International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(6s) 2026 pp. 1068-1086
https://doi.org/10.70917/ijcisim-2026-3018
Article

Identification of Adulterated Wheat Varieties Using Optimized Machine Learning Algorithms

Shridhar Chini[1,2], Rajesh Yakkundimath[1,2], Girish Saunshi[1,2]

[1]Department of Computer Science and Engineering, K.L.E institute of Technology, Hubballi-580027, Karnataka, India
[3]Visvesvaraya Technological University, Belagavi-590018, Karnataka, India
E-mail address: shridharchini@gmail.com, rajeshymath@gmail.com, girishsaunshi@gmail.com,

Abstract: Adulteration in crops widely affects the production, despite maximization of yield in agriculture industry. In the traditional system, farmers discriminate wheat grains of various varieties based on visual inspection, which is tedious task. Therefore, automatic approaches for identification of adulterant wheat varieties essential. This led to take cost-effective technologies like Computer Vision (CV) and Machine Learning (ML) for the detection of adulterated wheat varieties from bulk samples. The study presented machine learning based classifiers using Back Propagation Neural Network (BPNN) and Support Vector Machine (SVM) in classifying seven adulterated wheat varieties across five adulteration percentages of 10%, 15%, 20%, 25% and 30%. Combined color and texture features are extracted from the image samples and auto-encoder decoder method is applied to facilitate feature selection. To evaluate the obtained classification accuracies, a Deep Learning (DL) approach using hybrid CNN-SVM classifier is employed. The study compares performance metrics of the models tested in classification accuracy and computational time. The proposed non-destructive automatic methods are helpful for the testing quality and authenticity of wheat grain samples.
Keywords: Neural Networks, Support Vector Machine, Adulteration detection, Combined features, Feature reduction, Hybridization.

1. Introduction

In Asia, wheat is second highly consumed and cultivated crop after paddy. Wheat trade is an important industry and primary occupation that involves a significant role in the economy of India. The enormous demand for wheat on the international market has made it feasible to export different types of wheat varieties. Wheat is grown, cultivated and consumed in many varieties in different parts of India. The significance of wheat varieties in the livelihood of many farmers brings higher returns and income generating crop. As such, it is expected for wheat exporters to be cautious about wheat crop purity and willing to maintain food security in the country. Despite maximum wheat crop productivity in the country with limited resources, exporting companies face several challenges like adulteration which widely affect the production. Considering increase in wheat productivity for economic importance an accurate detection system is essential to assess the extent of adulteration. The adulteration of wheat grain significantly impact food security, market transparency, and consumer satisfaction. Wheat grains are commonly adulterated with fabricated wheat varieties to increase economic profit. Therefore, accurate interpretation, resulting in the detection and quantification of wheat grain adulteration is essential for quality evaluation.

Several destructive methods using chemical, physical and mechanical properties are applied in the recent past to identify and quantify adulterant wheat grains. However, such methods are time consuming and hard to implement due to several environmental factors. There is a need for automatic, non-destructive wheat grain adulteration detection system using CV techniques. This task of automatic inspection would assist the farmers, traders, inspectors in quality grading of wheat grains. In this study, a methodology is developed to automatically identify and classify adulteration levels from the wheat grain images.

Copyright: © 2026 by the authors

In order to know existing image analysis techniques related to identification and classification of wheat grain adulteration levels and connected works, a literature survey is reviewed. The outline of papers is presented. In the recent years, several analysis techniques have been effectively applied for detecting wheat grain adulteration. These identification methods are classified into molecular methods, spectroscopic techniques, and non-destructive imaging based approaches. Molecular based DNA techniques provide high specificity for identifying wheat varieties. However, these methods are expensive and require laboratory infrastructure, thus making difficult for real-time and large scale inspection. Spectroscopic based hyperspectral imaging system identified wheat grain adulteration by evaluating spatial and spectral information such as grain morphology and chemical composition. Verdú Amat et al. [1] described hyperspectral imaging to differentiate and analyze adulteration of wheat grains with other cereal grains by spectral variations associated with compositional differences. Although high success rate is obtained, expensive system cost and complexity in operation of hyperspectral systems is time consuming. Recent studies have combined hyperspectral imaging with ML and DL technologies to enhance classification accuracy.

Zhu et al. [2] proposed a detection model of adulterated wheat kernel using hyperspectral imaging combined with Convolutional Neural Network (CNN). The combined spectral spatial features extracted from wheat grains achieved high identification accuracy. Bai et al. [3] demonstrated the potential of hyperspectral imaging and ML techniques for non-destructive wheat adulteration detection. Various ML classifiers such as SVM, Random Forest (RF), Partial Least Squares Discriminant Analysis (PLS-DA), and Linear Discriminant Analysis (LDA) are combined with hyperspectral imaging to detect adulterated wheat grain [4, 5]. These models demonstrated promising results in identifying adulterated wheat kernels under controlled environment. Deep learning models utilizing spectral–spatial imagery have improved classification efficiency by learning complex non-linear adulterated relational patterns [6, 7]. Several researchers investigated the feasibility of using hyperspectral imaging and ensemble learning approaches in detection of multi-adulterant wheat grain mixtures [8]. Although these methods demonstrated high accuracy, most studies analyzed on laboratory scale datasets with limited variability of grain types in predicting adulteration levels. Most of the reported works focus on qualitative identification rather than quantitative estimation in wheat adulterant image samples. Despite significant progress in detection systems for wheat grain adulteration, challenges remain in developing cost-effective, fast, and scalable imaging systems. High equipment costs, data dimensionality and complex data processing restrict the widespread adoption of hyperspectral techniques in grain markets.

Vithu and Moses [9] presented depth introduction of machine vision system used in quality evaluation of food grain, highlighting the advantages of image based techniques in terms of speed, efficiency and reliability. The study provided that Computer Vision System (CVS) can effectively analyze shape, size, color, and texture of adulterant and defected grain. Payne et al. [10] discussed early detection and prevention methods for foreign materials which can significantly reduce contamination risks in food.

Various techniques have been developed over the years to preserve food quality and detect foreign material. Traditional approaches for detecting foreign materials in food depend on manual inspection, metal detectors, and sieving methods. However, these methods are inaccurate to detecting non-metallic contaminants. Edwards [11] and Edwards and Stringer [12, 15] identified patterns in foreign material and food hygiene incidents, revealing that a significant proportion of contamination cases are associated with raw agricultural commodities such as grains. These studies explored the need for automated inspection systems in food processing systems. Computer vision techniques based on DL have been successfully applied for foreign object detection in agricultural products. Rong et al. [13] developed a CVS using DL to detect foreign objects in walnuts, demonstrating high accuracy rate under varying conditions. Similarly, Region-based Convolutional Neural Networks (R-CNN) have been employed for foreign object debris detection in complex environments, harnessing potential of DL in automated inspection systems [14].

Advanced X-ray multi-spectral imaging techniques have also been explored in food quality evaluation. Einarsson et al. [16] proposed a Sparse Discriminant Analysis (SDA) classifier for foreign object identification in multispectral X-ray images of food items, achieving quality assessment between food and contaminants. Einarsdóttir et al. [17] further demonstrated multi-modal X-ray imaging method for novelty identification of foreign objects in food, highlighting its importance over conventional vision systems for detecting low-contrast contaminants.

Recent studies have focused on non-invasive automated detection of foreign bodies in food products. Mohd Khairi et al. [18] highlighted importance of non-invasive detection methods such as terahertz imaging, hyperspectral imaging, X-ray imaging, and CV in food inspection.

