Grenze International Journal of Engineering and Technology, June Issue

Few-Shot Learning for Binary Classification in Red Chilli Powder

Srividhya C.S[1], Siddesha S[2] and Rekha GR[3]
[1-3]Department of Computer Applications, JSS Science and Technology University, Mysuru, India
Email: srividhyacs@jssstuniv.in

***Abstract*— Detecting adulteration in red chilli powder is important for maintaining food quality and safety. One of the common adulterants that is added to red chilli powder is brick powder in global markets. This work evaluates the effective-ness of a few-shot learning approach to distinguish between pure and adulterated chilli powder. A small dataset comprising 30 images per class was used to simulate a low-data scenario, out of which 10 images were used for unseen data for testing purposes, which allowed to test how well these models perform under limited data conditions. Five pretrained networks -ResNet-18, MobileNetV3, EfficientNet-B0, DenseNet-121 and ConvNeXt-Tiny, were fine-tuned by keeping the feature extractor frozen and training only the final classification layer. Despite the small dataset, the models performed well, with DenseNet-121 and EfficientNet-B0 reaching validation accuracies of 95% and 90%. These results suggest that few-shot learning can be a practical and cost-effective method for creating quick, image-based tools to check for adulteration in spice powders.**

***Keywords*— Adulteration Detection, Red Chilli Powder, CNN, Few-Shot Learning, Food Safety and Quality Assessment, Transfer Learning, Spice Powder.**

I. INTRODUCTION

Adulteration of spice powders, particularly red chilli powder, remains a continuous concern for food safety and authenticity. Due to economic incentives and the visual similarity of adulterants (e.g., rice husk, brick dust, wheat bran, colourants, sawdust) are added to the genuine product [1]. Traditional analytical techniques give accurate results. But those methods are labor-intensive, require sophisticated laboratory setup, experts' interpretations, and involve costly chemicals for wet-lab analysis. This motivates the need for rapid, non-destructive screening approaches.
Recent studies have explored complementary sensing (e.g., e-nose) and vision-based machine learning to identify adulterants in spice and other foods, underscoring the feasibility of automated screening in laboratory and industrial settings [2]. Deep convolutional neural networks (CNNs) have advanced food quality assessment, including the detection of spice and milk adulteration, but typically require a large amount of labelled data [3]. Transfer learning mitigates data scarcity by reusing representational features learned on large-scale natural image corpora (e.g., ImageNet), often yielding strong performance with limited domain data [4, 5]. Comparative evidence indicates that fine-tuning mod-ern ImageNet-pretrained backbones transfers well to downstream visual tasks and can outperform handcrafted pipelines; however, benefits depend on task/domain similarity and careful regularization under low-data regimes [6].

Grenze ID: 01.GIJET.12.2.33_25

In parallel, the few-shot learning (FSL) literature targets extreme data scarcity by either (i) few-shot fine-tuning - adapting pretrained CNNs with strong regularization or (ii) meta-learning - training with episodic tasks to generalize rapidly to novel classes [7]. Reviews in agri-food contexts report that FSL can attain competitive accuracy with minimal supervision, particularly when classes are fixed or when episodic designs are feasible; nonetheless, meta-learning typically presumes many classes for meta-training, which may not hold in binary authenticity scenarios [8]. Against this backdrop, we evaluate a practical few-shot fine-tuning pipeline for binary chilli powder authenticity (pure vs adulterated) using a few pretrained CNNs such as ResNet-18 [9], MobileNetV3 [10], EfficientNet-B0 [4], DenseNet-121 [11], ConvNeXt-Tiny [12]. With only a few images per class, freeze most of the feature extractor and retrain compact classifier heads while employing augmentation, label smoothing, dropout, and early stopping to coun-ter overfitting, in line with best practices for small-sample transfer [13].

