



Detection and Classification of Plant Leaf Diseases

Mustafa Tortuk ^{1*}

¹ Bilgisayar Mühendisliği Bölümü, İskenderun Teknik Üniversitesi, Hatay, Türkiye.

ÖZET

Plant diseases pose a serious threat to global food production, leading to yield losses, increased production costs, and environmental damage. Plant pests and diseases have widespread negative impacts on economic, ecological, environmental, and human health, and calculating the true cost of these damages is extremely difficult. Traditional diagnostic methods also present significant limitations in terms of time and cost. This study adopts an image processing-based approach to this problem. Using the Crop Disease Detection dataset, plant disease detection was performed using CNN, DNN, K-NN, SVM, XGBoost, and Random Forest algorithms, employing both deep learning and machine learning methods. The study demonstrates that CNN architectures designed from scratch, without resorting to pre-trained models such as ResNet and MobileNet, can also exhibit high performance. The highest accuracy rate was obtained with the CNN model at 94.08%. In machine learning models, grid search was used for hyperparameter optimization, and the best results were achieved through this method.

MAKALE BİLGİSİ

Received 29.04.2026,

Accepted 13.06.2026,

Publication Date 30.06.2026

Keywords:

Object Classification,
Artificial Neural
Networks, Machine
Learning, Image
Processing, Plant Disease
Detection..

Distributed Under CC-BY 4.0



*Corresponding Author: Mustafa TORTUK, E-mail: mustafatortuk1@gmail.com

Bitki Yaprak Hastalıklarının Tespiti ve Sınıflandırılması

Mustafa Tortuk ^{1*}

¹ Bilgisayar Mühendisliği Bölümü, İskenderun Teknik Üniversitesi, Hatay, Türkiye.

ÖZET

Bitki hastalıkları, küresel gıda üretimi için ciddi bir tehdit oluşturmakta; verim kayıplarına, üretim maliyetlerinin artmasına ve çevresel zarara yol açmaktadır. Bitki zararlı ve hastalıklarının ekonomik, ekolojik, çevresel ve insan sağlığı açısından geniş ölçekli olumsuz etkileri bulunmakta olup bu zararların gerçek maliyetini hesaplamak son derece güçtür. Geleneksel teşhis yöntemleri ise zaman ve maliyet açısından ciddi kısıtlar oluşturmaktadır. Bu çalışmada söz konusu soruna görüntü işleme tabanlı bir yaklaşım benimsenmiştir. Crop Disease Detection veri seti kullanılarak; CNN, DNN, K-NN, SVM, XGBoost ve Random Forest algoritmaları ile hem derin öğrenme hem de makine öğrenmesi yöntemleri aracılığıyla bitki hastalıklarının tespiti gerçekleştirilmiştir. Çalışmada, ResNet ve MobileNet gibi önceden eğitilmiş (pretrained) modellere başvurulmaksızın, sıfırdan tasarlanan CNN mimarilerinin de yüksek başarımla sergileyebileceği gösterilmiştir. En yüksek doğruluk oranı, CNN modeli ile %94,08 olarak elde edilmiştir. Makine öğrenmesi modellerinde ise hiperparametre optimizasyonu için ızgara arama (grid search) yöntemi kullanılmış ve en iyi sonuçlara bu yol ile ulaşılmıştır.

MAKALE BİLGİSİ

Keywords:

Nesne Sınıflandırması, Yapay Sinir Ağı, Makine Öğrenmesi, Görüntü İşleme, Bitki Hastalığı Tespiti.

Distributed Under CC-BY 4.0



GİRİŞ

Tarım sektörü, küresel gıda güvenliğinin temel direğini oluştururken bitki hastalıkları üreticiler için her yıl telafi edilmesi güç kayıplara neden olmaktadır. Derin öğrenme ve makine öğrenmesi algoritmalarının gelişmesiyle birlikte bu hastalıkların görüntüler üzerinden otomatik olarak tespit edilmesi artık mümkün hale gelmektedir.

Vallabhajoshiya ve diğerleri (2021), bitki yaprak hastalıklarının tespitinde DenseNet, ResNet ve NasNet Mobile mimarilerini bir araya getiren transfer öğrenmesi tabanlı derin topluluk sinir ağı

yöntemi önermiş; PlantVillage veri seti üzerinde %100 doğruluk oranı elde etmişlerdir (Vallabhajosyula et al., 2021).

Wei ve diğerleri (2022), yaprak hastalıklarının sınıflandırılmasında açıklanabilir derin öğrenme yaklaşımını benimseyerek VGG, GoogLeNet ve ResNet mimarileri üzerinde karşılaştırmalı bir analiz yapmış; ResNet modeliyle PlantVillage veri seti üzerinde %99.89 doğruluk oranı elde etmişlerdir (Wei et al., 2022).

