INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

OPEN ACCESS

Received: 17-07-2023

Accepted: 25-08-2023

Published: 27-09-2023

A Novel Algorithm for Identifying Organic Cereals by Optimal Features and Intelligent Classifiers

Ami H Bhensjaliya[1]*, **Darshankumar C Dalwadi**[2]

1 Research Scholar, Gujarat Technological University, Ahmedabad, Gujarat, India

2 Associate Professor, EC Engineering Department, Birla Vishvakarma Mahavidyalaya, Vallabh Vidyanagar, Gujarat, India

Citation: Bhensjaliya AH, Dalwadi DC (2023) A Novel Algorithm for Identifying Organic Cereals by Optimal Features and Intelligent Classifiers. Indian Journal of Science and Technology 16(36): 2920-2928. https://doi.org/10.17485/IJST/v16i36.1796

* **Corresponding author**.

ahb2812@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To develop a suitable algorithm with artificial neural network, wavelet transforms for image classification problem using cereal dataset and to make a performance comparison against different features. The major objective of this study is to perform image classification on cereals dataset images. **Methods:** This study used Statistical classification and the fuzzy logic-based methods. Image classification is performed on the cereals dataset using morphological, color, and wavelet components with different features. 70 number of images used for testing and 30 number of images for training. The performance of the working of morphological, color, and wavelet components in classifying images from the cereals dataset is compared against different features namely major axis length, minor axis length, area, centroid, and perimeter. **Findings:** The study found that (Artificial Neural Network) ANN worked better with training accuracy of 95%, testing accuracy of 91% compared to MSVM (Multiclass support vector machine) and (K - Nearest neighbor) KNN algorithm. **Novelty:** This study presents a comparative aspect of image classification using morphological, color, and wavelet components using different features since not many studies or research articles showed the performance comparison of different classification methods along with different features. Since the real-world scenarios of today require enormous data to be processed, ANN can fit well to diversify applications since they highly reduce the number of parameters to be trained that speeds up the training process. Moreover, to be specific on image classification problems they require the best and most prominent features to be detected and uncovered. This can be achieved using ANN since it has the concept of classification using training and testing at its Core. Hence, ANN is highly recommended for such image classification applications than the traditional artificial-neural-networks because of the aforementioned reasons.

Keywords: Morphological; Wavelet transform; neural networks; Statistical classifier; Fuzzy logic

1 Introduction

Many researches were carried out to classify cereal grains. Characterization models were based on morphological features [1], color features [2], or textural features [3]. Other researchers [4] have tried to combine these features for the sake of improving the efficiency of classification. Recently, the wavelet technique was integrated with cereal grains characterization [5]. This technique, developed by Mallat [6], is used in textural image analysis to make object classification more precise. The present paper is divided into four main parts. The first one will deal with the cereal image acquisition system, the second part will be devoted to presenting the classification features with their morphological, color, and wavelet components [7]. The third section will focus on the different methods used in the classification process and the last one will compare the different methods accompanied by their performance evaluation.

In the section 2, we have discussed the methodology of the work in which we describe different steps like image acquisition, image processing, binarization, edge detection, feature extraction, segmentation. Also, we describe classification methods like statistical classification and fuzzy logic based classification method.

In the section 3 we have discussed result in which rates for Barley (98.9%), the ANN classifier lead using morphological features and Tender wheat (100%) using wavelet features; whereas for Hard wheat classification (98,7%) the statistical and fuzzy logic collective classifier was the best, then the first and second ones. These two methods gave healthier results.

Based on the classification method we have observed the Fuzzy Logic performs to run about 60% faster than the second fastest (Statistical classification method). Then the fastest method, the method important to the best recognition results is 4 times slower. The most execution method as it has a good recognition rate (94%), the Statistical and Fuzzy Logic combined classification method can be measured and it takes 50% less time than the method important to the optimal recognition rate.

2 Methodology

Here we take the total 112 images in which 56 images sample for wheat. Out of that 70% image required for training and 30% testing. The methodology used for segmentation and to classify the object. Figure 1 shows the block diagram of the proposed system model.

