Gazi Üniversitesi
Fen Bilimleri Dergisi
PART C: TASARIM VE TEKNOLOJİ

Gazi University
Journal of Science
PART C: DESIGN AND TECHNOLOGY

GU J Sci, Part C, 13(2): 653-664 (2025)

Dynamic Voting-Based Ensemble Deep Learning for Closely Resembling Crop Classification

Engin EŞME[1*] Muhammed Arif ŞEN[1] Halil ÇİMEN[1]

[1]*Konya Technical University, Artificial Intelligence Application and Research Center, Konya, Turkey*

Article Info

Research article
Received: 04/02/2025
Revision: 11/03/2025
Accepted: 17/06/2025

Keywords

Deep Learning
Dynamic Voting
Crop Classification
Smart Agriculture Technologies

Makale Bilgisi

Araştırma makalesi
Başvuru: 04/02/2025
Düzeltme: 11/03/2025
Kabul: 17/06/2025

Anahtar Kelimeler

Derin Öğrenme
Dinamik Oylama
Mahsul Sınıflandırma
Akıllı Tarım Teknolojileri

Graphical/Tabular Abstract (Grafik Özet)

To classify closely resembling agricultural crops, this study uses a dynamic voting approach. A meta-dataset, created from 17 model predictions, is trained with a CNN to select the most suitable models for each instance, enhancing accuracy and stability. / Bu çalışma, birbirine çok benzeyen tarımsal mahsulleri sınıflandırmak için dinamik bir oylama yaklaşımı kullanmaktadır. 17 modelin tahminlerinden oluşturulan bir meta-veri seti, her bir örnek için en uygun modelleri seçmek üzere bir CNN ile eğitilerek doğruluk ve kararlılık artırılır.

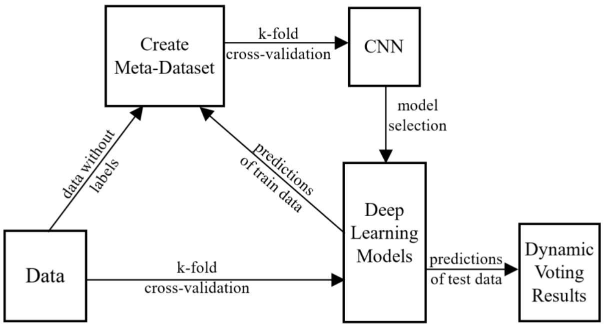

***Figure A**: Dynamic voting approach / **Şekil A:** Dinamik Oylama Yaklaşımı*

Highlights (Önemli noktalar)

- *This study addresses the problem of classifying closely resembling agricultural crops, which is a challenging task. / Bu çalışma birbirine çok benzeyen ve bu nedenle sınıflandırılması zor olan tarımsal mahsul türlerini ayırt etme problemini ele alır.*
- *The study presents a dynamic ensemble approach that intelligently selects the most expert models from among 17 candidates for each image to participate in the vote. / Çalışma, her görüntü için 17 model arasından en yetkin olanları akıllıca seçip oylamaya dahil eden dinamik bir topluluk yaklaşımı sunar.*
- *Dynamic voting addresses weaknesses in individual models, such as overfitting and performance inconsistency, to create a more robust, and stable, reliable classification system. / Dinamik oylama tekil modellerdeki aşırı öğrenme ve tutarsızlık gibi zayıflıkları gidererek daha sağlam, kararlı ve güvenilir bir sınıflandırma sistemi oluşturur.*

Aim (Amaç): *This research aims to make a significant contribution to the automatic recognition and classification of closely resembling agricultural crops by providing a theoretical foundation for the development of smart farming technologies. / Bu araştırma, akıllı tarım teknolojilerinin geliştirilmesi için teorik bir temel oluşturarak benzer görünümlü tarımsal ürünlerin otomatik olarak tanınmasına ve sınıflandırılmasına önemli bir katkı sağlamayı hedeflemektedir.*

Originality (Özgünlük): *The developed dynamic voting mechanism differs from fixed voting systems by selecting the most competent models for each test instance, thereby increasing classification accuracy and stability. / Geliştirilen dinamik oylama mekanizması, her bir test örneği için en yetkin modelleri seçerek sabit oylama sistemlerinden ayrışır ve bu yolla sınıflandırma doğruluğunu ve sistemin kararlılığını artırır.*

Results (Bulgular): *The proposed ensemble model achieved the best results across all performance metrics with 100% sensitivity and 99.7% specificity, minimizing the false positive rate and ensuring no positive instances were missed. / Önerilen topluluk modeli, %100 hassasiyet ve %99,7 özgüllük değerleriyle, yanlış pozitif oranını en aza indirip hiçbir pozitif örneği kaçırmayarak tüm performans metriklerinde en üstün sonucu elde etmiştir.*

Conclusion (Sonuç): *The proposed ensemble learning method addresses the weaknesses of individual models to provide a more robust classification system, preparing a practical foundation for the development of smart farming technologies. / Önerilen topluluk öğrenmesi yöntemi, bireysel modellerin zayıflıklarını gidererek daha sağlam bir sınıflandırma sistemi sunar ve akıllı tarım teknolojilerinin geliştirilmesi için pratik bir zemin hazırlar.*

*Corresponding author, e-mail: eesme@ktun.edu.tr
DOI: 10.29109/gujsc.1632938

GU J Sci, Part C, 13(2): 653-664 (2025)

Gazi Üniversitesi
Fen Bilimleri Dergisi
PART C: TASARIM VE TEKNOLOJİ

Gazi University
Journal of Science
PART C: DESIGN AND TECHNOLOGY

http://dergipark.gov.tr/gujsc

Dynamic Voting-Based Ensemble Deep Learning for Closely Resembling Crop Classification

Engin EŞME[1*] Muhammed Arif ŞEN[1] Halil ÇİMEN[1]

[1]*Konya Technical University, Artificial Intelligence Application and Research Center, Konya, Turkey*

Article Info

Research article
Received: 04/02/2025
Revision: 11/03/2025
Accepted: 17/06/2025

Keywords

Deep Learning
Dynamic Voting
Crop Classification
Smart Agriculture Technologies

Abstract

Deep learning-based approaches in developing autonomous machines that can perform smart agricultural applications such as spraying, irrigation, and harvesting per product exhibit successful applications, especially in tasks such as image classification and data analysis. Accurate identification of plant species and differentiation from weeds are crucial for increasing efficiency and ensuring sustainability in agricultural production. Classifying agricultural crops with similar appearances remains a challenging problem with existing methods. This research represents a significant step toward the automatic recognition and classification of closely resembling agricultural crops while providing a theoretical and practical foundation for the development of smart farming technologies. This study focuses on classifying agricultural crop images, which often bear close resemblance, employing 17 distinct deep learning models and dynamic voting. The dataset consists of 804 images representing five closely resembling crop types: jute, maize, rice, sugarcane, and wheat. To ensure robustness, 10-fold cross-validation is employed, and experiments are conducted consistently across all models using the same sample sets. The results report the models' classification performance on closely resembling agricultural crop images in terms of accuracy, training time, and disk space. According to the experimental findings, ShuffleNet achieved the highest individual accuracy of 98.63% on the test set, but the ensemble approach increased this value to 99.75%. The proposed ensemble approach improves accuracy and ensures greater robustness and stability. Further, the theoretical knowledge and results obtained can be integrated into smart agricultural machines to be developed in this field and will enable them to operate more efficiently.

