Pamukkale Univ Muh Bilim Derg, 27(5), 618-626, 2021

Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi

Pamukkale University Journal of Engineering Sciences

Wheat kernels classification using visible-near infrared camera based on deep learning

Derin öğrenmeye dayalı görünür yakın kızılötesi kamera kullanılarak buğday sınıflandırması

Kemal ÖZKAN[1], Erol SEKE[2], Şahin IŞIK[3]*

[1,3]Department of Computer Engineering, Faculty of Engineering and Architecture, Eskişehir Osmangazi Univ., Eskişehir, Turkey.
kozkan@ogu.edu.tr

[2]Department of Elt-Elk Engineering, Faculty of Engineering and Architecture, Eskişehir Osmangazi University, Eskişehir, Turkey.
eseke@ogu.edu.tr, sahini@ogu.edu.tr

Received/Geliş Tarihi: 07.05.2020 Revision/Düzeltme Tarihi: 01.09.2020 doi: 10.5505/pajes.2020.80774
Accepted/Kabul Tarihi: 16.10.2020 Research Article/Araştırma Makalesi

Abstract

This paper presents a smart machine learning system for classification of hyperspectral wheat data based on deep learning methodology. For this purpose, the performances of AlexNet and VGG16 models were investigated for the classification of hyperspectral wheat samples. In this study, the Support Vector Machine (SVM) and Softmax classifiers were carried out to predict labels of wheat kernels. In order to evaluate the system performance, a new hyperspectral wheat test dataset was constructed using Visible-Near Infrared images (VNIR) including 50 wheat species with 220 images per specimen, as 11000 samples in total. With experiments applied on newly created test dataset, overall approximated accuracy rates of 96.00% and 99.00% determined by linear SVM classifier, in case of fully connected layer (FC6 and FC7) features for AlexNet and VGG16, respectively. From the Softmax predictions, the 92% and 70% of samples were correctly discriminated based on trained VGG16 and AlexNet models, respectively. The obtained superior results show that using a deep Convolutional Neural Networks (CNN) architecture is more efficient by the means of accurate discrimination of wheat species. The proposed deep learning based categorization system promises high accuracy results for the quality analysis, classification and disease detection in food.

Keywords: Food engineering, Convolutional neural network, Classification, Visible-Near infrared, Wheat.

Öz

Bu makale, derin öğrenme metodolojisine dayalı hiperspektral buğday verilerinin sınıflandırılması için akıllı bir makine öğrenme sistemi sunmaktadır. Bu amaçla, hiperspektral buğday örneklerinin sınıflandırılması için AlexNet ve VGG16 modellerinin performansları araştırılmıştır. Bu çalışmada, buğday çekirdeklerinin türlerini tahmin etmek için Destek Vektör Makinesi (DVM) ve Softmax sınıflandırıcıları kullanılmıştır. Sistem performansını değerlendirmek için, Görünür Yakın Kızılötesi Görüntüleme (VNIR) kullanılarak 50 buğday türüne ait tür başına 220 görüntü toplamda 11000 örnek içeren yeni bir hiperspektral buğday test veri kümesi oluşturulmuştur. Yeni oluşturulan test veri seti üzerinde yapılan deneylerde, AlexNet ve VGG16 için tamamen bağlı katman (FC6 ve FC7) özellikleri kullanılması durumunda doğrusal DVM sınıflandırıcısı tarafından belirlenen yaklaşık %96.00 ve % 99.00'lık genel doğruluk oranları elde edilmiştir. Softmax sınıflandırıcı ile numunelerin sırasıyla %92 ve %70'i, eğitimli VGG16 ve AlexNet modellerine göre doğru bir şekilde ayırt edilebilmiştir. Elde edilen üstün sonuçlar, derin bir Evrişimsel Sinir Ağları (ESA) mimarisi kullanmanın, buğday türlerinin doğru bir şekilde ayırt edilmesi yoluyla daha verimli olduğunu göstermektedir. Önerilen derin öğrenme temelli sınıflandırma sistemi, gıdalarda kalite analizi, sınıflandırma ve hastalık tespiti için yüksek doğrulukta sonuçlar vaat etmektedir.

Anahtar kelimeler: Besin mühendisliği, Evrişimsel sinir ağları, Sınıflandırma, Görünür yakin kızılötesi, Buğday.

1 Introduction

Wheat is one of the main grain used in food production to supply human nutrition. This fact leads to initiatives at increasing reserved land size used for planting and production of wheat cultivars under various climate and soil conditions. Due to its valuable nutritional content, wheat grain is the main food-source of approximately 50 countries. Besides, improved technologies on convenient storage enable more intensive use of wheat in food industry as well as bakery products.

Wheat production in Turkey is also over its consumption as nearly 9.4 million hectares of wheat planted with some fluctuations in production related to climatic conditions. According to some statistics provided by a local agricultural research institute [1], Turkey's wheat consumption is approximately 19-21 million tons per year. Meanwhile, acquisition of high and qualified crops from the unit area depends on the timely struggles of farmers with good soil preparation, cultivation, conscious fertilization and sensible control on weeds, pests and diseases.

One of the most important factors in the production of wheat, among many, is the quality of the wheat plant itself. The quality decreases with the factors that reduce the seed value. It means that the quality of seeds may diminish due to contamination of foreign matter, diseases and inadequate storage. Type confusion can reduce quality of wheat; consequently, classification of the wheat plays an important role. Therefore, an unbiased recognition system is vital to classify the wheat samples prior to analysis and quality assessment. Automated classification methods reduce dependence on the human factor and labor. Automated classification techniques with the smart machine learning systems are more reliable, fairer and faster when compared with experts' decisions. Furthermore, identification of the varieties allows certain types of

*Corresponding author/Yazışılan Yazar

requirements in the storage process. Thus, such classification machines are advantageous in all aspects; much faster, highly reliable, more economical, lesser yield loss, and better production.

The rest of the paper is organized as follows. Section 2 presents the related previous works on wheat classification. In section 3, detailed information on the dataset and the utilized VGG16 and AlexNet models are given, including their internal mechanisms. Fourth section presents the experimental results. Conclusion and additional comments given in the last section.

