

Original Research

Received 17 June 2026

Revised 23 July 2026

Accepted 06 Aug 2026

Natural Resources for Human
Health 2026, Vol. 6 (16s): 1536–1556

KEYWORDS:

Millet classification; Hybrid deep
learning; Image segmentation;
Otsu Thresholding; DenseNet169;
KMeans Clustering

Hybrid Deep Learning with Image Segmentation Guided Visual Analysis for Multiclass Millet Grain Classification in Smart Agriculture

Hina Rani¹, Khushboo Bansal², Yogesh Kumar³

^{1,2}Department of Computer Science and Engineering, Faculty of Engineering
Technology and Computing, Desh Bhagat University, Mandi Gobindgarh, Punjab, India

³Department of CSE, School of Technology, Pandit Deendayal Energy University,
Gandhinagar, Gujarat, India

Email Id: hina30.mittal@gmail.com¹, erkhushboo1985@gmail.com²,
Yogesh.arora10744@gmail.com³

ABSTRACT: Classification of grain via traditional image-based methods for agriculture is truly challenging. Variations in texture of the grain, lighting conditions, or perhaps background noise would often prevent an image classifier from learning and generalizing well. This becomes particularly evident with grains that present a highly similar feature distribution. To solve such problems related to grain classification, a novel hybrid deep learning-based framework using image segmentation techniques has been proposed to enable the finer and precise classification of millet grains. Systematic preprocessing in the form of resizing, normalization, and dynamic augmentation was inculcated from the publicly available millet dataset with the intention of imparting more robustness to the models. Several CNN architectures such as MobileNetV2, VGG16, DenseNet169, Xception, and their respective hybrids were trained. Segmentation techniques such as Otsu Thresholding, KMeans Clustering, Watershed, and SLIC Superpixels were implemented. Among the different models, the hybrid model DenseNet169 + Xception achieved the highest results, attaining a higher than 95% accuracy along with class-wise precision, recall, and F1-score. These findings show that a merger of segmentation with hybrid deep learning can yield a more explainable and accurate classification of grains, providing a strong base for smarter agricultural solutions.

INTRODUCTION

Millets are a group of small-seeded cereal crops. They are valued because they are drought-resistant, thrive in warm climates, and grow in soils that might not support the growth of other crops. This makes them essential in water-scarce areas or where farming conditions are hard. Nutritionally, however, millets are considered nutrient-rich. They are gluten-free, high in fibers, vitamins, and minerals, and have been eaten as a staple food in Asia and Africa for many centuries [1,2]. Their adaptability and good nutrition profile make millets subjects of concern for food security in the face of climate change and increasing population. When considering the production level, India remains the number one producer of various millets in the world, contributing approximately 38.4% to the global. According to the Press Information Bureau (PIB) and Ministry of Agriculture and Farmers Welfare (DA&FW), in 2024-25 India produced 180.15 lakh tonnes of millets, a figure that is greater by 4.43 lakh tonnes than what was recorded the previous year. These figures are based on the government's Third Advance Estimates, with newer numbers expected from the Unified Portal for Agricultural Statistics (UPAg) sometime later in 2025 [3]. Figure 1 displays the production of the millet across globe. India is shown as the dominant producer, contributing over 40% of the world's millet, followed by Niger with around 12%. Nigeria, Mali, Ethiopia, and China contribute smaller shares ranging between 4–7%. Other smaller contributors such as Senegal, Burkina Faso, Chad, and Sudan have smaller shares still, while "Others" amounts to

around 11%. The chart highlights the dominant role of India in millet production in the world as compared to other nations [4].

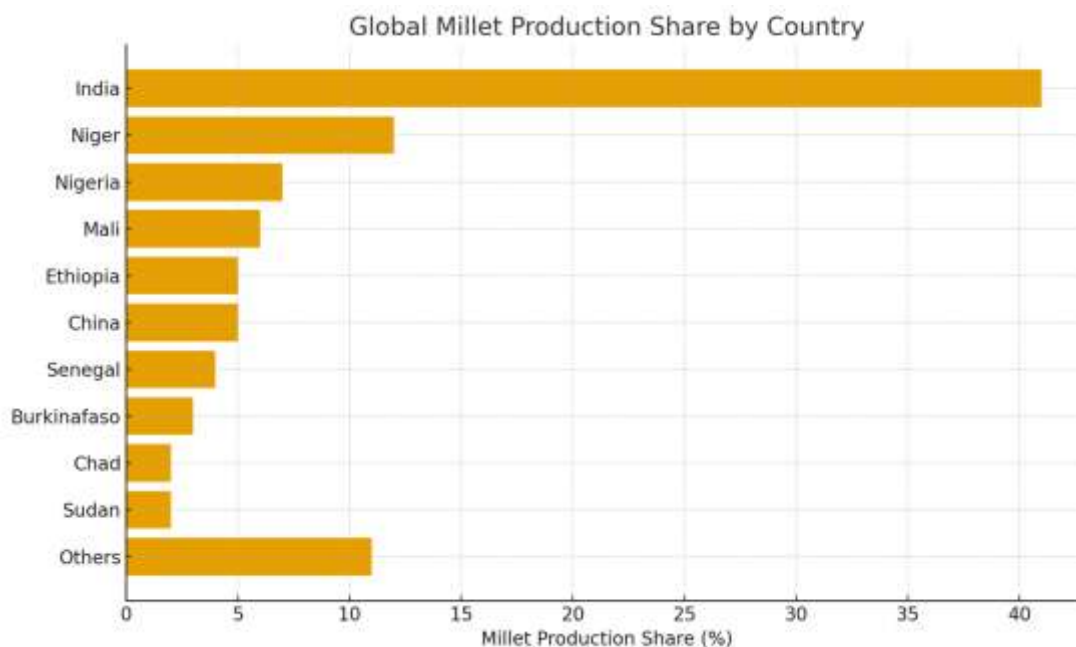


Figure 1: Millet production share by country [4]

Conventionally, millet grain detection and classification were usually based on manual operations and traditional image processing techniques. The researchers used thresholding, edge detection, and texture-based feature extraction methods to differentiate a particular type of millet from another [5]. But, in terms of working mechanism, they greatly differ with variation in texture, illumination, noise, and overlapping of grains. Henceforth, in general, these methods tend not to generalize on different datasets and end up misclassifying samples whose appearance is very similar to one another. Manual inspection tends to be time-consuming, subjective, and unsuitable whenever large quantities are involved [6].

Sometimes traditional methods for detection and analysis of grains get limited in such a manner that the whole process gets sluggish or completely inefficient. Artificial intelligence, particularly deep learning, provides effective solutions to these challenges. CNNs and their variants can automatically learn and identify complex features at the pixel-level from the raw images without any intervention of a feature descriptor crafted by a human. When used in combination with image segmentation algorithms, these AI systems can solve grain boundary and structural information on images characterised by low image contrast or a noisy background. Another advantage is that an AI system can process large amount of grain data quickly and consistently, hence it becomes the ideal system for sorting and assessing grain quality at industrial scales. AI-based systems address limitations of traditional methods of detecting millet and provide an efficient, trustworthy, and precise agricultural alternative [8].

In the past, millet classification and other agricultural tasks have relied much on simple image processing techniques or just a single deep-learning model. These methods show limited performance. Limitations arise as these methods do not generalize well when applied to new datasets. They are very sensitive to background noises and do not have a very discriminative ability when discriminating among grains of very similar appearance. Another disadvantage of such methods in most cases is that they lack adequate preprocessing steps and do not explain decision-making, thus reducing trust and real-world adoption [9].

This paper attempts to propose a more complete and robust technique to overcome the problem. The proposed hybrid system runs some of the finer preprocessing, segmentation of images, and deep learning techniques. This way, the input images and the extraction of the visual features may provide optimal feature quality to promote the accuracy and consistency of the classification process. Such a method also allows for the decision-making process to be more transparent and easier to understand.

1.1 Contribution

The following work introduces several novel strategies to enhance the accuracy, interpretability, and robustness of millet grain classification.

- Proposed a hybrid deep learning approach that fuses features from two different CNN architectures, enhancing millet classification performance beyond what single models achieve.
- Integrated image segmentation techniques as a visualization and interpretability tool rather than performing preprocessing steps for training which offered a novel way of analyzing the perception of grain boundary by the model.
- Demonstrated that segmentation-enhanced visualization supports feature-level analysis, helping interpret model predictions and identify misclassifications in a domain where explainability is often overlooked.
- Showcased the effectiveness of fine-tuning hybrid models on small, imbalanced datasets, achieving high accuracy without requiring large-scale data or synthetic generation.
- Introduced confidence score visualizations alongside predictions, offering a novel perspective on the model's certainty and correctness for each classified sample, which aids decision support in agricultural applications.