The proposed study clarifies that fine-tuning is sufficient to support recent findings that CNN-based methods can accurately detect spice adulteration and other forms of food fraud. Few-shot learning, enabled through transfer learning and heavy augmentation, provides a mechanism to train robust models using minimal labelled data. Despite recent advances, few-shot deep learning for spice adulteration remains underexplored.

This paper presents a systematic evaluation of five state-of-the-art convolu-tional neural networks in a few-shot setting for binary classification of pure and adulterated chilli powder images. Our contributions are twofold: (1) a rigorously regularised, few-shot fine-tuning baseline for chilli adulteration under severe data constraints; (2) a comparative assessment across five modern CNN back-bones using standard metrics like accuracy, precision/recall/F1, and confusion matrices. Collectively, the results support few-shot fine-tuning as a resource-efficient, deployable screening strategy for binary classification. The future di-rections in sensor fusion and domain-specific pretraining.

II. Related Works

Deep learning has become an important tool for food authentication and quality assessment. The trained models can capture subtle visual or spectral differences that are difficult to detect with traditional methods. Most existing deep-learning studies in food sciences rely on supervised training with moderately large da-tasets. These models often perform well but the challenge is to apply them in practical laboratory environments where obtaining hundreds of labelled samples is not always achievable. This has led to growing interest in learning methods that perform reliably even when only a small amount of training data is available. Several studies have explored few-shot learning, especially in the context of hyperspectral imaging. In one example, the work on camel milk powder adultera-tion detection, where the authors used a metric-based meta-learning approach to handle domain shifts in hyperspectral data. Their results showed that the method could maintain stable accuracy even when the acquisition conditions changed, and only a few labelled samples were available. This work reveals the usefulness of few-shot learning for food adulteration problems where distributions shift across batches [14].

Another line of research applied prototype-based models to address class im-balance in food hyperspectral datasets. The Proto-DS framework combined self-supervised learning with prototypical classification to improve the detection of minority adulteration classes and was validated on citrus peel, coffee beans, and herbal ingredients. This study showed that prototype construction can support meaningful classification even when limited annotated samples are present [15].

Few-shot capability has also been supported through architectural innova-tions. CRSSNet integrates convolutional residual learning with a Siamese atten-tion mechanism and achieved improved performance on standard hyperspectral few-shot benchmarks. Although not directly applied to food, the method is rele-vant due to the shared challenges of scarce labels and spectral variability [16]. A related hyperspectral work proposed using pseudo-labels from unlabelled sam-ples to stabilise learning when only very few labelled samples are available. This reduced overfitting and improved results in both one-shot and few-shot settings [17].

Few-shot learning has also been explored in image-based food quality analy-sis. A notable example is pork-freshness grading using a lightweight network called BBSNet, which combines a compact CNN backbone with attention and improved normalisation. The model performed well in both one-shot and five-shot tasks, showing that small-sample learning can succeed with only RGB imag-es [18]. Another study used a prototypical few-shot adaptation of a pretrained Transformer model for avocado-ripeness detection. This research work shows good accuracy 80–96% with very limited labelled samples [19].

Alongside these few-shot and meta-learning approaches, several studies have used pretrained CNNs transfer well to HSI tasks, improving few-shot perfor-mance when only limited labels are available. One hyperspectral study showed that pretrained models such as EfficientNet, ResNet, VGG, and DenseNet can be fine-tuned effectively even when only limited labelled samples are available [20]. Another related paper proposed a contrastive-learning and transfer-learning framework for few-shot hyperspectral classification, where a pretrained backbone is fine-tuned on very small support sets [20]. In plant pathology, fine-tuned EfficientNet-B0 was successfully applied to leaf-disease detection using small datasets and achieved strong performance [21]. A companion methodolog-ical work discusses how hybrid fine-tuning strategies can stabilise learning when sample sizes are very small [22].