Russel ve Selvaraj (2022), çok ölçekli paralel derin CNN mimarisi kullanarak hem yaprak türü hem de hastalık sınıflandırması gerçekleştirmiş; Law's Mask ve MultiScale teknikleriyle destekledikleri ResNet modeli PlantVillage veri seti üzerinde %98.61 doğruluk oranına ulaşmıştır (Russel & Selvaraj, 2022). Başka bir çalışmada, YOLO tabanlı çok sınıflı bitki meyve sınıflandırması gerçekleştirerek ön eğitilmiş modellerin performansını karşılaştırmaktadır (Chowdhury ve ark., 2026). Gajjar ve diğerleri (2022), gömülü platformlarda gerçek zamanlı bitki yaprak hastalığı tespiti için SSD MobileNet ve CNN mimarilerini kullanmış; kendi derledikleri veri seti üzerinde %96.88 doğruluk oranı elde etmişlerdir (Gajjar et al., 2022). Pandian ve diğerleri (2022), beş evrişim katmanlı derin CNN (DCNN) modeli tasarlayarak Conv-3, Conv-4 ve Conv-5 mimarileri ile hiper parametre optimizasyonunu karşılaştırmış; PlantVillage veri seti üzerinde %98.41 doğruluk oranına ulaşmışlardır (Pandian et al., 2022). Ashwinkumar ve diğerleri (2022), domates yaprak hastalığı veri seti üzerinde Emperor Penguin Optimizer (EPO) algoritmasıyla optimize edilmiş MobileNet tabanlı bir CNN modeli (OMNCNN) önermişler ve %98.77 doğruluk oranı elde etmişlerdir (Ashwinkumar et al., 2022). Harakannanavar ve diğerleri (2022), bilgisayarlı görü ve makine öğrenmesi algoritmalarıyla bitki yaprak hastalığı tespitinde Ayrık Dalgacık Dönüşümü (DWT), Gri Düzey Eş Oluşum Matrisi (GLCM) gibi özellik çıkarma yöntemlerini SVM, K-NN ve CNN sınıflandırıcılarıyla birleştirmiş; en yüksek doğruluk oranına (%99.09) DWT+PCA+GLCM+CNN kombinasyonu ile ulaşmışlardır (Harakannanavar et al., 2022). Kundu ve diğerleri (2022), görüntü işleme yöntemiyle bitki yaprak hastalığı tespitinde AlexNet mimarisini kullanmış; modelin virüs, bakteri ve mantar kaynaklı hastalıkların tespitinde başarılı sonuçlar verdiğini bildirmiş, ancak sayısal doğruluk oranı raporlanmamıştır (Kundu et al., 2022). Andrew ve diğerleri (2022), tarımsal uygulamalar için derin öğrenme tabanlı yaprak hastalığı tespitinde DenseNet-121, ResNet-50, VGG-16 ve Inception V4 modellerini transfer öğrenmesi ve ince ayar (fine-tuning) tekniğiyle PlantVillage veri seti üzerinde değerlendirmiş; DenseNet-121 %99.81 doğruluk oranıyla en iyi performansı sergilemiştir (Andrew et al., 2022). Russel ve Selvaraj (2022) ile Wei ve diğerleri (2022) başta olmak üzere literatürdeki pek çok çalışmanın önceden eğitilmiş (pre-trained) transfer öğrenmesi modelleri üzerine yoğunlaştığı görülmektedir; bu çalışmalarda sıfırdan eğitilen modellere kıyasla daha yüksek doğruluk oranları raporlanmıştır. Yapay zekanın bitkilerdeki hastalıkları tesbitinde bir çok uygulama gerçekleştirilmiştir. Bu uygulamalarda farklı hastalık tesbitinin yanısıra başarımı arttırmak için parametre analizleri de gerçekleştirilmiş be sınıflandırma başarımına etkileri incelenmiştir. (Atasoy ve Kutlu, 2022; Camgözlü ve Kutlu, 2019; Camgözlü ve Kutlu, 2020; Camgözlü ve Kutlu, 2023).

Bu çalışmada, crop_disease_detection_dataset kullanılarak sıfırdan tasarlanan CNN, DNN, Random Forest, SVM ve XGBoost makine öğrenmesi yöntemleriyle bitki hastalıklarının görüntü tabanlı tespiti amaçlanmaktadır. Önceden eğitilmiş modellere başvurulmaksızın geliştirilen bu yaklaşımda, CNN modeli %94.08 doğruluk oranıyla en yüksek performansı sergilemiş; Grid Search ile hiperparametre optimizasyonu uygulandıktan sonra XGBoost modeli %85.09, Random Forest ise %81.52 doğruluğa ulaşmıştır. Çalışmanın temel hedefi, literatürdeki transfer öğrenmesi odaklı yöntemlerle kıyaslanabilir sonuçların özgün bir model mimarisiyle elde edilebileceğini ortaya koymaktır.

MATERYAL VE METHOD

Veri Seti

Bu çalışmada, elma, muz, fasulye, soya fasulyesi, biber, üzüm, mısır, pamuk, patates, pirinç, domates ve buğday bitkilerine ait sağlıklı ve hastalıklı yaprak görüntülerinden oluşan crop_disease_detection_dataset kullanılmıştır. Toplamda 21.000 görüntü içeren veri seti, farklı hastalık türlerini (bakteriyel, kuruma vb.) kapsayan sınıflara ayrılmış olup modelin geniş bir bitki ve hastalık çeşitliliği üzerinde eğitilmesine olanak tanımıştır.