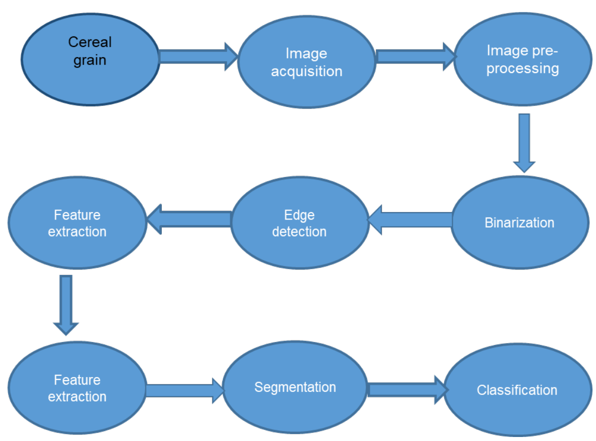


Fig 1. Block diagram of the proposed system model

2.1 Image acquisition system

2.1.1 Image acquisition device

To obtain grain images a high-resolution color camera by a USB 2.0 cable was used. The acquired images stayed at a 3.1-megapixel resolution. The grains are spread out by Light sources and were placed correspondingly over and over a glass plate. All the samples were taken at continuous camera settings, i.e., exposure time, saturation, and gamma. To reduce background pixels with the use of image subtraction the images gained are located and pre-processed. Indeed, compared to the image containing a background, the active image containing a grain sample is compared. The image we got covers the grains and a uniform background (black). This step of pre-processing makes the grains segmentation at ease and more well-organized.

2.2 Classification features

For each grain type, from the color images of the database, 122 more phonological, 18 color, and 12 wavelet features and 152 parameters are extracted.

2.2.1 Morphological features

Around the boundary of the edge, the region of interest was selected after separating the grain. From the binary images, the morphological features were acquired, containing only pixels of the grain edge. We can categorize these features as follows:

a) Grain size measurements: Length (L), width (l), width by length ratio (R1), area (S), perimeter (P), area by perimeter ratio (R), angles (GrA, PtA), and radius of curvature (Rr, Rl) of the two extremities, likelihood between the grain and the nearest ellipsis for the grain (E), mean (Sx, Sy), and standard deviation (σx, σy) of horizontal and vertical symmetry.

b) Freeman code: Features into four regions after dividing the grain image as shown in Figure 2**(a)**. We perform the freeman code for every region[8]. It's the very old contour descriptor and mostly used today. It's mainly established on the position of the pixels set, that are the very near neighbors (NN-set) of the actual pixel. From a given origin, every region is coded and starts consistency with the directions of the nearest neighbor that are symbolized in 8-connexity (coded on 3 bits) as verified in Figure 2 **(b)**. From the Freeman code 32 features are extracted; eight for every region. These features are summed up in Table 1.

Table 1. Freeman code features and their abbreviations

Region	Direct1	Direct2	Direct3	Direct4	Direct5	Direct6	Direct7	Direct8
Region1	V_{Z11}	V_{Z12}	V_{Z13}	V_{Z14}	V_{Z15}	V_{Z16}	V_{Z17}	V_{Z18}
Region2	V_{Z21}	V_{Z22}	V_{Z23}	V_{Z24}	V_{Z25}	V_{Z26}	V_{Z27}	V_{Z28}
Region3	V_{Z31}	V_{Z32}	V_{Z33}	V_{Z34}	V_{Z35}	V_{Z36}	V_{Z37}	V_{Z38}
Region4	V_{Z41}	V_{Z42}	V_{Z43}	V_{Z44}	V_{Z45}	V_{Z46}	V_{Z47}	V_{Z48}

