

Article

Quality Grading of Dried Abalone Using an Optimized VGGNet

Yansong Zhong ¹, Hongyue Lin ¹, Jiacheng Gan ², Weiwei You ^{2,3}, Jia Chen ^{1,4} and Rongxin Zhang ^{5,*}

¹ School of Electronic Science and Engineering (National Model Microelectronics College), Xiamen University, Xiamen 361005, China; 33320211150274@stu.xmu.edu.cn (Y.Z.); linhongyue20@stu.xmu.edu.cn (H.L.); chenj@xmu.edu.cn (J.C.)

² College of Ocean and Earth Sciences, Xiamen University, Xiamen 361005, China; jiachenggan@stu.xmu.edu.cn (J.G.); wwyou@xmu.edu.cn (W.Y.)

³ Fujian Key Laboratory of Genetics and Breeding of Marine Organisms, Xiamen University, Xiamen 361102, China

⁴ Innovation Laboratory for Sciences and Technologies of Energy Materials of Fujian Province (IKKEM), Xiamen University, Xiamen 361005, China

⁵ Key Laboratory of Underwater Acoustic Communication and Marine Information, Xiamen University, Xiamen 361005, China

* Correspondence: zhrx@xmu.edu.cn

Abstract: As living standards have improved, consumer demand for high-quality dried abalone has increased. Traditional abalone grading is achieved through slice analysis (sampling analysis) combined with human experience. However, this method has several issues, including non-uniform grading standards, low detection accuracy, inconsistency between internal and external quality, and high loss rate. Therefore, we propose a deep-learning-aided approach leveraging X-ray images that can achieve efficient and non-destructive internal quality grading of dried abalone. To the best of our knowledge, this is the first work to use X-ray to image the internal structure of dried abalone. The work was divided into three phases. First, a database of X-ray images of dried abalone was constructed, containing 644 samples, and the relationship between the X-ray images and the internal quality of the dried abalone was analyzed. Second, the database was augmented by image rotation, image mirroring, and noise superposition. Subsequently, a model selection evaluation process was carried out. The evaluation results showed that, in a comparison with models such as VGG-16, MobileNet (Version 1.0), AlexNet, and Xception, VGG-19 demonstrated the best performance in the quality grading of dried abalone. Finally, a modified VGG-19 network based on the CBAM was proposed to classify the quality of dried abalone. The results show that the proposed quality grading method for dried abalone was effective, achieving a score of 95.14%, and outperformed the competitors, i.e., VGG-19 alone and VGG-19 with the squeeze-and-excitation block (SE) attention mechanism.

Keywords: dried abalone; internal quality grading; deep learning; attention mechanism; X-ray image



Citation: Zhong, Y.; Lin, H.; Gan, J.; You, W.; Chen, J.; Zhang, R. Quality Grading of Dried Abalone Using an Optimized VGGNet. *Appl. Sci.* **2024**, *14*, 5894. <https://doi.org/10.3390/app14135894>

Academic Editors: Natalia Casado Navas, Isabel Sierra Alonso, Sonia Morante Zarcero and Roberto Romaniello

Received: 15 April 2024

Revised: 14 June 2024

Accepted: 2 July 2024

Published: 5 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As one of the most popular seafood products in East Asia, abalone has high nutritional and medicinal value. However, it is susceptible to bacterial contamination and spoilage in the air [1]. By drying abalone, most of the water can be removed, affecting the environment required by bacteria and prolonging the expiration date. Dried abalone has a unique and delicious flavor. Consequently, dried abalone products have become popular and have significant economic value. The quality of dried abalone varies due to differences in the product processing, storage, and transportation conditions. Poor-quality dried abalone may be a danger to human health. To meet the requirements of different market demands and food safety, it is necessary to be able to determine the quality of dried abalone.

Both the external quality (or appearance quality) and the internal quality of dried abalone are of importance. The external quality can be judged manually or automatically

by optical image processing. At present, the research focused on internal quality has mainly relied on texture characteristics analysis [2–4] and nutrient composition analysis [5–7]. These methods require equipment such as texture analyzers and microscopes for quality detection, which are destructive to dried abalone and are not suitable for quality classification during production. During production, the quality of dried abalone is assessed using the light transmission method. If there are silhouetted internal cavities, the dried abalone is considered poor quality and will be screened out. However, there is a high rate of missed detection, and the internal quality of dried abalone without silhouetted internal cavities cannot be determined. The production location and processing year of dried abalone are also evaluation indicators of the internal quality of abalone, but these evaluation methods are too subjective. In order to meet the requirements of the rapid expansion of the dried abalone industry, we must urgently develop a more effective and nondestructive method for evaluating the internal quality of dried abalone.

X-ray is an electromagnetic wave with an extremely short wavelength and high frequency, which allows nondestructive and rapid imaging [8]. Since Roentgen captured the first X-ray images of human hands, X-ray technology has been applied in a wide range of fields, including materials science [9], medicine [10], Earth science [11], and more. X-ray has different absorption rates for materials with different densities, and this difference in absorption rate is reflected in the gray value of X-ray images, which can distinguish bones, soft tissues, and cavities with different densities. In addition, well-designed X-ray systems have been widely used for the identification and classification of various foods [12–14]. Due to the lack of a direct correlation between the internal quality and the external appearance of dried abalone, it is necessary to assess its internal structure. Traditional methods for analyzing the internal quality of dried abalone rely on histological analysis, combined with experiential judgment. This method not only has a low accuracy rate in identification, it also causes damage to the dried abalone samples. On this basis, we considered using X-ray technology to determine the internal quality of dried abalone. By utilizing X-ray technology, we were able to quickly conduct non-destructive testing on abalone. However, due to the complexity and diversity of information reflected in X-ray images of dried abalone, manual processing is challenging. Therefore, a method was required to extract and classify the features of the X-ray images of dried abalone.

Deep learning performs well in image classification tasks. In the ImageNet Large-scale Visual Recognition Challenge (ILSVRC), a convolutional neural network, VGGNet [15], was proposed for image classification tasks. At the last ILSVRC in 2017, the winner, SENet [16], used a squeeze-and-excitation (SE) module to help the network focus on the important features. Networks with attention mechanisms, such as SE modules and the convolutional block attention module (CBAM) [17], can increase attention to channel characteristics. Due to its excellent performance, deep learning has been used for agricultural product category classification and food rating. For instance, the improved InceptionV3 model and Transfer Learning method enhanced the accuracy of soybean seed classification [18]. The MobileNetV2 architecture boosted the precision of corn seed disease classification [19]. The TL-MobileNetV2 model achieved high-accuracy fruit image classification [20]. Therefore, we adopted deep learning methods to classify the X-ray images of dried abalone.

The main contributions of this research are threefold:

- (1) It is the first work to solve the problem of efficiently and non-destructively carrying out internal quality grading of dried abalone.
- (2) It is the first work to apply X-ray images to dried abalone detection.
- (3) It is the first work to introduce deep learning into the dried abalone industry.

