Journal of High School Science

FreshU-GAN: A Deep Learning model for determining beef freshness

Zhang Y[1,3], Sun J[1,3], Ding E[2,3], Hou R[2,3,*]

Submitted: May 25, 2025, Revised: version 1, June 30, 2025, version 2, July 11, 2025, verison 3, July 17, 2025, verison 4, August 7, 2025
Accepted: July 8, 2025

Abstract
Meat freshness is an important aspect of food safety. As one of the most common types of meat, accurate assessment of beef freshness helps protect consumers' health and prevent potential health risks. To provide a convenient and accessible method for consumers to evaluate beef freshness based solely on visual information, we propose a novel deep learning framework that creatively integrates U-Net and Generative Adversarial Networks (GANs). Specifically, U-Net serves a dual purpose: as the generator within the GAN to produce realistic samples, and as a feature extractor for freshness classification. The discriminator in the GANs compels the U-Net to learn meaningful and discriminative features that improve classification performance. To validate the robustness and adaptability of our model, we executed our model on three different individual datasets, as well as the pooled dataset, to demonstrate the effectiveness and versatility of our proposed model across various imaging conditions.

Keywords
Food Safety, Beef freshness classification, Convolutional Neural Network, U-Net, Generative Adversarial Network, Meat freshness, Food freshness detection, Spoiled meat, Foodborne disease, Public health

Youran Zhang, Jiawei Sun, [1]Rangitoto College, [3]Rangitoto & Vineyards AI Innovation Centre, 564 East Coast Road, Mairangi Bay, Auckland, 0630, New Zealand. 183872@cloud.rangitoto.school.nz , 184217@cloud.rangitoto.school.nz
Emily Ding, Corresponding author: Robert Hou[*], [2]Vineyards AI lab, [3]Rangitoto & Vineyards AI Innovation Centre, 564 East Coast Road, Mairangi Bay, Auckland, 0630, New Zealand. emily.ding@vineyardsailab.com , robert.hou@vineyardsailab.com

https://doi.org/10.64336/001c.143179

Introduction

Food safety is a critical concern for human health. According to the World Health Organization (WHO), foodborne disease affects approximately 600 million people each year and causes 420,000 deaths (1). Because meat is perishable, its freshness affects the risk of foodborne illnesses caused by bacterial contamination, such as Escherichia, Salmonella, and Listeria (2). Among various types of meat, beef is one of the most commonly consumed worldwide, making it a representative choice for studying meat freshness. While customers often rely on visible and olfactory cues such as colour, texture, and smell to assess meat freshness, these indicators can be subjective and unreliable. Therefore, innovative technologies for freshness detection are needed to safeguard public health and maintain consumer confidence in meat products (3).

The rapid development of artificial intelligence(AI) has found application across various industries, including food safety. AI-based technologies, particularly deep learning and computer vision, have shown significant promise in automating and enhancing the detection process. Various studies (4,5) have demonstrated the effectiveness of AI in assessing meat quality and demonstrated its advantages and potential applications. Some of these studies, such as those by Taheri-Garavand et al. (4) are based only on the visual appearance of meat, aiming to simulate the real-world scenario where customers in a supermarket can use a visual-based algorithm to determine freshness simply by capturing an image. Similarly, Elangovan et al. (6) proposed an intelligent system combining machine learning algorithms and visual analysis to assess meat safety, marking a significant step forward in food quality control using visual data alone.

In this paper, we propose a novel deep learning framework for beef freshness classification, since beef is one of the most commonly consumed types of meat. The proposed model integrates two complementary models, U-Net and GANs. U-Net's unique capability lies in capturing fine-grained pixel-level features (7). Part of the bottleneck can learn abstract, discriminative and semantic representation. In our framework, we employed U-Net as the generator, replacing the traditional generator in GANs. Through adversarial training, the discriminator encourages the bottleneck to refine its feature extraction process, enhancing its focus on classification-critical details.

2 Related work

2.1 U-Net

U-Net, proposed in 2015 for biomedical image segmentation, is a fully convolutional neural network with a U-shaped architecture (7). It consists of a contracting path for feature extraction and a symmetric expanding path for precise localization. With skip connections to retain spatial information, U-Net can be trained end-to-end and delivers excellent performance across various segmentation tasks, especially in biomedical imaging.

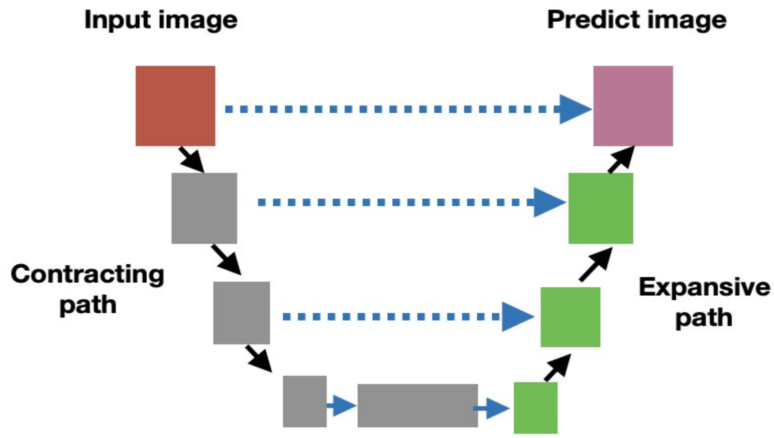

Figure 1. Architecture of U-Net

U-Net's, input is the original image, and the output is the segmented image, where each pixel is assigned a class label. It learns to map the input image to a segmentation mask while preserving both important features and fine details.