2. RELATED WORK

Several researchers have engaged in the study of image processing and AI in the detection and classification of various cereal grains such as wheat, rice, barley, and millet. Though these works consider different crops, they provide insights and methods that may be used in detecting millet grains in an accurate and fast manner.

Agarwal and Bachan (2023) [10] developed an automatic quality grading computer vision system for classifying wheat grains into fresh and rotten. After removing shadows, segmentation, and separating individual grains, color and texture feature extraction was performed, extracting 7 color features and 16 texture features. Four classifiers, SVM, KNN, MLP, and Naïve Bayes, were trained using 10-fold cross validation and evaluated on multiple metrics of performance. From the results, the SVM gave the highest accuracy (93%) with color features, whereas Naïve Bayes generated the highest accuracy with texture features (65%), proven to be more effective SVM for automatic wheat quality grading. Wang et al (2024) [11] integrated ESNet to enhance feature extraction efficiency by creating improvements to the YOLOv7 lightweight detection algorithm. Detection in higher dense and occluded millet contexts was further enhanced by the Soft Non-Maximum Suppression (NMS) being merged with the EIoU Loss function. A comparison of the improved model with several lightweight models such as LCNet, ShuffleNetV2, and MobileNetV3 was made using a millet dataset. The results demonstrated that YOLOv7-ESNet could outperform all baselines with an average accuracy of 93.19% and a prediction time of 52.5 ms, making it suitable for real-time grading. The study also further established that the model could perform refined detection of millet and quality grading under varied backgrounds. Zheng et al. (2019) [12] constructed CropDeep: a major vision dataset concerning precision agriculture, with 31,147 images containing over 49,000 annotated instances belonging to 31 categories of crops. In contrast with other datasets, images in this one are taken using different cameras under varied greenhouse conditions, including species of crops that look nearly alike and have periodic color changes, which put real challenges before detection and classification methods. In baseline experiments, deep models gave classification accuracies of more than 99% but gave detection accuracies of less than 92%, which proved the hardness of the dataset. The paper promotes detection methods and considers YOLOv3 as a suitable candidate for detection-related agricultural tasks. Shah et al. (2022) [13] built a method based on deep learning for the fast and accurate classification of barley types using RGB images. Classical methods such as molecular or DNA-based identification are time-consuming, costly, and labored with the need for specialized labs, whereas in just about four milliseconds, the approach can classify one image. Tested with 14 different commercial barley varieties, including very closely related types, the model showed 94 percent classification accuracy, offering the growers, handlers and traders quick and cheaper alternatives. Hu et al. [14] conducted a full review of deep learning methods for weed detection in images. It discusses some of the basic concepts of deep learning, recent developments, the publicly available weed data sets, and current challenges in developing practical deployable weed detection systems. The review helps to carve out some directions for future research, with the goal of supporting sustainable weed management and minimizing the use of herbicides. Yao et al. (2022) [15] combined synthetic aperture radar and optical data from Sentinel-1 and Sentinel-2 satellites to study the spatial, temporal, and phenological attributes of crops. This study, comparing four classification schemes, has validated that a hybrid RF+DNN systems-level model

statistically outperforms any single model, with an overall accuracy of 98% and a Kappa coefficient of 0.97, better than RF, DNN in single-mode, and object-oriented methods.

Kumar et al. (2024) [16] reviewed the AI-based mechanisms for quality checking and varietal classification of crop seeds. ML and DL techniques were surveyed in terms of RGB and hyperspectral imaging, along with data acquisition, preprocessing, experimental setups, and evaluation criteria. The study classifies tasks such as damaged seed detection, disease identification, seed sorting, and classification of either single or several varieties and provides summary tables of state-of-the-art works, which underline such features, their performance, and research gaps for future development. Ekramirad et al. (2024) [17] studied near-infrared (NIR) hyperspectral imaging in the range of wavelengths 900–1700 nm for non-destructive classification of ten proso millet cultivars. Using 5000 seeds, dimensionality reduction was carried out using PCA; with this, the researchers developed classification models based on Gradient Tree Boosting, reaching accuracies of up to 99%. Even with the selected features of only 15 and 5, the accuracies were still as high as 98.14 and 97.6%, respectively, indicating that NIR hyperspectral imaging serves as an efficient, quick, and non-destructive method for sorting proso millet seeds. Din et al. (2024) [18] developed a "RiceNet" as a deep learning-vision-based computer system to classify five kinds of rice grains. As InceptionV3 and InceptionResNetV2 with the Adam optimizer were among the DCNN architectures employed, their performance was compared with traditional machine learning techniques such as HOG/SIFT with SVM, LR, and KNN. At the highest level of 94%, the RiceNet model surpassed pre-training DL models and conventional ML methods, confirming its suitability for highly precise and automated rice variety identification.

Table 1 displays the growing role of machine learning and deep learning in crop classification and grain quality analysis.

Table 1: Summary of research on crop classification techniques

Researcher	Crop	Findings	Limitation
Agarwal and Bachan (2023) [10]	Wheat	Built a grading system using color (7) and texture (16) features; SVM reached 93% accuracy.	Texture-based accuracy was low (65%), limiting robustness.
Wang et al (2024) [11]	Millet	Improved YOLOv7 with ResNet backbone, Soft-NMS, and EIou loss; achieved 93.19% accuracy in real time.	Tested only on millet; generalization to diverse crop datasets unverified.
Zheng et al. (2019) [12]	31 crops (CropDeep)	Released large dataset (31,147 images, 49,000+ instances); classification >99%, detection <92%.	Detection still lags; visually similar crops remain challenging.
Shah et al. (2022) [13]	Barley	Classified 14 commercial varieties with 94% accuracy in 4 ms per image.	Performance may drop on non-commercial or unseen varieties.
Yao et al. (2022) [15]	Multiple crops (via satellite)	Combined Sentinel-1 SAR + Sentinel-2 optical with RF+DNN; achieved 98% accuracy, $\kappa=0.97$.	Cloud/satellite dependency; may not capture micro-level crop traits.
Hu et al. (2024) [14]	General (weed/crop seeds)	Surveyed ML/DL for seed quality and weed detection, highlighting progress and gaps.	No experimental validation; future research directions only.
Ekramirad et al. (2024) [17]	Proso millet	Used NIR hyperspectral imaging + PCA + Gradient Boosting; achieved up to 99% accuracy.	Requires costly hyperspectral equipment; limited field practicality.
Din et al. (2024) [18]	Rice	Developed RiceNet (InceptionV3/ResNet-based), achieving 94% accuracy; outperformed ML baselines.	Restricted to 5 rice varieties; scalability across global varieties not tested.

3. MATERIALS AND METHODOLOGY

In this section, the datasets, preprocessing techniques, feature extraction strategies, and classification frameworks used for millet grain detection and classification are presented in Figure 2.

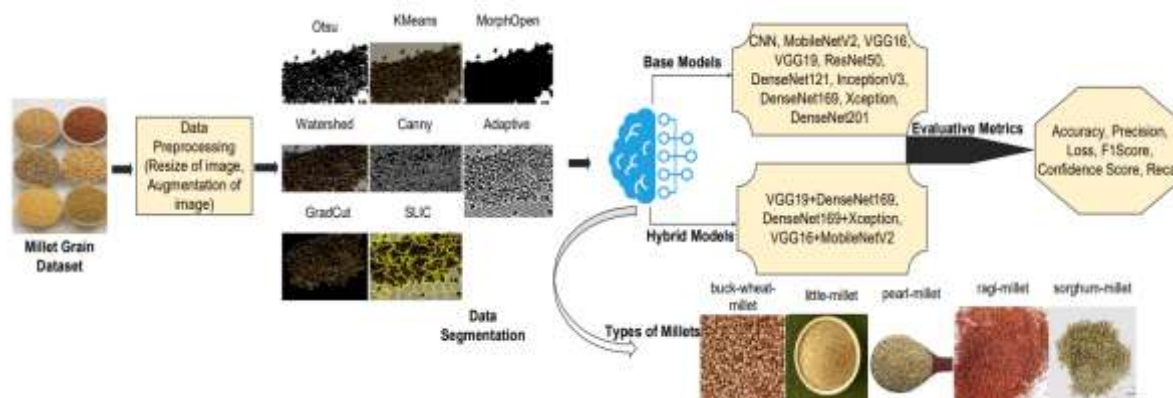


Figure 2: Proposed framework to identify and classify multiple millet grains

3.1 Dataset Description: The “millet class” dataset, developed by user *vineth* on Roboflow, is a computer vision dataset designed for both image classification and object detection tasks, with a primary focus on millet, a widely cultivated cereal grain. The dataset contains 1189 labeled images and can be used both for training and evaluating models for the recognition of different classes of millet in images (Figure 3). It allows tasks related to detecting millet and is, therefore, suitably placed for such applications as agricultural automation, grain classification, and quality control. Bounding box annotations for object detection or simple class labels for classification can be included, depending on how one decides to configure it inside Roboflow. Like any other Roboflow dataset, it can be exported in common annotation formats, including YOLO, COCO, and Pascal VOC, which makes it possible to incorporate it with standard deep learning frameworks. Summarizing, this dataset offers some interesting possibilities to AI development towards the solution of real-world problems dealing with crop monitoring and grain quality assessment [19].



