

1-) Verdiğiniz taskta belirtilen dataset kullanıldı. Dataset içeriğindeki ultrason fotoğraflarını kullandım. Fotoğrafları sınıflandırmak için tirads özneliğini kullandım.

3-) Elde edilen ultrason fotoğrafları tirads değerlerine göre malign(kötü huylu) ve malign olmayan olarak ikiye ayrıldı. Fotoğrafları sınıflandırmak için xml dosyaları ayrıştırıldı ve tirads değerleri okundu. Elde edilen fotoğraflara fazlalık kısımdan kurtulmak için kesme işlemi uygulandı. Etiketlere göre görselleştirme yapılarak verinin dağılımı gözlemlendi verinin dağılımına göre aşırı öğrenme olabileceği ön görüldü ve eşitleme işlemi yapıldı. Veriler eğitim test ve kontrol verileri olarak bölündü. Veri sayısı yeterli olmayacağı düşünülerek veri arttırma işlemi uygulandı. Verileri işlemek için kullanılacak VGG16 modeli import edildi ve gerekli eklemeler yapıldı. Çıkış kökünde aktivasyon fonksiyonu olarak 'sigmoid' tercih edildi. Compile aşamasında kayıp değeri fonksiyonu olarak 'binary-crossentropy' kullanıldı, optimizasyon fonksiyonu olarak 'adam', değerlendirme metriği olarak kesinlik tercih edildi.

Eğitim sırasında alınan çıktılarda sistemin 0.84 kesinlik verdiği gözlemlenmesine rağmen kontrol setimizle tahmin yaptığımızda ve metrikleri incelediğimizde 0.54 kesinlik değerini gözlemliyoruz. Bu sonuca ulaşmamıza neden olan etkenler veri setinde tirads skoruna göre yaptığımız etiketlemenin dengesizliğe neden olması ve aşırı öğrenmenin önüne geçmek için yapılan eksiltme işleminin verideki bazı tirads skorlarının tamamen kaybolmasına neden olmuş olabilir.

Daha iyi bir model sonucu elde etmek için verilere uygulanan ön işleme adımlarında değişikliğe gidilebilir. Modele diğer öznelilikler bir şekilde eklenebilir ve ya başka bir modelle bu öznelilikler işlenip mevcut modelle birleştirilebilir. Tirads skorlarının tahmini üzerine bir model geliştirilebilir.

4-) Bu problemin çözümünde eğitim ve kontrol aşamalarında verinin miktarının yetersiz olması önemli bir etken oldu. Ön işleme adımlarında kesme işlemi uygularken tiroid bölgesine daha yakın kesikler atılması başarıyı arttırabilecek bir işlem olmasına rağmen nasıl uygulayacağımı bulamadığım için dış kısımları kesmekle yetindim.

Bu çalışmada VGG-16 modeli kullanıldı bunun yerine daha basit ve her katmanına müdahale edilebilir olması için kendi CNN mimarimizi oluşturabiliriz. Görsel sınıflandırmada aktif olarak kullanılan AlexNet ve GoogleNet te kullanılabilir güçlü modellerdir. Bunlar denenip en yüksek değerleri veren model üzerinden işlemimiz devam edebiliriz.

5-) Eğittiğimiz modelin kontrol aşamasında kesinlik değerleri düştü bu düşüşün nedeninin tirads değeri üzerinden atadığımız malign değerinin sınıflandırma için yetersiz olduğu düşünüyorum. Dengesizliği yaratan temel şeyin sınıf eşitsizliğini gidermek adına yapılan işlemlerin verinin dağılımını etkilediğini ve bazı görsellerin öğrenilmesini engellediğini düşünüyorum. Bunun önüne geçebilmek için sistemi tirads değerlerini sınıflandırmayı öğrenecek şekilde tasarlar ve sonrasında çıktı üzerinden "2" ve "3" değerlerine sağlıklı diğer sınıflara hasta etiketini atayan bir sistem yazar isek daha doğru bir sistem tasarlamış oluruz.