2.2 Generative Adversarial Networks

Generative Adversarial Networks (GANs) are a class of AI models based on deep neural networks, designed to generate new data samples which have not existed before, such as images, videos, and speech [8]. A GAN consists of two key components: a generator and a discriminator. The generator takes random noise as input and produces synthetic data. The discriminator, acting as a binary classifier, distinguishes between real and generated data. Through continuous adversarial training, the generator improves its ability to create realistic outputs, while the discriminator refines its ability to differentiate them. This dynamic competition drives GANs to generate highly convincing and high-quality data.

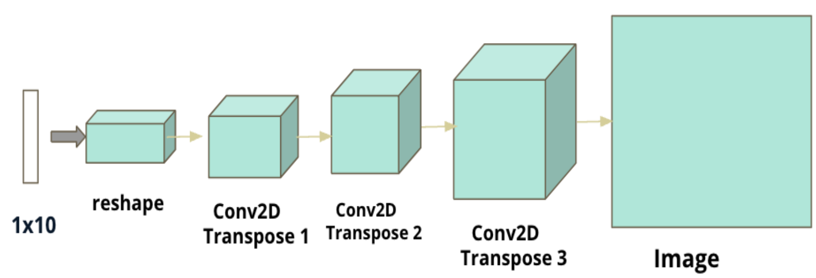

Figure 2. Architecture of the generator in GANs

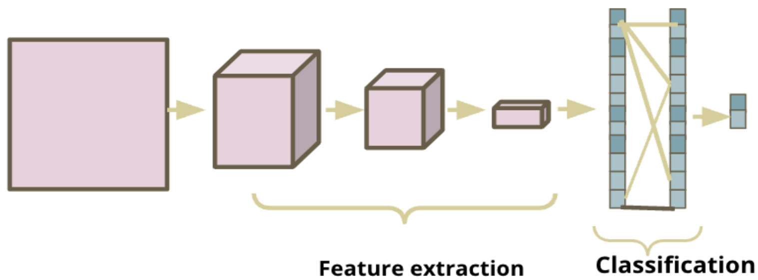

Figure 3. Architecture of Discriminator in GAN

The advantage of U-Net lies in its ability to extract features and preserve details effectively. In GANs, by adversarial training, the generative ability is enhanced while also improving classification performance. Hence, we explored the possibility of creatively combining these two powerful models to identify meat freshness.

3 Proposed model

In our model, we combined the advantages of U-Net, which has shown to be effective in extracting textures and details from biomedical images, with adversarial learning from GANs, to enhance classifier performance and further improve the feature representation capability within U-Net. We propose a model called FreshUGAN, where U-Net functions as the "Generator" for GANs and " Feature Extractor" for the classifier. Through adversarial learning, the model enables more effective feature extraction and classification.

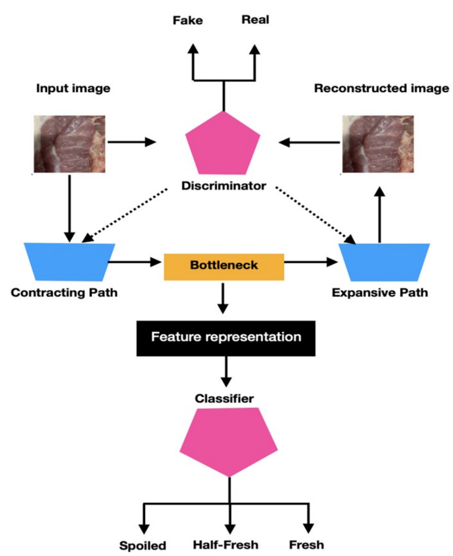

Figure 4. Framework of the proposed model

3.1 Dual role of U-Net: feature extractor and generator

U-Net achieves above average segmentation results due to its symmetric U-shaped architecture and skip connections, which enable it to capture both global and local features while preserving high-resolution information. The expanded path reconstructs the input image by the bottleneck layer with skip connections. The bottleneck layer, set at the bottom of the U-Net, is adjacent to the output of the final contracting layers. It serves as a feature representation of the input image.

Therefore, a perfect bottleneck representation results in a high-quality reconstructed image. Conversely, a well-reconstructed image must be from a valuable and well-learned bottleneck layer.

The input image can be represented as X, in U-net, the contracting path maps it to the bottleneck layer Z, hence $Z=f_{con}(X,\Theta_{con})$ (eq.1), where f_{con} is the mapping function of the contracting path, and Θ_{con} refers to the contracting path parameters. In the expanding path, Z is an input of the decoder function, $\widehat{X}=f_{\exp}(Z,S_1,S_2\cdots,S_n;\Theta_{\exp})$ (eq.2), where $\Theta_{\exp}$ represents the expanded path parameters, and $\widehat{X}$ is the reconstructed image, S_i which refers to the $i-th$ skip connection of the contracting path. If U-Net is well-designed and satisfies the equality $\widehat{X}\approx X$ (eq.3), this would be indicative that the bottleneck layer Z is a low-dimensional but effective feature representation of the input image X. Furthermore, a more accurately reconstructed image indicates a more effective and expressive bottleneck representation.The mean-squared error for image reconstruction is $L_{MSE}=\frac{1}{N}\sum_{i=1}^{N}(X_i-\widehat{X}_i)$ (eq.4) .

3.2 A Discriminator for adversarial feature refinement

To obtain a well-learned feature representation, we employed a discriminator through adversarial learning to enforce the expanded path to reconstruct an image that closely resembles the input. This process encouraged the bottleneck to capture a more meaningful and representative feature representation.

