

Derin Öğrenme Tabanlı CNN Modelleriyle Pirinç Yaprağı Hastalıklarının Sınıflandırılması ve Performans Karşılaştırması

İlhan Emre ADAK

Bursa Teknik Üniversitesi Mühendislik ve Doğa Bilimleri Fakültesi

Bilgisayar Mühendisliği Bölümü

adak.ie@hotmail.com

Berkan PARLAK

Bursa Teknik Üniversitesi Mühendislik ve Doğa Bilimleri Fakültesi

Bilgisayar Mühendisliği Bölümü

parlakberkanz@gmail.com

Proje Kaynak Kodu

Bu çalışmada kullanılan tüm modellerin eğitim süreçleri, veri ön işleme adımları ve değerlendirme analizleri, aşağıdaki GitHub deposunda paylaşılmıştır:

GitHub Bağlantısı: https://github.com/Ilhanemreadak/rice_leaves_daisies

1.ÖZET

Bu projede, pirinç yapraklarında görülen hastalıkları sınıflandırmak amacıyla dört farklı derin öğrenme modeli eğitilmiş ve performansları karşılaştırılmıştır. Kullanılan modeller MobileNet, VGG16, Xception ve VGG19'dur. Veri kümesi 5932 adet etiketlenmiş görüntüden oluşmaktadır ve Bacterial Blight, Blast, Brown Spot ve Tungro olmak üzere dört sınıfa ayrılmaktadır. Görüntüler eğitim ve test olmak üzere %80–%20 oranında bölünmüş, eğitim sırasında veri artırma tekniklerinden (döndürme, yakınlaştırma, yatay çevirme) yararlanılmıştır. Modeller transfer öğrenme yaklaşımıyla önceden eğitilmiş ImageNet ağırlıkları kullanılarak oluşturulmuştur. Eğitim sonrasında yapılan değerlendirmelerde, doğruluk, kesinlik (precision), duyarlılık (recall) ve F1 skoru gibi metrikler hesaplanmıştır. Sonuçlara göre en yüksek sınıflandırma başarımına MobileNet modeli ulaşmıştır. Bu çalışma, derin öğrenme modellerinin tarımsal hastalık tespiti gibi pratik alanlarda başarıyla kullanılabileceğini göstermektedir.

2.GİRİŞ

Tarımda bitki hastalıklarının erken teşhisi, verimliliği artırmak ve ürün kayıplarını en aza indirmek açısından büyük önem taşımaktadır. Özellikle pirinç gibi temel besin kaynaklarında görülen yaprak hastalıklarının hızlı ve doğru şekilde teşhis edilmesi, sadece çiftçilerin ekonomik kayıplarını azaltmakla kalmaz, aynı zamanda gıda güvenliğine de katkı sağlar. Geleneksel yöntemlerle hastalık tanısı, genellikle uzman kişiler tarafından görsel incelemeye dayanmakta ve bu durum zaman alıcı ve hataya açık olabilmektedir. Günümüzde yapay zekâ ve görüntü işleme tekniklerinin tarım alanında kullanımı giderek yaygınlaşmakta ve özellikle bitki hastalıklarının erken teşhisi için önemli fırsatlar sunmaktadır.

Bitki hastalıklarının zamanında ve doğru bir şekilde tespit edilmesi, verim kayıplarının önlenmesi ve sürdürülebilir tarım uygulamalarının desteklenmesi açısından kritik öneme sahiptir. Geleneksel yöntemlerle yapılan hastalık tespiti genellikle zaman alıcı ve uzmanlık gerektiren bir süreçtir. Bu nedenle, otomatik sınıflandırma yöntemlerinin geliştirilmesi hem zaman hem de maliyet açısından avantajlar sağlamaktadır.

Bu bağlamda, projede dört farklı önceden eğitilmiş derin evrişimli sinir ağı (CNN) modeli olan MobileNet, VGG16, Xception ve VGG19 kullanılmıştır. Modellerin eğitiminde transfer öğrenme ve veri artırma teknikleri uygulanmış, sonuçlar doğruluk, kesinlik, duyarlılık ve F1 skoru gibi metriklerle karşılaştırılmıştır. Böylece farklı modellerin performansları tarımsal görüntü sınıflandırma alanındaki etkinlikleri açısından değerlendirilmiştir.

3.LİTERATÜR İNCELEMESİ

Tarım alanında bitki hastalıklarının erken tespiti, ürün kalitesi ve verimliliği açısından kritik bir öneme sahiptir. Son yıllarda derin öğrenme teknikleri, özellikle konvolüsyonel sinir ağları (CNN), bu alanda yaygın olarak kullanılmaya başlamıştır. Farklı CNN mimarileriyle yapılan çalışmalar, pirinç yaprağı hastalıklarının otomatik sınıflandırılmasında yüksek doğruluk oranları elde ettiğini göstermektedir.