Terahertz non-invasive imaging system has shown feasibility for foreign body detection. Lee et al. [19] demonstrated the potential of continuous wave terahertz imaging for detecting foreign bodies in food products, including non-metallic contaminants. Although terahertz and X-ray systems are effective techniques, application of these techniques is limited by widespread adoption in grain markets due to high equipment costs and safety considerations.

Several studies have also explored the usage of CV techniques for defect inspection in agricultural commodities. Zhang and Li [20] highlighted applications of CV for foreign matter inspection in cotton and methodologies relevant to wheat grain inspection. Toyofuku and Haff [21] discussed the role of CVS to detect foreign body in the food industry, emphasizing their advantages for real-time processing and automation.

In addition to foreign material detection, recent work has addressed quality evaluation at the grain level. He et al. [22] proposed an online detection system for naturally contaminated wheat grains using visible–near infrared (Vis–NIR) spectroscopy combined with CV, demonstrating the feasibility of integrating spectral and spatial features for wheat grain inspection.

It is revealed from the literature that much work has been carried out in identification of adulterant food variety. However, limited work has been addressed for real-time detection and quantification of wheat grain adulteration using digital images under practical operating conditions. These limitations motivate the present work, which aims to develop an automated and efficient approach for detection and quantification of wheat grain adulteration using image processing and CV techniques. In most of the existing work, color, shape, and texture are utilized as features for classification of adulterant grains. There is scope to improve the accuracy reported to identify adulterant wheat varieties using combined features with optimized ML algorithms. The proposed system is developed to assist farmers and agricultural experts to identify and quantify adulterant wheat varieties technologically. The proposed methodology, feature extraction and classification algorithms with experimental results and discussion, and conclusion sections are presented in the remaining part of the paper.

2. Proposed work

The proposed work consists of major phases such as image acquisition, feature extraction and selection, and classification of adulterated wheat images. The image data set of adulterated wheat images is created and each image is resized to 400 x 400 pixels in order to simply the task of computation and storage. Features are extracted from image data set and significant features are selected to improve the efficiency of identification. Suitable classifiers are employed for classification of unknown wheat adulterated images using training and testing methods. The methodology developed is shown in the Figure.1.

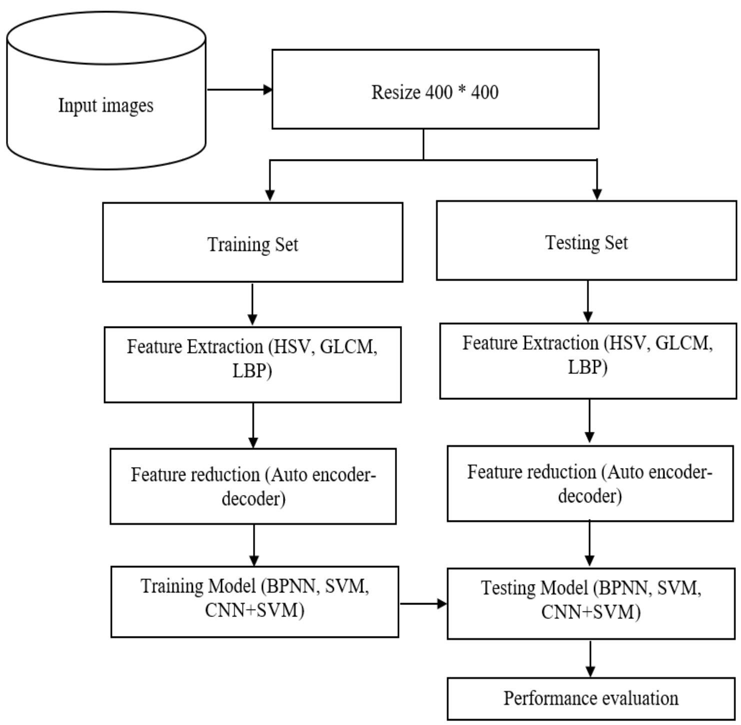


Figure.1: Schematic work flow of classifying adulterated wheat varieties

2.1 Image data set

The experiments in this study are carried out using two categories of images such as premium wheat varieties and commercially inferior wheat varieties. Due to the non-availability of a publicly available images, new data set is collected from five wheat varieties commonly grown in Dharwad region, Karnataka state, India. The images are captured using Canon EOS 1500D DSLR camera with 24.1 MP resolution. The captured images of wheat varieties considered in the proposed work are UAS-304, Emmer, DDK, Bansi, Durum. For adulteration study, seven adulterated bulk wheat samples are considered and each of the samples is prepared by mixing an identical looking and commercially inferior wheat variety with the premium wheat variety at five different adulteration levels of 10%, 15%, 20%, 25% and 30% according to commercial practice.

A total of 10,500 images is included at five different adulteration levels in each of the bulk variety, considering 2,100 images per adulteration level. The study considers a total of 10,500 (5 x 300 x 7= 10,500) from all the wheat samples at five different adulteration levels to form seven adulterated wheat varieties. The Table.1 provides summary of the adulterated wheat image dataset. The Figure .2 shows the adulterated wheat images at five percentage levels.

Table.1: Adulterated wheat image dataset description

Sl. No.	**Premium + Adulterant wheat variety**	**No. of image samples/adulteration level (10%, 15%, 20%, 25% and 30%)**	**Particulars**
1	UAS-304+Sharbati	300 x 5	Image file format: Joint Photographic Experts Group (.JPG), Image Dimension: 400 x 400, Bit depth:24
2	UAS-304+Badam	300 x 5	
3	UAS-304+Kapali	300 x 5	
4	Emmer+Kapali	300 x 5	
5	DDK+Kapali	300 x 5	
6	Bansi+Pusa Mangal (HI 8713)	300 x 5	
7	Durum+Pusa Mangal (HI 8713)	300 x 5	
Total sample images =10,500/5=2,100 images per adulteration level			

10% 15% 20% 25% 30%

Figure. 2: Wheat images adulterated at five percentage levels

2.2 Feature extraction

The color and texture features of different wheat varieties are altered significantly by adulteration. In this study, color and texture are extracted as features and used for training and testing ML models. The low level color and texture features deployed for training the ML models show high performance and suitable for image classification task. The color features are extracted using Hue, Saturation, Value (HSV) histogram representing similar color components of the human visual system, providing effective spatial color modelling. The texture features are characterized using GLCM and LBP descriptors , which encode local texture information in natural images and significantly increase classification results. A total of 9 color features is extracted by computing mean, standard deviation and skewness from each of the HSV channels.

The texture non uniformity of bulk wheat samples occurs from variations in individual grain visual appearance, scale, orientation, and illumination. To address these variations, texture feature extraction is performed with invariance to orientation, spatial scale, and gray level intensity. Texture features are extracted from grayscale images of adulterated wheat samples using Gray Level Co-occurrence Matrix (GLCM) and basic Local Binary Pattern (LBP) descriptors with an 8-neighborhood configuration. A total of 42 texture features is extracted comprising 10 GLCM based and 32 LBP based features. The LBP Histogram Group (HG1 thru HG25) features is extracted using 256 gray values from gray matrix and grouped into 25 histogram bands and the number of pixels in each band is counted. For LBP analysis, HG (HG1–HG25) features are obtained by grouping the 256 gray-level values into 25 histogram bins and computing the pixel count within each bin. The details of extracted color and texture features is given in the Table.2.