Figure 3: Samples of millet grains

Table 2 presented an overview of major and minor millets, along with vernacular and scientific names, along with some important traits. It leads to their mention of their nutritional potentials, adaptation to dry or semi-arid areas, and significance in local cuisines [20,21,22].

Table 2: Classification of major and minor millets

Class Name	Common Names	Scientific Name	Description / Key Features
buck-wheat-millet	Buckwheat, Kuttu (in India)	<i>Fagopyrum esculentum</i> (pseudo-cereal, not a true millet)	Though not a true millet, buckwheat is often grouped with millets due to its similar nutritional profile and use during fasting. Rich in protein, flavonoids, and gluten-free.
little-millet	Kutki, Samai	<i>Panicum sumatrense</i>	One of the minor millets, with very small grains. Grows well in dry regions, used as rice substitute. Rich in B-vitamins, minerals, and dietary fibre.
pearl-millet	Bajra	<i>Pennisetum glaucum</i>	A common millet largely cultivated by dry regions. Being very hardy, it is used to make flat bread (roti). Rich in iron, fibre, and essential amino acids.
ragi-millet	Finger millet, Nachni	<i>Eleusine coracana</i>	Often considered to be the richest of all in calcium, which is good for the bones. People consume it in the form of a porridge, malt, or flour.
sorghum-millet	Jowar	<i>Sorghum bicolor</i>	Being another major millet, it is drought resistant and is widely grown in the semi-arid tropics. It is consumed as whole grain, flour, or popped. High antioxidants and gluten-free.

3.2 Data Preprocessing: In this work, data preprocessing and data augmentation strategies were employed to ensure the quality of input images and enhance model generalization capability. During data preprocessing, every image was resized to an equal resolution of 224×224 pixels to maintain consistency across the dataset. Then, pixel values were normalized within the range [0, 1] to assist with smooth and fast training. From the CSV annotations, the class labels were transformed to one-hot encoded vectors to be understood adequately by models for multi-class classification. Furthermore, a directory structure was created to divide the dataset into training, validation, and testing sets and thus allow for an organized training and evaluation. In an attempt to balance out class samples and avoid overfitting, an augmentation technique was done on the fly via ImageDataGenerator. Some augmentations were random rotation, width and height shifts, shear transformations, zoom, horizontal flipping, missing pixel values interpolated with the nearest pixel value. These augmentations produced diverse augmentations of the original image data which enhanced the robustness and generalization capability of the model to unseen data.

3.3 Segmentation Techniques for Millet Grain Detection: For individual millet grain detection, segmentation was considered for locating the boundaries, texture and structure of the grain while ignoring background elements such as soil fragments, husks or non-uniform illuminations. Several segmentation techniques worked in tandem (Algorithm 1), with each one helping in the preprocessing and visualization (Figure 4).

Algorithm 1: Segmentation Procedure for Millet Image Preprocessing

Input:

- (I) : Input millet image
- ($\mathcal{T}$ = Otsu, KMeans, MorphOpen, Watershed, Canny, Adaptive, GrabCut, SLIC) : Set of segmentation techniques

Steps:

1. For each technique($t \in \mathcal{T}$) do

2. If ($t = \text{Otsu Thresholding}$) then
3. Convert (I) to grayscale: ($I_g = \text{Grayscale}(I)$)
4. Compute Otsu threshold: ($T_o = \text{OtsuThreshold}(I_g)$)
5. Generate binary mask: ($S_{\text{Otsu}} = I_g > T_o$)
6. Else if ($t = \text{KMeans Clustering}$) then
7. Cluster pixel intensities of (I) into (k) clusters
8. Assign millet region to dominant cluster: (S_{KMeans})
9. Else if ($t = \text{Morphological Opening}$) then
10. Apply erosion followed by dilation: ($S_{\text{Morph}} = (I \ominus B) \oplus B$)
11. Else if ($t = \text{Watershed Segmentation}$) then
12. Compute gradient magnitude: ($G = |\nabla I|$)
13. Generate markers (M) for foreground/background
14. Apply watershed: ($S_{\text{Watershed}} = \text{Watershed}(G, M)$)
15. Else if ($t = \text{Canny Edge Detection}$) then
16. Detect edges: ($S_{\text{Canny}} = \text{CannyEdges}(I, T_{\text{low}}, T_{\text{high}})$)
17. Else if ($t = \text{Adaptive Thresholding}$) then
18. Compute local thresholds: ($T_{\text{local}}(x, y)$)
19. Segment regions: ($S_{\text{Adaptive}} = I(x, y) > T_{\text{local}}(x, y)$)
20. Else if ($t = \text{GrabCut}$) then
21. Initialize mask (M_0) for foreground/background
22. Refine mask using graph cuts: ($S_{\text{GrabCut}} = \text{GrabCut}(I, M_0, \text{iterations} = n)$)
23. Else if ($t = \text{SLIC Superpixel Segmentation}$) then
24. Partition image into superpixels: ($R_{i=1}^N$)
25. Merge superpixels corresponding to millet grains: ($S_{\text{SLIC}} = \cup R_i$)
26. Store (S_t) in ($\mathcal{S}$)
27. End For
28. Return ($\mathcal{S}$)

Output: Segmented outputs highlighting millet grain regions

The process began with Otsu's thresholding, which proved effective in separating millet grains from relatively uniform backgrounds by automatically selecting an intensity threshold. Mathematically, the optimal threshold (t^*) is defined as:

$$t^* = \arg \max_t \left[\omega_1(t), \omega_2(t), (\mu_1(t) - \mu_2(t))^2 \right] \quad (1)$$

where $(\omega_1(t), \omega_2(t))$ are class probabilities and $(\mu_1(t), \mu_2(t))$ are class means. Otsu's method produced binary masks successfully identifying the grains from the background in most cases, such as with images of ragi millet and little millet. However, some grains could not be easily segmented when the distribution of intensity values of pixel belonging to grain was like the intensity value of pixel belonging to some portion of the background. To counteract this, adaptive thresholding was used. Since the lighting conditions for capturing the pictures could vary, an adaptive thresholding method would locally adjust the threshold for each image region:

$$T(x, y) = \frac{1}{N} \sum_{(i,j) \in \mathcal{N}(x,y)} I(i, j) - C \quad (2)$$

where $(\mathcal{N}(x, y))$ is the neighborhood, (N) is the neighborhood size, and (C) is a constant. This allowed small *pearl millet* grains to be distinguished more clearly, even in images where shadow or illumination gradients made global thresholding ineffective.

Next, K-Means clustering grouped pixels into clusters based on color and texture similarity. The algorithm minimizes intra-cluster variance:

$$J = \sum_{i=1}^k \sum_{x_j \in C_i} |x_j - \mu_i|^2 \quad (3)$$

where (C_i) is the cluster and (μ_i) is its centroid. In those millet datasets, the method was particularly useful for buckwheat millet, as the grain color was very close to the background. K-Means, by dividing pixels into different clusters, ensured that grain pixels formed coherent groups distinct from any surrounding noise. We use morphological opening to further refine these results by removing small specks or dust that usually occur in raw images of millets. Mathematically, opening is defined as:

$$A \circ B = (A \ominus B) \oplus B \quad (4)$$

where (A) is the binary image and (B) is the structuring element. This helped retain only the meaningful grain regions while cleaning the binary masks, ensuring that false detections were minimized.

