trained models in each cross-validation. This approach aimed to provide a broad assessment of the models' performance. Subsequently, the classifiers were retrained and re-evaluated using similar procedures to those adopted by previous literature analyzing these datasets, as they differ in the percentage of data allocated for training and testing. The Meat Quality Assessment Dataset was classified using 20-fold cross-validation, which leaves 5 % of the data for testing, similar to the approach used by Amani and Sarkodie (2022). For the Meat Freshness Image Dataset and Chicken Meat Dataset, classification was performed using the data splits as originally provided by Büyükarıkan (2024) and Calvin and Prakasa (2020) (Table S1).

2.3.1. Evaluation metrics

Five metrics were used to evaluate classification performance, namely accuracy, precision, recall, F1-score and Area Under the Receiver Operating Characteristic Curve (AUC-ROC), defined by Eqs. (1), (2), (3), (4):

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 = 2.\frac{Precision^{*}\ Recall}{Precision + Recall} \tag{4}$$

where TP stands for true positive, TN for true negative, FP for false positive, and FN for false negative.

Accuracy represents the ratio of correct predictions (TP and TN) to the total number of instances (Eq. 1). Precision refers to the proportion of TP among all positive predictions (Eq. 2), indicating how many predicted positives are correct. Recall (or sensitivity) measures the proportion of actual positives correctly identified by the model (Eq. 3), reflecting how many TP were captured. The F1-score (Eq. 4) is the harmonic mean of precision and recall, and it emphasizes the balance between false positives and false negatives.

The AUC-ROC measures the model's ability to distinguish between classes across all classification thresholds. It is calculated by plotting the true positive rate (TPR) or recall (Eq. 3) against the false positive rate (FPR, Eq. 5) at various threshold settings and computing the area under the resulting curve.

$$FPR = \frac{FP}{FP + TN} \tag{5}$$

3. Results

3.1. DCNN for feature extraction

Table 1 shows the average F1-score obtained across the three datasets for each DCNN tested: ConvNeXt Nano, ConvNeXt Small, ResNet18, and ResNet50. Each model was evaluated under various RADAM configurations, including output stride (8, 16, 32), number of RAEs (1, 2, 4, 8), projections per RAE (1, 2, 4, 8), and classifier (MLP, RF, SVM).

Across datasets, classification performances were generally similar among the four DCNNs, especially in the first two datasets. In the Meat Quality dataset, ResNet18 obtained the highest average F1-score (0.995). In the Meat Freshness dataset, ConvNeXt-Nano achieved the best score (0.941), grouped with ConvNeXt-Small and ResNet50. In the Chicken Meat dataset, ConvNeXt-Small scored highest (0.926), while ResNet18 and ResNet50 lower scores (0.873 and 0.886).

Considering the average results across datasets, ConvNeXt-Nano and ConvNeXt-Small presented the highest F1 scores (0.949 and 0.948), followed by ResNet50 (0.937) and ResNet18 (0.930).

In terms of runtime (Table 1), ConvNeXt-Nano and ResNet18 were the most efficient models, with average processing times of 44.6 s and 40.4 s, respectively. In contrast, ConvNeXt-Small and ResNet50 demanded significantly more time (85.5 s and 104.3 s, respectively), reflecting their increased architectural complexity (49.45 M and 23.51 M parameters, respectively).

3.2. 4.2 Spatial reduction by output stride

Fig. 3 displays the mean F1-scores for each output stride (OS) value.

Table 1
Performance summary of DCNNs across datasets, including F1-score, model size, and processing time.

DCNN	Meat quality assessment dataset	Meat freshness image dataset	Chicken meat dataset	Average F1-Score across datasets	N° Parameters (Mi)	Average F1-Score	Time (s)
	Average F-score						
ConvNeXt-Nano	0.993	0.941	0.911	0.949	14.95	0.947	44.6
ConvNeXt-Small	0.985	0.930	0.926	0.947	49.45	0.948	85.5
ResNet18	0.995	0.920	0.873	0.930	11.18	0.930	40.4
ResNet50	0.995	0.923	0.886	0.937	23.51	0.937	104.3

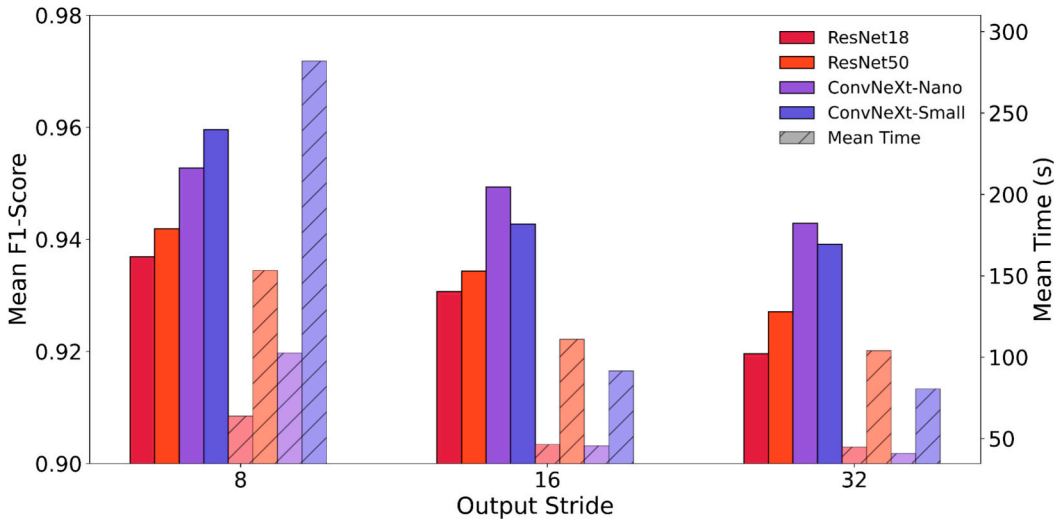


Fig. 3. Impact of output stride on model accuracy and runtime.

Lower strides, such as OS = 8, yielded slightly higher scores (mean F1 = 0.947), while OS = 16 and 32 led to marginally lower performance (0.940 and 0.933, respectively). However, this pattern was not consistent across all DCNNs.

Regarding processing time, lower OS values did not uniformly result in higher computation times. For example, ConvNeXt-Small with OS = 8 required 281.9 s on average, while OS = 32 required only 80.5 s, indicating that time complexity was more influenced by other tested parameters. Considering the trade-offs, ConvNeXt-Nano and ResNet18 emerged as the most balanced options for practical deployment.