Şekil 1. Örnek Görseller



Veri Hazırlama

Veri kümesi üzerinde CNN ve DNN gibi derin öğrenme algoritmalarında ön işlem adımları yapılmamıştır, ancak makine öğrenmesi adımlarında, ham piksel verisi yerine anlamlı özellik vektörlerinin kullanılması, hem hesaplama maliyetini düşürmekte hem de model performansını artırmaktadır. Bu doğrultuda her görüntü önce 64×64 piksel boyutuna yeniden ölçeklendirilmiş, ardından şekil ve doku bilgisini yakalamak amacıyla gri tonlamalı görüntü üzerinde HOG (Histogram of Oriented Gradients) algoritması uygulanmıştır. Renk tabanlı hastalık belirtilerini temsil etmek için ise görüntüler HSV renk uzayına dönüştürülerek 8×4×4 boyutlu normalize renk histogramları çıkarılmıştır. Elde edilen HOG ve renk histogramı özellikleri birleştirilerek her

görüntü için sabit uzunlukta tek bir özellik vektörü oluşturulmuş ve bu vektörler Random Forest, SVM ile XGBoost algoritmalarına girdi olarak sunulmuştur.

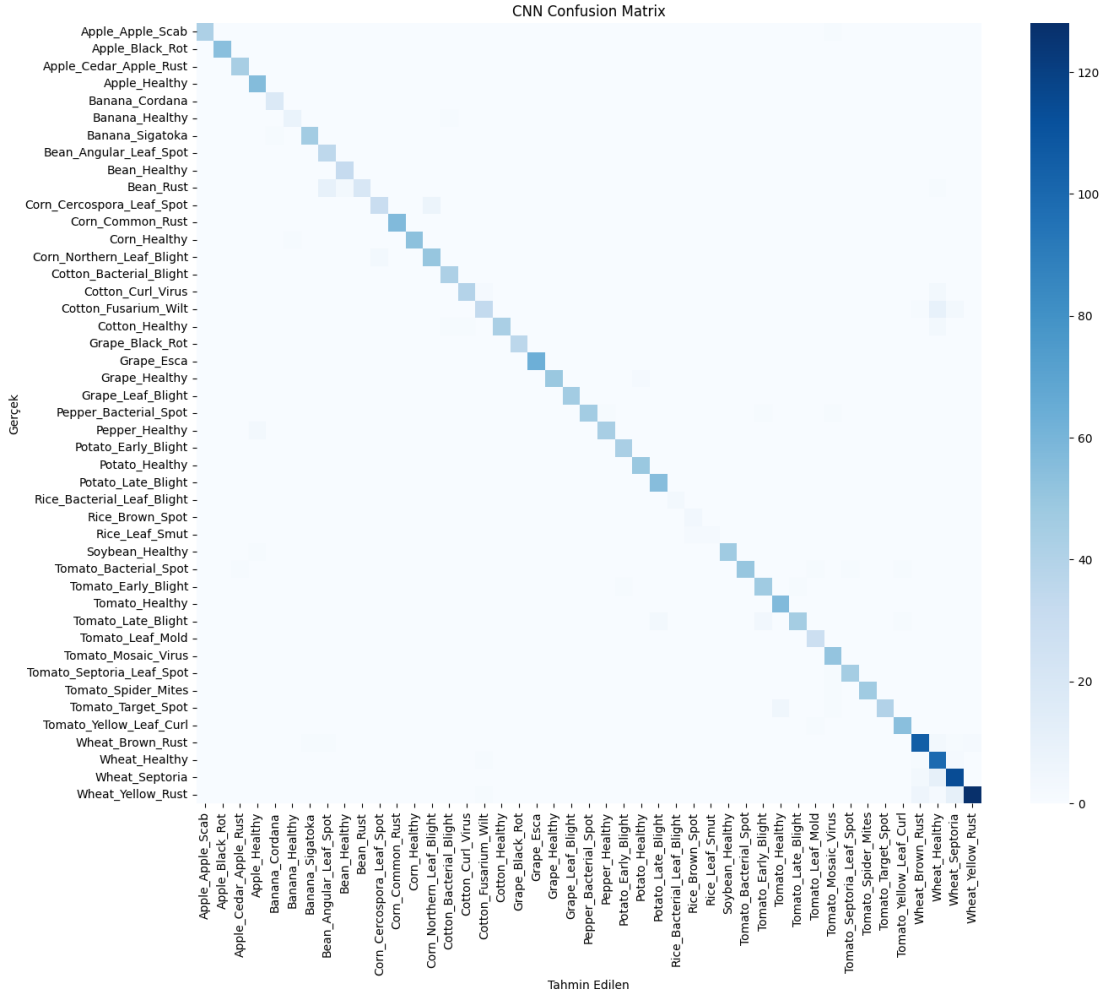
Modeller

Bu çalışmada, elma, muz, fasulye, soya fasulyesi, biber, üzüm, mısır, pamuk, patates, pirinç, domates ve buğday bitkilerine ait sağlıklı ve hastalıklı yaprak görüntülerinin sınıflandırılması amacıyla önceden eğitilmiş modellere başvurulmaksızın sıfırdan tasarlanan CNN ve DNN derin öğrenme mimarileri ile Random Forest, SVM ve XGBoost makine öğrenmesi algoritmaları geliştirilmiş ve karşılaştırmalı olarak test edilmiştir.