In parallel, deep learning has been widely used in spice and herb adulteration detection through conventional training approaches. These include DenseNet-based adulteration detection in chilli powder with explainable visualization [23]. Reviews on hyperspectral and agricultural imaging consistently highlight la-belled-data scarcity as one of the major obstacles in food imaging and spectros-copy. Few-shot learning, meta-learning and transfer learning are identified as promising directions for dealing with limited data [8].

III. METHODOLOGY

A. Dataset Description

The dataset [24] used in this study consists of images of pure and images of red chilli powder adulterated with brick powder in various percentage levels ranging from 5 percent to 50 percent. To reflect a typical few-shot learning scenario, the dataset is utilised partially. In total, 30 images represent the pure red chilli pow-der class, while 30 images correspond to different percentage mixtures of brick powder to the pure red chilli powder, out of which 10 images in each category were used for testing unseen data through inference code for each model.

B. Pre-processing

In the proposed study, for experimental purposes, the dataset was manually di-vided into two sets: Training set (10 pure, 10 adulterated) and Validation set (10 pure, 10 adulterated). Few-shot samples from both pure and adulterated classes were augmented to produce diversity. All images are resized to 224 × 224 pixels to ensure uniformity across all deep learning architectures.

Since the training samples are fewer in number, intense but still realistic data augmentation methods, such as random horizontal flipping, random rotation, random resized cropping with scaling in between, and colour jittering, were cru-cial for reducing overfitting. These augmentations applied only to the training images. The augmentation idea has ensured the improved sample diversity while maintaining the visual characteristics of red chilli powder textures.

C. Model Architectures

The proposed method, as shown in Figure 1, is experimented with five advanced Convolutional Neural Network architectures, such as DenseNet121, EfficientNet-B0, ConvNext-Tiny, ResNet18 and MobileNetV3-Small. To ensure consistency, all five architectures, ResNet18, MobileNetV3-Small, EfficientNet-B0, Dense-Net121 and ConvNeXt-Tiny were initialised using transfer-learning weights, which is particularly useful when working with small datasets.

DenseNet121: DenseNet is a CNN where every layer is connected to all previ-ous layers, helping the model learn better features, avoid redundancy, and train more smoothly. Here, every layer receives features from all previous layers, thereby retaining all valuable information[19]. Among the tested models, DenseNet121 proved to be the most effective. It handled fine-grained texture differences well, even with minimal training data. Its dense connections allowed the network to make better use of learned features, which in turn helped maintain steady and high validation accuracy. Throughout training, the loss and accuracy curves progressed smoothly, indicating strong generalisation in the few-shot learning.

EfficientNet-B0: The model EfficientNet builds its neural network in a bal-anced manner by expanding depth, width, and image size concurrently such that all three are used. This method provides improved accuracy with lower number of parameters[9]. The efficiency of the EfficientNet-B0 model could be seen from how balanced was its performance with regards to precision, recall, and accuracy for both classes. This is due to the compound scaling technique imple-mented in order to achieve good performance even with low computation capaci-ty. Despite minor changes during the latter part of training, the model performed very consistently.

ConvNeXt-Tiny: The ConvNeXt neural network is an upgraded version of CNN architecture. It takes into consideration the design principles of the trans-former design[20]. All components have been re-designed such that they can achieve the level of performance like a transfer but in terms of efficiency of the convolution. The ConvNeXt-Tiny neural network proved itself to have good fea-ture-learning capabilities because of its upgraded convolution design inspired by transformer design principles. It had an accuracy very close to EfficientNet-B0, and the training/validation curves were relatively stable.

ResNet18: ResNet makes training deep models quicker and more reliable by introducing shortcut (skip) connections that aid the network in learning residu-als[16]. For our baseline model, the ResNet18 was great in terms of consistently achieving poor performance. Though the residual connections helped with gradi-ent flow, the network struggled to differentiate between texture details in the few-shot setting because of the lack of depth. However, we can note that the training and validation performance were very well-regulated according to the curve alignment.