3.3. Randomized encoding parameters

The joint effects of the number of RAEs and the number of projections per RAE on RADAM performance were evaluated in terms of both feature extraction time and classification accuracy (Fig. 4).

The number of RAEs influenced on feature extraction time. Using a single RAE resulted in the slowest performance, whereas employing two RAEs made the process approximately three times faster. Beyond two RAEs, increases in their number led to progressively smaller improvements, with runtime tending to stabilize.

The number of projections per RAE had a limited influence on runtime, particularly when more than two RAEs were used. Time variation due to this parameter was only substantial under the single-RAE condition.

Regarding classification performance, the mean F1-score was higher when using one projection per RAE. In general, lower projection counts produced better performance, especially when the number of RAEs was

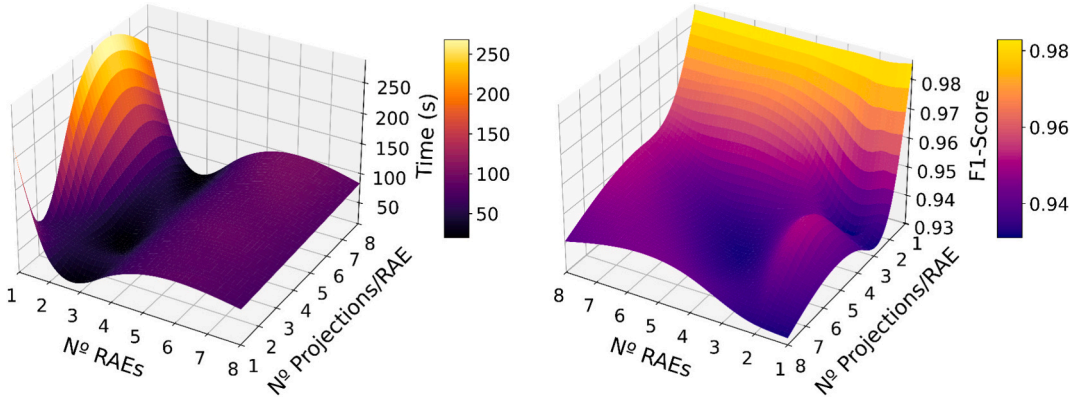


Fig. 4. Response surfaces for feature extraction time and f1-score based on RADAM parameters.

low. Increasing the number of projections tended to reduce accuracy and increase variability. Larger numbers of RAEs mitigated this effect, stabilizing performance across projection configurations.

Overall, the results suggest that increasing the number of RAEs leads to more stable classification performance, while using fewer projections per RAE helps preserve discriminative power. These findings reinforce the importance of carefully balancing redundancy and dimensionality in random encoding strategies.

3.4. Classifiers settings

The average F1-score of 5-fold classification was evaluated using RF, SVM, MLP based on features extracted from each DCNN setting defined by Output Stride, number of RAEs, and projections per RAE. For both ConvNeXt-Nano and ResNet18 architectures, the best-performing configuration corresponded to an Output Stride of 32, 2 RAEs, and 1 projection per RAE. With those features extracted from each dataset by the pre-trained DCNNs, t-SNE was applied for dimensionality reduction. Fig. 5 illustrates the two-dimensional embedding of observations for each dataset.

Fig. S1 presents the classification performance obtained using the extracted features across different classifier combinations on the three datasets. Classification was performed using the best hyperparameters from the Random Search procedure described in Section 2.3, and results are detailed in supplementary material (Table S2).

For the Meat Quality Assessment dataset, all classifiers achieved very high performance (mean F1-score $>$ 0.99) with low variability across folds (standard deviation $\leq$0.0053), indicating high class separability (Fig. 5A and D). ResNet18 combined with SVM, ResNet18 with MLP, and ConvNeXt-Nano with SVM all achieved the same mean F1-score (0.997), with ResNet18-based models exhibiting lower variability ($\pm$ 0.002) compared to ConvNeXt-Nano + SVM ($\pm$ 0.004), and ResNet18 + SVM also achieving the lowest search time (0.20 s). In contrast, MLP-based models required substantially longer training times (4.00 s with ConvNeXt-Nano and 8.63 s with ResNet18), reflecting the increased computational cost associated with deeper hidden layers and the absence of early stopping.

RF achieved slightly lower performance (F1 = 0.995), requiring a large number of estimators (up to 497 trees with ResNet18), which significantly increased fitting time (7.54 s). In the Meat Freshness Image dataset, ConvNeXt-Nano combined with SVM again yielded the higher scores (F1 = 0.996 $\pm$ 0.003). Optimized parameters indicated a simpler SVM model (C = 1.86, RBF kernel with gamma = 'auto', degree 2), suggesting that less complex configurations suffice for this dataset. ResNet18 achieved slightly lower performance (F1 = 0.990) but required half the training time. RF showed considerably worse performance on this dataset (F1 between 0.952 and 0.966) with long search times (up to 34.20 s). MLP was competitive (F1 up to 0.993) but incurred higher computational cost compared to SVM.

The Chicken Meat dataset exhibited greater variability (standard deviation up to 0.020). The highest performance was achieved by ConvNeXt-Nano + SVM (F1-score of 0.986 $\pm$ 0.014). Although t-SNE visualization indicated lower class separability (Fig. 5C and F), the SVM achieved high performance using a linear kernel and exhibited short training time (0.01 s), suggesting that the extracted features allow for effective linear discrimination despite the apparent complexity.

3.5. Final setting and comparison with literature

The mean scores for accuracy, precision, recall, F1-score and AUC achieved by SVM, RF and MLP, using features extracted with ConvNeXt-Nano and ResNet18, using 5-fold cross-validation for each dataset are presented in Table 2.

