Çanakkale Onsekiz Mart University

Journal of Advanced Research in Natural and Applied Sciences

dergipark.org.tr/tr/pub/jarnas

Open Access

e-ISSN: 2757-5195 | 2024, Vol. 10, Issue 4, Pages: 771-785 | doi.org/10.28979/jarnas.1512352

Classifying Sunn Pest Damaged and Healthy Wheat Grains Across Different Species with YOLOV8 and Vision Transformers

Melike Çolak[1], Özgü Özkan[2], Nergis Pervan-Akman[3], Ali Berkol[4], Murat Olgun[5], Zekiye Budak-Başçiftçi[6], Nazife Gözde Ayter-Arpacıoğlu[7], Okan Sezer[8], Murat Ardıç[9]

[1,2,3,4]BİTES Defence and Aerospace Technologies, Ankara, Türkiye

[5,6,7]Department of Field Crop, Faculty of Agriculture, Eskişehir Osmangazi University, Eskişehir, Türkiye

[8,9]Department of Biology, Faculty of Science, Eskişehir Osmangazi University, Eskişehir, Türkiye

Article Info

Received: 08 Jul 2024

Accepted: 27 Oct 2024

Published: 31 Dec 2024

Research Article

Abstract − Sunn pest damage is one of the most crucial types of agricultural damage. Authorities and farmers are working together to find a cost-effective solution for separating the damaged crops from the healthy ones. This challenge can be tackled cost-effectively with emerging technology. Over time, the number of researchers focusing on this problem by using various machine learning algorithms and image processing techniques has increased. This paper presents an approach using a recurrent neural networks-based transformer to identify different varieties of wheat grain that have been sunn pest-damaged and healthy. First, wheat grains were separated from each other using YOLOv8. Then, the dataset was enriched with different data augmentation techniques, and data-efficient vision transformers were used to classify sunn pest-damaged and healthy grains. Conversely, a high accuracy score of 98.61% was achieved on the augmented dataset while surpassing the accuracy score of 93.36% in the raw dataset. This paper's contributions to literature can be divided into three categories. In contrast to the previous research, perfectly shaped, broken, and half-wheat grains are used to better fit findings in real-life environments such as factory production lines. Moreover, this study employs a combination of augmentation techniques, implying that two separate augmentation techniques, texture-based and one morphological, were applied to the same image. Finally, no study in the available literature uses a vision transformer to classify healthy and sunned pest-damaged wheat grains. That leads to using a data-efficient vision transformer algorithm and achieving a high accuracy score of 98.61%.

Keywords − *Sunn pest, sunn pest detection, wheat cultivars, wheat grain segmentation, crop quality*

1. Introduction

The most traditional and significant grain crop is wheat. About 20% of the world's total nutritional calories and proteins come from this crop [1] due to being the raw material of various foods. Among the most crucial types of wheat are common wheat (Triticum aestivum), used for bread production; durum wheat (Triticum durum), used for pasta; and club wheat (Triticum compactum), a softer variety used for making cakes, crackers, cookies, pastries, and flours. A small amount of wheat is also utilized by industry to make items, including starch, paste, malt, dextrose, gluten, and alcohol [2]. Moreover, wheat is one of the most popular foods consumed worldwide. China leads the world in wheat consumption, with India close behind [3].

Sunn pests are one of the primary insects that harm wheat grains. The sunn pest causes significant qualitative

[1]melike.colak@bites.com.tr (Corresponding Author); [2]ozgu.ozkan@bites.com.tr; [3]nergis.pervan@bites.com.tr; [4]ali.berkol@bites.com.tr; [5]molgun@ogu.edu.tr; [6]zbudak@ogu.edu.tr; [7]gayter@ogu.edu.tr; [8]oksezer@ogu.edu.tr; [9]mardic@ogu.edu.tr

and quantitative damage in the Middle East and numerous other regions. Damage caused by sunn pests is characterized by the yellowing and death of stems and leaves and stunted growth of tips and buds. These pests typically feed on other parts of the plant and do not commonly affect flower formation and coloration. When sunn pests feed on seeds after they have matured, they may become shriveled, discolored (usually white), or hollow. However, if these pests feed on the seeds before they fully develop, it can lead to seed abortion. Both adults and nymphs can also feed on dry grains when moisture is present [4]. It has infested about 15 million hectares of wheat crops across the Middle East. The sunn pest damages wheat grain during various phases of development. The sunn pest sucks out the protein in the wheat grain at this stage, as the grains will lose their germination power. Additionally, they lose their qualities as flour. Economic devastation results from this loss of quality; in Iran alone, damage to wheat and barley crops exceeding 9 million tons has been reported [5].

Integrating advanced technology into the agricultural sector has spurred numerous research endeavors within wheat grain analysis, leveraging Machine Learning (ML) methodologies. Farmers eliminate damaged wheat grains manually because the machines that do this automatically are prohibitively expensive. Thus, the need for cheaper, faster, and more accurate quality control in wheat grains has paved the way for sunn pest damage detection [6-8] and classification [9] studies in wheat grains. In addition, since wheat species are used to produce different foods, they must not mix. In light of this motivation, researchers have conducted classification studies with datasets containing many different wheat species [10-19].

Researchers have tackled classification studies using datasets comprising several wheat varieties. The effective analysis of wheat grains is facilitated through the implementation of segmentation techniques, particularly in intricate tasks such as isolating individual wheat grains from image backgrounds [14, 20] or from each other [9], mainly when presented in bulk as in our dataset [21]. Besides, researchers have used segmentation algorithms to classify different species of wheat grains [22]. In Computer Vision (CV), the variety and abundance of data are one of the main factors affecting model training performance. Therefore, researchers have diligently expanded and enriched their datasets by strategically employing data augmentation techniques [8, 22-24] specifically tailored to address the intricacies of wheat grain analysis. The culmination of these extensive studies has yielded compelling results, demonstrating the model's capacity to distinguish between damaged wheat attributed to sunn pests and healthy wheat with impressive accuracy. Such accomplishments hold profound significance within the agricultural sector, where the reliable classification of wheat grains is pivotal.

In this study, our primary goal was to achieve a highly accurate classification of sunn pest-damaged and healthy grains in various wheat grain species with the Data-efficient Image Transformer (DeiT) algorithm. To conduct our experiments, the wheat grain dataset previously introduced in our research was leveraged [21]. Additionally, the aim to deepen our comprehension of wheat grain features by employing the You Only Look Once version 8 (YOLOv8) model for segmentation on a dataset containing bulk wheat grains was pursued. As part of this segmentation process, each wheat kernel was individually recorded as a separate image, leading to the generation of a novel and enriched dataset through diverse data augmentation techniques. Our study contributes to the literature in various ways by improving the existing studies on wheat grain species and sunn pest-damaged wheat grain classification.