As mentioned in section 2.2 *Generative Adversarial Networks* , a discriminator D is trained to distinguish between the real image X and the reconstructed image $\widehat{X}$. The objective was to make the attributes of $\widehat{X}$ as close as possible to those of X so that discriminator cannot distinguish betweeen the two. The adversarial loss is defined as $L_{adv}=min[-[\log D(X_{x\ p_{data}})]+\log[1-D(\widehat{X}_{x\ p_{u-net}})]]$ (eq.5), where $D(X_{x\ p_{data}})$ represents the probability that the discriminator classifies X as a real image, and $D(\widehat{X}_{x\ p_{u-net}})$ represents the probability that $\widehat{X}$ is a real image. Therefore, the combined loss function for the U-Net and adversarial learning is $L=L_{MSE}+L_{adv}$ (eq.6).

3.3 Freshness classification

The bottleneck representation would hence contain the features which can be used to identify the freshness of meat. We built the classifier with a fully connected neural network. The initial classifier layer was flattened and was followed by two dense layers with different neurons, both activated by

ReLU. The final output layer used a softmax activation with three (number of categories) neurons, producing class probabilities for multi-class classification.

This classifier was then integrated with a U-Net architecture, meaning that the final model used the original U-Net input and produced a classification output based on the extracted features. The model was compiled with the Adam optimizer and categorical cross-entropy loss, making it suitable for multi-class classification tasks.

4 Methods

In our work, we utilized the train_test_split function (from sklearn.model_selection) to ensure randomness and maintain class distribution. Furthermore, during the model training process, data was fed in batches, and these batches were randomly sampled in each epoch. As a result, the model was exposed to different batch compositions across epochs, which enhanced generalization and reduced overfitting. Table 1 lists the model parameters. These remained the same for all the three datasets as well as for the pooled dataset.

Table 1. Model Parameters

Module	Parameter	Description
U-Net (Generator)	Input Shape	(128, 128, 3)
	Convolutional layers (encoder)	64 →128 →256 →512
	Pooling Layers	MaxPooling2D(2,2)
	Convolutional Layers (Decoder)	256 →128 →64
	Upsampling Layers	UpSampling2D (2,2)
Discriminator	Input Shape	(128,128,3)
	Convolutional Layers	64→128
	Output Layer Activation	Sigmoid
Classifier	Input Shape	(512, 16, 16)
	Fully Connected Layers	128 → 64
	Output Layer Activation	Softmax
	Neuron Number	3

The first and the third dataset were already split into training and testing sets with balanced class distributions. For the second dataset, we applied the train_test_split function with the stratify parameter enabled to perform stratified sampling. This function preserved the class distribution consistency and prevented biases that could adversely affect model performance.This ensured that each subset maintained a representative proportion of each class, facilitating fair comparison and reproducibility among researchers.

We did not perform k-fold cross validation. Instead, we validated the robustness of our method using several approaches. First, the first and third datasets we used are publicly available standard datasets with training and testing splits carefully defined by the original publishers. These splits ensured balanced class distributions and representativeness and are widely accepted within the academic community. Second, for the combined dataset, we applied an 80/20 random split for training and testing, a ratio commonly recognized as standard and reasonable in both industry and academia. These measures collectively ensure that our model performed consistently and reliably across different data partitions, thus demonstrating its robustness.

5 Results

To evaluate the effectiveness of our proposed model, we performed experiments using three related meat freshness classification datasets, and finally, on the pooled dataset.

5.1 First dataset

The first dataset was provided by the Roboflow Team (11), which contained 2,266 images. This dataset was originally collected for food quality assessment, machine learning-based freshness detection, and shelf-life prediction. Each image is labeled as one of three categories: Fresh, Half-Fresh, and Spoiled. Figure. 5 shows representative samples from the dataset.

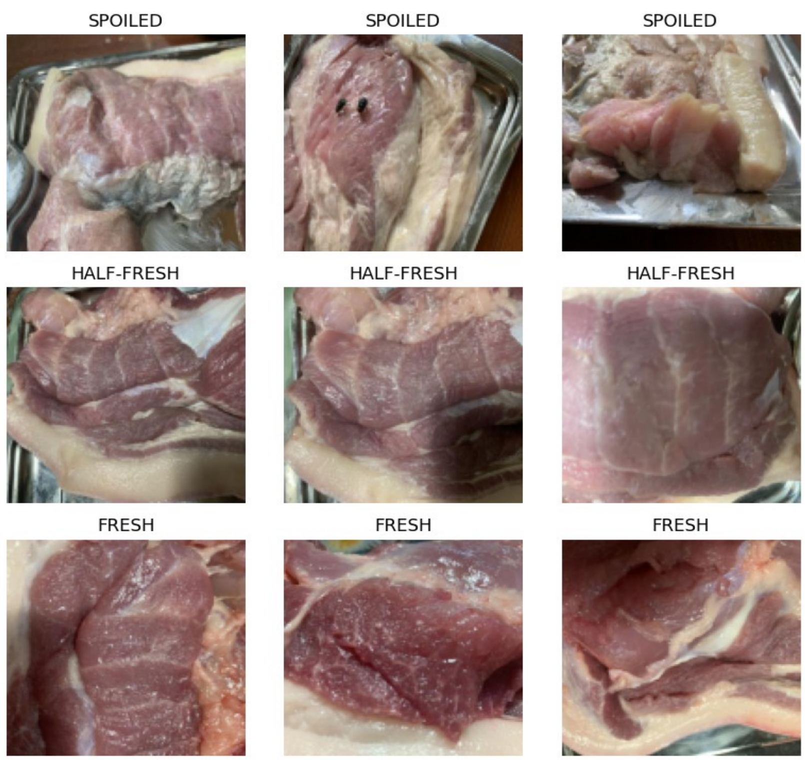

Figure 5. Representative samples for each category

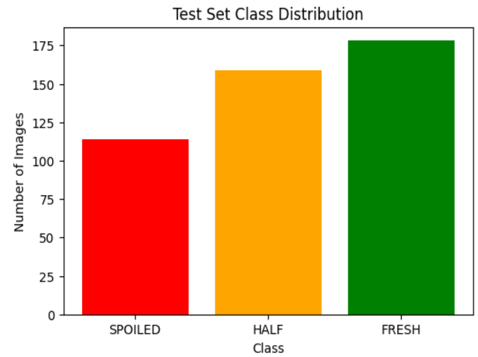

Figure 6. Category distributions in the training and test datasets

As shown in Figure 5, the images in both the training and testing datasets were captured under various lighting conditions. This diversity helped improve the model's generalization ability. Additionally, the training and testing sets were randomly selected, ensuring that the evaluation results were meaningful and robust.