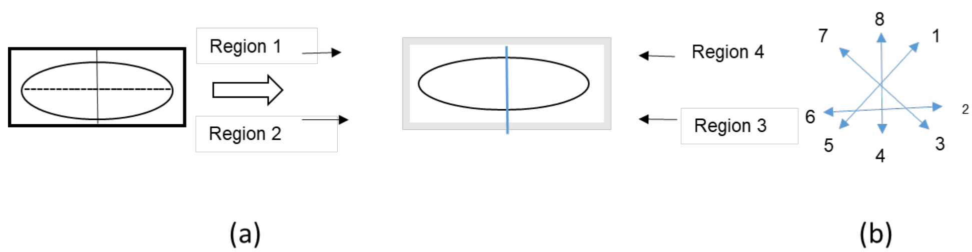


Fig 2. Freeman code extraction, (a) Dividing image into four regions to compute the Freeman code, (b) Direction codes

c) Fourier transform features: To decompose an image into its sine and cosine components, the Fourier Transform is an important image-processing tool used. This transforms on the contour pixels and generates a set of complex coefficients that characterize the shape of the contour. From these coefficients using different signatures, we extract the morphological descriptors (Equation (1)).

$$a(u) = \frac{1}{N}\sum_{K=0}^{N-1} S(K)\, exp\left(\frac{-j2puk}{N}\right] \tag{1}$$

Where,

u ∈ [0, N − 1] (N: number of points in contour)

s(k): the chosen signature

a (u): harmonic descriptors

To use the signatures that are complex, radial distance, and polar we selected first 25 harmonic coefficients from each signature that can be added to the set of morphological features.

For invariance by translation, consequently their Fourier descriptors (FD) and three signatures are used. But it was verified that they are sensitive to rotation. Invariance by rotation is then recognized by ignoring the FD phase and by as only modules of these Fourier descriptors.[9]

For the complex signature, indexing the form all descriptors except the first DC component are needed. For the contour position DC component describes, and it is unusable with the form description. By the one of second descriptors, the descriptors standardization consists of dividing their modules. The vector which indexes the form is given by the (Equation (2)).

$$F = \left[\left| \frac{FD2}{FD1} \right|, \frac{(FD3|}{(FD1|}, \ldots\ldots, \frac{(FDN-1|}{(FD1|} \right] \tag{2}$$

2.3 Color features

An isolated grain covers for each color image, on values of pixels we perform statistical parameters going to the grain. Color parameters involved: Mean value, Mean square value, Variance, Standard Deviation, Kurtosis, and Skewness of the Red, Green, and Blue primaries. We present these parameters and their abbreviations in Table 2.

Table 2. List of color parameters and their abbreviations

Color Components	Mean value	Mean square value	Variance	Standard deviation	Kurtosis	Skewness
Red	RM1	RM2	RV	RSD	RM3	RM4
Green	GM1	GM2	GV	GSD	GM3	GM4
Blue	BM1	BM2	BV	BSD	BM3	BM4
List of wavelet parameters and their abbreviations						
Matrix type		**Average value**		**Variance**		**Standard deviation**
Matrix of approximation image		MVAP		VAP		SDAP
Matrix of horizontal details		MVHD		VHD		SDHD
Matrix of vertical details		MVVD		VVD		DVD
Matrix of diagonal details		MVDD		VDD		SDDD

2.4 Wavelet features

By linear operators, the wavelet examination of an image is a multi-resolution analysis that is distinct by permitting analyzing a signal on many frequencies. Indeed, on a scale function, the signal is projected that gives a representation of the original signal at a higher scale. This projection causes a back zoom of the original signal, where the approximation is performed[(10)]. Starting from approximate coefficients the signal is remade. During the first projection we must also project the original signal on a wavelet to recover information lost. The second projection covers the details of the original signal.

In Table 2 the details of wavelet features have been reported and it resumes the chosen features. To extract the best parameters leading to an optimal classification they were statistically tested.

