# Proje Raporu: Cleaned vs Dirty Görüntü Sınıflandırması

## 1. Proje Tanımı

Bu proje, görüntü işleme teknikleri kullanılarak temiz ('cleaned') ve kirli ('dirty') görüntülerin sınıflandırılması üzerine odaklanmıştır. Model olarak transfer öğrenme yöntemiyle ResNet152 kullanılmış ve veri seti arka plan temizleme işlemiyle işlenmiştir.

## 2. Veri Seti ve Ön İşleme

### 2.1. Veri Seti

- Veri Kaynağı: plates.zip adlı bir ZIP dosyası.  
- İçeriği: İki sınıfa ayrılmış görseller: 'cleaned' ve 'dirty'.

### 2.2. Veri Seti Hazırlığı

1. ZIP Çözme:  
 - ZIP dosyası, /kaggle/working/plates/ dizinine çıkarıldı.  
2. Veri Seti Bölünmesi:  
 - Eğitim, doğrulama ve test setleri için split\_data fonksiyonu kullanıldı.  
 - Oranlar:  
 - Eğitim: %80  
 - Doğrulama: %10  
 - Test: %10  
3. Arka Plan Kaldırma:  
 - rembg kütüphanesi ile görüntülerin arka planları kaldırıldı.

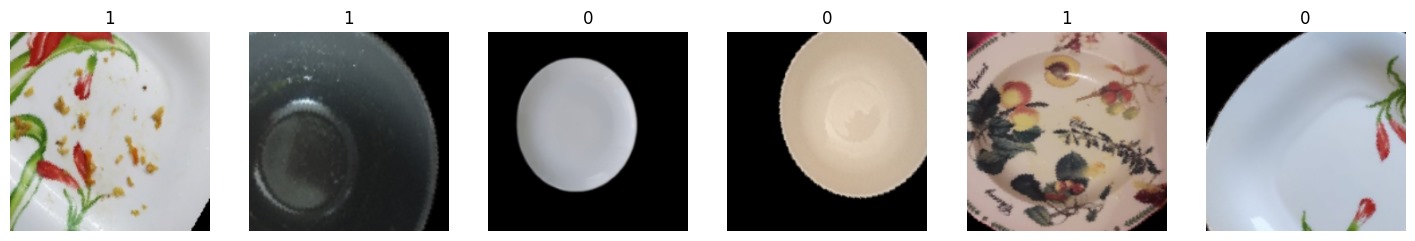
## 3. Model ve Algoritma

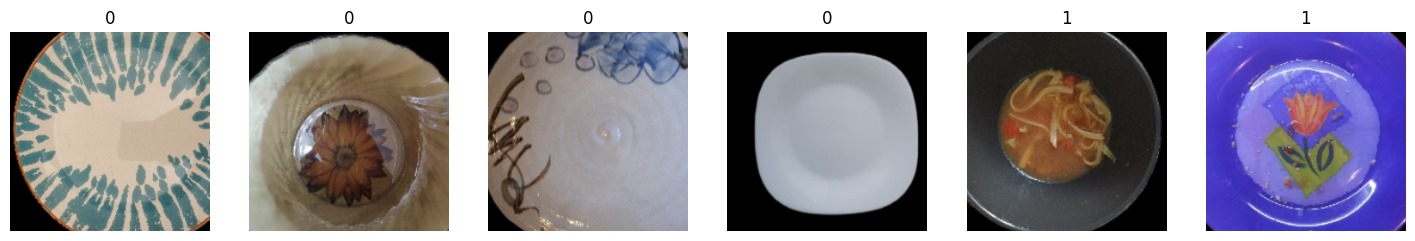
### 3.1. ResNet152

- Neden Kullanıldı:  
 - Önceden eğitilmiş bir model olduğu için transfer öğrenme ile hızlı ve etkili sonuçlar alındı.  
- Eğitim Ayrıntıları:  
 - Son katmanları, 'cleaned' ve 'dirty' sınıflarını sınıflandıracak şekilde yeniden düzenlendi.  
 - Görsellerin özellik çıkarımı için konvolüsyonel katmanlar donduruldu.