Dinamik Oylama Tabanlı Topluluk Derin Öğrenme ile Benzer Mahsullerin Sınıflandırılması

Makale Bilgisi

Araştırma makalesi
Başvuru: 04/02/2025
Düzeltme: 11/03/2025
Kabul: 17/06/2025

Anahtar Kelimeler

Derin Öğrenme
Dinamik Oylama
Mahsul Sınıflandırma
Akıllı Tarım Teknolojileri

Öz

Ürün başına ilaçlama, sulama ve hasat gibi akıllı tarım uygulamalarını gerçekleştirebilen otonom makinelerin geliştirilmesinde derin öğrenme tabanlı yaklaşımlar, özellikle görüntü sınıflandırma ve veri analizi gibi görevlerde başarılı uygulamalar sergilemektedir. Tarımsal üretimde verimliliği artırmak ve sürdürülebilirliği sağlamak için bitki türlerinin doğru tanınması ve yabancı otlardan ayırt edilmesi kritik bir öneme sahiptir. Birbirine benzer görünüme sahip tarımsal mahsullerin sınıflandırılması, mevcut yöntemlerle zorlu bir problem olmaya devam etmektedir. Bu araştırma, benzer görünümlü tarımsal ürünlerin otomatik olarak tanınması ve sınıflandırılması konusunda önemli bir adım olmakla birlikte, akıllı tarım teknolojilerinin geliştirilmesine yönelik teorik ve pratik bir temel oluşturmaktadır. Bu çalışma, genellikle birbirine çok benzeyen tarımsal mahsul görüntülerini sınıflandırmayı amaçlayarak 17 farklı derin öğrenme modeli ve dinamik oylama yöntemini kullanmaktadır. Veri seti, kenevir, mısır, pirinç, şeker kamışı ve buğday olmak üzere beş benzer görünümlü mahsul türüne ait toplam 804 görüntüden oluşmaktadır. Güvenilirliği sağlamak amacıyla 10 katlı çapraz doğrulama kullanılmış ve tüm modellerde aynı örnek setleriyle tutarlı deneyler gerçekleştirilmiştir. Elde edilen sonuçlar modellerin birbirine çok benzer tarımsal mahsül görüntülerini sınıflandırma performanslarını doğruluk, eğitim süresi ve disk alanı açısından rapor etmektedir. Deneysel bulgulara göre, ShuffleNet test setinde %98,63 ile en yüksek bireysel doğruluğa ulaşmış, ancak topluluk yaklaşımı bu değeri %99,75'e yükseltmiştir. Önerilen topluluk yaklaşımı doğruluğu artırmakla kalmayıp daha fazla sağlamlık ve kararlılık sağlamaktadır. Ayrıca elde edilen teorik bilgi ve sonuçlar, bu alanda geliştirilecek akıllı tarım makinelerine entegre edilebilecek ve daha verimli şekilde çalışmasını sağlayacaktır.

*Corresponding author, e-mail: eesme@ktun.edu.tr
DOI: 10.29109/gujsc.1632938

1. INTRODUCTION (GİRİŞ)

It is known that agriculture began 13.000 years ago in the northern region of present-day Iraq (between the Tigris and Euphrates rivers). Agriculture, which in the early periods was limited to the cultivation and collection of spontaneously growing wheat and wild plants with primitive methods, has now become a strategic issue of great importance [1]. While the number of people to feed was quite limited in the early periods, today agriculture plays a critical role in feeding more than seven billion people worldwide. Moreover, considering the United Nations (UN) estimate that the world population will be 10 billion in 2050, it can be said that agricultural technological developments and digital transformation efforts are of more strategic and vital importance [2].

Efficiency and quality of agricultural products, which are the main food source of humanity, are decreasing due to many problems such as urban expansion, lack of irrigation, decrease in agricultural lands, desertification, and agricultural diseases. Crop disease detection is crucial for global food security, requiring accurate and timely identification for effective intervention. A recent study reviewed various machine learning and deep learning models, highlighting the challenges of dataset imbalances and emphasizing the effectiveness of Vision Transformers and hybrid approaches in improving classification accuracy [3]. In addition, the COVID-19 epidemic and international political problems that we have experienced in recent years negatively affect the sustainable agricultural economy and pose a serious threat to food security and supply [4]. Overcoming all these problems in agriculture depends on improving agricultural equipment technology, increasing digital conversions in agricultural areas, and applying today's modern techniques such as artificial intelligence and machine learning more in agricultural areas. Thus, it will be possible to achieve more efficient and reliable production and increase food supply by using fewer human resources and more technological equipment.

Findings of recent studies indicate that deep learning methods provide high accuracy and more performance than traditional image-processing methods. Image data collected from agricultural areas usually allows the definition of the agricultural environments and variety of crops. For this reason, imaging collection and analysis with intelligent learning techniques are important to the classification of crops or the detection of anomalies in agricultural areas [5]. Image analysis in the agricultural domain includes three main issues such as image recognition [6], image classification [7], and anomaly detection [8]. The most commonly used sensing methods are satellite-based and used for multi and hyperspectral imaging. Using synthetic aperture radar systems and thermal/near-infrared cameras has become increasing extent, also X-ray and optical techniques are being implemented [9]. In image analyses, some of the most popular methods include machine learning techniques such as K-means and Support Vector Machines, Artificial Neural Network approaches, linear polarization, and regression analysis [10]. Accurate crop mapping is crucial for agricultural production and food security, especially amid climate change and population growth. A study utilizing an Attention-based Bidirectional GRU (A-BiGRU) model on Sentinel-2 time-series images demonstrated superior performance over traditional classifiers, achieving an overall accuracy of 98.04% in identifying rice, maize, and soybean [11]. Pratama et al. (2024) proposed an integrated voting classifier for multiple classification of dry bean varieties with a machine learning approach based on Gaussian Naive Bayes, Decision Tree and Logistic Regression. According to the results, they stated that the accuracy rates of the models varied significantly according to different data subsets, which showed the performance of the classifier under certain conditions and also highlighted areas for improvement [12].