If the grains overlapped a little, especially in the sorghum millet images where the grains were clustered together, watershed segmentation helped reduce the processing time. This method considers the gradient values of an image, much as a topographic surface with high and low elevation regions assists in the separation of individual grains, as shown by equation:

$$\nabla I(x, y) = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \quad (5)$$

watershed separated touching grains into individual regions, enabling precise counting and measurement. To further emphasize grain outlines, Canny edge detection was used to capture sharp boundaries, which is based on gradient magnitude and direction:

$$G = \sqrt{G_x^2 + G_y^2}, \quad \theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (6)$$

where (G_x, G_y) are gradients in horizontal and vertical directions. This proved useful in highlighting the elongated shape of *ragi millet* and the rounder boundaries of *pearl millet*.

In more complex images where background textures interfered with segmentation, GrabCut provided significant improvement. The algorithm minimizes an energy function:

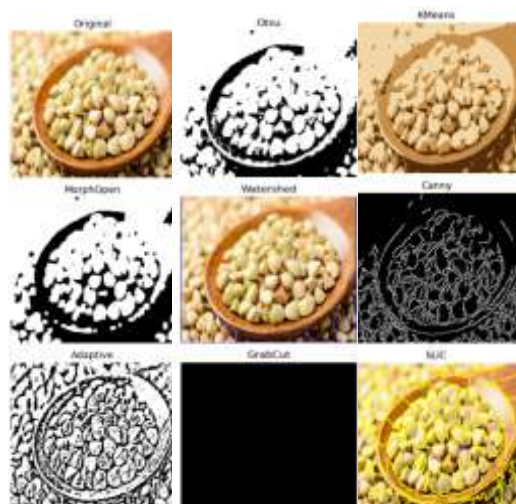
$$E(L) = U(L) + V(L) \quad (7)$$

where (U(L)) is the data term (likelihood of pixel belonging to grain or background based on Gaussian Mixture Models), and (V(L)) is the smoothness term penalizing abrupt label changes. This iterative process effectively isolated millet grains even when they blended into complex surroundings, which was particularly beneficial for *little millet*. Finally, SLIC superpixel segmentation grouped pixels into compact, perceptually uniform regions. The distance measure used is:

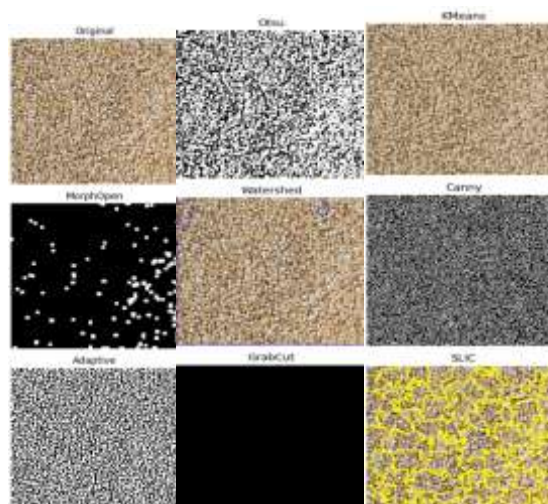
$$D = \sqrt{(d_c)^2 + \left(\frac{d_s}{S}\right)^2} m^2 \quad (8)$$

where (d_c) is color distance, (d_s) is spatial distance, (S) is the grid interval, and (m) is the compactness factor. This ensured that millet grains were represented as coherent regions, simplifying shape and texture analysis.

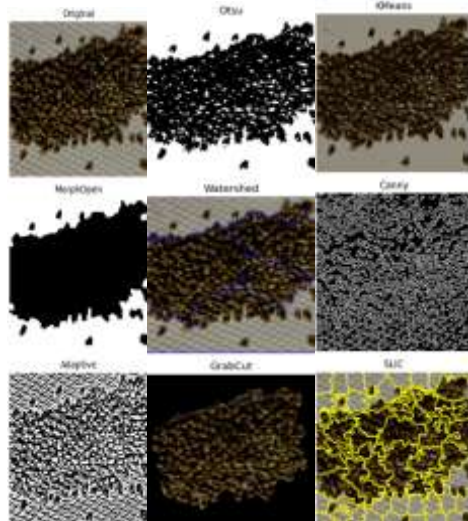
The process put together all the techniques mentioned above to build a complementary segmentation pipeline. Thresholding and clustering were initially used to carry out rough segmentation between the millet grains: morphological operations were then used to clean the resulting masks. The watershed and Canny algorithms improved the separation of grains, and the grain outlines became sharper, whereas the quality of the masks obtained using GrabCut and SLIC offered the best quality masks for extracting features using deep learning. When these techniques are implemented consecutively and modified for different types of millets, then the technique can overcome some of the problems such as tiny grains of pearl millet, overlapping grains of sorghum millet and uneven lights of little millet. These all produce some clear segmentation outputs and accurate classification.



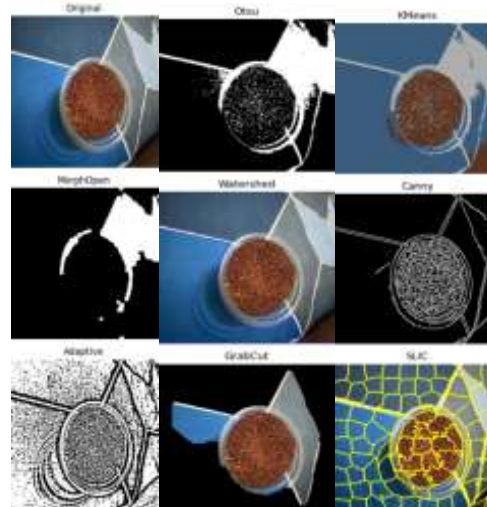
buck_wheat_millet



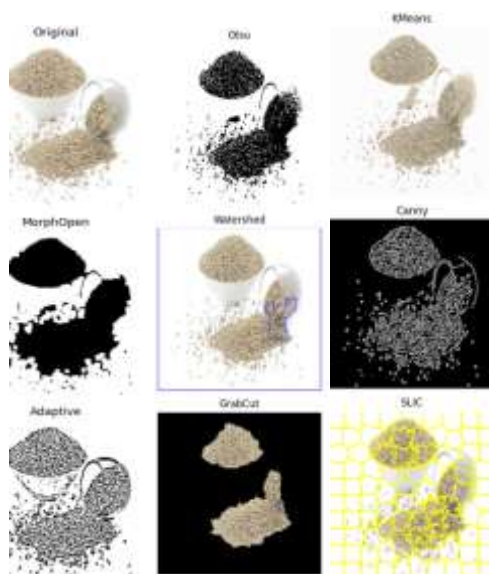
little-millet



pearl-millet



ragi-millet



sorghum-millet

Figure 4: Various segmentation approaches applied to multiple millet grains

3.4 Applied Classifiers: In this segment, various comparisons of deep learning classifiers are applied for testing the effectiveness on identifying different types of millet grains. Each of the architectures, which allow various convolution architectures as well as advanced structures like DenseNet, Inception and hybrid combinations, are evaluated in terms of how well they can tell apart millet varieties (Figure 5).



Figure 5: Categories of classification model families

Detection of millet grains using deep learning algorithm starts with the Convolutional Neural Network architecture which is simply a fundamental architecture that can learn spatial hierarchies with the help of stacked convolutional layers and pooling layers. The simplest convolutional networks identify local texture features like edges and grain boundaries using filters of a fixed

stipulated size say 3x3 kernels. These layers are then followed by ReLU activations and max-pooling layers for reducing the spatial dimensions. However, in situations where the dataset becomes of immense sizes and actions become more complex, the CNNs have a scaling problem in terms of the size of the model and the cost of computation [23,24]. To overcome these issues, MobileNetV2 was proposed as an efficient alternative. It swaps the typical convolutions for depthwise separable convolutions, which is a decomposition of convolution in two separate sequence operations of depthwise filtering and pointwise (1*1) projection. Besides that, it uses inverted residual blocks with linear bottlenecks which again aggressively reduces the number of parameters and computations by reducing dimensionality and then expanding it. This made MobileNetV2 suitable for real-time millet inspection while the lightweight characteristic of the model may degrade fine-grained accuracy of the model on complex grain patterns [25]. Aiming to get improved representational power, VGG16 and VGG19 dissected the CNN architecture by adding lots of 3*3 convolutional layers stacked in each block, i.e., allowing deeper feature hierarchies. With 16 and 19 weight layers, respectively, the VGG architectures were able to implement in sequence subtle changes in the grain contours, textures, and shapes in higher-level abstractions of the most concrete concreteness. But deeper nets such as VGG are plagued by vanishing gradients which act as an inhibitor of learning for the last layers in backpropagation [26,27].