Convolution Neural Network - CNN

Çalışmada derin öğrenme mimarisi olarak sıfırdan tasarlanan bir CNN modeli kullanılmıştır. Model 64x64x3 boyutunda girdi alacak şekilde tasarlanmış olup, sırasıyla 32, 64, 128, 256 ve 512 filtreden oluşan 5 evrişim katmanından (Conv2D) meydana gelmektedir; her evrişim katmanının ardından aşırı öğrenmeyi önlemek ve eğitimi hızlandırmak amacıyla BatchNormalization ve boyutsal indirgeme için MaxPooling2D uygulanmıştır. Tam bağlantılı katmana geçişten önce Flatten işlemi gerçekleştirilmiş, ardından 512 nöronlu bir Dense katman eklenmiş ve bu katmanın çıkışına BatchNormalization ile 0.5 dropout oranı uygulanmıştır. Tüm katmanlarda aktivasyon fonksiyonu olarak ReLU, optimizasyon algoritması olarak ise Adam kullanılmıştır. Model, sabit 50 epoch boyunca eğitilmiştir.

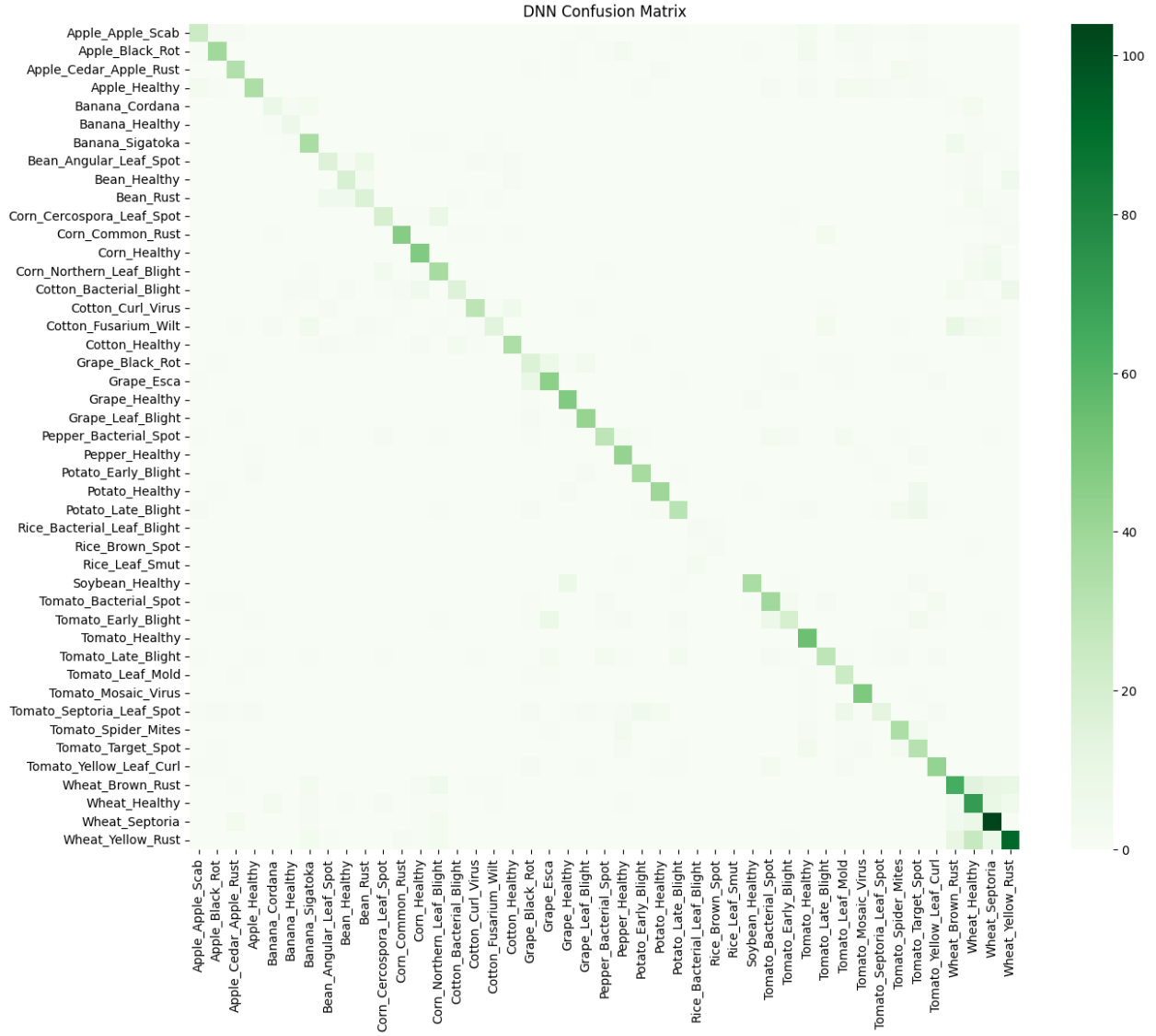
Şekil 1. Karmaşıklık matrisi CNN için



Deep Neural Network - DNN

Çalışmada kullanılan DNN modeli, $64 \times 64 \times 3$ boyutunda görüntü girdisi alacak şekilde tasarlanmıştır. Model sırasıyla 1024, 512 ve 256 nörondan oluşan üç gizli katmandan meydana gelmekte olup her katmanın ardından BatchNormalization uygulanmış ve sırasıyla 0.4, 0.3 ve 0.2 dropout oranlarıyla düzenlenmiştir. Tüm gizli katmanlarda aktivasyon fonksiyonu olarak ReLU kullanılmış, çıkış katmanında ise sınıf sayısı kadar nöron içeren Softmax aktivasyonu tercih edilmiştir, optimizasyon algoritması olarak ise Adam kullanılmıştır. CNN modeliyle tutarlı bir karşılaştırma sağlamak amacıyla sabit 50 epoch boyunca eğitilmiştir.