Figure 6, shows that the "FRESH" and the "SPOILED" categories contained the most and the least samples respectively. The number of samples across the three categories was relatively balanced, with no significant class imbalance. Using the parameters in Table 1, our model achieved an accuracy of 90.91%. The labels were assigned as follows: *Spoiled* = 0, *Half-Fresh* = 1, and *Fresh* = 2.

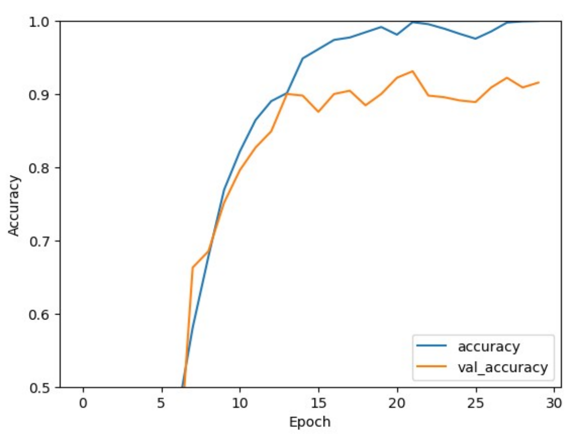

Figure 7. Training and Validation Accuracy over Epochs

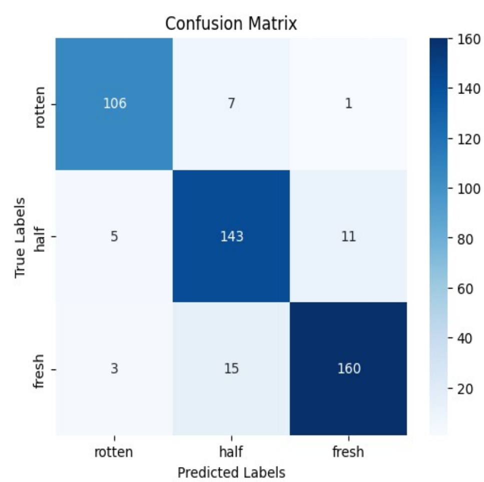

Figure 8. Confusion matrix

Figure 7 shows that the training and validation accuracy curves begin to stabilize from the 20th epoch onward. This indicates that the model has largely converged and is no longer experiencing significant performance fluctuations. Additionally, the performance gap between the training and validation sets remained < 10%. This suggested that the model generalized well on the data and did not suffer from severe overfitting.

A Confusion matrix was constructed that summarized the performance of the classification model by comparing its predictions against the actual annotated ground-truth. The diagonal elements represent correct predictions, while off-diagonal elements represent misclassifications between different classes. To maintain consistency with the subsequent experimental results, we used the word "rotten" to represent the “SPOILED" category in the confusion matrix.

Figure 8 shows that for the category SPOILED (rotten) out of 114 samples, there were 106 samples that were correctly classified, while 8 were misclassified, where 7 were predicted as HALF-FRESH (half), and 1 as FRESH (fresh).

For the HALF-FRESH (half) category, 159 samples were correctly identified, however 11 were incorrectly predicted as fresh, which could pose a potential food safety risk. This was likely due to the high visual similarity between the two categories in terms of colour and texture, as illustrated in Figure 5.

Regarding the FRESH (fresh) category, 160 samples were correctly classified, with 3 misclassified as SPOILED (rotten), which is a conservative error but may still lead to unnecessary waste. Figure 9, shows several randomly selected samples from the testing set along with their predicted results. "T" denotes the true label, and "P" indicates the predicted label.

A classification report is a commonly used tool to evaluate a model’s performance. It includes key metrics, such as precision, which refers to the number of true positives divided by the total number of positive predictions (i.e., the number of true positives plus the number of false positives).; recall, measures how often the model correctly identifies positive instances (true positives) from all the actual positive samples in the dataset.; and the F1-score, which is the harmonic mean of precision and recall, offering a balanced assessment of the model’s performance. The "support" column indicates the number of true samples for each class in the dataset, that is, the actual occurrences of that class. The different averages are ways to calculate metrics for multiple classes. The Macro average calculates the metric for each class first, then take the average. All classes have equal weight. The Weighted average calculates the metric for each class, then take a weighted average based on the number of samples in each class. Classes with more samples have more influence. The Micro average combines all classes’ prediction results and calculates the overall metric. This is useful as an indicator of the overall performance when the classes are imbalanced.