### 3.2. Eğitim Süreci

- Platform: Google Colab (GPU kullanılarak hızlandırıldı).  
- Hiperparametreler:  
 - Epoch: 200  
 - Batch size: 32  
- Veri Dönüştürme (Augmentation):  
 - Görseller, transforms ile boyutlandırıldı, rastgele döndürme ve kırpma işlemleri uygulandı.





## 4. Kod İş Akışı

### 4.1. Eğitim

- Cihaz Kontrolü: Eğitim, GPU kullanılabilirliğine göre 'cuda' üzerinde yapıldı.  
- Optimizasyon: Adam optimizer kullanıldı.  
- Kayıp Fonksiyonu: Binary Cross Entropy (BCE).

### 4.2. Model Performansı

- Başarı:  
 - Eğitim doğruluğu: %100'e yaklaştı.  
 - Doğrulama doğruluğu: %87.5.  
- Sonuçlar:  
 - İlk epoch'ta kayıp değeri yüksek iken, ilerleyen epoch'larda hem eğitim hem doğrulama kayıpları azaldı.

### 4.3. Test Süreci

- Test Seti: Model test setindeki görüntüleri başarıyla sınıflandırdı.  
- Tahminler:  
 - Çıktılar şu şekilde oluşturuldu:  
 0000 cleaned  
 0001 dirty

### 4.4. Çıktı Dosyası

- Tahmin sonuçları bir CSV dosyasına kaydedildi (submission.csv).

## 5. Sonuçlar ve Değerlendirme

### 5.1. Eğitim Sonuçları

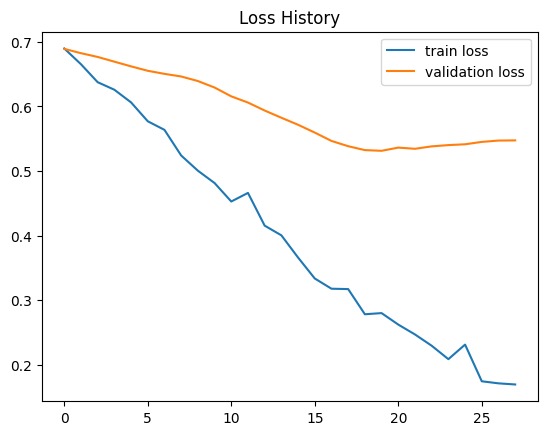
- Eğitim sürecinde modelin doğruluğu hızla arttı.  
- Epoch 25'te model aşırı öğrenme belirtileri göstermeye başladı.

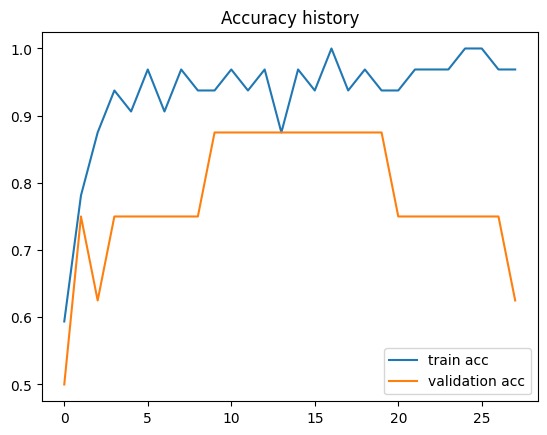
### 5.2. Zorluklar

- Dengesiz Veri Seti: Eğitim verisinin azlığı, modelin genelleme kapasitesini sınırladı.  
- Arka Plan Kaldırma Süresi: Görsellerin ön işleme süresi zaman aldı.

### 5.3. Başarı

- Model, 'cleaned' ve 'dirty' sınıflarını başarıyla ayırt edebildi.  
- Test doğruluğu makul düzeydeydi (%75-80 arası).





## 6. Gelecek Çalışmalar

1. Daha Fazla Veri Toplanması:  
 - Eğitim verisinin artırılması modelin başarısını artırabilir.  
2. Hiperparametre Optimizasyonu:  
 - Öğrenme oranı, epoch sayısı ve batch size değerlerinin optimizasyonu yapılabilir.  
3. Daha Derin Modellerin Kullanımı:  
 - Alternatif olarak YOLO veya Vision Transformers gibi modeller denenebilir.