



T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ ANABİLİM DALI

İŞARET DİLİNİN DERİN ÖĞRENME
YÖNTEMLERİYLE METİN HALİNE
DÖNÜŞTÜRÜLMESİ

YÜKSEK LİSANS TEZİ

MELEK ECE ÇİFTÇİ

Danışman

Doç. Dr. Ramazan TEKİN

Ocak-2024

BATMAN

Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Melek ECE ÇİFTÇİ tarafından hazırlanan “İşaret Dilinin Derin Öğrenme Yöntemleriyle Metin Haline Dönüştürülmesi” adlı tez çalışması 24/01/2024 tarihinde aşağıdaki jüri tarafından oy birliği ile Batman Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgi Teknolojileri Anabilim Dalı’nda YÜKSEK LİSANS olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Doç. Dr. Melih KUNCAN

.....

Danışman

Doç. Dr. Ramazan TEKİN

.....

Üye

Doç. Dr. Yılmaz KAYA

.....

Yukarıdaki sonucu onaylarım.

Dr. Öğr. Üyesi Ömer Murat ÖTER
Lisansüstü Eğitim Enstitüsü Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

İmza

Melek ECE ÇİFTÇİ

Tarih:

ÖZET

YÜKSEK LİSANS TEZİ

İŞARET DİLİNİN DERİN ÖĞRENME YÖNTEMLERİYLE METİN HALİNE DÖNÜŞTÜRÜLMESİ

Melek ECE ÇİFTÇİ

Batman Üniversitesi Lisansüstü Eğitim Enstitüsü
Bilgi Teknolojileri Anabilim Dalı

Danışman: Doç. Dr. Ramazan TEKİN

2024, 75 Sayfa

Jüri

Doç. Dr. Ramazan TEKİN

Doç. Dr. Yılmaz KAYA

Doç. Dr. Melih KUNCAN (Siirt Üniversitesi)

İşitme ve duyma engelli bireyler, çevreleriyle iletişim kurarken ciddi zorluklar yaşamaktadırlar. Bu tez, gerçek zamanlı işaret dili tanıma sistemi ile bu soruna çözüm getirmeyi hedeflemektedir. Çalışma kapsamında, literatürde sıkça rastlanan 17 kelimedenden oluşan bir veri seti, 5 denek tarafından özel olarak hazırlanmıştır. MediaPipe kitaplığındaki insan poz tahmin sistemi kullanılarak, kelimelerin anahtar noktaları başarıyla çıkarılmıştır. Bu anahtar noktalar, derin öğrenme modellerinden LSTM mimarisi ile sınıflandırılmıştır. Çalışma kapsamında kişi bağımlı ve kişi bağımsız iki ayrı uygulama yapılarak test veri seti üzerinde sırasıyla %94,71 ve %87,65 oranında doğruluk elde edilmiştir. Bu önerilen sistemin, işitme ve konuşma engelli bireylerin günlük iletişimlerini iyileştirmek adına literatürdeki diğer çalışmalara kıyasla yeterli seviyede performansa sahip bir çözüm ortaya konulduğu değerlendirilmektedir.

Anahtar Kelimeler: Anahtar nokta, Derin Öğrenme, İşaret Dili, LSTM

ABSTRACT

MS THESIS

CONVERTING SIGN LANGUAGE INTO TEXT USING DEEP LEARNING METHODS

Melek ECE ÇİFTÇİ

**Batman University Graduate Education Institute
Department Of Information Technologies**

Advisor: Doç. Dr. Ramazan TEKİN

2024, 75 Pages

Jury

Doç. Dr. Ramazan TEKİN

Doç. Dr. Yılmaz KAYA

Doç. Dr. Melih KUNCAN (Siirt Üniversitesi)

Individuals with hearing and speech impairments experience serious difficulties in communicating with their surroundings. This thesis aims to address this issue with a real-time sign language recognition system. Within the scope of the study, a dataset consisting of 17 frequently encountered words in the literature was specially prepared by 5 subjects. The key points of the words were successfully extracted using the human pose estimation system in the MediaPipe library. These key points were classified using the LSTM architecture of deep learning models, and the obtained results provided 99% accuracy. This proposed system offers an effective solution to improve the daily communication of individuals with hearing and speech impairments.

Keywords: Deep Learning, Key Point , LSTM , Sign Language,

ÖNSÖZ

Tez sürecimdeki özverili rehberliği, sabrı ve bilgeliği ile her aşamada bana destek olan ve bu zorlu yolculukta yanımda olduğunu hissettiren sayın hocam Doç. Dr. Ramazan TEKİN' e içten teşekkürlerimi sunuyorum. Ayrıca, tez için oluşturduğum veri setinde bana yardımcı olan Veysel ÇİFTÇİ, Elif ECE AKIN, Hamdullah AKIN ve Fidan AKIN' a katkılarından dolayı teşekkür ederim.

Melek ECE ÇİFTÇİ
BATMAN-2024



İÇİNDEKİLER

1. GİRİŞ	1
1.1. İşaret Dili.....	1
1.2. Türk İşaret Dili Tarihi	5
1.3. Türk İşaret Dili Yapısı	7
1.3.1. İşaretler	8
1.3.2. Parmak alfabesi	10
1.3.3. Ağızla söyleme (Dudak hareketleri)	11
1.3.4. Jest ve mimikler	11
1.4. Araştırmanın Amacı	12
2. LİTERATÜR TARAMASI	14
2.1. Elektronik Eldiven Tabanlı Sistemler	14
2.2. Sensör Tabanlı Çalışmalar	16
2.3. Video Kamera Tabanlı Çalışmalar	18
2.4. İnsan Poz Tahmin Tabanlı Çalışmalar	19
3. MATERYAL VE YÖNTEM.....	22
3.1. Veri seti	22
3.1. Önerilen Mimari.....	25
3.2. Yapay Zekâ ve Tarihçesi.....	26
3.3. Yapay Sinir Ağları	27
3.3.1. İleri beslemeli yapay sinir ağları	31
3.3.2. Geri beslemeli yapay sinir ağları	31
3.4. Derin Öğrenme	32
3.5. Derin Öğrenme Mimarileri	34
3.5.1. Konvolüsyonel sinir ağları (CNN).....	34
3.5.2. Tekrarlayan sinir ağları (RNN).....	35
3.5.3. Uzun kısa süreli hafıza ağları (LSTM)	38
3.6. MediaPipe Kütüphanesi	40
3.7. Model Performans Değerlendirme Metrikleri.....	44
4. ARAŞTIRMA BULGULARI.....	46
4.1. Kişi Bağımlı Sonuçlar.....	46
4.2. Kişi Bağımsız Sonuçlar	49
5. TARTIŞMA	53
6. SONUÇ	58
KAYNAKLAR	60

SİMGELER VE KISALTMALAR

Kısaltmalar

BILSTM	Çift Yönlü Uzun Kısa Süreli Bellek
CNN	Konvolüsyonel Sinir Ağlarıyla
DBN	Derin İnanç Ağları
DSÖ	Dünya Sağlık Örgütü
GRU	Kapılı Tekrarlayan Birimler
KNN	En Yakın Komşuluk
LSTM	Uzun Kısa Süreli Hafıza Ağları
MLP	Çok Katmanlı Algılayıcılar
MOPGRU	Kapılı Tekrarlayan Birim
RNN	Tekrarlayan Sinir Ağları
TDK	Türk Dil Kurumu
TİD	Türk İşaret Dili
YSA	Yapay Sinir Ağları

1. GİRİŞ

1.1.İşaret Dili

İletişim verici ve alıcı arasındaki mesaj aktarımı olarak adlandırılır. Kişiler arasındaki anlam aktarımı olarak tanımlanabilir. İletişim, bireylerin duygularını, tutumlarını ve davranışlarını, konuşarak veya sessizce, yazılı olarak ya da elektronik platformlarda başkalarına aktarmalarını içeren kapsamlı bir olgudur. İletişim sözlü ve sözsüz olarak üzere iki ana grupta incelenebilir (Tutar vd., 2017). Sözlü iletişim, kelimelerin kullanılması yoluyla gerçekleşirken, sözsüz iletişim sözcüklerin kullanılmadığı, ancak beden dili, jestler, mimikler gibi sözsüz davranış unsurların önemli olduğu iletişim türüdür. Karaca (Karaca,2018) iletişim türlerinin sayısının üçe çıkarılabileceği belirtilmektedir. Bu durumda iletişim, temel olarak bireylerin birbirleriyle etkileşimde bulunmalarını sağlayan çeşitli yöntemlere ayrılabilir. Bu yöntemler sözlü, sözsüz ve yazılı iletişim olarak üç ana kategoride toplanır. Yazılı iletişim, metin, mektup, e-posta gibi yazılı biçimde gerçekleşen iletişimi içerir. Karaca iletişim türlerine elektronik iletişim de dâhil edilerek toplamda dört ayrı iletişim türü olabileceğini de vurgulamıştır. Elektronik iletişim, teknolojik araçlar kullanılarak gerçekleştirilen iletişimi ifade etmektedir (Karaca, 2018).

İletişim, Türk Dil Kurumu sözlüğüne göre, bireyler arasında duygu, düşünce ya da bilgilerin her türlü yolla birilerine iletilmesi olarak tanımlanmaktadır. Yani, iletişim, insanların birbirleriyle haberleşmeleri, duygusal ve zihinsel içerikleri paylaşmaları, bilgi alışverişinde bulunmaları sürecini ifade eder (TDK,2022). İnsanlar arasındaki en çok kullanılan etkileşim türü, genellikle işitme ve konuşmaya dayalı olarak gerçekleşen sözlü iletişimdir. Ancak, konuşma yeteneği veya işitme kapasitesiyle ilgili kısıtlamalar yaşayan bireyler, bu yaygın iletişim biçimini tam anlamıyla kullanmakta zorluklarla karşılaşabilirler. Bu durumda, farklı iletişim stratejileri ve yöntemleri, bireylerin etkileşimde bulunmalarını ve iletişim kurmalarını destekleyebilir. Doğuştan gelen ya da çeşitli sebeplerle sonradan gelişen işitme kaybı, bu bireylerin günlük iletişimlerinde zorluk yaşamalarına sebep olabilir. İşitme kaybının şiddetine bağlı olarak, bireylere işitme yeteneği kazandırmak amacıyla kohlear implant veya işitme cihazları kullanılabilir. İşitme kaybı yaşayan bireylerin işitme yeteneklerini artırmak için kullandıkları işitme cihazı ya da çoğunlukla işitme siniri hasarı veya işitme kaybının neden olduğu diğer sorunlardan dolayı geleneksel işitme cihazlarının etkili olamadığı

durumlarda kullanılan kohlear implant bu bireyler tarafından, genellikle sesli ve sözel iletişim yöntemlerini tercih etmektedirler. Bu cihazlar, işitme engeli yaşayan bireylerin çevreleriyle daha etkili bir şekilde iletişim kurmalarını sağlayarak, onlara işitsel dünyayı daha yakından deneyimleme fırsatı sunmaktadır. Bununla birlikte, geç tanı alıp ya da geç cihaz takılan ve konuşma becerilerini tam anlamıyla geliştiremeyen işitme kayıplı bireyler, iletişimde işaret dilini kullanma eğilimindedir (Ertan ve Arıcı, 2022).

Günümüzde sayıları azımsanmayacak kadar fazla olan işitme engelli bireylerin bu iletişim yöntemlerini kullanamadıkları ya da sınırlı şekilde kullandıkları görülmektedir. Yazarak anlaşma, işaret dili, dudak okuma gibi yöntemler, işitme engeli olan kişilerin iletişim ihtiyaçlarını karşılamak adına sözlü iletişimden farklı stratejilerdir. Bu yöntemler, işitme engeli olan bireylere, dünya ile etkileşimde bulunma ve bilgi alışverişi yapma konusunda alternatif yollar sunar. Yani, bu yöntemler işitme engelli bireylerin iletişimdeki katılımlarını artırmak ve onlara daha etkili bir şekilde iletişim kurma fırsatı sağlamak amacıyla kullanılır (Karaca ve Bayır, 2018). İşaret dili işitme engelli bireylerin başvurdukları en yaygın iletişim tekniğidir. Bu teknik; el hareketleri, mimikler, jestler ve tamamlayıcı olarak kullandıkları ağızla tekrar etmeden oluşur. Sözsüz iletişimde çoğunlukla kullanılan jest ve hareketler işaret dilini oluşturan temel unsurlardır (Tutar vd., 2014). İşitme engelli bireylerin işaret dilini kendi aralarında kolaylıkla kullandıkları görülse de sağlıklı bireylerin işaret diline hâkim olmamaları veya işaret dilini bilmemelerinden dolayı toplumla iletişimleri sınırlı kalmaktadır. Toplumun işitme engelli bireylerle sağlıklı iletişim kurabilmesi için işaret dilinin öğrenilmesi gerekmektedir (Karaca ve Bayır, 2018). İşaret dili, Türk Dil Kurumu sözlüğüne göre, duyma engelli bireylerin birbirleriyle diyalogları sırasında jest ve mimiklerini kullanarak meydana getirdikleri görselleştirilmiş bir dil olarak nitelendirilmektedir. İşitme engelliler arasında yaygın olarak kullanılan bu dil, konuşma ve işitme yeteneğinden yoksun olan bireyler arasında etkili bir iletişim aracı olarak öne çıkar. İşaret dili, jest, mimik ve görsel semboller aracılığıyla duygu, düşünce ve bilgi iletimini mümkün kılar, böylece işitme engelli bireyler arasında anlamlı bir etkileşim sağlar (TDK,2022). İşaret dilleri, Arapça, Çince, İngilizce, İspanyolca, Türkçe gibi konuşma dillerle karşılaştırılabilecek özelliklere sahiptir. Ancak, işaret dillerini diğer dillerden ayıran en önemli özellik, konuşulan diller olmamalarıdır; bunun yerine, işaretlenen dillerdir. Bu diller, genellikle el ve vücut hareketleriyle birlikte eş zamanlı olarak kullanılan yüz ifadeleriyle gösterilirler. Bu biçimde kendini anlatmaya "işaretleme" denir. Günlük hayatta kullanılan dillerde olduğu gibi, kullanıcılar arasında doğal olarak gelişen dil özelliği işaret dillerinde de

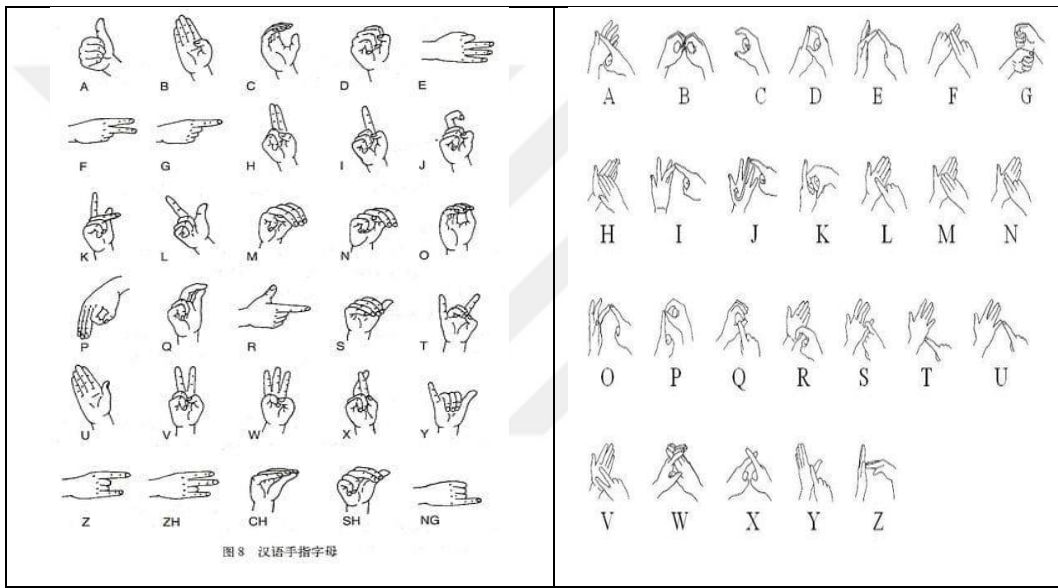
bulunmaktadır. Bundan ötürü, işaret dilleri de doğal dil özelliklerini taşırlar (Eryiğit, 2017). İşaret dilleri, tarihsel olarak günlük hayatta kullandığımız diller gibi eski olmalarına rağmen, bu dillerle ilgili bilimsel araştırmaların daha sonradan başlamış olması nedeniyle daha geç keşfedilmiş olduğu düşünülmektedir (Arık, 2013).

Dünya genelindeki işitme engellilerin genellikle tek bir işaret dilini kullandığı yaygın bir yanlışlıktır. Ancak gerçekte, dünya çapında 100'den fazla işaret dili mevcuttur. Dolayısıyla, sadece belirli bir işaret dili bilen bir bireyin, dünya genelindeki tüm işitme engelli bireylerle iletişim kurabileceği düşünülmemelidir. Ülkelerin kendi benzersiz işaret dillerine sahip olmaları, işitme engelliler arasında dil bariyerlerine neden olmaktadır. Bu durum, uluslararası düzeyde işaret dilleri arasında bir standardizasyonun eksikliğini vurgular ve iletişimi zorlaştırır. İşaret dilleri arasındaki çeviri zorluğunu aşmak niyetiyle toplumlar arası bir işaret dili oluşturulmuştur. Ancak, ülkeler kendi sınırları içinde kullandıkları ve benzersiz olduğunu düşündükleri işaret dilini kullanmayı sürdürmüşlerdir. Çağımızda, toplumlar arası iletişimde kullanılan işaret dili özellikle çeşitli işaret dilleri arasında tercüme yapılrken aracı bir dil olarak kullanımına başvurulmaktadır. Bu yaklaşım, işitme engellilerin dünya genelinde daha geniş bir iletişim ağına sahip olmalarını desteklemek amacıyla geliştirilmiştir. Ancak, her bir ülkenin kendi kültürel ve dil yapısını koruma isteği, ulusal işaret dillerinin varlığını sürdürmesine neden olmaktadır. Bu durum, uluslararası düzeyde işitme engelliler arasında daha etkili iletişim kurma çabalarını karmaşık hale getirmektedir (Ertan ve Arıcı, 2022).



Şekil 1.1. Uluslararası işaret dili alfabesi (Birk vd., 1997)

Görsel iletişim türleri arasında yer alan işaret dili işitmede sıkıntı yaşayan bireylerin jest ve mimiklerini kullanmasından oluşmaktadır. Dilin temelini oluşturan yapısal özellikler işaret dilinde de bulunmaktadır ve konuşma ve yazma dilinden farklı olarak işaret dilinin kendine has yapısal özellikleri vardır. Evrensel dil olarak kabul edilmeyen işaret dilleri bulundukları ülkelerin özelliklerine göre farklılıklar göstermektedirler (Selda ve Vasif, 2006). Diğer bir ifadeyle ülkelerin kullandıkları işaret dilleri farklılık göstermektedir ve ülkelerin işaret dillerini oluşturan kurallar ülkeye has özellikler taşımaktadır.



Şekil 1.2. Çin işaret dili ve İngiliz işaret dili alfabesi (“AI Media”, 2023)

Ülkelerin farklı bölgelerinde bile işaret dilleri farklılıklar taşıyabilmektedir. Konuşma dillerinde kullanılan dil bilgisi kurallarıyla işaret dilinin dil bilgisi kuralları aynı değildir. İşaret dilleri konuşma dillerinden türememiştir ve erken yaşlarda işaret dili öğrenmek engelli bireyler için önemlidir (Karaca ve Bayır, 2018). İşitme engelli bireylerin işaret dili öğrenme süreci genellikle aileleri veya çevreleriyle doğrudan ilişkilidir. Ancak, işitme engelli bireyler arasında bu dili öğrenme şansına erişimi olmayanlar da vardır. Özellikle, işitme engelli bireylerin kendi ülkelerine özgü işaret dillerini bilmemeleri ve iletişim kurabilmek adına ev içinde geliştirdikleri ev işaretlerini kullanmaları, bu bireylerin iletişim ihtiyaçlarını karşılamak adına yaratıcı çözümler bulduklarını gösterir. Ancak, bu durum aynı zamanda işitme engelli bireylerin işaret diline erişimde yaşadıkları

eşitsizliklere de işaret eder. İşaret dili eğitimi ve erişim, işitme engelli bireylerin daha geniş bir iletişim ağına sahip olmalarını sağlamak adına önemlidir (Arık, 2016). Ev işaret dili, temelde ileri dilbilgisi özelliklerine sahip olmayan bir iletişim yöntemi olarak tanımlanır. Bu tip işaret dili, genellikle işitme yeteneği olan ebeveynlerle işitme engelli çocuklar arasında kurulan iletişimi ifade eder. İletişimde, özel semboller ve işaretler (pseudo signs) kullanılarak oluşturulan bir dil meydana gelir. Bu, işitme engelli bireylerin ve ailelerinin, anlamak ve anlaşılmak adına kendilerine özgü bir iletişim yöntemi geliştirmelerinin bir yolu olarak ortaya çıkar (Dikyuva vd., 2015). İşaret tabanlı iletişim genellikle duyma engelli kişilerin yaşadığı toplumlarda kendiliğinden ortaya çıkan ve evrilen iletişim araçlarıdır. Bu diller, konuşulan dillerin sosyal ve zihinsel fonksiyonlarını başarıyla yerine getirebilir. İşaret dilleri, doğal bir dil evrimi süreci içinde, işitme engelli bireylerin kendi topluluklarında iletişim kurma ihtiyacından türemiş olabilir. Bu dillerin öğrenilmesi, çocuklar için konuşulan dillerin öğrenilmesine benzer bir süreç izleyebilir. İşitme engelli çocuklar, işaret dillerini, aileleri veya toplulukları tarafından kullanıldığı için doğal bir şekilde öğrenebilirler. Her ne kadar bireyler farklı dil yeteneklerine sahip olsalar da işaret dillerinin karmaşık bir dil sistemi oluşturabilmesi için bir topluluk içinde paylaşılan kurallara ve normlara ihtiyaç duyar (Ertan ve Arıcı, 2022).