MobileNetV3-Small: MobileNet is small, light and fast, making it ideal for mobile devices. It uses lightweight depth-wise separable convolutions to achieve fast, efficient inference on mobile and edge devices[17]. MobileNetV3-Small showed early signs of underfitting in this study. Both the training and validation accuracy plateaued quickly, indicating that the model lacked sufficient capacity to separate fine-texture differences between the two classes. Its inference speed is an advantage, but it struggled to extract the level of detail needed in this few-shot scenario.

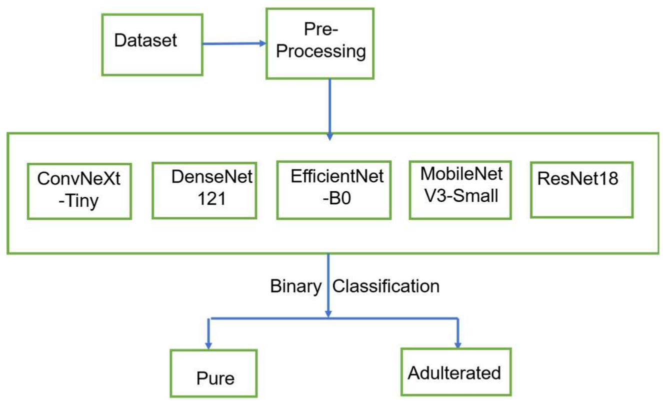

Fig. 1. Block Diagram of the Proposed Method

D. Experimental Setup

To adapt each network for binary classification, the original final layers were replaced with a custom head consisting of a Dropout layer ($p = 0.4$) followed by a fully connected layer with two output neurons representing the pure and adul-terated classes. This ensured that all five models shared the same output structure while also providing regularisation through dropout. Every model was fine-tuned using an identical training setup to maintain fairness in comparison.

The training configuration included Cross-Entropy Loss with label smoothing (0.1), the AdamW optimiser with a learning rate of 1e-4 and weight decay of 1e-4, and a One-Cycle Learning Rate schedule with a maximum learning rate of 3e-4. Training was carried out with a batch size of 4 for up to 40 epochs, along with early stopping based on validation loss, using a patience of 8 epochs. Label smoothing was intentionally applied to prevent the models from becoming over-confident. This pre-processing is also to support better generalisation. Better generalisation is an important consideration when working with small datasets. The One-Cycle LR strategy further helped stabilise convergence and avoided premature stagnation of the learning rate. In each epoch, the models were trained on the augmented dataset and then evaluated on the validation set. The version achieving the lowest validation loss with accuracy as a secondary check was saved for the final evaluation. Early stopping, data augmentation, label smooth-ing and dropout were employed to mitigate overfitting.

In this study, the model performance was assessed using standard classifica-tion metrics suitable for imbalanced and low-sample scenarios. The metrics are Accuracy, Precision, Recall, F1-Score and Confusion Matrix. This ensured that results captured not only overall accuracy but also class-wise performance. This approach is important in adulteration studies where false negatives (mislabelled adulterated samples) have serious implications.

E. Results and Analysis

A comparative study was conducted on five deep learning models. They are ResNet18, MobileNetV3-Small, EfficientNet-B0, DenseNet121, and ConvNeXt-Tiny. Table 1 reports the evaluation of their performance in a few-shot learning setup for detecting adulteration in red chilli powder. All models were fine-tuned using the same training conditions, including identical augmentation steps, opti-miser settings, learning-rate schedule, and early-stopping rules. By keeping these factors constant, the variations in performance more accurately reflect how each architecture learns from very limited data and how well it can generalise under such constraints. EfficientNet-B0, DenseNet-121 and ConvNeXt-Tiny showed the most stable behaviour across multiple runs. Evaluation metrics confirmed that the models generalised well to unseen images within the same distribution.