Table.2: Description of color and texture features

Sl. No.	Feature selection	Feature type	Feature set size
1	HSV histogram	Mean, Standard deviation, Skewness	9
2	GLCM	Mean, Variance, Entropy, Uniformity, Homogeneity, Inertia, Cluster shade, Cluster prominence, Maximum Probability, Correlation	10
3	LBP	Mean, Standard deviation, Smoothness, Third moment, Uniformity, Entropy, Gray level range, Histogram features (HG1 thru HG25)	32

2.3 Feature selection

An autoencoder based dimensionality reduction framework is employed for the selection of optimal feature subsets automatically from the combined 51 color and texture features. The experimental results have shown that the combined color and texture features are suitable for identification of adulteration levels in wheat grain samples.

The autoencoder architecture consists of a symmetric encoder decoder structure designed to learn a low dimensional latent representation while minimizing reconstruction error [24]. The encoder is composed of three fully connected layers with 32, 16, and 5 neurons, respectively, employing ReLU activation functions. The bottleneck layer produces a five-dimensional feature vector. The mean and standard deviation of these features are calculated and combined to obtain a final 10-dimensional feature vector. The resulting feature set is used for classification using BPNN and SVM classifiers. The decoder mirrored the encoder structure and reconstructed the original input using linear activation at the output layer. The network is trained in an unsupervised manner by minimizing the mean squared reconstruction error between the input and reconstructed output.

The network receives 51 features extracted from HSV, GLCM, and LBP descriptors. The encoder reduces these features through three fully connected layers ($51 \rightarrow 32 \rightarrow 16 \rightarrow 5$) and generates a five-dimensional representation. The decoder reconstructs the input by minimizing the reconstruction error. The mean and standard deviation of the five features are then calculated and combined to form a 10-dimensional feature vector. These features are used as input to the BPNN and SVM classifiers for adulteration level classification. The Algorithm.1 gives the steps involved in the selection of image features using the auto encoder decoder method.

Algorithm.1: Color & texture features extraction & selection in wheat images

Parameters: I= Input image dataset, R_i=i^{th} input image used for feature extraction, b=bias vectors, w=weights, h=hidden layers, L=Reconstruction loss function, $\sigma(\cdot)$ =Non-linear activation function (ReLU), ε=epoch, η=learning rate, n= number of training examples for every iteration, N=10, 500 is the total number of images.

// Input: Bulk samples of adulterated wheat varieties dataset

// Output: Reduced feature vector
Step 1.Load input dataset I={1, 2…..N}
Step 2. Prepare the labeled datasets sets for training

Step 3. Extract Relevant Features
for i= 1, 2, 3…..N
F_{hsv}= Extract_HSV_Feature(R_i)
F_{glcm}= Extract_GLCM_Feature(R_i)
F_{lbp}= Extract_LBP_Feature(R_i)
end
Step 4. Combine Features $F_V = [F_{HSV}, F_{GLCM}, F_{LBP}]$
Step 5. Optimize feature vectors using auto encoder decoder.
for $\mathbf{x}_i \in Fv^{51}, i = 1,2,\ldots,N$
Encoder: Map high-dimensional input into a low-dimensional latent space using Expression (1), (2), & (3).

$$h_1 = \sigma(W_1 x_i + b_1), W_1 \in R^{32\times 51} \quad (1)$$

$$h_2 = \sigma(W_2 h_1 + b_2), W_2 \in R^{16\times 32} \quad (2)$$

$$z_i = W_3 h_2 + b_3, W_3 \in R^{5\times 16} \quad (3)$$

where, $\mathbf{z}_i \in R^5$ denotes latent feature vector
Decoder: Reconstructs the input vector using equations…

$$\hat{h}_2 = \sigma_3(W_4 z_i + b_4)$$

$$\hat{h}_1 = \sigma_4(W_5 \hat{h}_2 + b_5)$$

$$\hat{x}_i = W_6 \hat{h}_1 + b_6$$

Where, $\hat{x}_i \in R^{51}$ is the reconstructed feature vector.
end
Step 6. Train auto encoder by minimizing the reconstruction error using Expression (4)

$$L = \frac{1}{N}\sum_{i=1}^{N} \| x_i - \hat{x}_i \|_2^2 \quad (4)$$

Step 7. Obtain latent representation for each image

$$\mathbf{z}_i \in R^5$$

Step 8. Aggregate optimized feature vectors using Expressions (5) & (6)

$$\mu = \frac{1}{N}\sum_{i=1}^{N} z_i \quad (5)$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (z_i - \mu)^2} \quad (6)$$

Step 9. Obtain final feature vector using Expression (7)

$$f = [\mu, \sigma] \in R^{10} \quad (8)$$

Stop.

2.4 Classification

The methodology is experimentally validated using classifiers for the considered 10 features obtained from auto encoder decoder. All the algorithms used in the proposed work are implemented using python based DL tool. The experiments are executed with Nvidia GeForce RTX 2060 GPU and 8GB RAM on Intel Core i5 system.

2.4.1 BPNN and SVM based classification

The BPNN and SVM classifiers are deployed to classify the adulteration levels from the images of adulterated wheat gain samples. The BPNN classifier with gradient descent based backpropagation algorithm is used to classify image samples. The SVM classier is trained with Sequential Minimal Optimization (SMO) algorithm, and classification is done using a kernel based decision function. The classifiers are trained and tested with a total of 10, 500 adulterated bulk wheat grain sample images considering seven mixed bulk wheat samples with five different percentage levels of adulteration. The parameter values used in the BPNN and SVM implementation is listed in the Table.3 & 4.

Table.3: Parameters for BPNN classifier

Parameter	Value
No. of input neurons	10
No. of hidden neurons	8
No. of output neurons	5
Hidden layer activation	ReLu
Output layer activation	Log sigmoid
Network type	Feed forward backpropagation
Training function	Gradient descent backpropagation
Learning rule	Gradient descent momentum
Performance metric	Mean Square Error (MSE)
Batch Size	32
Epochs	50
No. of hidden layers	1
Learning rate	0.01
Momentum constant	0.9

Table.4: Parameters for SVM classifier

Parameter	Value
Input features	10
Output classes	5
Kernel type	RBF
Regularization parameter (c)	1.0
Kernel scale (γ)	0.1
Margin type	Soft margin
Optimization algorithm	SMO
Loss function	Hinge loss
Feature normalization	Z-score normalization
Maximum iterations	1000
Convergence tolerance	

The dataset was randomly divided into training and testing sets with a 50:50 ratio. A total of 5,250 images were used for training and 5,250 images were used for testing. The features obtained from auto encoder decoder are used to train and test BPNN and SVM classifiers. The classification accuracy (%) is calculated using Expression (9) & (10). The process of adulteration levels classification is given in the Algorithm 2. The performance evaluation of BPNN and SVM classifiers is done with four parameters such as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative obtained using confusion matrix. Further, different performance metrics such as precision, recall, specificity, and F1-score is calculated using Expression. (11), (12), (13), & (14)[26, 27].

$$\text{Classification efficiency (\%)} = \frac{\text{Correctly classified sample images}}{\text{Total sample images for testing}} \text{ X } 100 \tag{9}$$

$$\text{Average classification efficiency (\%)} = \frac{\text{Classified correctly sum of sample images}}{\text{Total sample images}} \text{ X } 100 \tag{10}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{11}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{12}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{13}$$

$$\text{F1} = 2.\frac{\text{precision}.\text{recall}}{\text{precision} + \text{recall}} \tag{14}$$

Algorithm 2: Adulteration classification in wheat images using BPNN, SVM

Description: Feature vector ‘f’ with 10 features extracted from the images are given as input to the Algorithm 2. The significant features are computed using input layers, hidden layers and output layers with BPNN and SVM constructs an optimal maximum margin decision boundary using kernel functions to classify the images. Finally, the algorithm returns the consolidated confusion matrix of classification. The consolidated confusion matrix includes the performance of classification for the test image data set [28].