Generally, deep learning models have significantly enhanced predictive accuracy in machine learning applications across a broad spectrum of areas. In the field of machine learning, ensemble learning in deep learning has demonstrated superior performance compared to traditional algorithms. Despite the ability of various deep learning algorithms to automatically extract features and address complex challenges, the primary difficulty lies in the necessity for substantial expertise and experience to fine-tune optimal parameters, making the process time-consuming. For this reason, recent research efforts have sought to merge ensemble learning with deep learning to mitigate this challenge. Ensemble learning encompasses techniques that combine multiple base models within a unified framework to create a more robust model. The effectiveness of an ensemble method depends on various factors, including the training of the baseline models and the manner in which they are integrated [13–15].

Deep neural networks, as a flexible training method for learning, can represent more complex nonlinear structures. However, this flexibility often results in higher variance in deep models. To mitigate the high variance in deep models, ensemble deep learning approaches can be employed. These approaches involve training multiple deep models on the same problem and aggregating their predictions. The primary goal of ensemble techniques is to enhance predictive performance by effectively combining the strengths of various deep learning models [16], [17].

Recent literature commonly focuses on the application of majority fusion methods, particularly voting algorithms, to improve prediction performance in classification or regression problems involving ensemble deep models. This is due to their straightforward and intuitive nature. The most popular voting methods include Max Voting [18], Averaging Voting [19], Weighted Average Voting [20], and hybrid approaches [21]. Each of these methods has its own advantages and disadvantages, which must be carefully considered during implementation.

The classification of crops is very important to increase agricultural productivity. Distinguishing crops from harmful weeds in agricultural activities such as spraying and irrigation ensures better growth and efficiency of the product. For this reason, studies on the classification of agricultural products using deep learning-based image classification and ensemble methods are a current and popular topic in the literature due to their ability to automatically learn data-dependent features. Ayan et al. (2020) adapted and re-trained seven distinct pre-trained convolutional neural network (CNN) models using suitable transfer learning and fine-tuning methods on a publicly available dataset that includes images of insects harmful to crops. Subsequently, the three highest-performing CNN models (InceptionV3, Xception, and MobileNet) were ensembled using a strategy based on the sum of maximum probabilities to enhance classification performance. They stated that the weights of CNN models were adjusted using a genetic algorithm, and the proposed model achieved the highest classification accuracy [22]. Chen et al. (2024) proposed two weight-based ensemble deep learning methods constructed from vector- and matrix-based weights for the detection and classification of crop pests. To address the challenge of weight design, which is critical for the effectiveness of ensemble methods, they formulated the weight design problem as a quadratic convex optimization problem. The solution to this problem has a closed-form expression and can be computed efficiently. They demonstrated that the proposed approach is competitive with other leading methods, achieving high accuracy [23]. To tackle the challenges of early detection and effective crop management, Shahid et al. (2024) proposed a framework for classifying healthy and unhealthy cotton plants. The framework leverages advanced techniques, particularly deep learning, computer vision, and artificial intelligence. They employed feature extraction techniques, including continuous wavelet transform (CWT) and fast Fourier transform (FFT), in their strategy, which utilizes an averaging method to combine the classification scores [24]. Hyder and Talpur (2024) investigated the use of CNNs for the early detection of cotton leaf diseases. The study classified bacterial blight, curl virus, Fusarium wilt, and healthy leaves. It highlighted that CNNs and image processing techniques are effective in diagnosing diseases. The proposed approach simplifies the detection of cotton leaf diseases, contributing to the preservation of crop productivity [25]. Nanni et al. (2020) proposed an automatic classifier for the detection of pests for crop protection by integrating CNN with saliency methods. They developed three different saliency methods for image preprocessing. They stated that they obtained high accuracy results by testing their proposed method on both large and small datasets [26]. Leaf classification is a challenging task, particularly when distinguishing between crop plants of similar size. A study using Deep Learning models achieved a maximum test accuracy of 94.3% on augmented data from a dataset consisting of 570 high-resolution images of agricultural plant leaves organized into 21 categories [27].

This study provides a comprehensive evaluation by comparing the performance of 17 different deep learning models for the classification of closely resembling agricultural crop images. Since an ensemble learning approach is adopted, model diversity plays a crucial role in this study. Utilizing 17 models with different architectures ensures diversity in ensemble learning based on dynamic voting, allowing the strengths of each model to be leveraged while compensating for their weaknesses. The obtained results thoroughly report the models' classification performance for closely resembling agricultural crops in terms of accuracy, training time, and disk space. This diversity not only enhances overall performance but also provides a more robust and stable classification system. In addition to being a significant step toward the automatic recognition and classification of closely resembling agricultural crops, this research establishes a theoretical and practical foundation for the development of smart farming technologies. Initially, the paper provides an overview of the dataset, the deep learning methodologies utilized, and the proposed dynamic voting approach.

Subsequently, it details the training and testing processes, all of which were conducted using a dataset comprising 804 images representing five distinct crop types: jute, maize, rice, sugarcane, and wheat.

2. MATERIALS AND METHODS (MATERYAL VE METOD)

In this study, a deep learning-based ensemble model is proposed for classifying images of agricultural products. In the first stage, 17 different deep learning models are trained on a dataset of agricultural product images consisting of five classes. The architectures and key details of these 17 models are summarized in Table 1. Then, a dynamic voting-based ensemble approach is developed by utilizing the outputs of these models. In this method, only models with high expertise on the relevant class are involved in the decision-making process for classifying each new example, while models with low performance are excluded from the voting. This adaptive decision-making mechanism aims to increase classification accuracy.

2.1. Proposed Ensemble Model (Önerilen Topluluk Modeli)

Dynamic voting is an ensemble learning method and represents a process in which predictions from different models are evaluated. This method enables the identification of the most suitable models for each instance and considers only the predictions of models that perform with higher accuracy on the specific instance. By leveraging the strengths of each model, overall performance is enhanced. Unlike a fixed voting mechanism, dynamic voting takes into account the unique characteristics of each instance and provides an adaptive decision-making process. A diagram illustrating the proposed ensemble approach based on dynamic voting is presented in Figure 1 of this study.