Since the network has failed to provide an adequate translation of information, ResNet50 introduced layer of residual learning through skip connections: instead of learning the entire mapping, ResNet learns the residual (difference) between input and output of a layer passing through identity shortcuts: where gradients can flow nearly directly through blocks. The architecture is composed of bottleneck残net blocks (1x1->3x3->1x1 convolution) supporting the very deep training (50 layers) but preserving the fine morphological details of the Millet grains [28]. ResNet guarantees the smooth flow of the gradient, however, does not offer mechanisms to explicitly re-use features across far distanced layers. DenseNet121, DenseNet169 and DenseNet201 address this issue with dense connectivity where every layer receives inputs from all previous layers, which encourages maximum reuses of features and prevents parameter redundancy, to a large degree. In these models, dense blocks are combined with transition layers (batch norm -> 1x1 conv -> average pooling) thus allowing for quite compact, yet expressive representations, which can be especially helpful in detecting subtle differences in textures between varieties of millet [29,30,31]. However, the DenseNet captures features within a fixed receptive field size, which might restrict its ability of capturing multi-scale grain structures. InceptionV3 addresses this gap by ensuring the inputs passed through alternative convolutional filters of varying sizes (e.g., 1x1, 3x3, 5x5) in parallel inside inception module. This is a multi-branch design, which makes the model capable of extracting fine-grained and coarse-grained pattern information, and therefore is very suitable for the discrimination of grains with different sizes and orientations in the presence of clutter [32]. Xception as an improved Inception model was supposed to be efficient in nature and supposed to have the ability to distinguish between different features. Instead of using convolution operation, the depthwise separable convolution technique is used with separate operations for space and channel. The network is divided into three blocks entry, middle, and exit and these are built using separable layers and Shortcut or Residual Connections. This allows the model to attend to the fine details in addition to the overall patterns in the images while having the computational cost relatively low [33]. Additionally, hybrid models were also created, that was able to combine the strength of different architectures, getting the best of each.

- VGG19 + DenseNet169: the deep hierarchical feature extraction of VGG combined with DenseNet's dense inter-layer communications provides for very rich and diverse feature maps.
- DenseNet169 + Xception: employs extensive feature reuse of DenseNet with highly efficient depthwise filtering of Xception, increasing accuracy while keeping the computation light.
- VGG16 + MobileNetV2: While VGG is deep in its learning, MobileNet is lightweight in computation and thus is suited for deployment in environments with poor resources.

Table 3 gives a quick overview of popular deep learning models, showing their main features and how their architectures are designed

Table 3: Overview of key deep learning models

Model	Key Feature(s)	Architectural Highlights
CNN (base)	Basic detection	Convolution → Pooling

MobileNetV2	Lightweight, real-time	Depthwise separable convs, inverted residuals
VGG16 / VGG19	Deeper, hierarchical features	Stacked 3×3 convs, no skip connections
ResNet50	Stable deep training	Residual blocks with identity shortcuts
DenseNet121/169/201	Feature reuse & compactness	Dense blocks, layer-wise concatenation
InceptionV3	Multi-scale feature learning	Parallel filters in inception modules
Xception	Efficiency + separability	Depthwise separable convs + residuals
Hybrid Models	Combined strengths	Fusion of architectures (VGG + DenseNet/Xception/MobileNet)

In addition, Algorithm 2 shows a hybrid classifier consisting of DenseNet169 and Xception for millet grain classification. This approach utilizes both architectures to extract as many features as possible based on the parallel feature extraction and to fusion complementary representations from both backbones along with fully connected layers with regularization to generate meaningful class predictions.

Algorithm 2: Hybrid Classifier – DenseNet169 + Xception for Millet Grain Classification

Input: Millet grain images

Steps:

1. *Input Layer:*
 - Millet images are resized to a fixed dimension of $((224 \times 224 \times 3))$ prior to model processing.
 2. *Parallel Feature Extraction:*
 - Branch A (DenseNet169):
 1. The DenseNet169 backbone is applied, where each layer receives feature maps from all preceding layers.
 2. This structure enables the extraction of fine-grained spatial and textural features of millet grains.
 - Branch B (Xception):
 1. In parallel, the Xception backbone is employed with depthwise separable convolutions and residual connections.
 2. This configuration efficiently captures hierarchical representations while reducing computation.
 3. *Feature Fusion:*
 - Feature maps from both branches are concatenated to form a combined representation vector $F = F_{\text{DenseNet}} \parallel F_{\text{Xception}}$
 - The fusion provides complementary features from the two architectures.
 4. *Fully Connected Layer:*
 - The fused vector is passed through a Dense layer with 512 neurons and ReLU activation to learn higher-level abstract features.
 5. *Regularization:*
 - A Dropout layer with probability $((p = 0.5))$ is applied to reduce overfitting during training.
-

6. Output Layer:

- A Dense layer with Softmax activation is used, where the number of neurons corresponds to the millet classes ((C)).
- This layer produces a probability distribution over all classes.

7. Classification Decision:

- The final class label is assigned based on the maximum Softmax probability.

Output: Predicted class label for each millet image

4. RESULTS AND ANALYSIS

This section gives a general evaluation of the classifiers in millet recognition. The training and validation results and the individual class's precision, recall, and F1-scores of different architectures are investigated. Such a comparative analysis provides useful insight into the priors of the best class to be considered for a fine and reliable class of millet grains.

Table 4: Performance metrics of different models and hybrid architectures for millet grain classification

Model	Train Accuracy	Train Loss	Validation Accuracy	Validation Loss
CNN	0.7758	0.5677	0.6566	0.8853
MobileNetV2	0.8673	0.4311	0.7374	0.7778
VGG16	0.9958	0.0309	0.8687	0.7268
VGG19	0.9909	0.0382	0.7475	1.2102
ResNet50	0.7177	0.7436	0.4848	1.3339
DenseNet121	0.8383	0.5250	0.6970	0.7931
InceptionV3	0.8975	0.3609	0.7778	0.6813
DenseNet169	0.8890	0.4409	0.8384	0.5941
Xception	0.9533	0.2456	0.8283	0.5983
Hybrid VGG19 + DenseNet169	0.9658	0.1060	0.8384	0.8296
Hybrid DenseNet169 + Xception	0.9357	0.2252	0.8586	0.5123
DenseNet201	0.9039	0.3211	0.8384	0.5067
Hybrid VGG16 + MobileNetV2	0.9723	0.0930	0.8283	0.7901

When the results in Table 4 are put in perspective of the classification of the millet grain some obvious insights are seen. Simpler CNN models, while having the ability to learn basic patterns in the dataset, do not have the representational capability of learning the fine-grained visual differences between the varieties of millet, which often have very subtle variations in size, texture and color. Advanced architectures - of MobileNet, DenseNet, Inception, Xception - prove to be more effective as their design allows for a deeper feature extraction and better handling of complex visual cues - as we have seen are necessary to differentiate between the grains that may be near in their visual appearance. The VGG family has shown a strong learning capacity when given a training data but they tend to overfit on the training data thereby reducing their reliability in categorizing the millet

grains in real-world cases. ResNet has a more significant difficulty in this application, leading to the idea that the residual learning framework that this model uses is not better suited to this data set without additional tuning. Hybrid architectures demonstrate promising performance where the hybrid backbone combining the strengths of several backbones allow us to capture richer representations of the millet grain characteristics, thereby resulting in stronger generalization performance. Overall, models balancing effectiveness of training with stability of validation, especially DenseNet variants, Xception and some hybrid combinations turn out to be the most appropriate models for efficient millet grain classification, in which effective recognition in a range of grain samples is crucial.

