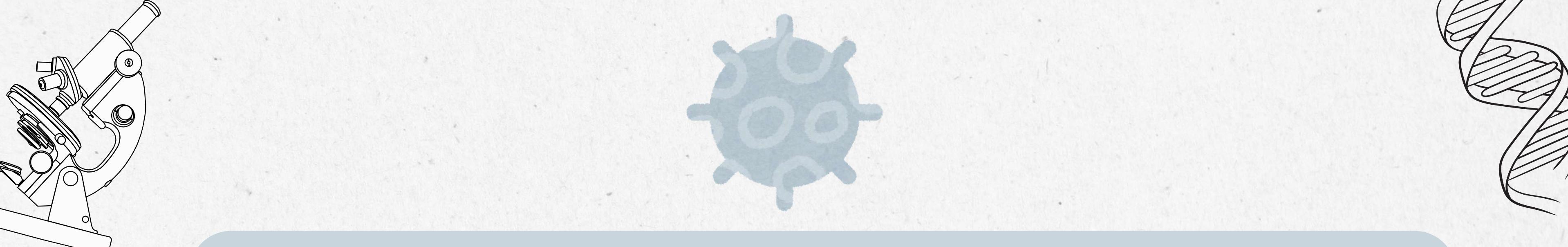


Final projesi Makine Öğrenmesi

Bitki Yaprığı Hastalığı Sınıflandırma

Sümeyye AKGÜL-2321021005



Problem Tanımı ve Motivasyon

Tarımsal Zorluk

Bitki hastalıklarının erken teşhisini, küresel gıda güvenliği ve ürün verimliliği için kritik öneme sahiptir. Geleneksel yöntemler uzman değerlendirmesine bağlı olup zaman alıcı ve maliyetlidir.

Otomatik hastalık teşhis sistemleri, çiftçilere hızlı ve güvenilir çözümler sunarak hasadı koruma potansiyeline sahiptir.

Proje Hedefi

Yaprak görüntülerinden hastalık türünü otomatik olarak belirleyen bir sınıflandırma modeli geliştirmek.

- Makine öğrenmesi ve derin öğrenme yöntemlerini karşılaştırma
- Gerçek zamanlı teşhis için uygulanabilir bir çözüm sunma
- Tarım teknolojilerinde yapay zeka kullanımını yaygınlaştırma



Problem Türü ve Proje Hedefi

Bu çalışma, denetimli öğrenme tabanlı çok sınıflı bir görüntü sınıflandırma problemidir.

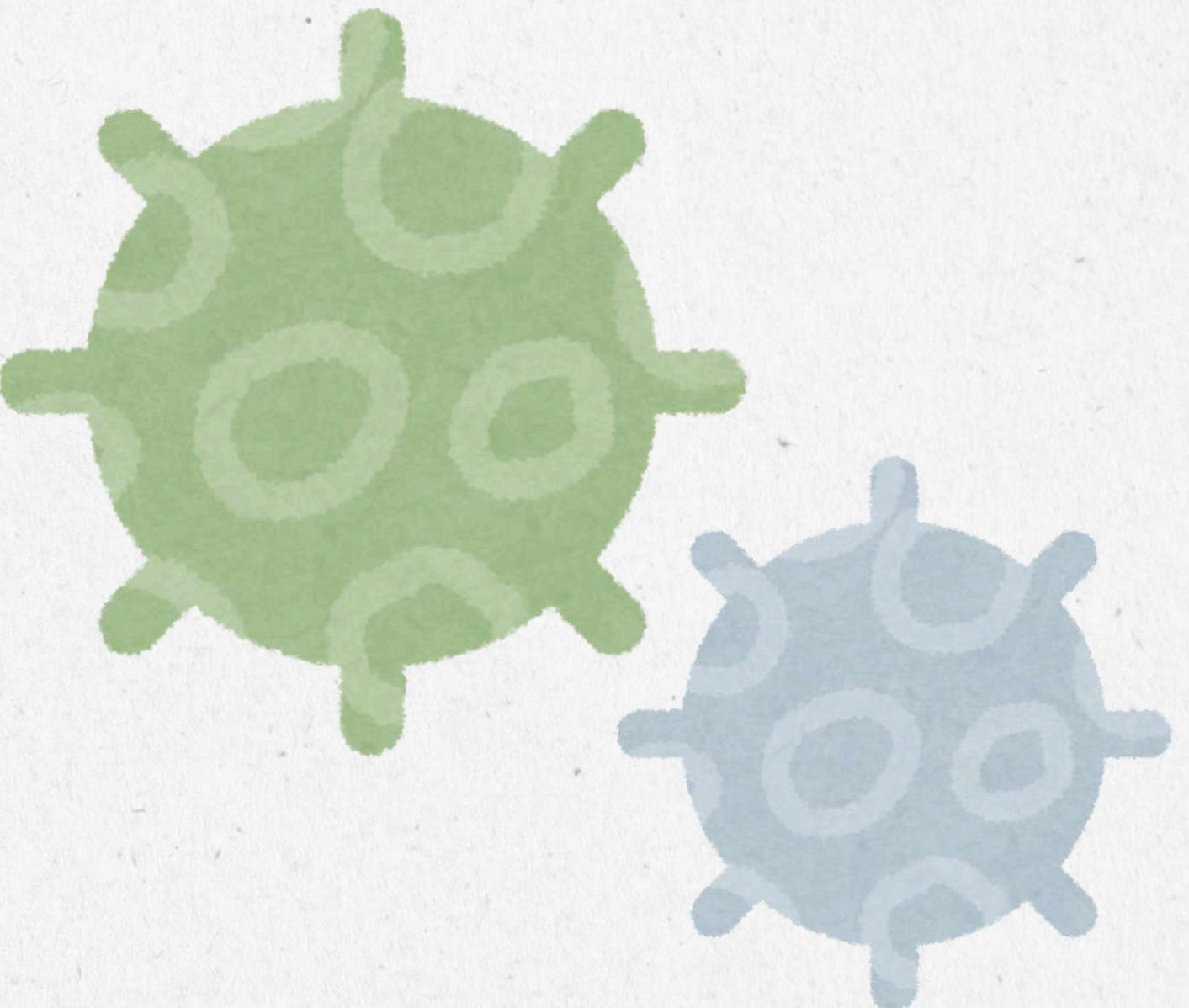
Amaç, yaprak görüntülerinden yaprağın:

- Sağlıklı
- Lekeli hastalık
- Mantar hastalığı

sınıflarından hangisine ait olduğunu otomatik olarak belirlemektir.

Bu kapsamda hedeflenenler:

- Random Forest, CNN ve MobileNetV2 yöntemlerini karşılaştırmak
- En yüksek doğruluk ve güvenilirliğe ulaşmak
- Gerçek dünyada uygulanabilir bir hastalık teşhis sistemi geliştirmek



Makine Öğrenmesi Yöntemleri

Yöntemler

01.

- Random Forest (Klasik Makine Öğrenmesi)
- Basit CNN (Convolutional Neural Network)
- MobileNetV2 (Transfer Learning – Derin Öğrenme)

Neden Bu Yöntemler

02.

- Random Forest**
- Açıklanabilir ve hızlı çalışır
 - Küçük–orta ölçekli veri setlerinde stabil sonuç verir
- Basit CNN**
- Görüntüden otomatik özellik çıkarma kapasitesine sahiptir
 - Görsel örüntülerin daha iyi öğrenir
- MobileNetV2**
- Daha önce büyük veri setlerinde eğitilmiş bir modeldir
 - Az veriyle yüksek başarı sağlar
 - Gerçek zamanlı uygulamalara uygundur

İmplementasyon

03.

- Random Forest**
- HOG + renk özellikleri çıkarıldı
 - Sklearn RandomForestClassifier ile eğitildi

CNN

 - 128×128 giriş boyutunda bir CNN mimarisi kuruldu
 - Conv2D → MaxPooling → Dense katmanları kullanıldı

MobileNetV2

 - Pretrained ImageNet ağı temel alındı
 - Üst katmanlar çıkarılıp yeni sınıflandırıcı eklendi
 - MobileNetV2'nin büyük kısmı donduruldu, son katmanları fine-tuning ile yeniden eğitildi