*i.*Our study provides an adaptable model for grains not in the whole wheat form in a production line by including broken and partial wheat grains in the dataset.

*ii.*In addition to the data augmentation techniques applied to the dataset separately in the literature, the model's performance was increased by combining various techniques to create a dataset that can be adapted to many environmental conditions, such as lighting conditions, image contrast, and the presence of shadows. This versatile approach allows our model to perform effectively in various scenarios.

*iii.*Our study is the first in the available literature to examine sunn pest damage to different wheat species.

*iv.*Finally, the approach we presented is a pioneer in applying Transformer-based models in wheat grain analysis, with an accuracy score of 98.61%, and it is the first study of such an approach.

The remainder of this paper is structured as follows: Section 2 presents related works from reviewing the literature in the subsections Augmentation, Segmentation, and Classification. Section 3 describes the dataset used in this study and the YOLOv8 model used for the wheat grain segmentation. Section 4 encompasses a detailed explanation of the data augmentation techniques used. Section 5 presents an exploratory data analysis of the augmented dataset. Section 6 explains the sunn pest damage classification in different wheat species. Section 7 contains the analysis and comparison of the results of the presented experiment in the subsections Segmentation, Augmentation, and Classification. Section 8 encapsulates the conclusions drawn from the study and outlines potential avenues for future research. Last of all, the final section explains the future work.

2. Related Work

Upon reviewing the literature, it became evident that several methods exist for resolving the classification issue between sunn pest-damaged and healthy grain. The authors have used various segmentation models and augmentation techniques to achieve a higher accuracy score. This section focuses on various augmentation techniques, segmentation methods, and classification problems.

2.1. Augmentation

The reviewed literature demonstrates the necessity for more enormous datasets, which led to various augmentation methods. The authors conducted diverse studies about this specific problem.

Shen et al. [22] utilized a dataset of 130 images that included damaged but beneficial wheat grains, such as unsound grains, injured grains, speckled grains, broken grains, germinated grains, and moldy grains. Subsequently, they applied five data augmentation methods to process the images to enhance recognition accuracy and deliberately highlight local characteristics and regions of interest. These methods included brightness reduction, noise addition, random point insertion, translation, and flipping. Bernandes et al. [23] sampled 1200 seeds, with 600 seeds infected by Fusarium Head Blight (FHB) and 600 healthy seeds. In this study, signs of fungal colonization in wheat seeds due to FHB infection were characterized by roughness, a wrinkled appearance, and pink-colored tissues. To avoid overfitting during hyperparameter optimization, they randomly rotated images by up to 30° and applied a 20% variation in height, width, offset, and zoom.

Furthermore, they rotated half of the images horizontally and used the closest strategy to fill in any newly created pixels that might appear after rotation or width/height shifting. Unlersen et al. [24] applied rotation and zoom operations to wheat grain bulk samples. First, they randomly determined an angle to rotate the image between -30 to +30 degrees. Then, the zoom operation was executed with a randomly determined rate between -10% and +10% on the rotated images. After the data augmentation, they obtained 3200 images, 800 from each wheat variety. In their study, Sabanci et al. [8] rotated and applied Gaussian noise to cropped wheat grain images. They also adjusted the zoom, size, and positioning of the images. This process expanded the dataset, yielding 1200 wheat grain images, 600 of which are healthy samples and 600 of which are sunned pest-damaged samples. Out of these, 1000 images were allocated for training deep learning models, while the remaining 200 were set aside for testing.

2.2. Segmentation

Segmentation is an image processing technique that divides an image into meaningful parts, enabling detailed analysis. In agriculture, this technique helps to identify different wheat species and detect damage from sunn pests or other bugs. This technique outlines grain boundaries to distinguish healthy and damaged grains. While common in corn, beans, and coffee grain analysis, specific studies on wheat are less frequent.

Researchers have studied various segmentation techniques to better examine the characteristics of wheat grains. Gao et al. [25] present a study to detect unsound wheat kernels based on an improved Residual Neural Network (ResNet), six kinds of wheat, including the sound kernel, broken kernel, sprouted kernel, injured kernel, moldy kernel, and spotted kernel are considered as the samples. The designed two-kernel adhesion wheat segmentation algorithm based on a concave mask exhibits high accuracy, with an error rate of 0.93% for 9988 wheat grains. Sharma and Singh [15] utilized Artificial Neural Networks (ANN), Support Vector Machine (SVM), partial least squares discriminant analysis, Random Forest (RF), and K-Nearest Neighbor (KNN) to classify wheat seed varieties using near-infrared hyperspectral imaging. The ANN model with Savitzky-Golay second derivative preprocessing achieved the highest accuracy of 97.77%.

2.3. Classification

There are two different kinds of wheat classification issues: The classification of wheat species and the classification of damaged or healthy grains. This section investigates both approaches.

Considering damage classification problems, the research conducted by Motie et al. [26] utilized SVM with a radial basis function kernel to achieve an accuracy exceeding 90% in differentiating sunn pest-infected wheat clusters from healthy plants using near-infrared images. Abbaspour-Gilandeh et al. [27] achieved 100% accuracy in discriminating healthy wheat grains from grains infected with Fusarium using SVM. Fazel-Niari et al. [28] studied various classification algorithms, including linear and quadratic statistical discriminant analysis and SVM, and achieved an average accuracy of 97.6% in classifying wheat grain groups. Shedole et al. [17] developed a Convolutional Neural Networks (CNN) based classification system for wheat grain using 900 images. The Decision Tree and Multilayer Perceptron classifiers achieved 98.7% accuracy, while the Naïve Bayes classifier had a lower precision of 94.22%. Additionally, the validation accuracy of the CNN model showed strong performance, ranging from 94% to 96%. Kaya and Saritas [29] developed a classification system using ANN for classifying type-1252 durum wheat kernels based on their clarity, achieving a maximum classification accuracy of 93.46%. In a study conducted by Erkinbaev et al. [30] achieved overall classification accuracies ranging from 83% to 100% using spectral data and ML techniques with a unified heuristic approach. Zhang and Ji [31] classified wheat grains into different states using hyperspectral imaging, with an SVM model achieving an average recognition accuracy of 98.5%.