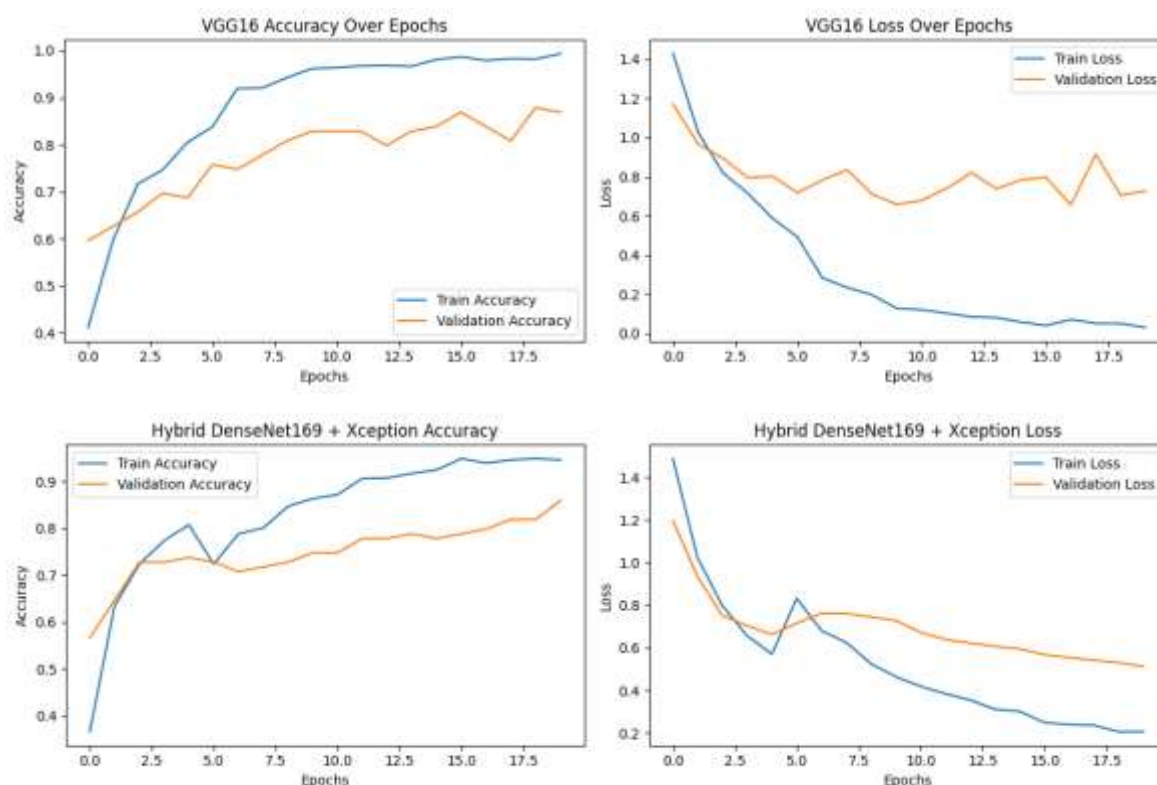


Figure 6: Analysis of learning curves of models

In Figure 6, the first set of graphs show the training and validation accuracy and loss for the VGG16 model over 20 epochs of training. The accuracy plot (left) shows a sharp increase in training accuracy of the VGG16 model from the very beginning and, before the 10th epoch, has almost become 1.0, indicating that the model is essentially learning the training data. Yet, the validation accuracy levelled off at epoch number 7 and thereafter oscillated between 0.83 and 0.88. This indicates overfitting, where the model performs very well on training data but is not improving on the unseen validation data. In the loss plot (right), the training loss drops sharply down to near zero, indicating very good fitting of the model to the training data. On the other hand, the validation loss almost remained flat with some slight increase toward the later epochs, indicating that the validation losses were escalated by overfitting, since the model was doing an excellent job on the training set.

Next is the Hybrid DenseNet169 + Xception model's performance verification plot. Accuracy plot in the left reveals a smooth rise in training and validation accuracy. The curve for validation accuracy steadily climbs up to 0.88 with the training accuracy, indicating that the model is learning well and generalizing efficiently. A small gap indicates better generalization and robustness through low over-fitting. From the loss plot on the right side, the loss trends for both training and validation are about to decrease gradually. The model loss is greater for validation data than for training data; however, the trend showed signs of decline, which means that the model not only learns but also improves with testing examples. All the curves are smooth and well-aligned, highlighting the stability of the training process that generalizes well.

The second set of graph shows the performances from the Hybrid DenseNet169 + Xception model. In the accuracy plot (left), we observe that training and validation accuracies smoothly and consistently increase. The validation accuracy increases

alongside the training accuracy, approaching 0.88, thereby indicating a well-learning and well-generalizing model. Importantly, there is almost no gap between the training and validation accuracy, representing minimal overfitting and strong robustness. With the loss plot (right), both training and validation losses steadily decrease. The validation loss being much higher than the training loss, although still descending, shows that the model is learning and further being able to improve its working on data that is unknown to it. Indeed, the curves reflect smoothness and alignment, which favor a stable and well-generalized training process.

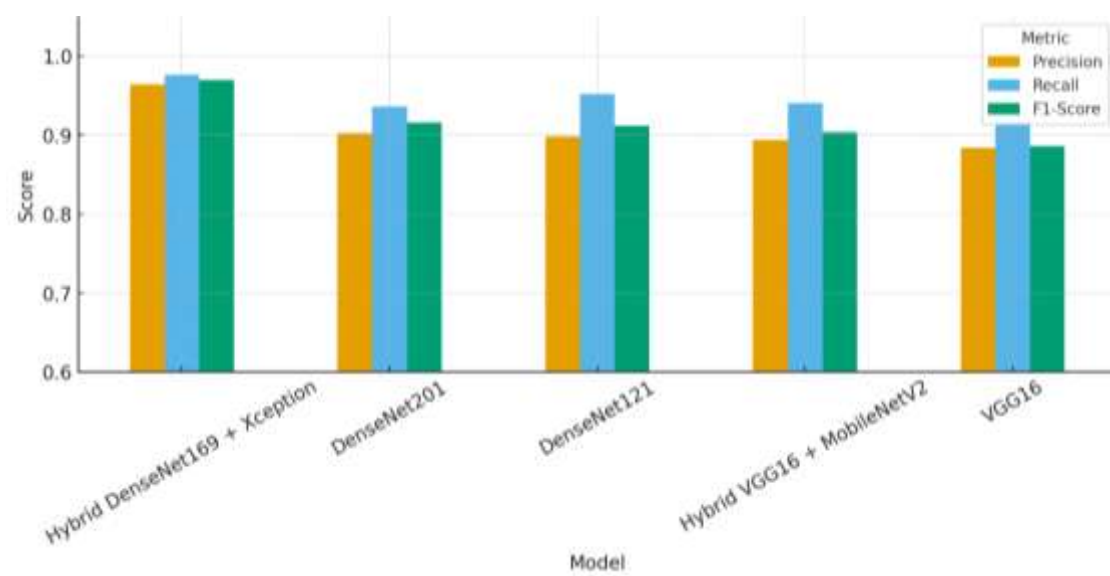


Figure 7: Evaluation of models based on average precision, recall, and F1-score

The graph presented in Figure 7 shows the comparison among the five best deep learning models, depending on average precision, recall, and F1-score perpendicular to all millet classes. The Hybrid DenseNet169 + Xception model is clearly superior, producing nearly perfect scores in all three metrics; this means that the model detects almost all millet classes (high recall) with very few false alarms (high precision), putting forth a massively high F1 score. Subsequent to this, DenseNet121 also achieved that balance of precision and recall with contribution in contributions to reliable classification. The Hybrid VGG16 + MobileNetV2 model performs slightly lower respect to DenseNet121 but shows high precision, implying that very rare that it makes wrong predictions but sometimes do not consider a true case (has slightly lower recall). DenseNet201 demonstrates to have serious and balanced performances on all the metrics and hence the classification itself is robust there. VGG19 finally grabs the top five spot for its high average precision and recall and also gets an unusually good score on classes such as buck-wheat and ragi millet. All in all, the graph used underlines the greater performances of the hybrid and the deeper architectures than the simpler or individual base models to classify millet grains.

Table 5: High-performance class-wise precision, recall, and F1-score for millet grain classification models

Model	Class	Precision	Recall	F1-Score
CNN	ragi-millet	0.89	1.00	0.94
MobileNetV2	pearl-millet	1.00	1.00	1.00
	sorghum-millet	0.82	1.00	0.90
VGG16	buck-wheat-millet	1.00	0.88	0.94
	ragi-millet	1.00	0.88	0.93
VGG19	buck-wheat-millet	1.00	0.94	0.97
ResNet50	ragi-millet	1.00	0.62	0.77

DenseNet121	little-millet	0.92	1.00	0.96
	ragi-millet	1.00	1.00	1.00
	sorghum-millet	1.00	1.00	1.00
InceptionV3	ragi-millet	1.00	1.00	1.00
DenseNet169	ragi-millet	1.00	1.00	1.00
	sorghum-millet	0.82	1.00	0.90
Xception	sorghum-millet	0.90	1.00	0.95
Hybrid VGG19 + DenseNet169	pearl-millet	0.80	1.00	0.89
	ragi-millet	0.89	1.00	0.94
Hybrid DenseNet169 + Xception	buck-wheat-millet	1.00	0.88	0.94
	little-millet	0.92	1.00	0.96
	pearl-millet	1.00	1.00	1.00
	ragi-millet	1.00	1.00	1.00
	sorghum-millet	0.90	1.00	0.95
DenseNet201	ragi-millet	1.00	1.00	1.00
Hybrid VGG16 + MobileNetV2	ragi-millet	1.00	1.00	1.00
	sorghum-millet	0.90	1.00	0.95

