North Asian International Research Journal of Multidisciplinary

IRJIF IMPACT FACTOR: 5.001

ISSN: 2454-2326 Vol.10, Issue-1 January-2024

Index Copernicus Value: 58.12 Indian Citation Index

NAIRJC UGC-Aligned | HEI-Approved Peer Reviewed Refereed Journal (Under New UGC Guidelines)

DOI: doiglobal.org/10.2024/NAIRJCMD.032

INTELLIGENT DETECTION AND CLASSIFICATION OF FOOD GRAIN ADULTERATION USING DEEP LEARNING AND COMPUTER VISION

DR. JYOTI. Y. KULKARNI

Associate Professor Govt. First Grade College, Kalagi.

ABSTRACT

Food-grain adulteration affects quality assurance, consumer confidence and food safety. Conventional inspection based on manual sorting, visual examination and laboratory testing remains important, but it can be slow, subjective or unsuitable for rapid high-volume screening. This paper proposes a computer-vision framework for intelligent detection and classification of visible adulteration in food grains using deep learning. The proposed workflow covers image acquisition, annotation, preprocessing, augmentation, convolutional neural-network (CNN) feature learning, transfer learning, classification and explainability. The framework is designed for grains such as rice, wheat and pulses and for visually observable foreign matter or abnormal kernels, including stones, husk, damaged grains, discoloured grains and selected look-alike adulterants. The paper distinguishes clearly between visual adulteration that can reasonably be detected from RGB images and chemical or microscopic adulteration that requires laboratory or multispectral methods. A reproducible experimental protocol is proposed using stratified train-validation-test partitions, class-wise precision, recall, F1-score, confusion matrices and macro-averaged metrics. MobileNetV3, EfficientNet-B0 and ResNet50 are proposed as comparative transfer-learning baselines, with an ensemble or optimized lightweight model considered for deployment. No fabricated experimental accuracy is reported; instead, the results section defines the tables and statistical evidence that should be populated after actual model training. The study concludes that deep learning can serve as a rapid screening and decision-support layer for grain quality control, but it should complement—not replace—validated food-analysis methods and regulatory laboratory testing.

***KEYWORDS**: food grain adulteration; deep learning; computer vision; CNN; transfer learning; food quality; grain classification; image processing*

1. INTRODUCTION

Food grains are central to food security and everyday nutrition. Their quality can be reduced by foreign matter, damaged kernels, inferior grains, non-permitted colouring, accidental contamination or deliberate adulteration. In India, the Food Safety and Standards Authority of India (FSSAI) provides standards and analytical methods for cereal and cereal products and also publishes rapid consumer-oriented tests for common adulterants.

Traditional grain inspection includes visual examination, physical separation and laboratory analysis. Visual methods are inexpensive but depend on human attention and expertise. Laboratory methods are more authoritative for chemical and compositional analysis, but they require equipment, reagents, trained personnel and time. This creates a practical need for rapid, non-destructive screening tools that can identify suspicious samples before confirmatory testing.

Computer vision offers such a possibility when adulteration produces visible differences in shape, colour, texture or morphology. Deep convolutional neural networks can learn discriminative image features directly from data, reducing reliance on manually designed descriptors. Transfer learning further allows models pretrained on large image datasets to be adapted to specialized agricultural tasks with comparatively smaller datasets.

This paper develops a research framework for applying deep learning to food-grain adulteration. The emphasis is methodological rigor: clearly defined classes, controlled image acquisition, leakage-free dataset splitting, robust metrics, explainability and explicit separation between visual screening and regulatory confirmation.

2. PROBLEM STATEMENT

Manual inspection of grain samples can become inconsistent when adulterants resemble genuine kernels or when large volumes must be examined rapidly. Lighting conditions, observer fatigue, grain overlap and subtle colour differences can further affect visual judgment. At the same time, a model trained under laboratory imaging conditions may fail when used with ordinary cameras or different backgrounds.