2.5 Classification methods

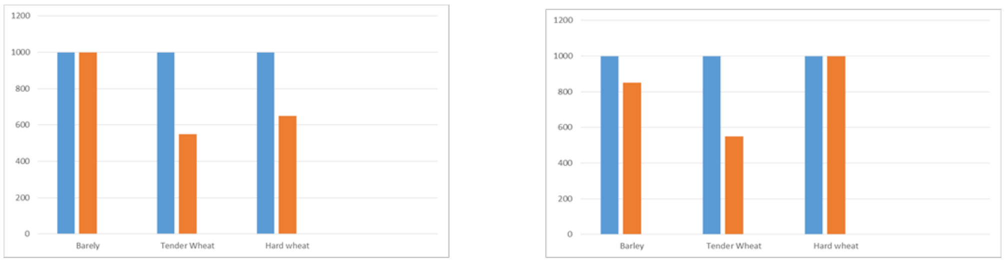
By use of different methods, we developed many methods for classification of feature extraction. The first method is a statistical classification method that uses only morphological and color features. The second method is a classification using a fuzzy logic-based method. The third is a combination between the first and the second. The last method is an artificial neural network classification method that exploits all features leading to the best classification result. For the classification of cereal grains, we present these different methods and their support.

2.5.1 Statistical classification method

From the set of samples, we attained statistics related to morphological and color features extracted from color images of grains. For all features from these statistics, we gained a distribution curve. This method operates openly on the distribution intervals of the morphological and color parameters. According to their ranks, the classification is made by successive tests on parameters. From treated samples, a mixture of grains is collected and the comprehended algorithm has been tested on images containing.

Classification results for the grain types using this method demonstrated the morphological and color features to categorize the grain types using a statistical method[(11)]. In Figure 3 classification results for the grain types using this method are demonstrated.

To work with morphological or color features (morph. 56%; color 51%) we observed that the gratitude rate for TW is weak. For HW, color features gave the best recognition rate exceeding 99.4%, but does not exceed 67% when working with morphological features. The morphological features provided us with a good classification result reaching 98.7% for Barley grains, due to their form that is dissimilar from further types of grains. The global gratitude rate for the statistical classification method is limited to 76%. When working with all the morphological and color features this is clarified by the covering that exists between the dispersal curves of grain classes.



Morphological parameters

Color parameters

Fig 3. Results of the statistical classification method when applied to morphological and color features

2.5.2 ***Fuzzy logic-based classification method***

To progress the recognition rate issued from the statistical classification method we applied a classification method based on fuzzy logic techniques due to the overlaps of sharing curves of grains types.

Classification using fuzzy logic is made according to the following steps:

- Classes' definition.
- Generation of the membership functions for every parameter.
- Development of inference rules.
- Decision making.

To the dissimilar grain types, it results from considering three parallel classes. The dissimilar parameters of every grain type Membership function are inferred from the distribution curves. By normalization of the curves and then by a Gaussian method for every curve the membership functions were conceived. The number of parameters considered depends on the number of rules[12]. The chosen norm is the max-prod. Then, the rules form is: "IF (condition1) AND (condition2) THEN (decision)".

Based on a test of identification parameters the choice of entries is considered. Table 3 demonstrates the test of the four best parameters for the classification from the set of morphological and color features associated with the fuzzy logic method.

Table 3. Test of best parameters

Parameters	Barley	Hard wheat	Tender wheat	Total		
Lsb	82.43%	77.02%	90.28%	80.94%		
E	77.82%	20.75%	90.91%	46.28%		
GrA	72.80%	70.57%	67.13%	68.62%		
RM2	64.02%	50.52%	39.18%	50.25%		
Number of neurons in the hidden layer						
	Morphological features			**Features AMF**	**Color features**	**Wavelet features**
	SM	**FC**	**FT**			
Number of features	15	32	25	122	18	10
Number of neurons	3	7	5	5	10	3

According to its recognition rates of the possible combinations of the four parameters, we select the best ones. The combinations selected are illustrated (Lsb and GrA: 83, 42%; Lsb, GrA, and RM2: 72, 71%; Lsb, GrA, E, and RM2: 68, 23%).