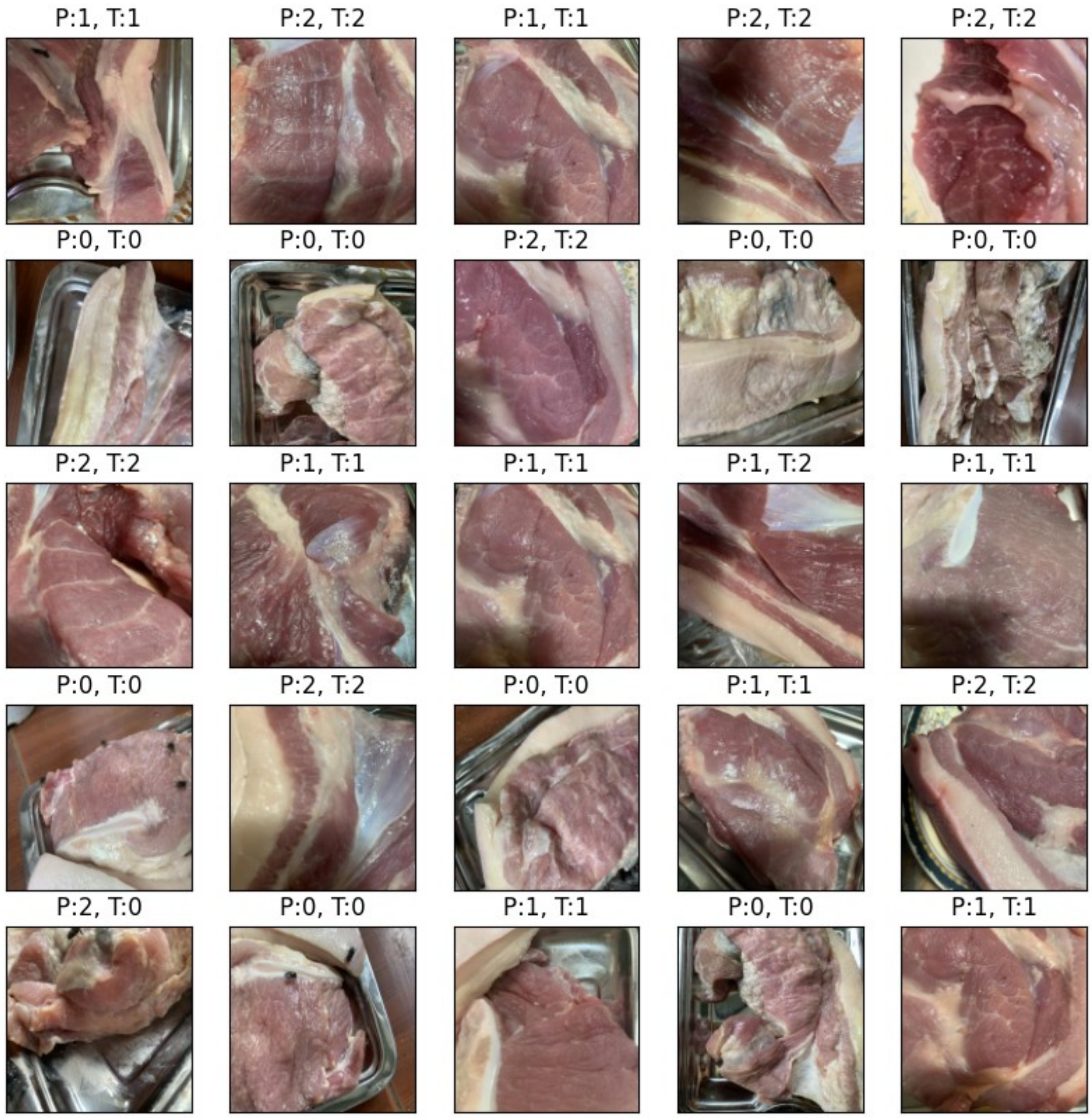

Figure 9. Several samples from the testing set along with their predicted results

The results in Table 2 and Figure 8 suggest that the model performs well overall, demonstrating a strong ability to distinguish between classes, with precision and recall numbers of ~ 90%. However, the overlap between the Fresh and Half-Fresh categories, as well as the examples of Spoiled samples being misclassified as Fresh, shown in Figures 8 and 9 highlights the need for either a more refined feature extraction or better data augmentation to improve classification accuracy in misclassified cases.

Table 2. Classification report

	Precision	Recall	f1-score	support
Spoiled (0)	0.90	0.93	0.91	114
Half fresh (1)	0.89	0.92	0.90	159
Fresh (2)	0.95	0.88	0.91	178
Micro avg	0.91	0.91	0.91	451
Macro avg	0.91	0.91	0.91	451
Weighted avg	0.91	0.91	0.91	451
Samples avg	0.91	0.91	0.91	451

5.2 Second dataset

The second dataset, referred to as "Images of Fresh and Non-Fresh Beef Meat Samples," is taken from Sanchez et al. (12). This dataset contains images of thirty beef meat samples across three different cuts: inside skirt, knuckle, and sirloin. For each cut, ten pieces of meat (each measuring 5 cm × 5 cm) were used. Images were captured on two different days. The first day after purchase were labeled as *fresh,* and the fifth day after purchase were labeled as *non-fresh*. This resulted in a total of 60 meat images, 30 fresh and 30 non-fresh, evenly distributed across the three cuts. Additionally, the dataset included corresponding segmentation images for each of the 60 samples, which could directly be used for classification purposes. Therefore, to adapt our model to this binary classification task, we made modifications to the output layer. As shown in Table 1, the classifier should have a single output neuron with a sigmoid activation function. Representative samples of the dataset are shown in Figure 10.