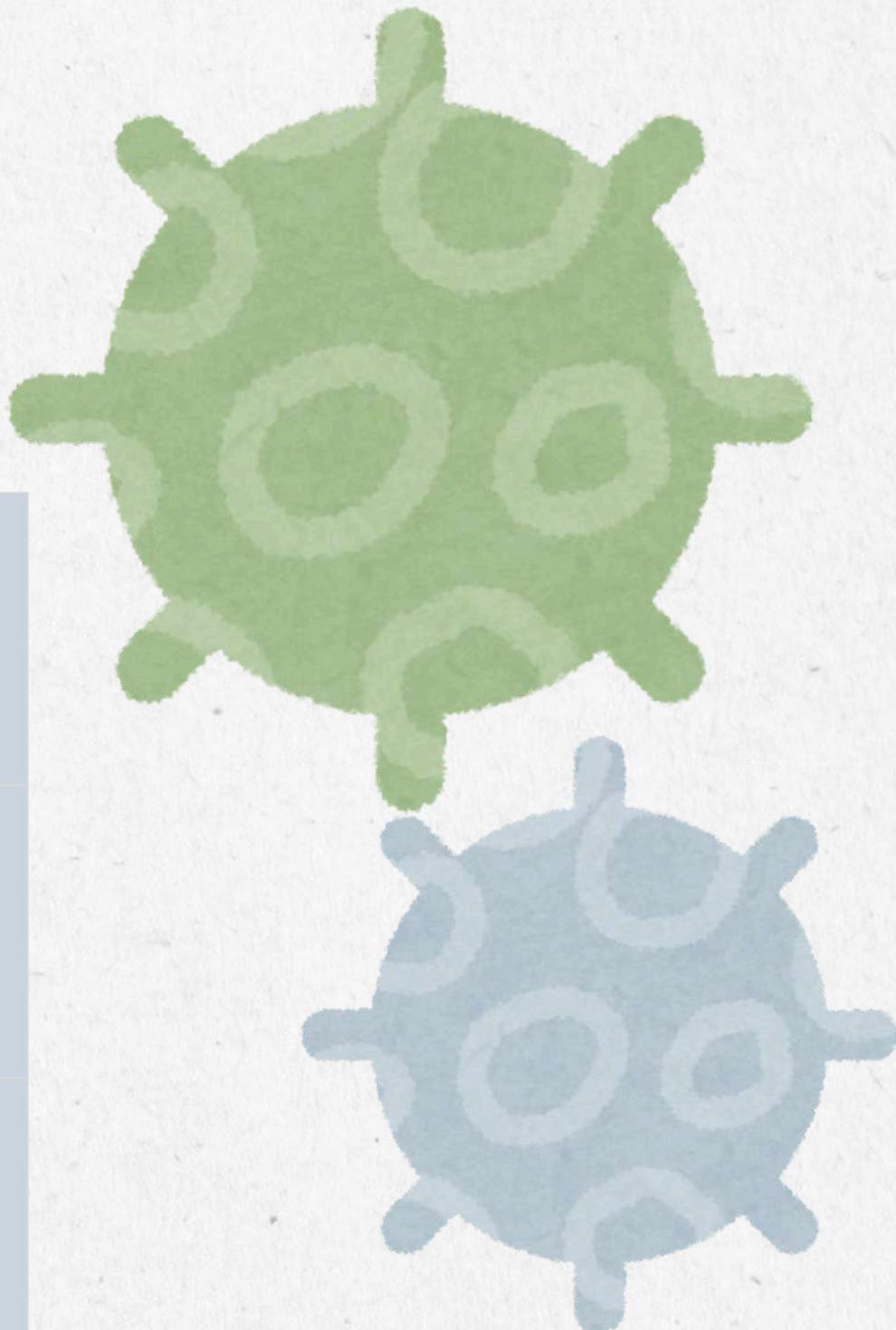
Parametreler

04.

- Random Forest**
- n_estimators
 - max_depth
- CNN / MobileNetV2**
- Learning rate
 - Batch size
 - Epoch sayısı
 - Dropout oranı
 - Optimizer (Adam, SGD vb.)
 - MobileNet'te base katmanın açılıp kapanması

Yöntemlerin Karşılaştırılması

Model	Avantajlar	Dezavantajlar
Random Forest	Hızlı çalışır, açıklanabilir sonuçlar üretir, küçük veri setlerinde stabil performans sağlar	Görüntüde derin ve karmaşık görsel özellikleri öğrenme kapasitesi sınırlıdır
CNN	Görsel örüntüleri ve detayları otomatik olarak öğrenir, görüntü tabanlı problemlerde yüksek performans gösterir	Eğitim süresi daha uzundur ve daha fazla veri gerektirebilir
MobileNetV2	Önceden eğitilmiş yapısı sayesinde daha az veriyle yüksek doğruluk sağlar, güçlü genelleme yeteneğine sahiptir	Daha yüksek donanım ihtiyacı olabilir, model karmaşıktır



Deneysel Çalışmalar

01

Veri Kümesi

- Bitki yaprağı görüntülerinden oluşturulmuş bir veri seti kullanılmıştır.
- 3 sınıf her sınıfın 500 örnek:
- Sağlıklı Yaprak – Lekeli Hastalık – Mantar Hastalığı
- veri seti %80 eğitim – %20 test olacak şekilde bölünmüştür.