Using pre-trained models such as Resnet, Alexnet, or Densely Connected Networks (DenseNet), which are trained with large datasets, is often more advantageous than training a model from scratch. In the literature, authors used a transfer learning approach to achieve higher accuracy scores in wheat grain classification problems. They used pre-trained models directly or combined different pre-trained model architectures with their models [8, 9, 24, 32]. Table 1 summarizes related work, including whether the paper is used as an augmentation, segmentation, or classification reference. Then, there is a studies column, a dataset column that shows a summary of the dataset, a methods column, and a column that explains the essential outcomes.

Table 1. Literature summary

Category	Studies	Dataset	Methods Used	Key Outcomes
Augmentation	[22]	130 wheat grain images (damaged and unsound)	Brightness reduction, noise addition, random point insertion, translation, and flipping	Enhanced recognition accuracy by emphasizing local characteristics.
	[23]	1200 wheat grain images (600 FHB infected, 600 healthy)	Random rotation (up to 30°), height/width shifting (20%), horizontal flip, closest pixel strategy	Avoided overfitting, characterized fungal colonization.
	[24]	3200 wheat grain bulk images	Rotation (-30° to 30°) and zoom (-10% to 10%)	Processed 800 images per wheat variety.
	[8]	1200 wheat grain (600 healthy, 600 damaged)	Rotation, Gaussian noise, zoom, resizing, positioning adjustments	Expanded dataset for training deep learning models.

Table 1. (Continued) Literature summary

Category	Studies	Dataset	Methods Used	Key Outcomes
Segmentation	[25]	9988 wheat grain images (6 types of kernels)	Improved ResNet, concave mask-based two-kernel adhesion segmentation	Achieved an error rate of 0.93% in segmentation.
	[15]	Near-infrared hyperspectral imaging (15 different wheat varieties; 936 wheat from each variety)	ANN, SVM, Random Forest, KNN with Savitzky-Golay second derivative preprocessing	ANN model achieved the highest accuracy of 97.77%.
Classification	[26]	Near-infrared images of wheat clusters (Sunn-pest damaged and healthy wheat grains)	SVM with radial basis function kernel	Over 90% accuracy in classifying sunn pest-infected wheat clusters.
	[27]	Wheat grains (300 healthy vs. 300 Fusarium infected)	SVM	Achieved 100% accuracy.
	[28]	Wheat grain groups (7 different wheat types, a total of 21000 single wheat grains)	SVM, linear, and quadratic discriminant analysis	Achieved an average accuracy of 97.6%.
	[17]	900 wheat grain images (dataset includes four classes: good grain, damaged grain, broken grain, and foreign particles.	CNN, Decision Tree, Multilayer Perceptron and Naïve Bayes	CNN model accuracy between 94% and 96%, Decision Tree/MLP 98.7%, Naïve Bayes 94.22%.
	[29]	Durum wheat kernels	ANN	Maximum classification accuracy of 93.46%.
	[30]	Spectral data (180 Healthy kernels, 60 Fusarium-damaged kernels, and 60 Ergot-damaged kernels)	ML techniques with a unified heuristic approach	Classification accuracy ranges from 83% to 100%.
	[31]	Hyperspectral imaging of wheat grains (220 healthy wheat grains, 220 germinated wheat grains, 220 mildewed wheat grains, and 220 shriveled wheat grains (wheat grains infected with fusarium head blight)	SVM	Achieved average accuracy of 98.5%.

3. Data Collection

The Related Work section explains that the authors handle wheat grain differently. These datasets vary in many aspects, including being created in controlled or uncontrolled environments, camera models and equipment in the shooting environment, camera angle, direction and type of wheat, and the focused problem such as species classification or damaged region detection. The dataset used in this study was prepared for the classification stage after preprocessing and segmenting different types of wheat grains affected by the sunn pest, presented in [21].

3.1. Raw Wheat Grain Dataset

In [21], a new dataset was introduced, which was used to develop an image classification model for classifying wheat grain species as Damaged and Healthy. The dataset includes a wide variety of wheat and covers six species: Bezostaja, Müfitbey, Nacibey, Sönmez-2001, Tosunbey, and Ekiz, the species made in Türkiye. Wheat grains differ in various parameters, such as width, length, color, stain condition, and wrinkled texture.

The dataset comprises 83 images of sunn pest damage and 87 images of healthy wheat grains. Additionally, 2502 healthy and 1063 damaged wheat grains were extracted from 170 bulk wheat grain images. Wheat grains vary in width, length, color, stain status, and wrinkled texture. The distribution of the cultivars is displayed in Table 2.

Table 2. Distribution of wheat grains

Cultivar	Healthy	Damaged
Bezostaja	427	107
Müfitbey	311	120
Nacibey	342	203
Sönmez-2001	370	120
Tosunbey	322	183
Ekiz	730	330

Due to the wheat grain dataset being multiclass and may be used for variety classification, it stands out from others due to its diverse species and condition, including broken, sunn pest damaged, and healthy wheat grain. However, the dataset is appropriate for detecting sunn pest damage when the condition of sunn pest damage is considered. Furthermore, the dataset contains grains that come into contact with one another, increasing its applicability to real-world issues. As can be seen, the dataset includes a range of wheat species that are impacted by sunn pests, demonstrating that the promised dataset is appropriate for many ML tasks, including segmentation, detection, and classification.

3.2. Wheat Grains Segmentation with YOLOv8

The dataset used in our study is the most crucial contribution that distinguishes the study from other studies in the literature. Although studies conducted with wheat grains in a particular order, facing the same direction, and of the same size give promising accuracy scores, it is not expected that the ML model will be able to solve the problem correctly in the face of wheat grains scattered in the real-life production line. For this reason, some processes were applied to the raw dataset to bring a new perspective to academic studies and to offer realistic solutions that can be adapted to real life in wheat agriculture.

In the segmentation process, the dataset was split into 60%, 20%, and 20% as train, test, and validation, respectively. The split 60%-20%-20% is commonly used to ensure a balanced model training and evaluation approach. The 60% training set provides enough data for the model to learn underlying patterns effectively. The 20% validation set helps tune the model's hyperparameters and avoid overfitting. The 20% test set is reserved to evaluate the model's generalization performance on unseen data, ensuring a reliable assessment of its accuracy. Then, to transfer the data to the model during training, the file paths of the train, test, and validation sets were given to the file data.yaml, and the class name was set as wheat.

The model YOLOv8 was utilized for the training process. YOLOv8, known for its efficiency and accuracy in object detection tasks, is particularly suitable for our problem due to its ability to handle varying object scales and dense environments. The default parameters of YOLOv8 were used, which include an initial learning rate of 0.01, momentum of 0.937, and weight decay of 0.0005.

YOLOv8's architecture consists of a backbone network that extracts essential features from input images and a head network that predicts bounding boxes and class probabilities. The model is pre-trained on the Common Objects in Context (COCO) dataset, allowing it to leverage transfer learning for better initial performance on our wheat grain dataset.