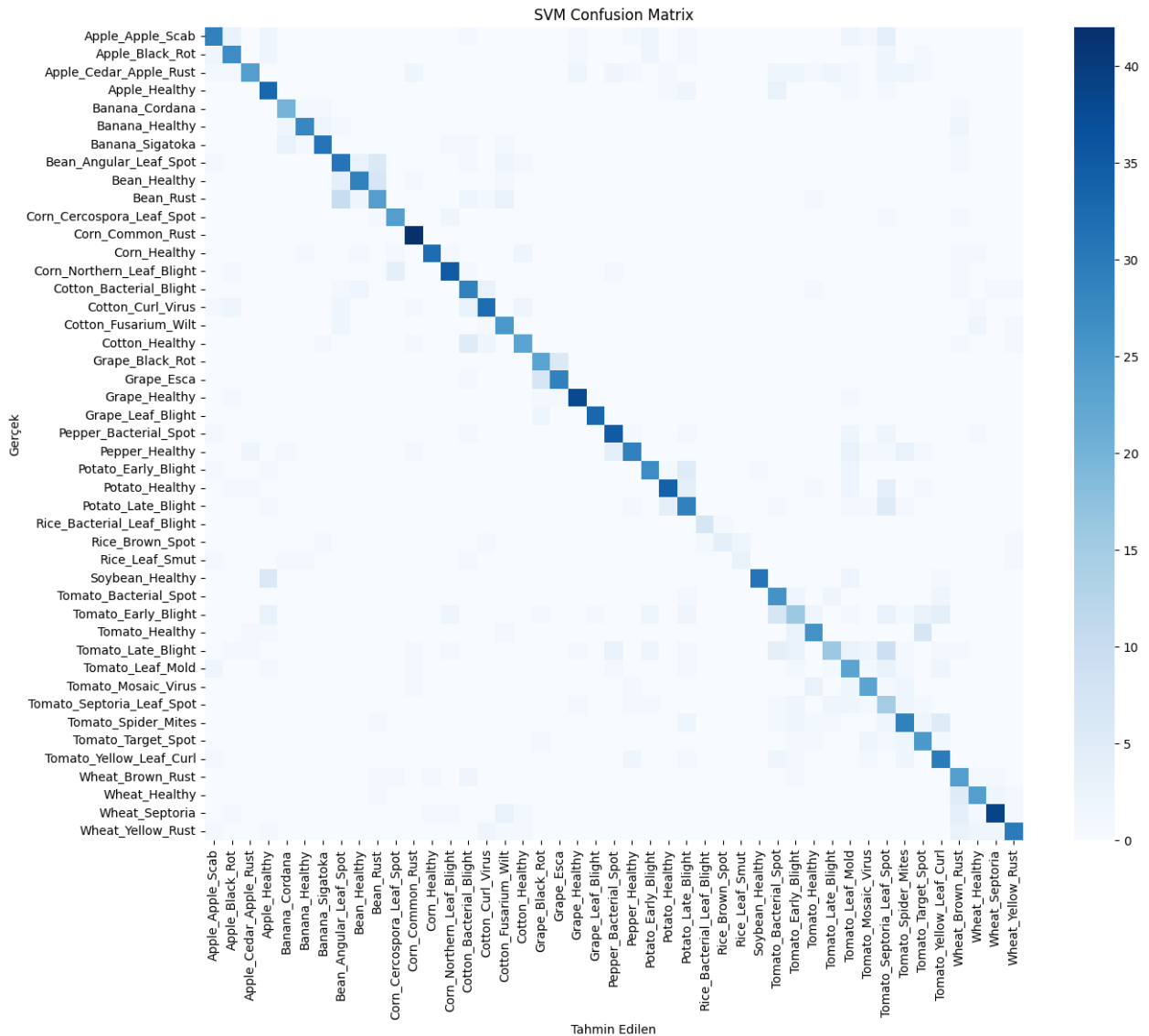
Şekil 2. Karmaşıklık matrisi DNN için



Support Vector Machine - SVM

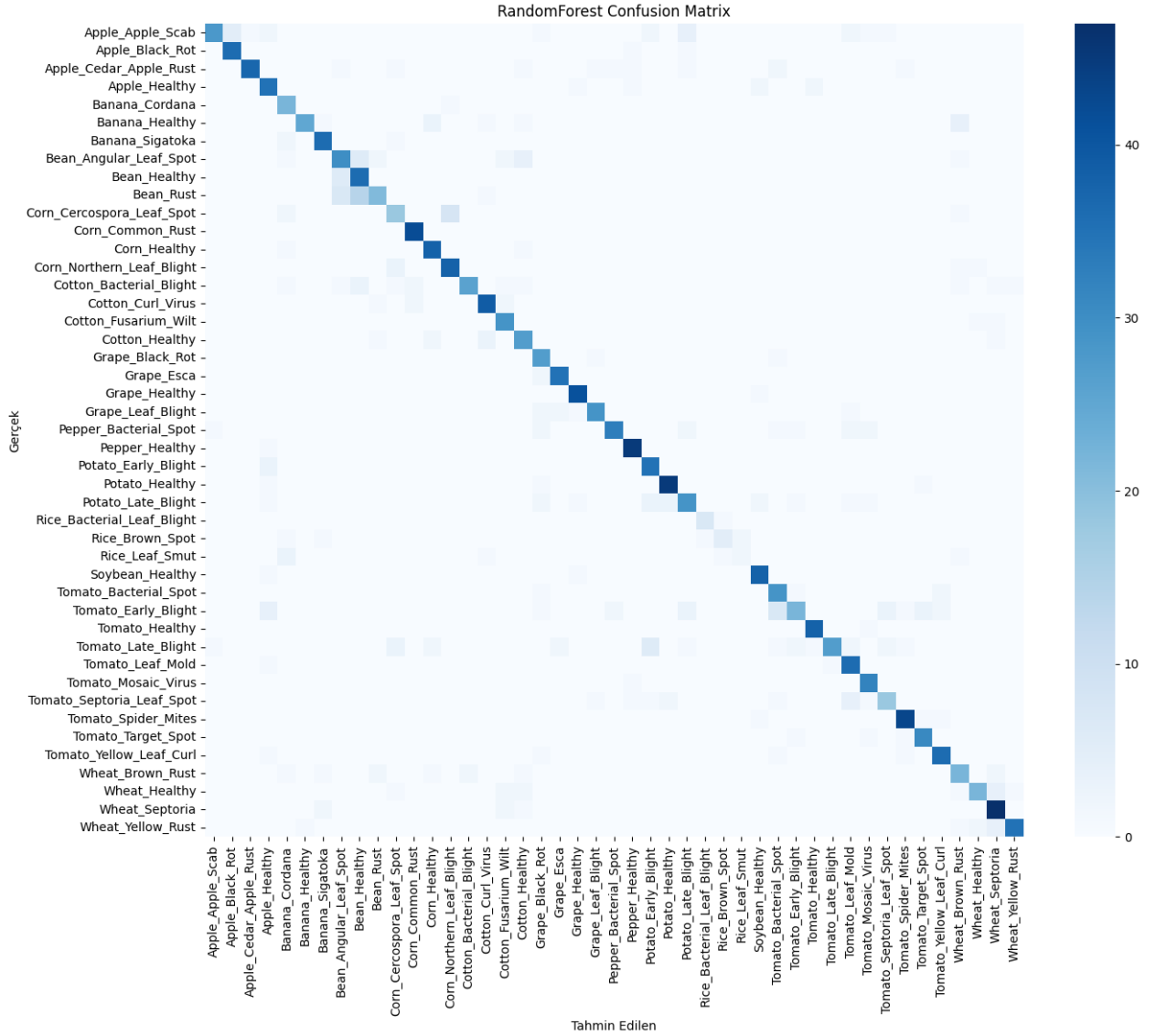
Çalışmada kullanılan SVM modelinin hiperparametre optimizasyonu, Grid Search ve 3 katlı çapraz doğrulama (cross-validation) yöntemiyle gerçekleştirilmiştir. Aday parametre uzayı olarak $C \in \{1, 10\}$ ve $\text{kernel} \in \{\text{linear}, \text{rbf}, \text{poly}\}$ değerleri denenmiş; toplam 6 aday kombinasyon için 18 model eğitimi gerçekleştirilmiştir. En iyi performans, $C=10$ ve RBF çekirdeği kombinasyonu ile elde edilmiş olup bu parametrelerle model %70.47 doğruluk, %72.18 hassasiyet, %70.47 duyarlılık ve %70.52 F1-skoru değerlerine ulaşmıştır.