From this test, we chose the parameters Lsb and GrA since when joining them it gives the best recognition rate. The result of this method using the grouping OF Lsb and GrA is shown in Figure 4**(a) .**

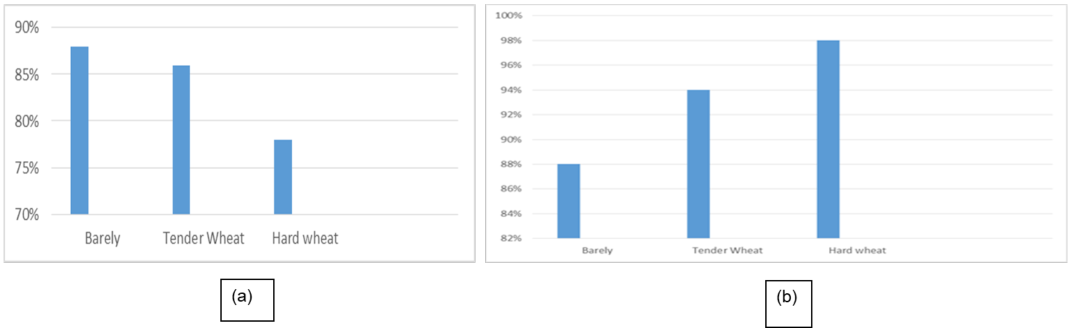


Fig 4. (a) Results for Fuzzy Logic Based classification method; (b) Different image results for Statistical and Fuzzy Logic combined classification method

The statistical method used for the hard wheat and tender wheat grains, gives us the best recognition rate. The first method is more reliable. On the other hand, for the barley grains we use a method that combines the two previous methods that gives the best recognition rate.

2.5.2 Statistical and Fuzzy Logic combined classification method:

The fuzzy logic method is decisive on the grain type in cases where the statistical method cannot make a decision. In the cases of overlaps of all morphological and color parameters, fuzzy logic is used in the combined method. The barley grains possess an optimal recognition rate. The improvement concerns the hard wheat and tender wheat grains only. The results of this method are illustrated in Figure 5.

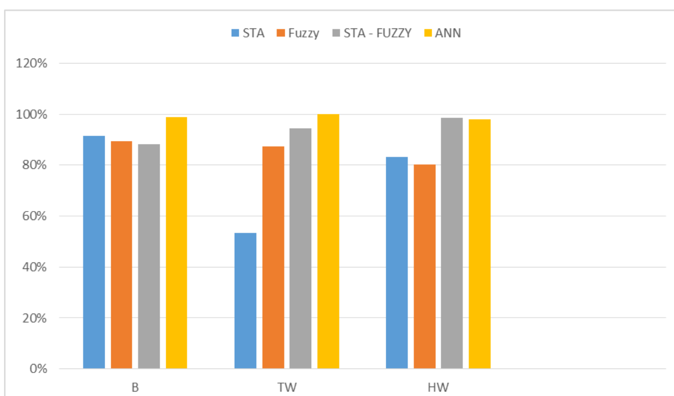


Fig 5. Comparison of the classification methods based on average recognition rates

2.6 Artificial Neural network classification method (ANN)

2.6.1 Training phase and network architecture

The network architecture is a multi-layer neural network MLP. Consuming the function "TRAINLM" the training is done from the Matlab neural network toolbox. The hyperbolic tangent and linear Matlab functions "tang" and "pure" are Activation functions[8]. We decide on the training error during the training phase, we mixed the neurons' numbers in the hidden layer. To a minimum training error, we chose 40000 as the training iterations number since this value leads. The number of neurons in the hidden layer depends on the type of features measured as entries of the network Table 3. When using different types of features (for the morphological features SM means Size Measurements, FC: Freeman Code, FT: Fourier Transform, and AMF: All Morphological Features) we prove the variant of the number of neurons in the hidden layer.

2.6.2 Classification results:

We used 3000 grains (1000 grains of each class) for this test, for characterization 600 grains, and validation 400 grains. By 1800 grains the training of each class is done, to acquire the true membership, 600 grains will be used and the other 1200 will be used to learn the system of the false membership to the class. To expand the classification space this technique appears to be very original and will make it possible, to refine space collates, and reduce the conflict rate between various classes.

Thus, this test will decide the conflict rate (CR), the rejection rate (RJR), and the recognition rate (RCR). During the first test Table 4 represents the results acquired.

