

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 denyesel cercevede birlestiren, goruntu tabanli bitki yapragi hastaligi siniflandirma calismasini sunar. Histogram of Oriented Gradients ve Destek Vektor Makinasi tabanli bir el yapimi ozellik cizgisi, ResNet-18 ile transfer ogrenme kullanan derin ogrenme cozumune karsilastirilir. Deneyler PlantVillage veri kumesi üzerinde net bir egitim, doğrulama ve test protokoluyle yurutulur. Model performansi doğruluk, makro ortalamali F1, precision-recall analizi ve karmasiklik matrisleri ile degerlendirilir. Sonuclar, derin ogrenmenin daha yüksek ve daha dengeli performans sagladigini, klasik temel cizginin ise yorumlanabilirlik ve verimlilik acisindan degerli kaldigini gosterir. Tekrarlanabilirlik ve denyesel 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. Birçok bitki hastaligi yapraklarda renk degisimi, doku bozulmasi veya duzensiz sekil gibi gorunur belirtilerle ortaya cikar. Geleneksel olarak bu belirtilerin belirlenmesi uzman bilgisi ve sahada tekrarli kontroller gerektirir; bu da maliyetli ve olceklenmesi zor bir yaklasimdir.

Yapay zeka bakis acisindan 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 tanimlamaktir. Bilgisayar gorusu ve derin ogrenmedeki ilerlemeler, otomatik teshisi mumkun kilmakta ve tarimsal uygulamalara olceklenebilir ve nesnel karar destegi saglamaktadir.

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

II. VERI KUMESI TANIMI

Deneysel, bitki hastaligi siniflandirma arastirmalarinda yayin olarak kullanılan PlantVillage veri kumesi üzerinde yurultmustur. Veri kumesi, kontrollu kosullarda cekilmis saglikli ve hastalilik yaprak goruntulerinden olusur; bu durum arka plan degiskenligini azaltir ve ogrenmeyi hastalikla ilgili gorsel oruntulere odaklar.

Bu projede kullanılan islenmis veri kumesi 24 farkli sinif icermektedir. Bir sinif yapraksız arka plan goruntulerini temsil ederken, diger siniflar belirli urun-hastalik kombinasyonları veya saglikli urun kategorileridir. Guvenilir ve tarafsız performans degerlendirmesi icin veri, sinif dagilimlarini koruyan stratified bir bolme stratejisiyle egitim, doğrulama ve test alt kumesine ayrlmistir.

TABLE I
VERI KUMESI BOLUNMELERI OZETI

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

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

III. YONTEM

Onerilen sistem, veri hazırlama, temel modelleme ve derin ogrenme tabanlı siniflandirmayı iceren yapıli ve tekrarlanabilir bir makine ogrenmesi is akisina sahiptir. Tasarim, yaklasimlar arasında net ve adil bir karsilastirma yapılmasını hedefler.

A. Veri On Isleme

Tum goruntuler egitim veya degerlendirme oncesinde sabit boyutlara yeniden olceklenir. Klasik ve derin ogrenme boru hatları icin, hesaplama verimliliği ile temsil gecesini degelemek amacıyla farklı goruntu boyutları kullanılır. Veri seti,

herhangi bir model egitiminden once egitim, dogrulama ve test alt kumesine ayrilir ve veri sizintisi onlenir. Tekrarlanabilirlik saglamak icin Python, NumPy ve PyTorch kitaplıklarında rastgele tohumlar sabitlenir.

B. Klasik Makine Ogrenmesi Temel Cizgisi

Referans model olarak elle tasarlanmis ozelliklere dayali bir goruntu siniflandırma boru hatti uygulanmistir. Histogram of Oriented Gradients (HOG), yaprak goruntulerindeki yerel kenar ve yapisal bilgiyi yakalamak icin kullanilirken, renk histogrammlari hastalik belirtileriyle iliskili kromatik ozellikleri kodlar. Cikarilan ozellik vektorleri standardize edilir ve do-grusul Destek Vektor Makinası (SVM) ile siniflandirilir.

SVM hiperparametre secimi, egitim verisinin stratified bir alt kumesi üzerinde capraz dogrulama ile rastgele arama yoluyla yapilir. Hesaplama yukunu azaltmak icin cikarilan ozellikler diske onbelleklenir ve deneyler boyunca tekrar kul-lanilir; bu sayede boru hatti orta seviye donanimlarda daha verimli calisir.

C. Derin Ogrenme Modeli

Derin ogrenme yaklasimi icin transfer ogrenme kullanilarak ResNet-18 evrisimsel sinir agi secilmistir. Ag, ImageNet uzerinde onceden egitilmis agirliklarla baslatilir ve son tam baglanti katmani hedef sinif sayisina gore yeniden duzenlenir. Iki konfigurasyon incelenmistir: evrisimsel omurganin dondurulmesi ve ust duzey katmanların kismen ince ayar edilmesi.