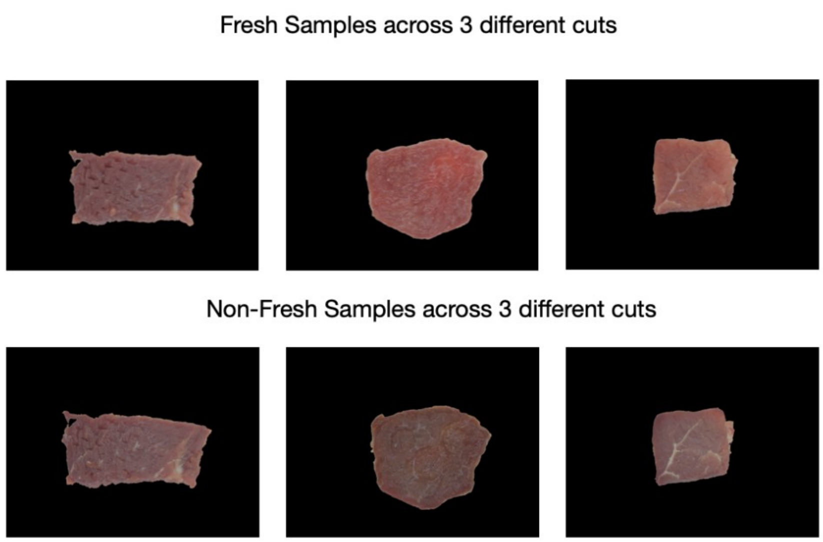

Figure 10. Sample images from three different meat cuts, each with corresponding fresh and non-fresh examples

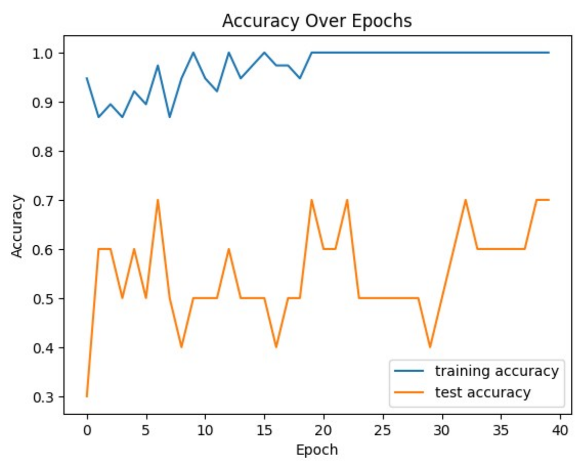

Figure 11. Training and validation accuracy over epochs

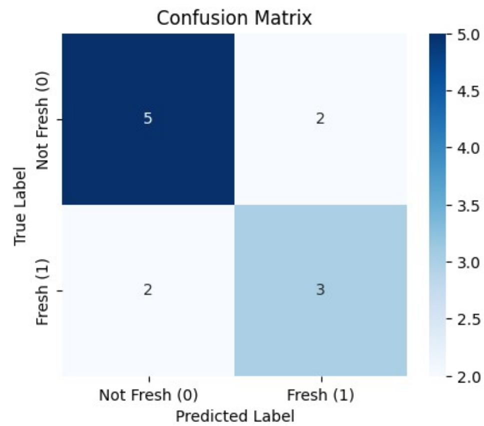

Figure 12. Confusion matrix

Table 3. Classification report

	Precision	Recall	fi-score	Support
Not fresh	0.71	0.71	0.71	7
Fresh	0.60	0.60	0.60	5
Accuracy			0.67	12
Macro avg	0.66	0.66	0.66	12
Weighted avg	0.67	0.67	0.67	12

As shown in Table 3, for this dataset, the model achieved a classification accuracy of ~ 70%. This low accuracy could have resulted from the limited size of the training dataset significantly constraining the model's performance. From the training curves shown in Figure 11, it can be seen that the model is prone to overfitting, which means that it achieved a high accuracy on the training set but relatively poor accuracy on the test set.

5.3 Third dataset

The third dataset we utilized was the Beef Quality Image Dataset for Deep Learning, referred to as LOCBEEF, obtained from Dharma et al. (13). This dataset was organized into two main directories, train and test. Each of these directories contained two subfolders corresponding to the categories fresh and rotten. In total, the dataset included 2,288 images for training and 980 images for testing. Similar to the second dataset, LOCBEEF was a binary classification dataset with two classes, fresh and rotten. Therefore, we applied the same model architecture and parameter settings as those used for the second dataset.

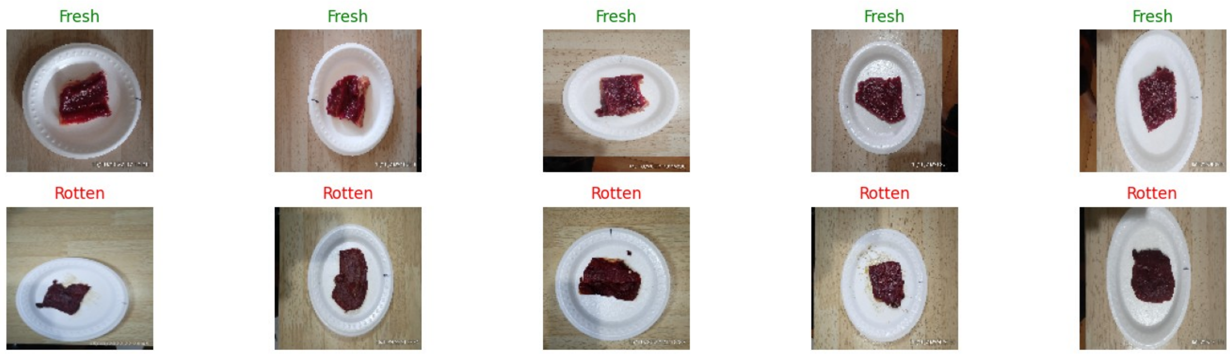

Figure 13. Representative images for each category

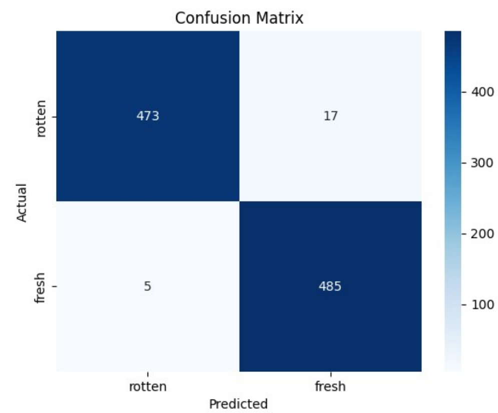

Figure 14. Training and validation accuracy over epochs

Figure 15. Confusion matrix

Table 4. Classification report

	Precision	Recall	f1-score	Support
Rotten	0.99	0.97	0.98	490
Fresh	0.97	0.99	0.98	490
Accuracy			0.98	980
Macro avg	0.98	0.98	0.98	980
Weighted avg	0.98	0.98	0.98	980