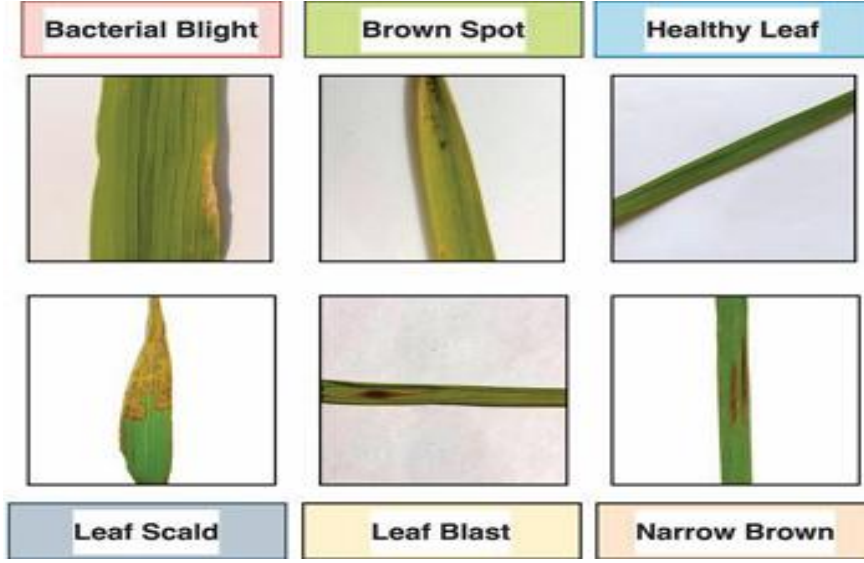
Örneğin, VGG19 tabanlı transfer öğrenme yöntemleri ile yapılan çalışmalarda %96 civarında doğruluk sağlanmıştır. Bu tür derin modeller, farklı hastalık türlerini ayırt etmekte başarılı olsa da, yüksek hesaplama ihtiyacı nedeniyle uygulama esnasında bazı kısıtlamalarla karşılaşabilmektedir. Öte yandan, MobileNet gibi daha hafif ve optimize edilmiş modeller, benzer doğruluk oranları verirken daha hızlı ve verimli sonuçlar sunmaktadır. Bu da tarımsal alanlarda taşınabilir cihazlarla gerçek zamanlı analiz yapmak için büyük avantaj sağlamaktadır.

Kendi deneyimlerimizde de MobileNet modeli yaklaşık %99,8 doğrulukla en yüksek performansı gösterdi. Xception modeli ise %98,9 gibi yüksek bir doğrulukla başarılı olurken, VGG16 ve VGG19 modelleri sırasıyla %89 ve %86 civarında doğruluk elde etti. Bu durum, literatürde öne çıkan modellerin genel eğilimlerini yansıtmakla birlikte, eğitim süresi ve veri seti özelliklerinin model başarısına etkisini de göstermektedir. Özellikle daha derin ve karmaşık yapılar daha uzun eğitim süreleri ve daha fazla hesaplama kaynağı gerektirmektedir.

Bununla birlikte, bazı araştırmalar hiperspektral görüntüleme gibi ileri tekniklerle doğruluk oranlarını %98'in üzerine çıkarabilmekte, ancak bu tür yöntemler veri toplama ve işleme açısından daha karmaşık sistemler gerektirmektedir. Bizim uygulamamızda ise standart görüntü verisiyle ve ADAM optimizasyonu ile %99'a yakın yüksek doğruluklar elde edildi.

Özetle, literatürde ve deneyimlerimizde MobileNet gibi hafif CNN modellerinin, pirinç yaprağı hastalıklarının sınıflandırılmasında hem yüksek doğruluk hem de pratiklik açısından uygun olduğu görülmektedir. Daha ağır modeller performans açısından bazı avantajlar sunsa da, hesaplama süresi ve kaynak tüketimi bakımından optimize edilmeleri gerekmektedir. Bu bulgular, tarımda yapay zekâ uygulamalarının geliştirilebilmesi için uygun model seçiminde yol gösterici olabilir.