// Input: X_{train}: Bulk samples of adulterated wheat varieties training dataset, X_{test}: Bulk samples of adulterated wheat varieties testing dataset.

// Output: Consolidated assessment metrics corresponding to the classification performance.

Step 1. Load the input dataset

Step 2. Prepare the labeled datasets sets for training and testing, by randomly selecting the samples.

Step 3. Apply Algorithm.1 on the combined feature set to obtain optimized feature vector

Step 4: Train the BPNN and SVM classifiers using optimized feature vector.

Step 5: With test dataset $X_{test,}$ repeat Step 3

Step 6: Classify test images and compute model assessment metrics.

Stop.

Adulterated wheat varieties have been classified using BPNN and SVM, and their classification accuracy is depicted in the Table.5. The classification accuracies of 88.4%, 89.4%, 92.8%, 88.6%, 88.6%, 88.25%, and 89% are obtained using SVM for UAS-304+Sharbati, UAS-304+ Badam, UAS-304+ Kapali, Emmer+ Kapali, DDK+ Kapali, Bansi+ Pusa Mangal, and Durum+ Pusa Mangal adulterated wheat varieties, respectively. The classification accuracies of 82.2%, 90%, 86.2%, 81.6%, 80.2%, 81%, and 84.4% are obtained using BPNN for UAS-304+ Sharbati, UAS-304+

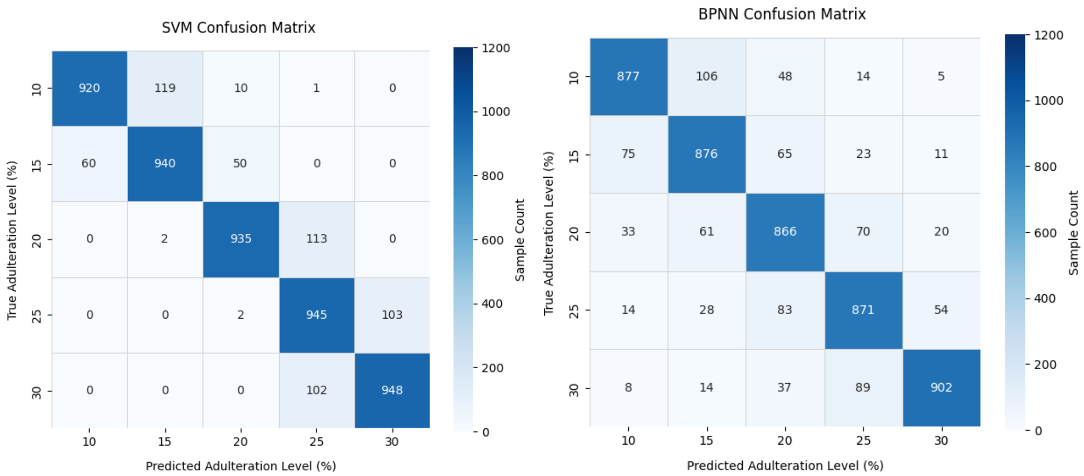
Badam, UAS-304+ Kapali, Emmer+ Kapali, DDK+ Kapali, Bansi+ Pusa Mangal, and Durum+ Pusa Mangal adulterated wheat varieties, respectively. It is observed from the Table.5 that SVM obtained best average accuracy value of 89.29% and BPNN obtained best average accuracy value of 83.66%. It is found that, the classification accuracies using SVM are more when compared to BPNN.

Table.5: Classification accuracy of BPNN and SVM classifiers

Sl. No.	Premium wheat variety	Adulterant wheat variety	Adulteration level (%)	Classifier			
				SVM		BPNN	
				Classification accuracy (%)	Average classification accuracy (%)	Classification accuracy (%)	Average classification accuracy (%)
1	UAS-304	Sharbati	10	88	88.4	82	82.2
			15	82		87	
			20	93		77	
			25	92		74	
			30	87		91	
2	UAS-304	Badam	10	88	89.4	83	90
			15	90		95	
			20	93		90	
			25	91		94	
			30	85		88	
3	UAS-304	Kapali	10	90	92.8	87	86.2
			15	94		88	
			20	96		83	
			25	93		82	
			30	91		91	
4	Emmer	Kapali	10	88	88.6	81	81.6
			15	84		77	
			20	88		79	
			25	90		83	
			30	93		88	
5	DDK	Kapali	10	91	88.6	81	80.2
			15	94		80	
			20	88		75	
			25	89		81	
			30	81		84	
6	Bansi	Pusa Mangal (HI 8713)	10	82	88.2	75	81
			15	89		77	
			20	79		83	
			25	94		88	
			30	97		82	
7	Durum	Pusa Mangal (HI 8713)	10	88	89	86	84.4
			15	89		79	
			20	89		82	
			25	88		88	
			30	91		87	

Overall average adulteration level classification accuracy (%)	**89.29**	**83.66**

The confusion matrix obtained from SVM and BPNN classifiers is shown in the Figure.3. The SVM model exhibits sharper diagonal clustering lower confusion between levels, correlating to its higher overall accuracy of 89.29%, where as BPNN has more distributed of-diagonal predictions, correlating to its lower overall accuracy of 83.66%. The Table.6 & 7. gives classification performance metrics of SVM and BPNN classifiers based on values computed using confusion matrix. The average value of performance measures precision, recall, F1-score, and specificity have been found 89.57%, 89.30%, 89.35%, and 97.33%, respectively, for the SVM across different percentage levels of wheat adulteration. The average value of performance measures precision, recall, F1-score, and specificity have been found 83.84%, 83.66%, 83.71%, and 95.91%, respectively, for the BPNN across different percentage levels of wheat adulteration. It is observed that SVM outperforms BPNN classifier.



(a) (b)

Figure.3: Confusion matrix: (a)SVM (b)BPNN

Table.6: Performance evaluation of SVM classifier

Adulteration Level	Precision (%)	Recall (%)	F1-score (%)	Specificity (%)
10%	93.88	87.62	90.64	98.56
15%	88.60	89.52	89.06	97.15
20%	93.78	89.05	91.35	98.52
25%	81.40	90.00	85.48	94.86
30%	90.20	90.29	90.24	97.55
Average (%)	89.57	89.30	89.35	97.33

Table 7: Performance evaluation of BPNN classifier

Adulteration Level	Precision (%)	Recall (%)	F1-score (%)	Specificity (%)
10%	87.09	83.52	85.27	96.90
15%	80.74	83.43	82.06	95.02
20%	78.80	82.48	80.60	94.45
25%	81.63	82.95	82.29	95.33

30%	90.93	85.90	88.34	97.86
Average (%)	83.84	83.66	83.71	95.91

2.4.2 CNN-SVM based classification

In vision based analysis and classification problems DL methods such as CNN have shown significant accuracy levels. The CNN is an Artificial Neural Network (ANN) used for the classification of adulterated wheat image samples. In present work, a CNN-SVM [28,29] hybrid model is introduced to corroborate the classification results obtained using individual BPNN and SVM classifiers. The CNN method is used for feature extraction coupled with SVM for classification allowing the network to benefit from both hierarchical feature learning and robust classification method.

The CNN accepts adulterated wheat images of size 32 x 32 x 1 as input and extracts deep features (weights) from the images. The deep features are extracted at different layers of CNN such as input layer, convolution layer, rectified linear unit (ReLU) layer, max-pooling layer and fully connected layer or output layer [30]. The Convolutional layer filter kernel of image size is 5 x 5 and the convolution is carried out with 64 filters. The convolutional layer has rectified linear unit (ReLU) nonlinear as activation function.