02

Ön İşleme Adımları

- Görüntüler RGB formatta okunmuş ve normalize edilmiştir (0–1 aralığı).
- Random Forest için görüntüler gri tona çevrildi
- HOG + renk özelliklerini çıkarıldı
- CNN ve MobileNetV2 için direkt görüntü tabanlı öğrenme yapıldı.

03

Değerlendirme Metrikleri

- Accuracy: Doğru tahmin edilen örneklerin tüm örneklerle oranı
- Precision: Doğru pozitif / (Doğru pozitif + Yanlış pozitif)
- Recall: Doğru pozitif / (Doğru pozitif + Yanlış negatif)
- F1-Score: Precision ve Recall'un harmonik ortalaması
- Confusion Matrix: Her sınıf için doğru–yanlış sınıflandırmalarının dağılımı

Deneysel Çalışmalar

04

Eğitim ve Hiper-parametreler

- Görseller normalize edilip modele verildi
- CNN & MobileNetV2 için:
- Learning Rate: 0.0001
- Epoch: 30
- Batch Size: 16
- Optimizer: Adam
- Random Forest için:
- n_estimators = 150

05

Veri Çoğaltma Adımları

- Veri Çoğaltma
- Rotation
 - Zoom
 - Horizontal Flip
 - Amaç: Veri çeşitliliğini artırmak ve modelin genelleme performansını iyileştirmek.

06

Hiper-parametre optimizasyonu

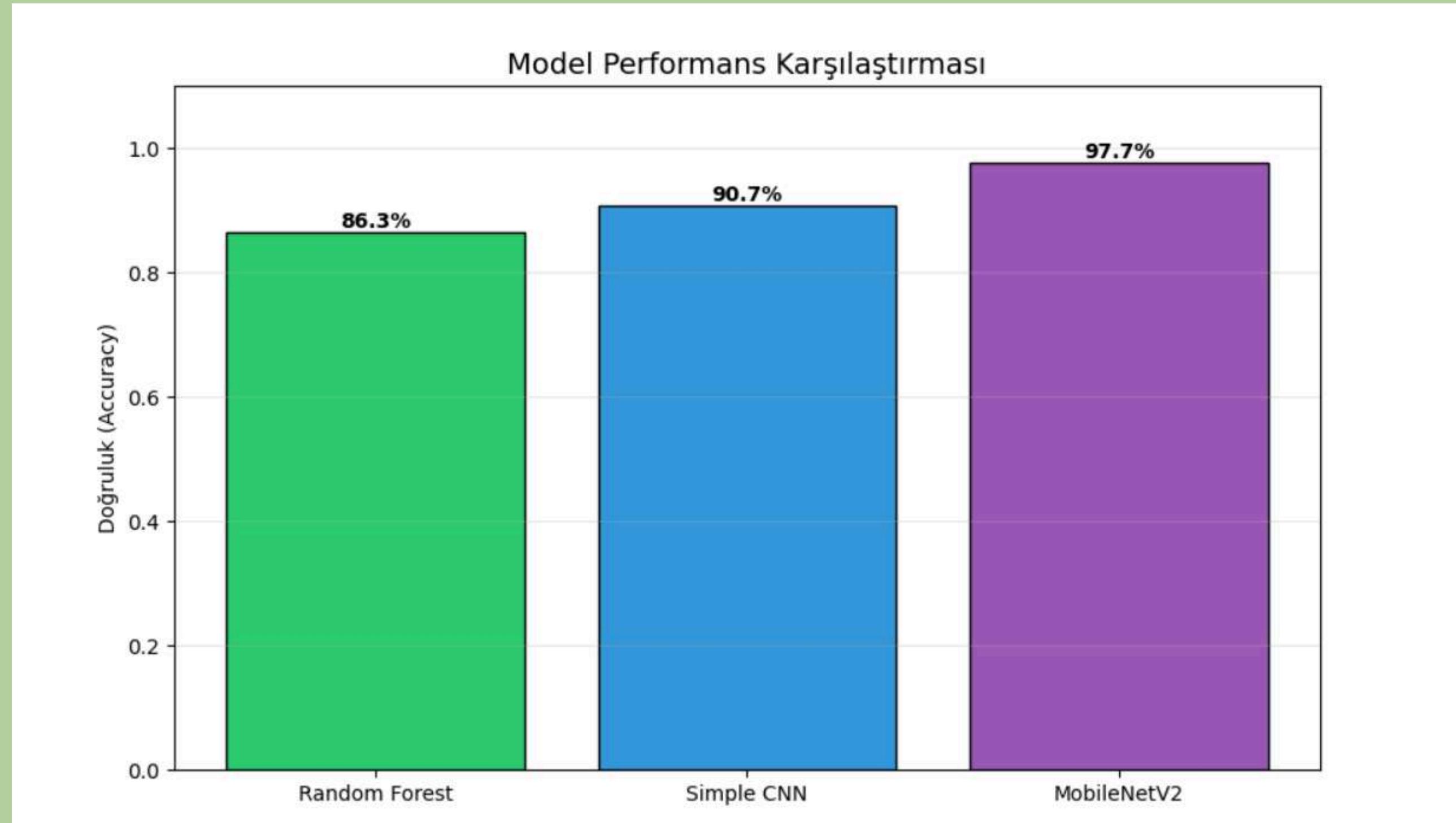
- Temel seviyede tuning yapıldı.
- Modeller için en verimli değerler deneysel olarak belirlenmiştir.
- Ayarlanabilir parametreler:
 - Learning rate
 - Epoch
 - Batch size
 - Random Forest: n_estimators, max_depth