Table 4. Classification

Rates (%)									
	Morphological features			Color features			Wavelet features		
Grain type	**CR**	**RJR**	**RCR**	**CR**	**RJR**	**RCR**	**CR**	**RJR**	**RCR**
B	1.1	0	98.9	2.8	0	97.2	2.5	0.7	96.8
HW	1.6	0.5	97.9	7.6	0.7	91.7	1.7	1	97.3
TW	7.7	2.8	89.5	3.5	1.3	95.2	0	0	100
Mean rates	3.5	1.1	95.4	4.6	0.7	94.7	1.4	0.6	98
Confusion Matrix (%) for the statistical method									
	B	**TW**	**HW**						
B	91.4	1.5	7.1						
TW	5.2	53.5	41.3						
Confusion Matrix (%) for the fuzzy logic-based method									
	B	**TW**	**HW**						
B	89.5	0.9	9.6						
TW	0.5	87.5	12.0						
HW	4.1	15.7	80.2						

3 Result and Discussion

Figure 5 illustrates the classification recognition rates of the four developed methods. The best acknowledgment rates for Barley is (98.9%). The ANN classifier lead using morphological features and Tender wheat (100%) using wavelet features whereas for Hard wheat classification (98,7%) the statistical and fuzzy logic collective classifier was the best, then the first and second ones. These two methods gave healthier results.

Four developed classifiers Tables 4 and 5 present the confusion matrices. This is because of the comparisons that exist in the morphology and the texture of these two cereal grain classes between Tender Wheat and Hard wheat. In the Statistical classification method (41, 3% for HW and 13, 9 TW), we note that the major confusions Fuzzy Logic classification based method (12% for HW and 15, 7% for TW), and Statistical and Fuzzy Logic Combined classification method (5% for HW and 0, 8% for TW). Consuming the ANN classification method (0% for HW and 1, 6% for TW) this problem is determined.

With Hard Wheat (STA: 7, 1%; FUZZY: 9, 6%; STA+FUZZY: 10, 5% and ANN: 0,6%) Barley grains are more disorganized than with Tender Wheat (STA: 1, 5%; FUZZY: 0, 9%; STA+FUZZY: 1, 1% and ANN: 0, 5%) that is more than Tender Wheat and color features which is because of the size.

We count the time in seconds to estimate the time performance for each classification method to categorize grains that take every algorithm in a sample of 300 grains holding 100 grains of each type.

Based on the classification method we have observed that the Fuzzy Logic performs to run about 60% faster than the second fastest (Statistical classification method). Then the fastest method the method important to the best recognition results is 4 times slower. The most execution method as it has a good recognition rate (94%). The Statistical and Fuzzy Logic combined classification method can be measured and take 50% less time than the method important to the optimal recognition rate. We note that we have used Matlab 2007 while the informed execution times depend on the implementation language.

In comparison, study made by [8], the previous author used SVM classifier. SVM has limited to two class classifiers. For higher number of classes, SVM is not recommended. To overcome the limit of SVM we proposed statistical classification method and

Table 5. Confusion Matrix (%)

For the statistical and fuzzy logic method			
	B	**TW**	**HW**
B	88.4	1.1	10.5
TW	0.6	94.4	5.0
HW	0.5	0.8	98.7
For the ANN method			
	B	**TW**	**HW**
B	98.9	0.5	0.6
TW	0	100	0
HW	0.5	1.6	97.9
Time performance of the different methods			
Method		**Time (s)**	
STA		72	
FUZZY		43	
STA+FUZZY		84	
ANN		177	

fuzzy logic based classification method.

Since the real-world scenarios of today require enormous data to be processed, ANN can fit well to diversify applications since they highly reduce the number of parameters to be trained that speeds up the training process. Moreover, to be specific on image classification problems they require the best and most prominent features to be detected and uncovered. This can be achieved using ANN since it has the concept of classification using training and testing at its Core. Hence, ANN is highly recommended for such image classification applications than the traditional artificial-neural-networks because of the aforementioned reasons.

4 Conclusion

The classification of wheat grain types (morphological, color, and wavelet) was well achieved by testing on different classification methods like Statistical classification; a fuzzy logic-based classification, Statistical and Fuzzy Logic combined classification, Artificial Neural network classification method (ANN). The statistical classification method gave an average recognition rate of 76%. The second method based on fuzzy logic techniques gave an average recognition rate of 85.73%. The hybrid method, which is a mixture of the two fore declared methods gave an average recognition rate of 93.83%. Finally, on all features the ANN classification method was tested and gave the best recognition rate of 98%.