1.2. Türk İşaret Dili Tarihi

Dünya üzerinde kullanılan en köklü işaret tabanlı iletişim şekillerinden biri de Türk İşaret Dili (TİD) olarak bilinmektedir (Zeshan, 2003). Osmanlı döneminde işitme engelli bireylere iş imkânı sunulduğu ve Topkapı Sarayı'nda işitme engelliler arasında işaret dilinin resmi kullanımına dair belgelerin bulunduğu konusundaki tespitler, Türk İşaret Dili'nin kökenlerine ışık tutmaktadır. Ayrıca, işitme engellilerin Osmanlı'da kamusal hizmetlerde ve mahkemelerde görev alarak topluma katkıda bulunduklarına dair bilgiler, işaret dilinin sadece saray ortamında değil, aynı zamanda büyük şehirlerde ve toplumun çeşitli kesimlerinde yaygın olarak kullanıldığını göstermektedir. Bu, Türk İşaret Dili'nin sadece bir elit kesim tarafından değil, geniş bir topluluk içinde benimsenip geliştirildiğini gösteren önemli bir noktadır. Tarihi belgeler ve hükümdarların da işaret dili bildiğine dair bilgiler, Türk İşaret Dili'nin tarih boyunca nasıl evrildiği konusunda önemli ipuçları sunmaktadır. Bu veriler, işaret dilinin sadece iletişimde değil, aynı zamanda sosyal ve kamu hizmetlerinde de etkili bir şekilde kullanıldığını göstererek, işitme engelli bireylerin toplumsal entegrasyonunda etkili bir rol oynadığını ortaya

koymaktadır. Osmanlı döneminde bulunan önemli şehirlerde bu bireyler, devlet dairelerinde bulunmuş ve davalarda görev almıştır. Mahkemelerde görev alan ve işaret dili bilen çalışanlarla yine burada görev alan işitme engelli birey arasında işaret dili kullanılarak iletişim sağlanmaktaydı. Bu durum, işitme engelli bireylerin mahkeme ortamlarında etkili bir şekilde hizmet verebilmeleri ve iletişim kurabilmeleri için toplumsal bir uyum çabası olarak değerlendirilebilir. İşaret dilinin bu tür ortamlarda kullanılması, işitme engelli bireylerin adil bir yargı sürecine katılımını destekleyerek toplumsal eşitlik ve erişilebilirlik ilkelerine uygun bir yaklaşımı yansıtmaktadır (Ertan ve Arıcı, 2022). Osmanlı İmparatorluğu'nda, 1500-1700 yılları arasında, mahkemelerde gizliliği korumak amacıyla görevlendirilen bir işitme engelli bireyler topluluğu bulunmaktaydı. İşitme engellilerin gizlilik gerektiren görevlerde sıkça yer aldığı belirlenmiştir (Çakır vd., 2013). Yapılan araştırmalar göz önünde bulundurulduğunda Osmanlı mahkemelerinde kompleks karmaşık bir işaret dili kullanımının olduğu ve bu dili bilenlerin bir statü sahibi olduğu anlaşılmaktadır. Osmanlının son yarıyılı ile Cumhuriyet ten sonra işitme engelli bireylere eğitiminin başlangıcına kadarlık dönemde işaret dili eğitimiyle ilgili yeterli kaynak bilgi bulunmamaktadır. Bu sebepten ötürü Osmanlı döneminde kullanılan işaret dili ile modern TİD arasında kesin bir bağlantı kurulamamaktadır. Fakat modern TİD ile Avrupa işaret dili arasında bir ilişki olmadığı kesin olarak söylenebilmektedir. Bu yargıda modern TİD ile Osmanlı döneminde en yakın ilişki içinde olduğu Fransa devletinin kullandığı işaret dilinin birbirine hiç benzememesinden yola çıkarak oluşturulmuştur. Osmanlı Devleti'nde kullanılan işaret dili ile modern Türk İşaret Dili (TİD) arasındaki benzerlikler ve sürekliliklerin belgelenmesi, dünya üzerinde kullanılan ilk işaret dilinin Türk İşaret Dili olduğunu kanıtlamak açısından önemli olabilir (Zeshan, 2003). 1953 yılında işaret dili işitme engelli bireylerin konuşma yetilerini geciktirdiği düşüncesiyle işitme engelli bireyler ve dilsiz okullarında yasaklanmıştır. Bunun sonucunda, işaret dili ile ilgili gerekli kaynakların oluşturulması ve yaygınlaşması belli bir süre engellenmiştir (Akalin, 2013). Milli eğitim bakanlığınca çıkarılan kanun ile okullarda işaret dili kullanımı öğrenci yararına olacağı düşüncesiyle yasaklanmıştır. İşaret dili yerine sözel eğitim kullanılarak erken yaşta işitme engelli bireyler olan öğrencilerin konuşması amaçlanmıştır. 2005 yılında bu tekniğin yanlış olduğu fark edilince bu yasak kaldırılmıştır (Kın, 2019). Türk işaret dili ile ilgili ilk kanuni düzenlemeler 2005 tarihli ve 5378 sayılı kanunla yapılmıştır. 5378 sayılı kanun ile işitme engellilerin eğitim ve iletişimleri devlet yükümlülüğü altına alınarak engelliler için oluşturulan işaret dili sistemi onların hayatları için dönüm noktası

olmuştur (Erten ve Arıcı, 2022). Ülkemizde başarılı sağlık politikaları neticesinde gerek yeni doğan işitme taramaları artışı gerekse erken tanı sonrası yapılan müdahaleler sonucu iyileşmeler, işitsel sözel eğitime katılım oranını artırmıştır. Çeşitli sebeplerden ötürü konuşma yetisi kazanamayan çocuklar işitme engelliler okuluna yönlendirilirken, konuşabilenler ise kaynaştırma eğitimi ile eğitimlerini sürdürmektedirler. Bu tutum sebebiyle özel eğitime gereksinim duyan öğrenci sayısı azalırken, eğitimlerini başarılı şekilde bitiren öğrenci sayısında artış meydana gelmiştir (İlkbaşaran, 2016). TİD, duyabilen bireylerin işitme engelliler için oluşturdukları yapay bir dil değil; kendi benzersiz yapısına ve dilbilgisi kurallarına sahip, işitme engelli bireylerin kendi ihtiyaçlarına dayalı ana dilleridir. İçerdiği öğelerle doğal bir dilin gerekliliklerini karşılamaktadır (Demir, 2010).

1.3. Türk İşaret Dili Yapısı

Türk İşaret Dili (TİD), Türkçe konuşulan bir bölgede gelişmiş olmasına rağmen, dil yapısı açısından farklı kurallara sahip olan ayrı bir dildir. Genel bir ifadeyle, Türk İşaret Dili (TİD), Türkçe kelimeleri karşılayan özgün işaretlere sahiptir. Fakat bu işaretlerin Türkçe kelimelerden oluştuğunu düşündürse de işitme engelli topluluğunun kendine özgü bir iletişim sistemi bulunmaktadır. Her dil gibi, Türk İşaret Dili'nin de kendi sözcük dağarcığı ve dilbilgisi kuralları vardır. Türk İşaret dili, Türkçe cümledeki kelimelerin işaretlerle ifade edildiği bir sistem gibi görünse de bu sadece yüzeyde bir benzerlik olabilir. İşaret dilleri, sözcüklerin yanı sıra mimikler, el ve yüz hareketleri gibi görsel unsurları içeren tamamen farklı bir dil sistemine sahiptir. Bu nedenle, Türk İşaret Dili'nin kendi benzersiz yapısına ve dilbilgisi kurallarına dikkat etmek önemlidir. TİD, Türkçede olduğu gibi kelime köklerine ya da temel kelimelerine eklemelerin, takıların, eklerin veya diğer morfolojik unsurların eklenmesi yoluyla yeni kelimelerin oluşturulduğu bir dil değildir, bu yönüyle farklı bir dilbilgisel yapısına sahiptir. Türkçe 'deki bazı eklerin TİD' deki karşılıkları, farklı işaretlerle veya temel olarak mimiklerle ifade edilir. Örneğin, "gelmemek" ifadesi "gelmek + değil" şeklinde değil, kendi özgün yapısında ifade edilir. Bir diğer önemli farklılık, Türkçe 'deki zaman kiplerinin tümü TİD' de kullanılmamaktadır. TİD, geniş (şimdiki), geçmiş ve gelecek zamana ait kavramlara odaklanır. Bu, işaret dilinin zaman kavramlarını nasıl ifade ettiği konusunda kendi benzersiz yaklaşımını yansıtır. Türk İşaret Dili, Türkçe ile benzer kelimelere sahip olabilir, ancak bu kelimelerin ifade ediliş biçimi, dilbilgisi kuralları ve zaman kavramları

gibi temel özellikleri açısından farklı bir dil olarak kabul edilir. Ayrıca, TİD'de Türkçe 'de olmayan kavramların da işaretlerle ifade edilmesi mümkündür. Fakat, parmak alfabesinin TİD'de kullanım oranlarına dair spesifik bir çalışma olmamasına rağmen, genel olarak sohbetlerde %5-%6 oranında kullanıldığı ifade edilmiştir. Bu durum, TİD'nin kendine özgü ifade biçimleri ve dilbilgisel yapılarının önemli bir parçasını oluşturmaktadır (Erten ve Arıcı, 2022). İşaret dilleri genellikle belirli hareketlerin harfler ya da kelimelerle eşleştirildiği sistemlerdir. Bu sistemler, kullanım durumlarına göre genellikle üç ana gruba ayrılır: parmak alfabesi, kelimeye dayalı işaret dili ve konuşan kişinin dudak hareketlerini ve yüz ifadelerini izleyerek, söylenen kelimeleri veya cümleleri anlamaya çalıştığı dudak okumadan oluşan işaret dilleridir (Selda ve Vasif, 2006). TİD'i oluşturan 4 önemli unsur bulunmaktadır. Bunlar; işaretler, parmak alfabesi, söyleme (dudak okuma) jest ve mimiklerdir (Karaca ve Bayır, 2018).

1.3.1. İşaretler

Kelime tabanlı işaret dilinde işaretin anlatımında vücudun yanı sıra jest ve mimiklerin kullanımı da oldukça önemlidir. İşaret dili kullanan bireyler kelime tabanlı işaret dilinin yanı sıra dudak okuma yöntemiyle de çevredekilerle iletişim kurmaktadır (Ezel, 2018). İşaretler kelimeleri sembolize etmektedir (Bayrak, 2009). Kelimelerin karşılığında bir işaret bulunmaktadır. İşaret ile kelime arasın da her zaman bağ kurulamayabilir. Bundan işaretin anlamsız hareketlerden oluştuğu fikri oluşmamalıdır. Örneğin Otobüs sözcüğüne bakıldığında onu temsile eden hareket anlamsız görünse de hareket dikkatle incelendiğinde ayakta duran yolcunun tutma yerini temsil ettiği görülmektedir. İşaret dili işaretleri 3 şekilde bulunabilmektedir. Yapıla işaret ile kelime arasında doğrudan bir bağ var ve işareten kelime tahmin edilebiliyorsa bu işarete şeffaf, işareten kelime tahmininin yapılamadığı fakat kelime ve işaret bilindikten sonra anlamlı bağın kurulduğu işaretler yarı şeffaf olarak isimlendirilir. İşaret ile kelime arasında tahminin yapılamadığı ve anlamlı bağın kurulamadığı işaretler ise şeffaf olmayan işaretler denir (Karaca ve Bayır, 2018).



Şekil 1.3. 'Altı' işaretinin gösterimi (TİD, 2015)

Şekil1.1'de Altı işaretinin gösterimi TİD' göre yapılmıştır. Şekil1.3'e göre sağ el baş parmak açık diğer dört parmak ise kıvrıktır. Kelime tabanlı işaretler durağan ve durağan olmayan olarak sınıflandırılmaktadırlar. İşaret yapılırken el hareketlerinde herhangi bir değişiklik olmuyorsa bu durağan, el hareketlerinin ilk başladığı an ile bitiminde farklılık varsa bu da durağan olmayan işaret olarak isimlendirilir (Karaca ve Bayır, 2018).



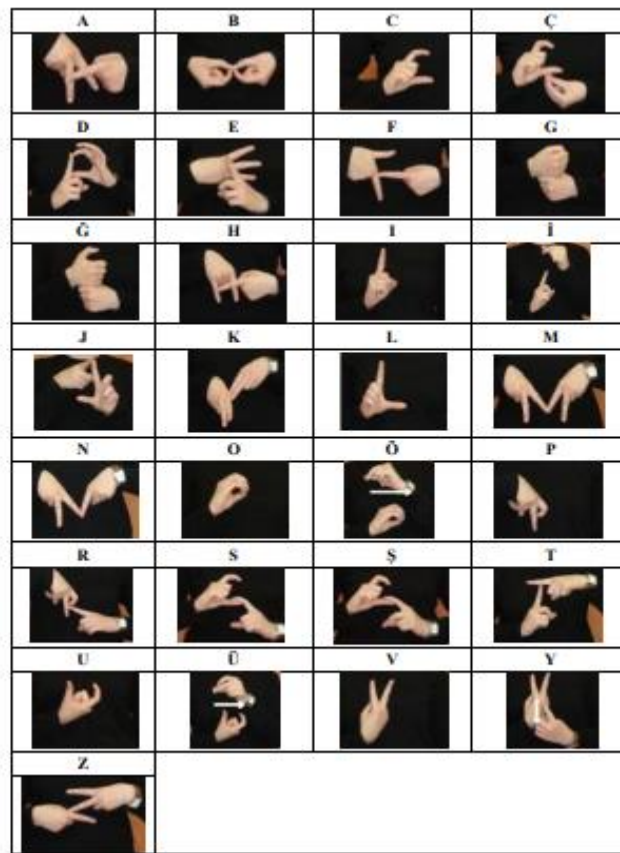
Şekil 1.4. 'Altı' işaretinin gösterimi (Karaca ve Bayır, 2018)

Şekil1.4'te bulunan ve durağan olmayan işaretlerden olan göz, durak ve hasta kelimelerin TİD'e göre gösterimi yer alırken; alttaki resimlerde ise birkaç hareketin birleşiminden oluşan değil işaretinin gösterimi bulunmaktadır. Bazı kelimeler tek bir hareket ile gösterilirken bazıları ise birkaç hareketin birlikte gösterilmesiyle oluşmaktadır

(Uysal, 2010). Bazen de sadece bir hareket tüm bir cümleyi anlatmakta kullanılabilir (Memiş, 2013).

1.3.2. Parmak alfabesi

Parmak alfabesi olarakta adlandırılan Türk işaret dilinde her bir işaretle karşılık 29 harf kullanılmaktadır. TİD alfabesinde işaretlerin gösterimi diğer ülkelerden farklı olarak iki el kullanılarak yapılmaktadır. Ancak, bu kuralın tüm harfler için geçerli olmadığını belirtmek önemlidir. Bazı harflerde sadece tek bir el kullanılabilirken, bazı harflerin gösterimi içinse çift el kullanımı söz konusudur. Örneğin; v, u, p, l, o, ı, c harflerinin gösteriminde iki el kullanılmazken, bu harfler dışında kalan harfler için iki el kullanımı gereklidir. Bu, işaret dili kullanıcılarının belirli harfleri ifade ederken ellerini nasıl kullandıklarını ve işaretledikleri harflerin görsel temsilini belirler. Parmak alfabesinin gösterimi göğsün hemen önünde yapılır ve çoğunlukla kısaltmalarda, özel isimlerde, eklerde ve yabancı dillerden alınan kelimelerde parmak alfabesi sıklıkla kullanılmaktadır (Karaca ve Görgünoğlu, 2016).



Şekil 1.5. Türk işaret dili parmak alfabesi (Kın, 2019)

Her harfin karşılığının olduğu TİD alfabesinde, hareketin yeterli gelmediği durumlarda veya özel isimlerin anlatımında parmak alfabesine başvurulmaktadır (Yasan, 2015).

1.3.3. Ağızla söyleme (Dudak hareketleri)

İşaret diliyle bir kelimenin veya ifadenin günlük hayattaki karşılığının ağızla söylenmesi, farklı şekillerde adlandırılabilir. Bu sürecin adları arasında "ağızla söyleme", "dudak hareketleri", ve "ağızlama" gibi terimler bulunabilir. Bu terimler, işareti yapılan sözcüğün işitme engelli birey tarafından ağız yoluyla ifade edilmesini tanımlar. Hareketle eş zamanlı yapılan ağızla söyleme, isimlere kıyasla fiillerde daha çok kullanıldığı görülmektedir (Karaca ve Bayır, 2018). Diğer işaret dillerinde olduğu gibi TİD kullanan kişilerin karşısındaki kişilerin dil kullanımı, işaret dilini anadil olarak kullanmasına veya sonradan öğrenmesine bağlı olarak değişmektedir. Bu olgu toplum dil bilimsel bakış açısından değerlendirildiğinde, günlük konuşmada karşındaki kişiye göre sözcük değişimi ile benzerlik göstermektedir. Bu da işitme engelli bireyler arasında TİD' i anadil olarak edinenlerle, TİD' i sonradan öğrenen ya da konuşma yeteneği olan bireyler arasında işaretleşme (işaret dili kullanımı) arasında farklılıklar görülebilir. TİD'i sonradan öğrenen kişilerde, iletişim hızı daha yavaş, parmak alfabesi kullanımı daha sık, dudakla işaretlerin Türkçesini söyleme daha fazladır. Çoğu zaman TİD kuralları göz ardı edilerek Türkçe'nin dilbilgisi kurallarına benzer bir yapıya sahip "işaretlenmiş Türkçe" adı verilen bir kodlama kullanılmaktadır. Özellikle bu durum haber bültenlerinde kullanılan işaretçilerde gözlemlenmektedir. Genellikle hem dudak okuma hem de ağızla söyleme yöntemi, Türkçe konuşan ve işaret dili bilmeyenlerle iletişimde geçildiğinde sıklıkla başvurulmaktadır (Dikyuva, 2015).

1.3.4. Jest ve mimikler

TDK' ya göre jest karşısındaki kişiye bir şey açıklarken genellikle el, kol ve başın kullanılarak yapıldığı iradeli ve içgüdüsel hareketlerken, düşünce ve hisleri karşımızdakine belirtecek şekilde yüzde yaptığımız hareketlerde mimik olarak adlandırılır (TDK, 2022). Dikyuva' nın belirttiği gibi, jest ve mimikler, işitme engelli olanların kullandığı görsel ve hareketli bir dil özelliği taşımaktadır. Jest ve mimikler,

işitme engelli bireylerin duygularını, düşüncelerini ve iletişim ihtiyaçlarını ifade etmelerine yardımcı olan önemli araçlardır (Dikyuva vd., 2015). Dikyuva' nın değindiği gibi işitme engelli bireylerin kullanımına daha uygun olan jest ve mimiklere TİD'de oldukça sık rastlanmaktadır. Jest ve mimikler işitme engelli bireyler tarafından sıklıkla kullanıldığı için TİD'in temelini oluşturan yapılardan olduğu söylenebilir (Karaca ve Bayır, 2018).

1.4. Araştırmanın Amacı

Dünya üzerindeki işaret dili sayısı 300'den fazla iken bu dili kullanana işitme engelli birey sayısı 70 milyon civarındadır. Türkiye verilerine göz atıldığında 1,5 milyon konuşma ve işitme engelli bireyin olduğu görülmektedir (Gül, 2023). İletişim teknolojilerinin hayatımızdaki önemi gün geçtikçe artmakta ve bu teknolojiler, insanlar arasındaki iletişimi hızlandırmak, kolaylaştırmak ve daha erişilebilir kılmak adına sürekli olarak geliştirilmektedir. Ancak, işitme ve konuşma engellilerin bu teknolojileri kullanımında karşılaştıkları sınırlamalar gerçek bir sorundur (Gül, 2023). Türkiye Cumhuriyeti Aile ve Sosyal bakanlığınca yapılan engelli ve yaşlı istatistik bültenine göre Türkiye'de 863.000 kişi (406.000 erkek, 429.000 kadın) işitmede zorluk yaşarken, 507.000 kişi de (278.000 erkek, 229.000 kadın) konuşmada zorluk yaşadığı raporlamıştır (EYGHM ,2022). Güncel rakamlar, işitme ve konuşma engelli birey sayısında bir artış olduğunu göstermektedir. Bu nedenle işaret dili tanıma sistemlerine ciddi ihtiyaç duyulmaktadır (Çınar vd., 2023). Dünya çapında çeşitli işaret dilleri kullanılırken, gerçek zamanlı işaret dili tanıma sistemlerinin geliştirilmesi, işaret dili kullanıcıları ile işitme kaybı olmayanlar arasındaki iletişimi kolaylaştırmak için hayati önem taşımaktadır. Bu sistemler, işitme engelli bireyler ve işitme güçlüğü çeken kişiler için yardımcı teknoloji oluşturmak üzere el ve vücut hareketleri, yüz ifadeleri ve genel hareket dahil çok modlu verileri analiz etmektedir. Bu sistemler, işaret dilini anlayarak, kullanıcıların ifadelerini anında çözümleyip yanıt verebilme potansiyeline sahip yapılardır. Gerçek zamanlı işaret dil tanıma sistemleri, bu bireylerin doğal bir şekilde işaret dili kullanmalarını ve bu yolla çevreleriyle iletişim kurmalarını sağlar. Bu, işitme engelli bireylerin kendilerini daha etkili bir şekilde ifade etmelerine ve çevreleri tarafından anlaşılmasına olanak tanır. Ayrıca işitme engelli bireylerin eğitim ve iş hayatına daha etkin bir şekilde katılmalarını sağlar. Eğitim kurumlarında ve iş yerlerinde bu teknoloji kullanılarak, işitme engelli bireylerin derslere, toplantılara veya mesleki etkinliklere tam katılımlarını

desteklemektedir. İnternet ve dijital medyanın önemli bir yer tuttuğu günümüzde, gerçek zamanlı işaret dil tanıma, işitme engelli bireylerin dijital içeriklere daha kolay erişim sağlamalarına yardımcı olur. Video konferanslar, online eğitimler ve diğer çevrimiçi etkileşimlerde bu sistemler, işaret dilini anlamaya ve üretmeye dayalı iletişimi mümkün kılar. Sonuç olarak, gerçek zamanlı işaret dil tanıma sistemleri, işitme engelli bireylerin toplumsal, eğitimsel ve profesyonel açıdan daha aktif ve etkin bir rol üstlenmelerine yardımcı olmakla birlikte bu teknolojilerle, dâhil edicilik ve çeşitlilik anlayışının bir parçası olarak, toplumun her kesiminin tam katılımını teşvik etmeye yönelik önemli bir adımdır. Gerçek zamanlı işaret dil tanıma sistemleri, işitme engelli bireylerin iletişimini kolaylaştırma, katılım ve erişimlerini artırma konusunda önemli bir rol oynamaktadır.

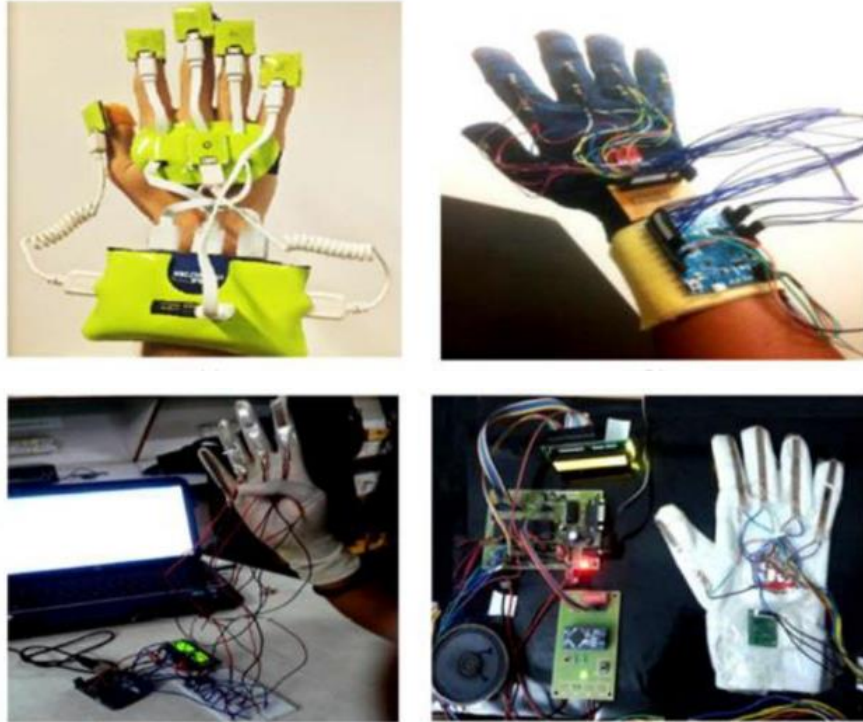
Bu amaçla bu çalışmada işaret dili tanıma sistemlerine katkı sağlamak hedeflenmiştir. Bu çalışmada amacımız, işitme engelli bireyler için erişilebilirliği ve iletişim imkânlarını artıracak verimli ve etkili işaret dili tanıma sistemlerinin geliştirilmesine yardımcı olmaktır. Bunun için önerilen mimaride önce veri seti literatürde sıklıkla kullanılan 17 işaret dili kelimeleri seçilmiştir. Seçilen işaret dili kelimelerine ait anahtar noktalar, Google tarafından geliştirilen açık kaynaklı bir yazılım kütüphanesi olan MediaPipe kullanılarak çıkarılmıştır. Bu kütüphane, bilgisayarlı görü (computer vision) ve makine öğrenimi (machine learning) tekniklerini kullanarak, çeşitli görsel ve ses tabanlı uygulamalar geliştirmek isteyen geliştiricilere yönelik bir araç seti sunmaktadır. Çıkarılan noktalar .npy dosyası olarak saklanmıştır. Sonrasında derin öğrenme modellerinden bir olan uzun kısa süreli hafıza ağları (LSTM) kullanılarak modelin eğitilmesi sağlanmıştır. Sonrasında Test için ayrılan verilerle modelin başarısı ölçülmüştür.