The research problem is therefore twofold: first, to determine whether deep-learning models can learn stable visual representations that distinguish pure grains from visible adulterants; and second, to design an evaluation protocol that demonstrates generalization beyond the exact images used during training.

3. OBJECTIVES

- To develop a computer-vision pipeline for detection and classification of visually observable food-grain adulteration.
- To construct a reproducible image-acquisition and annotation protocol for rice, wheat and selected pulses.
- To compare transfer-learning architectures such as MobileNetV3, EfficientNet-B0 and ResNet50.
- To evaluate model performance using accuracy, precision, recall, F1-score, confusion matrices and macro-averaged metrics.

- To investigate explainability methods such as Grad-CAM to verify whether predictions are based on relevant grain regions.
- To propose a lightweight deployment architecture suitable for laboratories, warehouses, procurement centers or mobile-assisted screening.
- To define the limitations of RGB vision and identify cases requiring chemical, microscopic, hyperspectral or regulatory testing.

4. RESEARCH QUESTIONS

- Can deep CNNs distinguish pure grain samples from visually detectable adulterated samples under controlled imaging conditions?
- Which transfer-learning architecture offers the best trade-off between classification performance and computational efficiency?
- How strongly do lighting, background and camera variation affect model generalization?
- Can explainability maps demonstrate that the network is focusing on grains and adulterants rather than irrelevant background cues?
- How can an AI screening system be integrated responsibly with established food-safety testing procedures?

5. FOOD-GRAIN ADULTERATION: SCOPE OF VISUAL DETECTION

The term adulteration covers heterogeneous phenomena. Some adulterants are visually observable: stones, pebbles, husk, weed seeds, damaged grains, insect-affected grains, abnormal colouring and foreign kernels. FSSAI's consumer guidance includes visual or simple physical tests for extraneous matter in food grains, Dhatura in food grains, colour adulteration, turmeric in sella rice and related cases.

Other forms of adulteration cannot be reliably inferred from a normal RGB photograph. Chemical dyes, pesticide residues, toxins, compositional substitution in flour and microscopic contamination may require chemical analysis, microscopy, spectroscopy or other validated methods. A scientifically responsible vision system must therefore state that a 'pure' image prediction means no targeted visible adulterant was detected—not that the sample has been proven chemically safe.

This distinction prevents overclaiming and defines the proposed system as a screening instrument.

6. REGULATORY AND FOOD-SAFETY CONTEXT

FSSAI specifies methods of analysis under the Food Safety and Standards framework and maintains manuals for different food categories. Its methods page lists a revised second edition of the Manual of Methods of Analysis of Foods—Cereal and Cereal Products uploaded in May 2026. FSSAI also maintains food-product standards and contaminant regulations.

For public awareness, FSSAI's Detect Adulteration with Rapid Test (DART) resources compile simple household tests across multiple food categories. These resources demonstrate that adulteration detection already uses a tiered logic: some abnormalities can be screened simply, whereas authoritative assessment may require laboratory methods.

Deep learning fits naturally into this tiered model as a rapid pre-screening layer. It can flag samples for closer inspection, quantify visible defects and prioritize confirmatory analysis.

7. LITERATURE REVIEW

Computer vision has long been applied to agricultural quality assessment using handcrafted features such as colour histograms, shape descriptors, texture matrices and edge statistics. These approaches require explicit feature engineering and may struggle when within-class variation is large.

Deep learning changed this paradigm by allowing hierarchical features to be learned directly from images. CNNs progressively combine local edges, textures and shapes into higher-level representations. In grain analysis, this makes it possible to learn subtle morphological differences that may be difficult to express as fixed rules.

Transfer learning is especially relevant because specialized food-adulteration datasets are usually much smaller than general computer-vision datasets. A pretrained backbone can be fine-tuned on grain images, reducing training time and improving feature quality. Lightweight networks such as MobileNet are attractive for mobile deployment, while EfficientNet and ResNet families provide strong baselines for accuracy-oriented comparison.