4.MATERYAL VE YÖNTEM



[Şekil 1]: Veri setindeki pirinç yaprağı hastalıkları ve sağlıklı yaprağın sınıfları

4.1. Veri Kümesi

Bu çalışmada kullanılan veri kümesi, dört farklı pirinç yaprağı hastalığını içeren toplam 5932 görselden oluşmaktadır. Görseller 224x224 piksel çözünürlüğünde, RGB formatındadır ve dört sınıfa ayrılmıştır: Bacterial Blight, Blast, Brown Spot ve Tungro. Her sınıfın örnek sayısı birbirine yakın olacak şekilde dengelenmiştir.

Veri kümesi, eğitim ve test olarak ikiye bölünmüştür; eğitim için %80, test için %20 oranı tercih edilmiştir. Sınıflar arasında dengeli bir dağılım sağlanması için rastgele seçme yöntemi uygulanmıştır. Bu dengeleme, model performansının sağlıklı değerlendirilmesi açısından önem taşımaktadır.

4.2. Veri Ön İşleme

Model eğitimi öncesinde tüm görseller 224x224 piksele yeniden ölçeklendirilmiş ve piksel değerleri [0,1] aralığına normalize edilmiştir. Ayrıca, modelin genelleme kabiliyetini artırmak amacıyla veri artırma teknikleri kullanılmıştır. Bu teknikler arasında yatay çevirme, rastgele döndürme, yakınlaştırma ve kaydırma gibi işlemler yer almıştır. Böylece, eğitim verisinde çeşitlilik artırılarak modelin farklı görsellere karşı dayanıklılığı yükseltilmiştir.

4.3. Kullanılan Modeller

Projede dört farklı önceden eğitilmiş CNN modeli değerlendirilmiştir: MobileNet, VGG16, Xception ve VGG19. Bu modeller, ImageNet veri kümesi üzerinde önceden eğitilmiş ağırlıklarla Keras kütüphanesinden alınmıştır.

Modellerin üst katmanları kaldırılarak yerine projenin dört sınıfına uygun olarak Flatten, Dense ve Dropout katmanları eklenmiştir. Böylece, önceden öğrenilen özellikler sabitlenirken sınıflandırma kısmı proje veri kümesine özgü olarak yeniden öğrenilmiştir.

4.4. Eğitim Süreci

Eğitimde categorical_crossentropy kayıp fonksiyonu ve Adam optimizasyon algoritması kullanılmıştır. Öğrenme oranı varsayılan 0.001 olarak belirlenmiştir. Ayrıca, aşırı öğrenmeyi engellemek ve en iyi model ağırlıklarını kaydetmek amacıyla EarlyStopping ve ModelCheckpoint callback'leri kullanılmıştır.

Her model 16 epoch boyunca eğitilmiş ve her epoch sonunda doğrulama performansı ölçülmüştür. Eğitim süresi açısından karşılaştırıldığında, A100 GPU üzerinde ortalama 24 dakika gibi kısa bir sürede tamamlanmıştır.

Model bazında elde edilen ortalama doğruluk ve performans değerleri ise şu şekildedir:

- **MobileNet:** %99.83 doğruluk, yüksek precision, recall ve F1 skorları ile en başarılı model olarak öne çıkmıştır.
- **Xception:** %98.90 doğrulukla ikinci sırada yer almıştır.
- **VGG16:** %89.20 doğruluk ile orta düzey performans göstermiştir.
- **VGG19:** %85.90 doğrulukla diğer modellere göre daha düşük performans kaydetmiştir.

Bu sonuçlar, MobileNet'in hem yüksek doğruluğu hem de hızlı eğitim süresiyle pratik ve etkili bir çözüm olduğunu göstermektedir. VGG tabanlı modeller ise daha karmaşık yapılar olmasına rağmen beklenen başarıyı tam anlamıyla sağlayamamış, özellikle eğitim süresi ve doğruluk açısından daha geride kalmıştır.

5. SONUÇLAR ve DEĞERLENDİRMELER

Bu çalışmada, pirinç yaprağı hastalıklarının sınıflandırılması amacıyla dört farklı transfer öğrenme tabanlı derin öğrenme modeli olan MobileNet, Xception, VGG16 ve VGG19 karşılaştırmalı olarak incelenmiştir. Tüm modeller, aynı eğitim-veri kümesi üzerinde eğitilerek doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1-score gibi temel sınıflandırma performans metrikleri açısından detaylı biçimde değerlendirilmiştir. Bunun yanında, modellerin eğitim süreleri de dikkate alınarak hem başarımlar hem de işlem verimliliği açısından bütüncül bir analiz gerçekleştirilmiştir.