TABLE I. COMPARATIVE PERFORMANCE OF FIVE DEEP LEARNING MODELS IN FEW-SHOT CHILLI POW-DER BINARY CLASSIFICATION

Model	Accuracy (%)	Precision	Recall	F1-Score
DenseNet121	95	0.95	0.95	0.95
EfficientNet-B0	90	0.92	0.90	0.90
ConvNeXt-Tiny	90	0.92	0.90	0.90
ResNet18	80	0.86	0.80	0.79
MobileNetV3-Small	45	0.45	0.45	0.44

DenseNet121 stood out as the most dependable model, giving the highest accu-racy along with strong precision, recall and F1-score. It was clearly better at picking up the fine texture differences in the images compared to the other net-works. EfficientNet-B0 and ConvNeXt-Tiny also performed well, with balanced results and steady feature learning, placing them in the middle but still reliable. ResNet18 offered a modest yet steady baseline, while MobileNetV3-Small found it difficult to separate the two classes, which shows that its smaller design is not sufficient for this task.

Looking at the training curves, DenseNet121 showed a clean and steady drop in both training and validation loss, and the two curves stayed close to each oth-er. This suggests that the model learned in a constant manner without overfitting. EfficientNet-B0 and ConvNeXt-Tiny had a few small variations in the later stag-es of training, but the validation accuracy continued to be generally close to the training accuracy. ResNet18 followed a similar pattern but at a lesser perfor-mance level. MobileNetV3-Small, however, showed clear signs of underfitting, as both training and validation accuracy flattened early, and the losses did not decrease much. These observations are consistent with the final metrics reported for each model.

Overall, the comparison shows that the size and capacity of a model play an important role in a few-shot learning approach towards adulteration detection. Larger networks like DenseNet121, EfficientNet-B0 and ConvNeXt-Tiny could handle the task more effectively, while lighter models such as MobileNetV3-Small lacked the representational strength needed for reliable results. This high-lights the need to choose architectures that can offer a good balance between generalisation and feature-learning ability when working with very small la-belled datasets in food-quality assessment.

After training, we evaluated each model on a separate set of unseen images to verify its generalisation ability. We loaded the trained models for each architec-ture and used identical inference code to ensure a reasonable comparison. This step allowed us to examine how well the models could make the binary classifi-cation of pure and adulterated samples outside the training and validation distri-bution of data. The predictions from this unseen dataset confirmed the consisten-cy of the reported performance metrics.

Figure 2 depicts the Confusion Matrix of the best performing model, that is, Densenet121 and Figure 3 conveys the comparison of performance metrics across models