The convolutional layer is followed by max-pooling layer to reduce the number of parameters. These parameters are down-sampled with max-pooling of 2 x 2 regions. The fully connected layer combines all the features learnt by previous layers. We found through experimentation that soft max function at the fully connected layer fails to construct sharp decision boundaries leading to more number of misclassification rates and affecting overall accuracy. The SVM classifier produced sharp decision boundaries with margin maximization compared to sofmax function in CNN at last layer of fully connected. For classification of large dataset, SVM employs a strong kernel based function and linear learning techniques. Hence, at the last layer of fully connected instead of softmax function with the cross entropy, SVM is used to classify the images. The feature extraction is done at different layers of CNN. The extracted deep features from layers of CNN model is provided as input to the SVM classifier to perform classification of wheat adulterated images. The weight parameters are learned using Adam optimizer. The architecture of proposed CNN is shown in the Figure. 4.

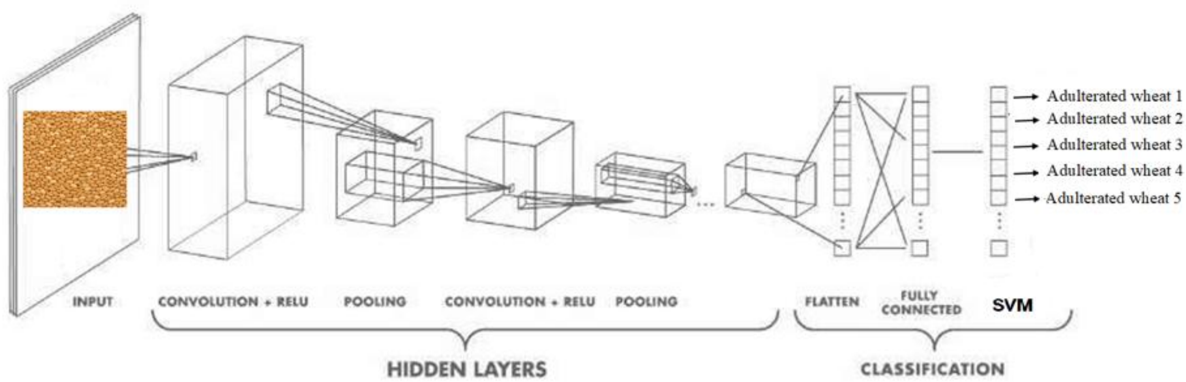


Figure. 4: Hybrid CNN-SVM architecture

In the experimentation, the number of images in the dataset is divided into 50% from each class are used for training and remaining 50% from each class is used for testing. The CNN training is performed using Adam optimization algorithm by varying number of images, learning rate, epochs, batch size. The parameter values used in the CNN-SVM implementation is listed in the Table.8.

Table.8: Parameters for CNN-SVM classifier

Parameter	Value
Input image size	32 x 32 x 1
Convolution layer 1	5 x 5 size, 64 filters, stride 1
Activation function	ReLU

Pooling layer 1	Max pooling, 2 x 2 size
Convolution layer 2	5 x 5 size, 64 filters, stride 1
Activation function	ReLU
Pooling layer 2	Max pooling, 2 x 2 size, stride 1
Fully connected layer	1024 hidden neurons
Dropout rate	0.5
Output layer	Fully connected
Output classes	05
Loss function	Categorical cross-entropy
Optimization algorithm	Adam
Learning rate	0.001
Batch size	32
Number of epochs	50

Algorithm 3: Adulteration classification in wheat images using CNN-SVM

Description: Features are extracted from the raw images that are given as input to the algorithm. The significant features are computed using base layers and upper layers to classify the images. Finally, the algorithm returns the consolidated confusion matrix of classification. The consolidated confusion matrix includes the performance of classification for the test image data set.

Parameters : ε: epochs, ξ: Iteration step, β: batch size; η:learning rate; w: weights, n: number of training examples for every iteration, $P(x_i = k)$: probability of input x_i associated to the predicted k^{th} class, k: classes index [30].

// Input: X_{train}: Bulk samples of adulterated wheat varieties training dataset, X_{test}: Bulk samples of adulterated wheat varieties testing dataset.

// Output: Consolidated assessment metrics corresponding to the classification performance.

Step 1. $I=\{i_1, i_2, i_3 \ldots\ldots i_k\}$ //Load input dataset, where k is the number of images in the dataset I.

Step 2. Convert each image in dataset into 32 x 32 size.

Step 3. Prepare the labeled training and testing datasets by randomly selecting the samples.

Step 4. Initialize CNN parameters (Learning rate, epochs, images covered in one iteration)

Step 5. Train the CNN and compute the initial weights

for ξ= 1 to size of dataset

for each epoch 1 to n

do

Randomly select a mini-batch from (size: β) from δ1

Forward propagation and compute the loss via μ

Back-propagation and update ω ∗with Adam optimization

end

Step 6. Extract deep feature vector using trained CNN

Step 7: Replace Softmax with SVM classifier at fully connected layer of CNN

Step 8: With test dataset X_{test}, repeat Steps 3 through 7.

Step 9: Classify test images and compute model assessment metrics.

Stop.

Adulterated wheat varieties have been classified using CNN-SVM, and their classification accuracy is depicted in the Table.5. The classification accuracies of 92.4%, 91.2%, 94.4%, 95%, 94%, 92.6%, and 93.8% are obtained for UAS-304+ Sharbati, UAS-304+ Badam, UAS-304+ Kapali, Emmer+ Kapali, DDK+ Kapali, Bansi+ Pusa Mangal, and Durum+ Pusa Mangal adulterated wheat varieties, respectively. It is observed from the Table.9 that CNN-SVM obtained best average accuracy value of 93.28%. It is found that the classification accuracies using CNN-SVM are more when compared to SVM and BPNN classifiers.

Table.9: Classification accuracy of CNN-SVM classifier

Sl. No.	Premium wheat variety	Adulterant wheat variety	Adulteration level (%)	Classifier	
				CNN-SVM	
				Classification accuracy (%)	Average classification accuracy (%)
1	UAS-304	Sharbati	10	98	92.4
			15	92	
			20	90	
			25	93	
			30	89	
2	UAS-304	Badam	10	88	91.2
			15	96	
			20	91	
			25	93	
			30	88	
3	UAS-304	Kapali	10	98	94.4
			15	98	
			20	93	
			25	95	
			30	88	
4	Emmer	Kapali	10	93	95
			15	93	
			20	97	
			25	96	
			30	96	
5	DDK	Kapali	10	95	94
			15	92	
			20	96	
			25	92	
			30	95	
6	Bansi	Pusa Mangal (HI 8713)	10	98	92.6
			15	92	
			20	94	
			25	89	
			30	90	

7	Durum	Pusa Mangal (HI 8713)	10	89	93.8
			15	98	
			20	94	
			25	92	
			30	96	
Overall average adulteration level classification accuracy (%)				**93.28**	

The confusion matrix obtained from CNN-SVM classifiers is shown in the Figure.5. The Table.10. gives classification performance metrics of CNN-SVM classifier based on values computed using confusion matrix. The average value of performance measures precision, recall, F1-score, and specificity have been found 93.48%, 93.47%, 93.47%, and 98.36%, respectively, for the CNN-SVM classifier across different percentage levels of wheat adulteration. It is observed that CNN-SVM outperforms SVM and BPNN classifiers.