Aşağıda, dört modelin temel performans metriklerine ait sonuçlar tablo olarak sunulmuştur:

Model	Accuracy	Precision	Recall	F1-Score
MobileNet	0.998312	0.998315	0.998312	0.998312
VGG16	0.891983	0.891576	0.891983	0.891583
Xception	0.989030	0.989180	0.989030	0.989051
VGG19	0.859072	0.857220	0.859072	0.857707

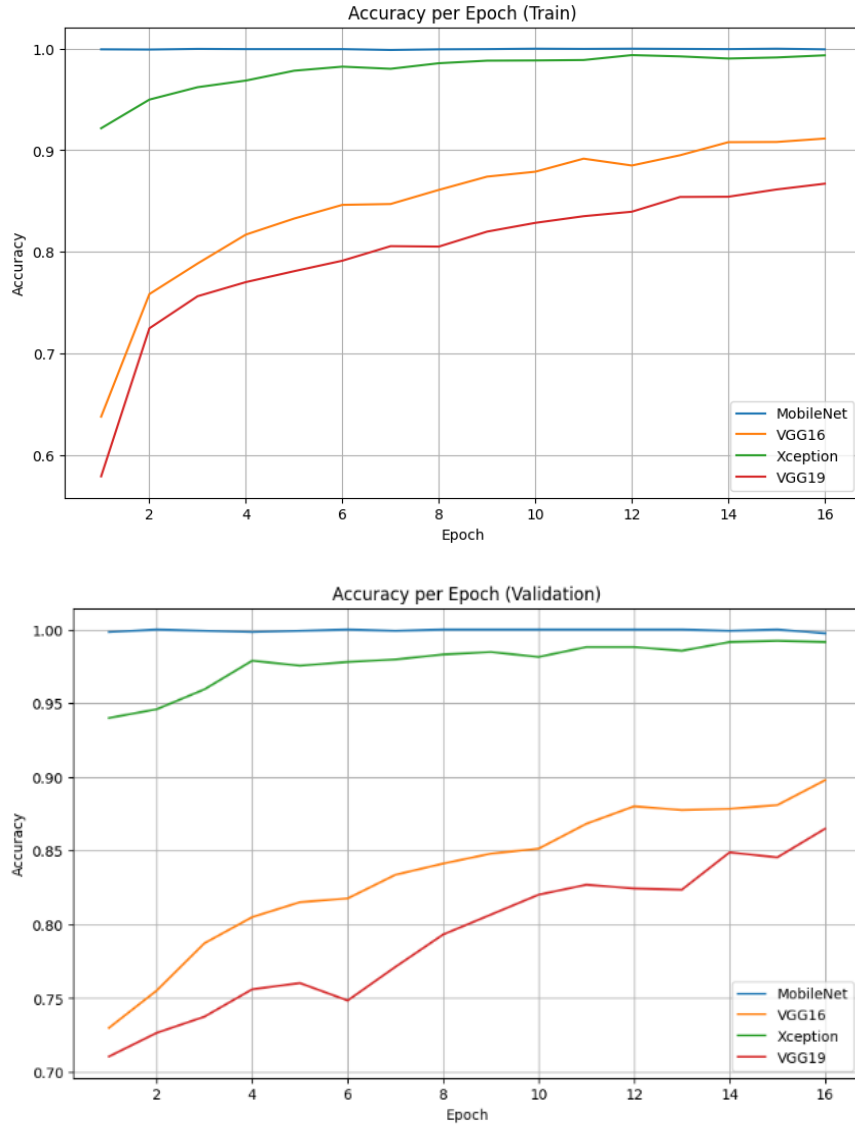
[Tablo 1. Modellerin Performans Karşılaştırması]

Tablo verileri incelendiğinde, MobileNet modelinin dört metrikte de en yüksek başarıyı gösterdiği görülmektedir. Özellikle %99.83 doğruluk oranı ile neredeyse hatasız sınıflandırma yapabilen bu model, hem sınıf içi hem sınıflar arası ayrımı etkili biçimde gerçekleştirmiştir. MobileNet'in bu başarısı, daha hafif yapısı sayesinde hızlı öğrenme kapasitesi ve düşük overfitting riski ile açıklanabilir.

İkinci en yüksek performans Xception modeline aittir. %98.9 doğruluk oranı ile oldukça başarılı bir sonuç sergileyen bu model, daha derin ve karmaşık bir mimariye sahip olmasına rağmen, dengeli metrik sonuçları ile dikkat çekmektedir. Xception, özellikle transfer öğrenme süreçlerinde güçlü özellik çıkarımı yapmasıyla bilinir; bu da veri kümesinin yapısıyla iyi bir uyum gösterdiğini düşündürmektedir.