2. LİTERATÜR TARAMASI

Teknolojinin hızla gelişmesiyle işitme kaybı yaşayan bireyler için çeşitli dil tanıma sistemleri tasarlanmaktadır. Yapılan çalışmalar dikkatle incelendiğinde, dil tanıma sistemlerinin pozitif yönde ivme kazandığı görülmektedir. Bu bölümde İşaret dili tanıma sistemleri 4 başlıkta incelenmiştir. Bu başlıklar işaret dili tanıma sistemlerinde kullanılan veri setlerini oluşturmak için kullanılan araçlar dikkatle alınarak oluşturulmuştur.

2.1. Elektronik Eldiven Tabanlı Sistemler

Yapılan ilk çalışmalarda kullanılan elektronik eldiven tabanlı sistemler, kullanıcıları maliyet, kullanım kısıtlamaları ve pahalılık gibi dezavantajlarla karşı karşıya bırakan teknolojik çözümlerdi. Bu sistemler, özellikle pahalı donanım gereksinimleri nedeniyle geniş kullanıcı kitlesine ulaşmada sınırlamalara neden olmuştur. Ancak, son yıllarda bu dezavantajların üstesinden gelmek ve daha geniş bir erişim sağlamak amacıyla İşaret dilini görsel verilere dayalı olarak tanıyan sistemler ön plana çıkmıştır (Erkuş, 2020). Yapılan ilk çalışmalarda kullanılan eldiven tabanlı sistemler, işaret diline ait hareketlerin yapılmasıyla oluşan dinamik tanıma sistemlerinde kullanılmıştır (Katılmış,2021).



Şekil 2.1. Veri eldivenleri kullanılarak yapılan çalışma örnekleri (Çaylı, 2020)

Japonca işaret dili üzerine saklı Markov modeli kullanılarak yapılan çalışmada veri eldiveni kullanılmıştır. Veri eldiveni olarak tasarlanan sistemde çeşitli renklerde bantlar bulunmaktadır. Elin farklı duruşlarının farklı renklerde olmasına dikkat edilen çalışmada ellerin duruşlarından ziyade, elin aldığı harekete göre görünen bantların renkleri ve yoğunlukları araştırılmıştır (Iwai vd., 1996). Saengsri ve ekibi (2012) tarafından ortaya atılan bir öneri, Tayland İşaret Dili'ni tanıma sisteminin geliştirilmesidir. Bu sistem, geleneksel olarak kullanılan elektronik eldiven ve hareket izleme cihazlarına dayanmaktadır. Bu öneri, işitme engelli bireylerin işaret diliyle iletişim kurmalarını desteklemek ve teknoloji aracılığıyla daha etkili iletişim sağlamak amacıyla yapılmıştır. Saengsri ve diğerleri, Tayland İşaret Dili'nin tanınması için bu teknolojik çözümlerin kullanılmasının potansiyel avantajlarını inceleyerek, işitme engelli bireylerin günlük hayatta daha fazla bağımsızlık elde etmelerine katkıda bulunmayı amaçlamışlardır. Sistem, Tayland İşaret Dilinde tek elle yapılan parmakla yazılan alfabe hareketlerini tanıma amacıyla özel olarak tasarlanmıştır. Saengsri ve ekibi, bu çalışmalarında elektronik eldiven ve dört farklı hareket takip cihazı kullanarak elde ettikleri toplam 1000 örnek değeri üzerinden 14 farklı noktadan alınan 20 ölçümü kullanmışlardır. Sistem, Yapay Sinir Ağları (YSA) ve veri bölümlleme tekniklerini kullanarak %94'lük bir başarı oranına ulaşmıştır (Saengsri vd.,2012). Tangsuksant ve arkadaşlarının 2014 yılında yapmış oldukları çalışmada, 6 farklı renkli işaretleyici ve 2 kameraya sahip bir elektronik eldiven kullanılarak Amerikan İşaret Dili'ndeki (ASL) hareketleri tanımaya yönelik statik bir tanıma sistemi geliştirilmiştir. Renkli işaretleyicilerin tespiti dairesel Hough dönüşümü tekniği kullanılarak sağlanmıştır. İşaretleyicilerin 3 boyutlu koordinat noktaları Doğrudan Doğrusal Dönüşüm (DLT) algoritması kullanılarak 2 kamera eklenerek doğrudan tespit edilmiştir. ASL'nin 24 harfine odaklanılan çalışmada YSA sınıflandırıcısı ile %95 doğruluk oranına ulaşılmıştır (Tangsuksant vd., 2014). Wang ve ekibi (2006), gerçek zamanlı bir Amerikan İşaret Dili (ASL) tanıma sistemi geliştirmek amacıyla çok boyutlu Gizli Markov Modelleri' ne (HMM) dayalı bir çalışma gerçekleştirmiştir. Araştırmada, kullanılmak üzere elektronik bir eldiven ve 3 boyutlu hareket izleyiciden yararlanılmıştır. 26 alfabe ve 36 adet başlıca el şekli içeren ASL veri setinde , elde edilen ortalama doğruluk oranı %95 olarak belirlenmiştir. Bu çalışma, HMM tabanlı sistemlerin işaret dillerinin gerçek zamanlı tanınması için etkili bir yaklaşım sunduğunu göstermektedir (Wang vd., 2006). ASL kelime tanımaya yönelik, Yapay Sinir Ağlarına (YSA) dayalı bir sistem ve olasılıksal bir

model de 2015 yılında Sarawat ve arkadaşları tarafından geliştirilmiştir. Elin şekli, konumu, yönelimi ve hareketi dahil olmak üzere dört temel özelliği, ileri beslemeli sinir ağlarını içeren farklı işlevlerden yola çıkarak tanımlanmıştır. Bu özelliklerin kombinasyonu, ASL'de olasılıksal bir modele dayalı olarak belirli bir işaret ya da kelime oluşturmak amacıyla kullanılmış ve sistem, ASL işaretlerini anlık olarak tanıma ve veri tabanı ile etkileşimli olacak şekilde güncelleme yeteneğine sahiptir. Yapılan çalışmada 40 ASL kelimesini kullanarak %95 doğruluk oranına ulaşılmıştır (Sarawat vd., 2015).

2.2. Sensör Tabanlı Çalışmalar

İnsan hareketlerinin tespiti için geçmişte yapılan çalışmalarda stereo ve tof teknolojileri kullanılırken günümüzde bu teknolojilerin yerini alan şekil 2.2 de bulunan Kinect sensörü kullanılmaya başlanmıştır. Kinect sensör diğer kameralara oranla daha yüksek çözünürlüğe sahiptir. İnsan vücudunun bölümlerini 20 eklem noktasına ayırarak iskelet yapısını çıkaran Kinect sensörü bu verilerle yüz tanıma, vücut tanıma ve robotik uygulamalarında tercih edilmiştir. Bu alandaki çalışmalarda daha önce kameralar kullanılarak derinlik algısı yakalanmaya çalışılırken Kinect sensör ile farklı bir yaklaşım yakalanmıştır. İnsan duruşu ve hareketlerini algılamamak için kullanılan diğer bir teknoloji Leap Motiondur. Bu sensörde stereo-kızıl ötesi görüşe sahip 3 kızıl ötesi ışık ve 2 kamerası bulunmaktadır. Belli bir uzaklıkta sensör üstüne eller konulduğunda el hareketlerini algılayarak bilgisayara ortamına yansıtma özelliği bulunmaktadır (Duran ve Kaya, 2018).



Şekil 2.2. Microsoft Kinect 2 sensör ve Leap Motion sensör (Nabulsi, 2018)

Porfirio ve diğerlerinin 2013 yılındaki çalışması, Kinect sensörleri kullanılarak Brezilya İşaret Dili harflerini tanımak amacıyla geliştirilmiş bir yöntemi içermektedir. Çalışmada, işaret dilindeki farklı harflerin tanınması için elin 3 boyutlu görüntüsü ve 2

boyutlu gösterimi kullanılmıştır. Video karelerinde elin ön ve yan çerçevelere bölünmesi, her çerçeve çiftinin 3 boyutlu kafes silüet yöntemi ile oluşturulması, işaretlerin tanınmasında etkili bir yaklaşım sunmaktadır. Bu çalışma, Kinect sensörlerinin işitme engelli bireylerin işaret dilini tanıma konusunda kullanışlı bir araç olabileceğini göstermektedir. Ayrıca, RGBD video kareleri kullanılarak elde edilen verilerin, görüntü işleme teknikleriyle işaretlerin doğru bir şekilde sınıflandırılmasına katkı sağlayabileceğini ortaya koymaktadır. Özniteliklerin çıkarılması için küresel harmonik yöntem kullanılırken, geliştirilen sistem, Brezilya İşaret Dili harflerinin yer aldığı toplam 610 video görüntüsü üzerinde beş farklı imzacı tarafından test edilmiştir. Deneysel sonuçlara göz atıldığında, %96 ile en yüksek sınıflandırma oranına sahip olan mimari Destek Vektör Makinelerinin (SVM) olduğu vurgulanmıştır (Porfirio vd., 2013). Sun ve meslektaşlarının gerçekleştirdiği çalışma, Kinect sensörleri aracılığıyla elde edilen renk, derinlik ve iskelet gibi dataları içeren ASL görsellerinin olduğu bir veri seti oluşturmayı amaçlamıştır. Bu veri seti, otomatik olarak tanıma yeteneğine sahip bir sistem geliştirmek için kullanılmıştır. Veri seti çıkarılırken, 2D kameraların görüntü sistemleri özelliklerini kullanmak yerine Kinect cihazlarının derinlik çıkarabilme yeteneğini kullanılmıştır. Oluşturulan veri setinde 73 ASL kelimesinin dokuz kişi tarafından üç kez tekrarlanmasıyla bir veri seti oluşturulmuş ve toplam 1971 örnek elde edilmiştir. Bu sistemde kullanılan metod Gradyan Odaklı Histogram (HOG) tanımlayıcısını kullanarak özellik çıkarımına dayanmaktadır. Deneysel sonuçlar, Destek Vektör Makinelerinin (SVM) %86 'lık bir başarı oranı yakalanmıştır (Sun C.vd., 2013). Kinect V2'sini kullanarak ArSL işaretlerini tanıyabilen model Hisman ve arkadaşı tarafından 2019 da yapılan bir çalışmayla literatüre kazandırılmıştır. Yapılan çalışmada vücut bütünüyle yapılan hareketli işaretlerin yanı sıra el hareketlerinin detaylı tespiti için önemli bir yöntem önerilmiştir. Bu çalışmada, 3 kişi tarafından oluşturulan 40 Arapça tıbbi terimden oluşan ArSL adlı veri seti toplamda 1200 örneklemeden oluşmaktadır. Bu veri seti, Arapça tıbbi terimlerin işaret dilindeki karşılıklarını içermektedir. Sınıflandırma sonuçlarına göre, K-Nearest Neighbors (KNN) algoritması %89 başarı elde ederken, Support Vector Machine (SVM) algoritması %79 başarı göstermiştir. Bu sonuçlar, kullanılan sınıflandırma yöntemlerinin, elde edilen veri setindeki işaret dillerini başarıyla tanıma yeteneğini yansıtmaktadır (Hisham ve Hamouda, 2019).

2.3. Video Kamera Tabanlı Çalışmalar

İşaret dili tanıma sistemlerinde veri seti oluşturmak için kullanılan araçlardan diğeri de video kameralardır. Şekil 3.3'te stereo ve tof kamera görselleri yer almaktadır. Bu kameralardan stereo kamera binoküler yapıdadır. Binoküler yapı iki kaynaktan aldığı verileri üst üste koyarak derinlik algısı oluşturma işlemidir. Stereo kameralarda 2 veya daha fazla kamera ile değişik açılardan alınan görüntülerin 3 boyutlu görüntüsünü çıkarmada kullanılır. Bu işlemin zor olması sebebiyle stereo kameraların dezavantajı olarak kabul edilir. Tof kameralarda derinlik algısının yakalanması için ışınlardan faydalanılmaktadır. Bu ışınlar nesnelere gider nesnelerden geri dönerken kamera ve nesne arasındaki mesafeyi hesaplar. Görüntü kalitesinin düşük olması sebebiyle her ne kadar ucuz ve hızlı olsa da bu kameralar için sorun teşkil etmektedir (Duran ve Kaya, 2018).



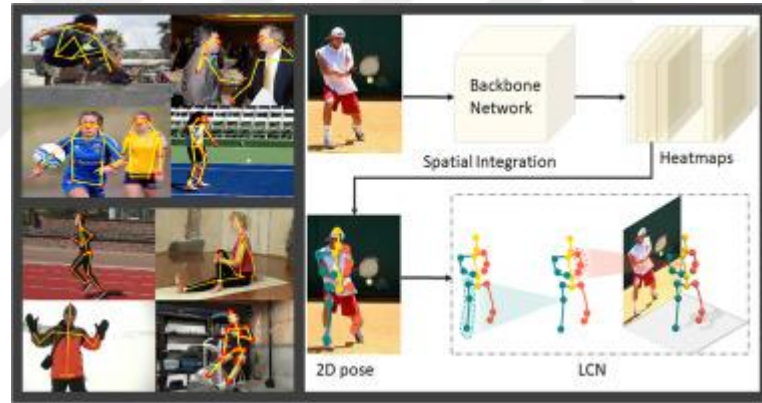
Şekil 2.3. Stereo ve Tof kamera (Duran ve Kaya, 2018)

Özcan ve ekibinin gerçekleştirdiği çalışma, evrimsel sinir ağı modelinde hiperparametre optimizasyonu sağlayan sistem kullanarak, 49 engelliden elde edilen 25 farklı sözcük videolarından oluşturulan bir veri setini incelemiştir. Bu veri seti, engellilerin tıbbi terimleri anlamalarına yardımcı olmak amacıyla tasarlanmıştır ve içinde hastane acil servisinde sıkça kullanılan kelimeleri barındırmaktadır. Çalışmanın önemli sonuçlarından biri, CNN tabanlı GoogLeNet modeliyle yapılan sınıflandırma işleminin %93,93'lük bir başarı oranına ulaşmasıdır. Bu, tıbbi terimlerin işaret dilinde doğru bir şekilde tanınması ve sınıflandırılmasının, engellilere sağlık hizmetlerine daha etkili bir şekilde erişmelerine yardımcı olabileceğini göstermektedir (Özcan ve Baştürk, 2020). Pacal ve Alptekin tarafından 2023 yılında yapılan çalışmada (Pacal ve Alaftekin, 2023), Türk işaret diline ait rakamlar CNN tabanlı güncel ve popüler mimariler tarafından sınıflandırılmıştır. 1500 örnekleme yapılan sınıflandırmada CNN tabanlı mimarilerin ortalama %90 üzerinde başarı yakalayarak, işaret dilini sınıflandırmada iyi sonuç verdiği gözlemlenmiştir. Bangla işaret dili sınıflandırma üzerine yapılan çalışmada Tasmere ve arkadaşı CNN kullanarak çalışmalar yapmışlardır. Bangla dilinin 37 harfi 6 kişi tarafından yapılarak toplamda 3219 örneklem ile bir veri seti oluşturulmuştur. Elde edilen

veri kümesi oluşturulan CNN ile sınıflandırıldığında %99 oranında bir doğruluk elde edilmiştir (Tesmer ve Ahmed, 2020).

2.4. İnsan Poz Tahmin Tabanlı Çalışmalar

İnsan poz tahmini, son birkaç on yılın en önemli bilgisayarlı görme görevlerinden biridir ve anahtar noktalar olarak da bilinirler. Belirli bir RGB görüntü veya videoda K vücut eklemlerini (dirsekler, bilekler vb. gibi) konumlandırarak ve ayrıca uzuvların yönünü tanımlayarak insan duruşunu otomatik olarak tahmin etme ve izleme görevini üstlenir. Poz tahmini, insan hareketinin otomatik izlenmesi ve değerlendirilmesi için kullanılan bir tekniktir. Spor analizi, insan-bilgisayar etkileşimi ve artırılmış gerçeklik benzeri uygulamalarda kullanılmaktadır. Bu teknik, bir kişinin vücut pozisyonunu ve hareketlerini belirleyerek, bu bilgileri analiz etmeyi ve yorumlamayı sağlar (Gamra ve Akhloufi, 2021).



Şekil 2.4. Steore ve Tof kamera (Gamra ve Akhloufi, 2021)

Bilgisayar alanında önemli bir yere sahip olan insan poz tahminini ile ilgili çalışmalarda araştırmacılar genellikle, insan bilgisayar etkileşimi, aksiyon tanıma, gözetim, resim anlama gibi konularda önemli uygulamalar yapmışlardır. Bununla birlikte uygulama alanının çeşitliliği aslında oldukça fazladır (Muradli, 2021). İnsan poz tahmini kullanılarak yapılan işaret dili tanıma çalışmalarına bakıldığında 2022 yılında ASL'nin LSTM ile sınıflandırılması adlı çalışmada hem statik hem dinamik işaretleri tanımak için MediaPipe kütüphanesi kullanılmıştır. El işaretlerinden oluşan veri setini oluşturmak için Google tarafından geliştirilen MediaPipe kütüphanesi açık kaynak kodlu bir uygulamadır. Dört farklı insanın el hareketlerinden oluşturulan veri seti 93.000 adet numpy dizisi haline

getirilmiştir. Veri seti el eklemlerinin koordinatlarının yanı sıra zaman parametresi de içermektedir. LSTM modeli kullanılarak sınıflandırılan veri seti %99 bir oranla doğru tahminlerde bulunmuştur (Sundar ve Bagyammal, 2022). Halder ve Tayede (Halder ve Tayede, 2021) tarafından yapılan çalışmada Amerikan, Hint, İtalyan ve Türkiye gibi çoklu işaret dili veri kümeleri kullanılmıştır. Çalışmada ilk olarak MediaPipe kullanılarak veri setleri oluşturulmuş ardından çalışma veri temizleme ve normalizasyon işlemi ile devam etmiştir. Çeşitli makine öğrenme teknikleri ,yapay sinir ağları ve çok katmanlı algılayıcılar kullanılarak sınıflandırma yapılan çalışmada ortalama %99'lık bir başarı oranı yakalanmıştır. 2022 yılında Subramanian ve arkadaşları (Subramanian vd., 2022) tarafında, Hint işaret dilinin sınıflandırılmasında Kapılı Tekrarlayan Birim Modeli (MOPGRU) kullanılmıştır. Çeşitli yönlerde ve farklı aydınlatma koşullarında bir web kamerası kullanılarak her biri 640×480 boyutunda 30 çerçeveli her bir işaret hareketi için 30 videodan oluşan veri havuzu oluşturulmuştur. Oluşturulan veri havuzu eğitim, test ve doğruluk olacak şekilde 3 kategoriye ayrılmıştır. Oluşturulan veri seti üzerinde MOPGRU modeli ile ortalama %95 oranında bir doğruluk yakalamıştır. Sundar ve Bagyammal (Sundar ve Bagyammal ,2022), yapmış oldukları çalışmada dinamik işaret dili sorununu ele almıştır. MediaPipe ellerin, vücudun ve yüzün kilit noktalarını çıkararak konumu, şekli ve yönü belirlemek için kullanılmıştır. Amerikan İşaret Dili (ASL) alfabelerinin tanınmasına yönelik basit ve etkili, vizyona dayalı bir yaklaşım tartışılmıştır. Önerilen sistemde Amerikan işaret dili alfabesinin 26 harfi kullanılmıştır. LSTM ile yapılan sınıflandırmada %99 doğruluk oranına ulaşılmıştır. 2021 yılında OpenPose uygulaması kullanılarak Bangladeş İşaret Dilinin Tanınması amacıyla 10 Bengalce kelime seçilerek çalışma yapılmıştır. Veri seti sınıflandırılması içinse LSTM-RNN, K-neighbors, SVM ve Gaussian Naive Bayers modelleri kullanılmıştır. Çalışmada en başarılı olan sınıflandırma modeli %96 gibi bir oranla LSTM-RNN ile elde edilirken Gaussian Naive Bayers modeli diğerlerine oranla daha düşük bir başarı yakalamıştır (Dewanjee, 2021). Türk işaret dilinin insan poz tahmin sistemleriyle sınıflandırılmasına yönelik bir çalışma Kırıcı ve arkadaşları 2022 yılında bir çalışma yapmışlardır. Çalışmada Mediapipe Holistic kullanarak Türk işaret diline ait 29 harften oluşan veri setinin anahtar noktaları çıkartılarak yapay sinir ağ mimarilerinden olan LSTM ile sınıflandırılması sağlanmıştır. LSTM ile oluşturulan modelde eğitim sırasında verimin artırılması için Keras tarafından geliştirilen callback fonksiyonları kullanılmıştır (Kırıcı vd., 2023). Ko ve arkadaşları tarafından (Ko vd., 2019), OpenPose kütüphanesini kullanarak kendi oluşturdukları ve KETI adı verdikleri veri setinin anahtar nokta tespitini yapmışlardır.

Anahtar noktaları çıkarılan görüntülerin çeşitli istatistiksel ölçüleri alınarak normalleştirme işlemi uygulanmıştır. Normalleştirilen görüntüler diziden diziye dayalı çeviri modeline veri girişi olarak kabul edilmiş ve sonrasında RNN (tekrarlayan sinir ağları), LSTM ve GRU (kapılı tekrarlayan birimler) gibi modellerde sınıflandırılma işlemine tabi tutulmuştur. Santhalingam 'da Ko gibi çalışmasında OpenPose kütüphanesini kullanarak videodan anahtar nokta çıkarımı yapmıştır. Videolardan alına el görüntüleri manuel olarak seçilip LSTM'de eğitilmiştir. Görüntülerin sınıflandırılması işleminde ise RNN modelinden faydalanılmıştır (Santhalingam vd., 2020). Kim ve arkadaşlarının yaptığı çalışmada, işaret dilinin belirli bir sensör veya giyilebilir cihaz yardımı olmadan tanınma oranını artırmayı amaçlamıştır. Bunun için derin öğrenme yöntemiyle Kore işaret dilinin sınıflandırmasında, insan poz tahmininde OpenPose kütüphanesi kullanılmıştır. Veri seti için oluşturulan kelimelere OpenPose kütüphanesi uygulanarak anahtar kilit noktaları çıkartıldıktan sonra CNN ile sınıflandırma yapılmış ve oldukça başarılı sonuçlar elde edildiği gözlemlenmiştir (Kim ve Jung, 2021).

3. MATERYAL VE YÖNTEM

Bu çalışmada, LSTM tabanlı bir derim öğrenme modeli TİD sınıflandırıcı olarak önerilmektedir. Veri seti tez kapsamında oluşturulmuş tamamen özgün bir veri setidir. Veri setinde 5 denekten alınmış 17 farklı işaret dili hareketi yer almaktadır. Her hareket aynı denek tarafında 30 kez tekrar edilmiştir. Dolayısıyla her denegin örnek sayısı $17 \times 30 = 510$ ve veri seti toplamda 2550 adet örnek içermektedir.

Çalışmada kullanılan LSTM modeli, Python programla dili ile Tensorflow ve Keras paketleri kullanılarak gerçekleştirilmiştir. Modeller ve analizler AMD Ryzen 7 5800H (16 CPU, ~3,2GHz) işlemciye, Nvidia GeForce RTX 3050 GPU (4GB GDDR6, ~1,5GHz) ve 16 GB RAM'e sahip bir kişisel bilgisayar üzerinde gerçekleştirilmiştir. Modelde, ardışık 3 LSTM katman yer almakta ve her bir katmanda çıktı büyüklüğü sırasıyla 64, 128, 64 olarak belirlenmiştir. Sonraki aşamada 3 Dense katmanlar yer almakta ve her bir katmanda çıktı büyüklüğü sırasıyla 64, 32 ve 17 dir.