In conclusion, in order to realize the non-destructive quality classification of dried abalone, we proposed an X-ray-aided approach using deep learning technology. Specifically, we validated the feasibility of using X-ray images for the quality classification of dried abalone, constructed a database of dried abalone X-ray images with samples of three quality types, and used VGG-19, VGG-19+SE, and VGG-19+CBAM for performance evaluation.

This work provides a new and feasible solution for the internal quality classification of dried abalone and might provide a guideline for market standardization.

2. Database Construction

2.1. X-ray Imaging of Dried Abalone

There is currently a lack of publicly available databases of X-ray images of dried abalone. Therefore, we constructed a database using an X-ray acquisition device. The experimental device is shown in Figure 1. The X-ray transmitter was the MH-214 dental X-ray emitter by Foshan Meihao Medical Device Co., Ltd. (Wenzhou, China), which operates at a tube voltage of 60 kV, a tube current of 1 mA, and an exposure time of 0.3 s. The sensor was the HDR-600A digital X-ray sensor, also manufactured by Foshan Meihao Medical Device Co., Ltd., with an effective detection area of $26 \times 36 \text{ mm}^2$. The software accompanying the system was the XVSensor v2.1.6.1 X-ray image processing software, currently under development by Shenzhen Keshi Technology Co., Ltd. (Beijing, China).



Figure 1. Database sample collection device.

The subsequent steps illustrate the process of X-ray image acquisition. First, a dried abalone sample was secured above the sensor. Subsequently, X-rays were emitted towards the sample through the X-ray transmitter. Following this, the signals sensed by the sensor were converted into X-ray images of the dried abalone using the accompanying software. Finally, by repeating these operations, a database of X-ray images was constructed for dried abalones.

The dried abalone samples for database construction were provided by Xiamen Qicheng Marine Technology Co., Ltd. (Xiamen, China). A total of 644 samples were provided, including 103 and 263 samples produced in Japan in 2019 and 2020, respectively, and 278 samples produced in China in 2020. All samples were processed from hybrid abalone (female *Haliotis discus hannai* crossed with male *Haliotis fulgens*), and the net weight of each collection was 40–50 g. In our database, dried abalone with lower quality were randomly distributed across all samples, with no direct correlation to specific batches.

2.2. X-ray Image Analysis of Dried Abalone

By analyzing the X-ray images of the dried abalone, we were able to roughly divide the images into three categories (Type A, Type B, and Type C), as shown in Figure 2, where Figure 2b is a zoomed-in view. For Type A X-ray images, the pixel values exhibited a more uniform distribution. For Type B, the distribution was less uniform than that of Type A due to noticeable white spots in the X-ray images. Type C X-ray images featured a prominent dark gray area in the center.

The aforementioned analysis had a degree of subjectivity. Therefore, to mitigate the influence of human factors, we computed histograms of the three types of X-ray images, as illustrated in Figure 3. For the Type A X-ray images, the pixel distribution was relatively uniform; the pixel value mainly ranged from 80 to 170. Type B images exhibited a concentrated pixel distribution in the central area, indicating over-brightness corresponding to white spots in the X-ray, as shown in Figure 2. Conversely, Type C X-ray images had a more concentrated distribution of pixels in the low-value region, indicative of the existence of excessively dark areas that corresponded to dark gray regions in the X-ray, as shown in Figure 2.

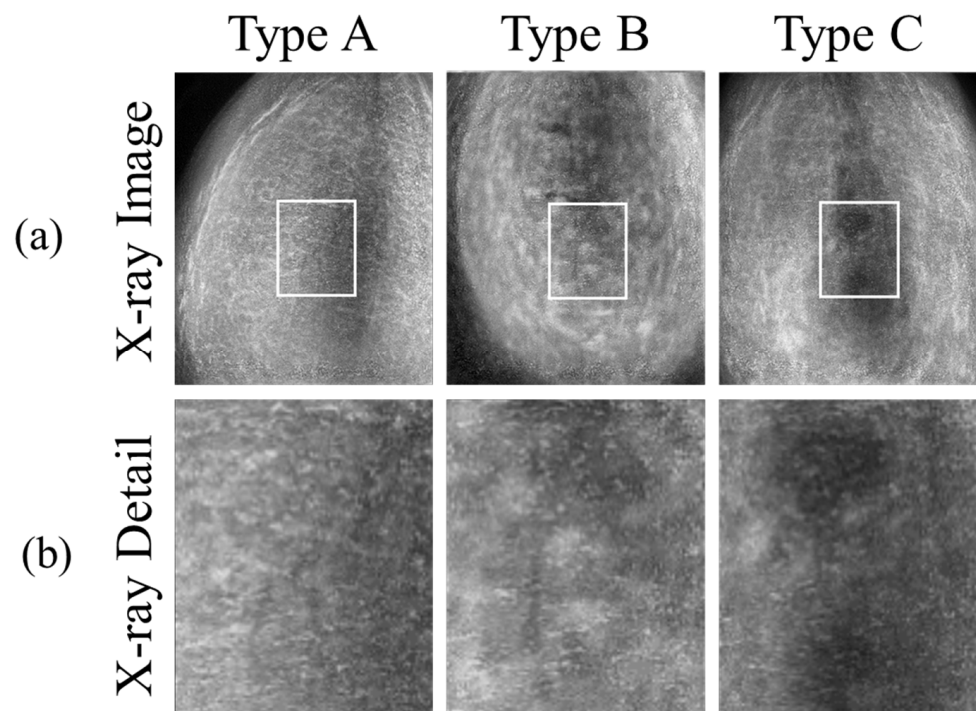


Figure 2. X-ray images and internal structures of different qualities of dried abalone samples (a) X-ray images; (b) X-ray image details.

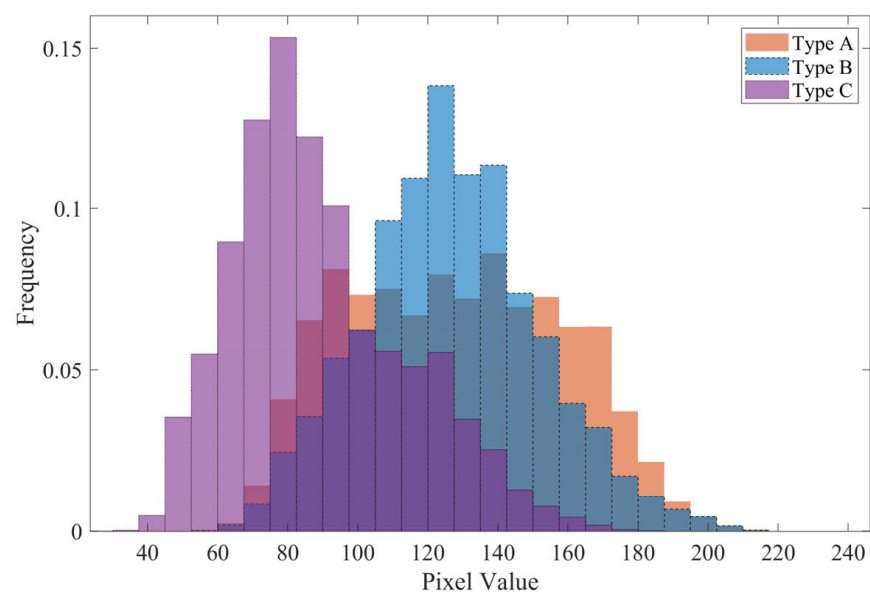


Figure 3. Pixel-value histogram for X-ray images of dried abalone.