The models showed consistently high and uniform performance across all evaluation metrics, with similar average values of accuracy, precision, recall, and F1-score for each configuration. Precision and recall were notably high for all tested combinations, while AUC values were close to or equal to 1, reflecting excellent class separability in the

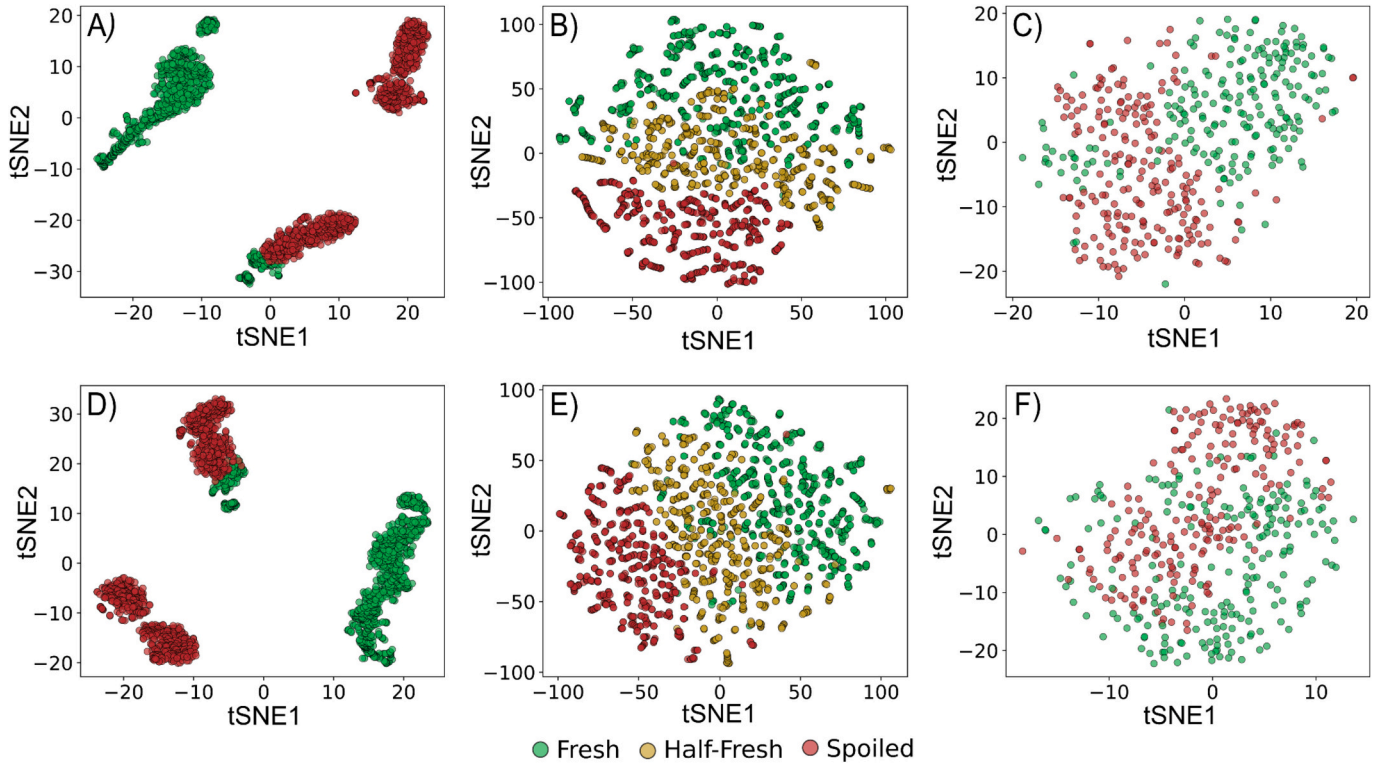


Fig. 5. Two-Dimensional t-SNE Visualization of Features Extracted by Pre-trained DCNNs. (A–C) show features extracted using ConvNeXt-Nano; (D—F) show features extracted using ResNet-18. (A, D): Meat Quality Assessment Dataset; (B, E): Meat Freshness Image Dataset; (C, F): Chicken Meat Dataset.

Table 2
Classification performance achieved with 5-fold cross-validation.

Dataset	Algorithm	Algorithm	Mean ± Standard Deviation				
			Accuracy	Precision	Recall	F1-score	AUC
Meat Quality Assessment Dataset $n = 1896$	ConvNeXt-Nano	SVM	**0.997 ± 0.004**	**0.997 ± 0.004**	**0.997 ± 0.004**	**0.997 ± 0.004**	**1.000 ± 0.000**
		RF	0.993 ± 0.009	0.993 ± 0.009	0.993 ± 0.009	0.993 ± 0.009	1.000 ± 0.000
		MLP	0.995 ± 0.007	0.995 ± 0.007	0.995 ± 0.007	0.995 ± 0.007	1.000 ± 0.000
	ResNet18	SVM	**0.997 ± 0.002**	**0.997 ± 0.002**	**0.997 ± 0.002**	**0.997 ± 0.002**	**1.000 ± 0.000**
		RF	0.994 ± 0.005	0.994 ± 0.005	0.994 ± 0.005	0.994 ± 0.005	1.000 ± 0.000
		MLP	**0.997 ± 0.002**	**0.997 ± 0.002**	**0.997 ± 0.002**	**0.997 ± 0.002**	**1.000 ± 0.000**
Meat freshness image dataset $n = 2266$	ConvNeXt-Nano	SVM	**0.996 ± 0.003**	**0.996 ± 0.003**	**0.996 ± 0.003**	**0.996 ± 0.003**	**1.000 ± 0.000**
		RF	0.965 ± 0.010	0.966 ± 0.010	0.965 ± 0.010	0.965 ± 0.010	0.997 ± 0.001
		MLP	0.993 ± 0.007	0.993 ± 0.007	0.993 ± 0.007	0.993 ± 0.007	1.000 ± 0.000
	ResNet18	SVM	0.990 ± 0.004	0.990 ± 0.004	0.990 ± 0.004	0.990 ± 0.004	1.000 ± 0.000
		RF	0.944 ± 0.009	0.945 ± 0.010	0.944 ± 0.009	0.944 ± 0.009	0.992 ± 0.001
		MLP	0.984 ± 0.004	0.984 ± 0.004	0.984 ± 0.004	0.984 ± 0.004	0.999 ± 0.000
Chicken Meat Dataset $n = 433$	ConvNeXt-Nano	SVM	**0.986 ± 0.014**	**0.986 ± 0.014**	**0.987 ± 0.013**	**0.986 ± 0.014**	**0.998 ± 0.003**
		RF	0.965 ± 0.021	0.965 ± 0.021	0.967 ± 0.020	0.965 ± 0.021	0.994 ± 0.006
		MLP	0.981 ± 0.019	0.981 ± 0.019	0.982 ± 0.018	0.981 ± 0.019	0.997 ± 0.004
	ResNet18	SVM	0.963 ± 0.015	0.964 ± 0.015	0.963 ± 0.015	0.963 ± 0.015	0.993 ± 0.005
		RF	0.926 ± 0.016	0.927 ± 0.016	0.926 ± 0.016	0.926 ± 0.016	0.978 ± 0.010
		MLP	0.956 ± 0.018	0.957 ± 0.018	0.956 ± 0.018	0.956 ± 0.018	0.991 ± 0.006

feature space extracted by the DCNNs. Among the classifiers, SVM yielded slightly superior results, contributing to state-of-the-art performance in meat freshness identification, with accuracy ranging from 92.6 % (ResNet18 + RF on the Chicken Meat Dataset) to 99.7 % (various combinations in Meat Quality Assessment). Notably, ConvNeXt combined with SVM consistently achieved the best results across all datasets (Table 2), and was therefore selected as the reference configuration for comparison with findings in the literature.