Recent food-quality research also increasingly uses object detection and segmentation when individual defects must be localized. Classification answers whether an image belongs to a class; detection additionally identifies where adulterants occur. A mature system may therefore combine image-level classification with object-level detection.

8. RESEARCH GAP

Several gaps motivate the proposed work. First, many agricultural vision studies focus on grain variety identification, disease classification or quality grading rather than deliberate adulteration. Second, high reported accuracy on randomly split images can be misleading if multiple photographs of the same physical sample appear in both training and test sets.

Third, models are often evaluated under a single background and lighting condition, allowing the network to learn environmental cues rather than genuine grain morphology. Fourth, few studies explicitly distinguish visual adulteration from chemical adulteration. Fifth, explainability and deployment efficiency are not always evaluated alongside predictive performance.

The proposed methodology addresses these gaps through sample-level splitting, controlled and variable imaging conditions, multiple architectures, explainability, external-condition testing and a carefully bounded claim of visual screening.

9. PROPOSED SYSTEM ARCHITECTURE

The proposed system contains seven stages: sample preparation, image acquisition, annotation, preprocessing, augmentation, deep-learning inference and decision reporting. Images are captured using a fixed overhead camera rig as well as a secondary smartphone condition for robustness testing. Each sample receives a unique physical-sample identifier before any image is taken.

The image is passed through a preprocessing pipeline that standardizes resolution and colour representation. A CNN backbone extracts features. A classification head estimates probabilities for pure grain and adulterant categories. Optionally, an object detector identifies individual foreign particles and produces bounding boxes.

The output interface should display the predicted class, confidence score, visual explanation heatmap and a warning that regulatory confirmation may be necessary.

10. PROPOSED DATASET DESIGN

A defensible dataset should be built from physically independent samples. For example, rice, wheat and pulses can be collected from multiple suppliers, batches and locations. Pure samples should be inspected before labeling. Adulterated samples should be prepared using documented, safe reference materials for research imaging; hazardous adulterants should not be created merely for dataset construction.

Recommended visible classes include pure grain, stone/pebble contamination, husk/straw, damaged or discoloured kernels, foreign grain/seed, and other visually verified extraneous matter. Grain-specific subclasses can be introduced only when sufficient samples are available.

Multiple images may be taken from each physical sample at different orientations, but all images from one sample must remain within a single data split. This prevents information leakage.

11. IMAGE ACQUISITION PROTOCOL

A standardized acquisition box can use diffuse LED illumination, a matte neutral background and a fixed camera-to-sample distance. Images should be captured at a resolution high enough to preserve kernel texture. A colour reference card can be included during calibration and removed or masked before model training.

To test real-world robustness, a second acquisition set should intentionally vary smartphone model, illumination intensity and background within predefined limits. The test set should include conditions not seen during training.

Metadata should record grain type, adulterant class, batch, supplier category, camera, lighting condition, acquisition date and physical-sample identifier.

12. PREPROCESSING

Preprocessing should be minimal enough to avoid destroying diagnostically useful information. Images can be resized to 224×224 or the native input resolution of the selected backbone. Pixel values are normalized according to the pretrained model's requirements.

Background segmentation can be evaluated but should not automatically be assumed beneficial. If segmentation is used, the same algorithm must be applied consistently during deployment. Colour correction may reduce lighting variation, but aggressive normalization could remove colour cues associated with adulteration.

13. DATA AUGMENTATION

Augmentation improves robustness by exposing the model to plausible visual variation. Suitable transformations include small rotations, horizontal/vertical flips where physically meaningful, mild zoom, translation, brightness variation and limited contrast changes.

Augmentations must preserve the class label. Strong colour transformations should be avoided when colour is itself diagnostic. Augmentation is applied only to the training set, never to validation or test images.