To investigate the relationship between the X-ray images of dried abalone and their actual internal qualities, we dissected the three types of dried abalone samples along the side. Due to the high density and hard texture of dried abalone, we used a toothless cutting saw blade (thickness of 5 mm, diameter of 12 inches) to cut the dried abalone. The anatomical images are shown in Figure 4.

Through dissection, we found that the internal color of Type A dried abalone was uniform, the light transmission was good, and the number of white pitted spots were fewer. Type B displayed an uneven internal color, poor light transmission, and many white pitted

areas. Type C muscle was different from normal muscle tissue; it had many cavities and cotton-like areas, and there was muscle tissue corruption in the cavity area.

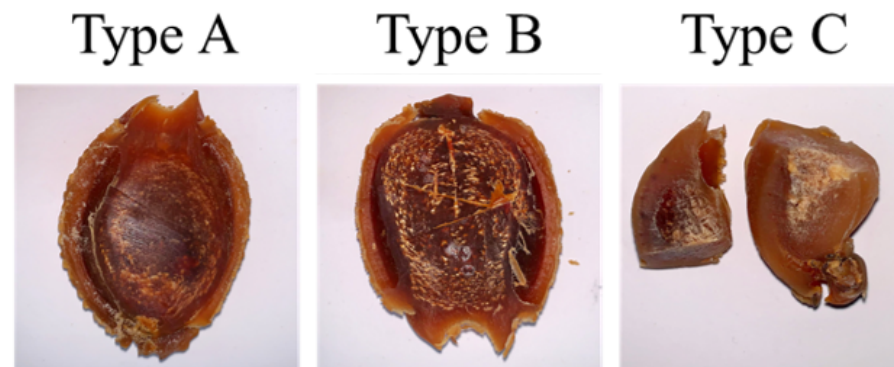


Figure 4. Internal anatomical diagrams of the three types of dried abalone.

2.3. Microscopic Analysis of Quality Differences between Dried Abalone

The cavity area of Type C dried abalone was mostly caused by deterioration due to improper storage, leading to defective products. The quality of Type A and Type B dried abalone cannot be distinguished by technical means in the industry, and both are considered qualified products. To further understand the microscopic differences in quality between Type A and Type B, the dried abalone samples were scanned using an electron microscope (Phenom Pro X desktop scanning electron microscope from Phenom Scientific). Figure 5 depicts the shell muscle and foot muscle of the dried abalone, as observed under the electron microscope. According to Gao [21], the electron scanning microscope method can be modified for dried abalone. The sample size was $3 \times 3 \times 3 \text{ mm}^3$ and the acceleration voltage was 10 kV.

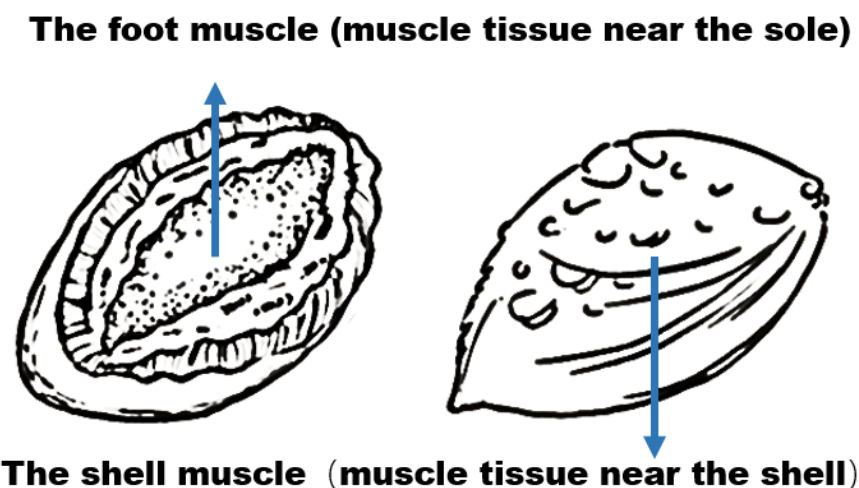


Figure 5. Diagram of abalone foot muscle and shell muscle.

The muscle tissue structure of abalone primarily comprises collagen fibers and myofibrils, interwoven to create a dense network structure [22]. Figure 6 depicts electron microscope images of different muscles. In the Type A dried abalone, the myofibrils were slender and exhibited an intact and dense structure, with no apparent breakages. Conversely, Type B dried abalone fibers appeared relatively loose and disordered, with a larger space between the fibers. In general, the collagen content was positively correlated with the hardness, chewability, and resilience of the abalone [23]. A denser myofibrillar fiber structure corresponded to a lower collagen content, indicating a more favorable taste and internal quality of abalone.

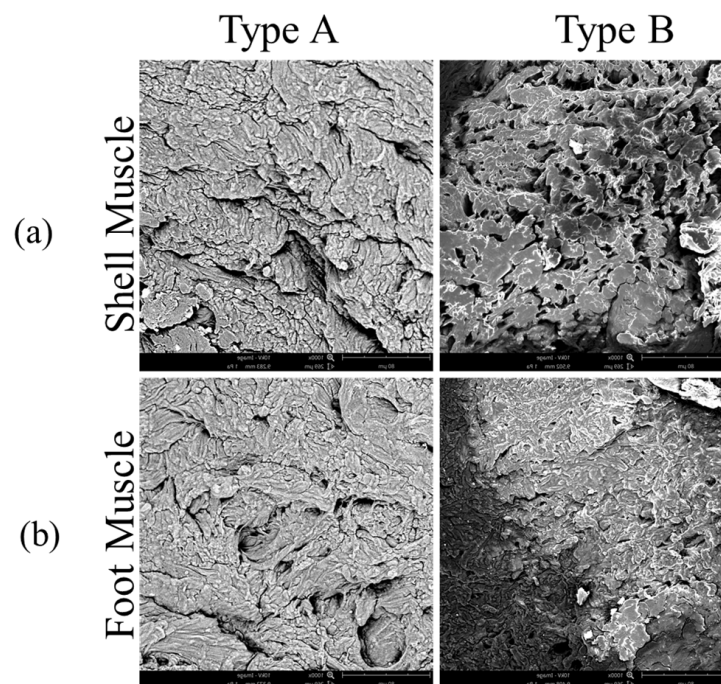


Figure 6. Micrographic views of dried abalone with different qualities. (a) Shell muscle; (b) foot muscle.