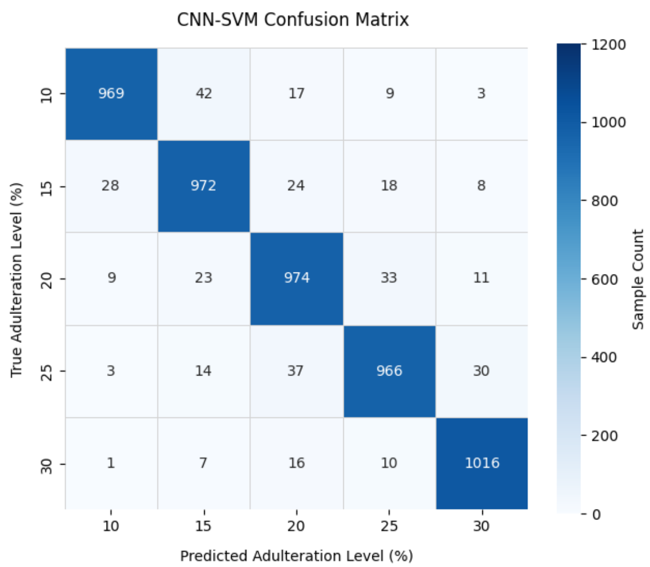


Figure.5: Confusion matrix of CNN-SVM

Table.10: Performance evaluation of CNN-SVM classifier

Adulteration Level	Precision (%)	Recall (%)	F1-score (%)	Specificity (%)
10%	95.98	93.24	94.59	98.99
15%	91.87	92.57	92.22	97.97
20%	91.20	92.76	91.97	97.80
25%	93.24	92.00	92.62	98.31
30%	95.13	96.76	95.94	98.72
Average (%)	93.48	93.47	93.47	98.36

3 Discussion

The BPNN, SVM, and CNN-SVM classifiers are trained and tested by considering seven adulterated wheat varieties across five percentage adulteration levels. Overall classification accuracy across 5 percentage adulteration levels using SVM, BPNN, CNN-SVM classifiers is shown in the Figure.6. Classification accuracy of 85.06%, 90.44%, and 94.14% is achieved using BPNN, SVM, and CNN-SVM classifiers, respectively with 10% wheat adulteration. Classification accuracy of 84.37%, 89.75%, and 94.43% is achieved using BPNN, SVM, and CNN-SVM classifiers, respectively with 15% wheat adulteration. Classification accuracy of 82.58%, 88.54%, and 93.57% is achieved using

BPNN, SVM, and CNN-SVM classifiers, respectively with 20% wheat adulteration. Classification accuracy of 83.45%, 89.47%, and 92.86% is achieved using BPNN, SVM, and CNN-SVM classifiers, respectively with 25% wheat adulteration. Classification accuracy of 92.86%, 94.27%, and 91.71% is achieved using BPNN, SVM, and CNN-SVM classifiers, respectively, with 30% wheat adulteration. It is observed from the classification results that CNN-SVM outperformed SVM and BPNN classifiers in identification of adulterated wheat based on percentage levels.

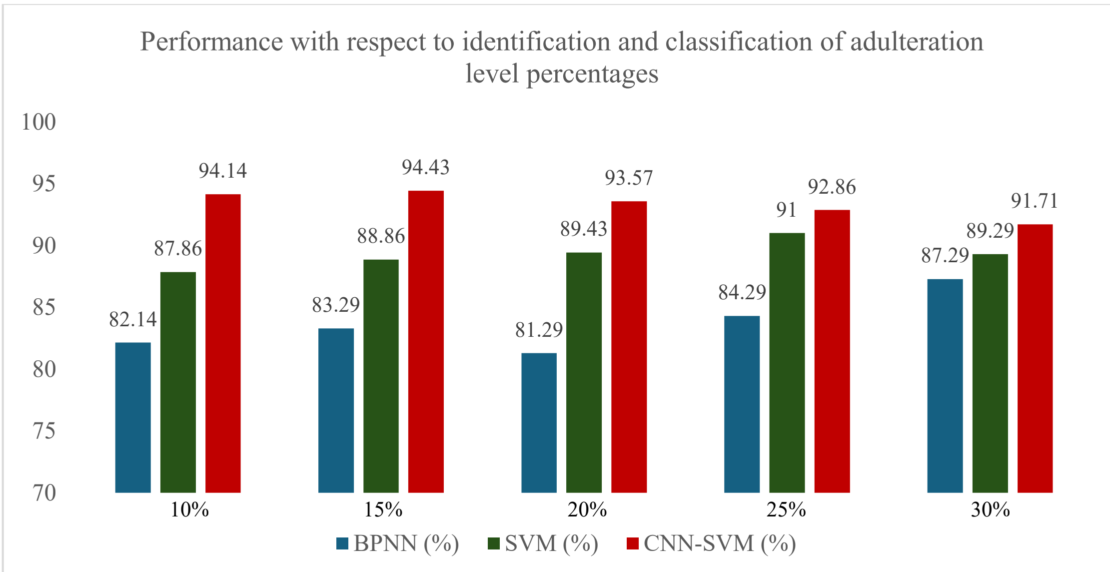


Figure.6: Classifiers performance with respect to adulteration level percentages

Overall classification accuracy considering seven adulterated wheat varieties using SVM, BPNN, CNN-SVM classifiers is shown in the Figure.7. Classification accuracy of 82.2%, 88.4%, and 92.4% is achieved using BPNN, SVM, and CNN-SVM classifiers, respectively, with UAS-304+Sharbati adulteration. Classification accuracy of 90%, 89.4%, and 91.2% is achieved using BPNN, SVM, and CNN-SVM classifiers, respectively, with UAS-304+Badam adulteration. Classification accuracy of 86.2%, 92.8%, and 94.42% is achieved using BPNN, SVM, and CNN-SVM classifiers, respectively, with UAS-304+Kapali. Classification accuracy of 81.6%, 88.6%, and 95% is achieved using BPNN, SVM, and CNN-SVM classifiers, respectively with Emmer+Kapali. Classification accuracy of 80.2%, 88.6%, and 94% is achieved using BPNN, SVM, and CNN-SVM classifiers, respectively, with DDK+Kapali. Classification accuracy of 81%, 88.2%, and 92.6% is achieved using BPNN, SVM, and CNN-SVM classifiers, respectively with Bansi+Pusa mangal. Classification accuracy of 84.4%, 89%, and 93.8% is achieved using BPNN, SVM, and CNN-SVM classifiers, respectively with Durum+Pusa Mangal. It is observed from the classification results that CNN-SVM outperformed SVM and BPNN classifiers in identification of adulterated wheat varieties.

