

# Klasik Makine Ogrenmesi ve Derin Ogrenme Yaklasimlariyla Bitki Yaprak Hastaligi Siniflandirmasi

Omar A. M. Issa  
Artificial Intelligence Engineering  
Ostim Technical University  
Ankara, Turkey  
Student No: 220212901

**Abstract**—Bitki hastaliklarinin erken ve dogru tespiti verim kaybini azaltmak ve tarimsal uretkenligi artirmak icin kritiktir. Uygulamada teshis hala buyuk oranda elle yapilir; bu da zaman alici, oznel ve olceklenmesi zor bir surectir. Bu rapor, veri hazirlama, klasik makine ogrenmesi ve derin ogrenme tekniklerini tek bir deneysel cercevede birlestiren, goruntu tabanli bitki yapragi hastaligi siniflandirma calismasini sunar. Histogram of Oriented Gradients ve Destek Vektor Makinasi tabanli bir el yapimi ozellik temel cizgisi, ResNet-18 ile transfer ogrenme kullanan derin ogrenme cozumune karsilastirilir. Deneyler PlantVillage veri kumesi uzerinde net bir egitim, dogrulama ve test protokoluyle yurutulur. Model performansi dogruluk, makro ortalamali F1, precision-recall analizi ve karmasiklik matrisleri ile degerlendirilir. Sonuclar, derin ogrenmenin daha yuksek ve daha dengeli performans sagladigini, klasik temel cizginin ise yorumlanabilirlik ve verimlilik acisinden degerli kaldigini gosterir. Tekrarlanabilirlik ve deneysel titizlik calisma boyunca vurgulanir.

**Index Terms**—Bitki hastaligi tespiti, goruntu siniflandirma, makine ogrenmesi, derin ogrenme, ResNet, HOG, SVM

## I. GIRIS VE PROBLEM TANIMI

Bitki hastaliklari dunya genelinde tarimsal uretkenlik icin onemli bir zorluktur. Erken tespit edilmezse hastaliklar ciddi verim kayiplarina ve urun kalitesinde dususlere yol acar. Bir-cok bitki hastaligi yapraklarda renk degisimi, doku bozulmasi veya duzensiz sekil gibi gorunur belirtilerle ortaya cikir. Geleneksel olarak bu belirtilerin belirlenmesi uzman bilgisi ve sahada tekrarli kontroller gerektirir; bu da maliyetli ve olceklenmesi zor bir yaklasimdir.

Yapay zeka bakis acisinden bitki yapragi hastaligi tespiti, denetimli cok sinifli goruntu siniflandirma problemi olarak tanimlanabilir. Bir yaprak goruntusu verildiginde amac, dogru hastalik sinifini atamak veya yapragi saglikli olarak tanimlamaktır. Bilgisayar gorusu ve derin ogrenmedeki ilerlemeler, otomatik teshisi mumkun kilmakta ve tarimsal uygulamalara olceklenebilir ve nesnel karar destegi saglamaktadır.

Bu projenin amaci, veri on isleme, temel modelleme, derin ogrenme, sayisal optimizasyon ve degerlendirme adimlarini bir araya getiren uctan uca bir siniflandirma boru hattı tasarlamak ve degerlendirmektir. Klasik ve derin ogrenme yaklasimlarini karsilastirarak, muhendislik acisinden guclu ve sinirli yonlerini ortaya koymak hedeflenmistir.

## II. VERI KUMESI TANIMI

Deneyler, bitki hastaligi siniflandirma arastirmalarinda yaygin olarak kullanan PlantVillage veri kumesi uzerinde yurutulmustur. Veri kumesi, kontrollu kosullarda cekilmis saglikli ve hastalikli yaprak goruntulerinden olusur; bu durum arka plan degiskenligini azaltir ve ogrenmeyi hastalikla ilgili gorsel oruntulere odaklar.

Bu projede kullanan islenmis veri kumesi 24 farkli sinif icermektedir. Bir sinif yapraksiz arka plan goruntulerini temsil ederken, diger siniflar belirli urun-hastalik kombinasyonlari veya saglikli urun kategorileridir. Guvenilir ve tarafsiz performans degerlendirmesi icin veri, sinif dagilimlarini koruyan stratified bir bolme stratejisiyle egitim, dogrulama ve test alt kumesine ayrilmistir.