14. DEEP LEARNING MODELS

Three transfer-learning backbones are proposed. ResNet50 provides a widely used residual architecture and a strong conventional baseline. EfficientNet-B0 balances parameter efficiency and representational capacity through compound scaling. MobileNetV3 is optimized for resource-constrained inference and is suitable for a smartphone or embedded prototype.

The pretrained classification layer is replaced with a task-specific head containing global average pooling, optional dropout and a final dense layer with softmax activation for multiclass classification. Fine-tuning can proceed in two stages: train the new head with the backbone frozen, then unfreeze selected upper layers at a lower learning rate.

15. MATHEMATICAL FORMULATION

For an input image x, the network produces logits z_k for K classes. Softmax converts these values to class probabilities: $p_k = \exp(z_k) / \Sigma_j \exp(z_j)$. The predicted class is the class with maximum probability.

Multiclass cross-entropy can be used as the primary loss: $L = -\Sigma_k y_k \log(p_k)$, where y_k is the one-hot encoded target. If class imbalance is substantial, weighted cross-entropy or focal loss can be evaluated.

Accuracy alone is insufficient. For each class, Precision = TP/(TP+FP), Recall = TP/(TP+FN), and F1 = 2PR/(P+R). Macro-F1 averages class-level F1 values equally and is therefore useful when class frequencies differ.

16. TRAINING STRATEGY

The dataset should be partitioned by physical sample, preferably using approximately 70% training, 15% validation and 15% held-out testing, with stratification by class where possible. An additional external-condition test set can assess robustness to different cameras or lighting.

AdamW or stochastic gradient descent can be used as the optimizer. Early stopping should monitor validation loss or macro-F1. Learning-rate reduction and checkpointing should be predetermined before the final test evaluation.

The test set must remain untouched during architecture selection. Hyperparameters are chosen using training and validation data only. Final performance is reported once on the locked test set.

17. EVALUATION FRAMEWORK

The primary evaluation should report overall accuracy, macro-precision, macro-recall, macro-F1 and class-wise metrics. A confusion matrix reveals which adulterants are systematically confused. Receiver-operating characteristics may be included in one-vs-rest form when appropriate.

Confidence intervals can be estimated through bootstrap resampling at the physical-sample level. This is preferable to resampling individual images when several images belong to the same sample.

For deployment, model size, inference latency and memory consumption should be reported alongside predictive metrics. A slightly less accurate lightweight model may be more useful than a computationally expensive model if the target device is a smartphone.

18. COMPARATIVE RESULTS FROM PUBLISHED STUDIES

Study / Task	Method	Input	Accuracy	Other reported metric	Use in this paper
Yang et al. (2022) – sorghum adulteration in Baijiu brewing	ResNet-based CNN	RGB sorghum grain images	93.34%	Adulteration ratio also estimated	Direct adulteration benchmark
Yang et al. (2022) – same task	SqueezeNet-based CNN	RGB sorghum grain images	87.98%	Lightweight architecture	Direct adulteration benchmark
Natarajan & Ponnusamy (2025) – rice adulteration/quality analysis	CNN	Thermal rice images	95.83%	Thermal method reported 100% in stated comparison	Related rice-adulteration benchmark
Akbarpour et al. (2026) – five maize seed genotypes	Custom CNN	Standard RGB seed images	96.67%	Precision 96.74%; sensitivity 96.67%; specificity 99.18%	Seed-purity / visual-classification benchmark

Table 1. Literature-based benchmark results relevant to grain adulteration and visual grain/seed classification. These values are reproduced as reported benchmarks from the cited studies and are not claimed as results of the present proposed model.

19. CONFUSION MATRIX AND ERROR ANALYSIS

After testing, the confusion matrix should be analyzed rather than presented only as a figure. Misclassification between pure grain and adulterated grain is particularly important because false negatives may allow suspicious samples to pass screening. False positives increase unnecessary manual or laboratory inspection.