Table 3 presents the accuracy obtained using a methodology consistent with the literature. All metrics besides accuracy were left out as they were identical, and this was a common metric between our study and the mentioned literature. For the Meat Quality Assessment Dataset, a 20-fold cross-validation was applied to replicate the 5 % test set split used by Amani and Sarkodie (2022), while for the other two datasets the split is summarized in Table S1. Detailed results for each cross-validation are also available in the supplementary material (Table S3).

The data split utilized in the referenced literature improved the classification performance. The mean accuracy for the Meat Quality Assessment Dataset was basically the same. It is worth mentioning that with this procedure, the employed method achieved an accuracy of 100 % in fifteen of the 20-folds (Table S3). Regarding both Meat Freshness Image Dataset and Chicken Meat Dataset, the provided split resulted in a perfect classification by the SVM, with 100 % accuracy.

4. Discussion

4.1. Dataset quality

The classification performance achieved using the Meat Quality Assessment Dataset was superior compared to the other datasets evaluated. This dataset offers two key advantages. First, it contains high-resolution images (1280 × 720 pixels), whereas the other datasets include images with lower resolutions (416 × 416 and 400 × 400 pixels, respectively. Although all images were resized to 224 × 224 pixels prior to feature extraction using DCNNs, higher original resolution tends to preserve more relevant visual information after preprocessing (Rukundo, 2023).

The second advantage pertains to the content of the images. In the Meat Quality Assessment Dataset, Ulucan et al. (2019) monitored meat trays over time that had been pre-classified by industry experts as fresh or spoiled based on environmental and visual parameters. These images depict entire trays containing multiple meat portions, often exhibiting varying stages of spoilage. The presence of even a single deteriorated section could suffice for accurate classification. In contrast, the other two datasets comprise images of individual meat pieces, making spoilage detection subtler and challenging.

In contrast, the Chicken Meat Dataset posed greater challenges for meat freshness classification. Although its image resolution is comparable to that of the Meat Freshness Image Dataset, the images are limited to small regions of the meat surface. This close-up perspective may enhance surface detail capture but restricts variability across different areas of the cut. Moreover, the spoilage protocol adopted by Calvin et al. (2020) involved exposing fresh meat to ambient conditions for nearly 24 h.

While such exposure likely initiated degradation processes, visual inspections suggest that spoilage effects were subtler in this dataset compared to the others. This case exemplifies the value of complementary sensor-based methods (e.g., Raman spectroscopy, electronic nose), which provide richer chemical and structural information (Schreuders et al., 2021; Wu et al., 2022). Integrating visual inspection with such modalities in multimodal frameworks may therefore offer a more comprehensive basis for meat freshness classification, addressing attributes not detectable through external appearance alone.

Regarding the Meat Freshness Image Dataset, it is noteworthy that t-SNE visualization of the extracted features revealed a gradual transition from fresh to half-fresh and spoiled classes (Fig. 5B and E). The samples were distributed in relatively cohesive clusters, though not clearly separated, suggesting a continuous spoilage process over time, consistent with expected biological degradation.

4.2. Classification performance

Ulucan et al. (2019) achieved a maximum accuracy of 99.58 % in the classification of Meat Quality Assessment Dataset. They employed

Table 3
Classification accuracy achieved with a split similar to original studies.

Dataset	Meat Quality Assessment Dataset			Meat Freshness Image Dataset		Chicken Meat Dataset	
Approach	ConvNext-Nano+SVM	Ulucan et al. (2019)*	Amani and Sarkodie (2022)*	ConvNext-Nano+SVM	Büyükarıkan (2024)*	ConvNext-Nano+SVM	Calvin and Prakasa (2020)*
	99.74 %	99.58 %	100.00 %	100.00 %	99.20 %	100.00 %	92.90 %

DCNN composed of 24 layers, including the input layer, followed by five sequential blocks of convolutional, activation (ReLU), normalization, and pooling layers, culminating in a fully connected layer for final classification. Amani and Sarkodie (2022), who also classified the Meat Quality Assessment Dataset, reported an accuracy of 100 % using DCNN optimized with a particle swarm optimization algorithm - which is a significantly more complex and computationally intensive approach compared to the one proposed here.

Given the limited size of this test subset, however, such a split may not fully capture the variability present in the remaining 95 %. This unexplored variability could pose a potential risk to the model's generalizability and lead to misclassifications under different conditions. In fact, this was observed in five of the folds, where the model failed to correctly classify the entire dataset (Table 2 and S3).

As can be found in Table 3, ConvNeXt + SVM also achieved accuracies of 100 % in fifteen of the 20 folds when performing classification using 5 % of the dataset for testing the trained model. This approach achieved accuracy comparable to existing methods while reducing computational cost. By extracting features from activation maps of a pre-trained DCNN using RADAM, extensive model training was avoided, allowing for evaluation across multiple cross-validation folds. Additionally, the implementation proposed by Scabini et al. requires minimal hyperparameter tuning, as its original configuration represents a good initial estimate that is close to optimal performance (Section 3.3, Fig. 4).

When classifying the Chicken Meat Dataset Calvin et al. (2020) trained three well-known DCNN models - AlexNet, VGGNet, and GoogLeNet (also known as Inception-v1) – alongside a custom model named Ayam6Net, to classify meat freshness. The maximum accuracy achieved was 92.9 % with Ayam6Net, while GoogLeNet reached an accuracy of 85.7 % and both AlexNet and VGGNet performed poorly, with accuracies near to 50.0 %. The differences in performance of these models can be partially associated with the complexity of their architectures. AlexNet and VGGNet have over 200 million parameters each, whereas Ayam6Net and GoogLeNet have approximately 19 and 6 million parameters, respectively. Therefore, AlexNet and VGGNet require significantly larger datasets for adequate training to effectively learn the relevant patterns needed to discriminate meat freshness, compared to Ayam6Net and GoogLeNet.

By leveraging the pre-trained ConvNeXt-Nano model with SVM as the classifier, it was possible to surpass the accuracies reported by Calvin et al. (2020) without facing the challenge of balancing dataset size and model complexity. This was demonstrated both through 5-fold cross-validation (Table 2) and using the dataset split provided by the authors, which resulted in perfect classification (Table 3).