3.1. Veri seti

Gerçek zamanlı başarılı bir işaret dili tanıma sistemi geliştirmek için, Türk İşaret Dilinin (TİD) jestlerini ve hareketlerini yakalayan yüksek kaliteli görüntü çerçeveleri elde etmek çok önemlidir. Önerilen mimarimizde, yetişkin 5 denek tarafından gerçekleştirilen çeşitli TİD hareketlerinden (kelimelerinden) oluşan TİD veri kümesini kullanılmıştır. Veri setimizde 17 farklı kelime bulunmaktadır. Veri seti; “Hata”, “arkadaş”, “iyi”, “seni seviyorum”, “beğenmek”, “konum”, “tanışmak”, “dikkatli olmak”, “teşekkür etmek”, “düşünmek”, “sen”, “polis”, “kafa”, “aşağı”, “yukarı”, “baba” ve “saygı” kelimelerini içermektedir. Hazırlanan modelin TİD’i başarılı bir şekilde öğrenmesi ve analiz edebilmesi için nitelikli veri setlerinin oluşturulması gerekmektedir. Bu nedenle, önerilen mimaride, yetişkin 5 denek tarafından gerçekleştirilen çeşitli TİD hareketlerinden oluşan geniş bir veri kümesi kullanılmıştır. Bu veri kümesi, TİD’ in çeşitli yönlerini kapsayan 17 farklı kelimeyi içermektedir. Aşağıda bulunan tabloda veri setinde kullanılan TİD işaretleri yer almaktadır.

Tablo 3.1. Veri setinde kullanılan işaretler

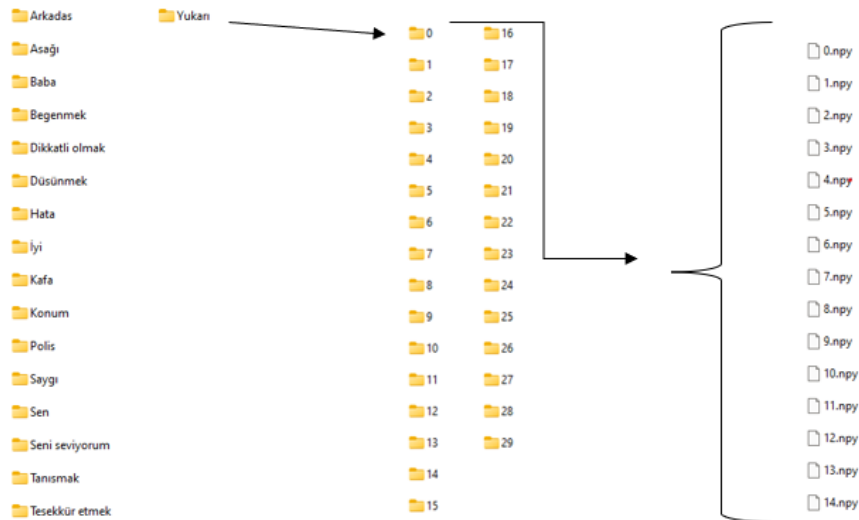
Hata	İyi	Arkadaş
Polis	Yukarı	Aşağı
Saygı	Seni Seviyorum	Baba
Sen	Beğenmek	Düşünmek
Dikkatli Olmak	Kafa	Tanışmak
Teşekkür Etmek	Konum(Yer)	

Gerçek zamanlı işaret dili tanıma sistemi için önerdiğimiz mimaride, doğru el takibi gerçekleştirmek için Google tarafından geliştirilen açık kaynaklı bir proje olan Mediapipe modülü kullanılmaktadır. MediaPipe modülü, sağlam ve etkili el poz tahmini sunarak her iki elin hareketlerini ve konumlarını gerçek zamanlı olarak takip etmemizi sağlar. Şekil 3.1’ de görüldüğü gibi, Mediapipe el takip modülü, her el için toplam 21 yer işareti çıkararak onların mekânsal konfigürasyonlarını ve hareketlerini yakalamaktadır. Bu yer işaretleri, işaret dili tanıma modelinin sonraki aşamaları için temel özellikler olarak hizmet eder.



Şekil 3.1. Mediapipe el işaret noktaları

Her işareti MediaPipe 15 farklı kare olacak şekilde yakalamıştır ve sonrasında her kelime için bu işlem 30 defa tekrar edilerek veri seti çeşitliliği artırılmıştır. Görüntülerin anahtar noktaları alınarak, x, y ve z koordinatları “.npy” uzantılı dosyalar içerisinde NumPy dizileri halinde saklanmıştır. LSTM mimarisinde eğitilmek üzere oluşturulan veri setini %80’i eğitim, %20’u ise test olacak şekilde bölünmüştür.

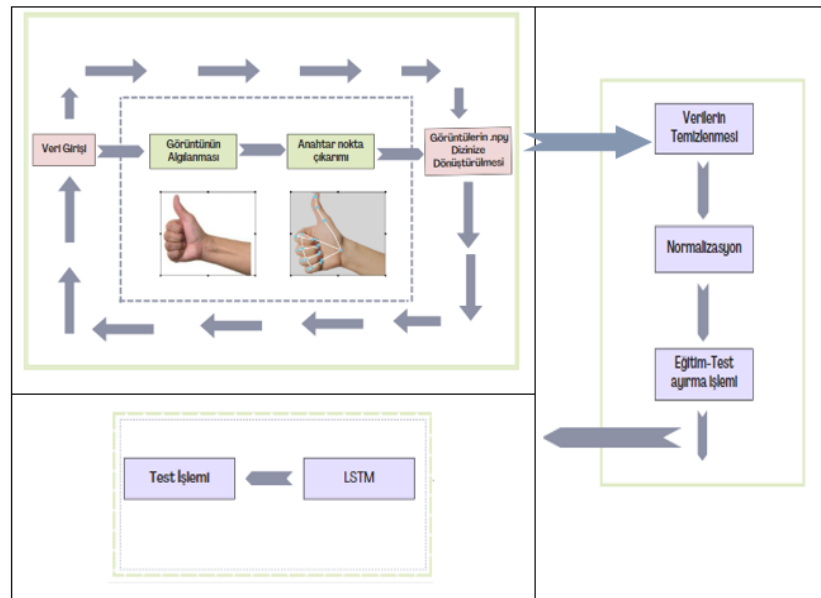


Şekil 3.2. Modelde kullanılan kelimelere ait .npy dosyaları

Her sınıf, sınıf başına 510 örnekle önemli sayıda görüntüyle temsil edilir. Bu büyük veri kümesi boyutu, önerdiğimiz mimarinin eğitim süreci sırasında önemli miktarda veriden yararlanmasını sağlar. Sonuç olarak model, TİD hareketleriyle ilişkili karmaşık el hareketlerini ve ince varyasyonları etkili bir şekilde öğrenebilir. Bu büyüklükte bir veri kümesinin kullanılması, modelimizin TİD hareketlerinin hem ince nüanslarını hem de farklı özelliklerini yüksek hassasiyetle yakalamasına olanak tanımaktadır.

3.1.Önerilen Mimari

Gerçek zamanlı işaret dili tanıma sistemi için önerilen mimari, TİD hareketlerini doğru bir şekilde yorumlamayı ve sınıflandırmayı amaçlamaktadır. Bunu başarmak için görüntü çerçevesi edinimi, elle izleme, özellik çıkarma ve sınıflandırmayı içeren çok adımlı bir süreç kullanılmaktadır. Geniş bir TİD veri kümesinden ve en son teknoloji tekniklerden yararlanan mimarimiz, modelin TİD hareketlerinin karmaşık ayrıntılarını ve hareketlerini hassasiyetle yakalamasını sağlar. Şekil 3.3'te önerilen İşaret dilinin derin öğrenme modelleriyle metin haline dönüştürülmesi adlı çalışmaya ait genel mimarinin akışı göstermektedir. Bu diyagram, sistemimizde yer alan sıralı adımların görsel bir temsilini sağlar.



Şekil 3.3. Önerilen mimari

Mimarinin ilk uygulama aşamasının başlayabilmesi için veri girişi gerekmektedir. MediaPipe tarafından kamera görüntüdeki elleri algılamaya başlamakta, sonrasında görüntü de bulunan ellerin kilit noktalarının belirlenmesiyle işlem devam etmektedir. Görüntüye ait belirlenen kilit noktalarda (düğümlerin) x, y, z eksenlerine ait koordinat bilgileri bulunmaktadır. Bu bilgiler MediaPipe tarafından .npy formatına dönüştürülerek saklanmaktadır. Kelimeleri temsil eden her hareket 30 kez tekrar edilerek veri setinin tümü oluşturulmuştur. MediaPipe tarafından .npy formatına dönüştürülüp saklanan verilerde, herhangi boş veri girişi olup olmadığını kontrol etmek için Pandas kütüphanesinden yararlanılmıştır. Görüntünün bulanık olması, ekranda el olmaması vb. sebeplerden dolayı, hatalı verilerin temizlenmesi gerekmektedir. Bu boş girdileri içeren satırlar aranır ve bulunan boş dizinler Pandas Kütüphanesi tarafından kaldırılmıştır. İstenmeyen hatalı veriler çıkarıldıktan sonra verilerin normalizasyonu adımına geçilmiştir. Normalizasyon, verileri 0 ve 1 arasında yeniden ölçekleme işlemidir. Normalizasyon işlemi bittikten sonra veri seti 2'ye ayrılmıştır. Verilerin %80' i eğitim ve geri kalanı işe test için bölümlendirilmiştir. Sonrasında oluşturulan veri setinin LSTM modelinde eğitime işlemi gerçekleştirilmiştir. LSTM modeli yinelemeli sinir ağlarından biridir. Veri setini oluşturan kelimeler durağan hareketler olmadığından LSTM modeli seçilmiştir ve sırasıyla veri temizleme ve normalizasyon işleminden sonra LSTM modelinin eğitilmesi için kullanılmıştır. Oluşturulan eğitilmiş LSTM model, hareketlerin gerçek zamanlı olarak tahmin edilmesi için saklanmıştır. Sistemde eğitim verilerine bağlı olarak ağ ağırlıklarının tekrar tekrar güncellenmesi için, Adam Optimizer tekniği kullanılmıştır.

3.2.Yapay Zekâ ve Tarihçesi

Antik Roma uygarlığına uzanan yapay zekâ, insana dair düşünebilme, muhakeme etme, öngörülü olma gibi kompleks işlemlerin makinelere aktarılması üzerine çalışmaların yapıldığı bir bilim dalıdır (Çalışkan ve Demir,2022). Genel anlamda yapay zekâ, insan zekasına ait tüm bilişsel ve otonom davranışları taklit etmesi istenen bir sistemdir (Wikipedia, 2023). 1943 yılında, Walter Pitts ve Warren McCulloch tarafından geliştirilen "Beynin Boolean Devre Modeli" adlı çalışma, sinir ağlarının temelini atmıştır. Bu, sinir ağlarının bilgi işleme süreçlerini matematiksel olarak modellemeye yönelik önemli bir adımdır. 1950 senesinde, Alan Turing'in "Bilgi işleyen makineler ve zekâ" adlı

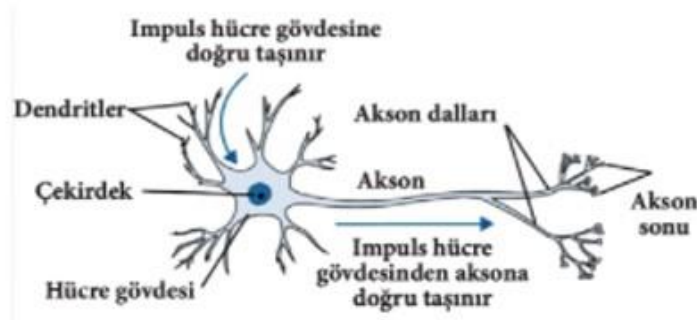
makalesinde, "Makineler Düşünebilir mi?" sorusu gündeme gelmiştir. Turing aynı zamanda "Turing Testi" ni önererek bir makinenin zekâya sahip olup olmadığını belirleme amacını taşımıştır. 1956 yılında, Dartmouth Koleji'nde bir grup bilim insanı ve matematikçi, "yapay zekâ" terimini ilk kez resmi olarak kullandı. Bu konferans, yapay zekâ araştırmalarının başlamasında önemli bir kilometre taşı oldu. 1950'lerde, birkaç önemli yapay zekâ programı geliştirildi. Bunlar; Samuel'in Kontrol Edici Programı: Oyun teorisine dayanan bu program, bilgisayarın satranç oynayarak öğrenmesini sağladı.

Newell ve Simon'un Mantık Teorisi: Bu teori, insan düşünme süreçlerini simüle etmeye çalışan bir mantık teorisini içeriyordu. Gelernter'ın Geometri Motoru: Bu program, geometrik problemleri çözmek için tasarlanmıştı. IBM (International Business Machines; Uluslararası İş Makineleri), 1952-1969 yılları arasında satranç oynayabilen ilk programı yazmıştır. 1965 yılında Gray Robinson mantıklı düşünebilen bir algoritma geliştirmiştir. 1966-73 yılları arasında yapay zekânın hesapsal karmaşayla karşılaşması sonucu araştırmalara uzun bir süre ara verilmiştir. 1969-79'da Bilgiye dayalı sistemlerin ilk adımları atılmıştır. Yapay zekanın endüstri haline gelmesi 1980'leri bulmuştur. 1986'da popülerliğini tekrar kazanan yapay zekâ 1987'de bilime dönüştürülmüştür. Zekâ ajanları kavramı 1995 yılında ortaya çıkmış, 1999' de ise Deep Blue Kasparov'u yenmiştir. 1998'de internetin dünya üzerinde geniş kitlelere yayılması ile yapay zekâ programları kitlelere ulaşmıştır. Robot uygulamaların pazara sürülmesi 2000'li yılları bulmuştur. Günümüzde katlanarak artan yapay zekâ popüleritesi birçok alanda kendini göstermektedir (Pirim, 2006). Bilginin hızla yayılmasının ardından ortaya çıkan veri çokluğu bilgisayar teknolojilerinin oldukça hızlı ilerlemesinin sonuçlarından biridir. Bu veri çokluğu içerisinde kullanılmak istenen verilerin hızlı ve doğru bir yöntemle elde edilmesi amaçlanmaktadır. Bu soruna çözüm bulunması amacıyla birden fazla teknik geliştirilmiştir. Bunlar; veri madenciliği, veri analitiği, makine öğrenmesi gibi tekniklerdir. Günümüzde derin öğrenme büyük ilgi görmektedir ve birçok alanda başarıyla kullanılmaktadır. Derin öğrenme, büyük veri setlerinden öğrenme ve karmaşık görevleri yerine getirme yeteneği ile öne çıkmaktadır (Alzubaidi vd., 2021).

3.3.Yapay Sinir Ağları

Beyindeki nöronların işlevsel mekanizmalarına benzer bir şekilde tasarlanan yapay sinir ağları beyni taklit ederek bilgisayar yazılımları aracılığıyla veri toplama, hatırlama, genelleme yapma ve yeni veri üretme gibi temel işlevleri gerçekleştiren yapay

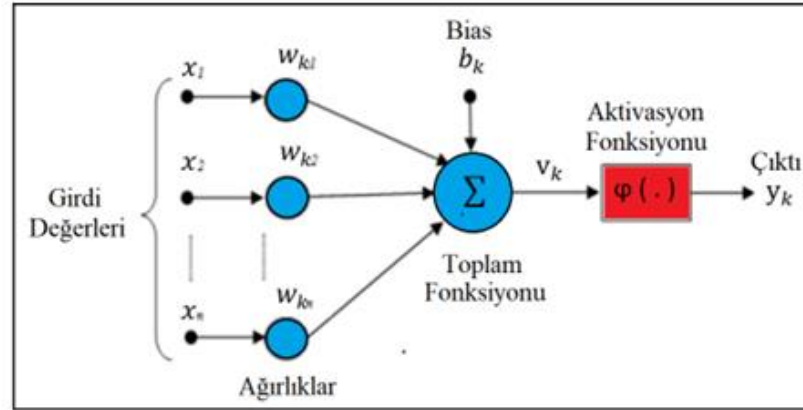
zekâ teknolojileridir. Yapay sinir ağı, insan beyninden esinlenerek matematiksel olarak modellenmiş öğrenme sürecidir. Yapay sinir ağı, farklı türlerdeki mimarileri ve yapıları ifade etmek için çeşitli adlar altında tanımlanabilir. Bazı yaygın olarak kullanılan tanımlamalar şunlardır. Yapay Sinir Ağı; bu terim, genel anlamda yapay sinir ağını ifade eder. Yapay sinir ağı, birçok bağlantılı nöronun bir araya gelmesiyle oluşan bir bilgisayar modelidir. Derin öğrenme, yapay zeka alanında kullanılan bir öğrenme yöntemi olup, özellikle çok katmanlı yapay sinir ağlarının kullanıldığı bir alt dalı simgeler. Daha derin ve karmaşık yapıdaki ağlar, genellikle büyük veri setleri üzerinde karmaşık desenleri öğrenme yeteneğine sahiptir. Paralel Dağıtılmış Ağlar; birçok bağımsız işlemci veya nöronun aynı anda çalıştığı bir mimariyi ifade eder. Bu, hızlı öğrenme ve işleme yetenekleri sağlar. Bağlantılı Ağlar; nöronların birbirine bağlı olduğu ve bilginin bu bağlantılar aracılığıyla iletilip işlendiği yapay sinir ağını tanımlar. Nöromorfik Ağlar; biyolojik nöronların işleyişini daha yakından taklit etmeye çalışan yapay sinir ağlarıdır. Bu tür ağlar, sinir sistemine daha benzer bir mimari kullanarak öğrenme yeteneklerini geliştirmeye çalışır (Öztürk ve Şahin, 2018). Yapay sinir ağı modeli ilk defa 1943 yılında Warren McCulloch ve matematikçi Walter Pitts tarafından isimlendirilmiştir. Warren McCulloch ve Walter Pitts biyolojik bir sinir hücresini nöron olarak adlandırırken 4 temel yapıdan oluştuğunu ve bu yapıların dendrit, hücre gövdesi, aksonlar ve sinapslar olduğunu belirtmişlerdir (McCulloch ve Pitts, 1943).



Şekil 3.4. Nöron'un biyolojik modeli (Karpathy, 2018)

Frank Rosenblatt, 1957 yılında "Perceptron: An electronic device for pattern recognition" adlı makalesinde, biyolojik sinir hücrelerinden esinlenerek perceptron kavramını tanımlamıştır. Perceptron, yapay sinir ağlarının temelini oluşturan basit bir öğrenme modelidir. (Rosenblatt, 1957). Şekil 3.2'de gösterilen bir nöronun matematiksel gösterimi yer almaktadır. Görselde McCulloch-Pitts modeline göre sinir hücresi n adet

girdi değeri (x_j) alır ve sonrasında her girdi değerine ağırlık değişkeni (w_j) eklenir. Modelde yer alan x_j ve w_j değerleri doğrusal olarak toplandıktan sonra aktivasyon fonksiyonuna (ϕ) çarpılır. Bir yapay sinir ağında, nöronun çıkışı genellikle bir aktivasyon fonksiyonu kullanılarak belirlenir. Bu çıkış, genellikle "yk" veya "y" ile gösterilir. Bu işlemler sırası aşağıda bulunan denklem 1.1' deki gibi formül kullanılarak gösterilebilir. (McCulloch,1943).



Şekil 3.5. Nöron 'un matematiksel modeli (Karpathy, 2018)

$$u_k = \sum_{j=1}^n w_{kj} x_j \quad (3.1)$$

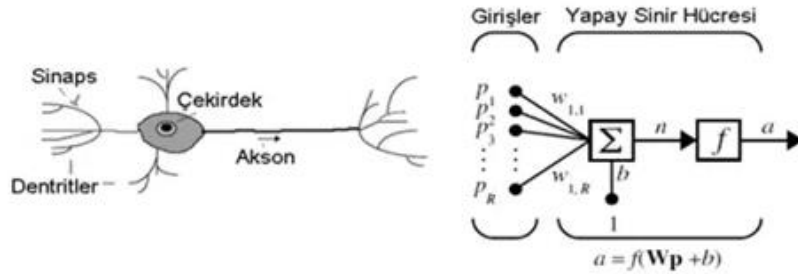
$$y_k = \phi(u_k + b_k) \quad (3.2)$$

$$v_k = u_k + b_k \quad (3.3)$$

$$y_k = \phi(v_k) \quad (3.4)$$

İstenilen çıkış elde edilinceye kadar sistemin ağırlıkları değiştirilerek nöronun eğitimi gerçekleştirilir. Yapay sinir ağının oluşturulması birden fazla nöronun kullanılmasıyla elde edilir (Rosenblatt, 1957). Warren McCulloch ve WA Pitts, 1940'lı yılların başında Yapay Sinir ağlarının (YSA) temellerini attılar. İnsan beyninden ilham alan YSA, mevcut verilere dayanarak genellemeler yapma ve bu genellemelere dayanarak bilinmeyen veriler hakkında kararlar verme yeteneğine sahiptir. Bir Yapay Sinir Ağındaki nöronlar arasındaki bağlantılar, ağırlıklar aracılığıyla ifade edilir. Bir nöronun çıkışını belirlemek için girişlere ağırlıklar uygulanır, bu değerler toplanır ve ardından bir aktivasyon fonksiyonuna geçirilir. Bu şekilde, bir katmandaki nöronların çıkışları, diğer katmanında girdi değerleri olarak kullanılır. Bu bağlantılar ve ağırlıklar, YSA' nın öğrenme yeteneğini sağlar. Yapay Sinir Ağları, genellikle öğrenme süreci sırasında

ağırlıkları ayarlayarak bir veri seti üzerindeki belirli bir görevi öğrenir. (Şahin, 2018). Biyolojik sinir sisteminin her bir elemanının yapay sinir ağı sisteminde bir dengi bulunmaktadır. Yapay sinir ağlarında biyolojik sinir hücresinin bir modeli genellikle şu şekilde ifade edilir. Dendritler -girişler; biyolojik sinir hücresinin dendritleri, yapay sinir ağlarında girişleri temsil eder. Bir yapay sinir ağında her bir giriş bir dendrite karşılık gelir. Toplama fonksiyonu-Hücre gövdesi; dendritlerden gelen girişler, bir toplama fonksiyonu ile bir araya getirilir. Bu adım, ağırlıklı girişlerin toplamını hesaplamayı içerir. Her dendrite bağlı bir ağırlık bulunur ve toplama fonksiyonu bu ağırlıklı girişlerin toplamını üretir. Transfer Fonksiyonu-Hücre gövdesi; toplama işleminden elde edilen toplam, bir transfer fonksiyonu aracılığıyla işlenir. Bu fonksiyon, genellikle nöronun çıkışını belirlemek için kullanılır. Örneğin, yaygın bir transfer fonksiyonu olan sigmoid fonksiyon, çıkışı belirlemek için kullanılır. Ağırlıklar -Sinapslar, her bir dendrite bağlı olan ağırlıklar, sinaptik güçleri temsil eder. Bu ağırlıklar, girişlerin toplamını belirlemek için kullanılır. Çıkış -Akson; transfer fonksiyonundan geçirilen toplam, yapay sinir hücresinin çıkışını oluşturur. Bu çıkış, bir sonraki katmandaki nöronlara giriş olarak kullanılabilir (Öztürk ve Şahin, 2018).