Error analysis should inspect difficult images and classify failure causes: overlap, occlusion, small adulterant size, lighting, blur, background leakage or genuine visual similarity. These observations can guide targeted data collection.

20. EXPLAINABLE AI

Grad-CAM can visualize image regions contributing to a CNN prediction. In a valid grain classifier, high-importance regions should generally overlap grains or foreign particles rather than background corners, labels or acquisition artifacts.

Explainability does not prove that a model is correct, but it can reveal spurious shortcuts. If a network consistently attends to background colour associated with a particular class, the dataset should be redesigned before deployment.

21. OBJECT DETECTION EXTENSION

Image classification gives a sample-level decision, but industrial inspection may require localization and counting. A second-stage detector such as a YOLO-family model can be trained with bounding-box annotations for stones, husk, damaged kernels and foreign seeds.

Detection allows calculation of the number or approximate area of visible adulterant objects. However, percentage-by-weight adulteration cannot be inferred reliably from pixel area without a validated calibration model. The system should therefore avoid converting image counts directly into regulatory concentration values unless experimentally validated.

22. DEPLOYMENT ARCHITECTURE

A practical prototype can operate in three modes. A desktop laboratory mode accepts images from a fixed camera; a mobile mode captures a photograph and runs a lightweight model locally; and a centralized mode sends images to a secured server for analysis.

Offline inference is attractive for warehouses and rural procurement locations with unreliable connectivity. MobileNetV3 or a quantized model can be exported to TensorFlow Lite or ONNX-compatible runtimes. The interface should show class confidence and permit the operator to mark uncertain results for manual review.

23. RESULTS AND DISCUSSION: REPORTING PROTOCOL

No original model training was performed for this manuscript; therefore, the numerical results in Table 1 are not presented as new experimental findings. They are published benchmarks selected to demonstrate the performance range reported for related grain-adulteration and seed-vision tasks. Yang et al. (2022) reported 93.34% test accuracy with a ResNet-based model and 87.98% with SqueezeNet for sorghum adulteration detection. Akbarpour et al. (2026) reported 96.67% test accuracy, 96.74% precision, 96.67% sensitivity and 99.18% specificity for five-class RGB maize-seed classification.

The published results suggest two consistent patterns. First, stronger CNN backbones can substantially outperform lighter networks on visually subtle grain classes, as seen in the ResNet–SqueezeNet comparison reported by Yang et al. (2022). Second, controlled RGB imaging can support high classification performance without hyperspectral hardware, as demonstrated by Akbarpour et al. (2026). These benchmarks justify the proposed comparison of MobileNetV3, EfficientNet-B0 and ResNet50, but they cannot substitute for training and testing on a new, independently collected adulteration dataset.

The benchmark values also reinforce the need for careful validation. Reported accuracies in the high-80% to mid-90% range are plausible for related grain-vision tasks, but direct comparisons are limited because datasets, classes, cameras and test protocols differ. Any future experimental version of this paper should report its own locked-test results separately from these literature values and should explicitly test for duplicated images, sample leakage and background shortcuts.

24. EXPECTED SCIENTIFIC CONTRIBUTION

The principal contribution is a reproducible framework specifically bounded to visible food-grain adulteration. It integrates food-safety context with modern computer vision, prevents sample leakage, includes deployment metrics and requires explainability checks.

A second contribution is the explicit separation of screening from confirmation. The model is not presented as a replacement for FSSAI-prescribed analytical methods. Instead, it can reduce inspection workload by rapidly identifying samples that warrant closer examination.

25. PRACTICAL APPLICATIONS

- Rapid screening at grain procurement centers and warehouses.
- Quality-control assistance for rice mills, flour mills and pulse-processing units.
- Educational demonstrations in food-technology and agricultural laboratories.
- Smartphone-assisted preliminary inspection for field officers or trained quality personnel.
- Automated sorting research when integrated with conveyor cameras and actuators.
- Dataset-assisted research on grain quality, defects and foreign-matter detection.