2.4. Construction of a Database for Dried Abalone

Anatomical and electron microscope observations of dried abalone samples showed that the quality of Type A dried abalone was good, the quality of Type B dried abalone was average, and the quality of Type C was poor. Therefore, the quality of dried abalone can be graded using X-ray gray images.

Incorporating the variations between the aforementioned dried abalone of various qualities, the gathered X-ray images were divided according to their qualities, as shown in Figure 7. Three grades, “good”, “average”, and “poor” were determined based on this. The dimensions of each image were $344 \times 417 \times 1$, where 344 and 417 were the length and width of the picture, and 1 was the number of channels of the picture. The original number of images with “good”, “average”, and “poor” qualities were 303, 268, and 73, respectively.

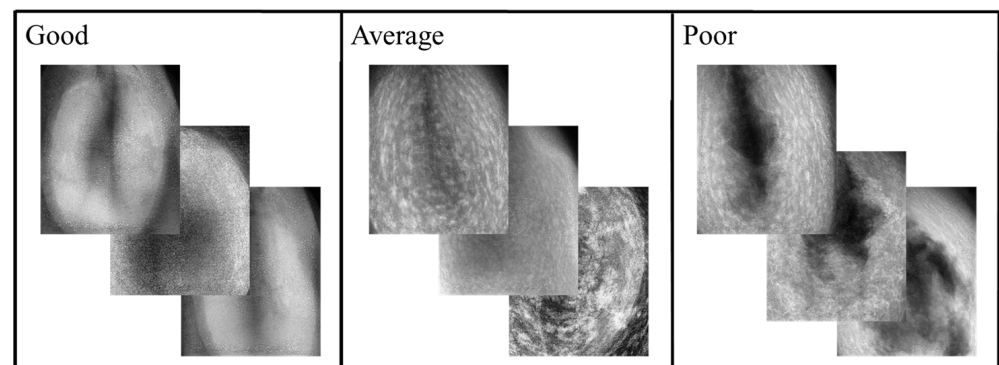


Figure 7. Samples of dried abalone with different internal quality labels.

2.5. Database Augmentation

Due to the high price of dried abalone samples, our X-ray database was limited in size. To mitigate the risks of overfitting and diminished generalizability associated with limited sample sizes, the size of the database was expanded. In addition, since the “poor” quality

dried abalone in production is mostly screened out, the number of such samples in the database was small. Therefore, sample imbalance needed to be considered.

Data augmentation was performed on this database through random brightness changes, rotations, flips, and blurring. Brightness changes can simulate the appearance of the same sample under varying levels of X-ray exposure. Given the diversity of shooting angles in actual imaging, flips and rotations can simulate samples captured from different perspectives. High noise levels inherent in the X-ray imaging process degrade image quality below standard expectations, allowing random noise to emulate samples containing noise from the actual shooting process [24]. By employing these methods for sample expansion, we were able to enhance the model's generalization capabilities and alleviate the sample imbalance problem through data augmentation. After augmentation, the total number of samples increased to 5736. The quantities of the images labeled as "Good", "Average", and "Poor" were augmented to 2424, 2144, and 1168, respectively, which was 8 times, 8 times, and 16 times the size of the original categories, respectively.

3. Establishment of a Quality Grading Model for Dried Abalone

3.1. Model Selection

There are numerous image classification models; VGG-16, VGG-19, MobileNet, AlexNet, and Xception are some of the most well-known architectures in the field of deep learning. VGG-16 and VGG-19 were developed by the Visual Geometry Group at the University of Oxford; they consist of deep stacks of 3×3 convolutional layers, with VGG-19 being capable of learning more complex feature representations due to its greater number of layers [15]. MobileNet was designed as a lightweight network to address the issue of limited computational resources, significantly reducing the model's computational load and size through depthwise separable convolutions [25]. AlexNet, one of the early deep CNNs, introduced ReLU activation functions and dropout techniques, achieving breakthrough performance in the 2012 ImageNet competition [26]. Xception, meaning "extreme Inception", is a model developed by Google that employs depthwise separable convolutions and residual connections, aiming to balance performance and efficiency [27]. Table 1 presents the distinct characteristics of the aforementioned classification networks.

Table 1. Comparison of image classification networks.

Feature/ Network	VGG-19	VGG-16	MobileNet	AlexNet	Xception
Network Depth	19 layers	16 layers	Lightweight, fewer layers	5 conv + 3 fc layers	Deep separable conv, many layers
Computational Efficiency	Low	High	High, mobile-friendly	Medium	High, uses deep separable conv
Recognition Accuracy	High, complex tasks	High	Moderate, lightweight	Moderate, early model	High, balanced accuracy and efficiency
Applicable Scenarios	High-precision image classification	High-precision image classification	Mobile and embedded vision	Early application in image classification	High-resolution image processing

Given the need for high-precision classification in the grading of dried abalone's internal quality, this study used the VGG-19 network. Compared to other networks, its deeper architecture enables it to capture more nuanced features, thereby enhancing the accuracy of classification.

3.2. Model Overview

VGG-19 comprises 16 convolution layers and 3 fully connected layers, organized into 5 groups. The initial two groups each contain two convolution layers, while the last

three groups each consist of four convolution layers. Once an image is fed into VGG-19, the convolutional layer extracts features from the feature map, enhances specific features, and simultaneously reduces noise. The inserted maximum pooling layer between convolutional layers reduces data processing requirements, suppresses noise, and eliminates redundancy.

In the last fully connected layer, each neuron establishes a full connection with every neuron in the preceding layer, executing a weighted sum of input features to act as a classifier. Following model learning, the output value undergoes normalization via Softmax, yielding probabilities corresponding to each category. The classification result predicted by the network corresponds to the category with the maximum probability value.

However, deep learning networks, including VGG-19, face challenges such as slow training speed and difficulty in model optimization. This is because the distribution of inputs to each layer changes continuously during training. This phenomenon is known as “internal covariate shift”. This constant change forces the network to continuously adjust learning parameters, thereby slowing down the training speed. Introducing batch normalization (BN) into a deep learning network can provide an effective strategy to address these issues. BN mitigates internal covariate shifts by standardizing the data of each mini-batch, making the training process more stable and faster. The incorporation of BN layers makes the network less sensitive to initial weights, enhancing model stability. This implies a broader range of choices for initial parameter settings, reducing the risk of training failure. Since BN layers reduce the bias and variance of input values, higher learning rates can be used without causing model divergence, thus accelerating convergence. The standardization process in BN layers introduces a degree of noise, akin to a mild form of data augmentation, which can prevent overfitting to some extent and improve the generalization capability. In summary, integrating BN layers into the VGG-19 model not only speeds up the training process and enhances model stability and generalizability, but also allows for the use of higher learning rates, further improving model performance.