Table 1. Used deep learning models (Kullanılan derin öğrenme modelleri)

AlexNet	AlexNet, a deeper and wider CNN model compared to the traditional LeNet method, was proposed by Alex Krizhevesky et al. in 2012 [28]. AlexNet can provide very successful results in large-scale image recognition compared to common traditional machine learning and computer vision approaches. Thus, it is considered as a significant development that the interest in deep learning-based image recognition is rapidly increasing [29].
DarkNet19 DarkNet53	DarkNet19 is a CNN with 19 layers usually used for object detection. Owing to the pre-trained of the network, it can classify a wide range of images and provide rich feature representations [30]. DarkNet53, a CNN technique, is a basic method used to extract features from images, classify images, and verify them by detecting specific elements. DarkNet53 has a ReLu layer as part of its design in its architecture. Due to its fully connected layers with adjustable number of neurons, it can perform feature synthesis and nonlinear transformations more easily [31].
DenseNet201	Dense Convolutional Network is a CNN architecture proposed by Huang et al. in 2017. It can scale to hundreds of layers by providing direct connections between two layers with the same feature map size. It can achieve the performance of advanced network architectures with fewer parameters and less computation and tend to provide more consistent improvements [32].
EfficientNetB0	EfficientNet model was proposed by Tan and Quoc in 2019 to provide a simple and efficient way to easily scale a basic CNN to any target resource constraint in a more principled manner. EfficientNets are neural architectures that can provide much better accuracy and efficiency with fewer parameters compared to ConvNets [33].
GoogLeNet	A specific example of the Inception architecture, GoogLeNet, is a convolutional neural network proposed by Szegedy et al. in 2015, with a depth of 22 layers. The main advantage of the GoogLeNet architecture is that it can provide a significant increase in quality, despite a reasonable increase in computational requirements, compared to shallower and narrower architectures. It can provide similar quality results in identification and classification with more expensive non-Inception architectures of similar depth and width [34].

Inceptionv3	Inception-v3 is a 48-layer deep convolutional neural network, proposed by Szegedy et al. in 2016. Inception-v3 can provide high-performance image networks at an acceptable computational cost compared to simpler and more monolithic architectures. It allows training relatively smaller training sets with higher performance with lower parameter count and batch-normalized auxiliary classifiers [35].
MobileNetv2	MobileNetV2, an advanced mobile architecture, is a highly usable neural network proposed by Sandler et al. in 2018. It shows high performance especially for standard operations of mobile applications due to its simple and high-throughput structure [36].
NASNet-Large	NASNet architecture was proposed by Zoph et al. in 2018 to design a new search space that can separate the complexity of a neural architecture from the depth of a network and ensure transferability. The proposed method has a highly flexible architecture that can be scaled in terms of computational cost and parameters to easily address many different problems [37].
ResNet18 ResNet50 ResNet101	ResNets (Recently proposed residual networks) is a widely used CNN architecture developed by Kaiming He et al. in 2015, which allows training of high deep networks up to 1000 layers. ResNets architecture is a method that introduces the concept of residual connections, addressing the problem of vanishing gradients in deep networks. ResNets are neural network models that can be designed in different structures according to the number of layers, can be easily implemented without computational burden, and can generalize standard CNNs [38], [39].
ShuffleNet	ShuffleNet, a computationally efficient CNN, was introduced and developed by Zhang et al. in 2017 to greatly reduce the computational cost without decreasing the accuracy performance. Designed specifically for mobile devices with very limited processing power, ShuffleNet architecture can provide approximately 13 times the speedup compared to AlexNet on an ARM-based mobile device with similar accuracy value [40].
SqueezeNet	SqueezeNet, proposed by Iandola et al. in 2016, requires less communication and bandwidth compared to other CNN architectures due to its smaller architecture structure. SqueezeNet, which is suitable for many hardware with limited memory, can achieve the same level of accuracy with 50 times fewer parameters compared to AlexNet [41].
VGG16 VGG19	It was developed by the VGG team to investigate the effect of network depth on accuracy in large-scale image recognition and classification. Compared to a traditional ConvNet architecture, VGG models were able to provide more generalized performance by identifying more complex structures with less deep image representations [42].
Xception	The Xception architecture is a convolutional neural network model proposed by Chollet in 2017 to improve Inception modules. It has been stated that an Xception architecture with a similar number of parameters to Inception V3 performs better on a large image classification dataset and is easier and more efficient to use [43].

In this study, an ensemble is formed using the majority of state-of-the-art deep learning models in the literature. In the first stage, 17 deep learning models are trained on the dataset using a 10-fold cross-validation method. As shown in Figure 1, in the second stage, a meta-dataset is created based on the prediction results of these models on the training set. For each training instance, the models are labeled with 1 or 0 depending on whether their predictions are correct. Thus, a 17-element binary vector is obtained for each instance. These binary vectors are labeled with the class label of the corresponding instances. Consequently, in the meta-dataset, each instance's label consists of a binary vector representing the prediction results of the 17 models. This approach ensures that each instance in the meta-dataset reflects the performance of the models in correctly or incorrectly predicting that specific instance. In the third stage, the constructed meta-dataset is trained using a CNN model. ShuffleNet is chosen for this process due to its superior performance on crop data. ShuffleNet provides a dynamic selection mechanism for each crop instance, determining which deep learning models' predictions should be considered.

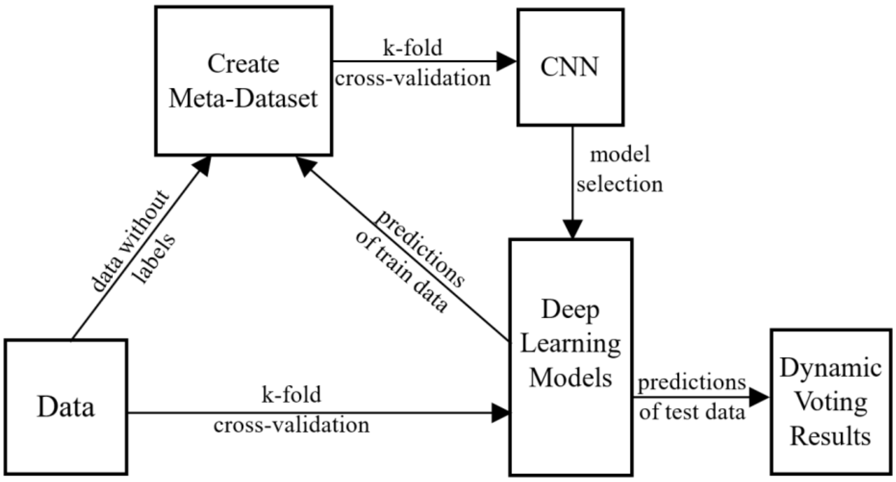

Figure 1. Proposed dynamic voting scheme (Önerilen dinamik oylama şeması)

This method allows the models participating in the voting process to be dynamically selected. Consequently, the predictions of models that are more specialized for a specific instance are prioritized, while the predictions of models that lack sufficient expertise on that instance are disregarded. This approach aims to maximize the contributions of expert models and enhance the accuracy of the predictions resulting from the ensemble voting process.

2.2. Experimental design (Deney tasarımı)

In this paper, a dataset of five different types of crops - namely jute, maize, rice, sugarcane, and wheat - is used for crop classification. The dataset was collected by Jaiswal and is accessible on Kaggle [44]. Images in the jute class generally include thin and long plants with green and small leaves, while images in the maize class typically feature plants with green, long, and wider leaves. Rice class images consist of green and short plants with seeds on them. When sugarcane images are examined, it can be observed that their leaves are green, while their stems are somewhat gray. Unlike the other classes, images in the wheat class mostly depict yellow plants. An example from each class is illustrated in Figure 2.