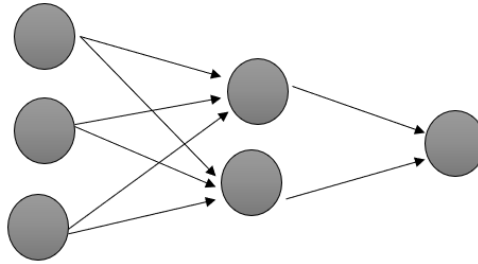
Şekil 3.6. Biyolojik sinir hücresi ile yapay sinir ağı (Görür, 2020)

YSA' lar, bir dizi özellikle avantajlı olan temel özelliklere sahiptir ve bu özellikler, onları birçok uygulama alanında kullanılabilir kılar. Eksik Verilerle Çalışma, yapay sinir ağları, eksik verilerle çalışabilme yeteneğine sahiptir. Bu, gerçek dünya veri setlerinde sıkça karşılaşılan bir durumdur. Çok Sayıda Değişken ve Parametre Kullanma, YSA' lar genellikle büyük ve karmaşık veri setlerini işleyebilir ve birçok değişkeni, parametreyi içerebilir. Uyarlanabilirlik, YSA' lar, öğrenme süreçleri ile uyarlanabilirler. Yeni veri ve deneyimlere göre modellenabilirler. Doğrusal Olmama, yapay sinir ağları, doğrusal olmayan ilişkileri ve desenleri modelleyebilir, bu da gerçek dünya problemlerine daha iyi uymalarını sağlar. Paralel Çalışma, YSA' lar, paralel işlem yetenekleri ile büyük veri

setlerini etkili bir şekilde işleyebilirler. Öğrenme ve Genelleme, yapay sinir ağları, öğrenme yetenekleri sayesinde veri setlerinden örüntüler çıkartabilir ve bu öğrenmeleri genellemeye yönlendirilebilirler. Hata Toleransı, yapay sinir ağları, belirli bir düzeyde gürültü ve hatalara karşı direnç gösterebilirler. Esneklik, yapay sinir ağları, birçok farklı problem ve görevde kullanılabilecek genel amaçlı modeller oluşturabilirler. Bu özellikler, YSA' ları çeşitli endüstri ve bilim alanlarında kullanılabilir hale getirir. Sağlık, otomotiv, finans, enerji ve birçok diğer sektörde YSA' lar, karmaşık problemleri çözmek, veri analizi yapmak, öngörülerde bulunmak ve daha birçok görevde etkili bir şekilde kullanılabilmektedir (Öztürk ve Şahin, 2018).

3.3.1. İleri beslemeli yapay sinir ağları

Sinir ağlarının en basit hali ile tanımlanmış şekli olan ileri beslemeli yapay sinir ağlarındaki nöronlar katmanlardan oluşmaktadır (Yadav ve Kumar, 2015). Katmanlarda yer alan hücrelerin çıktıları kendinden sonraki katmana ağırlık değeri eklenerek girdi olarak gönderilir. Girdi ve çıktı katmanları arasında gizli katmanlar bulunmakta ve bilgi akışı girdi katmanından gizli katmanlara doğru gerçekleşir. Gizli katmanlarda bulunan nöron sayısı problem durumuna göre değişiklik gösterirken, nöron miktarının optimal sonucunu gösteren bir metot şu ana kadar geliştirilmemiştir (Şenol ve Denizhan, 2023).

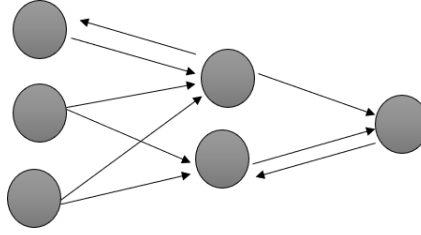


Şekil 3.7. İleri beslemeli yapay sinir ağı yapısı

3.3.2. Geri beslemeli yapay sinir ağları

Katmanlar arasından geri beslemenin yapıldığı bir sinir ağı çeşididir. Dinamik hafıza yapısına sahip olan geri beslemeli sinir ağının nöron çıktıları, yalnız o anki giriş parametrelerine bağlı değil kendinden önceki giriş değerlerinden de etkilenmektedirler. Bu sebepten ötürü özellikle tahmin uygulamalarında başarılı sonuçlar verdiği için tercih

edilmektedirler. Ayrıca çeşitli zaman serisi uygulamalarında da iyi sonuçlar verdiği gözlemlenmiştir (Aşkın vd., 2013).

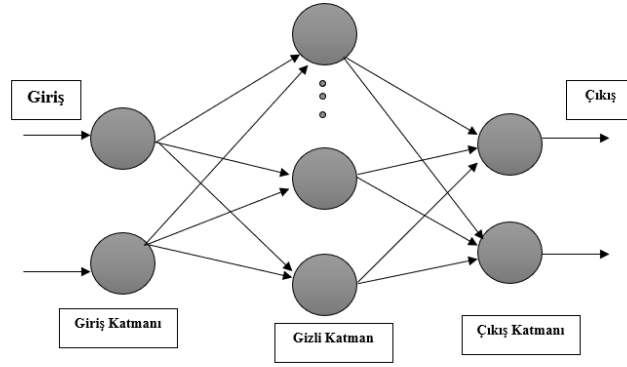


Şekil 3.8. Geri beslemeli yapay sinir ağı yapısı

Şekil 3.5 'te görüldüğü üzere bu sinir ağında nöronun çıktısı kendinden önce gelen nörona veya katmana ya da kendi ile aynı katmanda bulunan herhangi bir nöronun giriş değeri olarak bağlanabilir. Böylelikle giriş değerleri iki yönde aktarılmış olur. Bu özelliğinden ötürü non-lineer dinamik davranış sergilemektedirler (Şahin, 2023).

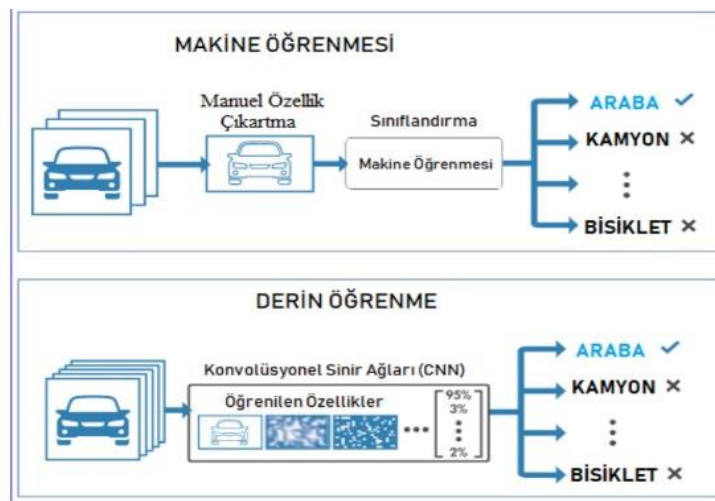
3.4. Derin Öğrenme

Makine öğrenme tekniklerinden biri olan derin öğrenme 1986 yılında ilk defa Dechter tarafından yapılan çalışmada bu kavramdan bahsedilmiştir. Bir sinir ağı olan derin öğrenme kavramını ise ilk kez 2000'lerde Igor Aizenberg ve ekibi ortaya atmıştır. Derin öğrenme mimarileri birden fazla yapay sinir ağının bir araya gelmesinden oluşmuştur. Derin sinir ağı olarakta isimlendirilen derin öğrenme yöntemleri, bünyesinde birden fazla sinir ağı yapısı barındırdığından dolayı bu şekilde de isimlendirilmektedirler (Ezel, 2018). Yapay sinir ağlarının aksine çok fazla katmandan meydana gelen derin öğrenme 1990 ile 2000 yılları arasında veri ve donanım yetersizliğinden ötürü istediği etkiyi yakalayamamıştır. Sinir ağı modelinin karmaşıklığı işlem yükünü artırdığından o dönemde kullanılan donanımsal özelliklerin yetersiz olması derin öğrenme yöntemlerinin ilerlemesine engel olmuştur (Erkuş, 2020).



Şekil 3.9. Derin sinir ağı yapısı

Derin sinir ağlarında katmanların gösterim şekli aynı anda tüm katmanların bütünleşik olduğu bir şekilde gerçekleşir. Kontrol tek bir geri bildirim sinyali tarafından olurken hedefe varmak için modelde sürekli bir değişim meydana gelir (Chollet, 2021). Derin öğrenmeyi klasik makine öğrenmesi tekniklerinden farklı yapan en önemli özelliği ham veriyi işlemeden önce ön işleme tabi tutarak veriyi model için kullanılır hale getirmesidir. Girdilerin belirgin olmasını sağlamak için özniteliklerini çıkararak problemin çözümünü kolaylaştırır. Derin öğrenmenin bu özelliğinden ötürü makine öğrenmesine kıyasla problemlerin çözümü daha kolaydır. Şekil 3.4’te derin öğrenme ve makine öğrenmesinin girdilerin tahmin edilmesinde izledikleri yol gösterilmiştir (Erkuş,2020).



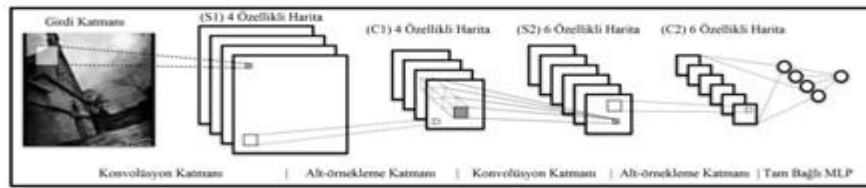
Şekil 3.10. Makine öğrenmesi ve derin öğrenme yaklaşımı (Erkuş, 2020).

Derin öğrenme metotları 2 temel gruptan oluşmaktadır. Bunlar danışmanlı öğrenme ve danışmansız öğrenmedir. Veri setinde bulunan veriler etiketli değil ise buna danışansız aksine veri setine ait verilerde etiket varsa ve bu şekilde sistemin öğrenmesi sağlanıyorsa buna danışmanlı öğrenme denir (Çalışkan ve Demir, 2022).

3.5. Derin Öğrenme Mimarileri

3.5.1. Konvolüsyonel sinir ağları (CNN)

Çok katmanlı algılayıcılar (MLP), genellikle derin öğrenme veya yapay sinir ağları bağlamında kullanılan bir terimdir. Hayvanların görme merkezinden yola çıkarak tasarlanan ileri yönlü basit bir CNN modelinde gerçekleşen matematiksel konvolüsyon işlemi, bir nöronun uyarı alanında belirli bir bölgeye odaklanan bir filtre (kernel) tarafından gerçekleştirilen işlemdir. Bu, özellikle görüntü işleme görevlerinde kullanılan bir operasyondur. CNN genel olarak 3 temel ana katmandan oluşmaktadır. Bunlar; sayısı değişiklik gösterebilen konvolüsyonel katman, alt örnekleme (subsampling) katmanı ve bağlı katmandan oluşmaktadır (Şeker vd.,2017). CNN mimarisinin geliştirilmesiyle AlexNet, GoogLeNet, VGG, LeNet gibi modeller geliştirilmiştir (Tan vd., 2021). Şekil 3.5'te 1988 yılında LeCun tarafından bulunan ve 1998 yılına kadar yenilenerek iyileştirmeleri yapılan ilk CNN ağı olan LeNet modeline ait mimar gösterilmiştir (Cun vd., 1989).



Şekil 3.11. LeNet ağının mimarisi (Şeker vd., 2017)

CNN modelinin kullanım alanları oldukça geniştir. Bunlar; doğal dil işleme (NLP), görüntü-ses işleme ve biyomedikal benzeri birçok farklı alan olsa da en iyi sonuçları görüntü işlemede elde etmiştir. ImageNet Yarışmasında çok sayıdaki resim ve nesne sınıflandırılması kategorilerinde en iyi sonuçları CNN algoritmalarının varyasyonları almıştır. Görüntü işleme dışında NLP’ de kullanılan CNN modelleri

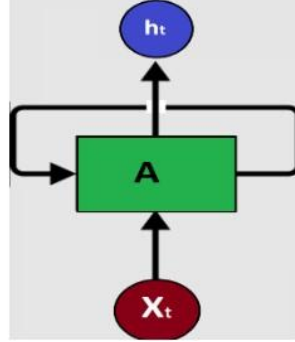
başarılı sonuçlar almıştır. NLP ye ait anlamsal ayrıştırma, arama sorgusu elde etme, sınıflandırma, cümle modelleme ve tahmin problemleri gibi dallarda iyi performans sergilemiştir (Şeker vd.,2017).

3.5.2. Tekrarlayan sinir ağları (RNN)

Geleneksel sinir ağlarının aksine tekrarlayan sinir ağlarında her giriş çıkış ile bağlantılıdır. Bir cümle yapısında sonraki kelimenin tahmin edilebilmesi kendinden önceki kelimeye bağlıken klasik sinir ağlarının aksine bunu hatırlayabilen sinir ağı RNN'lerdir. Farklı bir örnek ile anlatmak gerekirse izlenen filmin sonunun tahmin edilebilmesi için, belli bir yere kadar izlenen filmin olay akışı hatırlanarak tahminde bulunulur. RNN bu tarz durumlar için tasarlanmış sinir ağıdır. Standart ağlarda fark edilen hafıza yapısı eksikliğinden ötürü bu yeni mimari önerilmiştir. Geleneksel sinir ağlarına eklenen gizli katman ve yapısındaki döngülerle tasarlanan RNN sayesinde bu sorun çözülmüştür (Sabaz, 2022). RNN'lerin geliştirilmesi sürecinde birçok türetilmiş model ve iyileştirilmiş mimariler ortaya çıkmıştır. Bu mimari ve modelleri, RNN' nin performansını artırmak, uzun vadeli bağımlılık sorunlarını çözmek ve daha karmaşık sıralı veri problemleriyle başa çıkabilmek için bulunmuşlardır (Tan vd.,2021).

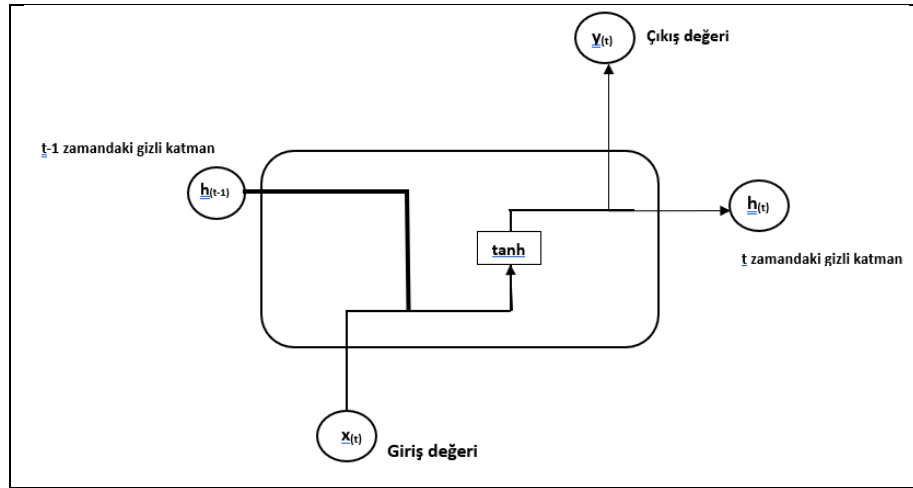
Jeff Elman tarafından geliştirilen Simple Recurrent Network (SRN), RNN alanında önemli bir ilerlemedir. SRN, RNN'lerin girişlerle çıkışlar arasında geçmiş bilgiyi koruyabilen bir türüdür. Cümle yapısı simülasyonunu kullanan Elman her bir kelimeyi isim ve fiil kategorilerine ayırmıştır. Ayrıca çalışmasında canlı-cansız, insan-hayvan ve hayvanlar bazında avcı-yırtıcı şeklinde gruplandırmıştır (Mikolov,2010). Tekrarlayan Sinir Ağı (Recurrent Neural Network RNN) birimleri arasında bulunan bağlantılar yönlendirilmiş döngülerden oluşmaktadır. Döngü yapısının oluşturduğu ağ iç yapısı sayesinde dinamik zamansal davranışlar sergilenir. İleri beslemeli sinir ağları ile arasındaki en önemli fark RNN'lerin kendi giriş belleğini, girdilerin gelişigüzel dizilerini işleyebilmeleridir (Mikolov,2010). RNN'lerdeki amaç sıralı bilgileri kullanabilmek iken; görüntü tabanlı verilerde giriş-çıkış birbirlerini etkilemediği kabul edilmektedir. Doğal dil işleme alanlarında (NLP) dilin zamansal bir sıra içinde dizilimi nedeniyle giriş ve çıkışlar birbiriyle bağlantılıdır. Çünkü doğal dil işlemede cümle içinde kullanılan kelimeden sonra hangi kelimenin geleceğini tahmin etmek gerekir. RNN yinelenen olarak

isimlendirilmesinin nedeni cümledeki sözcükler gibi dizinin her ögesini önceki çıktılara bakarak yerine getirmesidir (Şeker vd.,2017).



Şekil 3.12. Standart RNN yapısı (Sabaz, 2022)

Şekil 3.12 incelendiğinde sinir ağı ünitesi olarak görev yapan A ünitesinde X_t giriş değeri olurken h_t ise çıkış değeri olarak yer almaktadır. Döngü bilginin bir adımdan başka adıma geçerek bilginin sinir ağında dolaşımının aktarılmasını sağlamaktadır. T zamanında h_t çıktı değeri bir sonraki zaman diliminde (t+1) girdi olarak kullanılır. RNN sinir ağı yapısı bu sebeplerden ötürü metin, video gibi zaman serisinden oluşan ve adımlar arasında ilişkilerin bulunduğu durumlarda kullanılmaktadır (Sabaz, 2022).



Şekil 3.13. Basit bir RNN nöronunun yapısının detaylı gösterimi

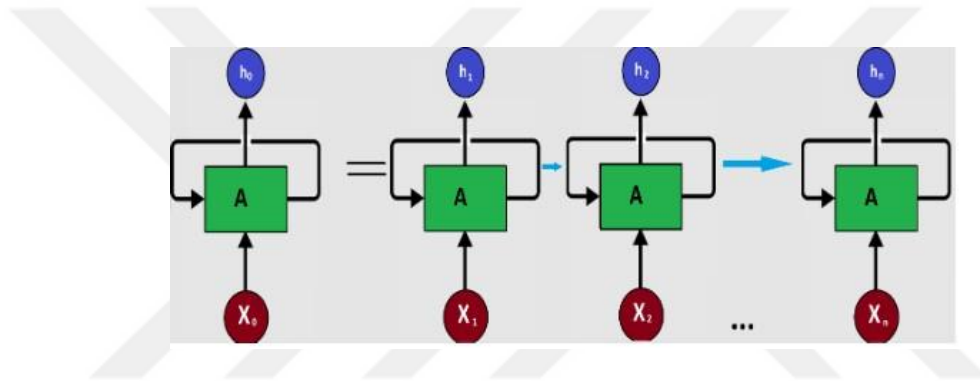
Şekil 3.13 baz alınarak RNN hücresine gelen ve sonrasında çıkış olarak kabul edilen değerın hesaplanması aşağıda verilen fonksiyon işlemleri ile gerçekleştirilir.

Formülde yer alan U_h , V_h giriş değerleri W_h ağırlık değeri, b_h bias değeri, h_t gizli katman, y_t çıkış değerini temsil etmektedir.

$$h_t = \tanh(U_h * W + V_h * h_{t-1} + b_h) \quad (3.2)$$

$$y_t = (W_y * h_t + b_y) \quad (3.3)$$

Geleneksel sinir ağlarına çok fazla benzeyen RNN'ler birden fazla klasik yapay sinir ağlarının zamanla tekrarlamasıyla elde edilir. Sinir ağı ünitesi olarak görev yapan A ünitesi çıktı üretir. Her çıktı değeri bir sonraki ünitelere giriş değeri olarak yollarır (Sabaz, 2022).



Şekil 3.14. RNN döngülerinin birleştirilmesi (Sabaz, 2022)

Şekil 3.14 denklem 3.4 - 3.5 ve denklem 3.6 işlemlerin standardize edilmiş halini temsil eder. Döngülerin birleştirilmesiyle ortaya çıkan çıktıların adım basamakları aşağıda belirtilmiştir.

$$h_t = f(h_{t-1} - X_t) \quad (3.4)$$

$$h_t = \tanh(W_h h_{t-1} + W_{xh} X_t) \quad (3.5)$$

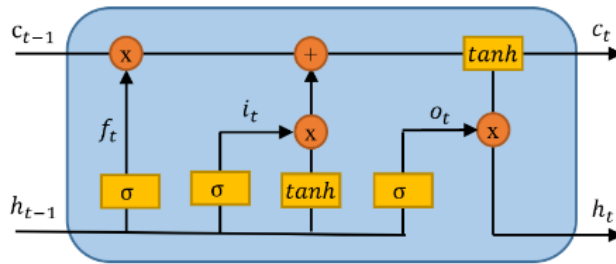
$$y_t = W_h y_t \quad (3.6)$$

RNN'lerin öğrenme kabiliyeti olumlu bir gelişme olmakla birlikte, bazı önemli sorunlarla karşılaşılır. Bu sorunlardan biri, kaybolan gradyan problemidir. Aktivasyon fonksiyonları kullanımının amacı, sonuç değerlerini çoğunlukla belirli bir aralığa değerine indirmektedir. Bu dar aralık, girdideki büyük değişimlere karşı aktivasyon fonksiyonunun tepkisinin sınırlı olmasına neden olabilir. Bu durumda türev değeri küçük

olur ve ilgili katman yeterli öğrenmeyi gerçekleştiremez. Tekrarlayan ağlarda bu problem, katmanların yeterince öğrenememesine ve uzun metin veya videolarda gördüklerini unutmaya başlamasına yol açabilir. Bu nedenle, hafıza süreleri kısa vadeli olur ve geçmişteki olayların veya girdilerin uzun vadeli etkilerini hatırlamakta zorlanırlar. Bu sorunları çözmek için LSTM kullanılır (Sabaz, 2022).

3.5.3. Uzun kısa süreli hafıza ağları (LSTM)

LSTM, geleneksel RNN'lerin uzun vadeli bağımlılık problemleriyle başa çıkamaması sorununu aşmak için tasarlanmıştır. Bu kapılar, hangi bilgilerin hatırlanması ve hangilerinin unutulması gerektiğini ayarlar. Eğer gönderilen girdi önemsizse, bu bilgi unutulur; değilse bir sonraki duruma aktarılır. Bu işlemlerin gerçekleştiği yerler, kapılar ve taşıyıcı hücre bantlarıdır. LSTM kapıları, önceki bilgileri dikkate alarak gelecekteki tahminleri geliştirebilen bir mekanizmayı temsil eder, bu da uzun vadeli bağımlılık sorununu çözer ve kaybolan gradyan problemine karşı direnci artırır (Sabaz, 2022). 1997 yılında Hochreiter ve Schmidhuber bu özel ağı bulan kişilerdir. LSTM mimarisinde 3 kapı bulunmaktadır. Bunlar; giriş, unutma ve çıkış kapılarıdır. Tasarlanan ilk LSTM modeli günümüzde kullanılan mimariden farklıdır. Tasarlanan ilk yapıda gözetleme bağlantıları ve unutma kapısı bulmamaktaydı. LSTM mimarisindeki unutma kapısı kendi belleğini sıfırlamada kullanılırken, kesin zamanlamaları bellekte tutmak için gözetleme bağlantıları mimariye dahil edilmiştir (Şeker vd., 2017).



Şekil 3.15. Standart LSTM hücresi (Tanışman vd., 2021)

Unutma kapısının temel işlevi kendisine gelen bilginin unutulup unutulmadığına karar vermekken giriş kapısı gelen bilgilerden hangisinin hafızada tutulacağına karar verir. Çıkış kapısında ise tahmin ve gelen bilgilerden hangisinin çıktı olacağı kararı verilir. LSTM algoritmasında ilk iş olarak unutma kapısına yeni bilgi olarak gelen X_t ve

önceki katmandan iletilen h_{t-1} bilgileri bir dizi işleme tabi tutulur. İşlem sonucunda hangisinin silineceğine karar verilir. Bunun için aktivasyon fonksiyonu olarak sigmoid fonksiyonu kullanılır (Yılmaz ve Orman, 2021).

$$f_t = (W_f * X_t + W_{f,h} * h_{t-1} + b_f) \quad (3.7)$$

Sonraki işlem giriş kapısında gerçekleşir. i_t eşitliği ile sigmoid fonksiyonu kullanılarak bilgiler güncellenir ve güncellenen bilgiler tanh fonksiyonuna koyularak hangilerinin depoda tutulacağına karar verilir.

$$i_t = (W_i * X_t + W_{i,h} * h_{t-1} + b_i) \quad (3.8)$$

$$C_t = \tanh(W_{c,x} * X_t + W_{c,h} * h_{t-1} + b_c) \quad (3.9)$$

Yeni bilgilerin oluşturulması aşaması taşıyıcı bantlarda aşağıda eşitlikle gerçekleşir.