Şekil 3. Karmaşıklık matrisi **SVM** için



Random Forest

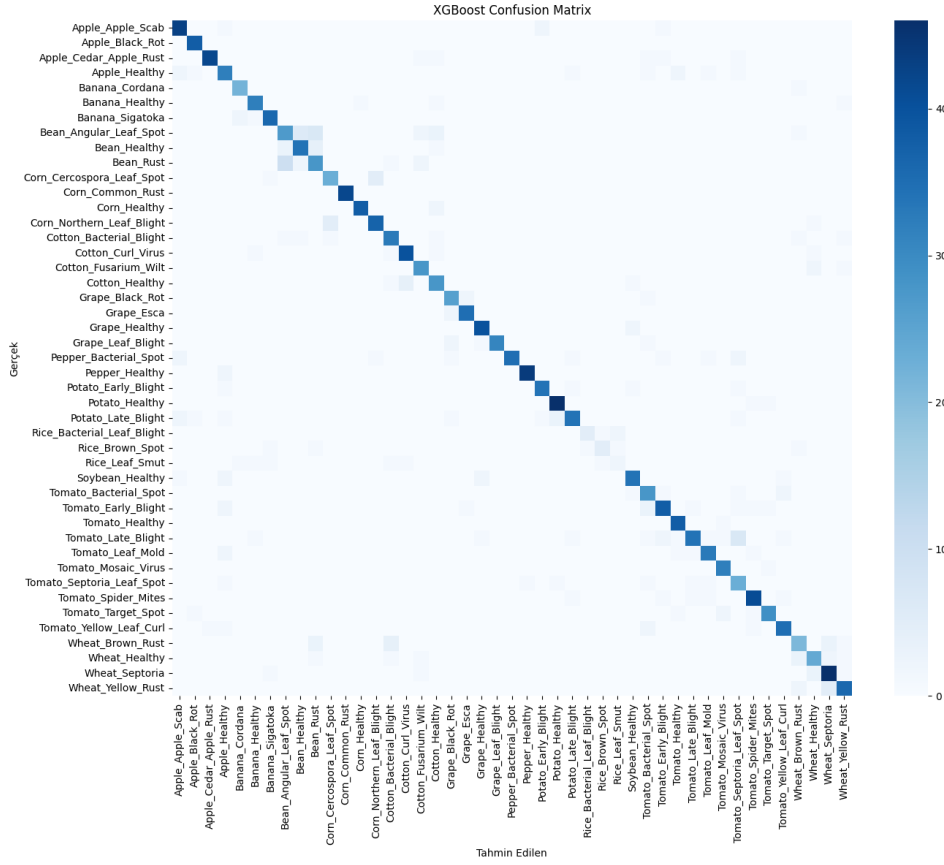
Random Forest modelinin hiperparametre optimizasyonu, Grid Search ve 3 katlı çapraz doğrulama yöntemiyle gerçekleştirilmiştir. Aday parametre uzayı olarak $n_estimators \in \{100, 200\}$, $max_depth \in \{None, 10, 20\}$ ve $min_samples_split \in \{2, 5\}$ değerleri denenmiş; toplam 8 aday kombinasyon için 24 model eğitimi gerçekleştirilmiştir. En iyi performans, $n_estimators=200$, $max_depth=None$ ve $min_samples_split=2$ kombinasyonu ile elde edilmiş olup bu parametrelerle model %81.52 doğruluk, %82.55 hassasiyet, %81.52 duyarlılık ve %81.00 F1-skoru değerlerine ulaşmıştır.

Şekil 4. Karmaşıklık matrisi **Random Forest** için

XGBoost

XGBoost modelinin hiperparametre optimizasyonu, Grid Search ve 3 katlı çapraz doğrulama yöntemiyle gerçekleştirilmiştir. Aday parametre uzayı olarak $n_estimators \in \{100, 200\}$, $max_depth \in \{3, 6\}$ ve $learning_rate \in \{0.1, 0.01\}$ değerleri denenmiş; toplam 8 aday kombinasyon için 24 model eğitimi gerçekleştirilmiştir. En iyi performans, $n_estimators=200$, $max_depth=3$ ve $learning_rate=0.1$ kombinasyonu ile elde edilmiş olup bu parametrelerle model %85.09 doğruluk, %85.34 hassasiyet, %85.09 duyarlılık ve %84.99 F1-skoru değerlerine ulaşmıştır.

Şekil 5. Karmaşıklık matrisi XGBoost için



SONUÇLAR ve TARTIŞMA

Bu çalışmada, `crop_disease_detection_dataset` kullanılarak bitki yapraklarındaki hastalıkların tespiti amacıyla CNN, DNN, Random Forest, SVM ve XGBoost algoritmaları geliştirilmiş ve karşılaştırmalı olarak değerlendirilmiştir. Elde edilen sonuçlar Tablo 6'da sunulmaktadır.

DeneySEL sonuçlar incelendiğinde, sıfırdan tasarlanan CNN modelinin %94.08 doğruluk oranıyla en yüksek performansı sergilediği görülmektedir. Bu sonuç, evrişim katmanlarının görüntüdeki uzamsal özellikleri hiyerarşik biçimde öğrenme kapasitesinin bitki hastalığı tespiti gibi görsel örüntü tanıma görevlerinde belirgin bir üstünlük sağladığını ortaya koymaktadır. Makine öğrenmesi algoritmaları arasında ise Grid Search optimizasyonu sonrasında XGBoost %85.09 ile en iyi performansı göstermiş, bunu %81.52 ile Random Forest izlemiştir. SVM, RBF çekirdeği ve C=10 parametreleriyle %70.47 doğruluğa ulaşırken DNN modeli %68.27 ile en düşük performansı sergilemiştir. DNN'nin görece düşük başarımı, tam bağlantılı katman yapısının görüntülerdeki uzamsal ilişkileri CNN kadar etkin biçimde modelleyememesinden kaynaklandığı değerlendirilmektedir.