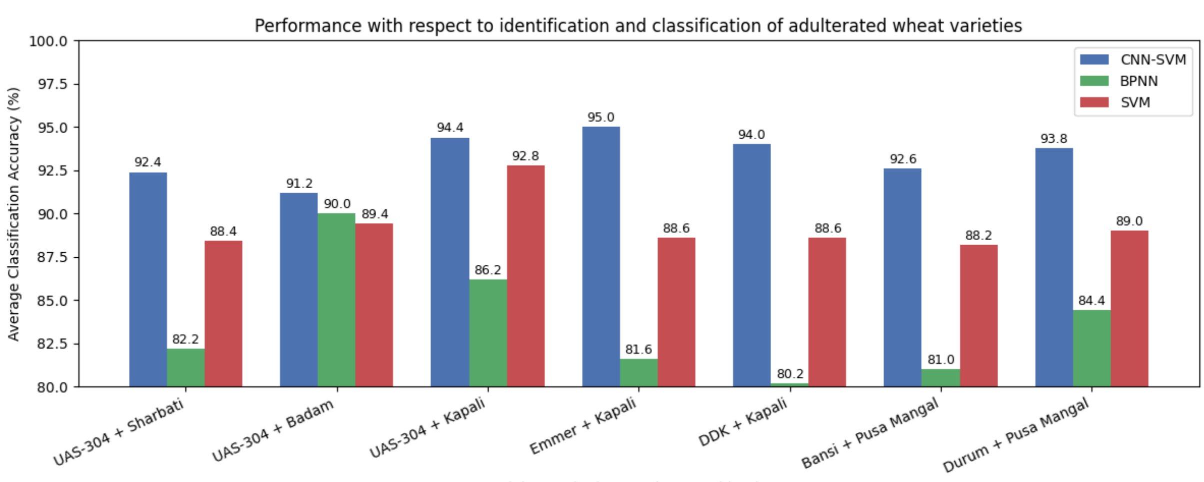


Figure.7: Performance comparison based on different classifiers

The Figure.8 shows classification performance of classifiers with respect to adulteration of wheat varieties. The value of performance measures for SVM precision, recall, F1-score have, and specificity have been found 89.29%, 90.49%, 89.87%, and 97.64%, respectively. The value of performance measures for BPNN precision, recall, F1-score, and specifity have been found 83.66%, 85.66%, 84.64%, and 96.48%, respectively. The value of performance measures for CNN-SVM precision, recall, F1-score, and specifity have been found 93.28%, 93.28%, 93.27%, and 98.32%, respectively. It has been observed that the CNN-SVM outperforms the other classifiers.

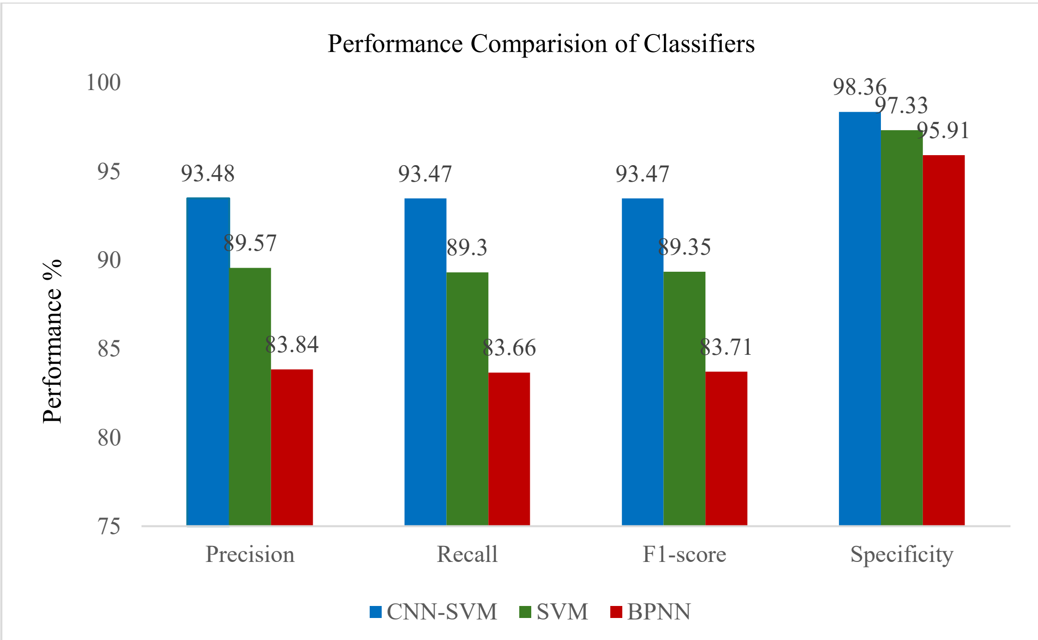


Figure.8: Average performance evaluation of classifiers

3.1 Comparative Study

The methodology is tested for a total of 10,500 images consisting of 2,100 images of each wheat adulterated percentage level. We have considered seven different wheat adulterated varieties of varying sizes. Support vector machine and BPNN is trained and tested using combined color, GLCM, and LBP features. However, after a thorough examination of input 51 features, it is found that many features have an overlapping range for a specific class, adversely affecting the classification performance.

Furthermore, the large dimensions of the input feature vector cause the curse of dimensionality, which adversely affects the generalization ability of ML algorithms. Therefore, auto encoder decoder based feature selection technique to select the optimal features for classification is utilized. The significant improvement in the performance is due to the fact that the auto encoder decoder based feature selection removes the non-contributing features and helps in finetuning the hyperparameter to build a generalized model. The selected optimal features are further employed to re-train the algorithms, and the results are presented. From the obtained results it is observed that SVM outperformed BPNN for classification of adulterated wheat varieties. The promising performance of the SVM classifier is due to the fact that the extracted color and texture feature efficiently encode the abnormal responses of the adulterated patterns in the wheat images.

Besides the promising performance of SVM and BPNN, the experiments are conducted on the dataset using hybrid CNN with SVM to corroborate the results obtained from individual classifiers. The hybrid CNN-SVM method gives significant classification performance compared to SVM and BPNN.

The comparative study based on experimentation results obtained using BPNN and SVM with reduced 10 features from auto-encoder decoder and CNN-SVM hybrid model is given in the Table.11. The SVM achieves a training accuracy of 97% and a testing accuracy of 89.29%, with an average training and testing time of 0.006–0.010

s and 0.002–0.003s per image, respectively. The BPNN achieves a training accuracy of 90% and testing accuracy of 83.66%, with an average training and testing time of 0.003–0.006s and 0.001–0.002s per image, respectively. The CNN-SVM achieves a training accuracy of 100% and testing accuracy of 93.28%, with an average training and testing time of 0.020–0.035s and 0.004–0.008s per image

Among the evaluated methods demonstrated in the present study, the CNN–SVM hybrid model is the most effective for the classification of wheat adulteration. The SVM and BPNN using auto-encoder decoder based method are computationally efficient with satisfactory results. Training accuracy was computed using the training subset, whereas testing accuracy was computed using the independent testing subset.

Table.11: Comparative study using SVM, BPNN and CNN-SVM

Methods	**(Training+Tesing [per image]) Average computation time in seconds**	**Training accuracy (%)**	**Testing accuracy (%)**
Color+Texture+Auto encoder decoder+SVM	0.006 – 0.010+0.002 – 0.003	97	89.29
Color+Texture+Auto encoder decoder+BPNN	0.003 – 0.006 + 0.001 – 0.002	90	83.66
CNN-SVM	0.020 – 0.035+0.004 – 0.008	100	93.28

4 Conclusion

The classification task is performed on seven adulterated wheat varieties with five classes image dataset containing 2100 images of each class, giving a dataset of 10,500 images. The identification of adulterated wheat grain bulk samples based on color, LBP and GLCM features are examined. Based on the selected features using autoencoder decoder, state-of-the-art ML classifiers like SVM and BPNN effectively classified image samples. The experimentation shows acceptable results with combined color and texture features using SVM and BPNN classifiers. In order to address complexities of high dimensional space, hybrid CNN-SVM approach is proposed integrating CNN capabilities for feature extraction and SVMs strength for classification. Overall classification accuracies obtained using BPNN, SVM, and CNN-SVM are 83.66%, 89.29%, and 93.28%, respectively. The classification accuracies of considered features are acceptable. It is observed that the training and testing accuracy of hybrid CNN-SVM classifier is more which in turn takes more time for classification compared to SVM and BPNN classifiers. There is scope for improvement in overall classification accuracies with the recent advances in DL techniques implemented on larger dataset with less time consumption. The work can also be extended by using different feature extraction and selection methods to perform adulteration task on other different types of agriculture crops and food products focusing on improving acceptability and reliability of the system.