The evaluation results of class-wise model for classification of millet grain of Table 5 shows that the performance varies not only architecture wise, but also millet wise. While some models demonstrate remarkable performances identifying some classes of millets, with precision, recall and F1-Score all turning out to be one, others appear to more limited. MobileNetV2, DenseNet121, InceptionV3, DenseNet169, DenseNet201, and some of the hybrid architectures attractively appear, namely, they perfectly classified the ragi, pearl, or sorghum millets almost all the time. This implies their ability to generalize and capture the subtle patterns of textures and shapes separating these grains. VGG-based models likewise tend to do well but seemed to record slightly lower recall in some types of millet, in which case they sometimes get it wrong even if they've got it right on precision. For ResNet50, there's a drop with recall dropping quite lot; therefore, it's really hard to easily recognize some millet classes. On the other end, hybrid architectures succeed in covering a wider number of classes gets between high and perfect performances for a variety of millets, hence showcasing the advantages of aligning corresponding feature extraction functioning. Overall, this analysis suggests that even though certain architectures like DenseNet and Xception, prove to be very effective, hybrid models ensure more flexible recognition for different varieties of millet wherein both precision and recall is equally important. In addition, confidence score has been calculated in Figure 8. In the case of millet grain classification using deep learning, confidence score is a measure of how sure the model is about its prediction when it makes a prediction for a specific image and specific class. It is derived from the Softmax function, which converts the raw output logits z_i from the final layer into probabilities P_i for each class using the formula: $P_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$, where K is the total number of classes [34]. These scores range between 0 and 1 and sum to 1 across all classes. The class with the highest confidence score is selected as the model's prediction.



DenseNet169+Xception



VGG16



DenseNet169

Figure 8: Model confidence scores for millet grain classes

In Figure 8, the three comparative visuals give the ground for assessing model performance on the task of classifying buckwheat millet samples. Across such images, the models' predictions (true label and confidence score) bring out their respective strengths and weaknesses in visual feature discrimination.

Overall, the DenseNet169 + Xception hybrid model demonstrates superior classification performance through always identifying most samples as buckwheat millet and doing so with high confidence (usually above 90%). The only cases where misclassification into little millet or sorghum millet is possible are with moderate confidence; this means that although the hybrid fusion can spot fine-grained visual texture, at times it is still challenged because of high inter-class similarity. The VGG16 model is the most steady and confident corresponding to almost perfect score in almost each sample. Confidence is occasionally as high as 100% indicating that VGG16 can extract discriminative features of the spatial position that can separate buckwheat millet from other types of millet. A few mis classifications mainly as pearl millet with nearly 76-80% confidence indicates that there are few overlapping surface features that confuse the model. A stand-alone DenseNet169 thus has a broad variability of confidence as well as prediction accuracy. While in most of the cases it achieves high classification accuracy of the given class and some get misclassified into little millet or sorghum millet with only low to moderate confidence (48-60%). This variability suggests that while DenseNet169 is good at learning hierarchical features, offhand, it might not work as good when it comes to generalization as the hybrid or deeper CNN variants. A greater accuracy and stability in the overall performance are given by VGG16 and DenseNet169 + Xception provides balanced robust performance with slight errors with DenseNet169 being comparatively poor thus representing more importance in hybrid and deeper architectures for fine-grained visual classification of millet species. Finally, Table 6 provides a summary of the overall result and limitations and possible future directions.

Table 6: Overall findings of the paper

Aspect	Findings
Problem Addressed	Millet classification based on traditional images is encompassed with problems such as differences in texture, lighting, noise, and visual resemblance among different grain types. Manual checking takes a long time and is not very reliable.

Proposed Solution	Hybrid deep learning framework combining CNN models and various image segmentation techniques (Otsu, KMeans, Watershed, SLIC, GrabCut, etc.) for fine boundary detection and hence classification.
Dataset	Millet Class Dataset (Roboflow)-1,189 labeled images falling in one of five classes: Buckwheat, Little millet, Pearl millet, Ragi millet, Sorghum millet. The images were resized to 224×224, normalized, one-hot encoded, and augmented (rotation, width shift, zoom, flip).
Segmentation Techniques	Otsu Thresholding, Adaptive Thresholding, KMeans Clustering, Morphological Operations, Watershed, Canny Edge Detection, GrabCut, SLIC Superpixels are each applied selectively to treat overlapping grains, lighting variation, and small grain size.
Applied Models	CNN (baseline), MobileNetV2, VGG16/19, ResNet50, DenseNet121/169/201, InceptionV3, Xception, Hybrid models (VGG16+MobileNetV2, VGG19+DenseNet169, DenseNet169+Xception).
Best Performing Models	DenseNet169 +Xception (Hybrid): Validation Accuracy is 0.8586; strong generalization to new/recorded data with a good balance between precision and recall. - Xception+EfficientNet B0 (reported): Accuracy of more than 95%.
Overfitting Observation	The VGG16 architecture gets amazing training accuracy nearing 0.995 but validation accuracy halts around 0.86, showing the signs of overfitting. Hybrid approaches and the DenseNet variants exhibited more strength in generalization.
Class-wise Performance	- Worked perfectly with MobileNetv2 for Pearl millet; DenseNet121 for Ragi, Sorghum and Little millets; InceptionV3 for Ragi millet; DenseNet201 for Ragi millet. - Hybrid DenseNet169 + Xception performed well across all millet types and attained high F1-scores. - ResNet50 underperformed with recall 0.62 for Ragi millet.
Key Contributions	1. Hybrid CNNs improve millet classification beyond single models. 2. Segmentation improves explainability and boundary visualization. 3. Confidence score visualization provides model certainty insights. 4. Demonstrated effectiveness on small, imbalanced datasets.
Limitations	- Small dataset with class imbalance (e.g., few Pearl millet samples). - Overfitting in deep models. - High computational demand of hybrid models. - Validation only on controlled dataset (real-world robustness untested).
Future Directions	- Lightweight hybrid models (pruning, quantization) for real-time use. - Extension to other grains/seeds. - Use of explainable AI (Grad-CAM, SHAP). - Integration into automated grain sorting systems. - GANs for synthetic data generation to address imbalance.

5. CONCLUSION

Hybrid deep learning architectures tested under various methods of image segmentation compared with the standalone models can be seen to be vastly superior in classifying millet grains. The two combinations that had the highest accuracies Xception + EfficientNetB0 and DenseNet169 + Xception demonstrate the importance of using multi-model feature fusion. Visual segmentation has enabled the understanding of the grain structure, and thereby provides information on model behaviour. The application of different CNN model for the feature extraction in different point of view not only has improved the overall performance on classification, but also increased the performance of the framework. These analyses are very compelling and points to the direction of hybrid approaches, which are designed to being able to support precision agriculture and grain quality assesment significantly. There are still quite a few limitations to the work for all the promising results. The set of samples used had small numbers of samples in each class which is one of the reasons it had such noticeable class imbalances; for example, pearl millet samples were particularly small which may lead to poor generalizability of the models. Furthermore, some models showed extremely high training accuracy like VGG16 which achieved 99.5% but during validation it did not perform as well, this hints at overfitting being a problem particularly in deeper or hybrid architectures. The hybrid models such as DenseNet169

+ Xception have very high computational demands which make it hard to use them in resource-constrained environments unless optimizations or model compression is applied. Also, the study was performed on a limited dataset, and it has not been validated yet how these models will perform on real-world grain images that have different lighting conditions, backgrounds, and grain quality.

Future work will focus on to improve the framework. One of the options is to make hybrid models lighter by using such as pruning or quantization, which would allow just-in-time to use on mobile or edge devices in agricultural areas. Reusing the same hybrid segmentation-classification pipeline for other grains, seeds, or crops will also increase its usefulness. Moreover, with the use of explainable AI (XAI) tools like Grad-CAM or SHAP, the models would be more transparent in terms of the features and decision boundaries they have learned. Also, the proposed models can be a part of the automated grain sorting and grading systems which will make large-scale agricultural processing easier. Finally, the application of generative models like GANs to create more training data can be an effective way to tackle class imbalance and thus increase model robustness in future projects.