Sonuçlar

Model Performans ve Çalışma Süresi Karşılaştırma Tablosu

Model	Doğruluk (Accuracy)	Hata Payı	Çalışma Süresi (Tahmini)	Karmaşıklık
Random Forest	%86.00	%14.00	Çok Kısa (Saniyeler)	Düşük
Simple CNN	%89.67	%10.33	Kısa/Orta (Dakikalar)	Orta
MobileNetV2	%99.00	%1.00	Uzun (Dakikalar)	Yüksek

Sonuçlar



Sonuçların Değerlendirmesi

Analizler, bitki hastalıkları teşhisinde derin öğrenme modellerinin klasik yöntemlere göre çok daha yüksek bir hassasiyet sunduğunu kanıtlamıştır. Özellikle aktarımılı öğrenme stratejisinin kullanılması, karmaşık görsel verilerin işlenmesinde manuel yöntemlerden daha güvenilir ve hızlı bir çözüm sağladığını doğrulamaktadır.

01

En İyi Performans: MobileNetV2 (Transfer Learning) modeli, hazır ağırlıkları ve gelişmiş mimarisi sayesinde en yüksek doğruluğu vererek projenin en başarılı modeli olmuştur.

02

Derin Öğrenme Üstünlüğü: CNN tabanlı modeller, Random Forest gibi klasik yöntemlerin aksine görsel özellikleri (lezyon, doku, damar) otomatik öğrenerek daha yüksek başarı sağlamıştır.

03

Transfer Learning Etkisi: Önceden eğitilmiş modelin son katmanlarının Fine-tuning ile güncellenmesi, kısıtlı veriyle bile çok hızlı ve hassas sonuçlar alınmasını sağlamıştır.

Sonuçların Değerlendirilmesi

Modelin eğitim sürecinde uygulanan ileri teknikler, sistemin sadece mevcut veriyi öğrenmesini değil, yeni verilere karşı yüksek uyum yeteneği kazanmasını sağlamıştır. Veri artırma yöntemleriyle desteklenen bu süreç, dinamik öğrenme mekanizmaları ve aşırı öğrenmeyi engelleyen denetimlerle birleşerek, modelin en zorlu görsel koşullarda bile kararlı ve hatasız sonuçlar üretmesine olanak tanımıştir.

04

Ezberlemeyi Önleme: Early Stopping ve Dropout kullanımı, modellerin eğitim verisini ezberlemesini (overfitting) engellemiştir ve yeni görüntülere uyum yeteneğini artırmıştır.

05

Veri Çeşitlendirme: Uygulanan Data Augmentation (döndürme, yakınlaştırma vb.) işlemleri, modelin farklı ışık ve açılardaki yaprak görsellerine karşı dayanıklı olmasını sağlamıştır.

06

Dinamik Öğrenme: ReduceLROnPlateau kullanımı, eğitim tıkanıklıklarında öğrenme hızını otomatik düşürerek modelin en iyi sonuç noktasına hatasız ulaşmasını sağlamıştır.

Teşekkürler !