In the training process, YOLOv8's advanced anchor-free detection mechanism plays a significant role in accurately predicting the locations and sizes of wheat grains, which can vary significantly in a production line setting. After each epoch, the model's performance was continuously monitored by evaluating the Mean Average Precision (mAP) on the validation set. Train results, including precision, recall, and mAP scores, will be discussed in the Results section, comprehensively evaluating our model's performance.

4. Data Augmentation

After the segmentation process, augmentation techniques were applied to segmented wheat grains to expand our data by a factor of seven. 15 different augmentation methods were applied. Initially, two distinct kinds of augmentation were used: morphological and texture-based.

The morphological augmentation techniques include;

*i.*Affine: Scales images to a value of 40 to 80% of their original size.

*ii.*Fliplr: Reflects photos horizontally.

*iii.*Flipud: Reflects photos vertically.

*iv.*Furthermore, texture-based augmentations contain;

*v.*Add to Brightness: This function adds a constant value to an image's brightness. The range is set from +30 to -30.

*vi.*Linear Contrast: Adjusts the contrast of each image by $127\alpha(v - 127)$, where v represents the pixel value, and alpha is uniformly sampled (once per image) from the interval [0.4, 1.6].

*vii.*Multiply Saturation: The augmenter first converts images to Hue Saturation Value (HSV) colorspace, doubling the pixel values in the H channel before converting back to RGB. The range is determined between 0.5 and 1.5.

In addition to these six techniques, all of them are combined. Firstly, one from morphological and one from texture-based are picked to acquire a combined augmentation technique. The first method is used, followed by the second. For instance, if both Fliplr and Linear Contrast are picked, the image is flipped horizontally first, and then linear contrast is applied. With this process, nine more augmentation techniques are obtained. In Table 3, the combined augmentation techniques are shown.

Table 3. Combined augmentation techniques

Morphological	Texture-Based	Combined
Affine (AF)	Add To Brightness (BR)	AFBR
Affine (AF)	Linear Contrast (LC)	AFLR
Affine (AF)	Multiply Saturation (MS)	AFMS
Flipud (FUD)	Add To Brightness (BR)	FUDBR
Flipud (FUD)	Linear Contrast (LC)	FUDLR
Flipud (FUD)	Multiply Saturation (MS)	FUDMS
Fliplr (FLR)	Add To Brightness (BR)	FLRBR
Fliplr (FLR)	Linear Contrast (LC)	FLRLC
Fliplr (FLR)	Multiply Saturation (MS)	FLRMS

Table 3 shows the new combined augmentation techniques and what is used to create them. When augmentation, morphological, and texture-based techniques are combined in Table 3, 15 augmentation techniques are obtained. None of these augmentation techniques were utilized on a single wheat grain image. Initially, six were selected for every image. For each image, 2 out of 3 morphological and texture-based augmentation techniques and 2 out of 9 combined techniques were selected. First, two morphological augmentations are performed, chosen at random, for each image. Subsequently, two texture-based techniques were picked randomly and applied, and finally, two combined techniques were selected randomly and implemented. Ultimately, each image has been featured seven times, including itself. In Figure 1, the rows represent texture-based augmentations, and the columns represent morphological augmentations. Their crossing image illustrates the combination of these two augmentation techniques. As a result, the raw image and 15 images representing the 15 different augmentation techniques were used.

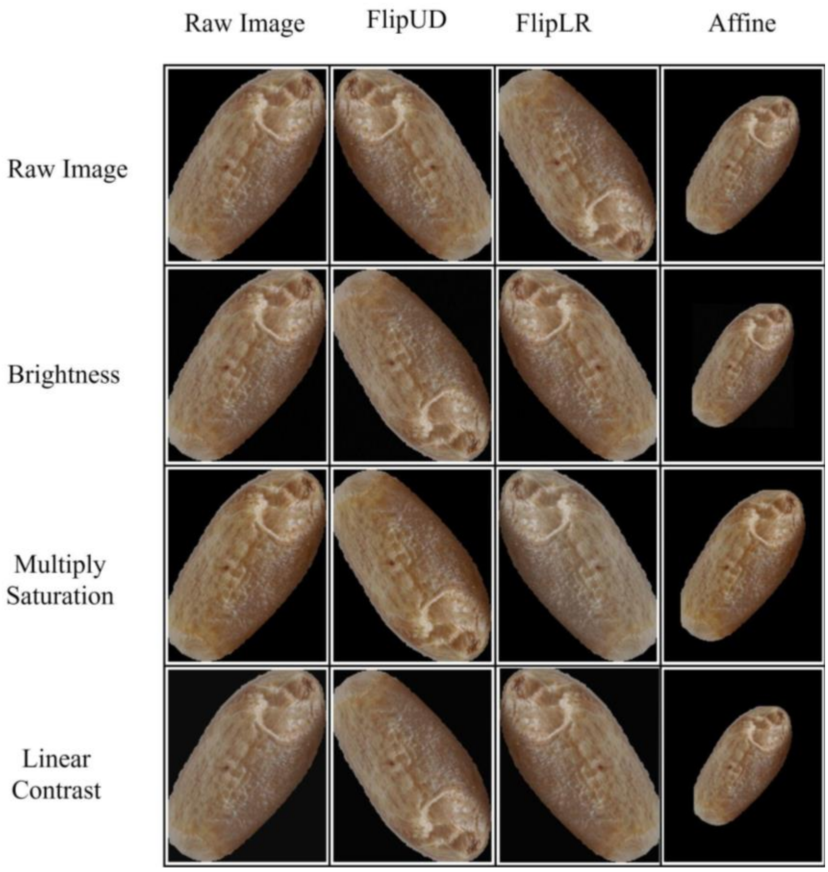

Figure 1. Image before and after each augmentation technique

Ultimately, when the data is augmented with the raw images, there are 24759 images. For each of the wheat varieties, the augmented image count without the raw images is as follows:

*i.*Bezostaja: 3144

*ii.*Müfitbey: 2538

*iii.*Tosunbey: 2988

*iv.*Nacibey: 3336

*v.*Ekiz: 6222

*vi.*Sönmez-2001: 2994

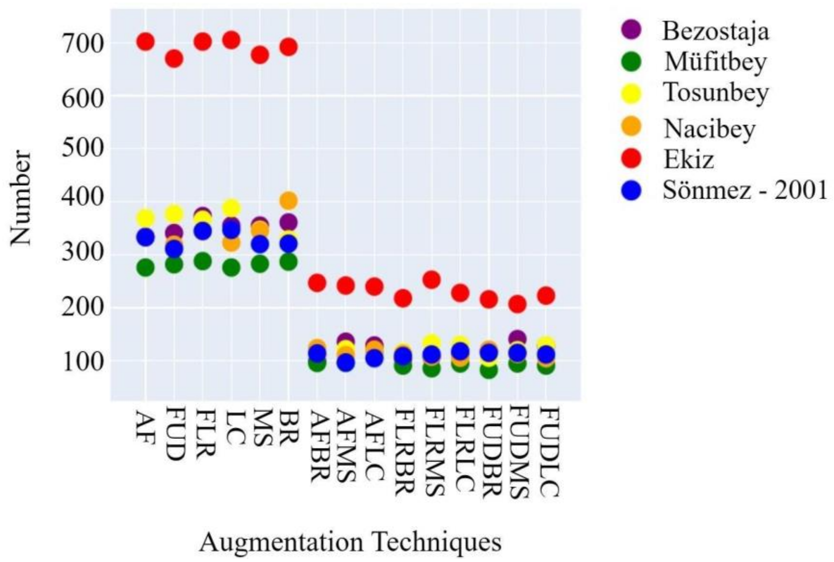

Figure 2. Augmented image count according to wheat varieties

In Figure 2 the distribution of all augmented image counts according to their relative wheat varieties has been illustrated. The combined augmentation techniques' counts are lower than the initial augmentation techniques' counts because 4 out of 6 in the initial ones were picked. However, in the combined techniques, 2 out of 15 were picked. So, the probability of choosing the same augmentation technique is less in combined augmentation methods. Conversely, the Ekiz varieties have a far larger count than the others since their raw version likewise has a higher count, and the difference widens when it is increased 6 times.

Wheat grains were segmented from the raw image, then these augmentation methods were applied. Following the augmentation process, zero padding was added based on the segmentation coordinates of the wheat grains. At the end of this process, all of the wheat grains were extracted one by one with a black background, leading to the final dataset.

5. Exploratory Data Analysis

In this section, exploratory data analysis is presented in our augmented dataset. In our research, graphs displaying brightness for wheat damaged by sunn pests against healthy wheat were carefully examined. As can be seen in Figure 3, a significant disparity in average brightness values was discovered. The damaged wheat grains had an average brightness of 85.23, compared to 116.40 for the healthy wheat grains. This significant difference shows that brightness may effectively distinguish between the two conditions. Moreover, uncommon brightness levels, or outliers, were examined in damaged and healthy wheat. Despite finding extreme levels, it is seen between the 0-200 range, they distribute similarly. This means that using only brightness might not be enough to separate damaged and healthy wheat, which presents a challenge, for instance, classification in wheat grain condition.

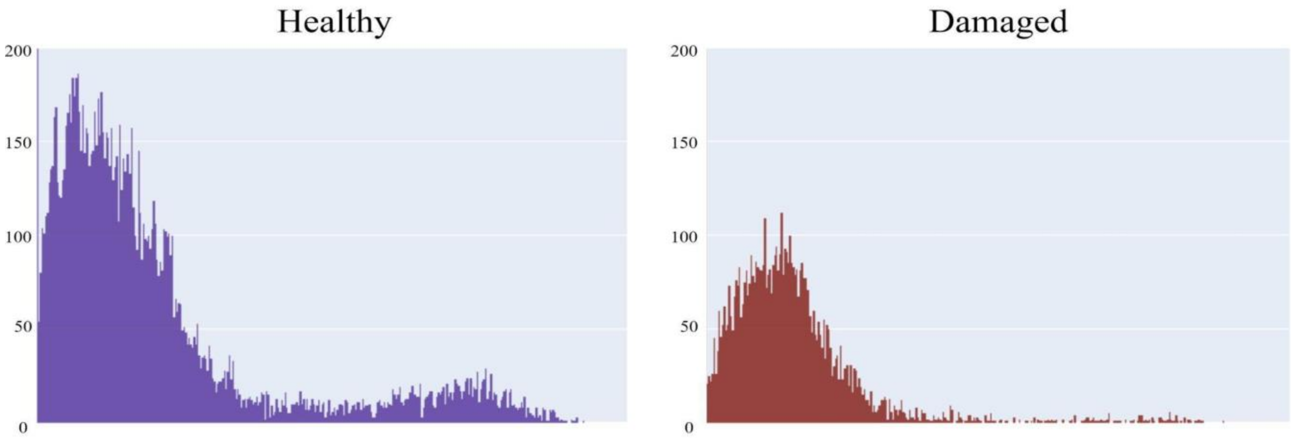

Figure 3. Brightness-frequency histogram

6. Sunn Pest Damage Classification in Different Wheat Species

When the literature is examined, it is seen that CNN and state-of-the-art models are preferred in studies that use image data. Since this transformer-based algorithm is not available in the literature, the DeiT model with the transfer learning approach was used in our research. Touvron et al. [33] assert that the DeiT represents a significant advancement in training Transformers to enhance CV performance. Despite CNN being the prevailing method for CV tasks in the past eight years and benefiting from numerous enhancements and adjustments, DeiT's performance is already on par with them. DeiT stands for data-efficient transformer, which focuses on making a convolution-free model trained on less data and can outperform CNN-based algorithms. DeiT model pre-trained and fine-tuned on ImageNet-1k at resolution 384x384. The last layer of the DeiT model was frozen, and seven layers were added with fine-tuning: Linear, Relu, and Dropout. Figure 4 shows the model architecture and classification process.