2 Related work

In the past decades, with fast and non-destructive analysis abilities provided by digital imaging, various imaging approaches have been carried out in order to determine the wheat class, quality, disease and foreign matter. Nowadays, use of hyperspectral imaging in the classification and identification tasks is broadening. Technically, a special camera that capture images in various regions in the electromagnetic spectrum rather than the wavelengths visible to human eye generates the Hyperspectral Image (HSI). Although the human eye is sensitive to wavelengths between ~400 to ~700 nm, many living creatures in nature have eyes with much wider spectral range that can be considered as hyperspectral. Hyperspectral imaging conveys analysis of the images obtained by combining multiple wavelength ranges. Wheat images with different wavelengths have been subject to research through various methods and classifications for estimation of its protein, gluten and moisture value, virtuousness as well as its identity [2, 3]. The following briefly discusses previous research on wheat.

In order to overcome identification challenges faced in the grain industry, the usefulness of NIR spectrometry combined with chemometry is explored for wheat verification purpose in [4]. The dataset is consisted of 400-2500 nm spectral images of 249 wheat cultivars. The capabilities of Principle of Component Analysis (PCA) and a particular Mahalanobis (GH) distance are investigated to construct an identification system for wheat varieties. Eventually, the robustness of system has reached to an accuracy of 99.5% for the calibration sample and the success rate was 94% for the validation sample set.

Despite the fact that there are high inter-class similarities along with large intra-class variations in wheat samples, their chemical structure is different. Considering this, researchers used Near-Infrared Spectrometry (NIR) with wavelengths ranging from 960 nm to 1700 nm in order to more accurately distinguish wheat classes in Canada [5]. Using a NIR camera, hyperspectral samples are scanned at 10 nm intervals and Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis methods (QDA) were applied on presented dataset and achieved a success ratio that accounts for 86% in the classification process.

In a study [6] on disease identification from wheat images, the authors investigated wheat plants suffering from insect damage by Near-Infrared Reflectance (NIR) hyperspectral imaging technique. The key motivation was to develop an alternative solution for insect damage, which is one of the destructive factors for wheat quality. 1000-1600 nm wavelength range was considered and 6 statistical and 10 histogram features were extracted from these images. Later, statistical discriminant classifiers (Linear, Quadratic and Mahalanobis) were used to evaluate the system capability on classification. Healthy and insect-damaged wheat kernels were spotted successfully in 85-100% range with linear and quadratic discriminant models.

In another study [7], some initiatives were attempted at improving the quality detection of the flour using parameters of protein, moisture, Zeleny sedimentation and farinograph water absorption by using NIR (400-2498 nm) camera. For this purpose, they considered flour samples of wheat varieties grown in 79 different regions of Turkey and analyzed them both chemically and on NIR spectrometer device. Artificial Neural Network (ANN) model was used in this analysis. The high accuracy rate obtained implies that the utilized framework is useful for detection of certain contents including moisture, protein, alveograph, dough tenacity and strength, farinograph water absorption and Zeleny sedimentation. Another reported point is that ANN classifiers were convenient to generate efficient predictions about wheat flour quality.

An automated machine vision system for classification of wheat and barley grains was developed by determining hand-crafted features in [8]. The target classification objects were divided into two classes for training and testing stages. In the test phase, Discriminant Analysis (DA) and K-nearest neighbors (K-NN) methods were used in the prediction system. Ten images were used in training for each wheat and barley grain. Each of these images contained about 60 pieces of wheat and barley. In total, 545 wheat and barley were used for the test. These visible wavelength images had a resolution of 5 megapixels (2,560 x 1,920 pixels). From each grain, 99 features were derived; 21 morphological, 6 color and 72 texture features. After the normalization and the classification with these features, they achieved approximately 99% success rate.

Efforts aiming at infusing a machine learning system into fungicide treatment on durum wheats [9], an identification system on 213 durum wheat samples grown on different conditions in four different regions of southern Spain is developed. The VIS + NIR (400nm-2500nm) and NIR (1100nm-2500nm) imaging techniques were carried out to construct the drum wheat dataset. Modified Partial Least Squares method (MPLS) was used on the obtained spectral image data. The numerical performance results indicated that they achieved 84% success with the MPLS method. Based on the results of their study, they suggested that NIR Spectroscopy could succeed in determining fungicide treatment on durum wheat samples.

In [10], the capability of PCA based discrimination was investigated by integrating a machine learning tool for three different types including vitreous, yellow berry and fusarium-damaged. A specific NIR camera was used to capture spectral images in the wavelength range of 1000-1700 nm. According to the results, the lowest false classification belongs to the vitreous class with less than 1%. The error rates were reported as low for the other two, yellow berry results being slightly better than fusarium damage.

To classify wheat seeds and flour samples produced in different regions of Chile and estimate their geographical origins [11], a Near-Infrared Spectroscopy (NIRS) and chemometric methods were combined to analyze the relationships between extracted features and identity of wheat samples. They used the data obtained from NIR spectroscopy in the model recognition method. The authors carried out the Discriminant Partial Least Squares (DPLS) method to classify the samples. They reportedly achieved 76% success in classifying wheat grains and classified geographical origins of flour samples with an

accuracy between 90% and 96%. Results show that the method can successfully classify wheat grain and flour samples.

PCA and Multiple Linear Regression (MLR) methods, once again, were utilized to obtain favorable results in the quality analysis of wheat seeds and fungal harm based on multispectral imaging and chemometry [12]. In the developed system, the wavelengths between 360 and 950 nm were considered to create a multispectral wheat dataset. For this purpose, a CCD camera and a LED that can emit light with 8 different wavelengths were utilized during image capturing stage.

In a recent study [13], a NIR camera was utilized for a faster and more powerful system. Authors made experiments with five kinds of wheat; bread, spelled, durum, emmer and einkorn. Using Partial Least Squares Discriminant Analysis (PLSDA) in the classification, they obtained 80-100% accuracies in their study.

For the identification and classification processes in industrial applications, there has been a great deal of interest on the deep learning ideas, especially CNN, due to higher accuracy rates and even higher accuracy they promise when compared with traditional methods. A series of question pertaining to plant phenotyping was addressed in study of [14] with two experimental evaluation on modern CNN architectures. In the first evaluation, the plant root tip was aimed to be identified from a part of given root, as root architecture directly reveal the important physiological functions. In the second evaluation, from a given section of plant shoot, biologically expressive features such as leaf tip, leaf base, ear tip and ear base were aimed to be classified with modern CNNs. In another work [15], the spatio-spectral data (900-1700 nm spectral range) related to rice types was processed for classification task. The hyperspectral data consists of two spatial dimensions and one spectral dimension for paddy and processed rice types. For model construction, performances of VGGnet [16], ResNET and ResNET-B [17] were compared with SVM classifier. In [18], two different spectral ranges from 441–948 nm (Spectral range 1) and 975–1646 nm (Spectral range 2) were extracted and forwarded to CNN method for recognition of single rice seed. In [19] though, the RGB images of maize kernels were used to train a deep convolutional neural network architecture, namely ResNet. Promising results were obtained for binary classification problem of qualified or defective.