Egitim, capraz entropi kaybi ve Adam optimizer ile yurtilir. Dogrulama kaybina dayali erken durdurma, asiri uyumu ve gereksiz hesaplamayı sinirlamak icin uygulanir. Bu ayarlar, modelin verimli bicimde yakinsamasini ve genelleme per-formansini korumasini saglar.

IV. DENEYSEL TASARIM

Deneysel tasarrum, modeller arasında adil ve seffaf bir kar-silastirma saglayacak sekilde planlanmistir. HOG+SVM boru hatti klasik temel cizgi olarak, ResNet-18 modeli ise gelis-tirilmis derin ogrenme yaklasimi olarak kullanilmistir. Tum modeller, ayni veri bolummeleri ve degerlendirme protokollerii ile egitilip test edilmistir.

Sayisal optimizasyon deneysel duzenin onemli bir parcasidir. Klasik modelde SVM hiperparametreleri rastgele arama ile optimize edilirken, derin ogrenme modelinde ogrenme orani ve agirlik azalma taramalari yapılmış ve erken durdurma tutarlı bicimde uygulanmistir. Bu tasarrum, en az bir temel model ve bir gelistirilmis yaklasim kullanma gereksinimini karsilar.

V. DEGERLENDIRME METRIKLERI

Model performansi birden fazla tamamlayici metrikle degerlendirilir. Genel dogruluk referans olarak raporlanır ancak tek basina yeterli kabul edilmez. Makro ortalamalı F1, siniflar arasi dengeli performansi olcmek icin kullanılır; precision-recall analizi ve karmasiklik matrisleri ise sinif bazlı davranislar hakkında ek gorus saglar.

VI. SONUCLAR

Deneysel sonuclar, derin ogrenme modelinin tum metriklerde klasik temel cizgiden daha iyi performans gösterdigini ortaya koyar. Ozellikle ResNet-18 daha yüksek makro F1 skorlarına ulasir; bu da sinif dengesizligine karsi daha sağlam bir performans anlamina gelir. Karmasiklik matrisi analizi, hatalarin cogu gorunsel olarak benzer hastalik kategorileri arasında ortaya cikarken, saglikli ve arka plan siniflarının genellikle yüksek guvenle siniflandirildigini gösterir.

VII. TARTISMA VE HATA ANALIZI

Yanlis siniflandirilan ornekler incelendiginde hatalarin siklikla hastalıklar arasındaki ince gorsel farklardan veya dusuk kaliteli goruntulerden kaynaklandigi gorulur. Notebook icinde yer alan gradyan tabanlı gorsellestirmeler, derin ogrenme modelinin yapragın hastalikla iliskili bolgelerine odaklandigini doğrulayarak niteliksel bir dogrulama saglar.

VIII. SAYISAL YONTEMLER VE OPTIMIZASYON KATKISI

Sayisal optimizasyon teknikleri bu projede onemli rol oynar. SVM icin hiperparametre araması ve derin ogrenme modeli icin ogrenme orani taramalari, yakinssama kararlilikini ve genelleme performansini iyilestirir. Erken durdurma, asiri uyumu ve gereksiz egitim iterasyonlarını azaltarak sayisal yon-temlerin uygulamali makine ogrenmesindeki etkisini gösterir.

IX. TEKRARLANABILIRLIK VE MUHENDISLIK KALITESI

Tekrarlanabilirlik, sabit rastgele tohumlar, modular kod yapisi ve deneysel ayarların acikça belgelendirilmesiyle saglanmistir. Veri bolummeleri diske kaydedilmiş ve ortam bilgileri gunluklere yazdirilarak deneylerin ayni kosullarda tekrarlanmasi kolaylastirilmistir.

X. SINIRLAMALAR VE GELECEK CALISMALAR

Guclu performansa ragmen onerilen sistemin sinirlamaları vardır. Veri kumesi buyuk oranda kontrollu kosullarda cekilmis goruntulerden olustugu icin gercek dunya tarimsal ortamlara genellenebilirlik sinirli olabilir. Gelecek calismalarda alan uyarlamasi, daha cesitli veri toplama ve gercek zamanlı uygulamalar ele alinabilir.

XI. SONUC

Bu rapor, bitki yaprak hastaligi siniflandirmasını hem klasik makine ogrenmesi hem de derin ogrenme yaklasimlarıyla kapsamlı bicimde ele almıştır. Derin ogrenme modelleri daha yüksek performans sunarken, klasik temel cizgiler yorumlanabilirlik ve verimlilik acisinden degerini korur. Proje, veri isleme, model tasarımı, sayisal optimizasyon ve degerlendirmeyi tekrarlanabilir bir muhendislik cercevesinde basariyla birelestirmistir.

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.