26. LIMITATIONS

The proposed system is limited by the visual information available to the camera. It cannot reliably identify invisible chemical adulterants, toxins, pesticide residues or compositional changes that do not alter appearance. Performance may also decline when grain is heavily overlapped, images are blurred or deployment conditions differ strongly from training.

Dataset representativeness is another limitation. A model trained on samples from a small number of suppliers may not generalize to different varieties, regions or harvest seasons. Continuous external validation is therefore necessary.

27. ETHICAL AND REGULATORY CONSIDERATIONS

Food-safety predictions can affect commercial reputation and consumer decisions. An AI system should therefore avoid definitive accusations of adulteration when it has performed only visual screening. Reports should use language such as 'suspected visible foreign matter detected' and recommend confirmatory analysis where required.

Training data should be documented, versioned and audited for labeling quality. Commercial deployment should include operator training, calibration checks, model-version tracking and a clear process for challenging erroneous predictions.

28. FUTURE SCOPE

Future work can extend RGB imaging with hyperspectral or multispectral sensing to capture spectral signatures unavailable to ordinary cameras. Microscopic imaging can support detection of finer contaminants, while sensor fusion may combine image features with weight, moisture or near-infrared measurements.

Self-supervised learning may reduce dependence on large labeled datasets. Domain adaptation can improve robustness across cameras and regions. Object detection and instance segmentation can support automated sorting, while federated learning could allow institutions to improve shared models without centralizing all raw images.

29. CONCLUSION

Deep learning and computer vision provide a promising foundation for rapid, non-destructive screening of visually detectable food-grain adulteration. Their value lies in consistency, speed and the ability to learn complex morphological patterns that are difficult to encode manually.

However, strong research design is essential. Dataset leakage, uncontrolled backgrounds and invented performance figures can make a system appear more capable than it is. The proposed methodology therefore emphasizes physical-sample-level splitting, external-condition testing, multiple architectures, macro-averaged metrics, error analysis and explainability.

The most defensible role for such a system is decision support. It can identify suspicious samples and assist quality personnel, while validated laboratory methods remain necessary for chemical, microscopic or regulatory confirmation. With a representative dataset and rigorous external validation, the proposed framework could contribute to faster and more accessible grain-quality screening.

REFERENCES

1. Food Safety and Standards Authority of India. (2026). Manual of methods of analysis of foods: Cereal and cereal products (2nd ed.). FSSAI.
2. Food Safety and Standards Authority of India. (2026). Check adulteration at home / Detect Adulteration with Rapid Test (DART). FSSAI.
3. Food Safety and Standards Authority of India. Food Safety and Standards (Food Products Standards and Food Additives) Regulations, 2011, as amended.
4. Food Safety and Standards Authority of India. Food Safety and Standards (Contaminants, Toxins and Residues) Regulations, 2011, as amended.
5. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770–778.
6. Howard, A., Sandler, M., Chu, G., et al. (2019). Searching for MobileNetV3. Proceedings of the IEEE/CVF International Conference on Computer Vision, 1314–1324.
7. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning, 6105–6114.
8. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE International Conference on Computer Vision, 618–626.
9. Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6, 60. https://doi.org/10.1186/s40537-019-0197-0
10. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
11. Yang, S., Lin, Y., Li, Y., Xu, D., Zhang, S., & Peng, L. (2022). Deep neural network-based sorghum adulteration detection in Baijiu brewing. IEEE Open Journal of Instrumentation and Measurement, 1, 1–8. https://doi.org/10.1109/OJIM.2022.3190024
12. Akbarpour, O., Akbari, S., Fanourakis, D., et al. (2026). Deep learning-based seed variety classification: A case study in maize. BMC Plant Biology, 26, 105. https://doi.org/10.1186/s12870-025-07919-3
13. Natarajan, S., & Ponnusamy, V. (2025). Rice quality analysis using thermal images and deep learning. Proceedings in CEUR Workshop Proceedings.