In [20], the Shallow and VGG16 models were trained on VNIR images of wheat samples, for 40 classes and 200 images per class. A more systematic and theoretical analysis were carried out in [21], for wheat identification from SWIR, VNIR and RGB channels, for 40 classes. Similarly in [16], different feature vector fusion methodologies were applied after extracting the features (1x4096) from FC6 layer of Visual Geometry Group (VGG16), for SWIR, VNIR and RGB channels. Performance improvements were observed by different combinations of RGB, SWIR and VNIR channels.

Technically, it is very challenging, if not impossible, to distinguish one class from others even with the representative samples through naked eye image analysis of wheat species. This is no surprise since human eye is relatively sensitive to few gray levels and is not able to perceive fine spatial details contained in images. Because of high interclass similarity, eye examination often results in incorrect categorization. Therefore, it is greatly required to develop algorithms for machine-aided decision making that incorporates features that human eye is not good at. Hence, the task of classification of hyperspectral images with VNIR wavelength, a type of input humans is not trained for either, is also a good candidate for machine-aided recognition/classification purposes.

In the presented study, new hyperspectral dataset has been created using images of wheat species captured by a Visible-Near Infrared (VNIR, wavelength: 400-1100 nm) camera. The dataset contains 50 wheat classes each containing 220 images (a total of 11,000 images). Details of the experiments are given in the related sections. Wheat sample images are fed to the AlexNet and VGG16 models for training massive number of filters with increasing depth in given CNN architectures. The performance of system is validated through further evaluations with accuracy and F-score metrics of Support Vector Machine (SVM) (linear kernel) and Softmax classifiers. Highly accurate classification results are achieved and compared with state of art methods.

3 Materials and methods

3.1 Summary of deep learning: CNNs features

Pattern recognition applications rely on some discriminative features that can be used to robustly distinguish one pattern from another, and therefore, researchers try to determine such features. However, practical implementation layouts present several challenges that may prevent successful feature acquisition. In image based pattern recognition, for example, illumination changes, scaling and occlusion can hinder the pattern's view and introduce vaguely defined problems, weakening the possibility of reaching ultimately successful decisions. With these in mind, the utilized feature extraction method should maximize the intra-class similarity while minimizing the inter-class similarity. Supporting this argument, the Local Binary Pattern (LBP) [22], Histogram of Gradient (HOG) [23], Scale Invariant Feature Transform (SIFT) [24] and Dense SIFT Features (DSIFT) are being used together with Bag of Words (BOW) model. BOW is the quantization step on extracted visual features (words) following a k-means clustering process. Quantization provides manageable size D words out of millions of visual words (called pool of visual words). A histogram for each image is built after determining similarity between visual features and related clusters. Thus, an image is represented with a histogram whose dimension is D. Training data is run through a linear/nonlinear classifier (e.g. Support Vector Machine (SVM)). A histogram is obtained for each test sample and their labels are determined by the classifier. This summarizes almost all machine-learning systems reported. In efforts to improve the performance of machine-learning systems for big datasets, the deep CNNs have been exploited with the study of Krizhevsky et. al., namely AlexNet. AlexNet earned its reputation by yielding impressive results in object recognition. With the successful implementations of deep CNNs, the era of traditional feature extraction NN strategies are almost ended, revolutionizing machine learning. No sooner than the reputation of AlexNet, the VGG16 was developed based on the confidence limit on the accuracy and enhanced depth of weights.

3.2 The featured deep learning models

The general flow of feed forward process of AlexNet was illustrated in Figure 1. The original AlexNet architecture accepts 227x227 color images and generates decisions for 1000 classes through 21 intermixed specially designed layers as illustrated. AlexNet details can be accessed through definitive references for image categorization like AlexNet [25].

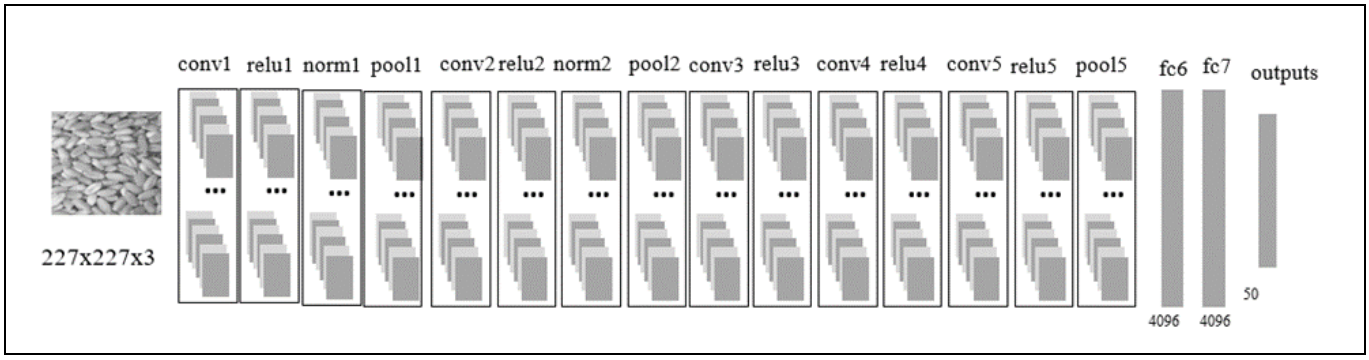


Figure 1. The overall flow of trained AlexNet model.

When comparing the internal mechanism of VGG16 and AlexNet, it is obviously clear that the VGG16 is much deeper as it consists of 16 layers whereas the AlexNet has eight layers. One may assume that a convolution filter refers to a linear classifier. In this regard, the VGG16 is more effective than AlexNet since focusing on the larger receptive fields directly improves the potential capability of identification process. However, the larger receptive fields also inevitably give rise to a trade-off between high accuracy rate and millions of parameters that are required to be trained. Therefore, from the comparison of hardware resources consumption and computational cost of both models, one may infer that the AlexNet is fastest and more efficient in terms of use of resources in offline or in runtime. However, the VGG16 dominates the performance of AlexNet when applied for wheat identification.