As seen in Figure 2, all images are different from each other. However, since there are sample images that are similar to each other in all classes, the classification process can be done with artificial intelligence to have a higher classification accuracy. Sample images that may cause errors during classification are presented in Figure 3.

Each class in the dataset consists of 40 images, and each image is 224*224*3 in size, has a resolution of 96 dpi and a depth of 24 bits. The dataset has also augmented images. The augmentation process is conducted by horizontally flipping, shifting, rotating, and vertically shifting the raw images. For each class, about 120 augmented images are obtained by using 40 raw images in a single class and the dataset contains 804 crop images in total

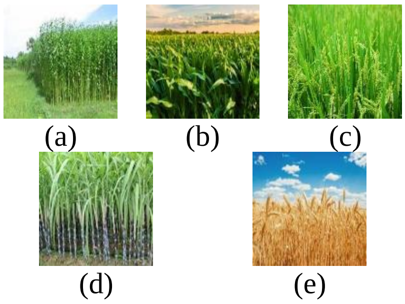

Figure 2. Sample images of each class: a) jute b) maize c) rice d) sugarcane e) wheat (Her sınıftan örnek görüntüler: a) jüt b) mısır c) pirinç d) şeker kamışı e) buğday)

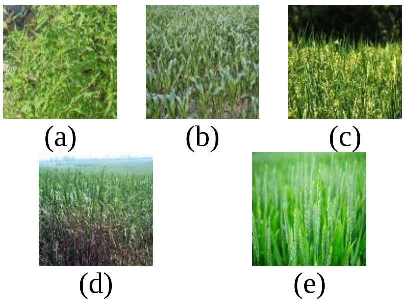

Figure 3. *Similar images from each class: a) jute b) maize c) rice d) sugarcane e) wheat* (Her sınıftan benzer görüntüler: a) jüt b) mısır c) pirinç d) şeker kamışı e) buğday)

3. RESULTS (BULGULAR)

17 deep learning models are trained on the dataset. The models are used in their original forms as introduced in the literature, with images resized to match the input dimensions and the final layers modified to classify into five categories. Transfer learning is not employed; the weights are initialized randomly, and the training process is conducted entirely on the agricultural dataset. Models are set parameters with MiniBatchSize 32, Max epoch 100,

Learning rate 0.0001. During training a randomly selected validation set, which is 30% of the training set is used. The accuracy of the models is evaluated by using 10-fold cross-validation. In addition, a fair classification environment is provided by fixing the training and test samples in all models. The experiments are conducted on a computer equipped with an AMD Ryzen 9 5950X CPU@3.40 GHz, 64 GB RAM, and an Nvidia GeForce RTX 3080 (12 GB) GPU. The models are designed and trained in the MATLAB environment using the Deep Learning Toolbox. The Parameters of the experiments are presented in Table 2.

Table 2. Parameters of the experiments (Deneylerin parametreleri)

Option	Value
kfold	10
Train/Validation Ratio	0.7 / 0.3
Shuffle	Each Epoch
SolverName	Adaptive moment est.
MiniBatchSize	32
MaxEpochs	100
LearnRate	0.0001

Table 3 presents the validation accuracy results for the models upon completion of training. The results are listed separately for each fold. To enable a general comparison of performance, the table also includes the average across all 10 folds, standard deviation and the corresponding ranking.

To detail the results presented in Table 3, the ShuffleNet model demonstrated the best performance with an average accuracy of 98.20%, ranking first. Notably, it achieved 100% accuracy in certain folds. NASNet-Large ranked second with an average accuracy of 97.83%, consistently delivering high performance across all folds. GoogLeNet secured third place with an accuracy of 97.69%, achieving over 98% accuracy in most folds. Among the other models, DenseNet201, ResNet18, and Xception showed strong performance with accuracy rates exceeding 96%. Models such as AlexNet, DarkNet19, DarkNet53, EfficientNetB0, Inceptionv3, ResNet50, and ResNet101 exhibited moderate performance, with accuracy ranging between 90% and 96%. Meanwhile, MobileNetv2, SqueezeNet, VGG16, and VGG19 were the lowest-performing models.

Overall, even models with relatively lower average performance demonstrated high accuracy in certain folds. This suggests that some models may specialize in specific examples. Therefore, it would be a suitable approach to dynamically select the models participating in majority voting based on the test sample. Additionally, models like MobileNetv2 and SqueezeNet, which rank at the bottom with average accuracies below 80%, could be considered for complete exclusion from majority voting.

Figure 4 shows the time taken to complete training for a fold during the training process, as well as the disk space used by the model file generated for that fold. NASNet-Large has the longest training time (445 minutes), while VGG19 requires the most disk space. In contrast, lightweight models like AlexNet and SqueezeNet have minimal training times (4-6 minutes) and low disk usage, making them resource-efficient. This comparison in Figure 4 highlights trade-offs between resource demands and practical model selection based on computational constraints.

Table 3. The validation accuracy results of deep learning models (Modellerinin doğrulama başarısı)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	Std.	Avg.	Ranks
AlexNet	97.24	97.24	93.09	94.47	94.04	97.22	98.14	94.93	94.50	94.44	1.66	95.53	10
DarkNet19	99.08	94.47	97.70	97.24	98.17	**99.07**	98.14	93.09	91.74	98.61	2.51	96.73	7
DarkNet53	97.24	95.85	94.47	93.55	96.79	98.15	93.95	**99.54**	96.33	92.13	2.16	95.80	9
DenseNet201	97.24	97.70	**99.54**	94.47	97.25	96.30	96.28	96.31	96.79	**99.07**	1.39	97.09	4
EfficientNetB0	94.47	94.47	92.17	89.86	90.37	94.44	87.91	93.55	97.25	92.59	2.60	92.71	13
GoogLeNet	98.62	**99.54**	97.70	98.16	98.17	92.13	98.14	97.70	98.62	98.15	1.92	97.69	3
Inceptionv3	96.31	94.93	99.08	96.31	97.71	97.22	97.21	94.93	97.71	93.52	1.56	96.49	8
MobileNetv2	78.34	81.11	84.33	78.34	82.11	79.63	75.81	83.41	77.98	78.70	2.55	79.98	16
NASNet-Large	**100.00**	97.24	98.62	97.70	96.33	97.69	97.67	96.31	98.17	98.61	**1.05**	97.83	2
ResNet18	94.47	91.71	97.24	97.24	98.62	97.69	**99.07**	99.08	97.25	97.22	2.15	96.96	5
ResNet50	96.31	99.08	97.24	95.39	93.12	96.30	92.09	97.24	94.50	93.98	2.02	95.52	11
ResNet101	88.94	92.63	96.31	94.01	96.33	91.20	95.81	94.47	93.58	92.13	2.26	93.54	12
ShuffleNet	98.62	97.24	**99.54**	**100.00**	**100.00**	96.76	97.21	98.62	97.71	96.30	1.28	**98.20**	1
SqueezeNet	80.18	58.99	81.11	77.88	77.52	81.48	75.35	82.95	80.28	75.00	6.52	77.07	17
VGG16	86.64	93.55	77.42	88.94	92.66	90.28	90.70	82.95	87.61	93.06	4.79	88.38	14
VGG19	88.48	82.95	88.48	85.25	74.77	91.67	84.65	87.56	88.53	87.04	4.38	85.94	15
Xception	94.47	96.31	97.24	96.77	99.08	96.30	98.14	96.77	**99.54**	94.91	1.55	96.95	6