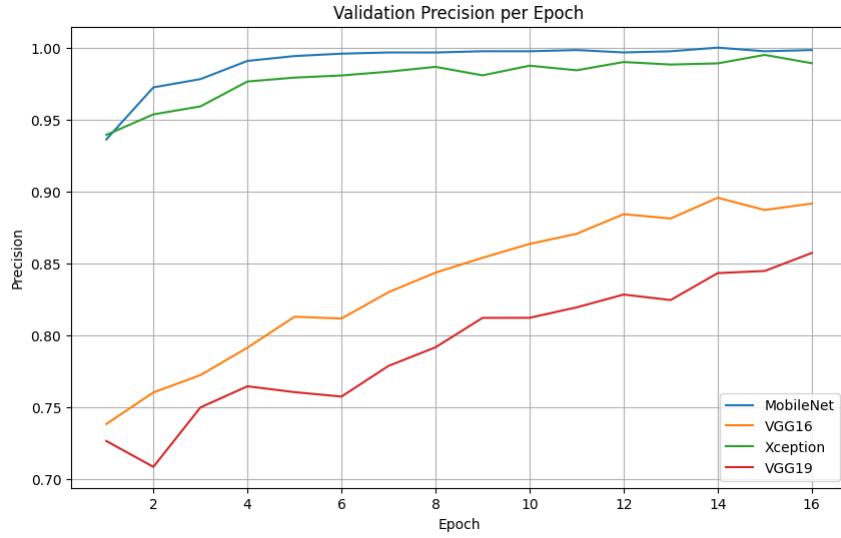
VGG16 ve VGG19 modelleri ise nispeten daha düşük doğruluk oranlarına sahiptir (%89.19 ve %85.91). Bu durum, bu modellerin daha eski mimarilere sahip olmaları ve veri setine yeterince genelleyici öğrenme sağlayamamalarından kaynaklanabilir. Ayrıca eğitim sürecinde overfitting belirtilerinin gözlemlenmesi bu modellerin genel başarımlarını düşürmüştür.

Sonuç olarak, hem sınıflandırma başarımları hem de hesaplama verimliliği açısından değerlendirildiğinde, MobileNet modeli bu çalışmada en başarılı yaklaşım olarak ortaya çıkmıştır. Xception ikinci sırayı alırken, VGG tabanlı modeller performans açısından geride kalmıştır. Bu çıkarımlar, ileride yapılacak çalışmalarda daha karmaşık veri kümeleri ve farklı derin öğrenme mimarilerinin denenmesi açısından yol gösterici olabilir.



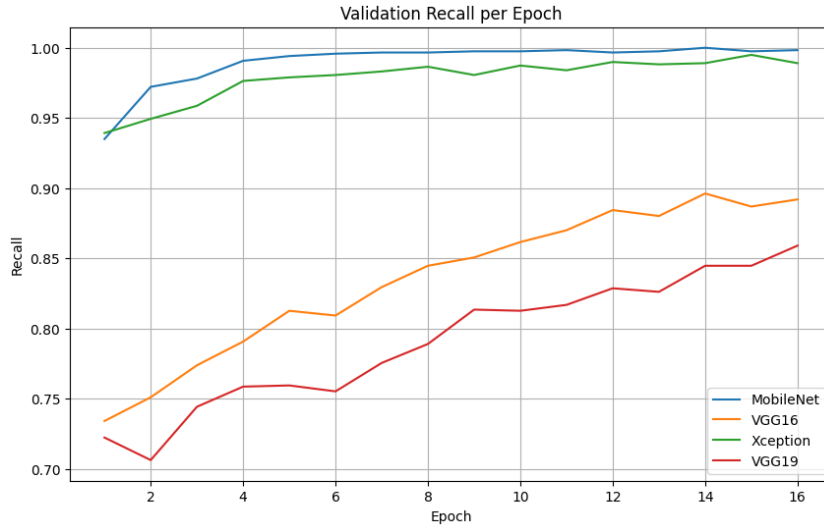
[Grafik 1 ve 2: Modellerin Accuracy Karşılaştırması]

Grafikte görüldüğü üzere MobileNet modeli %99.83 gibi oldukça yüksek bir doğruluk oranı yakalamıştır. Bu sonuç, MobileNet'in pirinç yaprağı hastalıklarını ayırt etmede üst düzey başarı sağladığını göstermektedir. Xception modeli de %98.90 doğruluk oranıyla benzer şekilde güçlü performans sergilemiştir. Buna karşın VGG16 ve VGG19 modelleri sırasıyla %89.20 ve %85.91 doğruluk oranları ile daha düşük başarı göstermiştir. Bu fark, MobileNet ve Xception gibi daha yeni ve optimize edilmiş mimarilerin, özellikle tarımsal görüntü sınıflandırması gibi karmaşık görevlerde geleneksel derin mimarilere kıyasla daha üstün olduğunu ortaya koymaktadır.