3.3 Hyperspectral dataset

The dataset, consisting of 50 classes, constructed from the controlled environment photos of widely cultivated wheat seeds in Eskişehir, Turkey. Figure 2 shows some visual representations along with associated labels. For each class, 55 instances of size 640x512 are captured and the number of images increased to 220 by dividing each into four quadrants. The images were captured using MT-CL-CAM-VNIR-6415 (Micro-Tasarim Co.) camera. The camera is Visible-Near Infrared (VNIR) and is sensitive to 400 nm to 1100 nm range in the spectral band, generating a gray-level image as a sum of irradiance fields of wavelengths in this range.

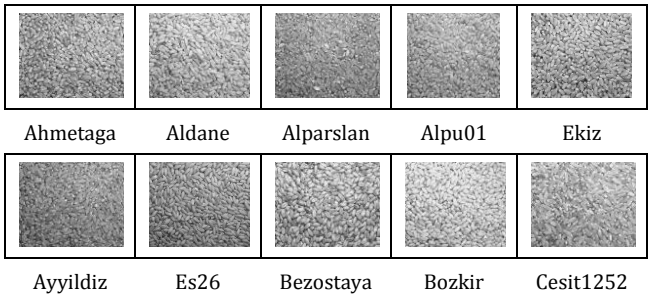


Figure 2. Visual demonstration of hyperspectral samples for some wheat species.

A controlled environment was set up, as shown in Figure 3(a-b). The VNIR Camera was positioned at about 30 cm above the samples, viewing approximately 9x7 cm^2 area on the sample plane. Two 100-W halogen light fixtures, having approximately 45º angle of incidence on both sides of the main axis, illuminate the samples. Halogen lamps were selected since their emissions are in the interested spectral range. Exposure time was calibrated to 17 ms and 1-point correction applied to enhance the contrast. Technically, our image capturing system relies on the VNIR imaging, which is superimposed 400-1100 nm of the spectral band. The system gives a unique image, which is the average of normalized recorded channels. Therefore, the prediction system is applied on obtained unique images rather than the data on all bands. From the visualized samples in Figure. 2, we can observe that the appearances of wheat samples are quite similar to each other and lead to misclassification of wheat species with human eye. Therefore, we unveiled the discriminative details by utilizing the VNIR imaging, which is the combination of visible and near infrared spectrums.

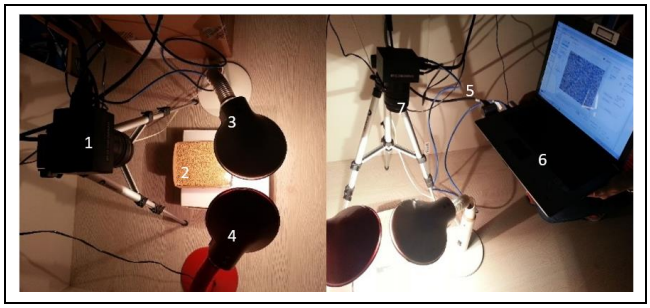


Figure 3. VNIR spectral image capturing environment. (1): VNIR camera. (2): Wheat samples. (3): Left light source (halogen lamp). (4): Right light source (halogen lamp). (5): Frame grabber cable. (6): Data acquisition computer and (7): Lens.

3.4 Proposed wheat verification system

Specifically, a CNN model provides high-level features in a gradual manner. This is the biggest advantage of CNN over classical NN. Most of the studies that are devoted for classification tasks, generally consider the outputs of FC7 layer for object classification, calling them the features. However, there are different possibilities for using a CNN architecture for object classification:

i. Training data with a newly developed CNN architecture,

ii. Using the concept of transfer learning as considering pre-trained CNN model with Softmax classifier,

iii. Using pre-trained CNN model and applying a fine-tuning procedure for the specified class labels,

iv. Retraining the weights of pre-trained CNN model and fitting a new classifier on extracted features,

v. Retraining the weights of pre-trained CNN model together with Softmax classifier.

In this study, we considered options (iv) and (v). For the case of (iv), we considered only FC7 and FC6 features obtained from AlexNet and VGG16 models that are trained on hyperspectral wheat dataset. For this purpose, FC8, the last head of CNN architecture, was omitted and FC6 and FC67 were considered as discriminative features to classify wheat kernels. By considering the concept of fine-tuning procedure, the SVM classifier, in addition to Softmax, was employed, on extracted features for determining labels of wheat kernels. The reason for removing the Softmax classifier is that the Softmax classifier suffers from weak discriminative capability when it comes to generating predictions in testing wheat samples. As a second experiment, we also compared the identification strength of Softmax and SVM classifiers to give general insights on their performances.

Data augmentation is common to increase the number of images in the datasets in deep learning studies. In this study, we have divided images into four quadrants, eventually, reaching to 55×4 = 220 images per class. Images are necessarily cropped to size 227x227, since AlexNet model is designed to work on such images. Similarly, for VGG16, the input samples are fed into the model in the form of 224x224. VGG16 and AlexNet models are evaluated on the wheat cultivars dataset of 50x220=11000 images for their classification performances. The proposed wheat identification framework is summarized in Figure 4. Technically, the SVM is a well-known classifier to generate good results for both linearly and nonlinearly separable classes. Decision on linearity is usually made on the characteristics of the data. However, for classification of FC6/FC7 features in this work, a linear kernel is implemented and tested. For SVM, 10-fold cross validation procedure is performed, checking for generalization ability of the utilized VGG16 and AlexNet models. For cross-validation, each partition contained 1100 testing and 9900 training (randomly selected half of the training set is reserved for validation) instances. Thus, randomly selected 90% of image dataset is utilized for training SVM classifier and the remaining ones are used for performance evaluation. For performance evaluation, the LIBSVM library [26] is used with the parameters of linear kernel and cost (C) being 10 in SVM simulations.

4 Results and discussion

To evaluate the performance of developed identification frameworks, the most common metric, accuracy, is used in practice to reveal the effectiveness of trained deep learning models. The accuracy rate is determined by using a resampling approach that is relied on the cross validation technique. The overall generalization capability of systems are measured by 10-fold cross validation.