The classification report (Table 4) shows that the model performed well on both classes, achieving an overall accuracy of 98%. The precision and recall values were both ~ 1, indicating that the model returned very few false positive or false negative predictions.

5.4 Merged data from the three datasets

To demonstrate the generalization and robustness of the model, we combined three different datasets for training. During this process we addressed two challenges, the first of which involved label harmonization. The first dataset contained three categories, while the second and third datasets only included two each. To enable consistent multi-class classification, we redefined and relabeled the categories across all datasets using the first dataset as a template. Therefore, we retained Fresh as label 2 and redefined not fresh as “HALF” (label 1) in the second dataset. In the third dataset, Fresh was also assigned label 2 for consistency, while rotten was relabeled as “SPOILED” (label 0).

The second challenge involved data reorganization. To construct a new combined dataset, we randomly selected a subset of images from each source. Specifically, the combined dataset consisted of 545 images from the first dataset, 60 images from the second, and 980 images from the third. In constructing the combined dataset, our image selection was based on three considerations; those of data availability, computational feasibility, and class distribution balance. For the first dataset, which contains 2,266 images (1,721 training + 545 testing), We included the widely adopted 545-image testing subset, which offered a balanced class distribution and contributed representative samples. For the second dataset, which contained only 60 images, we included all of them to ensure completeness. For the third dataset, which contained over 3,200 images (2,288 training + 980 testing), we selected the 980-image testing subset, just as we did for the first dataset. The reason for this selecting this particular combination of images was to avoid computational overload and data imbalance. In addition, the testing subsets chosen had balanced class distributions and diverse image conditions, making them more suitable for reliable training and evaluation.

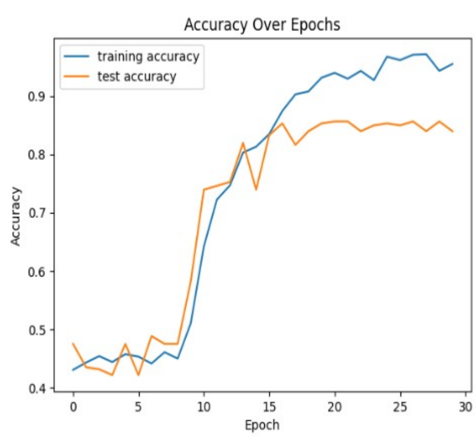

Figure 16. Training and validation accuracy over epochs

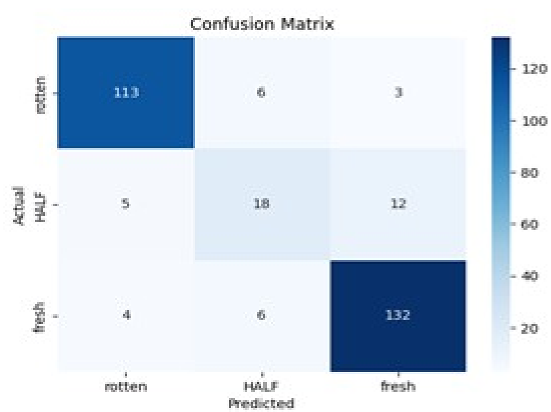

Figure 17. Confusion matrix

Table 5. Classification report

	Precision	Recall	f1-score	Support
Rotten	0.93	0.93	0.93	122
Half-fresh	0.60	0.51	0.55	35
Fresh	0.90	0.93	0.91	142
Accuracy			0.88	299
Macro avg	0.81	0.79	0.80	299
Weighted avg	0.87	0.88	0.88	299

The model achieved an accuracy of ~ 87% on the pooled dataset.

5.5 Minimizing false negatives

Rotten or half-fresh food classifed as fresh can have severe implication for food safety. Dataset 1 is one of the most authoritative and widely recognized public datasets in this field. Its data collection process and labeling standards are rigorous, closely reflecting real-world consumer scenarios when assessing meat freshness in supermarkets. Our model achieved a false negative rate of 1/106 ≈ 0.94% on this dataset, successfully staying below the 1% threshold, demonstrating the model's reliability under ideal data conditions. In contrast, Datasets 2 and 3 can be argued to be of noticeably lower quality. Their collection environments and labeling standards do not seem to align well with real-world consumer scenarios. For example, in Dataset 3 (17/473 ≈ 3.59%), the objects are not clearly visible or large enough, and the texture details are poorly defined. Additionally, both datasets support only binary classification, lacking intermediate categories such as 'half-fresh,' which increases the risk of misclassifying rotten samples as fresh and results in higher false negative rates. For the combined dataset, the false negative rate is approximately 2.65% (3 out of 113), which we consider acceptable given the varying labeling standards and image quality across the three datasets. Therefore, a model's effectiveness depends not only on the algorithm itself but also on whether the data is representative and accurately labeled. Since different image quality is expected to be encountered in real-life situations, it could be that device- independent, color space perceptual models, such as CIELAB and CIELUV; as defined by the International Comission on Illumination; could be embedded within models such as ours to improve stand-alone, image-only food-freshness predictability. When combined with a color-changing dye or equivalent as a surrogate for spoilage (see Limitations), the minimization of false-negatives using imaging alone becomes even more plausible.