Table 5. Sonuç Karşılaştırma Tablosu

Model	Accuracy	Precision	Recall	F1-Score
CNN	94.08%	94.62%	94.08%	94.04%
RandomForest	81.52%	82.55%	81.52%	81.00%
SVM	70.47%	72.18%	70.47%	70.52%
XGBoost	85.09%	85.34%	85.09%	84.99%
DNN	68.27%	69.14%	68.27%	67.61%

Makine öğrenmesi algoritmalarında ham piksel verisi yerine HOG ve HSV renk histogramı tabanlı özellik çıkarma yöntemlerinin kullanılması, özellikle Random Forest ve XGBoost modellerinin performansını belirgin şekilde artırmıştır. Nitekim Grid Search optimizasyonu öncesinde %43.28 doğrulukta kalan XGBoost modeli, optimizasyon sonrasında %85.09'a yükselmiş; benzer şekilde Random Forest %42.03'ten %81.52'ye ulaşmıştır. Bu bulgular, uygun özellik mühendisliği ve hiperparametre optimizasyonunun geleneksel makine öğrenmesi algoritmalarının performansı üzerindeki kritik etkisini açıkça göstermektedir.

Sonuç olarak, önceden eğitilmiş modellere başvurulmaksızın sıfırdan geliştirilen CNN mimarisinin literatürdeki transfer öğrenmesi tabanlı çalışmalarla rekabet edebilir düzeyde sonuçlar ürettiği görülmüştür. Gelecek çalışmalarda veri artırma tekniklerinin genişletilmesi, daha derin mimari tasarımları ve farklı özellik çıkarma yöntemlerinin denenmesiyle model performansının daha da iyileştirilebileceği öngörülmektedir.

KAYNAKÇA

- Atasoy H., Kutlu Y. (2022). Parameter Selection in Elliptical Fourier Series for Leaf Classification. Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi, 24(71), 375-381., Doi: 10.21205/deufmd.2022247104
- Ashwinkumar, S., S. Rajagopal, V. Manimaran, B. Jegajothi, Automated plant leaf disease detection and classification using optimal MobileNet based convolutional neural networks, Materials Today: Proceedings, Volume 51, Part 1, 2022, Pages 480-487, ISSN 2214-7853, <https://doi.org/10.1016/j.matpr.2021.05.584>
- Chowdhury, S., Kutlu, Y., Goongoon, N. (2026). Comparative Evaluation of YOLO-Based Pretrained Models for Multi-Class Vegetable Detection. Tethys Environmental Science, 3(1), 16-30, doi : 10.71350/tethysjournal.2

- Camgözlü Y, Kutlu Y. (2020). Examining the Difference Between Image Size, Background Color, Gray Picture and Color Picture in Leaf Classification with Deep Learning. *Akıllı sistemler ve uygulamaları dergisi.*, 3(2)
- Camgözlü Y, Kutlu Y. (2019). Analysis of Pooling Effect on CNN using Leaf Database. *Natural and Engineering Sciences*, 4(3), 115-120
- Camgözlü Y, Kutlu Y. (2023). Leaf Image Classification Based on Pre-trained Convolutional Neural Network Models. *Iskenderun Technical University*, 8(3), Doi: 10.28978/nesciences.1405175
- Eunice, J., A., J., Popescu, D. E., Chowdary, M. K., & Hemanth, J. (2022). Deep Learning-Based Leaf Disease Detection in Crops Using Images for Agricultural Applications. *Agronomy*, 12(10), 2395. <https://doi.org/10.3390/agronomy12102395>
- Gajjar, R., Gajjar, N., Thakor, V.J. *et al.* Real-time detection and identification of plant leaf diseases using convolutional neural networks on an embedded platform. *Vis Comput* **38**, 2923–2938 (2022). <https://doi.org/10.1007/s00371-021-02164-9>
- Kundu, R., U. Chauhan and S. P. S. Chauhan, "Plant Leaf Disease Detection using Image Processing," 2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM), Gautam Buddha Nagar, India, 2022, pp. 393-396, doi: 10.1109/ICIPTM54933.2022.9754170.
- Pandian, J. A., Kanchanadevi, K., Kumar, V. D., Jasińska, E., Goño, R., Leonowicz, Z., & Jasiński, M. (2022). A Five Convolutional Layer Deep Convolutional Neural Network for Plant Leaf Disease Detection. *Electronics*, 11(8), 1266. <https://doi.org/10.3390/electronics11081266>
- Russel, N.S., Selvaraj, A. Leaf species and disease classification using multiscale parallel deep CNN architecture. *Neural Comput & Applic* **34**, 19217–19237 (2022). <https://doi.org/10.1007/s00521-022-07521-w>
- Sunil S. Harakannanavar, Jayashri M. Rudagi, Veena I Puranikmath, Ayesha Siddiqua, R Pramodhini, Plant leaf disease detection using computer vision and machine learning algorithms, Global Transitions Proceedings, Volume 3, Issue 1, 2022, Pages 305-310, ISSN 2666-285X, <https://doi.org/10.1016/j.gltp.2022.03.016>
- Vallabhajosyula, S., Sistla, V., & Kolli, V. K. K. (2022). Transfer learning-based deep ensemble neural network for plant leaf disease detection. *Journal of Plant Diseases and Protection*, 129(3), 545–558. <https://www.jstor.org/stable/48706020>
<https://www.kaggle.com/datasets/ketangathibandhe/crop-disease-detection-dataset>
- Wei, K., Chen, B., Zhang, J., Fan, S., Wu, K., Liu, G., & Chen, D. (2022). Explainable Deep Learning Study for Leaf Disease Classification. *Agronomy*, 12(5), 1035. <https://doi.org/10.3390/agronomy12051035>