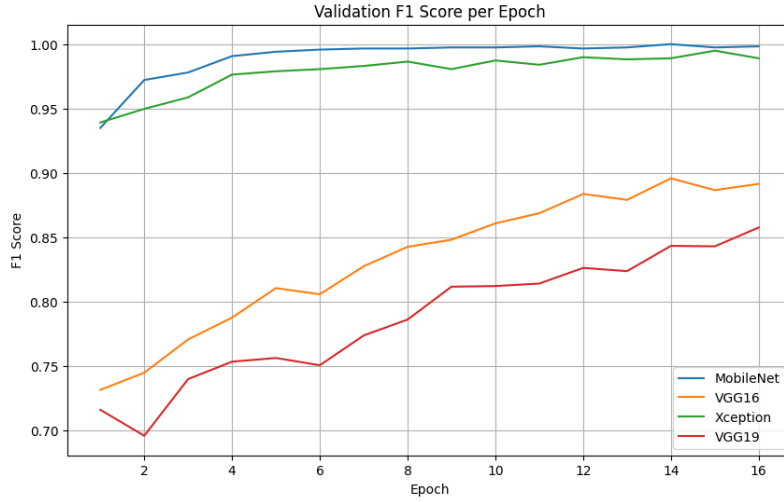
[Grafik 2: Precision Değerlerinin Modeller Arasında Dağılımı]

Precision metriği açısından MobileNet %99.83 ile en iyi sonucu vermiştir, yani modelin pozitif tahminlerinin neredeyse tamamı doğru olarak sınıflandırılmıştır. Xception modeli de yüksek kesinlik değerine ulaşmıştır. VGG16 ve VGG19 ise daha düşük precision değerleriyle, sınıflandırmada yanlış pozitif tahminlerin daha fazla olduğunu göstermektedir. Precision'ın yüksek olması, yanlış hastalık teşhislerinin önlenmesi açısından kritik önem taşır; özellikle tarımsal hastalık tespiti gibi uygulamalarda gereksiz müdahalelerin önüne geçer.



[Grafik 3: Recall Metriklerinin Karşılaştırılması]

Recall değerleri, gerçek hastalıkların ne kadar doğru tespit edildiğini gösterir. MobileNet ve Xception modellerinin recall oranları %99'un üzerinde seyretmiş, bu da hastalık belirtilerinin kaçırılmadığını ve güvenilir teşhis yapıldığını göstermektedir. Düşük recall, bazı hastalıklı yaprakların yanlış olarak sağlıklı sınıfa atanması anlamına gelir ki, bu durum tarımsal açıdan ciddi kayıplara neden olabilir. VGG16 ve VGG19 modelleri ise daha düşük recall değerlerine sahip olup, bu alanda istenilen başarıya ulaşamamıştır.