Table 1 exhibits the classification results obtained from the cross validation steps followed by training simulations on linear kernel of SVM for FC6/FC7 features. In AlexNet architecture, one can clearly observe that the linearly classified FC7 features give 95.8700% overall accuracy result, which is slightly better than that of FC7 alone, once replacing training and test data 10 times in regard to cross validation concept. On the contrary to VGG16, there is a substantial performance improvement for FC7 features of AlexNet in comparison to FC6 feature set. The observed performance decline of AlexNet may be related to underestimation tendency of FC6 features.

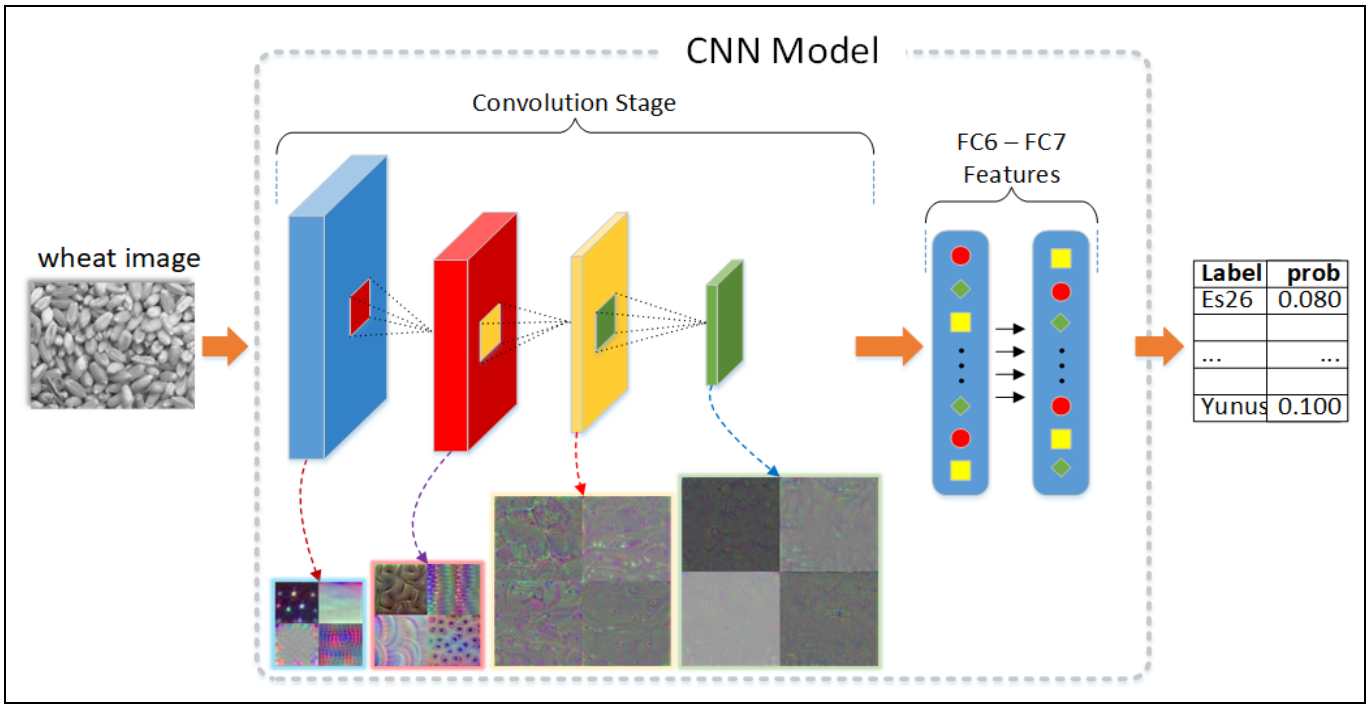


Figure 4. The utilized wheat verification frameworks.

Table 1. Linear kernel SVM scores on 1x4096 FC6 and FC7 features.

	FC6 Features		FC7 Features	
	Alexnet	VGG16	Alexnet	VGG16
CV01	0.9218	98.7909	95.8636	99.0473
CV02	0.9155	98.7891	95.9255	99.0164
CV03	0.9336	98.8418	95.7564	99.0582
CV04	0.9218	98.8691	95.9036	99.0127
CV05	0.9191	98.8164	95.8691	98.9782
CV06	0.9282	98.8455	95.8600	98.9673
CV07	0.9264	98.7382	95.9364	99.0418
CV08	0.9264	98.8655	95.8491	98.9545
CV09	0.9155	98.7818	95.7600	98.9545
CV10	0.9118	98.8455	95.9764	98.9491
Overall	0.9220	98.8184	95.8700	98.9980
Std	0.0064	0.0402	0.0675	0.0400

Moreover, the standard deviation between cross-validation accuracies of VGG16 are 0.0675 and 0.0400, for FC6 and FC7 features, respectively. The lower standard deviation rate validates the use of pre-trained deep learning models to scale for other classification tasks.

In general, the use of FC7 instead of FC6 features generates similar results in terms of accuracy and standard deviation. This can be explained by; (i) FC6 results are already accurate and cannot be numerically improved much, (ii) the number of classes is too small to require 4096 features and descriptive elements are concentrated in a few components. In case of AlexNet, most of FC7 features are essentially not valuable for a good classification and FC6 cannot take advantage of these.

Since the accuracy metric involves some limitations by means of exploring discrimination ability and biasing tendency of a utilized model, we also applied the F-score metric to assess the system performance. In particular, the F-score metric allows us to evaluate the precision and recall traits of proposed deep learning models.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

The precision and recall measures were computed with equation (1) and (2), respectively. Conceptually, the F-score is the harmonic mean of precision and recall rates as indicated with Equation (3). TP is the number of correctly classified wheat cultivar, whereas the FP refers to the number of incorrectly classified wheat cultivar. Again, the FN is the number of unexpected wheat cultivar that is misclassified as expected wheat cultivar. From the theoretical background, one can infer that the low rate of FP and FN versus high TP rate implies a good classification strength for models. A high value of F-score (1) indicates that learned weights are better to linearly separate wheat cultivars, in contrast, a low value of F-score (0) means that the weights have some weaknesses for discrimination.