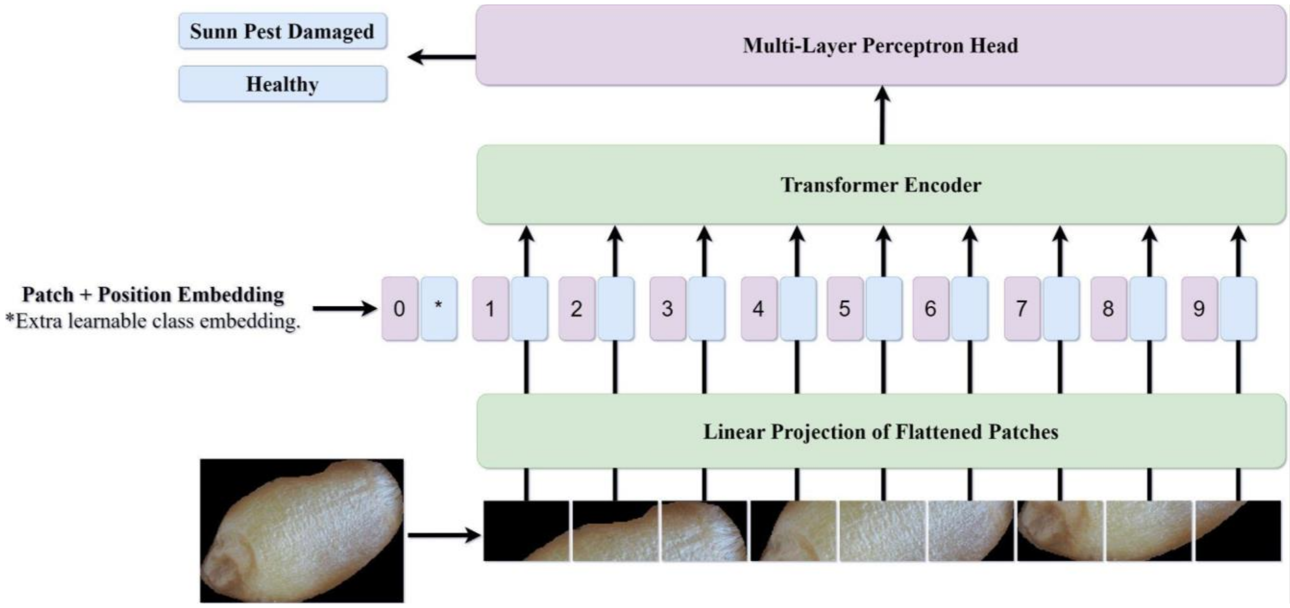

Figure 4. Our model architecture

Our experiments were conducted with two datasets: segmented wheat grains and expanded with augmentation techniques of these data. In both scenarios, the datasets' train, test, and validation sets were split, with proportions of 80%, 20%, and 20%, respectively. The learning rate is 0.00005, the batch size is 128, and the epochs are 200. During the training stage, early stopping was used to prevent the model from memorizing the training data and from overfitting the test data. Early stopping was controlled according to validation loss, with a patience coefficient of 3. The results of both scenarios were discussed comparatively in the Results section.

7. Results

7.1. Segmentation

In the wheat segmentation process, YOLOv8 was utilized to segment wheat grains from the images. This model was chosen for its advanced object detection and segmentation capabilities, making it well-suited for tasks involving dense and overlapping objects such as wheat grains. YOLOv8 represents a significant advancement in real-time object detection and segmentation, building upon the solid foundation laid by previous versions of the YOLO family. The defining feature of YOLO models is their ability to predict multiple bounding boxes and corresponding class probabilities in a single forward pass through the neural network. This design ensures that the model operates with exceptional speed and efficiency, making it suitable for real-time applications.

The YOLOv8 model incorporates several state-of-the-art enhancements for superior performance. The network architecture is significantly deeper and wider, capturing complex features and finer details. Advanced feature pyramid networks (FPN) combine features from different layers, improving precision and recall by detecting objects at various scales. YOLOv8 also uses an optimized anchor-free detection strategy to directly predict object centers and scales, reduce computational complexity, and enhance accuracy, especially for small, densely packed objects. Cross Stage Partial Network (CSPNet) also improves gradient flow and reduce computational load, leading to more efficient training and inference while maintaining high accuracy.

The high precision and recall values indicate that the YOLOv8 model is highly effective in correctly identifying and segmenting wheat grains within the test images. The precision score reflects the model's accuracy in detecting true positives among the identified grains. In contrast, the recall score indicates the model's ability to identify all actual grains in the images. The mAP values comprehensively measure the model's performance. The mAP50 value, which considers a single Intersection over the Union (IoU) threshold of 50%, shows an impressive score of 99.4%. This high score signifies that the model can accurately detect and segment wheat grains with a moderate overlap threshold. The mAP50-95 value averages the precision across multiple IoU thresholds from 50% to 95%, slightly lower at 93.3% for bounding boxes and 91.3% for segmentation masks. This indicates that while the model performs exceptionally well at moderate thresholds, there is a slight decrease in performance at higher thresholds. This decrease could be attributed to the challenges in segmenting wheat grains with fine details or densely packed regions.

7.2. Augmentation

When the segmentation process was finished, different augmentation techniques were performed. There are morphological augmentation techniques, texture-based augmentation techniques, and combined augmentation techniques consisting of one morphological and one texture-based technique. When the augmentation is considered, it is seen that performing augmentation techniques dramatically increased the accuracy by %5 -from %93.36 to %98.61-. The reason for the increase is not only the higher data number but also the chosen technique. Due to some of the augmentation methods, the grain color gets brighter, as a result, the sunn pest damage point has become more visible, and thus accuracy increased. In contrast, the sunn pest damaged point and the color of the grains both go darker in the Multiply Saturation method; nevertheless, the sunn pest damaged point becomes so much darker that it becomes more noticeable. On the other hand, in Linear Contrast, by enhancing the contrast of the images, the model can better distinguish between damaged and undamaged grains, leading to higher accuracy and more reliable predictions.

7.3. Classification

After the augmentation stage, two experiments were performed on the raw and augmented datasets. Determining optimal hyperparameters and layer values added to the end of the DeiT model was a critical step before model training. Parameters in the linear layer input sizes 1024, 512, 256, and 128 were tested to maximize classification accuracy. Additionally, dropout layer coefficients were tested as 0.20, 0.35, and 0.50 for both models. In both cases, the best parameters were specified as an input size 512 and a dropout coefficient of 0.50. The models trained on the raw and augmented datasets utilized a V100 GPU, with training durations of 10359.64 and 1644.41 seconds, respectively. As a result of the binary classification, a test accuracy score of 93.36% and 98.61% were achieved. The precision score was obtained as 99.39%, recall as 98.62%, specificity as 98.56%, F1-Score as 99.00%, and Matthews Correlation Coefficient (MCC) as 96.67%. In the model trained with the raw dataset, these results were 95.93%, 95.01%, 95.47%, 93.60%, and 84.58%, respectively. Comparative results between these models are shown in Figure 5.