6 Limitations

The training and testing data in this study were primarily based on photographs of beef cuts; therefore, the model's current capability is limited to assessing beef freshness. It is important to note that a deep learning model's classification ability largely depends on the type and quality of the training data. If a comprehensive dataset including other meat categories such as fish, chicken, lamb; among others; with appropriate annotations were available, the model could theoretically classify the freshness of meat from multiple meat categories. Since the current data only included beef, the scope and title of this study are limited to beef freshness classification to accurately reflect the application boundary.

Furthermore, for large cuts of meat, if the interior is rotten, but the surface is not, the model cannot accurately determine freshness, since it can only analyze the visible area captured by a camera. The algorithm can only detect spoilage caused by microorganisms that lead to visible changes; microorganisms or other factors that cause decay without altering visual features cannot be detected. If method(s) could be developed - using food safe dyes as an example – to allow any decay to migrate and translate into a subtle color or texture change throughout the entirety of the cut-meat, methods such as the one described here (using only images to detect freshness) could be used with no limitations. The other alternative would be to revert back to multimodal detection techniques using other environmental sensors and biosensors (including those that could detect olfactory, enzymatic and gaseous changes) in conjunction with visual image data.

7 Conclusion

In this paper, we proposed a novel model, which we term FreshU-GAN, which integrates the model architectural strengths of U-Net and Generative Adversarial Networks (GANs). The U-Net component functioned as an auto-encoder to extract detailed visual features from meat images, while the adversarial training mechanism in the GAN encouraged the model to reconstruct images that were indistinguishable from real samples, thereby enhancing feature learning.

To validate the effectiveness and robustness of our approach, we applied the model to three individual beef datasets, each containing images of beef cuts - one with three freshness categories and two with binary (fresh/rotten) classification. We also evaluated the model on a combined dataset, containing images from all the three individual datasets to demonstrate its generalization capability across different dataset settings.

This method was designed to be user-friendly, allowing consumers to simply upload a photo of the meat and receive an initial freshness assessment. The method is unique in that, it determines meat freshness only from visual image data without relying on inputs from any additional environmental or biosensor data. For future work, we aim to enhance the model's ability to detect subtle visual cues of spoilage, further increasing classification precision and reliability.

8. References

1. World Health Organization, "Estimating the burden of foodborne diseases," https://www.who.int/activities/estimating-the-burden-of-foodborne-diseases

2. Pennacchia, C., Ercolini, D.,Villani, F. (2011). Spoilage-related microbiota associated with chilled beef stored in air or vacuum pack. *Food Microbiology*, 28(1), 84–93. https://doi.org/10.1016/j.fm.2010.08.010

3. Luo, L., Xu, S., Fu, S., Tao, G. (2025). A Review of Factors Influencing the Freshness of Meat and Meat Products and Methods for Its Control. *Meat Research*, 39(2), 55–66. https://doi.org/10.7506/rlyj1001-8123-20240909-239

4. Taheri-Garavand, A., Fatahi, S., Omid, M., Makino, Y. (2019). Meat quality evaluation based on computer vision technique: A review. *Meat Science*, 156,183–195. https://doi.org/10.1016/j.meatsci.2019.06.002

5. Lin, Y., Ma, J., Sun, D.W., Cheng, J.H., Zhou C. (2024). Fast real-time monitoring of meat freshness based on fluorescent sensing array and deep learning: From development to deployment. *Food Chemistry*, 448, 139078. https://doi.org/10.1016/j.foodchem.2024.139078

6. Elangovan, P., Dhurairajan, V., Nath, M.K., Yogarajah, P., Condell, J. (2024). A Novel Approach for Meat Quality Assessment Using an Ensemble of Compact Convolutional Neural Networks. *Applied Sciences*, 14(14), 5979. https://doi.org/10.3390/app14145979

7. Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, 9351. Springer. https://doi.org/10.1007/978-3-319-24574-4_28

8. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B. (2014). Generative Adversarial Networks. *Advances in Neural Information Processing Systems (NeurIPS)*, 3(11), 2672–2680. http://dx.doi.org/10.1145/3422622

9. Ali, S., DiPaola, D., Lee, I., Hong, J., Breazeal, C. (2021). Exploring Generative Models with Middle School Students. *Proc. 2021 CHI Conf. Human Factors in Computing Systems*, Article no. 678, 1-13. https://doi.org/10.1145/3411764.3445226

10. He, Y., Ding, E.X., Hou, R.J. (2025). Hide-and-Seek GANs for Generation with Limited Data. In: Mahmud, M., Doborjeh, M., Wong, K., Leung, A.C.S., Doborjeh, Z., Tanveer, M. (eds) Neural Information Processing. ICONIP 2024. Lecture Notes in Computer Science,15288, Springer, Singapore. https://doi.org/10.1007/978-981-96-6582-2_18

11.[Dataset] Meat Freshness Dataset. *Roboflow Universe*. https://universe.roboflow.com/reinsgundar/meat-freshness-xpmq5

12. Sanchez, C., Dominguez-Soberanes, J., Orvañanos, T., Lara, M., García, E., Cisneros, J.P., Orozco, L.E., Rosales-Tavera, E. (2022). Images dataset of beef meat samples with different shelf life. *Data Brief*, 50, 109503. https://doi.org/10.1016/j.dib.2023.109503

13. Mulya Dharma, T., Adriman, R., Saddami, K. (2022). LOCBEEF: Beef Quality Image dataset for Deep Learning Models. *Mendeley Data*, V1. https://doi.org/10.17632/nhs6mjg6yy.1