TABLE I  
VERI KUMESI BOLUNMELERI OZETI

Bolum	Goruntu Sayisi	Yuzde
Egitim	19,862	70%
Dogrulama	4,256	15%
Test	4,257	15%

Sinif frekanslari incelendiginde hastalik kategorileri arasinda orta duzey dengesizlik gorulur. Bu nedenle, dogrulukla birlikte makro ortalamali F1 gibi sinif dengesizligini dikkate alan metrikler de vurgulanir. Ayrıntili gorsel analizler ve veri kesfi, eslik eden Jupyter Notebook icinde sunulmustur.

## III. YONTEM

Onerilen sistem, veri hazirlama, temel modelleme ve derin ogrenme tabanli siniflandirmayi iceren yapili ve tekrarlanabilir bir makine ogrenmesi is akisina sahiptir. Tasarim, yaklasimlar arasinda net ve adil bir karsilastirma yapilmasini hedefler.

### A. Veri On Isleme

Tum goruntuler egitim veya degerlendirme oncesinde sabit boyutlara yeniden olceklendirilir. Klasik ve derin ogrenme boru hatlari icin, hesaplama verimliliği ile temsil gucunu dengelemek amaciyla farkli goruntu boyutlari kullanilir. Veri seti,

herhangi bir model eğitiminden önce eğitim, doğrulama ve test alt kumesine ayrılır ve veri sızıntısı önlenir. Tekrarlanabilirliği sağlamak için Python, NumPy ve PyTorch kitaplıklarında rastgele tohumlar sabitlenir.

### B. Klasik Makine Öğrenmesi Temel Çizgisi

Referans model olarak elle tasarlanmış özelliklere dayalı bir görüntü sınıflandırma boru hattı uygulanmıştır. Histogram of Oriented Gradients (HOG), yaprak görüntülerindeki yerel kenar ve yapısal bilgiyi yakalamak için kullanılırken, renk histogramları hastalık belirtileriyle ilişkili kromatik özellikleri kodlar. Çıkarılan özellik vektörleri standardize edilir ve Doğrusal Destek Vektör Makinesi (SVM) ile sınıflandırılır.

SVM hiperparametre seçimi, eğitim verisinin stratified bir alt kumesi üzerinde capraz doğrulama ile rastgele arama yoluyla yapılır. Hesaplama yükünü azaltmak için çıkarılan özellikler diske onbelleklenir ve deneyler boyunca tekrar kullanılır; bu sayede boru hattı orta seviye donanımlarda daha verimli çalışır.

### C. Derin Öğrenme Modeli

Derin öğrenme yaklaşımı için transfer öğrenme kullanılarak ResNet-18 evrimsel sınır ağı seçilmiştir. Ağ, ImageNet üzerinde önceden eğitilmiş ağırlıklarla başlatılır ve son tam bağlantı katmanı hedef sınıf sayısına göre yeniden düzenlenir. İki konfigürasyon incelenmiştir: evrimsel omurganın dondurulması ve üst düzey katmanların kısmen ince ayar edilmesi.

Eğitim, capraz entropi kaybı ve Adam optimizör ile yürütülür. Doğrulama kaybına dayalı erken durdurma, asiri uyumu ve gereksiz hesaplamayı sınırlamak için uygulanır. Bu ayarlar, modelin verimli biçimde yakınsamasını ve genelleme performansını korumasını sağlar.

### IV. DENEYSEL TASARIM

Deneysel tasarım, modeller arasında adil ve seffaf bir karşılaştırma sağlayacak şekilde planlanmıştır. HOG+SVM boru hattı klasik temel çizgi olarak, ResNet-18 modeli ise geliştirilmiş derin öğrenme yaklaşımı olarak kullanılmıştır. Tüm modeller, aynı veri bölünmeleri ve değerlendirme protokolleri ile eğitilip test edilmiştir.