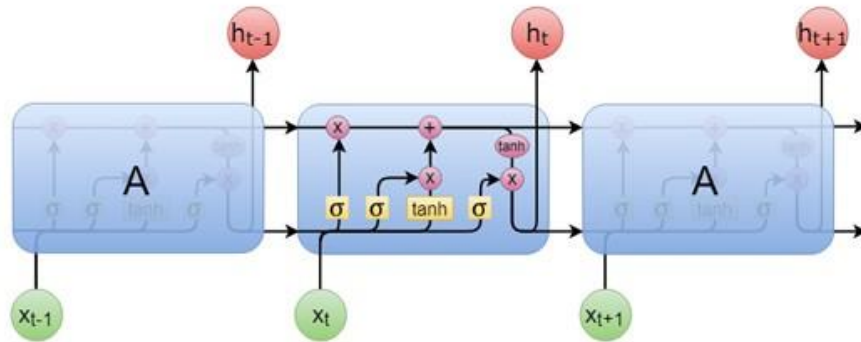
$$C_t = C_{t-1} * f_t + i_t * C_t \quad (3.10)$$

Son aşamada çıkış kapısında çıktı olarak sonraki katmana iletilecek değer bulunur.

$$o_t = (W_o * X_t + W_{o,h} * h_{t-1} + b_o) \quad (3.11)$$

$$h_t = o_t * \tanh(C_t) \quad (3.12)$$

Bu döngü şeklinde devam ederek ağırlık (w) ve bias (b) değerlerinin minimize edilmesi için gerçek eğitim ve çıktı değerleri farkı alınarak model eğitilmektedir.



Şekil 3.16. LSTM mimari yapısı (Özer, 2020)

LSTM yapısında bilginin hatırlanıp unutulmasına karar veren mekanizma Unutma kapısına aittir. Çalışma prensibi sayının büyüklüğünden bağımsız olan unutma

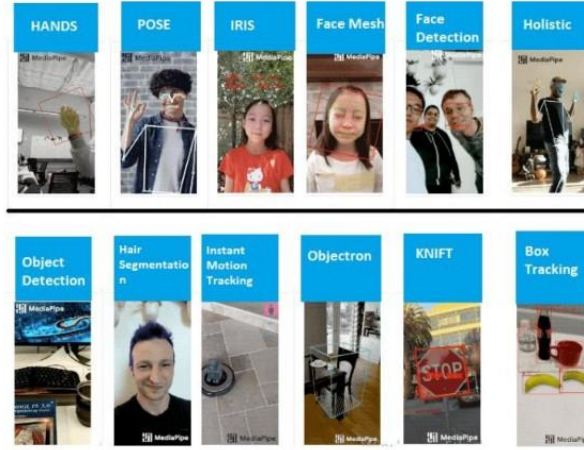
kapısında, unutulmak istenen değer 0 çarpanı ile çarptığında unutma işlemini basitçe gerçekleştirmiş olur. Bu işlemde önemli olan asıl faktör unutulması gereken bilgiye karar verme işlemidir. Bunu belirleme işlemi şu şekilde gerçekleşir; her hücre kendinden önce gizli katmandan gelen bilgiyi ve halihazırdaki verileri sigmoid fonksiyonuna yollayarak işleme tabi tutar. İşlem sonucunda bulunan değer 1 yakınsa bilgi hafızada tutulurken sıfıra yakın değerlerde bilginin unutulmasına karar verilir (Sabaz,2022). LSTM mimarileri konuşma/metin işleme, müzik bestesi üretimi, robotik kodlama ve otomasyon, kalp cerrahi sistemleri, el yazısı tanıma gibi konularında oldukça iyi sonuçlar vermektedir (Şeker vd., 2017).

3.6. MediaPipe Kütüphanesi

Görüntü ve videolarda yer alan insan vücudunun belli bölümlerinin ve eklemlerinin pozlarının tanımlanması ve sınıflandırılması işlemine insan poz tahmini adı verilmektedir. Kullanılan teknik insan vücudu pozlarını 2B ve 3B uzayda temsil etmek ve bu pozların çıkarımını yapmaya dayanmaktadır. Tam olarak kullanılan yöntem görüntü ve videolarda bulunan insana ait kilit noktalar olarak çeşitli eklemlerin tanımlanarak koordinatlarının çıkarılıp bir diziye aktarılması işlemidir. Dizi haline getirilen veriler modelle girdi olarak verilir. Model girdilerin koordinatları çıktı olarak tanımlanır ve işlem sonucunda tahminlerin kesinliğini belirten bir puan verir . İnsan Poz tahmini çalışmalarında sıklıkla kullanılan kütüphaneler şunlardır; MediaPipe, OpenPose, Pose Detection, DensePose, AlphaPose, HRNet ve YOLOv7 (Walia, 2022). MediaPipe, Google tarafından geliştirilen bir çözüm ve hizmet oluşturma çerçevesidir. Kaynak kodları Github'da paylaşılan bu servis, makine öğrenimi için kullanılan ve modüler bir yapıya sahip olan bir yapıdır. Yazılımcılara hızlı ve kolay geliştirilebilen bir mimari sunar. Bu özelliği sayesinde sadece masaüstü uygulamalar için değil, birçok platformda kullanılabilir olması büyük bir avantajdır (Sabaz, 2022). MediaPipe kütüphanesi Google araştırmacıları tarafından makine öğrenmesi esaslarına dayanarak geliştirilen, açık kaynak kodlu bir yazılımdır (Aydemir ve Arslan, 2021).

MediaPipe kütüphanesi, yazılım geliştiricilere geniş bir yelpazede uygulamalar oluşturmak için kullanabilecekleri çeşitli görüntü işleme ve makine öğrenimi modüllerini sunan etkileyici bir araçtır. Bu kütüphane, kullanıcıların karmaşık teorik konulara hâkim olmadan bile yüksek düzeyde işlevsellik eklemelerine olanak tanır. İşte MediaPipe kütüphanesinin sağladığı başlıca hizmetle. MediaPipe kütüphanesinin sunduğu hizmetler

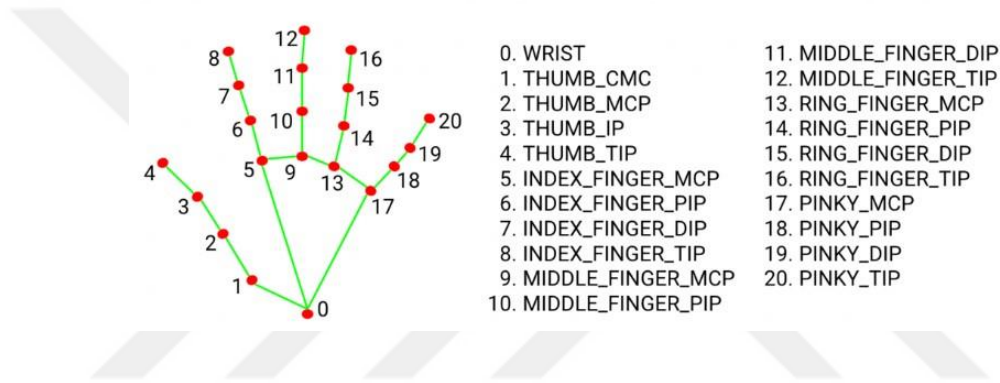
şunlardır; Yüz Tespiti ve yüzde nokta oluşturma, Saç Segmentasyonu, 3 Boyutlu Nesne Tespiti, Çoklu El Takibi, Poz Sınıflandırma, Nesne Tespiti ,Mobil Cihazlarda Takip, Yüz Tespiti, Nesne Takibi, Nesne Eşleştirme v.d (Sabaz, 2022).



Şekil 3.17. MediaPipe'in sunmuş olduğu hizmetlerin (Sabaz, 2022)

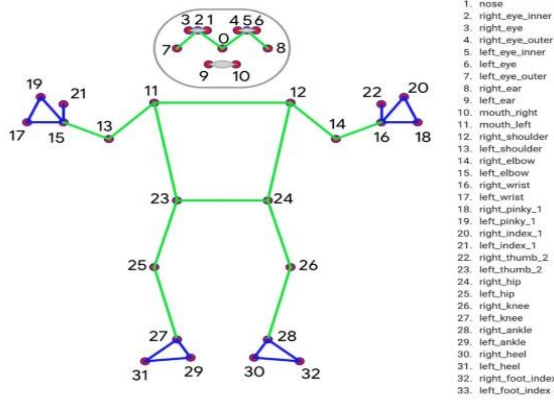
MediaPipe Holistic, optimize edilmiş bir işlem hattına sahip yeni bir insan pozu topolojisi sunarak, gerçek zamanlı olarak çalışan üç temel bileşeni içerir: poz, yüz ve el. Bu bileşenler arasında minimum bellek aktarımıyla çalışan çıkarım arka uçlarına sahiptir. MediaPipe Holistic, kalite ve hız değişimlerine bağlı olarak bu üç bileşenin birbiriyle değiştirilebilirliğini desteklemektedir. Çığır açan 540'tan fazla anahtar noktayı birleştiren bu üç bileşen, 33 poz, 21 el noktası başına ve 468 yüz işareti ile birleşik bir topoloji sunar. Mobil cihazlarda neredeyse gerçek zamanlı performansa ulaşabilen MediaPipe Holistic, hem mobil (Android, iOS) hem de masaüstü cihazlarda kullanılabilir. MediaPipe Holistic, MediaPipe ailesinin bir parçası olarak sunulmaktadır ve ayrıca erişimi kolaylaştırmak amacıyla araştırma (Python) ve web (JavaScript) için yeni kullanıma hazır API'leri de içermektedir. Bu gelişmiş özellikler, MediaPipe'in gücünü ve esnekliğini daha geniş bir kullanıcı kitlesiyle buluşturmayı hedeflemektedir (Grishchenko ve Bazarevsky, 2020). MediaPipe, çeşitli dillerde çalıştırılabilmesi sayesinde farklı platformlarda geliştirme yapılmasına imkân tanır. Android, iOS, Java Script, Python ve C++ gibi diller ve platformlar üzerinde çalıştırılabilen MediaPipe, 2022 itibariyle tüm modüllerin tüm platformlarda bulunmadığı bir durumu yansıtmaktadır. MediaPipe kütüphanesindeki bazı modüller, belirli platformlar veya dillerle sınırlı olabilir. Örneğin, "Face Detection" gibi

genel özellikler tüm platformlarda ve birçok programlama dilinde kullanılabilecek genel amaçlı modüllerdir. Ancak, "Hair Segmentation" gibi özel bir özellik belirli bir platforma veya programlama diline özgü olabilir. Bu durumda, "Hair Segmentation" modülü sadece Android platformunda ve C++ dilinde kullanılabilir (Sabaz, 2022). MediaPipe bünyesinde bulunan Hand Landmarker'ın görevi görüntüde veya videoda bulunan ellerin yer işaretlerini algılamayı sağlar. Ellerdeki anahtar noktaları göstermek ve görsel efektler oluşturmaya yardımcı olur. MediaPipe 'ta bulunan el yer işareti algılama modeli, avuç içi algılama modeli tarafından tanımlanan kırpılmış el görüntüsü üzerindeki belirli el yer işaretlerini tanımlar. Bir yer işareti model paketi, tespit edilen el bölgelerinde 21 el-boğum koordinatlarının anahtar nokta lokalizasyonunu tespit eder (Samaan vd., 2022).



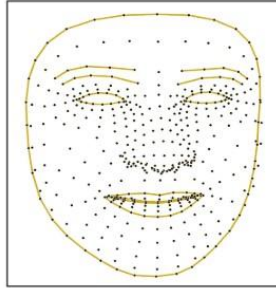
Şekil 3.18. MediaPipe el anahtar noktaları (Samaan vd., 2022)

MediaPipe Pose Landmarker, görüntü veya videodaki insan vücudunun anahtar noktalarını çıkarmayı sağlar. Bu sayede vücudun istenilen bölgesini belirlemek, duruş analizini yapmak ve hareketleri kategorize etmeyi yardımcı olmaktadır. Pose Landmarker, poz anahtar noktalarını tespit etmek için bir sıra işlem gerçekleştirir. İlk olarak model görüntü içindeki insan varlığını algılar, sonraki işlemde algılanan insan pozuna ait anahtar noktaları çıkarır. MediaPipe poz yer işareti modeli, şekil 3.19'de gösterildiği üzere vücut bölümlerinin yaklaşık konumunu temsil eden 33 anahtar noktadan oluşmaktadır (Samaan vd., 2022).



Şekil 3.19. Poz anahtar noktaları (Samaan vd., 2022)

MediaPipe yüz dedektörü ile video veya görüntüdeki yüzlerin anahtar noktalarını çıkarmayı sağlar. MediaPipe yüz dedektöründe yüzün çevresi, gözler, dudaklar ve kaşlar yer işaretlerini birleştiren çizgilerle temsil edilirken, geriye kalan yüzün her noktasına karşılık gelen 468 anahtar nokta bulunmaktadır (Samaan vd., 2022).



Şekil 3.20. Yüz anahtar noktaları (Samaan vd., 2022)

MediaPipe, yüz bölgeleriyle ilgili detaylı tahminler yapabilen özel modeller sunmaktadır. Yüz Landmark Modeli, yüzdeki belirli noktaları (landmark) tespit ederken, bu noktaların dudaklar, gözler ve iris çevresindeki detaylı bölgeleri içeren özel bir model daha fazla hesaplama maliyetiyle yüksek doğrulukla tahmin edebilir. Bu detaylı modeller, özellikle yüz analitiği uygulamalarında ve kullanıcı arayüzü tasarımlarında önemli bir rol oynayabilir. Örneğin, yüz ifadelerinin daha ayrıntılı bir şekilde analiz edilmesi veya yüz bazlı etkileşimlerin geliştirilmesi gibi senaryolarda bu modeller değerli olabilir. Bu tür modeller, görüntü işleme ve makine öğrenimi uygulamalarında MediaPipe kütüphanesini kullanarak geliştiricilere geniş bir yelpaze sunmaktadır (Sabaz, 2022).

3.7. Model Performans Değerlendirme Metrikleri

Makine öğrenmesi algoritmalarının başarılarını değerlendirmek metrikler ve bu metriklere dayalı olarak model performansını değerlendirmek amacıyla çeşitli parametreler kullanılmaktadır. Accuracy (doğruluk), genellikle tercih edilen bir metrik olmasına rağmen, veri seti dağılımının dengesiz olduğu durumlarda yanıltıcı olabilir. Bu nedenle, kesinlik (precision), duyarlılık (recall) ve bu değerlerin harmonik ortalaması olan F1-score gibi alternatif parametreler de kullanılmaktadır. Karışıklık matrisi, modelin sınıflandırma performansını değerlendirmek için gerçek ve tahmini değerler arasındaki ilişkiyi gösteren bir matristir. Bu matris, doğru ve hatalı tahminleri görselleştirerek modelin performansını anlamak için kullanılır. (Uğurlu vd., 2023).

Tablo 3.2. Karışıklık Matrisi Gösterimi

		Mevcut Durum	
		Pozitif	Negatif
Tahmin	Pozitif	TP	FP
	Negatif	FN	TN

True Positive (TP): Modelin pozitif olarak tahmin ettiği durumların gerçekte de pozitif olduğu durumları ifade eder. Yani, model doğru bir şekilde pozitif tahmin yapmıştır.

True Negative (TN): Negatif değer tahmininin olduğu durumların, gerçekte de negatif olduğu durumları ifade eder. Yani, model doğru bir şekilde negatif tahmin yapmıştır.

False Positive (FP): Modelin negatif olması gereken bir durumu pozitif olarak tahmin ettiği durumları ifade eder. Yani, model yanlış bir şekilde pozitif tahmin yapmıştır.

False Negative (FN): Modelin pozitif olması gereken bir durumu negatif olarak tahmin ettiği durumları ifade eder. Yani, model yanlış bir şekilde negatif tahmin yapmıştır.

Bu terimler, modelin performansını ölçen çeşitli metriklerin hesaplanmasında kullanılır. Örneğin, hassasiyet (precision), duyarlılık (recall), F1 skoru gibi metrikler, TP, TN, FP

ve FN değerlerine dayanarak hesaplanır. Bu denklemler, bir sınıflandırma modelinin performansını değerlendirmek için kullanılan temel metriklerdir.

Accuracy (Doğruluk): Accuracy, modelin doğru tahminlerinin toplam tahminlere oranını ölçer. İdeal durumda, bu değer 1'e yaklaşmalıdır.

$$\text{Doğruluk} = (TN+TP)/(TN+TP+FN+FP) \quad (3.13)$$

Recall (Duyarlılık): Recall, gerçekte pozitif olan tüm durumların ne kadarını doğru bir şekilde pozitif tahmin ettiğimizi ölçer.

$$\text{Duyarlılık} = TP/(TP+FN) \quad (3.14)$$

Precision (Kesinlik): Precision, modelin pozitif olarak tahmin ettiği durumların ne kadarının gerçekten pozitif olduğunu ölçer.

$$\text{Kesinlik} = TP/(TP+FP) \quad (3.15)$$

F-Score: F-Score, Kesinlik ve Duyarlılığın harmonik ortalamasıdır. Dengeleme metriği olarak kullanılır.

$$\text{F-Score} = 2 * \text{Duyarlılık} * \text{Kesinlik} / (\text{Duyarlılık} + \text{Kesinlik}) \quad (3.16)$$

Bu metrikler, bir sınıflandırma modelinin performansını çeşitli açılardan değerlendirmemize yardımcı olur. İdeal durumda, yüksek accuracy, recall, precision ve F-score değerleri istenir, ancak bu metrikler arasında bir denge kurmak önemlidir.

4. ARAŞTIRMA BULGULARI

Gerçek zamanlı işaret dil tanıma sistemleri, işitme engelli bireylerin iletişimini kolaylaştırmada önemli bir rol oynamaktadır. Bu sistemler, el, vücut hareketleri, yüz ifadeleri ve genel hareketler gibi çoklu verileri değerlendirerek işitme engelli kişilere yönelik yardımcı teknolojik araçlardır. Gerçek zamanlı işaret dil tanıma, işitme engelli bireylerin doğal bir şekilde iletişim kurmalarını ve eğitim ile iş hayatında daha etkin bir şekilde yer almalarını sağlar. Ayrıca dijital içeriklere erişimi kolaylaştırarak, internet ve dijital medyanın sunduğu imkanlardan faydalanmalarına yardımcı olabilirler. Bu teknoloji, işitme engelli bireylerin toplumsal, eğitimsel ve profesyonel açıdan daha etkin bir rol oynamalarına katkıda bulunmaktadır. Bu noktada yapay zekâ algoritmaları gerçek zaman işaret dili tanıma sistemlerinde oldukça fazla kullanılmaktadır.

Bu çalışmada, TİD'e özgü çok kullanılan kelimelere ait işaret dili hareketlerini tanımaya yönelik bir model geliştirilmiştir. Çalışmada kullanılan TİD kelime sayısı 17 olup, her bir kelime 5 denek tarafından 30 kez tekrar edilerek veri seti oluşturulmuştur. Daha sonra MediaPipe ile işaret noktaları çıkarılarak LSTM sınıflandırıcıya girilmiştir. LSTM sınıflandırıcı, 3 ardışık LSTM katman (çıkış boyutu sırasıyla 64, 128, 64) ve ardından 3 ardışık Dense katmandan (çıkış boyutu 64, 32, 17) oluşmaktadır. Yapılan analiz uygulamalarında için 80%-20% eğitim ve test veri seti oranları esas alınmıştır. LSTM sınıflandırıcıların başarıları 10 çapraz doğrulamaya göre elde edilmiş olup, aşağıda detaylı bir şekilde sunulmuştur. Sonuçlar her bir denek için ayrı ayrı kişi bağımlı ve tüm denekler bir arada olacak şekilde kişi bağımsız olarak iki ayrı bölümde sunulmuştur.

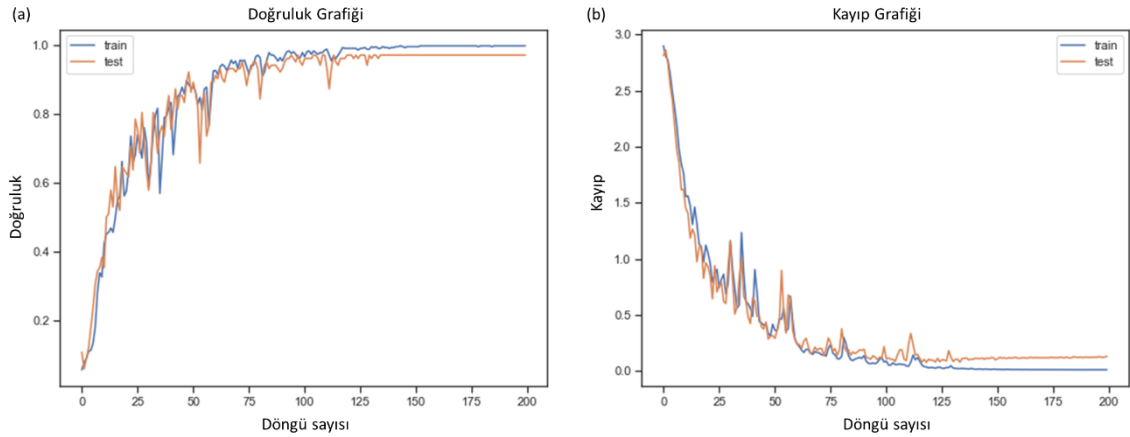
4.1. Kişi Bağımlı Sonuçlar

Kişi bağımlı sonuçlar ifadesi, bir modelin performansının değerlendirildiği durumu ifade eder ve bu değerlendirme denek (kişinin) özelliklerinden etkilendiği anlamını taşımaktadır. Bu tür bir analiz genellikle belirli bireylerin veya deneklerin üzerinde yapılan test sonuçlarını içerir. Bu bölümde deneklerin birbirinden bağımsız analiz sonuçları sunulmuştur. Aşağıdaki Tablo 4.1'de eğitim/train veri seti için her bir deneğe ait 10 çapraz doğrulamaya göre başarı ölçütlerinin değerleri sunulmuştur.

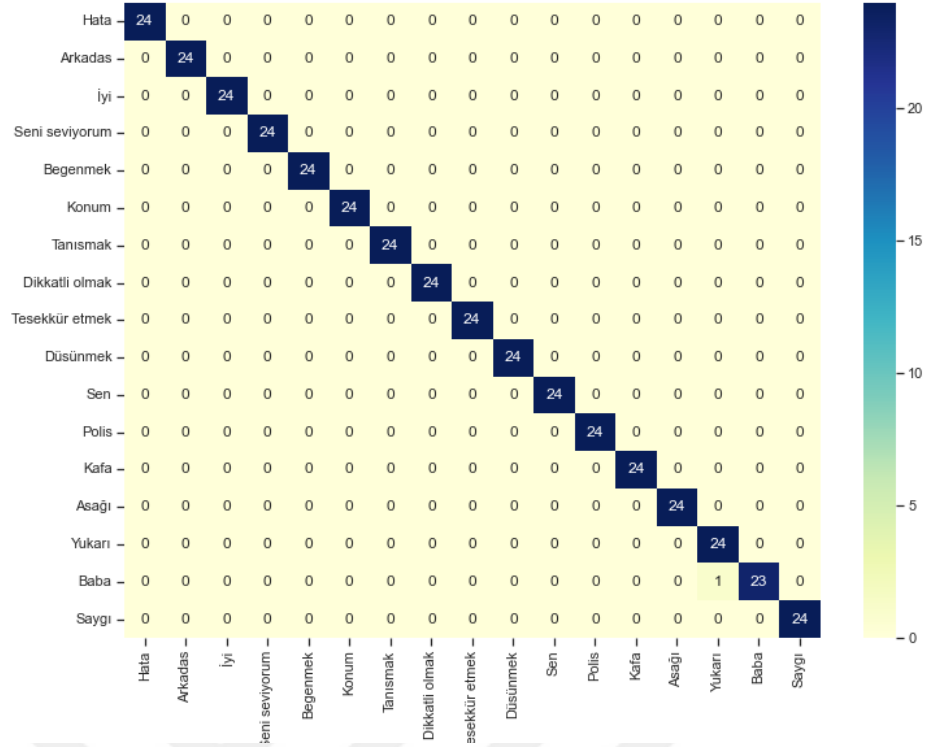
Tablo 4.1. Kişi bağımlı eğitim başarı ölçütleri

	<i>Doğruluk</i>	<i>Duyarlılık</i>	<i>Hassasiyet</i>	<i>F1-Skor</i>
<i>Denek 1</i>	98,75	98,79	98,75	98,75
<i>Denek 2</i>	98,31	98,37	98,31	98,31
<i>Denek 3</i>	98,77	98,79	98,77	98,76
<i>Denek 4</i>	98,75	98,86	98,75	98,75
<i>Denek 5</i>	99,19	99,24	99,19	99,19

Tablo 4.1’de belirtildiği üzere eğitim sonuçları kıyaslandığında elde edilen en yüksek doğruluk değeri olan %99,19 Denek 5’e ait veri setinde gözlemlenmiştir. Aşağıda Şekil 4.1’de bu deneğe ait 10 çapraz doğrulama arasında en yüksek doğruluk (%100) değerine sahip denemenin eğitim ve test Doğruluk (Accuracy) ve Kayıp (Loss) grafikleri görülmektedir.