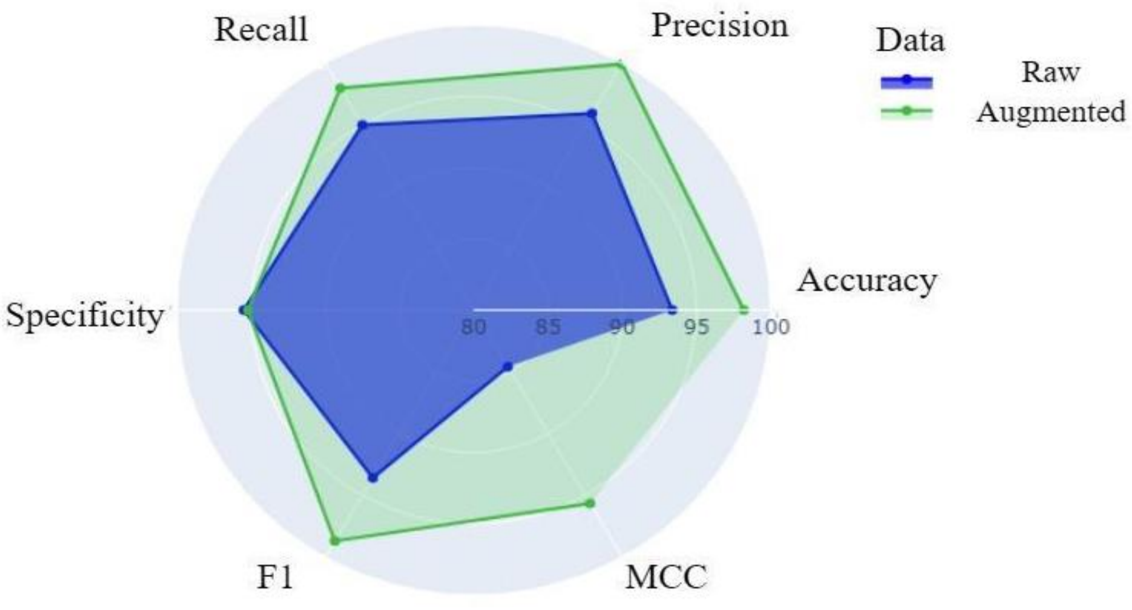

Figure 5. Raw and augmented results comparison

As shown in Figure 5, accuracy, precision, recall, specificity, F1 score, and Matthews Correlation Coefficient (MCC) are the six critical metrics used to compare the performance analysis between raw and augmented data. Each axis represents one of these measures, ranging from 80 to 100. The blue polygon shows the performance of the raw data, while the green polygon shows the augmented data. The raw data performs admirably with accuracy at 93.36%, precision at 95.93%, recall at 95.01%, specificity at 95.47%, F1 score at 93.60%, and MCC at 84.58%. The augmented data, however, greatly improves these metrics: MCC increases to 96.67%, accuracy to 98.61%, precision to 99.39%, recall to 98.62%, specificity to 98.52%, and F1 score to 99.00%. The enhanced performance of the augmented data across all measures is evident in this visualization, suggesting that data augmentation is a valuable technique for improving the predictive powers of the model. The green polygon highlights these benefits's more extensive area than that of the blue polygon, which supports the use of data augmentation approaches in ML problems.

The impressive performance of the self-attention mechanism in the Transformer-based Diet model lies behind the high accuracy score. Self-attention has been essential in identifying fine distinctions between sunn pest damage and healthy wheat grains. The self-attention mechanism helps the model handle long-range dependencies. This makes it more effective in understanding the relationships between different parts of wheat, ultimately leading to a more comprehensive classification structure.

8. Conclusion

One of the most critical forms of damage in agriculture is caused by sunn pests. Detecting sunn pest damage in wheat grains is a task that is both time-consuming and costly when done manually. Authors have studied this issue to make detecting sunn pest damage easier and classify healthy and damaged wheat grains, which is the purpose of our paper. This paper includes three critical processes: wheat grain segmentation, data augmentation, and classification with DeiT. The wheat grains were identified using YOLOv8 and cropped from the base image in the segmentation part. After that, 15 distinct augmentation methods were applied to these grains. These methods enrich the dataset and make it challenging for the model to understand the difference between healthy and damaged conditions. For this reason, significant success was achieved as results with a high accuracy score of 98.61% on the augmented dataset were yielded by our Transformer-based model, while 93.36% was obtained on the raw dataset.

Our contributions include a combination of augmentation techniques, implying that two separate augmentation techniques, texture-based and morphological, were applied to the same image. Additionally, whole and broken/half grains are included in the dataset, making it more suitable for real-life production bands. Furthermore, -to the best of our knowledge- using a multiclass dataset makes it the first study to work on a multiclass wheat grain dataset for sunn pest damage detection. Lastly, implementing the RNN-based algorithm in the evaluation process is one of the most important contributions of our study.

In future research, expanding the dataset and exploring alternative models and algorithms to enhance the model's accuracy would be beneficial. This exploration could include looking beyond RNN-based algorithms and considering architectures like CNNs or hybrid models that merge CNNs with Transformers. These alternatives might more effectively capture spatial features, complementing the Transformer's strength in handling temporal and sequential data. Additionally, exploring Graph Neural networks might offer a more intricate way to model the relationships between grain structures. Adopting such approaches could lead to new ways to improve the durability and efficiency of systems designed to classify damage caused by sunn pests.

Author Contributions

The first and second authors performed the data analysis, experiments, and statistical analysis and wrote the paper. The third author reviewed and edited the paper. The sixth, seventh, eighth, and ninth authors collected the data. The fourth and fifth authors supervised the project. All authors read and approved the final version of the paper.

Conflicts of Interest

All the authors declare no conflict of interest.

Ethical Review and Approval

No approval from the Board of Ethics is required.

References

[1] B. Shiferaw, M. Smale, H.-J. Braun, E. Duveiller, M. Reynolds, G. Muricho, *Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security*, Food Security 5 (2013) 291–317.

[2] Britannica, Wheat (Plant) (2024), https://www.britannica.com/plant/wheat, Accessed 30 June 2024.

[3] Statista, Wheat consumption worldwide in 2023/2024, by country (in 1,000 metric tons)* (2024), https://www.statista.com/statistics/1094065/total-global-wheat-consumption-by-country/, Accessed 30 June 2024.

[4] Plant Health Australia, Sunn pest (2024), https://www.planthealthaustralia.com.au/wp-content/uploads/2024/01/Sunn-pest-FS.pdf, Accessed 30 June 2024.

[5] M. Alamouti, M. Majdi, R. Talebi, M. Dastranj, A. Bandani, G. H. Salekdeh, M. R. Ghaffari, *Transcriptome wide identification of neuropeptides and G protein-coupled receptors (GPCRs) in Sunn pest, Eurygaster integriceps Puton*, Gene 893 (2024) 147911 14 pages.

[6] K. Sabanci, *Detection of sunn pest-damaged wheat grains using artificial bee colony optimization-based artificial intelligence techniques*, Journal of the Science of Food and Agriculture 100 (2) (2020) 817–824.

[7] Z. Basati, B. Jamshidi, M. Rasekh, Y. Abbaspour-Gilandeh, *Detection of sunn pest-damaged wheat samples using visible/near-infrared spectroscopy based on pattern recognition*, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 203 (2018) 308–314.