Acknowledgement

We gratefully acknowledge the assistance and cooperation extended by Agricultural Experts of UAS, Dharwad, Karnataka, INDIA, who kindly assisted us in gathering datasets and shared insightful agricultural knowledge for the present study.

Credit authorship contribution statement

Shridhar Chini: Data collection, Methodology **Dr. Rajesh Y**: Conceptualization, Writing, Reviewing & Editing. **Mr. Girish Saunshi:** Data Creation, Visualization. **Dr. Guruprasad K**: Software, Validation

Data availability

Data sets/materials generated during these findings are available on reasonable requests to the corresponding author.

Ethical responsibilities of authors

All authors have read, understood, and have complied as applicable with the statement on “Ethical responsibilities of authors” as found in the Instructions for Authors.

Funding

There was no specific grant from a funding agency in the public, commercial, or nonprofit sectors for this research.

Use of AI Tools

AI tools were used solely for language refinement and grammar. No AI was used for data collection, analysis, interpretation, or content generation.

Conflict of interest

The corresponding author certifies that none of the other authors have any competing interests.

References

1. Verdú Amat S, Vásquez F, et al. 2015. Detection of adulterations with different grains in wheat using hyperspectral imaging. Food Control. 50:371–379. https://doi.org/10.1016/j.foodcont.2014.09.014
2. Zhu J, Rao Z, Ji H. 2024. Fast and simultaneous detection of wheat kernel adulteration using hyperspectral imaging and deep convolutional neural network. J Food Sci. https://doi.org/10.1111/1750-3841.16972
3. Bai G, et al. 2025. Rapid and non-destructive detection of adulterated wheat grains using hyperspectral imaging and machine learning. J Food Compos Anal. https://doi.org/10.1016/j.jfca.2024.106250
4. Liu M, He Y, Zhao X. 2022. Identification of adulterated wheat kernels using hyperspectral imaging and random forest. Spectrochim Acta A Mol Biomol Spectrosc. 278. https://doi.org/10.1016/j.saa.2022.121316
5. Chen Y, Li Q, Pan Z. 2023. Detection of wheat grain adulteration based on hyperspectral imaging and support vector machines. Food Anal Methods. 16:1123–1134. https://doi.org/10.1007/s12161-022-02449-3
6. Wang L, Yang H, Li S. 2023. Deep learning-based hyperspectral analysis for wheat grain adulteration detection. Comput Electron Agric. 206. https://doi.org/10.1016/j.compag.2023.107610
7. Kumar R, Mishra P, Lohumi S. 2024. Spectral–spatial deep learning models for non-destructive detection of wheat grain adulteration. Food Control. 151. https://doi.org/10.1016/j.foodcont.2023.109816
8. Huang Z, Zhang X, Liu Y. 2024. Detection of multi-adulterant wheat grain mixtures using hyperspectral imaging and ensemble learning. LWT Food Sci Technol. 188. https://doi.org/10.1016/j.lwt.2023.115395
9. Vithu P, Moses JA. 2016. Machine vision system for food grain quality evaluation: A review. Trends Food Sci Technol. 56:13–20. https://doi.org/10.1016/j.tifs.2016.07.008
10. Payne K, O'Bryan CA, Marcy JA, Crandall PG. 2023. Detection and prevention of foreign material in food: A review. Heliyon. 9(9):e19574. https://doi.org/10.1016/j.heliyon.2023.e19574
11. Edwards M. 2014. Food hygiene and foreign bodies. In: Hygiene in Food Processing. 2nd ed. Woodhead Publishing; 441–464. https://doi.org/10.1533/9780857098634.4.441
12. Edwards MC, Stringer MF. 2007. Observations on patterns in foreign material investigations. Food Control. 18(7):773–782. https://doi.org/10.1016/j.foodcont.2006.03.007
13. Rong D, Xie L, Ying Y. 2019. Computer vision detection of foreign objects in walnuts using deep learning. Comput Electron Agric. 162:1001–1010. https://doi.org/10.1016/j.compag.2019.05.031
14. Cao X, et al. 2018. Region-based CNN for foreign object debris detection. Sensors. 18(3):737. https://doi.org/10.3390/s18030737
15. Einarsson G, et al. 2017. Foreign object detection in multispectral X-ray images of food items. In: Proc Scand Conf Image Anal. LNCS 10269. Springer. https://doi.org/10.1007/978-3-319-59129-2_35
16. Einarsdóttir H, et al. 2016. Novelty detection of foreign objects in food using multi-modal X-ray imaging. Food Control. 67:39–47. https://doi.org/10.1016/j.foodcont.2016.02.009
17. Mohd Khairi MT, et al. 2018. Noninvasive techniques for detection of foreign bodies in food: A review. J Food Process Eng. 41(6):e12808. https://doi.org/10.1111/jfpe.12808
18. Lee YK, et al. 2012. Detection of foreign bodies in foods using continuous wave terahertz imaging. J Food Prot. 75(1):179–183. https://doi.org/10.4315/0362-028X.JFP-11-264
19. Zhang H, Li D. 2014. Applications of computer vision techniques to cotton foreign matter inspection: A review. Comput Electron Agric. 109:59–70. https://doi.org/10.1016/j.compag.2014.09.001
20. Toyofuku N, Haff RP. 2012. Computer vision for foreign body detection and removal in the food industry. In: Computer Vision Technology in the Food and Beverage Industries. Woodhead Publishing; 181–205. https://doi.org/10.1533/9780857095770.2.181
21. He X, et al. 2021. Online detection of naturally contaminated wheat grains using Vis–NIR spectroscopy and computer vision. Biosyst Eng. 201:1–10. https://doi.org/10.1016/j.biosystemseng.2020.10.017
22. Anami BS, Malvade NN, Palaiah S. 2019. Automated recognition and classification of adulteration levels from bulk paddy grain samples. Inf Process Agric. 6(1):47–60. https://doi.org/10.1016/j.inpa.2018.09.002
23. Bank D, Koenigstein N, Giryes R. 2023. Autoencoders. In: Rokach L, Maimon O, Shmueli E, eds. Machine Learning for Data Science Handbook. Springer. https://doi.org/10.1007/978-3-031-24628-9_5
24. Hecht-Nielsen R. 1989. Theory of the backpropagation neural network. In: Proc Int Joint Conf Neural Netw. IEEE; 1:593–605. https://doi.org/10.1109/IJCNN.1989.118638
25. Platt JC. 1998. Sequential minimal optimization: A fast algorithm for training support vector machines. Technical Report MSR-TR-98-14. Microsoft Research.

26. Jadhav VS, Yakkundimath R, Saunshi G, Konnurmath G. 2025. Deep learning based classification of cervical cancer stages using transfer learning models. Indian J Med Paediatr Oncol. 46(5):480–488. https://doi.org/10.1055/s-0045-1809821
27. Yakkundimath R, Saunshi G. 2023. Identification of paddy blast disease field images using multi-layer CNN models. Environ Monit Assess. 195:646. https://doi.org/10.1007/s10661-023-11123-7
28. Karakaş B, Kulluk S. 2025. A hybrid CNN-SVM algorithm for detecting manufacturing defects. IEEE Access. 13:192173–192188. https://doi.org/10.1109/ACCESS.2025.3456789
29. Agarap AFM. 2017. An architecture combining convolutional neural network and support vector machine for image classification. arXiv. https://doi.org/10.48550/arXiv.1712.03541