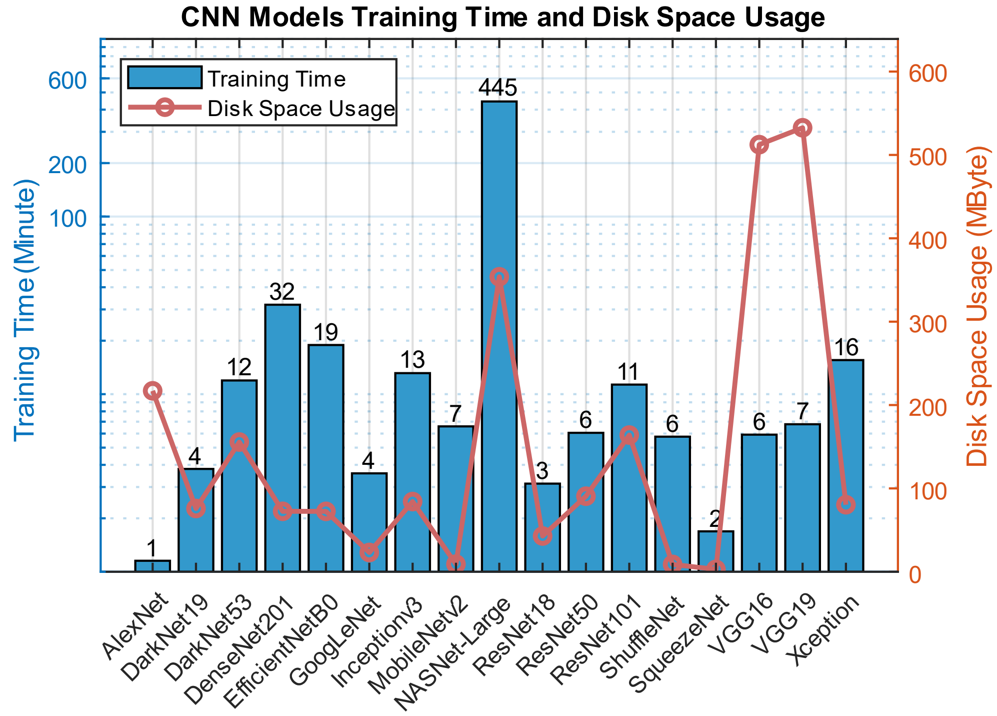

Figure 4. Training Time and Disk Space Utilization of Models (Modellerin Eğitim Süresi ve Disk Alanı Kullanımı)

Table 4 presents the accuracy rates of the models on the test set for each fold, along with their average accuracy and rankings across all folds.

To detail the results presented in Table 4 the ShuffleNet model achieved the second-highest performance on the test set, following the ensemble model, with an average accuracy of 98.63%, consistent with its performance on the validation set. Similarly, models like NASNet-Large, GoogLeNet, and Xception demonstrated stable performance on the test set, aligning with their validation set results, highlighting their robustness. In contrast, AlexNet and DarkNet53, which ranked mid-tier on the validation set, showed improved performance on the test set, indicating better generalization capabilities compared to other models. On the other hand, DenseNet201 and ResNet18, despite their high rankings on the validation set, dropped by five positions on the test set. This decline suggests weaker generalization ability and a tendency to overfit the training data. Models such as DarkNet19, EfficientNetB0, Inceptionv3, MobileNetv2, ResNet50, ResNet101, SqueezeNet, VGG16, and VGG19 exhibited greater variability in performance across folds, as indicated by their higher standard deviations.

Table 4. The test accuracy results of deep learning models (Modellerinin test başarısı)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	Std.	Avg.	Ranks
AlexNet	97.50	93.83	92.59	98.77	93.83	97.50	95.00	98.75	93.75	95.00	2.16	95.65	6
DarkNet19	98.75	90.12	95.06	97.53	**100.00**	95.00	**100.00**	85.00	87.50	90.00	5.13	93.90	9
DarkNet53	93.75	93.83	93.83	95.06	91.36	95.00	91.25	96.25	97.50	96.25	1.94	94.41	7
DenseNet201	88.75	91.36	93.83	90.12	96.30	92.50	95.00	91.25	95.00	96.25	2.50	93.04	10
EfficientNetB0	76.25	83.95	82.72	83.95	88.89	88.75	83.75	77.50	80.00	82.50	3.94	82.83	16
GoogLeNet	98.75	96.30	96.30	**100.00**	98.77	91.25	93.75	96.25	**100.00**	**100.00**	2.80	97.14	4
Inceptionv3	96.25	96.30	97.53	96.30	98.77	93.75	92.50	88.75	91.25	91.25	3.08	94.26	8
MobileNetv2	70.00	75.31	82.72	71.60	82.72	80.00	65.00	68.75	70.00	56.25	7.87	72.23	18
NASNet-Large	**100.00**	97.53	98.77	98.77	91.36	98.75	97.50	**100.00**	**100.00**	98.75	2.42	98.14	3
ResNet18	93.75	85.19	95.06	91.36	92.59	92.50	96.25	92.50	93.75	96.25	3.01	92.92	11
ResNet50	87.50	97.53	95.06	93.83	85.19	92.50	86.25	92.50	83.75	93.75	4.47	90.79	12
ResNet101	77.50	86.42	95.06	91.36	88.89	92.50	93.75	88.75	93.75	90.00	4.83	89.80	13
ShuffleNet	98.75	95.06	98.77	**100.00**	**100.00**	98.75	**100.00**	97.50	98.75	98.75	1.40	98.63	2
SqueezeNet	80.00	49.38	72.84	70.37	69.14	80.00	71.25	81.25	78.75	78.75	9.03	73.17	17
VGG16	88.75	85.19	81.48	90.12	88.89	86.25	93.75	82.50	83.75	92.50	3.95	87.32	14
VGG19	90.00	85.19	88.89	85.19	81.48	85.00	80.00	86.25	88.75	91.25	3.43	86.20	15
Xception	96.25	93.83	97.53	93.83	97.53	98.75	93.75	95.00	97.50	97.50	1.79	96.15	5
Ensemble	**100.00**	**98.77**	**100.00**	**100.00**	**100.00**	**100.00**	98.75	**100.00**	**100.00**	**100.00**	**0.50**	**99.75**	1

The ensemble model, with an average accuracy of 99.75%, outperformed all other models and secured the top position. Its low standard deviation indicates minimal performance differences across folds, demonstrating that it is a robust and reliable ensemble model compared to individual models.