The Figure 5 provides a graphical representation of F-score values returned from Softmax classifier for each processed class. The testing simulations are conducted on the FC6 features of VGG16 and AlexNet models by considering predicted probabilities determined from the last activation layer of models, namely Softmax classifier. On the basis of measured F-scores in Figure 5, one may note that VGG16 contributes more discriminative features for accurate verification of wheat varieties as compared to ones returned from AlexNet.

The VGG16 based discrimination system achieves very good F-score value with 92.06% performance rate which is considerably better than the AlexNet based system, which accounts for 69.74% overall. This observation matches up with the well-known argument that is "the deeper is the better" in terms of discrimination functionality. Focusing on the obtained results, the AlexNet fails on classification of c22 (harmankaya99) and c23 (hendrix) wheat cultivars. This is probably the result of AlexNet reducing the input image size down to 55x55x96 followed by first convolutional layer with 11x11 filters applied with stride 4. Also, performing a max-pooling process after the first convolutional layer leads to information loss and poor generalization ability.

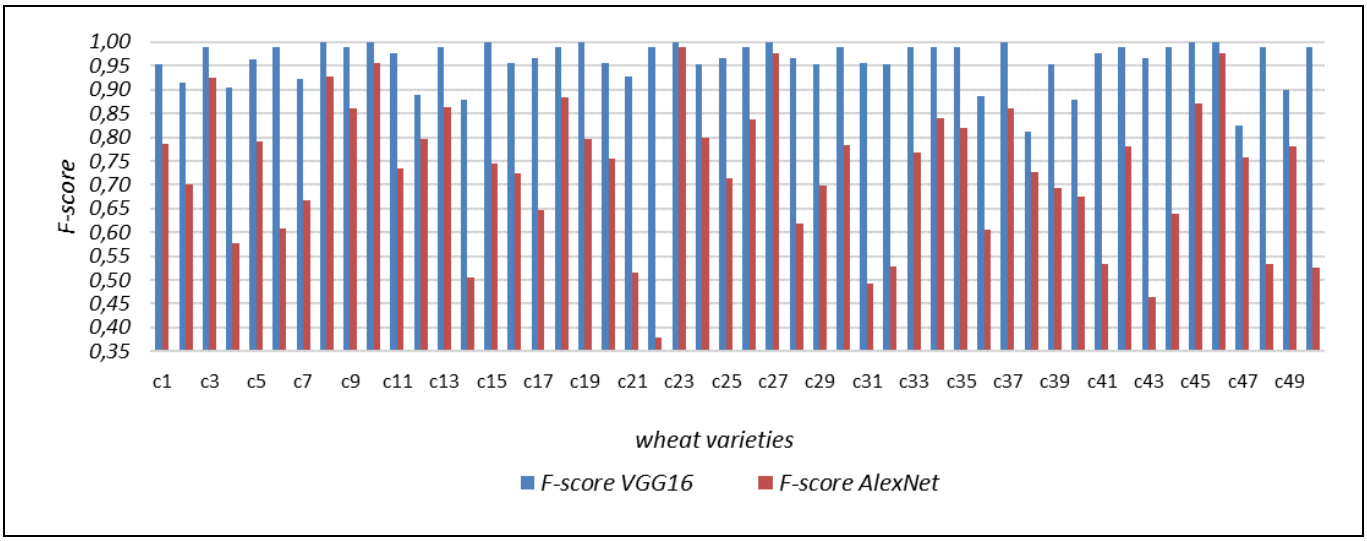


Figure 5. Analysis of measured F-score values for each wheat cultivar.

In order to avoid such losses, VGG16 gradually degrades the data size through the layers. Therefore, VGG16 achieves very pleasing *F*-score results for c31 (konya2002), c34 (meram) and c35 (meram2002) even though their origin is the same. This interesting point implies that deep learning based models can accurately identify the wheat varieties that share similar characteristics and are cultivated in the same region. Another comment can be made with respect to the results shown in Figure 6, is that there is a balance between precision and recall values of VGG16. One may also depict that the utilized image capturing process in a controlled environment conclusively helped to have this general result.

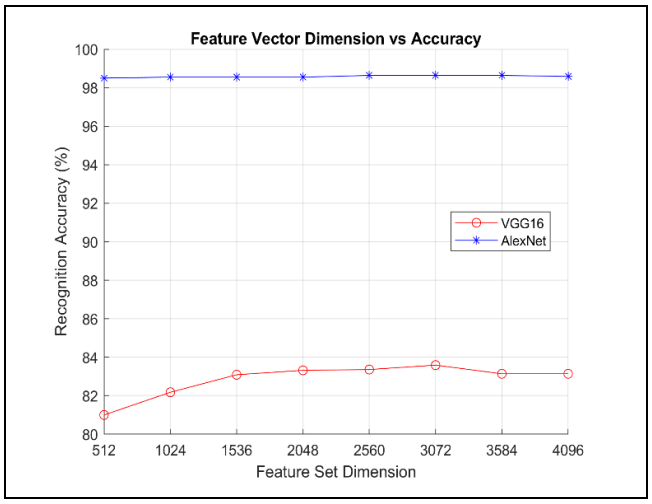


Figure 6. Analyzing impact of different dimension of CNN features for categorization of wheat cultivars.

To investigate the relationship between dimension of the feature set and effectiveness of models, we present a further analysis on performances of utilized AlexNet and VGG16 models together with SVM classifier. The impact of eight distinct dimensions (1x512, 1x1024, 1x1536, 1x2048, 1x2560, 1x3072, 1x3584, and 1x4096) of FC6 features has been evaluated to enhance credibility of the wheat verification framework. 1xU refers to the first consecutive U of 4096 FC7 features. For this experiment, 2200 wheat instances are randomly reserved for testing while 8800 are considered for training the models. In case of training simulations, the SVM one-against-all methodology is conducted to avoid under and over-fitting issues. The best accuracy values are around 98.64% for VGG16 and 83.59% for AlexNet architectures, when the dimensions of feature set chosen as 1x3072. Surprisingly, there is a little variation in accuracy values of 1x512 and 1x4096 feature sets, in case of both AlexNet and VGG16. It means that one can use 1x512 feature dimension within the confidence limits of recognition performance and computational cost.