**Şekil 4.1.** Denek 5’e ait Doğruluk ve Kayıp grafikleri.

Şekil 4.2’de eğitim veri seti için Denek 5’e ait 10 çapraz doğrulama arasında en yüksek doğruluk değerine (%100) sahip denemenin karmaşıklık matrisi görülmektedir.



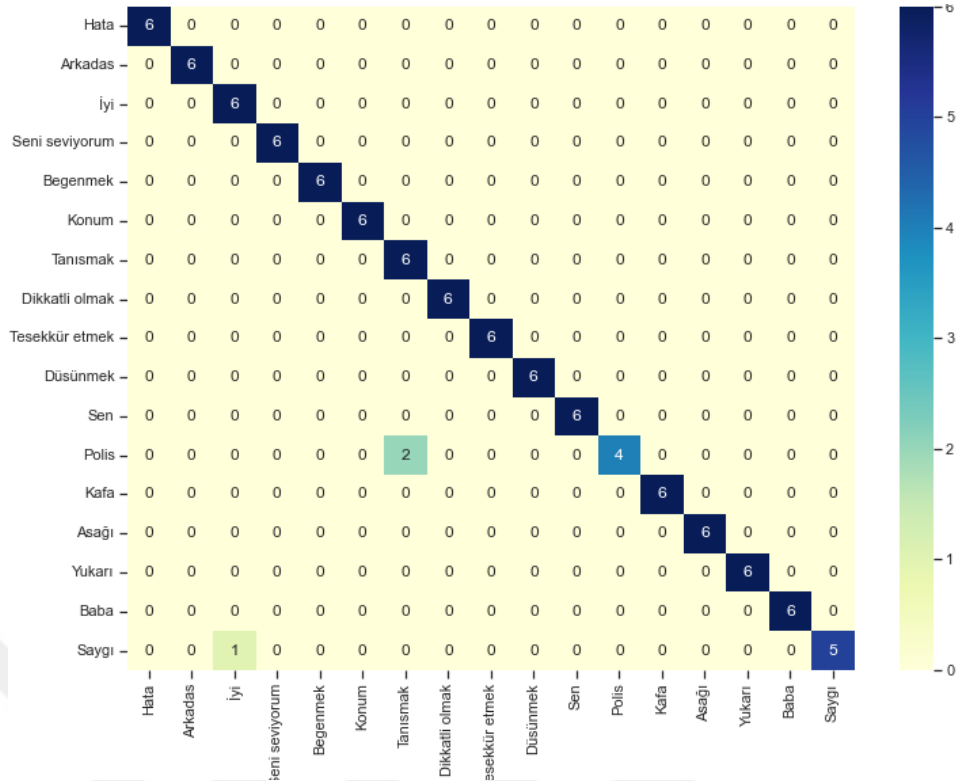
Şekil 4.2. Denek 5 için eğitim veri seti karmaşıklık matrisi.

Aşağıdaki Tablo 4.2’de test veri seti için her bir deneğe ait 10 çapraz doğrulamaya göre başarı ölçütlerinin değerleri sunulmuştur.

Tablo 4.2. Kişi bağımlı test başarı ölçütleri

	<i>Doğruluk</i>	<i>Duyarlılık</i>	<i>Hassasiyet</i>	<i>F1-Skor</i>
Denek 1	91,76	92,95	91,76	91,66
Denek 2	91,18	92,12	91,18	91,05
Denek 3	82,45	84,05	82,45	82,1
Denek 4	94,9	95,54	94,9	94,82
Denek 5	94,71	95,41	94,71	94,67

Tablo 4.2’de belirtildiği üzere test sonuçların kıyaslandığında elde edilen en yüksek doğruluk değeri %94,71 ile Denek 5’e ait veri setinde bulunmuştur. Şekil 4.3’te test veri seti için Denek 5’e ait 10 çapraz doğrulama arasında en yüksek doğruluk değerine (%97,06) sahip denemenin karmaşıklık matrisi görülmektedir.



Şekil 4.3. Denek 5 test karmaşıklık matrisi.

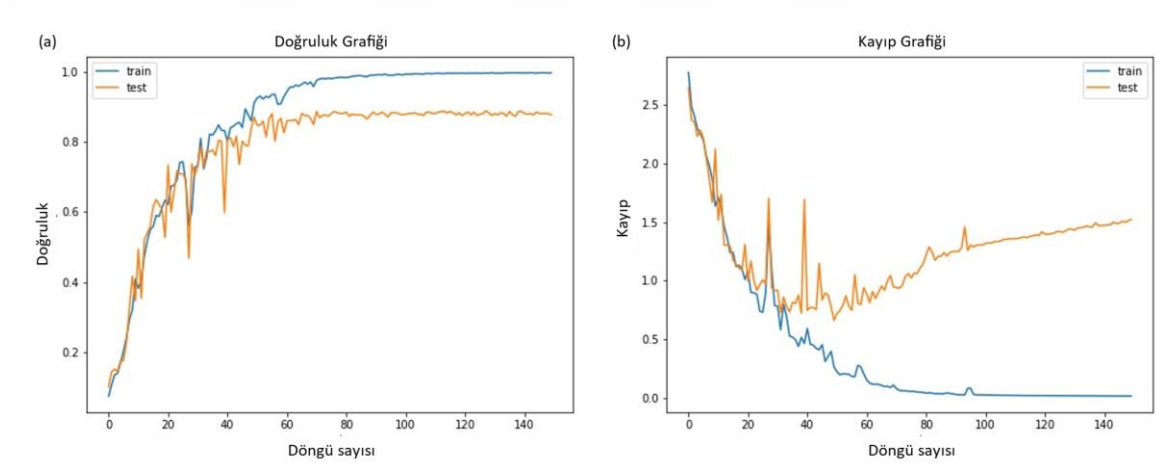
4.2. Kişi Bağımsız Sonuçlar

Bu bölümde tüm deneklerin bir arada yer aldığı kişi bağımsız analiz sonuçları sunulmuştur. Kişi bağımsız analiz sonuçları ifadesi, bir modelin performansının değerlendirildiği bir durumu ifade eder, bu değerlendirme denek (kişinin) özellikleriyle bağlantılı değildir. Kişi bağımsız analiz sonuçları, modelin genelleme yeteneğini değerlendirmek adına önemlidir. Model, eğitildiği veri setinden farklı bir veri seti veya farklı bireylerle karşılaştığında ne kadar etkili olduğunu görmek için kişi bağımsız testlere tabi tutulur. Aşağıdaki Tablo 4.3'te eğitim/train veri seti için 10 çapraz doğrulamaya ait her bir denemenin başarı ölçüt değerleri sunulmuştur. Çapraz doğrulama sonunda elde edilen ortalama doğruluk %98,22 olarak lede edilmiştir.

Tablo 4.3. Kişi bağımsız eğitim başarı ölçütleri.

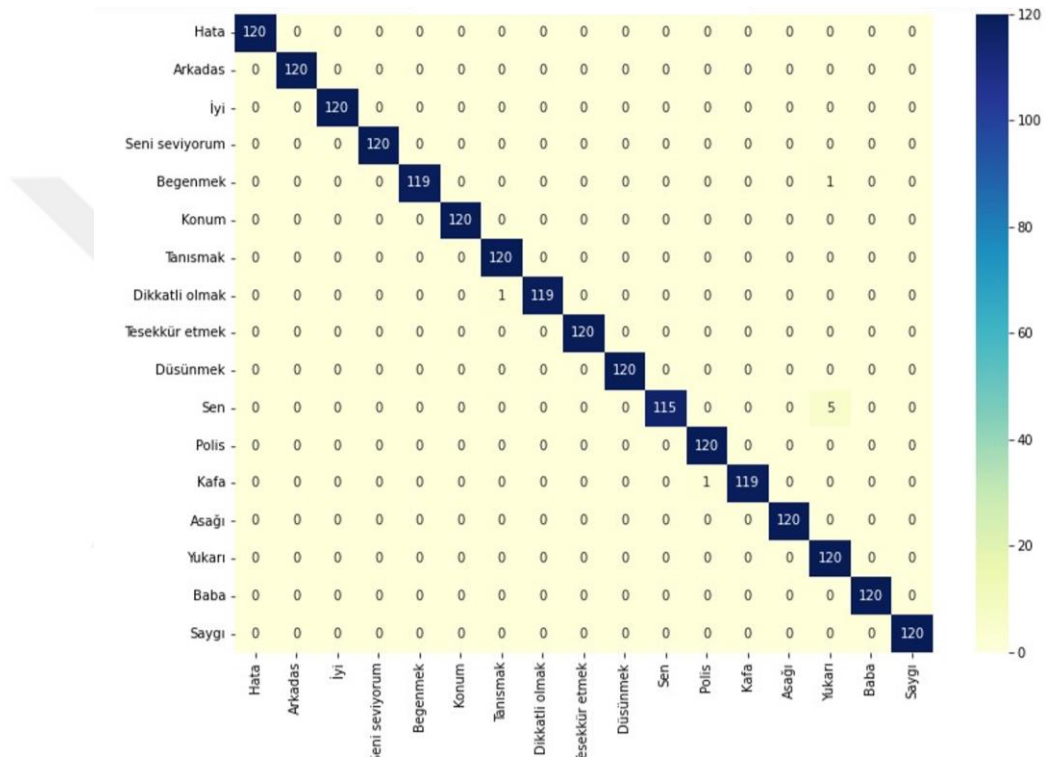
<i>S.No</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
<i>Deneme 1</i>	98,04	98,07	98,04	98,04
<i>Deneme 2</i>	91,72	91,83	91,72	91,72
<i>Deneme 3</i>	99,95	99,95	99,95	99,95
<i>Deneme 4</i>	96,86	96,91	96,86	96,85
<i>Deneme 5</i>	99,61	99,61	99,61	99,61
<i>Deneme 6</i>	99,46	99,47	99,46	99,46
<i>Deneme 7</i>	99,61	99,61	99,61	99,61
<i>Deneme 8</i>	97,65	97,73	97,65	97,66
<i>Deneme 9</i>	99,71	99,71	99,71	99,71
<i>Deneme 10</i>	99,61	99,62	99,61	99,61
<i>Ortalama</i>	98,22	98,25	98,22	98,22

Tablo 4.3'te yola çıkarak eğitim veri setleri için en yüksek doğruluk değeri %99,95 ile Deneme 3'te elde edilmiştir. Aşağıda Şekil 4.4'te bu denemeye ait Doğruluk (Accuracy) ve Kayıp (Loss) grafikleri görülmektedir.

**Şekil 4.4.** Deneme 3'e ait Doğruluk ve Kayıp grafikleri.

Bu bölümde sunulan kişi bağımsız analiz, modelin farklı denekler üzerindeki performansını değerlendirmek adına gerçekleştirilmiştir. Tüm deneklerin bir araya getirilmesi, genel model performansının bir özeti olarak adlandırılabilir. Tablo 4.3'te yer alan çapraz doğrulama denemeleri, eğitim veri seti üzerinde yapılan 10 farklı denemenin sonuçlarını içerir. Bu denemeler, modelin veri seti üzerinde farklı parçalara ayrılan alt küme üzerinde eğitilip test edilmesini anlamına gelir. Çapraz doğrulama, modelin farklı veri seti parçaları üzerinde nasıl performans gösterdiğini daha sağlam bir şekilde

değerlendirmeye yardımcı olur ve modelin genelleme yeteneğini ölçmede önemli bir araçtır. Çapraz doğrulama sonuçlarından elde edilen ortalama doğruluk değeri %98,22 olarak belirlenmiştir. Bu değer, modelin farklı alt küme kombinasyonlarında yüksek bir tutarlılıkla doğru tahminlerde bulunduğunu gösterir. Bu sonuçlar, modelin kişi bağımsız testlerde yüksek bir doğruluk elde ettiğini ve genel olarak güvenilir bir performans sergilediğini göstermektedir. Şekil 4.5'te eğitim veri seti için Deneme 3'e ait karmaşıklık matrisi görülmektedir.



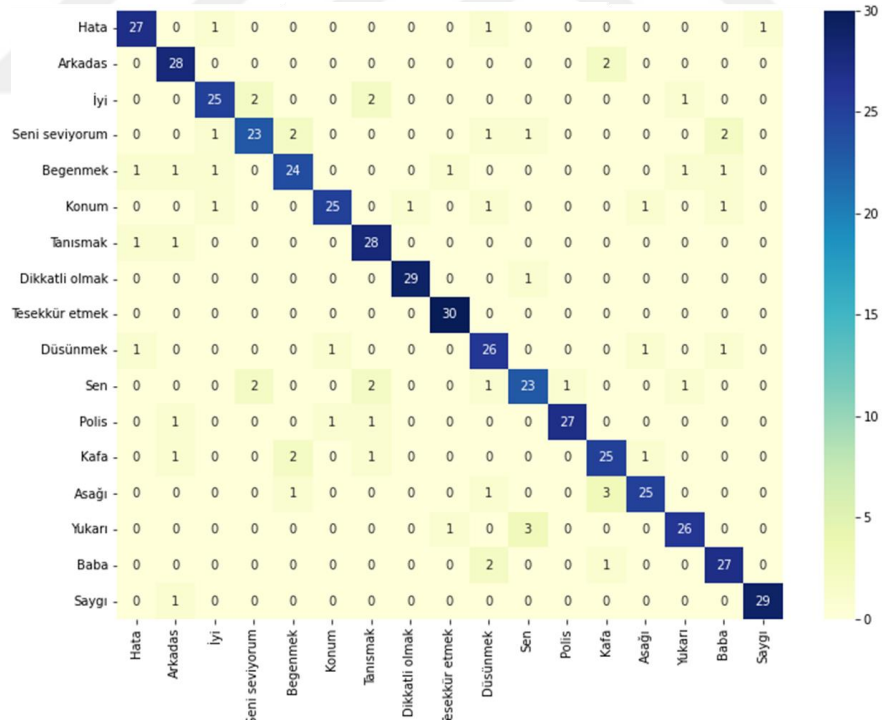
Şekil 4.5. Deneme 3 eğitim veri seti için karmaşıklık matrisi.

Aşağıdaki Tablo 4.4'te test veri seti için 10 çapraz doğrulamaya ait ayrı ayrı başarı ölçüt değerleri sunulmuştur. Çapraz doğrulama sonunda elde edilen ortalama doğruluk %84,75 olarak elde edilmiştir.

Tablo 4.4. Kişi bağımsız test başarı ölçütleri.

<i>S.No</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
<i>Deneme 1</i>	83,53	84,72	83,53	83,52
<i>Deneme 2</i>	82,55	82,93	82,55	82,53
<i>Deneme 3</i>	87,65	87,79	87,65	87,61
<i>Deneme 4</i>	81,96	82,1	81,96	81,92
<i>Deneme 5</i>	86,86	86,99	86,86	86,82
<i>Deneme 6</i>	85,88	86,37	85,88	85,87
<i>Deneme 7</i>	86,67	86,98	86,67	86,43
<i>Deneme 8</i>	82,35	82,83	82,35	82,32
<i>Deneme 9</i>	83,14	83,62	83,14	83,04
<i>Deneme 10</i>	86,86	87,26	86,86	86,85
<i>Ortalama</i>	84,75	85,16	84,75	84,69

Tablo 4.4'te gözlemlendiği üzere test veri sonuçlarından kıyaslandığında en yüksek doğruluk değeri %87,65 ile Deneme 3'te elde edilmiştir. Şekil 4.6'de test veri seti için Deneme 3'e ait karmaşıklık matrisi görülmektedir.

**Şekil 4.6.** Deneme 3 test veri seti için karmaşıklık matrisi.

5. TARTIŞMA

Teknoloji birçok birey için iletişimi, eğitimi ve günlük yaşamı kolaylaştırmada önemli bir rol oynasa da mevcut teknolojik çözümlerin çoğu işitme ve konuşma engelli bireyler düşünülerek tasarlanmamıştır. Bu eksik, onların topluma tam entegrasyonlarının önünde önemli bir engel teşkil etmektedir. Bu araştırma tezinde, bu topluluğun gereksinimlerine uyarlanmış teknolojinin yaratılmasına odaklanarak konu ele alınmıştır. Geleneksel görüntü işleme teknikleri kullanılan işaret dili modellerinin karmaşıklığı ve yüksek hesaplama gücü, göz önüne alındığında, bu çalışma ile geleneksel yöntemlerle işaret dili tanıma sistemlerine alternatif sunulmuştur. Çalışmada kullanılmak üzere hazırlanan TİD'e ait veri seti MediaPipe kütüphanesi kullanılarak oluşturulmuştur. MediaPipe kütüphanesinin karmaşık el hareketlerini doğru bir şekilde algılamada verimli bir araç olduğu, yapılan çalışma ve elde edilen verilerle ortaya konulmuştur. Literatürde sıklıkla kullanılan TİD'e ait 17 kelime 5 denek tarafından 30 defa tekrar edilerek bir veri seti oluşturulmuştur. Oluşturulan veri seti, sıralı verilerdeki uzun vadeli bağımlılıkları öğrenmede etkili olduğu kanıtlanmış LSTM'i mimarisinde eğitilmiş ve test edilmiştir. Kişi bağımlı ve kişi bağımsız olarak eğitim ve testte tabi tutulan veri seti sonuçları değerlendirildiğinde, sonuçların TİD kelimeleriyle uyumlu performans gösterdiği görülmektedir. Tablo5.1'de yer alan görsellerde eğitilmiş modele işaretler gösterildiğinde modelin işaretleri tahmin etme ve metne dönüştürme işlemi yer almaktadır. Kişi bağımlı ya da kişi bağımsız eğitim ve testler sonuçlarından yola çıkarak, ortaya konulan modelin TİD'e ait tüm kelimelere (işaretlere) etkili bir şekilde uygulanabilir olduğu görülmektedir.

Tablo 5.1. Modelin TİD'e ait işaretleri tahmin etme görselleri

Tanismak Arkadas Tanismak Polis ArıHata Yukari Hata Yukari Hata		Begenmek Sen Seni seviyorum Sen iy
Hata	Hata	Hata
Arkadas	Arkadas	Arkadas
iyi	iyi	iyi
Seni seviyorum	Seni seviyorum	Seni seviyorum
Begenmek	Begenmek	Begenmek
Konum	Konum	Konum
Tanismak	Tanismak	Tanismak
Dikkatli olmak	Dikkatli olmak	Dikkatli olmak
Tesekkür etmek	Tesekkür etmek	Tesekkür etmek
Dusunmek	Dusunmek	Dusunmek
Sen	Sen	Sen
Arkadaş	Hata	İyi



Kişi bağımlı eğitim başarı ölçütleri oldukça yüksek ve tutarlı bir performans sergilemektedir. Her beş denek için doğruluk, duyarlılık, hassasiyet ve F1-skorları %98'in üzerindedir. Özellikle, Denek 5'in %99,19 ile en yüksek başarı ölçütlerine sahip olması dikkat çekicidir. Bu bulgular, kullanılan eğitim verilerinin, modelin belirli işaretleri doğru bir şekilde tanıma yeteneğini artırdığını ve kişi bağımlı testlerde yüksek doğruluk ve performans elde ettiğini ortaya koymaktadır. Kişi bağımlı test sonuçlarına bakıldığında, ortalama doğruluk oranı %90'ın üzerinde olup, deneklerin duyarlılık, hassasiyet ve F1-skorları da tatmin edici düzeydedir. Özellikle Denek 4 ve Denek 5, yüksek doğruluk, duyarlılık, hassasiyet ve F1-skorları ile öne çıkmaktadır. Bu sonuçlar, oluşturulan modelin Türk İşaret Dili için etkili bir şekilde çalıştığını ve kullanıcı bağımlı testlerde güvenilir sonuçlar elde ettiğini göstermektedir. Ancak, Denek 3'ün düşük doğruluk oranı dikkat çekmektedir; bu durumun detaylı bir şekilde incelenmesi ve ihtiyaç olursa modelin iyileştirilmesi gerekmektedir. Kişi bağımsız eğitim başarı ölçütleri ortalama olarak, doğruluk, hassasiyet, duyarlılık ve F1-skoru %98 olarak ölçülmüştür. Bu, modelin kişi bağımsız eğitim verileri üzerinde genel başarı gösterdiğini ancak, Deneme 2'deki düşük performansından dolayı modelin zayıf noktalarında iyileştirmeler yapılması gerektiğini göstermektedir. Kişi bağımsız test başarı ölçütleri, modelin genelleme yeteneğini değerlendirmek için kullanılmıştır. Modelin doğruluk, hassasiyet, duyarlılık ve F1-skoru ortalama %84,75 olarak bulunmuştur. Bu, genelde modelin kişi bağımsız test verileri üzerinde başarılı bir performans sergilediğini gösterir. Ancak, bazı denemelerde daha düşük performanslar gözlemlenmiştir. Örneğin, Deneme 4'te doğruluk ve F1-skoru %81'in altındadır. Benzer şekilde, Deneme 2'de diğer denemelere kıyasla daha düşük bir performans görülmektedir. Bu düşük performans, deneklerin boyutları veya veri örnekleri üzerindeki farklılıklarından kaynaklı sonuçların değişebileceğine işaret etmektedir. Kişi bağımsız çalışma, daha büyük bir veri setiyle (2550 örneklem) çalışmış olmasına rağmen veri setinin büyüklüğü tek başına başarı oranını belirlemez; veri setinin

çeşitliliği ve temsil gücü de önemlidir. Kişi bağımlı çalışma, %94,71 başarı oranıyla kişi bağımsız çalışmadan daha yüksek bir doğruluk elde etmiştir. Bu durum, kişi bağımlı modelin, belirli kelimeleri içeren küçük bir veri setinde daha iyi performans gösterdiğini doğrulamaktadır. Karşılaştırma amacıyla bu tez çalışmasında önerilen model ile literatür yapılmış benzer çalışmalara ilişkin başarı değerleri Tablo5.2’de sunulmuştur.

Tablo5.2. İşaret tanıma sistemleri modellerinin karşılaştırması.