Büyükarıkan (2024) conducted a study methodologically similar to ours, albeit with greater depth. Rather than using RADAM, she employed the GlobalAveragePooling2D (GAP) layer to extract features from activation maps of pre-trained networks such as VGG16, DenseNet, and ResNet50. Her analysis further explored alternative color spaces beyond RGB, including HLS and HSV. When combined with bi-directional long short-term memory (bi-LSTM) networks, this approach led to improved classification performance, achieving accuracy, precision, and recall of 98.9 %, 99.0 %, and 98.9 %, respectively. Notably, our approach surpassed this performance without the need to explore alternative color representations, relying solely on features extracted from RGB images.

4.3. Impact of DCNN architecture on performance and efficiency

The selection of the DCNN plays a critical role in image classification tasks, as different architectures operate based on distinct design principles and computational budgets (Jeevan & Sethi, 2024). This impact is thoroughly explored in the study proposing the RADAM method (Scabini et al., 2023), where the DCNN model was treated as a key experimental variable across analyses. Each DCNN family varies in the way it encodes visual patterns, the number and depth of activation maps generated (i.e., channels), and the accepted input size - all of which influence the quantity and resolution of information available for convolutional processing (Liu et al., 2022; Mustapha et al., 2024; Rukundo, 2023). For instance, within the ConvNeXt family, the Nano variant has approximately 14.9 million parameters, while the XL variant has over 348 million, illustrating the wide spectrum of architectural complexity and resource demand (Liu et al., 2022). Such differences directly affect the network's capacity to extract discriminative features and, ultimately, classification performance and efficiency (Jeevan & Sethi, 2024).

ConvNeXt-Nano and ConvNeXt-Small outperformed the ResNet architectures in terms of average F1-score. While ConvNeXt-Nano and ConvNeXt-Small performed similarly in terms of average F1-score, ConvNeXt-Nano ConvNeXt-Nano did so with only one-fifth the number of parameters and nearly half the inference time, highlighting its practical scalability. ResNet50 showed marginally better performance than ResNet18 but incurred a substantial computational cost. Despite its lower accuracy, ResNet18 presents a viable alternative in resource-limited environments.

The OS parameter exhibited a limited impact on classification performance. There was an apparent trend of decreasing average F1-scores with increasing OS values, which seemed more pronounced in deeper and more parameter-rich architectures such as ConvNeXt-Small and ResNet50. This suggests that these models may be more sensitive to the reduced spatial resolution of activation maps resulting from high OS values, potentially impairing their ability to capture fine-grained visual patterns critical for accurate classification (Chen et al., 2018).

In contrast, no consistent relationship was observed between OS and processing time, which appeared to be more strongly affected by other hyperparameters like the number of RAEs and projections. Overall, OS is a secondary factor in the RADAM-based feature extraction pipeline, with model architecture and RAE configuration playing a more dominant role.

4.4. Influence of RAE quantity and projections on classification performance

Although Scabini et al. (2023) highlighted potential limitations in using a single-neuron RAE to encode sufficient information from activation maps (recommending ensembles of multiple RAEs to improve performance), our findings show that using only one RAE and one projection per RAE did not hinder classification quality, but did impact stability.

This configuration with a single RAE yielded high average F1-scores, comparable to more complex setups especially in terms of peak performance. However, consistent with previous findings, we observed significant gains in processing efficiency and classification stability when increasing the number of RAEs (Scabini et al., 2023). Notably, while higher numbers of RAEs generally improved consistency, the greatest reduction in feature extraction time occurred when moving from one to two RAEs. This configuration (two RAEs with one projection each) delivered an optimal balance between accuracy and speed across different backbone models and output stride settings.

Interestingly, increasing the number of projections per RAE did not lead to proportional improvements in classification performance. This observation supports the architectural choice of employing an ensemble of encoders, suggesting that model diversity – rather than merely increasing projection depth within a single encoder - plays a more critical role in enhancing the expressiveness and generalizability of the learned texture representations (Scabini et al., 2023; Wortsman et al., 2022).

Each additional projection linearly expands the output vector size, as every projection yields a feature vector whose length corresponds to the sum of activation map channel dimensions. Thus, it was initially hypothesized that more projections would extract complementary information from the aggregated activation maps, potentially enhancing

discriminative power. However, these results suggest that this additional information was either redundant or not sufficiently informative to justify the increased dimensionality.

This finding aligns with the notion that more features do not necessarily equate to better performance, especially in scenarios where the underlying signal is already well captured with fewer dimensions (Altman & Krzywinski, 2018; Parastar & Weller, 2025). In practice, this means that increasing projections might only inflate computational cost without yielding tangible benefits in classification accuracy (Jia et al., 2022). For the datasets evaluated, simpler configurations with fewer projections not only performed competitively but were also more computationally efficient, reinforcing the importance of parsimony in feature extraction strategies.

4.5. Texture features for classification

The application of the RADAM approach has proven to be a highly effective method for extracting features from meat images to assess freshness in both bovine and chicken samples. This effectiveness can be largely attributed to the ConvNeXt and ResNet architectures employed, as well as to their robust pre-training on the large-scale ImageNet dataset. The RADAM method also contributed to the strong results, outperforming other feature extraction methods such as the use of the GAP layer output, as employed by Büyükarıkan (2024), in side-by-side tests (Scabini et al., 2023).

Unlike other methods that use white-box models to assess meat quality, such as those described by Sánchez et al. (2023) and (Liu et al., 2016), the approach adopted here does not provide straightforward insights into which specific features are being used to classify meat freshness or how they contribute to this classification. For each image, RADAM generates a feature vector whose dimensionality corresponds to the number of channels in the activation maps selected during the extraction process - values that are inherently abstract due to the nature of deep representations. Each activation map undergoes encoding via a randomized autoencoder (RAE), and the resulting projections are aggregated (e.g., via mean or sum) to form a compact representation. This vector is then passed to a traditional classifier (here, SVM, RF or MLP) to perform the classification. While it is technically possible to trace back the specific information used for classification, doing so is extremely complex and resource-intensive. Apart from this limitation, the approach remains highly practical, delivering satisfactory results for meat freshness assessment.