For fairness, a more detailed comparison with approaches that were presented in recent 3 years is provided. A technical background about their datasets (i.e., number of grain types and number of samples in total) is also given, considering the fact that the accuracy score depends on the utilized number of samples in training and testing. In our previous work [27], we analyzed the effectiveness of hand-crafted features, namely histogram of descriptors around local regions of wheat images together with SVM classifier. Our previous system relied on the visible-infrared imaging (RGB) technique and achieved 88.30% accuracy rate for 40 classes comprising 6400 samples. Identification of five grain types including bread wheat, durum, aircorn, emmer and spelt was investigated by industrial NIR spectroscopy [13]. The overall accuracy rate was 84.56% over 1512 samples. Again, in another work [28], the potential capability of PCA was explored to categorize the three maize types (hard, medium and soft) with 93.00% accuracy rate. Sharing similar theoretical background with proposed methodology, the discriminative and predictive power of a deep learning method (CNN) was analyzed with image-based plant phenotyping in [14].

Over limited number of classes, the 97.30% success rate is reported through the prediction model following their developed CNN architecture. Apart from the CNN features, the contribution of shape descriptors are reported by using X-ray imaging along with the PCA methodology for discrimination of wheat varieties [2]. 89.97% accuracy is reported for particular three wheat varieties. In [18], four rice kernel varieties are attempted to be identified through a similar CNN with hyperspectral images with overall accuracy of 87% for 20907 samples. In another similar work [15], 86.30% recognition rate is reported after the integration of CNN into hyperspectral imaging based system for classification of the rice varieties. A comparable accuracy of 94.30% is achieved in [3] for discrimination of durum and common wheat seeds. Indeed, specific image acquisition techniques seem to play an important role in classification. However, the applied discrimination model, rather than imaging system, provides more contributions when it comes to develop a real time diagnostic system. Conducting experimental simulations on a few grain types is infeasible and inefficient by means of generalizability and transportability. Focusing on the presented results, one may observe that our proposed method gives accuracy values that are among the highest. The discrimination capability of VGG16 is reported with almost 99.00% accuracy rate, which dominates the strength of handcrafted features that were evaluated in our previous study [27].

Another comprehensive study [21] was designed to determine the effects of image fusion and feature fusion for wheat identification. It is found that combining FC6 features of RGB and VNIR images shows a significant increase in performance, which reaches to 100% accuracy. However, it is a computationally costly procedure to fuse the RGB and VNIR features. Comparison of the findings with study of [21] confirms that FC6 features of a CNN algorithm give remarkable identification scores (approximately 99.00%) when trained with linear SVM on VNIR image sets.

The key contribution of our method is that we reached similar detection scores on a more comprehensive dataset rather than limited number of grain types. This can be counted as weakness of other presented methods. We can observe from the Table 2 that most of the studies in the field of grain (wheat, maize or rice) classification have only focused on small number of classes, which is between the range of 2 and 5, whereas the proposed method focuses on 50 wheat types. The generalizability of research with insufficient data on the classification is rather difficult. In contrast to other CNN based methods, this research has investigated the contributions of FC6 and FC7 features for wheat identification. The findings of this study indicates that massive depths of a deep learning architecture generate successful discrimination of 50 wheat species combined with VNIR imaging.

Table 2. Performance analysis of some methods.

Target Grain	Imaging Mode	Wavelength Range (nm)	# of Samples	# of Classes	Accuracy (%)	Reference
wheat	VNIR	400-1100	11000	50	99.00	Proposed
wheat	VNIR	400-1100	8000	40	99.38	Ozkan et al. [21]
wheat	Visible	400-700	6400	40	88.30	Olgun et al. [27]
wheat	NIR	1200-2400	1523	5	84.56	Ziegler et al. [13]
maize	NIR	1101-2503	20-40	3	93.00	Williams et all [28]
wheat	Visible	400-700	62118	5	97.30	Pound et al. [14]
wheat	X-ray	-	288	3	89.97	Charytano et al. [2]
rice	VNIR - NIR	441-1646	20907	4	87.00	Qiu et al. [18]
rice	NIR	900-1700	1656	4	86.30	Chatnunt et al. [15]
wheat	NIR	1100-2400	4112	2	94.80	Vermeulen et al. [3]

5 Conclusion

Deep learning features have put many concepts in computer vision in place over the past few years and have built their own reality. High accuracy rates of the deep learning methods are attributed to cascaded convolution layers, that is, filters are evolved gradually through each layer. It seems that the next generation studies would be carried out based on newly developed deep learning ideas. From this fact, we have applied CNN including AlexNet and VGG16 features on hyperspectral images, as a scope of study. Eventually, wheat varieties are successfully parsed and achieved very high accuracy rates in the experiments on 50 wheat varieties with 11000 samples in total. By examining results, one can note that accuracies obtained from FC6 outperforms FC7, with a linear SVM classifier together with AlexNet architecture. Comparing VGG16 and AlexNet, we conclude that VGG16 provides superior identification performance with almost 99.00% score, whereas AlexNet reaches the approximately 96.00% accuracy rate when modelling SVM on returned features. Furthermore, the performances of VGG16 and AlexNet reported as 92% and 70% accuracy rates with Softmax classifier.

The proposed smart system provides important benefits in terms of economy and labor. At the same time, it promises high accuracy results for the quality analysis, classification and disease detection in food industry. Moreover, the VNIR imaging facilitates the categorization of wheat samples when deep learning techniques are applied for reliable identification [29]. Future studies would be based on the different bandwidth intervals or would be focused on more significant wavelength values on electromagnetic spectrum.

6 Acknowledgment

This work was granted support by the Turkey Scientific and Technological Research Projects Support Program with project number 1160576. This work was also supported by the Eskisehir Osmangazi University Scientific Research Project Commissions (Grant No.: 2016-1120). We would like to thank to Mikro-Tasarım Company and Eskişehir Commodity Exchange Laboratory for their contributions in material, equipment and experience.

7 Author contribution statements

In the scope of this study, the contributions of the authors are as follows; Kemal ÖZKAN, in formation of the idea, design and literature review; Erol SEKE, in reviewing the article structure; Şahin IŞIK, in collecting data, implementing utilized methods, assessment of obtained results, writing, supplying the materials used and examining the results.

8 Ethics committee approval and conflict of interest statement

No stage of this study required permission from the ethics committee.

There is no conflict of interest with any person / institution for this article and preparation stages thereof.

9 Kaynaklar

[1] Agricultural Research Institute. "Directorate of Trakya Agricultural Research Institute". https://Arastirma.Tarimorman.Gov.Tr/Ttae/Sayfalar/Detay.Aspx?Sayfaid=47 (26.03.2021).