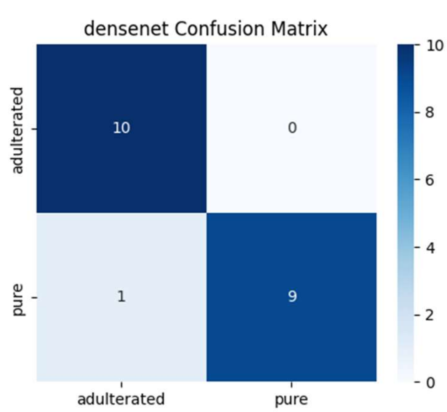

Fig. 2. – Confusion Matrix

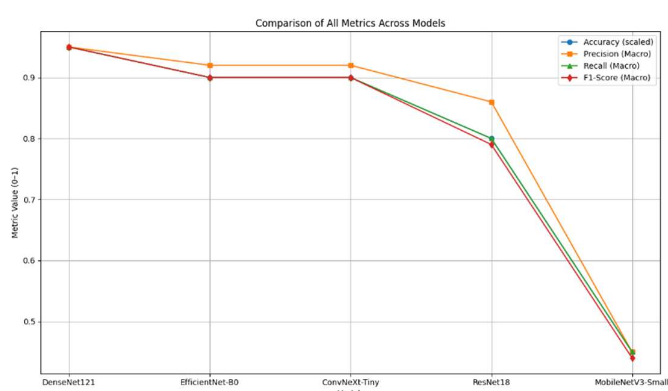

Fig. 3. Comparison of performance metrics across models

IV. Challenges and Discussion

The experiment revealed some challenges with using a few-shot learning ap-proach on a small image dataset of red chilli powder. The biggest limitation in this work was the small amount of data, as only 30 pure and 30 adulterated imag-es were available. Though good augmentation techniques were used, the models could not get enough real variation. This adaptability issue made a few models prone towards overfitting. This could be seen in the shifting validation losses, especially in ResNet18 and ConvNeXt-Tiny models. In these models, the per-formance initially improved but then started to decline. MobileNetV3-Small tended to underfit and often missed fine texture differences in the images. In contrast, it was observed that the deeper models demanded more computing power but generally produced better and more stable results.

The model's capacity and stability during training were found as the main is-sues. Although it needed more processing power, DenseNet121 produced the best results. EfficientNet-B0's performance varied slightly over subsequent epochs, and ConvNeXt-Tiny occasionally displayed overfitting spikes. Through-out the training process, MobileNetV3-Small was experiencing severe underfit-ting. This led to poor class separation. ResNet18 encountered difficulties as well because of its light architecture, which made it more difficult for the model to detect microscopic texture patterns in the chilli powder images.

V. Conclusion

This proposed work established that few-shot learning can be effective for mak-ing binary classification of pure and adulterated red chilli powder. Even with minimal training images, models such as DenseNet121, EfficientNet-B0, and ConvNeXt-Tiny were able to produce good results. Out of the 5 models experi-mented with, DenseNet121 performed top, giving about 95% accuracy with a perfect ROC–AUC score. Overall, it was observed from the investigation results that with good transfer-learning settings, proper regularisation, and strong aug-mentation used, the models can still make reliable predictions in spite of a small dataset.

But it was observed, all networks could give good results. MobileNetV3-Small struggled to capture delicate texture patterns. It exhibited weaker performance, underlining the importance of architecture choice, particularly for image-based adulteration tasks. Collecting more data would help the models generalise better, and the same approach can be tried for detecting different adulterants. Combin-ing image data with other instrumentation data, such as NIR or FTIR. This may also strengthen the system. Approaches such as self-supervised pretraining or meta-learning may further enhance few-shot performance. Overall, this method could lead to the development of fast, non-destructive, and user-friendly digital tools for detecting adulteration in spice powders and ensuring food quality and safety.

References

[1] Choudhary, A., Gupta, N., Hameed, F., & Choton, S. (2020). An overview of food adulteration: Concept, sources, impact, challenges and detection. International Journal of Chemical Studies, 8(1), 2564-2573.

[2] Roy, M., & Yadav, B. K. (2022). Electronic nose for detection of food adulteration: A review. Journal of Food Science and Technology, 59(3), 846-858.

[3] Li, Z., Wang, D., Zhu, T., Tao, Y., & Ni, C. (2024). Review of deep learning-based methods for non-destructive evaluation of agricultural products. Biosystems Engineer-ing, 245, 56-83.

[4] Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolu-tional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.

[5] Kornblith, S., Shlens, J., & Le, Q. V. (2019). Do better imagenet models transfer bet-ter?. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 2661-2671).

[6] Kolesnikov, A., Beyer, L., Zhai, X., Puigcerver, J., Yung, J., Gelly, S., & Houlsby, N. (2020, August). Big transfer (bit): General visual representation learning. In European conference on computer vision (pp. 491-507). Cham: Springer International Publishing.

[7] Wang, Y., Yao, Q., Kwok, J. T., & Ni, L. M. (2020). Generalizing from a few exam-ples: A survey on few-shot learning. ACM computing surveys (csur), 53(3), 1-34.