To further explore the interpretability of the feature representations generated by RADAM, one promising avenue is to investigate correlations between its extracted features and those obtained from established meat texture assessment methods. For instance, mechanical methods such as Texture Profile Analysis (TPA) and the Warner-Bratzler shear test are widely recognized for quantifying meat texture attributes like hardness and chewiness (Ruiz De Huidobro et al., 2005; Schreuders et al., 2021). Spectroscopic techniques, including Near-Infrared Spectroscopy (NIRS), offer non-destructive means to assess meat composition and tenderness (Schreuders et al., 2021; Wu et al., 2022). Exploring the relationships between RADAM's deep feature vectors and the outputs of these conventional methods could provide valuable insights into the biological relevance and interpretability of the learned representations.

Additional meat quality measurements can be integrated with the features provided by RADAM to enable a more holistic classification framework. For example, Huang et al. (2023) developed a flexible sensor capturing temperature and electrical impedance from lamb meat, and He et al. (2025) combined this with optical sensing. Using a 1D-CNN-BiLSTM-ATT model, they achieved 94.57 % accuracy in freshness classification. Integrating such multimodal data with RADAM features could improve classification robustness and generalization.

4.6. Challenges and future directions

Computer vision emerges as a promising alternative for large-scale monitoring of meat quality traits in industrial settings. However, it is not intended to completely replace conventional analytical methods. Rather, its integration with traditional and more recent advanced techniques through data fusion offers a complementary and potentially more robust approach. Still, these techniques remain essential for generating accurately labeled datasets.

In this context, foodomics offers a powerful framework for molecular-level characterization of food matrices. By integrating multiple omics platforms, it enables the precise identification of bioactive compounds and quality markers, while also providing insights into their functional roles in health, nutrition, and spoilage (Cajka, 2024; Shi et al., 2024). These capabilities are critical for improving traceability, ensuring safety, and supporting a deeper understanding of the biochemical determinants of food quality (Shen, Ge, et al., 2024; Xue et al., 2025).

The integration of heterogeneous data sources, such as omics data and computer vision, remains an emerging and complex research frontier (Tang et al., 2023). Data fusion strategies can be implemented at different levels: early fusion (e.g., concatenation of raw image features and omics data), intermediate fusion (joint dimensionality reduction or latent space alignment), and late fusion (combination of predictions from separate models) (Borràs et al., 2015) Recent studies have explored various fusion techniques to integrate omics and imaging data, highlighting the potential of such integrative approaches (Borràs et al., 2015; Gu et al., 2025; Shen, Wang, et al., 2024). Nonetheless, standardized pipelines for fusing omics and imaging data in food science are still under development, underscoring the need for continued research in this area.

The image acquisition process itself warrants particular attention, as it plays a critical role in the overall reliability of computer vision-based food quality assessments. Factors such as controlled illumination, precise camera positioning, and careful sample preparation are critical because they directly influence the uniformity of color representation, contrast, and spatial resolution of the images (Mahajan et al., 2015; Wu & Sun, 2013). Variations in these parameters can introduce noise, distort feature extraction, and compromise the reliability of classification models. Additionally, fat content is known to affect the visual appearance of meat (de Fernandes et al., 2019) and may influence both freshness perception and model performance. This aspect was not considered in the present study and future work should evaluate model performance on datasets stratified by fat content to explore its potential impact.

5. Conclusions

Our findings highlight the potential of integrating pre-trained Deep Convolutional Neural Networks (DCNNs) with the Random Encoding of Aggregated Deep Activation Maps (RADAM) method for classifying meat freshness. By applying RADAM to encode activation maps produced throughout the architecture of the DCNN and feeding the resulting feature vectors into traditional machine learning classifiers, we achieved state-of-the-art performance across two of beef and one of chicken image datasets. This approach reduced reliance on the need for training or fine-tuning deep networks, yet still outperformed previously reported methods in the literature. Notably, it enabled near-perfect classification in two out of the three datasets, a result not documented before. These outcomes underscore the feasibility of using lightweight, transferable deep representations combined with conventional ML techniques as a practical and computationally efficient solution for meat freshness assessment.

CRediT authorship contribution statement

Marcelo M. Hidalgo: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Robson C. Lima:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Elisabete A. De Nadai Fernandes:** Writing – review & editing, Supervision, Resources, Project administration. **Márcio A. Bacchi:** Writing – review & editing. **Gabriel A. Sarriés:** Validation, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT to assist with localization during translation. After utilizing this tool/service, the author(s) reviewed and edited the content as needed and take (s) full responsibility for the final published content.

Funding

This work was supported by The São Paulo Research Foundation (FAPESP, grant 22/12732–5) and National Council for Scientific and Technological Development (CNPq, grant 406637/2022–9).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2025.146430.

Data availability

All data analyzed here can be found on the mentioned articles, i.e. Ulucan et al. (2019), Büyükarıkan (2024) and Calvin et al. (2020).

References

Altman, N., & Krzywinski, M. (2018). The curse(s) of dimensionality. *Nature Methods, 15*(6), 399–400. https://doi.org/10.1038/S41592-018-0019-X;KWRD=LIFE+SCIENCES

Amani, M. A., & Sarkodie, S. A. (2022). Mitigating spread of contamination in meat supply chain management using deep learning. *Scientific Reports, 12*(1), 1–10. https://doi.org/10.1038/s41598-022-08993-5

Bekhit, A. E. D. A., Holman, B. W. B., Giteru, S. G., & Hopkins, D. L. (2021). Total volatile basic nitrogen (TVB-N) and its role in meat spoilage: A review. *Trends in Food Science & Technology, 109*, 280–302. https://doi.org/10.1016/J.TIFS.2021.01.006

Bevilacqua, M., Bro, R., Marini, F., Rinnan, Å., Rasmussen, M. A., & Skov, T. (2017). Recent chemometrics advances for foodomics. *Trends in Analytical Chemistry, 96*, 42–51. https://doi.org/10.1016/J.TRAC.2017.08.011

Borràs, E., Ferré, J., Boqué, R., Mestres, M., Aceña, L., & Busto, O. (2015). Data fusion methodologies for food and beverage authentication and quality assessment – A review. *Analytica Chimica Acta, 891*, 1–14. https://doi.org/10.1016/J.ACA.2015.04.042

Büyükarıkan, B. (2024). ConvColor DL: Concatenated convolutional and handcrafted color features fusion for beef quality identification. *Food Chemistry, 460*, Article 140795. https://doi.org/10.1016/j.foodchem.2024.140795

Cai, M., Li, X., Liang, J., Liao, M., & Han, Y. (2024). An effective deep learning fusion method for predicting the TVB-N and TVC contents of chicken breasts using dual hyperspectral imaging systems. *Food Chemistry, 456*, Article 139847. https://doi.org/10.1016/J.FOODCHEM.2024.139847