[2] Charytanowicz M, Kulczycki P, Kowalski PA, Łukasik S, Czabak-Garbacz R. "An evaluation of utilizing geometric features for wheat grain classification using x-ray images". *Computers and Electronics in Agriculture*, 144, 260-268, 2018.

[3] Vermeulen P, Michele S, Juan PFA, Vincent B. "Discrimination between durum and common wheat kernels using near infrared hyperspectral imaging". *Journal of Cereal Science*, 84, 74-82 2018.

[4] Miralbés C. "Discrimination of european wheat varieties using near infrared reflectance spectroscopy". *Food Chemistry*, 106(1), 386-389, 2008.

[5] Mahesh S, Manickavasagan A, Jayas D, Paliwal J, White N. "Feasibility of near-infrared hyperspectral imaging to differentiate canadian wheat classes". *Biosystems Engineering*, 101(1), 50-57, 2008.

[6] Singh C, Jayas D, Paliwal J, White N. "Detection of insect-damaged wheat kernels using near-infrared hyperspectral imaging". *Journal of Stored Products Research*, 45(3), 151-158, 2009.

[7] Mutlu AC, Boyaci IH, Genis HE, Ozturk R, Basaran N, Sanal T, Evlice AK. "Prediction of wheat quality parameters using near-infrared spectroscopy and artificial neural networks". *European Food Research and Technology*, 233(2), 267-274, 2011.

[8] Guevara FH, Gomez JG. "A machine vision system for classification of wheat and barley grain kernels". *Spanish Journal of Agricultural Research*, 9(3), 672-680, 2011.

[9] Soto MC, Gaitán AJ, Domínguez J. "Application of near infrared spectroscopy technology for the detection of fungicide treatment on durum wheat samples". *Talanta*, 97, 298-302, 2012.

[10] Serranti S, Cesare D, Bonifazi G. "The development of a hyperspectral imaging method for the detection of fusarium-damaged, yellow berry and vitreous italian durum wheat kernels". *Biosystems Engineering*, 115(1), 20-30, 2013.
[11] González MIM, Moncada GW, González CP, San NZM, López FG, Ortega IL, Hernández JMH. "Chilean flour and wheat grain: tracing their origin using near infrared spectroscopy and chemometrics". *Food Chemistry*, 145, 802-806, 2014.
[12] Jaillais B, Roumet P, Pinson-Gadais L, Bertrand D. "Detection of fusarium head blight contamination in wheat kernels by multivariate imaging". *Food Control*, 54, 250-258, 2015.
[13] Ziegler JU, Leitenberger M, Longin CFH, Würschum T, Carle R, Schweiggert RM. "Near-Infrared reflectance spectroscopy for the rapid discrimination of kernels and flours of different wheat species". *Journal of Food Composition and Analysis*, 51, 30-36, 2016.
[14] Pound MP, Atkinson JA, Townsend AJ, Wilson MH, Griffiths M, Jackson AS, Bulat A, Tzimiropoulos G, Wells DM, Murchie EH, Pridmore TP. "Deep machine learning provides state-of-the-art performance in image-based plant phenotyping". *Gigascience*, 2018. https://doi.org/10.1093/gigascience/giy042.
[15] Chatnuntawech I, Tantisantisom K, Khanchaitit P, Boonkoom T, Bilgic B, Chuangsuwanich E. "Rice classification using spatio-spectral deep convolutional neural network". *ArXiv,* 2018. https://arxiv.org/abs/1805.11491.
[16] Simonyan K, Zisserman A. "Very deep convolutional networks for large-scale image recognition". *ArXiv,* 2014. https://arxiv.org/abs/1409.1556.
[17] He K, Zhang X, Ren S, Sun J. "Deep residual learning for image recognition". *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, USA, 12 December 2016.
[18] Qiu Z, Chen J, Zhao Y, Zhu S, He Y, Zhang C. "Variety identification of single rice seed using hyperspectral imaging combined with convolutional neural network". *Journal of Applied Sciences*, 8(2), 1-12, 2018.
[19] Ni C, Wang D, Vinson R, Holmes M, Tao Y. "Automatic inspection machine for maize kernels based on deep convolutional neural networks". *Biosystems Engineering*, 178, 131-144, 2019.
[20] Özkan K, Isik S, Yavuz B, Demirez DZ. "Shallow and deep convolutional neural network models for classification of VNIR wheat samples". *Fifth International Conference on Engineering and Natural Sciences*, Prague, Czech Republic, 12-16 June 2019.
[21] Özkan K, Isik S, and Yavuz BT. "Identification of wheat kernels by fusion of RGB, SWIR, and VNIR samples". *Journal of the Science of Food Agriculture*, 99(11), 4977-4984, 2019.
[22] Ahonen T, Hadid A, Pietikainen M. "Face description with local binary patterns: application to face recognition". *IEEE Transactions on Pattern Analysis Machine Intelligence*, 28(12), 2037-2041, 2006.
[23] Dalal N, Triggs B. "Histograms of oriented gradients for human detection". *Fifth IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, USA, 20-25 June 2005.
[24] Lowe DG. "Distinctive ımage features from scale-invariant keypoints". *International Journal of Computer Vision*, 60(2), 91-110, 2004.
[25] Krizhevsky A, Sutskever I, Hinton GE. "Imagenet classification with deep convolutional neural networks". *Advances in Neural Information Processing Systems*, 25, 1097-1105, 2012.
[26] Chang CC, Lin CJ. "LIBSVM: a library for support vector machines". *Journal of ACM Transactions on Intelligent Systems Technology*, 2(3), 1-27, 2011.
[27] Olgun M, Onarcan AO, Özkan K, Isik S, Sezer O, Özgişi K, Ayter NG, Başçiftçi ZB, Ardiç M, Koyuncu O. "Wheat grain classification by using dense SIFT features with SVM classifier". *Computers and Electronics in Agriculture*, 122, 185-190, 2016.
[28] Williams PJ, Kucheryavskiy S. "Classification of maize kernels using NIR hyperspectral imaging". *Food Chemistry*, 209, 131-138, 2016.
[29] Fan Y, Ma S, Wu T. "Individual wheat kernels vigor assessment based on NIR spectroscopy coupled with machine learning methodologies". *Infrared Physics Technology*, 2020. doi:10.1016/j.infrared.2020.103213.