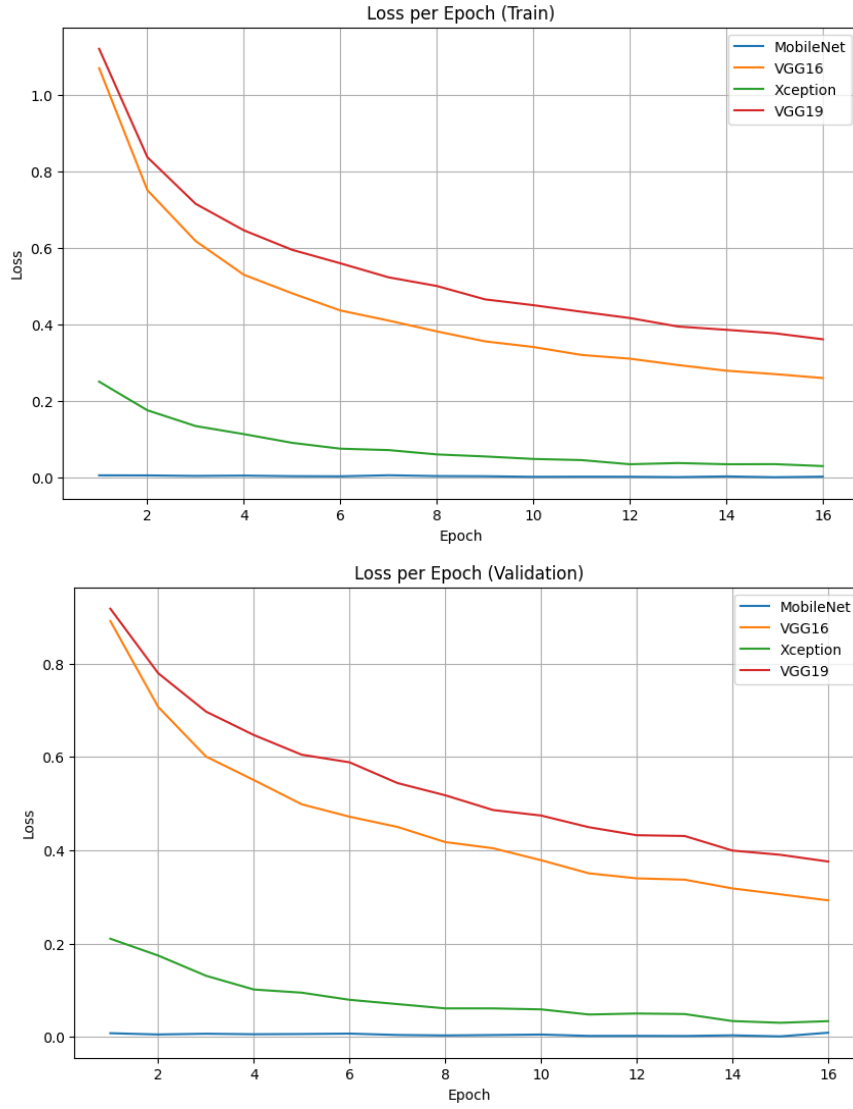
[Grafik 4: F1-Score ile Modellerin Genel Performansı]

F1-Score, bir modelin hem kesinlik (precision) hem de duyarlılık (recall) açısından dengeli performansını özetleyen önemli bir metriktir. Özellikle sınıflandırma problemlerinde, modelin doğru tahmin yaparken aynı zamanda yeterli sayıda doğru örneği yakalayabilmesini gösterir.

Bu açıdan bakıldığında, MobileNet %99.83 F1-score değeriyle oldukça başarılı bir sonuç sunmuştur. Hem yanlış pozitif hem de yanlış negatif tahmin oranlarının son derece düşük olması, modelin dengeli ve tutarlı bir performans sergilediğini göstermektedir.

Xception da %98.9'luk F1-score değeri ile güçlü bir performans göstermiştir. Karmaşık yapısıyla desen tanıma başarısı yüksek olan bu model, sınıflar arasında tutarlı sonuçlar vermiştir.

VGG16 ve VGG19 ise sırasıyla %89.15 ve %85.77 F1-score değerleriyle daha düşük bir başarı göstermiştir. Bu durum, her iki modelin özellikle sınıflar arasında daha fazla hata yaptığına ve genel genelleme kapasitesinin sınırlı olduğuna işaret etmektedir.



[Grafik 5 ve 6: Modellerin Eğitim ve Test Loss Değerleri]

Eğitim ve test loss grafiklerinin analizi, modellerin öğrenme süreçleri ve genelleme kapasiteleri hakkında önemli bilgiler sunar. Bu grafiklerde, her epoch sonunda hesaplanan kayıp değerlerinin zamanla nasıl değiştiği gösterilmiştir.

Grafiklere bakıldığında, MobileNet hariç tüm modellerin eğitim kayıp değerlerinde düzenli ve sürekli bir azalma gözlemlenmiştir. Bu durum, modellerin eğitim verisindeki örüntüleri başarılı şekilde öğrendiğini ve kayıp fonksiyonunu etkin bir biçimde minimize ettiğini göstermektedir. VGG16 ve VGG19 modellerinde, kayıp değerlerinin hızlı ve stabil bir şekilde düşmesi dikkat çekicidir. Bu, bu modellerin veri kümesindeki karmaşık özellikleri daha çabuk ve verimli yakaladığını işaret eder.

Test loss grafiklerinde ise MobileNet'in diğer modellere kıyasla daha düşük ve stabil kayıp değerleri sunduğu görülmektedir. Test kaybının düşmesi, modelin aşırı öğrenme (overfitting) yapmadığını ve yeni, görülmemiş veriler üzerinde de yüksek performans sergilediğini gösterir. Buna karşılık, VGG16 ve VGG19 modellerinde test kaybında zaman zaman dalgalanmalar ve daha yüksek değerler gözlemlenmiştir. Bu durum, bu modellerin sınırlı genelleme kabiliyeti ve overfitting riski taşıdığını düşündürmektedir.