Table 5 presents the 10-fold averages of all models and provides a detailed listing of their performance on the test set. The ensemble model achieved the highest classification accuracy, outperforming all other models with an accuracy of 99.75%. While models like SqueezeNet and MobileNetv2 showed very low accuracy in some samples, and models like DarkNet19 and EfficientNetB0 demonstrated relatively low accuracy in certain folds, the ensemble model compensated for these inconsistencies through dynamic majority voting. This approach allowed the ensemble model to achieve 100% sensitivity, ensuring no positive samples were missed, while also minimizing the false positive rate. The high F-Measure value indicates that both the ensemble and ShuffleNet models maintain a strong balance between sensitivity and precision.

Table 5. Average test performances of the models (Modellerin ortalama test performansları)

Model	Accuracy	Sensitivity	Specificity	Precision	F-Measure	G-mean
AlexNet	0.9565	0.9867	0.9488	0.8319	0.8999	0.9672
DarkNet19	0.9390	0.9214	0.9430	0.8064	0.8567	0.9317
DarkNet53	0.9441	0.9354	0.9473	0.8248	0.8707	0.9403
DenseNet201	0.9304	0.9298	0.9299	0.7711	0.8405	0.9290
EfficientNetB0	0.8283	0.8806	0.8165	0.5412	0.6681	0.8460
GoogLeNet	0.9714	1.0000	0.9645	0.8836	0.9344	0.9819
Inceptionv3	0.9426	0.9588	0.9398	0.7986	0.8682	0.9488
MobileNetv2	0.7223	0.8026	0.7012	0.4084	0.5377	0.7473
NASNet-Large	0.9814	0.9889	0.9798	0.9274	0.9564	0.9843
ResNet18	0.9292	0.9316	0.9273	0.7591	0.8342	0.9281
ResNet50	0.9079	0.9261	0.9054	0.7279	0.8058	0.9141
ResNet101	0.8980	0.9390	0.8871	0.6886	0.7911	0.9120
ShuffleNet	0.9863	0.9868	0.9862	0.9453	0.9649	0.9864
SqueezeNet	0.7317	0.7871	0.7193	0.4232	0.5456	0.7517
VGG16	0.8732	0.9180	0.8619	0.6312	0.7431	0.8884
VGG19	0.8620	0.8477	0.8671	0.6150	0.7072	0.8556
Xception	0.9615	0.9594	0.9612	0.8603	0.9055	0.9596
Ensemble	**0.9975**	**1.0000**	**0.9970**	**0.9866**	**0.9931**	**0.9985**

In conclusion, the ensemble model demonstrated the best performance across all metrics. This makes the ensemble model an ideal choice for applications requiring minimal false positives, no missed positive instances, and maximum accuracy.

4. DISCUSSION (Tartışma)

This paper investigates the task of classifying closely resembling agricultural crop images using deep learning models. The paper presents 17 deep learning model performances on the task of classifying and an ensemble learning approach based on dynamic voting. Dynamic voting doesn't use the same voting models on each test sample. For each test sample, the number of voting models and voters varies. The results demonstrate the robustness and effectiveness of this approach across all evaluated metrics, highlighting its potential as a reliable solution for classification problems.

The ensemble model achieved the highest average accuracy (99.75%), surpassing all individual models, including ShuffleNet. This success stems from its dynamic voting mechanism, which compensates for inconsistent performances across samples by prioritizing reliable models. For instance, it mitigates the weaknesses of models like SqueezeNet and MobileNetv2, which showed low accuracy in some folds, ensuring high sensitivity (100%) and a balanced F-Measure. Additionally, the improved generalization of AlexNet and DarkNet53 on the test set emphasizes the importance of evaluating models on diverse datasets. In contrast, the ensemble model addresses the overfitting issues observed in DenseNet201 and ResNet18, ensuring greater robustness and stability.

Overall, ensemble learning based on dynamic voting provides a promising framework for improving the performance of classifying closely resembling agricultural crop images and addressing the challenges of variability and overfitting in deep learning.

5. CONCLUSIONS (SONUÇLAR)

This study offers a comprehensive assessment by evaluating the performance of 17 deep learning models in classifying closely resembling agricultural crops. Given the adoption of an ensemble learning approach, model diversity is a key focus. The use of 17 models with distinct architectures enhances diversity, enabling the strengths of each model to be utilized while mitigating their weaknesses. The results provide a detailed analysis of the models' classification performance in terms of accuracy, training time, and disk space. This diversity not only improves overall performance but also contributes to a more robust and reliable classification system. Beyond advancing the automatic recognition and classification of closely resembling agricultural crops, this research lays a theoretical and practical foundation for the development of smart farming technologies. The study employs a 10-fold cross-validation approach and consistently utilizes the same samples across all models in the experiments. The results obtained substantiate the capability of the models to accurately detect and categorize agricultural crop images. Notably, ShuffleNet, NASNet-Large, GoogLeNet and DenseNet emerged as the top performers with an impressive between 97.09% and 98.20% accuracy on the test set, as indicated by the experimental outcomes. On the other hand, overfitting problems have been also observed in models such as MobileNet and SqueezeNet. Therefore, instead of traditional majority voting, dynamic voting provides more robustness and stability by addressing the observed overfitting issues. As a result, the dynamic voting approach has improved the accuracy in the problem of classifying very similar agricultural crops.

While the proposed dynamic voting mechanism proves to be a significant advancement, it comes with certain limitations. The computational complexity of dynamically selecting models for each sample may pose challenges for large-scale datasets or real-time applications. Future research focuses on designing and developing a smart vehicle which works autonomously and is able to collect data from the field to optimize this process and more real-time testing of the proposed approach.

ACKNOWLEDGMENTS (TEŞEKKÜR)

This study was conducted at the Artificial Intelligence Application and Research Center of Konya Technical University.

DECLARATION OF ETHICAL STANDARDS (ETİK STANDARTLARIN BEYANI)

The authors of this article declare that the materials and methods they use in their work do not require ethical committee approval and/or legal-specific permission.

AUTHORS' CONTRIBUTIONS (YAZARLARIN KATKILARI)

Engin EŞME: Conceptualisation, Methodology, Software, Formal Analysis, Investigation, Data Processing, Writing.
Muhammed Arif ŞEN: Conceptualisation, Literature Search, Validation, Investigation, Resources, Writing.
Halil ÇİMEN: Conceptualisation, Literature Search, Validation, Investigation, Resources, Writing.

