

Türk İşaret Dili ile Derin Öğrenme Temelli Algılama

İşitme kaybı, bireylerin sözel iletişim kurma yetisini önemli ölçüde etkileyen bir sağlık sorunu olup, bireylerin günlük yaşam aktivitelerini ve toplumsal katılımını kısıtlamaktadır. Dünya Sağlık Örgütü'ne (WHO) göre dünya nüfusunun %5'inden fazlası, yani yaklaşık 430 milyon kişi (34 milyonu çocuk olmak üzere) işitme kaybı nedeniyle rehabilitasyona ihtiyaç duymaktadır. 2050 yılına kadar bu rakamın 700 milyona ulaşacağı, yani dünya genelinde her 10 kişiden 1'inin işitme kaybı yaşayacağı öngörülmektedir (WHO, 2025). İşitme kaybının ortaya çıkmasında genetik faktörler, yaşlanma, doğum sırasında oluşan komplikasyonlar, çevresel etmenler ve sigara kullanımı gibi çeşitli nedenler etkili olmaktadır. İşitme kaybına sahip bireyler, iletişimde zorluklar yaşamakta ve bu durum, sosyal hayatlarına ve profesyonel yaşamlarına doğrudan etki etmektedir.

İşitme engelli bireyler için işaret dili çok büyük önem arz etmektedir. İşaret dili farklı ülkelerde farklı olmasıyla beraber anlaşılmayı da zorlaştırmaktadır. İşaret dili tanıma, bilgisayarla görme alanında önemli ve zorlu bir çalışma alanı olarak kabul edilir. Bunun başlıca nedenleri arasında işaret dillerinin karmaşık görsel ve jest-mimik yapısının olması ve otomatik tanıma ile ilgili az sayıda veri tabanı ve çalışmanın olmasıdır (Rezende ve ark., 2021).

İşaret dillerinden biri olan Türk İşaret Dili genellikle harfler ve kelimeler Türk alfabesini kullanır ancak TİD dilbilgisi olarak farklı bir yapıya sahiptir. Bu çalışma, Türk İşaret Dili ifadelerini anlama üzerine yoğunlaşmaktadır. Bu konu hakkında bazı önde gelen çalışmalar öne çıkmaktadır.

(Torun & Karacı, 2024) , çalışmalarında 41 basit kelime ile orijinal bir veri seti oluşturmuştur. Bu veri seti, 35'i dinamik 6'sı statik olacak şekildedir. Oluşturulan veri seti GRU, LSTM ve Bi-LSTM modelleri ile eğitilmiş ve hibrit mimarilere sahip olan CNN+GRU, GRU+LSTM, GRU+Bi-LSTM modelleri ile değerlendirilmiştir. Sonuçlara bakıldığında, hibrit modellerin klasik derin öğrenme modellerine kıyasla daha iyi performans gösterdiği gözlemlenmiştir.

(PACAL & ALAFTEKİN, 2023), literatürde ilk kez rakam temelli olarak yeni bir veri seti önermiştir. CNN yaklaşımları kullanılarak katmanlı bir yapı ve hiyerarşik bir öğrenme modeli oluşturulmuştur. Bu çalışmada CNN mimarileri olan VGG, ResNet, MobileNet,

DenseNet ve EfficientNet kullanılmıştır. Veri setinin özel olarak hazırlandığı ve az sayıda olduğu için modelin kabiliyetini artırmak adına veri artırma tekniklerine başvurulmuştur. Modellerin eğitilmesi ve test edilmesi için 1.500 görüntüden oluşan veri seti belli oranlarda ayrılmış olup %70 eğitim, %15 doğrulama ve %15 test için kullanılmıştır. Sınıflandırma metrikleri ve deneysel sonuçlara bakıldığında modellerin hepsi %90'ın üzerinde doğruluk oranlarına ulaşmış ve başarılı sonuçlar vermiştir.

(Alaftekin et al., 2024) diğer projelere kıyasla arka plan fark etmeksizin ve ön işleme gerek meksizin Türk İşaret Dili'ni anlamak üzerine yeni bir metod sunmuştur. Bu yöntem YOLOv4-CSP optimizasyonu ile geliştirilmiştir. YOLO ile beraber gerçek zamanlı tespit yapılması hedeflenmiştir. Bu model, %98.95 doğruluk oranına ulaşmış ve mevcut algoritmalara kıyasla daha başarılı bir performans sergilemiştir. Mevcut çalışmalarla karşılaştırıldığında, YOLO tabanlı sistemlerin diğer yöntemlere kıyasla daha hızlı ve doğru sonuçlar verdiği görülmektedir. Örneğin, Sincan & Keleş (2020), Türk İşaret Dili için CNN ve LSTM tabanlı bir model geliştirerek %96.11 doğruluk oranına ulaşmıştır. Rivera-Acosta ve arkadaşları (2021), YOLO ve LSTM tabanlı bir model ile ASL kelime tanıma sisteminde %98.07 doğruluk elde etmiştir. Bu çalışmada geliştirilen YOLOv4-CSP modeli, Türk İşaret Dili'nde %98.95 doğruluk oranına ulaşarak önceki çalışmalardan daha başarılı sonuçlar elde etmiştir.

(BANKUR & KAYA, 2022) TİD harfleriyle özgün bir veri seti oluşturmuş ve sonuç olarak toplamda 10.000 adet resimden oluşan özgün bir veri seti elde etmiştir. Hazırlanan özgün veri seti, TİD harflerinin hızlı bir şekilde tanınması amacıyla tek aşamalı nesne tespiti modelleri olan SSD MobileNet, EfficientNet ve YOLO modelleri ile eğitime tabi tutulmuştur. Yakın eğitim süreçleri ile YOLOv5 modelinin diğer nesne tespit modellerine oranla hızlı ve daha doğru sonuçlar verdiği gözlemlenmiştir. Böylelikle YOLOv5 modeli uzun süreli eğitime sokulmuş ve sonucunda TİD harflerini tanınması %92.3 oranında başarı ile gerçekleşmiştir.

Türk İşaret Dili tanıma alanındaki çalışmalar son yıllarda ilerleme kaydetmiş olup, yapay zeka ve derin öğrenme modelleri bu alanda etkin bir şekilde kullanılmaya başlanmıştır. Gelecekte, daha geniş kapsamda ve gerçek zamanlı uygulamalar geliştirilerek işitme engelli bireylerin sosyal hayata entegrasyonunun artırılması hedeflenmektedir.