Ayrıca, tüm modellerde eğitim ve test loss değerlerinin birbirine yakın olması, modelin dengeli bir öğrenme süreci geçirdiğine işaret eder. Büyük farklar genellikle overfitting veya underfitting gibi problemlerin göstergesidir.

Son olarak, loss grafiklerinin detaylı incelenmesi, hiperparametre ayarlarının (örneğin öğrenme hızı, batch büyüklüğü) uygun seçildiğini ve eğitim sürecinin kontrollü biçimde ilerlediğini göstermektedir. Bu grafikler, ileride model iyileştirmeleri ve yeni deneylerde yol gösterici olacaktır.

5.1. Eğitim Süresi ve Donanım Etkisi

Tüm modeller, Nvidia A100 GPU kullanılarak ortalama 24 dakika civarında eğitilmiştir; MobileNet için 24 dakika, Xception için 25 dakika gibi küçük farklılıklar gözlenmiştir. Bu yakın eğitim süreleri, yüksek hesaplama gücüne sahip donanımın (A100 GPU) sağladığı hız avantajını ortaya koymaktadır. Bu bağlamda, eğitim süresi artık donanım kapasitesiyle daha çok ilişkilendirilebilir hale gelmiş ve modeller arasındaki hesaplama maliyeti farkları donanım desteğiyle azaltılmıştır.

Buna rağmen, bazı modellerin mimari karmaşıklığı nedeniyle eğitim süresi farklarının hafifçe gözlemlenmesi mümkündür. MobilNet'in daha küçük ve optimize yapısı, daha hızlı eğitim imkanı sunarken, VGG16 ve VGG19 gibi daha derin ve parametre sayısı yüksek modeller, aynı donanımda bile biraz daha uzun sürebilmektedir. Ancak bu çalışmada farklar oldukça küçüktür.

5.2. Genel Değerlendirme ve İleriye Dönük Öneriler

Elde edilen sonuçlar, transfer öğrenme tabanlı derin öğrenme modellerinin pirinç yaprağı hastalıklarının sınıflandırılması için yüksek doğruluk ve güvenilirlik sunduğunu göstermektedir. Özellikle MobileNet modeli, hem yüksek performans hem de kısa eğitim süresi ile pratik uygulamalar için ideal bir çözüm olarak öne çıkmaktadır.

Bununla birlikte, sınıf bazlı performans farklılıkları ve bazı modellerin özellikle belirli hastalıklarda daha zayıf kalması, veri çeşitliliğinin ve model optimizasyonlarının önemini vurgulamaktadır. Gelecekte, daha geniş ve çeşitli veri kümeleriyle, farklı hiperparametre ayarlamaları ve hibrit model yaklaşımları denenebilir. Ayrıca ensembel yöntemlerle modellerin bir arada kullanılması, daha dengeli ve yüksek performanslı sınıflandırmaların elde edilmesine katkı sağlayabilir.

Sonuç olarak, bu çalışma tarımsal görüntü işleme alanında yapay zekâ uygulamalarının etkinliğini destekler nitelikte olup, hastalık tespiti konusunda gelişmiş modellerin tarım sektörüne doğrudan fayda sağlayabileceğini göstermektedir.

Proje Kaynak Kodu

Bu çalışmada kullanılan tüm modellerin eğitim süreçleri, veri ön işleme adımları ve değerlendirme analizleri, aşağıdaki GitHub deposunda paylaşılmıştır:

GitHub Bağlantısı: https://github.com/Ilhanemreadak/rice_leaves_daisies

KAYNAKLAR

- [1] A. Kaur, V. Kukreja, P. Tiwari, M. Manwal ve R. Sharma, “An Efficient Deep Learning-based VGG19 Approach for Rice Leaf Disease Classification,” in Proc. IEEE I2CT, 2024, ss. 1–6.
- [2] N. Thai-Nghe, N. T. Tri ve N. H. Hoa, “Deep Learning for Rice Leaf Disease Detection in Smart Agriculture,” in Artificial Intelligence in Data and Big Data Processing, Springer, 2022, ss. 659–670.
- [3] S. Hassan, A. K. Maji, M. Jasinski ve E. Jasinska, “Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach,” Electronics, c. 10, no. 12, s. 1388, 2021.
- [4] S. Sharma ve K. Guleria, “Pre-trained Deep Neural Network Based Features for Plant Disease Detection,” Agriculture, c. 13, no. 5, s. 936, 2023.
- [5] N. Sankalana, “Rice Leaf Disease Image Dataset,” Kaggle, 2023. [Çevrimiçi]. Erişim: <https://www.kaggle.com/datasets/nirmalsankalana/rice-leaf-disease-image>