CONFLICT OF INTEREST (ÇIKAR ÇATIŞMASI)

There is no conflict of interest in this study.

REFERENCES (KAYNAKLAR)

[1] Gray H, Nuri KR. Differing Visions of Agriculture: Industrial-Chemical vs. Small Farm and Urban Organic Production. American Journal of Economics and Sociology. 2020; 79(3): 813-832. doi:10.1111/ajes.12344
[2] Bouguettaya A, Zarzour H, Kechida A, Taberkit AM. Deep learning techniques to classify agricultural crops through UAV imagery: A review. Neural computing and applications. 2022;34(12):9511-9536.
[3] Ngugi HN, Akinyelu AA, Ezugwu AE. Machine Learning and Deep Learning for Crop Disease Diagnosis: Performance Analysis and Review. Agronomy. 2024;14(12):3001.
[4] Jámbor A, Czine P, Balogh P. The impact of the coronavirus on agriculture: first evidence based on global newspapers. Sustainability. 2020;12(11):4535.
[5] Kamilaris A, Prenafeta-Boldú FX. Deep learning in agriculture: A survey. Comput Electron Agric. 2018;147:70-90.
[6] Abdullahi HS, Sheriff R, Mahieddine F. Convolution neural network in precision agriculture for plant image recognition and classification. In: 2017 Seventh International Conference on Innovative Computing Technology (INTECH). Vol 10 ; 2017:256-272.
[7] Yu S, Jia S, Xu C. Convolutional neural networks for hyperspectral image classification. Neurocomputing. 2017;219:88-98.

[8] Mohanty SP, Hughes DP, Salathé M. Using deep learning for image-based plant disease detection. Frontiers in plant science. 2016;7:1419.

[9] Ishimwe R, Abutaleb K, Ahmed F, others. Applications of thermal imaging in agriculture—A review. Advances in remote Sensing. 2014;3(03):128.

[10] Saxena L, Armstrong L. A survey of image processing techniques for agriculture. Published online 2014.

[11] Feng F, Gao M, Liu R, Yao S, Yang G. A deep learning framework for crop mapping with reconstructed Sentinel-2 time series images. Comput Electron Agric. 2023;213:108227.

[12] Pratama IPA, Atmadji ESJ, Purnamasar DA, Faizal E. Evaluating the performance of voting classifier in multiclass classification of dry bean varieties. Indonesian Journal of Data and Science. 2024;5(1):23-29.

[13] Schmidhuber J. Deep learning in neural networks: An overview. Neural networks. 2015;61:85-117.

[14] Pan SJ, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data Eng. 2009;22(10):1345-1359.

[15] Mohammed A, Kora R. A comprehensive review on ensemble deep learning: Opportunities and challenges. Journal of King Saud University-Computer and Information Sciences. 2023;35(2):757-774.

[16] Amara J, Bouaziz B, Algergawy A. A deep learning-based approach for banana leaf diseases classification. Published online 2017.

[17] Mohammed A, Kora R. An effective ensemble deep learning framework for text classification. Journal of King Saud University-Computer and Information Sciences. 2022;34(10):8825-8837.

[18] Kim HC, Pang S, Je HM, Kim D, Bang SY. Constructing support vector machine ensemble. Pattern Recognit. 2003;36(12):2757-2767.

[19] Montgomery JM, Hollenbach FM, Ward MD. Improving predictions using ensemble Bayesian model averaging. Political Analysis. 2012;20(3):271-291.

[20] Latif-Shabgahi GR. A novel algorithm for weighted average voting used in fault tolerant computing systems. Microprocess Microsyst. 2004;28(7):357-361.

[21] Atay Y, Yildirim MO, Dogan CU. High performance classification of cancer types with gene microarray datasets: hybrid approach. Gazi University Journal of Science Part C: Design and Technology. 2021;9(4):811-827.

[22] Ayan E, Erbay H, Varç\in F. Crop pest classification with a genetic algorithm-based weighted ensemble of deep convolutional neural networks. Comput Electron Agric. 2020;179:105809.

[23] Chen M, Wang J, Chen Y, Guo M, Zheng N. Weight-based ensemble method for crop pest identification. Ecol Inform. Published online 2024:102693.

[24] Shahid MF, Khanzada TJS, Aslam MA, Hussain S, Baowidan SA, Ashari RB. An ensemble deep learning models approach using image analysis for cotton crop classification in AI-enabled smart agriculture. Plant Methods. 2024;20(1):104.

[25] Hyder U, Talpur MRH. Detection of cotton leaf disease with machine learning model. Turkish Journal of Engineering. 2024;8(2):380-393. doi:10.31127/tuje.1406755

[26] Nanni L, Maguolo G, Pancino F. Insect pest image detection and recognition based on bio-inspired methods. Ecol Inform. 2020;57:101089.

[27] Varma WK, Kumar V. Analysis of Crop Leaf Image Classification using Deep Learning Models over Novel Dataset. In: 2023 10th International Conference on Signal Processing and Integrated Networks (SPIN). 2023:14-19.

[28] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst. 2012;25.

[29] Alom MZ, Taha TM, Yakopcic C, et al. The history began from alexnet: A comprehensive survey on deep learning approaches. arXiv preprint arXiv:180301164. Published online 2018.

[30] Al-Haija QA, Smadi M, Al-Bataineh OM. Identifying phasic dopamine releases using darknet-19 convolutional neural network. In: 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). 2021:1-5.

[31] Rajkumar R, Gopalakrishnan S, Praveena K, Venkatesan M, Ramamoorthy K, Hephzipah JJ. Darknet-53 convolutional neural network-based image processing for breast cancer detection. Mesopotamian Journal of Artificial Intelligence in Healthcare. 2024;2024:59-68.

[32] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017:4700-4708.

[33] Tan M, Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning. 2019:6105-6114.

[34] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015:1-9.
[35] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:2818-2826.
[36] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:4510-4520.
[37] Zoph B, Vasudevan V, Shlens J, Le Q V. Learning transferable architectures for scalable image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:8697-8710.
[38] Targ S, Almeida D, Lyman K. Resnet in resnet: Generalizing residual architectures. arXiv preprint arXiv:160308029. Published online 2016.
[39] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:770-778.
[40] Zhang X, Zhou X, Lin M, Sun J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:6848-6856.
[41] Iandola FN. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size. arXiv preprint arXiv:160207360. Published online 2016.
[42] Simonyan K. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. Published online 2014.
[43] Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017:1251-1258.
[44] Aman Jaiswal. Agriculture crop images. 2020. https://www.kaggle.com/datasets/aman2000jaiswal/agriculture-crop-images