A proper attention mechanism can optimize the performance of networks, especially in terms of information processing effectiveness and precision [28]. Depending on the specific attention concerns, attention mechanisms are categorized into four types: spatial attention, temporal attention, channel attention, and branch attention. Two combinations of attention mechanisms—channel–spatial attention and spatial–temporal attention—were derived based on these types. The attention mechanism adopted in this study was the channel–spatial attention mechanism (CBAM) [17].

CBAM integrates the channel attention module (CAM) and the spatial attention module (SAM), allowing for a focused analysis of features in both the channel and spatial dimensions. The SAM module emphasizes the target’s location information and compresses the channel dimension, while the CAM module concentrates on the key details of the image, compressing the spatial dimension. Through global average pooling and global max pooling, CBAM is capable of capturing statistical information and salient differences among channels. Fully connected layers and activation functions further learn how to transform these statistical insights into weights for each channel, where these weights are learnable parameters.

Figure 8 illustrates the optimized structure of the VGG-19+CBAM network in this paper, with the dimensions of the feature map outputted by each layer marked below the layer’s name. The size of the feature map in the database was $344 \times 417 \times 1$, and the descriptions of the corresponding database operations are omitted here. The input feature map underwent several convolutional and maximum pooling layers, followed by processing through the CBAM module, three fully connected layers, before ultimately being fed into the Softmax for the final quality classification result. Note that the first convolutional layer, Conv1, resized the input image to a fixed size of $224 \times 224 \times 64$. All max-pooling layers had a stride of two, and all convolutional layers were zero-padded with a stride of one. The final output size of the fully connected layer was $1 \times 1 \times 3$, with the output of the first two layers being $1 \times 1 \times 4096$.

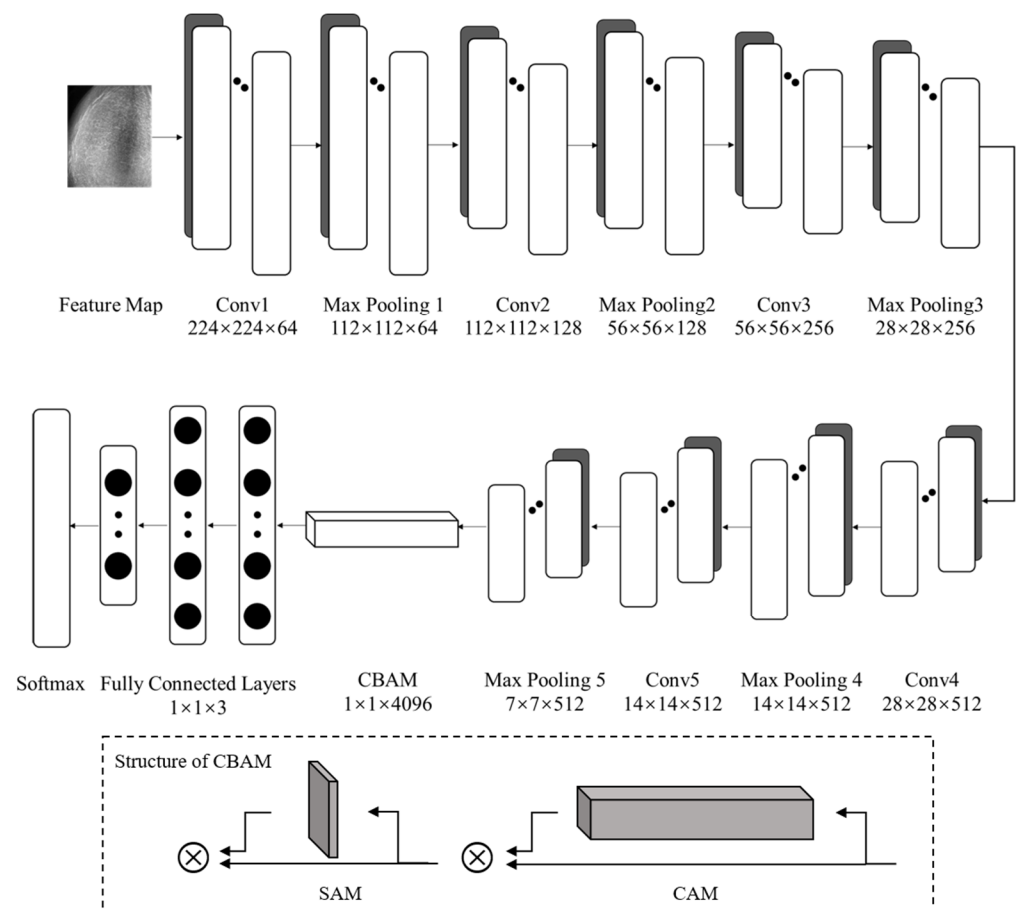


Figure 8. VGG-19+CBAM network structure. (Dots of larger size represent neurons, and dots of smaller size represent apostrophes).

3.3. Model Training

3.3.1. Experimental Software and Hardware

Model training and testing in this paper were conducted in a GPU environment. The CPU was an AMD Ryzen 5 3500U processor (by Advanced Micro Devices, Inc., Santa Clara, CA, USA), paired with an NVIDIA Tesla V100 GPU (by Nvidia Corporation, Santa Clara, CA, USA) and 16.00 GB RAM. The operating system was Windows 10, and the Anaconda version was 4.7.12. The CUDA version was “cuda_10.0.130_411.31_win10”, the cuDNN version was “cudnn-10.0-windows10-x64-v7.4.2.24”, and the TensorFlow version was “tensorflow-gpu 1.13.1”.

3.3.2. Database Processing

The database was randomly divided into the training set and the testing set, with a ratio of 2:1. The training set comprised 3824 samples, with 1616 labeled as “good”, 1429 as “average”, and 779 as “poor”. The testing set consisted of 1912 samples, including 808 labeled “good”, 715 labeled “average”, and 389 labeled “poor”. The training set samples were used to train the model, while the testing set samples were used for evaluation and comparison.

3.3.3. Loss Function

The loss function quantifies the disparity between predicted and actual results. In this work, cross-entropy served as the loss function, measuring the difference between

the predicted distribution and the real distribution. In multi-classification problems, cross-entropy can be expressed as

$$\text{Loss} = -\frac{1}{N} \sum_{i=0}^{N-1} \sum_{k=0}^{K-1} y_{i,k} \ln p_{i,k} \quad (1)$$

where N is the number of samples, K is the number of labels, $y_{i,k}$ indicates the i -th sample with the true label denoted by k , and $p_{i,k}$ indicates the probability that the i -th sample is predicted by k .

3.3.4. Evaluation Indicator

To assess the model's accuracy, we employed the confusion matrix, accuracy rate, and attention map. The confusion matrix visually evaluates the classification results, categorizing them into four types for binary classification: true positive, false negative, false positive, and true negative. True positive indicates both the model output and sample labels are positive, true negative indicates both are negative, false positive indicates the model output is positive while the sample is negative, and false negative indicates the model output is negative while the sample is positive.