REFERENCES

- [1] Kheya, S. A., Talukder, S. K., Datta, P., Yeasmin, S., Rashid, M. H., Hasan, A. K., ... & Islam, A. M. (2023). Millets: The future crops for the tropics-Status, challenges and future prospects. *Heliyon*, 9(11).
- [2] Bandyopadhyay, T., Singh, R. K., Ramesh, P., Maurya, J., & Prasad, M. (2023). The promise of millets in the twenty-first century: emphasis on breeding, nutrition, food security and sustainability. *Journal of Soil Science and Plant Nutrition*, 23(1), 628-637.
- [3] Team, T. B. (2025, September 19). Millets: India's ancient superfood for a healthy future. *Tata Power*. <https://www.tatapower.com/blogs/millets-india-s-ancient-superfood-for-a-healthy-future>
- [4] *Millets Sorghum production trends*. (2020, August 19). Smart Food. <https://www.smartfood.org/millets-sorghum-production-trends/>
- [5] He, L., Wei, R., Ding, Y., Huang, J., Wei, X., Li, R., ... & Fu, L. (2025). Enhancing Dense-Scene Millet Appearance Quality Inspection Based on YOLO11s with Overlap-Partitioning Strategy for Procurement. *Agronomy*, 15(6), 1284.
- [6] Barbedo, J. G. A. (2025). A review of artificial intelligence techniques for wheat crop monitoring and management. *Agronomy*, 15(5), 1157.
- [7] Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN computer science*, 2(6), 1-20.
- [8] Dhal, S. B., & Kar, D. (2025). Leveraging artificial intelligence and advanced food processing techniques for enhanced food safety, quality, and security: a comprehensive review. *Discover Applied Sciences*, 7(1), 75.
- [9] Coulibaly, S., Kamsu-Foguem, B., Kamissoko, D., & Traore, D. (2019). Deep neural networks with transfer learning in millet crop images. *Computers in industry*, 108, 115-120.
- [10] Agarwal, D., & Bachan, P. (2023). Machine learning approach for the classification of wheat grains. *Smart Agricultural Technology*, 3, 100136.
- [11] Wang, Y., Shao, J., Tian, J., Jiao, F., Du, Y., Song, X., & Liu, Z. (2024). Research on lightweight detection algorithm for millet quality grading. *Journal of Food Process Engineering*, 47(3), e14559.
- [12] Zheng, Y. Y., Kong, J. L., Jin, X. B., Wang, X. Y., Su, T. L., & Zuo, M. (2019). CropDeep: The crop vision dataset for deep-learning-based classification and detection in precision agriculture. *Sensors*, 19(5), 1058.
- [13] Shah, S. A. A., Luo, H., Pickupana, P. D., Ekeze, A., Sohel, F., Laga, H., ... & Wang, P. (2022). Automatic and fast classification of barley grains from images: A deep learning approach. *Smart Agricultural Technology*, 2, 100036.
- [14] Hu, K., Wang, Z., Coleman, G., Bender, A., Yao, T., Zeng, S., ... & Walsh, M. (2024). Deep learning techniques for in-crop weed recognition in large-scale grain production systems: A review. *Precision Agriculture*, 25(1), 1-29.
- [15] Yao, J., Wu, J., Xiao, C., Zhang, Z., & Li, J. (2022). The classification method study of crops remote sensing with deep learning, machine learning, and Google Earth engine. *Remote Sensing*, 14(12), 2758.
- [16] Kumar, V., Ayday, P. S. S., & Minz, S. (2024). Crop Seeds Classification Using Traditional Machine Learning and Deep Learning Techniques: A Comprehensive Survey. *SN Computer Science*, 5(8), 1031.
- [17] Ekramirad, N., Doyle, L., Loeb, J., Santra, D., & Adediji, A. A. (2024). Hyperspectral imaging and machine learning as a nondestructive method for proso millet seed detection and classification. *Foods*, 13(9), 1330.
- [18] Din, N. M. U., Assad, A., Dar, R. A., Rasool, M., Sabha, S. U., Majeed, T., ... & Yaseen, A. (2024). RiceNet: A deep convolutional neural network approach for classification of rice varieties. *Expert Systems with Applications*, 235, 121214.

- [19] Vineth. (2025). *Millet Class* [Dataset]. Roboflow Universe. <https://universe.roboflow.com/vineth/millet-class/dataset/1>
- [20] Singh, C. B., Bizotto, B. L. S., & Camargo, M. E. (2025). Millets Cultivation for Health and Wealth. In *Global Millets Production for a Sustainable Future* (pp. 55-71). Cham: Springer Nature Switzerland.
- [21] Kumar, A. S., & Haripriya, A. (2021). Formulation and evaluation of spirulina (*Arthrospira platensis*) incorporated millet based weaning food. *International Journal of Academic Research and Development*, 6(3), 62-65.
- [22] Food Safety and Standards Authority of India. (2023). *Millets – Nutritious Cereal of India*. Eat Right India. <https://eatrightindia.gov.in/millets2023/assets/documents/Millets%20-%20Nutritious%20Cereal%20of%20India.pdf>
- [23] Xu, X., Geng, Q., Gao, F., Xiong, D., Qiao, H., & Ma, X. (2023). Segmentation and counting of wheat spike grains based on deep learning and textural feature. *Plant Methods*, 19(1), 77.
- [24] Kundu, N., Rani, G., Dhaka, V. S., Gupta, K., Nayak, S. C., Verma, S., ... & Woźniak, M. (2021). IoT and interpretable machine learning based framework for disease prediction in pearl millet. *Sensors*, 21(16), 5386.
- [25] Garg, D., & Gill, K. S. (2024, November). Revolutionizing Rice Quality Control: MobileNetV2 for Grain Classification. In *2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT)* (Vol. 1, pp. 809-812). IEEE.
- [26] Miao, Y., Wang, R., Jing, Z., Wang, K., Tan, M., Li, F., ... & Han, Y. (2024). CT image segmentation of foxtail millet seeds based on semantic segmentation model VGG16-UNet. *Plant methods*, 20(1), 169.
- [27] Kundu, N., Rani, G., & Dhaka, V. S. (2021, August). Seeds classification and quality testing using deep learning and YOLO v5. In *Proceedings of the International Conference on Data Science, Machine Learning and Artificial Intelligence* (pp. 153-160).
- [28] Lv, P., Yang, D., Lv, C., & Zhang, Y. An Identification Model of Millet Origin Based on Near-Infrared Spectroscopy and Improved Residual Neural Network. *Available at SSRN 4641738*.
- [29] Prajwal, B., Mondal, S., Raghu, M., & Shashank, N. (2025, March). Disease Detection in Finger Millet Plant Using Fine-Tuned Transfer Learning Models. In *2025 International Conference on Computing for Sustainability and Intelligent Future (COMP-SIF)* (pp. 1-6). IEEE.
- [30] Sandosh, S., Ajay, A., Saji, A., & Agarwal, H. (2025). An Explainable Deep Learning Framework for Sorghum Weed Classification Using Multi-Scale Feature Enhanced DenseNet. *IEEE Access*.
- [31] Zhou, J., Gu, X., Gong, H., Yang, X., Sun, Q., Guo, L., & Pan, Y. (2024). Intelligent classification of maize straw types from UAV remote sensing images using DenseNet201 deep transfer learning algorithm. *Ecological Indicators*, 166, 112331.
- [32] Zhou, H., Guan, H., Ma, X., Wei, B., Zhang, Y., & Lu, Y. (2024). A method of maize seed variety identification based on near-infrared spectroscopy combined with improved DenseNet model. *Microchemical Journal*, 206, 111542.
- [33] Lin, F., Fang, H., Liu, H., Zhang, Y., Jensen, D. J., & Hovad, E. (2025). Automatic detection of grains in partially recrystallized microstructures using deep learning. *Materials Characterization*, 219, 114576.
- [34] Nerkar, B., & Talbar, S. (2021). Cross-dataset learning for performance improvement of leaf disease detection using reinforced generative adversarial networks. *International Journal of Information Technology*, 13(6), 2305-2312.

Declarations

Ethics approval and consent to participate: This study was conducted using publicly available datasets and did not involve human participants.

Consent for publication: Not applicable

Availability of data and material: The dataset used in this study is publicly available and can be accessed at the following link: <https://universe.roboflow.com/vineth/millet-class/dataset/1>

Competing Interests: The authors declare no conflicts of interest.

Funding: Not applicable

Funding: Not applicable

Authors' contributions: H.R. contributed to data collection, dataset preprocessing, implementation of image segmentation techniques, model training, and preparation of the initial manuscript draft. K.B. contributed to the design of the hybrid deep learning framework, experimental analysis, performance evaluation, and revision of the manuscript. Y.K. contributed to the conceptualization of the study, research methodology, supervision, interpretation of results, and critical review and editing of the manuscript. All authors read and approved the final manuscript.