Cajka, T. (2024). Liquid chromatography–mass spectrometry–based metabolomics approaches for foodomics research. *Current Opinion in Food Science, 58*, Article 101201. https://doi.org/10.1016/J.COFS.2024.101201

Calvin, P. G. B., & Prakasa, E. (2020). *Classification of Chicken Meat Freshness using Convolutional Neural Network Algorithms. 2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies, 3ICT 2020*. https://doi.org/10.1109/3ICT51146.2020.9312018

Chen, J., Zhang, J., Wang, N., Xiao, B., Sun, X., Li, J., Zhong, K., Yang, L., Pang, X., Huang, F., & Chen, A. (2024). Critical review and recent advances of emerging real-time and non-destructive strategies for meat spoilage monitoring. *Food Chemistry, 445*, Article 138755. https://doi.org/10.1016/J.FOODCHEM.2024.138755

Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). *Encoder-decoder with Atrous separable convolution for semantic image segmentation. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), 11211 LNCS, 833–851*. https://doi.org/10.1007/978-3-030-01234-2_49

de Fernandes, D. D. S., Romeo, F., Krepper, G., Di Nezio, M. S., Pistonesi, M. F., Centurión, M. E., … Diniz, P. H. G. D. (2019). Quantification and identification of adulteration in the fat content of chicken hamburgers using digital images and chemometric tools. *LWT, 100*, 20–27. https://doi.org/10.1016/j.lwt.2018.10.034

Filus, K., & Domańska, J. (2023). Global entropy pooling layer for convolutional neural networks. *Neurocomputing, 555*, Article 126615. https://doi.org/10.1016/J.NEUCOM.2023.126615

Gu, Z., Xu, Q., Wang, X., Lin, X., Duan, N., Wang, Z., & Wu, S. (2025). Food freshness prediction platform utilizing deep learning-based multimodal sensor fusion of volatile organic compounds and moisture distribution. *ACS Sensors.*. https://doi.org/10.1021/acssensors.5c00254

He, W., Huang, W., Wang, Y., Li, Z., Blanka, T., & Zhang, X. (2025). A lamb freshness detection model using a flexible optoelectronic in-situ sensing system and multi-input multi-label causal ensemble learning. *Food Chemistry, 471*, Article 142803. https://doi.org/10.1016/J.FOODCHEM.2025.142803

Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing, 70*(1–3), 489–501. https://doi.org/10.1016/J.NEUCOM.2005.12.126

Huang, W., Xia, J., Wang, X., Zhao, Q., Zhang, M., & Zhang, X. (2023). Improvement of non-destructive detection of lamb freshness based on dual-parameter flexible temperature-impedance sensor. *Food Control, 153*, Article 109963. https://doi.org/10.1016/J.FOODCONT.2023.109963

Jeevan, P., & Sethi, A. (2024). Which Backbone to Use: A Resource-efficient Domain Specific Comparison for Computer Vision. https://arxiv.org/pdf/2406.05612.

Jia, W., Sun, M., Lian, J., & Hou, S. (2022). Feature dimensionality reduction: A review. *Complex & Intelligent Systems 2022 8:3, 8*(3), 2663–2693. https://doi.org/10.1007/S40747-021-00637-X

Kasun, L. L. C., Zhou, H., Huang, G. B., & Vong, C. M. (2013). Representational learning with ELMs for big data. *IEEE Intelligent Systems, 28*(6), 31–34.

Kaushal, S., Tammineni, D. K., Rana, P., Sharma, M., Sridhar, K., & Chen, H. H. (2024). Computer vision and deep learning-based approaches for detection of food nutrients/nutrition: New insights and advances. *Trends in Food Science & Technology, 146*, Article 104408. https://doi.org/10.1016/J.TIFS.2024.104408

Lee, Y. S., Kim, T., Choi, S., & Kim, W. (2022). When does AI pay off? AI-adoption intensity, complementary investments, and R&D strategy. *Technovation, 118*, Article 102590. https://doi.org/10.1016/J.TECHNOVATION.2022.102590

Liao, Q., Gardner, B., Barlow, R., McMillan, K., Moore, S., Fitzgerald, A., … Nelis, J. L. (2025). Improving traceability and quality control in the red-meat industry through computer vision-driven physical meat feature tracking. *Food Chemistry, 480*, Article 143830. https://doi.org/10.1016/J.FOODCHEM.2025.143830

Liu, J., Cao, Y., Wang, Q., Pan, W., Ma, F., Liu, C., Chen, W., Yang, J., & Zheng, L. (2016). Rapid and non-destructive identification of water-injected beef samples using multispectral imaging analysis. *Food Chemistry, 190*, 938–943. https://doi.org/10.1016/J.FOODCHEM.2015.06.056

Liu, Z., Mao, H., Wu, C. Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). *A ConvNet for the 2020s. Proceedings of the IEEE computer society conference on computer vision and pattern recognition, 2022-June, 11966–11976*. https://doi.org/10.1109/CVPR52688.2022.01167

Luo, N., Xu, D., Xing, B., Yang, X., & Sun, C. (2024). Principles and applications of convolutional neural network for spectral analysis in food quality evaluation: A review. *Journal of Food Composition and Analysis, 128*, Article 105996. https://doi.org/10.1016/J.JFCA.2024.105996

Mahajan, S., Das, A., & Sardana, H. K. (2015). Image acquisition techniques for assessment of legume quality. *Trends in Food Science & Technology, 42*(2), 116–133. https://doi.org/10.1016/J.TIFS.2015.01.001

Modzelewska-Kapituła, M., & Jun, S. (2022). The application of computer vision systems in meat science and industry – A review. *Meat Science, 192*, Article 108904. https://doi.org/10.1016/J.MEATSCI.2022.108904

Mustapha, M. T., Ozsahin, I., & Ozsahin, D. U. (2024). Convolution neural network and deep learning. In *Artificial intelligence and image processing in medical imaging* (pp. 21–50). Academic Press. https://doi.org/10.1016/B978-0-323-95462-4.00002-9.