Sayısal optimizasyon deneysel düzenin önemli bir parçasıdır. Klasik modelde SVM hiperparametreleri rastgele arama ile optimize edilirken, derin öğrenme modelinde öğrenme oranı ve ağırlık azalma taramaları yapılmış ve erken durdurma tutarlı biçimde uygulanmıştır. Bu tasarım, en az bir temel model ve bir geliştirilmiş yaklaşım kullanma gereksinimini karşılar.

### V. DEĞERLENDİRME METRİKLERİ

Model performansı birden fazla tamamlayıcı metrikle değerlendirilir. Genel doğruluk referans olarak raporlanır ancak tek başına yeterli kabul edilmez. Makro ortalama F1, sınıflar arası dengeli performansı ölçmek için kullanılır; precision-recall analizi ve karmaşıklık matrisleri ise sınıf bazlı davranışlar hakkında ek görüş sağlar.

## VI. SONUÇLAR

Deneysel sonuçlar, derin öğrenme modelinin tüm metriklerde klasik temel çizgiden daha iyi performans gösterdiğini ortaya koyar. Özellikle ResNet-18 daha yüksek makro F1 skorlarına ulaşır; bu da sınıf dengesizliğine karşı daha sağlam bir performans anlamına gelir. Karmaşıklık matrisi analizi, hataların çoğu görsel olarak benzer hastalık kategorileri arasında ortaya çıkarken, sağlıklı ve arka plan sınıflarının genellikle yüksek güvenle sınıflandırıldığını gösterir.

## VII. TARTISMA VE HATA ANALIZI

Yanlış sınıflandırılan örnekler incelendiğinde hataların sıklıkla hastalıklar arasındaki ince görsel farklardan veya düşük kaliteli görüntülerden kaynaklandığı görülür. Notebook içinde yer alan gradyan tabanlı gorselleştirmeler, derin öğrenme modelinin yaprağın hastalıkla ilişkili bölgelerine odaklandığını doğrulayarak niteliksel bir doğrulama sağlar.

## VIII. SAYISAL YÖNTEMLER VE OPTİMİZASYON KATKISI

Sayısal optimizasyon teknikleri bu projede önemli rol oynar. SVM için hiperparametre araması ve derin öğrenme modeli için öğrenme oranı taramaları, yakınsama kararlılığını ve genelleme performansını iyileştirir. Erken durdurma, asiri uyumu ve gereksiz eğitim iterasyonlarını azaltarak sayısal yöntemlerin uygulamalı makine öğrenmesindeki etkisini gösterir.

## IX. TEKRARLANABİLİRLİK VE MUHENDİSLİK KALİTESİ

Tekrarlanabilirlik, sabit rastgele tohumlar, modüler kod yapısı ve deneysel ayarların açıkça belgelendirilmesiyle sağlanmıştır. Veri bölünmeleri diske kaydedilmiş ve ortam bilgileri günlüklere yazdırılarak deneylerin aynı koşullarda tekrarlanması kolaylaştırılmıştır.

## X. SINIRLAMALAR VE GELECEK ÇALIŞMALAR

Güçlü performansa rağmen önerilen sistemin sınırlamaları vardır. Veri kumesi büyük oranda kontrollü koşullarda çekilmiş görüntülerden oluştuğu için gerçek dünya tarımsal ortamlara genellenebilirlik sınırlı olabilir. Gelecek çalışmalarda alan uyarlaması, daha çeşitli veri toplama ve gerçek zamanlı uygulamalar ele alınabilir.

## XI. SONUÇ

Bu rapor, bitki yaprak hastalığı sınıflandırmasını hem klasik makine öğrenmesi hem de derin öğrenme yaklaşımlarıyla kapsamlı biçimde ele almıştır. Derin öğrenme modelleri daha yüksek performans sunarken, klasik temel çizgiler yorumlanabilirlik ve verimlilik açısından değerini korur. Proje, veri işleme, model tasarımı, sayısal optimizasyon ve değerlendirmeyi tekrarlanabilir bir mühendislik çerçevesinde başarıyla birleştirmiştir.

## KAYNAKLAR

- [1] D. Hughes and M. Salathe, "An open access repository of images on plant health to enable the development of mobile disease diagnostics," arXiv:1511.08060, 2015.
- [2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [3] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, 2011.
- [4] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in Advances in Neural Information Processing Systems, 2019.