Accuracy, a commonly used classification performance indicator, is derived from the confusion matrix. It represents the ratio of correctly identified samples to the total number of samples. The accuracy calculation can be expressed as

$$\text{acc} = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}} \quad (2)$$

where N_{TP} , N_{TN} , N_{FP} , and N_{FN} represent the number of samples in the true positive, the true negative, the false positive, and the false negative.

We utilized Grad-CAM to visually illustrate the impact of the attention mechanism. Grad-CAM [29] expresses attention by overlaying a heat distribution map on the input image, using colors to highlight the network's focal regions. The essence of Grad-CAM involves utilizing gradients to calculate the weight values of the activation map for a specific layer. This calculation involves a weighted sum, the result of which can be expressed as

$$L_{\text{Grad-CAM}}^c = \text{ReLU}\left(\sum_k \alpha_k^c A^k\right) \quad (3)$$

where A represents the feature layer, which is generally the feature layer output by the last convolutional layer, k represents the k -th channel in layer A , A^k represents the data in the k channel in layer A , c represents the type of label, and α_k^c represents the weight of A^k , defined as

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{i,j}^k} \quad (4)$$

where y^c is the score predicted by the network for c before Softmax activation, $A_{i,j}^k$ is the amount of data at the coordinate (i, j) in channel k of layer A , and Z is the dimension of the feature layer.

3.3.5. Training Hyperparameters

The model underwent 60 training epochs, with a batch size of 32 and an initial learning rate of 0.0001, and utilized the Adam optimizer with default values.

4. Results and Discussion

4.1. Comparison of Network Accuracy

To investigate the effect of classification networks on the internal quality grading of dried abalone, we conducted tests on the dried abalone database using five original networks: VGG-16, VGG-19, MobileNet, AlexNet, and Xception. Table 2 presents the

comparison result, VGG-19 network achieved an accuracy of 94.21% for the internal quality grading of dried abalone samples, which was significantly higher than that of the other networks. This outcome can be attributed to several factors: MobileNet is tailored for mobile and embedded devices, prioritizing speed over accuracy, which leads to a reduced number of layers and consequently lower recognition precision; Xception, while having more layers than MobileNet, incorporates depthwise separable convolutions that substantially decrease computational requirements, allowing it to maintain a reasonable level of accuracy. The VGG-19 model boasts the deepest architecture, which is conducive to high-precision image classification tasks. VGG-16 and AlexNet, with slightly fewer layers than VGG-19, exhibited slightly lower recognition accuracy in comparison to VGG-19. These results underscore the importance of network architecture in achieving accurate quality grading for dried abalone.

Table 2. Performance comparisons of the five networks.

Model	Accuracy (%)	Error Rate (%)
VGG-16	92.29	7.71
VGG-19	94.21	5.79
MobileNet	42.08	67.92
AlexNet	92.50	7.50
Xception	75.31	24.69

To further explore the impact of attention mechanisms and batch normalization (BN) on the VGG-19 network, we conducted a series of experiments to evaluate their effects on model performance. Attention mechanisms are expected to enhance feature extraction by focusing on relevant parts of the input data, potentially improving the accuracy and robustness of the model. Batch normalization, on the other hand, is known to stabilize the learning process and reduce the internal covariate shift, which can accelerate training and lead to better convergence.

Table 3 shows the accuracy of VGG-19 using different components. Integrating BN into the VGG-19 network effectively enhanced the accuracy of the internal quality classification of dried abalones. The accuracy of VGG-19+SE did not improve, demonstrating that separate channel attention did not improve the classification accuracy. However, the accuracy of the VGG-19+CBAM model increased to 95.14%. CBAM can improve the accuracy of VGG-19 in the classification of dried abalone X-ray images. Additionally, since SE and CBAM are lightweight and general-purpose modules with a minimal increase in parameter count and computational load, their impact on training and inference times is negligible, and thus, this paper does not include a comparative analysis of parameter counts.

Table 3. Ablation study of VGG-19 with different components.

Model	Original	Using BN
VGG-19	94.21	94.30
VGG-19+SE	93.48	93.88
VGG-19+CBAM	94.37	95.14

Consequently, by comparing the convergence speed of the loss value and accuracy among the three networks during training, VGG-19+CBAM demonstrated a faster convergence speed than VGG-19 and VGG-19+SE. Figure 9 depicts the confusion matrices of the testing set for VGG-19, VGG-19+CBAM, and VGG-19+SE. VGG-19 exhibited errors in two categories: low accuracy for “poor” samples, likely due to a small sample size causing an imbalance, and misrecognition errors between “good” and “average” samples, reflecting their fine-grained similarities.

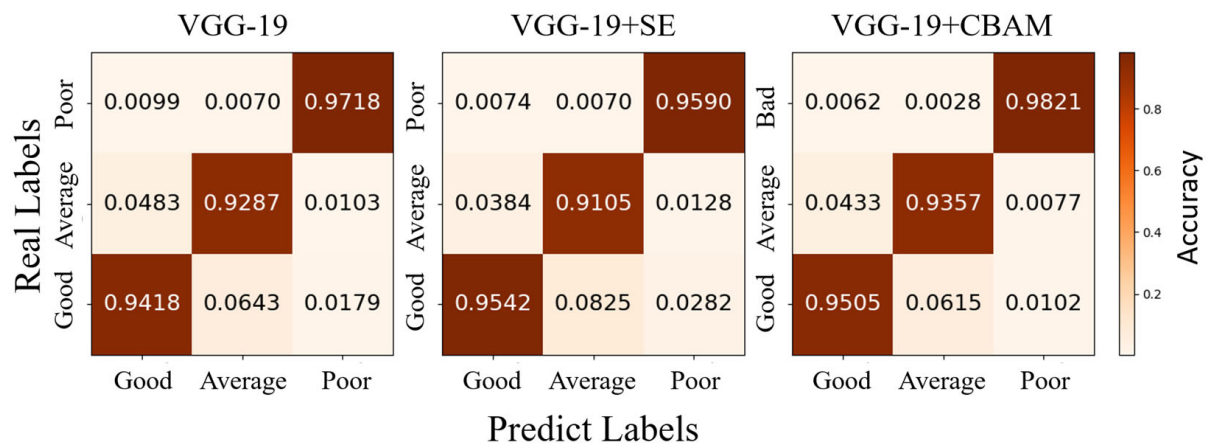


Figure 9. Confusion matrices of the three networks.

Compared with VGG-19, both attention mechanisms enhanced the accuracy for “good” samples, which means that VGG-19 suffered the highest misjudgment in that category. In general, VGG-19+CBAM outperformed VGG-19, and VGG-19+SE had the lowest accuracy. Notably, CBAM significantly reduced the fine-grained characteristics of “good” and “average” samples, concurrently addressing the imbalance issue with “poor” samples.