OECD/FAO. (2022). OECD-FAO agricultural outlook 2022-2031. *OECD Publishing.*. https://doi.org/10.1787/f1b0b29c-en

Parastar, H., & Weller, P. (2025). Feature selection and extraction strategies for non-targeted analysis using GC-MS and GC-IMS: A tutorial. *Analytica Chimica Acta, 1343*, Article 343635. https://doi.org/10.1016/J.ACA.2025.343635

Putri, L. A., Rahman, I., Puspita, M., Hidayat, S. N., Dharmawan, A. B., Rianjanu, A., … Wasisto, H. S. (2023). Rapid analysis of meat floss origin using a supervised machine learning-based electronic nose towards food authentication. *npj Science of Food, 7*(1), 1–15. https://doi.org/10.1038/s41538-023-00205-2

Rex, E. S., Annrose, C. R. J., & Jenifer Jose, J. (2022). Comparative analysis of deep convolution neural network models on small scale datasets. *Optik, 271*, Article 170238. https://doi.org/10.1016/J.IJLEO.2022.170238

Ruiz De Huidobro, F., Miguel, E., Blázquez, B., & Onega, E. (2005). A comparison between two methods (Warner–Bratzler and texture profile analysis) for testing either raw meat or cooked meat. *Meat Science, 69*(3), 527–536. https://doi.org/10.1016/j.meatsci.2004.09.008

Rukundo, O. (2023). Effects of image size on deep learning. *Electronics, 12, Page 985, 12 (4)*, 985. https://doi.org/10.3390/electronics12040985

Saha, D., Padhiary, M., & Chandrakar, N. (2025). AI vision and machine learning for enhanced automation in food industry: A systematic review. *Food and Humanity, 4*, Article 100587. https://doi.org/10.1016/j.foohum.2025.100587

Sánchez, C. N., Orvañanos-Guerrero, M. T., Domínguez-Soberanes, J., & Álvarez-Cisneros, Y. M. (2023). Analysis of beef quality according to color changes using computer vision and white-box machine learning techniques. *Heliyon, 9*(7), Article e17976. https://doi.org/10.1016/j.heliyon.2023.e17976

Scabini, L., Zielinski, K. M., Ribas, L. C., Gonçalves, W. N., De Baets, B., & Bruno, O. M. (2023). RADAM: Texture recognition through randomized aggregated encoding of deep activation maps. *Pattern Recognition, 143*, Article 109802. https://doi.org/10.1016/j.patcog.2023.109802

Schreuders, F. K. G., Schlangen, M., Kyriakopoulou, K., Boom, R. M., & van der Goot, A. J. (2021). Texture methods for evaluating meat and meat analogue structures: A review. *Food Control, 127*, Article 108103. https://doi.org/10.1016/j.foodcont.2021.108103

Shao, L., Chen, S., Wang, H., Zhang, J., Xu, X., & Wang, H. (2021). Advances in understanding the predominance, phenotypes, and mechanisms of bacteria related to meat spoilage. *Trends in Food Science & Technology, 118*, 822–832. https://doi.org/10.1016/J.TIFS.2021.11.007

Shen, Q., Ge, L., Lu, W., Wu, H., Zhang, L., Xu, J., … Cheng, K. (2024). Transplanting network pharmacology technology into food science research: A comprehensive review on uncovering food-sourced functional factors and their health benefits. *Comprehensive Reviews in Food Science and Food Safety, 23*(5), 1–50. https://doi.org/10.1111/1541-4337.13429

Shen, Q., Wang, S., Wang, H., Liang, J., Zhao, Q., Cheng, K., Imran, M., Xue, J., & Mao, Z. (2024). Revolutionizing food science with mass spectrometry imaging: A comprehensive review of applications and challenges. *Comprehensive Reviews in Food Science and Food Safety, 23*(4), Article e13398. https://doi.org/10.1111/1541-4337.13398

Shi, J., Liu, Y., & Xu, Y. J. (2024). MS based foodomics: An edge tool integrated metabolomics and proteomics for food science. *Food Chemistry, 446*, Article 138852. https://doi.org/10.1016/J.FOODCHEM.2024.138852

Tang, Q., Liang, J., & Zhu, F. (2023). A comparative review on multi-modal sensors fusion based on deep learning. *Signal Processing, 213*, Article 109165. https://doi.org/10.1016/J.SIGPRO.2023.109165

Ulucan, O., Karakaya, D., & Turkan, M. (2019). *Meat quality assessment based on deep learning. In 2019 innovations in intelligent systems and applications conference (ASYU). 1–5. IEEE*. https://doi.org/10.1109/ASYU48272.2019.8946388

Valdés, A., Álvarez-Rivera, G., Socas-Rodríguez, B., Herrero, M., Ibáñez, E., & Cifuentes, A. (2022). Foodomics: Analytical opportunities and challenges. *Analytical Chemistry, 94*(1), 366–381. https://doi.org/10.1021/acs.analchem.1c04678

Wojnowski, W., Majchrzak, T., Dymerski, T., Gębicki, J., & Namieśnik, J. (2017). Electronic noses: Powerful tools in meat quality assessment. *Meat Science, 131*, 119–131. https://doi.org/10.1016/J.MEATSCI.2017.04.240

Wortsman, M., Ilharco, G., Gadre, S. Y., Roelofs, R., Gontijo-Lopes, R., Morcos, A. S., Namkoong, H., Farhadi, A., Carmon, Y., Kornblith, S., & Schmidt, L. (2022). Model soups: Averaging weights of multiple fine-tuned models improves accuracy without increasing inference time (pp. 23965–23998). PMLR. https://proceedings.mlr.press/v162/wortsman22a.html.

Wu, D., & Sun, D. W. (2013). Colour measurements by computer vision for food quality control – A review. *Trends in Food Science & Technology, 29*(1), 5–20. https://doi.org/10.1016/J.TIFS.2012.08.004

Wu, X., Liang, X., Wang, Y., Wu, B., & Sun, J. (2022). *Non-destructive techniques for the analysis and evaluation of meat quality and safety: A review. Foods 2022, Vol. 11, page 3713, 11(22), 3713*. https://doi.org/10.3390/FOODS11223713

Xue, J., Wu, H., Ge, L., Lu, W., Wang, H., Mao, P., Liao, J., Zeng, X., Wang, S., Jiang, L., Liang, J., Huang, J., Wang, Z., & Shen, Q. (2025). Towards a safer food chain: Recent advances in multi-technology based lipidomics application to food quality and safety. *Trends in Food Science & Technology, 156*, Article 104859. https://doi.org/10.1016/J.TIFS.2024.104859

Zhao, Z., Wang, R., Liu, M., Bai, L., & Sun, Y. (2025). Application of machine vision in food computing: A review. *Food Chemistry, 463*, Article 141238. https://doi.org/10.1016/J.FOODCHEM.2024.141238

Zhu, Y., Gu, M., Su, Y., Li, Z., Xiao, Z., Lu, F., & Han, C. (2024). Recent advances in spoilage mechanisms and preservation technologies in beef quality: A review. *Meat Science, 213*, Article 109481. https://doi.org/10.1016/J.MEATSCI.2024.109481