[8] de Andrade Porto, J. V., Dorsa, A. C., de Moraes Weber, V. A., de Andrade Porto, K. R., & Pistori, H. (2023). Usage of few-shot learning and meta-learning in agriculture: A literature review. Smart Agricultural Technology, 5, 100307.

[9] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recogni-tion. In Proceedings of the IEEE conference on computer vision and pattern recogni-tion (pp. 770-778).

[10] Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., ... & Adam, H. (2019). Searching for mobilenetv3. In Proceedings of the IEEE/CVF international con-ference on computer vision (pp. 1314-1324).

[11] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely con-nected convolutional networks. In Proceedings of the IEEE conference on computer vi-sion and pattern recognition (pp. 4700-4708).

[12] Liu, Z., Mao, H., Wu, C. Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A convnet for the 2020s. In Proceedings of the IEEE/CVF conference on computer vision and pat-tern recognition (pp. 11976-11986).
[13] Chen, W. Y., Liu, Y. C., Kira, Z., Wang, Y. C. F., & Huang, J. B. (2019). A closer look at few-shot classification. arXiv preprint arXiv:1904.04232.
[14] Ruan, S., Di, R., Zhang, Y., Yan, T., Cang, H., Tan, F., ... & Xu, W. (2024). Metric-based meta-learning combined with hyperspectral imaging for rapid detection of adul-teration in domain-shifted camel milk powder. LWT, 206, 116537.
[15] Pang, K., Liu, Y., Zhou, S., Liao, Y., Yin, Z., Zhao, L., & Chen, H. (2024). Proto-DS: A Self-Supervised Learning-Based Nondestructive Testing Approach for Food Adultera-tion with Imbalanced Hyperspectral Data. Foods, 13(22), 3598.
[16] Xia, M., Yuan, G., Yang, L., Xia, K., Ren, Y., Shi, Z., & Zhou, H. (2023). Few-Shot hyperspectral image classification based on convolutional residuals and SAM Siamese networks. Electronics, 12(16), 3415.
[17] Ding, C., Li, Y., Wen, Y., Zheng, M., Zhang, L., Wei, W., & Zhang, Y. (2021). Boost-ing few-shot hyperspectral image classification using pseudo-label learning. Remote Sensing, 13(17), 3539.
[18] Liu, C., Zhang, J., Chen, K., & Huang, J. (2025). BBSNet: An Intelligent Grading Method for Pork Freshness Based on Few-Shot Learning. Foods, 14(14), 2480.
[19] Ahmed, M., Mustafa, H., Wu, M., Babaei, M., Kong, L., Jeong, N., & Gan, Y. (2024). Few shot learning for avocado maturity determination from microwave images. Journal of agriculture and Food Research, 15, 100977.
[20] Yang, G., & Wang, Z. (2025). A Deep Transfer Contrastive Learning Network for Few-Shot Hyperspectral Image Classification. Remote Sensing, 17(16), 2800.
[21] Ali, H., Shifa, N., Benlamri, R., Farooque, A. A., & Yaqub, R. (2025). A fine tuned Ef-ficientNet-B0 convolutional neural network for accurate and efficient classification of apple leaf diseases. Scientific Reports, 15(1), 25732.
[22] Zhao, L., Ou, Z., Zhang, L., & Li, S. (2022). Hybrid Fine-Tuning Strategy for Few-Shot Classification. Computational Intelligence and Neuroscience, 2022(1), 9620755.
[23] Brar, D. S., Singh, B., & Nanda, V. (2025). An XAI-enabled 2D-CNN model for non-destructive detection of natural adulterants in the wonder hot variety of red chilli pow-der. Sustainable Food Technology.
[24] Sarkar, Tanmay; Choudhury, Tanupriya; Bansal, Nikunj; V R, Arunachalaeshwaran (2022), " Dataset for Adulterated Red Chilli Powder with Brick Powder", Mendeley Da-ta, V1, doi: 10.17632/cpm7y44746.1