[8] K. Sabanci, M. F. Aslan, E. Ropelewska, M. F. Unlersen, A. Durdu, *A novel convolutional-recurrent hybrid network for sunn pest–damaged wheat grain detection*, Food Analytical Methods 15 (6) (2022) 1748–1760.

[9] K. Laabassi, M. A. Belarbi, S. Mahmoudi, S. A. Mahmoudi, K. Ferhat, *Wheat varieties identification based on a deep learning approach*, Journal of the Saudi Society of Agricultural Sciences 20 (5) (2021) 281–289.

[10] M. Olgun, A. O. Onarcan, K. Özkan, Ş. Işık, O. Sezer, K. Özgişi, N. G. Ayter, Z. B. Başçiftçi, M. Ardiç, O. Koyuncu, *Wheat grain classification by using dense SIFT features with SVM classifier*, Computers and Electronics in Agriculture 122 (2016) 185–190.

[11] K. Sabanci, M. F. Aslan, A. Durdu, *Bread and durum wheat classification using wavelet based image fusion*, Journal of the Science of Food and Agriculture 100 (15) (2020) 5577–5585.

[12] K. Sabanci, A. Kayabasi, A. Toktas, *Computer vision-based method for classification of wheat grains using artificial neural network*, Journal of the Science of Food and Agriculture 97 (8) (2017) 2588–2593.

[13] A. Taner, Y. B. Öztekin, A. Tekgüler, H. Sauk, H. Duran, *Classification of varieties of grain species by artificial neural networks*, Agronomy 8 (7) (2018) 123 14 pages.

[14] D. Agarwal, Sweta, P. Bachan, *Machine learning approach for the classification of wheat grains*, Smart Agricultural Technology 3 (2023) 100136 10 pages.

[15] A. Sharma, T. Singh, N. Garg, *Combining near-infrared hyperspectral imaging and ANN for varietal classification of wheat seeds*, in: Manu Malek (Ed.), 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT), Kannur, 2022, pp. 1103–1108.

[16] A. Khatri, S. Agrawal, J. M. Chatterjee, *Wheat seed classification: Utilizing ensemble machine learning approach,* Scientific Programming 2022 (1) (2022) Article ID 2626868 9 pages.

[17] S. Shedole, S. B. J, N. A. V. P, *A convolution neural network-based wheat grain classification system*, Journal of Scientific Research of The Banaras Hindu University 66 (2) (2022) 204 22–29.

[18] S. Lingwal, K. Bhatia, M. Tomer, *Image-based wheat grain classification using convolutional neural network,* Multimedia Tools and Applications 80 (28) (2021) 35441–35465.

[19] A. Kayabasi, *An application of ANN trained by ABC algorithm for classification of wheat grains*, International Journal of Intelligent Systems and Applications in Engineering 6 (1) (2018) 85–91.

[20] W. Zhao, S. Liu, X. Li, X. Han, H. Yang, *Fast and accurate wheat grain quality detection based on improved YOLOv5*, Computers and Electronics in Agriculture 202 (2022) 107426 10 pages.

[21] N. Pervan Akman, M. Çolak, Ö. Özkan, T. Tümer Sivri, A. Berkol, M. Olgun, Z. Budak Başçiftçi, G. Ayter, O. Sezer, M. Ardıç, Wheat dataset for species classification and sunn pest damage detection (2023), https://data.mendeley.com/datasets/gmw48bvxdz/1, Accessed 30 June 2024.

[22] R. Shen, T. Zhen, Z. Li, *Segmentation of unsound wheat kernels based on improved mask RCNN*, Sensors 23 (7) (2023) 3379 17 pages.

[23] R. C. Bernardes, A. De Medeiros, L. da Silva, L. Cantoni, G. F. Martins, T. Mastrangelo, A. Novikov, C. B. Mastrangelo, *Deep-learning approach for fusarium head blight detection in wheat seeds using low-cost imaging technology,* Agriculture 12 (11) (2022) 1801 14 pages.

[24] M. F. Unlersen, M. E. Sonmez, M. F. Aslan, B. Demir, N. Aydin, K. Sabanci, E. Ropelewska, *CNN–SVM hybrid model for varietal classification of wheat based on bulk samples,* European Food Research and Technology 248 (8) (2022) 2043–2052.

[25] H. Gao, T. Zhen, Z. Li, *Detection of wheat unsound kernels based on improved ResNet*, IEEE Access 10 (2022) 20092–20101.

[26] J. B. Motie, M. H. Saeidirad, M. Jafarian, *Identification of sunn-pest affected (eurygaster integriceps put.) wheat plants and their distribution in wheat fields using aerial imaging*, Ecological Informatics 76 (2023) 102146 15 pages.

[27] Y. Abbaspour-Gilandeh, H. Ghadakchi-Bazaz, M. Davari, *Discriminating healthy wheat grains from grains infected with fusarium graminearum using texture characteristics of image-processing technique, discriminant analysis, and support vector machine methods*, Journal of Intelligent Systems 29 (1) (2019) 1576–1586.

[28] Z. Fazel-Niari, A. H. Afkari-Sayyah, Y. Abbaspour-Gilandeh, I. Herrera-Miranda, J. L. Hernandez Hernandez, M. Hernandez-Hernandez, *Quality assessment of components of wheat seed using different classifications models*, Applied Sciences 12 (9) (2022) 4133 12 pages.

[29] E. Kaya, İ. Saritas, *Towards a real-time sorting system: Identification of vitreous durum wheat kernels using ann based on their morphological, colour, wavelet and gaborlet features,* Computers and Electronics in Agriculture 166 (2019) 105016 9 pages.

[30] C. Erkinbaev, M. Nadimi, J. Paliwal, *A unified heuristic approach to simultaneously detect fusarium and ergot damage in wheat*, Measurement: Food 7 (2022) 100043 6 pages.

[31] L. Zhang, H. Ji, *Identification of wheat grain in different states based on hyperspectral imaging technology,* Spectroscopy Letters 52 (6) (2019) 356–366.

[32] W. Zhang, H. Ma, X. Li, X. Liu, J. Jiao, P. Zhang, L. Gu, Q. Wang, W. Bao, S. Cao, *Imperfect wheat grain recognition combined with an attention mechanism and residual network*, Applied sciences 11 (11) (2021) 5139 12 pages.

[33] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, H. Jegou, *Training data-efficient image transformers distillation through attention,* in: Marina Meila, Tong Zhang (Eds.), 38th International Conference on Machine Learning, Austria, 2021, pp. 10347–10357.