Furthermore, we employed five-fold cross-validation to assess the VGG-19+CBAM network, as shown in Figure 10. This procedure helps ensure the robustness of the model and reduces reliance on specific training sets, thereby enhancing the model’s predictive power on unseen data. By implementing cross-validation, we can more reliably evaluate the model’s effectiveness across different subsets of the data, providing a comprehensive understanding of its classification capabilities and generalizability. The results of the five-fold cross-validation indicated that the VGG-19+CBAM network exhibited varying levels of performance across different subsets of the dried abalone database, with stability being a concern. This inconsistency was primarily attributed to the insufficient number of samples, which led to performance discrepancies. In the future, we plan to augment the database with more samples to enhance the generalizability and stability of the VGG-19+CBAM network.

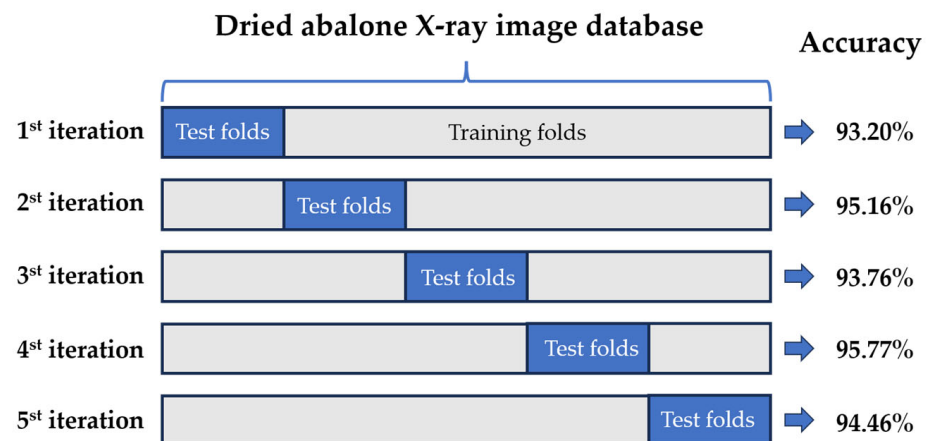


Figure 10. VGG-19+CBAM cross-validation of the dried abalone X-ray image database.

4.2. Comparison of the Attention Mechanism Effect

To compare the effects of the attention mechanisms and understand CBAM’s functionality, we used Grad-CAM with default values to visualize the results of the three networks, inserted after Conv5. Figure 11 illustrates the results of nine selected sample images, wherein increased redness indicates heightened attention. Compared with the other two networks, VGG-19+CBAM exhibited a broader range of network attention, focusing

more precisely on characteristic areas. With CBAM, “poor” samples received increased attention at central positions, whereas SE paid less attention to these regions. The attention regions for “average” and “good” samples were significantly expanded, suggesting that the network extracted more information from the images.

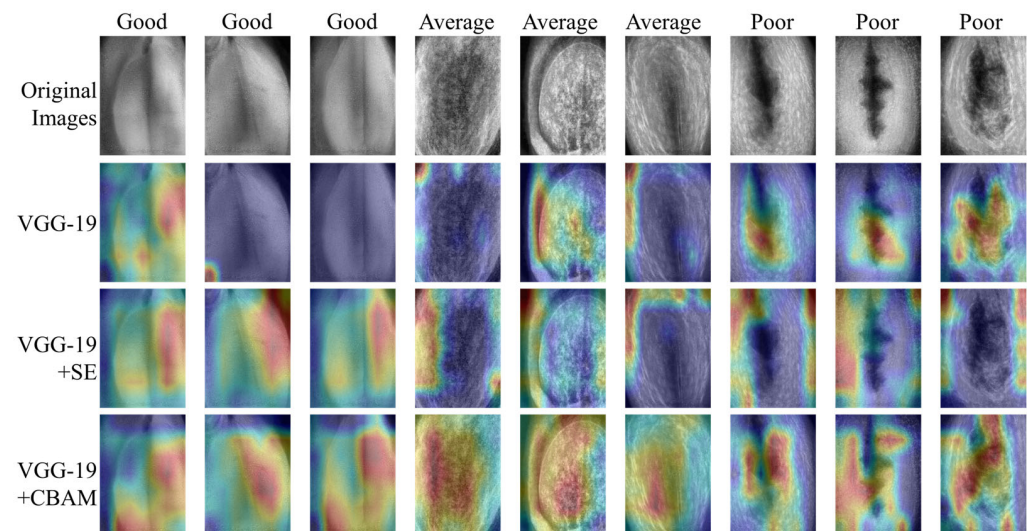


Figure 11. Grad-CAM visualization results.

5. Conclusions

To grade the internal quality of dried abalone quickly and non-destructively, this study adopted X-ray as input and optimized the VGG-19 with an attention mechanism. This method can be used for internal quality grading standardization for dried abalone, achieving high accuracy. The following are the main conclusions:

- (1) We verified and described the differences in the internal structure of dried abalone with different qualities. X-ray and electron microscope images were used to verify the internal differences between different qualities of dried abalone. The results provided a theoretical basis for the following quality identification of dried abalone and also provided a new perspective for exploring the upgrading of processing technology and the quality improvement of dried abalone.
- (2) At present, the classification of dried abalone in the industry is mainly based on human sensory judgment, and there is a lack of a classification basis for the internal quality of dried abalone. In this work, we explored the feasibility of X-ray as a nondestructive method for quality identification and grading of dried abalone. X-ray images were used to judge the quality of dried abalone. This method has good universality and can be extended to more types of foods to realize nondestructive and rapid inspection, identification, and classification.
- (3) We established the first dried abalone X-ray image quality classification database and used the original VGG-19 and its modified versions combined with SE or CBAM for automatic internal quality classification of dried abalone. Specifically, the accuracies of VGG-19, VGG-19+SE, and VGG-19+CBAM were 94.30%, 93.88%, and 95.14%, respectively.

However, our study has limitations. Initially, the precision of recognition accuracy did not reach the desired levels, possibly due to factors such as the scope of the dataset and model complexity. Additionally, the stability of the recognition process was lacking, with the model’s performance exhibiting variability, likely due to overfitting on the training data.

In the future, we will aim to resolve these limitations. We intend to improve the accuracy of dried abalone quality classification by delving into more sophisticated algorithms, fine-tuning hyperparameters, and incorporating further relevant features (origin, production year, specifications, etc.) to bolster the model’s predictive power. Moreover, we

will explore augmenting the dataset with a more diverse and balanced collection of dried abalone X-ray images to alleviate overfitting-related issues and strengthen the model's generalization across various data subsets. Furthermore, we will implement our algorithm alongside appropriate hardware and apply it within the dried abalone industry.