5 Future Scope

The proposed method was able to achieve high performances, but there are some issues that still make it too difficult to achieve better performance such as type of database and quality images in these databases and also, the technique that are used in pre-processing stage. The following recommendations can be considered in order to enhance the performance of the proposed system.

- Applying the proposed system on different databases.
- Using another feature extraction technique.
- Applying more features on different technique.

6 Acknowledgement

The research was carried out in Electronics and Communication Engineering department of Birla Vishvakarma Mahavidyalaya Engineering College, Vallabh Vidyanagar. The research is a part of Ph.D. work in Gujarat Technological University.

References

1) Çelik Y, Başaran E, Dilay Y. Identification of durum wheat grains by using hybrid convolution neural network and deep features. *Signal, Image and Video*

Processing. 2022;16(4):1135–1142. Available from: https://doi.org/10.1007/s11760-021-02094-y.

2) Zhou L, Zhang C, Taha MF, Wei X, He Y, Qiu Z, et al. Wheat Kernel Variety Identification Based on a Large Near-Infrared Spectral Dataset and a Novel Deep Learning-Based Feature Selection Method. *Frontiers in Plant Science*. 2020;11:1–12. Available from: https://doi.org/10.3389/fpls.2020.575810.
3) Wang L, Sun Y. Image classification using convolutional neural network with wavelet domain inputs. *IET Image Processing*. 2022;16(8):2037–2048. Available from: https://doi.org/10.1049/ipr2.12466.
4) Choudhary R, Paliwal J, Jayas DS. Classification of cereal grains using wavelet, morphological, colour, and textural features of non-touching kernel images. *Biosystems Engineering*. 2008;99(3):330–337. Available from: https://doi.org/10.1016/j.biosystemseng.2007.11.013.
5) Liu J, Li P, Tang X, Li J, Chen J. Research on improved convolutional wavelet neural network. *Scientific Reports*. 2021;11(1):1–14. Available from: https://doi.org/10.1038/s41598-021-97195-6.
6) Khatri A, Agrawal S, Chatterjee JM. Wheat Seed Classification: Utilizing Ensemble Machine Learning Approach. *Scientific Programming*. 2022;2022:1–9.
7) Natarajan S, Ponnusamy V. Classification of Organic and Conventional Vegetables Using Machine Learning: A Case Study of Brinjal, Chili and Tomato. *Foods*. 2023;12(6):1–21. Available from: https://doi.org/10.3390/foods12061168.
8) Ibrahim S, Kamaruddin SBA, Zabidi A, Ghani NAM. Contrastive analysis of rice grain classification techniques: multi-class support vector machine vs artificial neural network. *IAES International Journal of Artificial Intelligence (IJ-AI)*. 2020;9(4):616–622. Available from: http://doi.org/10.11591/ijai.v9.i4.pp616-622.
9) Bouguettaya A, Zarzour H, Kechida A, Taberkit AM. Deep learning techniques to classify agricultural crops through UAV imagery: a review. *Neural Computing and Applications*. 2022;34(12):9511–9536. Available from: https://doi.org/10.1007/s00521-022-07104-9.
10) Benos L, Tagarakis AC, Dolias G, Berruto R, Kateris D, Bochtis D. Machine Learning in Agriculture: A Comprehensive Updated Review. *Sensors*. 2021;21(11):1–55. Available from: https://doi.org/10.3390/s21113758.
11) Velesaca HO, Suárez PL, Mira R, Sappa AD. Computer vision based food grain classification: A comprehensive survey. *Computers and Electronics in Agriculture*. 2021;187:1–13. Available from: https://doi.org/10.1016/j.compag.2021.106287.
12) Walenczykowska M, Kawalec A, Krenc K. An Application of Analytic Wavelet Transform and Convolutional Neural Network for Radar Intrapulse Modulation Recognition. *Sensors*. 2023;23(4):1–18. Available from: https://doi.org/10.3390/s23041986.