Author Contributions: Conceptualization, Y.Z., H.L., J.G. and W.Y.; Data curation, Y.Z. and H.L.; Funding acquisition, J.C. and R.Z.; Methodology, Y.Z., H.L. and J.G.; Resources, W.Y.; Software, Y.Z. and H.L.; Validation, Y.Z.; Writing—original draft, Y.Z. and H.L.; Writing—review and editing, W.Y. and R.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Science and Technology Projects of Innovation Laboratory for Sciences and Technologies of Energy Materials of Fujian Province (IKKEM), grant number H RTP202231; the National Nature Science Foundation of China, grant number 62001404; the National Funding Program for Postdoctoral Researchers, grant number GZC20231397; and the China Postdoctoral Science Foundation, grant number 2023M742914.

Institutional Review Board Statement: Not applicable.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The databases generated for this study are available upon request to the corresponding author.

Acknowledgments: We wish to express our heartfelt gratitude to Xuan Luo and Jinqian Yang for their help with the experiment design, experimental data interpretation, and participation in the reclassification.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Song, Y.; Zang, X.; Kamal, T.; Bi, J.; Cong, S.; Zhu, B.; Tan, M. Real-time detection of water dynamics in abalone (*Haliotis discus hannai* Ino) during drying and rehydration processes assessed by LF-NMR and MRI. *Dry. Technol.* **2018**, *36*, 72–83. [\[CrossRef\]](#)
2. Cepero-Betancourt, Y.; Tabilo-Munizaga, G.; Lemus-Mondaca, R.; P'erez-Won, M.; Villalobos-Carvajal, R.; Moreno-Orsorio, L. High pressure impregnation-assisted drying of abalone (*Haliotis rufescens*) slices: Changes in protein conformation, thermal properties, and microstructure. *J. Food Process. Preserv.* **2023**, *46*, 10. [\[CrossRef\]](#)
3. Liang, J.; Zhao, G.H.; Li, Y.; Zang, J. Processing Optimization of Dried Abalone Based on Texture Characteristics and the Microstructure Comparison. *Food Res. Dev.* **2022**, *43*, 84–90.
4. Yu, M.M.; Fan, Y.C.; Zhang, X.R.; Li, D.Y.; Liu, Y.X.; Zhou, D.Y.; Zhu, B.W. Effect of boiling on texture of abalone muscles and its mechanism based on proteomic techniques. *Food Chem.* **2022**, *388*, 133014. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Sunxian, Z. Effect of Drying Methods to Nutritional Compositions of Abalone. *J. Food Sci. Technol.* **2018**, *36*, 78–83.
6. Li, Q.; Lu, J.; Chang, Y.; Shen, G.; Feng, J. Effect of different cooking methods on nutritional intake and different storage treatments on nutritional losses of abalone. *Food Chem.* **2022**, *377*, 132047. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Ma, Y.B.; Zou, W.G.; Ai, C.X.; Luo, X.; Liu, S.T.; Huang, H.Q.; Chen, Y.X.; You, W.W.; Ke, C.H. Interactive effects of diet and water temperature on the nutritional value, flavor, and sensory quality in hybrid abalone meat. *Aquaculture* **2023**, *567*, 739241. [\[CrossRef\]](#)
8. Mathanker, S.; Weckler, P.; Bowser, T. X-ray Applications in Food and Agriculture: A Review. *Trans. ASABE* **2013**, *56*, 1227–1239.
9. Stock, S.R. X-ray microtomography of materials. *Int. Mater. Rev.* **1999**, *4*, 44. [\[CrossRef\]](#)
10. Hoheisel, M. Review of medical imaging with emphasis on X-ray detectors, Nuclear Instruments Methods in Physics Research. Section A, Accelerators, Spectrometers. *Detect. Assoc. Equip.* **2006**, *563*, 215–224. [\[CrossRef\]](#)
11. Cnudde, V.; Boone, M.N. High-resolution X-ray computed tomography in geosciences: A review of the current technology and applications. *Earth-Sci. Rev.* **2013**, *123*, 1–17. [\[CrossRef\]](#)
12. Feng, X.; Zhang, H.; Yu, P. X-ray fluorescence application in food, feed, and agricultural science: A critical review. *Crit. Rev. Food Sci. Nutr.* **2020**, *6*, 1–11. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Lim, H.; Lee, J.; Lee, S.; Cho, H.; Lee, H.; Jeon, D. Low-density foreign body detection in food products using single-shot grid-based dark-field X-ray imaging. *J. Food Eng.* **2022**, *335*, 111189. [\[CrossRef\]](#)
14. Looverbosch, T.V.D.; He, J.; Tempelaere, A.; Kelchtermans, K.; Verboven, P.; Tuytelaars, T.; Sijbers, J.; Nicolai, B. Inline nondestructive internal disorder detection in pear fruit using explainable deep anomaly detection on X-ray images. *Comput. Electron. Agric.* **2022**, *197*, 106962. [\[CrossRef\]](#)
15. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
16. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-Excitation Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 2011–2023. [\[CrossRef\]](#) [\[PubMed\]](#)

17. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. *CBAM: Convolutional Block Attention Module*; Springer: Cham, Switzerland, 2018.
18. Gulzar, Y. Enhancing soybean classification with modified inception model: Atransfer learning approach. *Emir. J. Food Agric.* **2024**, *36*, 1–9. [[CrossRef](#)]
19. Alkanan, M.; Gulzar, Y. Enhanced corn seed disease classification: Leveraging MobileNetV2 with feature augmentation and transfer learning. *Front. Appl. Math. Stat.* **2024**, *9*, 1320177. [[CrossRef](#)]
20. Gulzar, Y. Fruit image classification model based on MobileNetV2 with deep transfer learning technique. *Sustainability* **2023**, *15*, 1906. [[CrossRef](#)]
21. Gao, Z. Study on Drying Technology and Product Characteristics of Abalone. Master's Thesis, Fujian Agriculture and Forestry University, Fuzhou, China, 2011.
22. Olaechea, R.P.; Ushio, H.; Watabe, S.; Takada, K.; Hatae, K. Toughness and Collagen Content of Abalone Muscles. *Biosci. Biotechnol. Biochem.* **1993**, *57*, 6–11. [[CrossRef](#)]
23. Xiao, G.; Dong, X.; Zhu, B. Abalone gastropod of different parts and structure study. *J. Food Sci. Technol.* **2010**, *35*, 155–159.
24. O'Keeffe, D. *Denoising of Carpal Bones for Computerised Assessment of Bone Age*; University of Canterbury, Department of Electrical Computer Engineering: Christchurch, New Zealand, 2010.
25. Howard, A.G.; Zhu, M.; Chen, B. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv* **2017**, arXiv:1704.04861.
26. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
27. Hollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 17–21 June 2016; pp. 1800–1807.
28. Niu, Z.; Zhong, Q.; Yu, H. A Review on the Attention Mechanism of Deep Learning. *Neurocomputing* **2021**, *452*, 48–62. [[CrossRef](#)]
29. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.