<i>Kaynak</i>	<i>Yaklaşım (Metot, Veri seti toplama aracı ve Kullanılan işaret Dili)</i>			<i>Veri kümesi</i>	<i>Sınıflandırıcı</i>	<i>%Başarı</i>
Saengsri vd.,2012	Dinamik Tanıma Sistemi	Elektronik Eldiven	Tayland işaret dili	20 kelime 1000 örneklem	YSA	%94
Tangsukant vd., 2014	Hough Dönüşümü Tekniği	Elektronik Eldiven	ASL	24 harf 2580 örneklem	YSA	%95
Wang vd., 2006	Gizli Markov Modelleri	Elektronik Eldiven Ve Leap Motion Sensör	Amerikan işaret dili	26 Harf	YSA	%95
Sarawat vd., 2015	Olasılıksal Bir Model	Elektronik Eldiven	Amerikan işaret dili	40 sözcük 6000 örneklem	YSA	%95
Porfirio vd., 2013	3D Kafes Ve 2D Gösterimi	Kinect Sensör	Brezilya işaret dili	61 farklı RGBD video görüntüsü	SVM	%96
Sun C. vd., 2013	Gradyan Odaklı Histogram Tanımlayıcı	Kinect Sensör	Amerikan işaret dili	73 kelime 1971 örneklem	SVM	%85
Hisham ve Hamouda, 2019	Dinamik Hareket Tanıma Sistemi	Kinect Sensör	Arap işaret dili	40 kelime 1200 örneklem	KNN, SVM	%89
Özcan ve Baştürk, 2020	Hiperparametre Optimizasyonu Yöntemi	Video Kamera	Türk İşaret dili	25 kelime 13186 örneklem	CNN	%93
Pacal ve Alaftekin, 2023	Derin Öğrenme Yaklaşımları	Video Kamera	Türk İşaret dili	10 rakam 1500 örneklem	VG ResNet MobileNet DenseNet EfficientNe	%97
Tesmer ve Ahmed, 2020	Derin Öğrenme Mimarisi	Video Kamera	Bangladeş işaret dili	37 harf 3219 örneklem	CNN	%99
Halder ve Tayede, 2021	İnsan poz Tahmin Sistemi	MediaPipe Kütüphanesi	Amerikan(alfabe) Hint(alfabe) İtalyan(alfabe) ASL (rakam) Türk (rakam)	26 ASL harfi (156000 örn),24 Hint harfi (4972 örn)	SVM, KNN, Random Forest, Decision Tree, YSA,	%99

Subramani an vd., 2022	İnsan Poz Tahmin Sistemi	MediaPipe Kütüphanesi	Hint işaret dili	13 kelime 1662 örneklem	RNN, GRU MOPGRU LSTM, BiGRU, BiLSTM	%95
Sundar ve Bagyamma I,2022	İnsan Poz Tahmin Sistemi	MediaPipe Kütüphanesi	Amerikan işaret Dili	26 harf	LSTM	%99
Kırıcı vd., 2022	İnsan Poz Tahmin Sistemi	MediaPipe Kütüphanesi	Türk işaret dili	29 harf 870 örneklem	LSTM	-
Dewanjee, 2021	İnsan Poz Tahmin Sistemi	OpenPose Kütüphanesi	Bangladeş işaret dili	10 kelime 3456 örneklem	LSTM, RNN, KNN, SVM, Gaussian Naive Bayers	%96
Ko vd., 2019	Diziden Diziye Çeviri Modeli	OpenPose Kütüphanesi	Kore İşaret dili (KETI dataset)	10 cümle 14.672 örneklem	RNN, LSTM, GRU	%93
Santhaling am vd., 2020	Videodan Anahtar Nokta Çıkarımı	OpenPose Kütüphanesi	GMU-ASL51	51 kelime 4333 örneklem	LST	%99
Kim ve Jung, 2021	İnsan Poz Tahmin Sistemi	OpenPose Kütüphanesi	Kore işaret dili	10 sözcük, 13200 örneklem	CNN	%81
Bu çalışma (Kişi bağımlı)	İnsan Poz Tahmin Sistemi	MediaPipe Kütüphanesi	Türk İşaret dili (en çok kullanılan kelimeler)	17 sözcük 510 örneklem	LSTM	%94,71
Bu çalışma (Kişi bağımsız)	İnsan Poz Tahmin Sistemi	MediaPipe Kütüphanesi	Türk İşaret dili (en çok kullanılan kelimeler)	17 sözcük 2550 örneklem	LSTM	%87,65

TİD kelimeleri kullanılarak yapılan çalışmayı literatürdeki diğer çalışmalardan ayıran en önemli özelliği modelin performansını değerlendirmek için hem kişi bağımlı hem de kişi bağımsız değerlendirmeleri içermesidir. Bu, modelin spesifik kişilere özgü performans değerlendirilmesinin yanı sıra tüm deneklerden oluşan veri setini de değerlendirmeye olanak tanır. Çalışma, oluşturulan veri setini farklı denekler ve ortamlarla zenginleştirilerek gerçekleştirilmiştir. Bu, modelin farklı kişilerin ve çeşitli ortam koşullarının etkisi altında nasıl performans gösterdiğini anlamak açısından önemlidir. Bu faktörler, modelin genelleme yeteneğini daha gerçekçi bir şekilde değerlendirmesine imkân sağlar. Hazırlanan veri seti, çeşitli deneklerin katılımıyla zenginleştirilmiş olup, modelin farklı vücut yapılarına, hareket tarzlarına ve çeşitli kullanıcı profillerine karşı etkinliğini değerlendirmek için geniş bir yelpazede bilgiler içermektedir. Daha geniş bir veri seti, modelin çeşitli senaryolara ve kullanıcı gruplarına

uyum sağlama yeteneğini daha iyi gösterebilir. Böyle bir veri setinin oluşturulması, çalışmanın sadece belirli kişilere özgü değil, aynı zamanda farklı kişiler ve çevresel koşullar altında da etkili olup olmadığının daha kapsamlı bir şekilde değerlendirmesine olanak tanıyacaktır.



6. SONUÇ

Bu araştırma, işitme ve konuşma engelli bireylerin teknoloji kullanımında karşılaştıkları zorluklara odaklanarak, özellikle işaret dili tanıma sistemleri üzerinde yeni bir bakış açısı sunmayı amaçlamıştır. Geleneksel görüntü işleme tekniklerine dayalı işaret dili modellerinin karmaşıklığı ve yüksek hesaplama gücü gereksinimi, bu bireylerin günlük yaşamlarına entegre olmalarını zorlaştırmaktadır. Bu kapsamlı araştırma tezi, işitme ve konuşma engelli bireyler için özel olarak tasarlanmış bir teknoloji yaklaşımını ele almaktadır. Bu kapsamda yapılan çalışmada, MediaPipe kütüphanesi ve LSTM mimarisi kullanarak geliştirilen alternatif bir yöntemle, işitme ve konuşma engelli bireyler için daha erişilebilir bir teknoloji önerilmektedir. İki önemli odak noktası bulunan çalışmada, TİD üzerinde çalışan bir İnsan Poz Tahmin Sistemi ve bu sistemin kişi bağımlı ve kişi bağımsız olarak değerlendirilmesi yer almaktadır. Çalışmada kullanılan veri seti, MediaPipe kütüphanesi aracılığıyla hazırlanmış ve geleneksel görüntü işleme teknikleri yerine derin öğrenme temelli bir yaklaşımla işaret dili tanıma sistemlerine alternatif sunmuştur. Elde edilen sonuçlar, kişi bağımlı eğitim ve test verilerinde yüksek doğruluk ve performans gösteren bir model ortaya koymaktadır. Bu, belirli işaretleri doğru bir şekilde tanıma yeteneğinin kişisel veri setleri üzerinde geliştirildiğini göstermektedir. Özellikle, Denek 5'in %99,19 başarı oranıyla dikkat çekici bir performans sergilemiştir. ancak, kişi bağımsız test sonuçları, genel başarı oranlarının %84,75 olduğunu göstermektedir. Bazı denemelerde düşük performans gözlemlenmiş olup, bu durumun detaylı bir şekilde incelenmesi ve modelin iyileştirilmesi gerekliliğini göstermektedir. Elde edilen bulgular, işitme ve konuşma engelli bireylerin işaret dilini tanımada geleneksel yöntemlere alternatif bir çözüm sunmanın olasılığını kanıtlamaktadır. Bu çalışmayla, işitme ve konuşma engelli bireylerin teknolojiye entegrasyonunu artırmak adına var olan çalışmalara ek katkı sunacağı varsayılmakta ve daha da geliştirilmesi gereken noktalarının olduğu düşünülmektedir.

Gelecekte, bu metodolojinin geniş bir kullanıcı kitlesi üzerinde test edilmesi ve geri bildirim alınarak geliştirilebileceği öngörülmektedir. Daha büyük ve çeşitli bir veri seti kullanımı, modelin farklı kişilere ve çeşitli çevresel koşullara nasıl tepki verdiğini öğrenmenin yanı sıra, modelin performansını geliştirmede de yardımcı olacaktır. Kişiden bağımsız eğitim ve test sonuçlarından yola çıkarak, var olan veri setindeki deneklerin boyutları veya özellikleri üzerindeki farklılıkları dengelemek için yöntemler kullanılması, modelin farklı boyuttaki denekler karşısında dengeli bir performans sergilemesine

yardımcı olacaktır. Gerçek dünyada test etmek adına çeşitli test ortamlarının oluşturulması örneğin, farklı aydınlatma koşulları, arka planlar veya çevresel gürültü seviyeleri gibi faktörleri dikkate alarak modelin performansının geliştirilmesi sağlanabilir ve modelin gerçek zamanlı bir uygulamada nasıl performans göstereceğini değerlendirebilir. Gerçek zamanlı uygulama durumları, modelin pratikteki kullanılabilirliği hakkında daha fazla bilgi sağlayacaktır. Modelin kullanılabilirliği ve kullanıcı dostu olma özelliklerini değerlendirmek adına doğrudan kullanıcılarla etkileşimde bulunması ve geri bildirim alınması, modelin kullanıcı ihtiyaçlarına daha iyi uyum sağlamasına yardımcı olabilir. Ayrıca Türk İşaret Dili dışında farklı işaret dillerini içeren veri setleri kullanarak modelin dil ve kültür bağlamında nasıl performans gösterdiğini değerlendirilebilir. Diğer taraftan, günümüzde bireylerin genellikle mobil cihazları sıkça kullanmaları nedeniyle, işaret dili tanıma sistemlerinin mobil uygulamalara entegrasyonu ve taşınabilir cihazlarda etkili bir şekilde çalışabilme yetenekleri üzerine çalışmalar gerçekleştirilebilir. Bu geliştirmeler, modelin daha geniş bir kullanıcı kitlesi için daha etkili ve kullanıcı dostu hale getirilmesine katkı sağlayabilir. Sonuç olarak, bu çalışma, işitme ve konuşma engelli bireyler için daha etkili ve kullanıcı dostu bir teknoloji çözümü sunma amacını taşımaktadır. Elde edilen sonuçlar, modelin belirli kişilere özgü ve genel bir kullanım durumunda başarılı olabileceğini göstermektedir.

KAYNAKLAR

- Akalın, Ş. H. (2013). Türk İşaret Dili. *Yeni Türkiye Dergisi* (Türkçe Özel Sayısı), 55, 1496-1502.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, 1-74.
- Arık, E. (2013). Türk İşaret Dili'nde Sınıflandırıcılar Üzerine Bir Çalışma. *Bilig Dergisi*, (67), 1-24.
- Arık, E. (2016). Türk İşaret dili araştırmaları. E. Arık (Ed.), *Ellerle Konuşmak* (s. 1-11). *Koç Üniversitesi Yayınları*.
- Aşkın, D., Mamızadeh, A., İskender, İ. (2013). Farklı Yapay Sinir Ağları Yöntemlerini Kullanarak Kuru Tip Transformator Sargısının Termal Analizi. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 26(4).
- Arslan, S., Aydemir, F. (2021). Covid-19 Pandemi Sürecinde Çocukların El Yıkama Alışkanlığının Nesnelerin İnterneti Tabanlı Sistem ile İzlenmesi. *Kütahya Dumlupınar Üniversitesi Mühendislik Bilimleri ve Araştırmaları Dergisi*, 3(2), 161-168.
- Bayrak, S. (2009). İşaret dilinin bilgisayarlı yorumlanması (Yüksek Lisans Tezi), *Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, 25-30
- Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, 2(1), 1-127.
- Birk, H., Madsen, C. B., Moeslund, T. B. (1997). Real-time recognition of hand alphabet gestures using principal component analysis. In *Proceedings of the Scandinavian conference on image analysis* (Vol. 1, pp. 261-268). Proceedings published by various publishers.
- Chollet, F. (2021). Deep learning with Python. Second Edition, *Shelter Island*, 55-60
- Çakır, H., Çetin, Ş., Abidin, B. A. Ş. (2013). İşitme engellilere yönelik dinamik web sayfasının geliştirilmesi. *Bilişim Teknolojileri Dergisi*, 6(2), 1.
- Çalışkan, D., Demir, Ö. (2022). Derin Öğrenme Yöntemleri ile Şüpheli Davranış Tespiti. *International Periodical Of Recent Technologies In Applied Engineering*, 3(1), 28-43.
- Çaylı, A. (2020). Engelliler için Türk işaret dilini öğrenen ve sınıflandıran duyuşal eldiven tasarımı ve uygulaması (Yüksek Lisans Tezi), *Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü*, Afyon, 10-12

- Çınar, A., Erkuş, M., & Tuncer, T. (2023). Türk İşaret Dilindeki Harflerin Tespiti İçin Derin Öğrenme Tekniğinin Kullanımı. *In International Conference on Scientific and Innovative Studies*, Konya, 360-365
- Dewanjee, T., Nuder, A., Malek, M. I., Nanjiba, R., & Rahman, A. A. (2021). Recognition of Bangladeshi sign language from 2D videos using openpose and LSTM based RNN (Doctora Tezi), *Department of Computer Science and Engineering Brac University Brac University*, Bangladeş, 10-71
- Dikyuva, H., Makaroğlu, B., Arık, E. (2015). Türk İşaret Dili dilbilgisi kitabı, *Aile ve Sosyal Politikalar Bakanlığı*, Ankara, 40-55
- Duran, F., & Alper, K. A. Y. A. (2018). Duruş ve Hareket Algılama Teknolojileri: Stereo, Time Of Flight ve Yapısal Işık Algılayıcılar. *Bilişim Teknolojileri Dergisi*, 11(1), 57-73.
- Erkuş, M. (2020). Türk işaret dilinde kelime tabanlı derin öğrenme uygulaması (Yüksek Lisans Tezi), *Başkent Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, 10-18
- Erten, N., Arıcı, H. (2022). İşaret Dilinin Tarihi Serüveni ve Türk İşaret Dili. *Afyon Kocatepe Üniversitesi Sosyal Bilimler Dergisi*, 24(1), 1-14.
- Eryiğit, C. (2017). Yazılı Türkçe Dilinden Türk İşaret Diline (TİD) Makine Çeviri Sistemi (Yayımlanmamış Doktora Tezi), *İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul, 5-32
- Ezel, E. (2018). Derin Öğrenme Yöntemi Kullanılarak Görüntü-Tabanlı Türk İşaret Dili Tanıma (Doctoral Tezi), *Selçuk Üniversitesi Fen Bilimleri Enstitüsü*, Konya, 5-10.
- Gamra, M. B., & Akhloufi, M. A. (2021). A review of deep learning techniques for 2D and 3D human pose estimation. *Image and Vision Computing*, 114, 104282.
- Görür, S., (2020). Yapay Sinir Ağları [online], Erişim adresi: <https://gorursevim.medium.com/yapay-sinir-a%C4%9F%C4%B1-ccf807f2cf82>, [Erişim tarihi: 18 Temmuz 2023]
- Grishchenko, I., Bazarevsky, V. (2020). MediaPipe Holistic Simultaneous Face, Hand and Pose Prediction, on Device [online], Erişim adresi: <https://ai.googleblog.com/2020/12/mediapipe-holistic-simultaneous-face.html>, [Erişim tarihi: 18 Haziran 2023].
- Halder, A., & Tayade, A. (2021). Real-time vernacular sign language recognition using mediapipe and machine learning. *Journal homepage: www. ijrpr. com* ISSN, 2582, 7421.
- Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504-507.
- Hisham, B., & Hamouda, A. (2019). Arabic Dynamic Gesture Recognition Using ClassifierFusion. *Journal Of Advances In Information Fusion*, 14 (1), s. 66-84.

- Iwai, Y., Watanabe, K., Yagi, Y., & Yachida, M. (1996, October). Gesture recognition by using colored gloves. In *1996 IEEE International Conference on Systems, Man and Cybernetics. Information Intelligence and Systems* (Cat. No. 96CH35929) (Vol. 1, pp. 76-81). IEEE.
- İlkbaşaran, D. (2016). Türkiye’deki Sağır Gençlerin İletişim Alışkanlıkları Ve Türk İşaret Dili’nin toplumsal Dilbilimi Açısından İncelenmesi. E. Arık (Ed.), *Ellerle Konuşmak: Türk İşaret Dili Araştırmaları, Koç Üniversitesi Yayınları*, İstanbul, 411–443.
- İlyas, Ö. (2020). Uzun Kısa Dönem Bellek Ağlarını Kullanarak Erken Aşama Diyabet Tahmini. *Bandırma Onyediy Eylül Üniversitesi Mühendislik Bilimleri ve Araştırmaları Dergisi*, 2(2), 50-57.
- Karaca, M. F. (2018). Üç Boyutlu Sanal Model İle Türk İşaret Dili Simülasyonu. (Yayımlanmamış Doktora Tezi), *Karabük Üniversitesi Fen Bilimleri Enstitüsü*, Karabük, 12-16.
- Karaca, M. F., & BAYIR, Ş. (2018). Türk işaret dili incelemesi: iletişim ve dil bilgisi. *Ulusal Eğitim Akademisi Dergisi*, 2(2), 35-58.
- Karaca, M. F., Görgünoğlu, S. Türkçeden Türk İşaret Dili’ne. *İleri Teknoloji Bilimleri Dergisi*, 5(1).
- Karpathy, A. (2018). Convolutional Neural Networks for Visual Recognition. *Stanford University CS231n*, 456-463
- Katılmış, Z. (2021). Makine öğrenmesi algoritmalarıyla Türk işaret dilinde harf ve dinamik sözcük tanıma (Yüksek Lisans Tezi), *Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Enstitüsü*, Bilecik, 3-12
- Kim, I. H., & Jung, I. H. (2021). A Study on Korea Sign Language Motion Recognition Using OpenPose Based on Deep Learning. *Journal of Digital Contents Society*, 22(4), 681-687.
- Kın, Z. B. (2019). Türk işaret dili alfabesinin derin öğrenme yöntemi ile sınıflandırılması (Yüksek Lisans Tezi), *Başkent Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, 5-55
- Ko, S. K., Kim, C. J., Jung, H., Cho, C. (2019). Neural sign language translation based on human keypoint estimation. *Applied sciences*, 9(13), 2683. *language recognition with Kinect. IEEE International Conference on Image Processing*, Melbourne, VIC, Australia, s.4190-4194.
- LeCun, Y., Jackel, L. D., Boser, B., Denker, J. S., Graf, H. P., Guyon, I., ... & Hubbard, W. (1989). Handwritten digit recognition: Applications of neural network chips and automatic learning. *IEEE Communications Magazine*, 27(11), 41-46.
- MathWorks introduces Release 2017a (R2017a) with a range of new capabilities in MATLAB and Simulink.

- McCulloch, W. (1943). A Logical Calculus Of The Ideas Immanent In Nervous Activity. *Bulletin of Mathematical Biophysics*, 5.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5, 115-133.
- Mehmet, G. (2023). İnsansı robot ile bazı Türkçe işaret dili ifadelerinin gerçekleştirilmesi. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 38(4), 2565-2578.
- Memiş, A. (2013). Kinect Rgb görüntülerinde ve derinlik haritalarında uzam-zamansal özellikleri kullanarak işaret dili tanıma (Yüksek Lisans Tezi), *Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul,20-25.
- Mikolov, T., Karafiát, M., Burget, L., Cernocký, J., & Khudanpur, S. (2010). Recurrent neural network based language model. *In Interspeech* (Vol. 2, No. 3, pp. 1045-1048).
- Muradli, F. (2021). Derin öğrenme kullanılarak görüntülerden insan duruş tespiti (Yüksek Lisans Tezi), *Sakarya Üniversitesi Fen Bilimleri Enstitüsü*, Sakarya,8-12.
- Nabulsi, A. A. (2018). Hand Gesture Recognition via Electromyographic (EMG) Armband for CAD Software Control. *University of Missouri-Kansas City*.
- Özcan, T., & Baştürk, A. (2020). ERUSLR: Yeni Bir Türkçe İşaret Dili Veri Seti ve Hiperparametre Optimizasyonu Destekli Evrişimli Sinir Ağı İle Tanınması. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 36(1), 527-542.
- Öztürk, K., & Şahin, M. E. (2018). Yapay Sinir Ağları Ve Yapay Zekâ 'Ya Genel Bir Bakış, *Takvim-i Vekayi*, 6(2), 25-36.
- Pacal, I., & Alaftekin, M. (2023). Türk İşaret Dilinin Sınıflandırılması için Derin Öğrenme Yaklaşımları. *Journal of the Institute of Science and Technology*, 13(2), 760-777.
- Pirim, A. G. H. (2006). Yapay zekâ. *Yaşar Üniversitesi E-Dergisi*, 1(1), 81-93.
- Porfirio, A. J., Wiggers, K. L., Oliveira, L. E., & Weingaertner, D. (2013). LIBRAS Sign Language Hand Configuration Recognition Based on 3D Meshes. *IEEE International Conference on Systems, Man, and Cybernetics*, Manchester, UK , s. 1588-1593.
- Recognition by Using 3D Geometric Invariant Feature and ANN Classification, Fukuoka, Japan. The 7th 2014 Biomedical Engineering International Conference (BMEICON) , s. 1-5.recognition system based on neural networks and a probabilistic model. *Turkish journal of electrical engineering and computer sciences* , 23, s. 2107-2123.
- Rosenblatt, F. (1957). The perceptron, a perceiving and recognizing automaton Project Para. *Cornell Aeronautical Laboratory*.

- Sabaz, f. (2022). Otomatik Türkçe Dudak Okuma İçin Bilgisayarlı Görü ve Derin Öğrenme Modellerinin Geliştirilmesi (Doktora Tezi), *Karabük Üniversitesi Fen Bilimleri Enstitüsü*, Karabük, 10-12.
- Saengsri, S., Niennattrakul, V., & Ratanamahatana, C. A. (2012). TFRS: Thai finger-spelling sign language recognition system. In *2012 second international conference on digital information and communication technology and it's applications (DICTAP)* (pp. 457-462). IEEE.
- Samaan, G. H., Wadie, A. R., Attia, A. K., Asaad, A. M., Kamel, A. E., Slim, S. O., ... & Cho, Y. I. (2022). Mediapipe's landmarks with rnn for dynamic sign language recognition. *Electronics Journal*, 11(19), 3228.
- Santhalingam, P. S., Pathak, P., Rangwala, H., & Košecká, J. (2020). Finehand: Learning hand shapes for american sign language recognition. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)* (pp. 700-707). IEEE.
- Sarawate, N., Leu, M. C., & Öz, C. (2015). A real-time American Sign Language word recognition system based on neural networks and a probabilistic model. *Turkish Journal of Electrical Engineering & Computer Sciences*, 23.
- Selda, B., & Vasif, N. V. (2006). Recognition of vowels letters of Turkish sign language by artificial neural networks. In *2006 IEEE 14th Signal Processing and Communications Applications*.
- Sign Language Alphabets From Around The World [online], Erişim adresi <https://www.ai-media.tv/knowledge-hub/insights/sign-language-alphabets/>, [Erişim Tarihi: 13 Ekim 2023].
- Subramanian, B., Olimov, B., Naik, S. M., Kim, S., Park, K. H., & Kim, J. (2022). An integrated mediapipe-optimized GRU model for Indian sign language recognition. *Scientific Reports*, 12(1), 11964.
- Sun, C., Zhang, T., Bao, B. K., & Xu, C. (2013). Latent support vector machine for sign language recognition with Kinect. In *2013 IEEE International Conference on Image Processing* (pp. 4190-4194). IEEE.
- Sundar, B., & Bagyammal, T. (2022). American Sign Language Recognition for Alphabets Using MediaPipe and LSTM. *Procedia Computer Science*, 215, 642-651.
- Şahin, E. (2018). Makine öğrenme yöntemleri ve kelime kümesi tekniği ile istenmeyen e-posta/e-posta sınıflaması (Yüksek lisans tezi), *Hacettepe Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, 8-13
- Şahin, M. E. (2023). Gaz Yakıtlı Kombi Sisteminin Yapay Sinir Ağı ile Yakıt Miktarı Tahmini Isparta Örneği. *Uluslararası Teknolojik Bilimler Dergisi*, 15(1), 11-18.

- Şeker, A., Diri, B., & Balık, H. H. (2017). Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. *Gazi Mühendislik Bilimleri Dergisi*, 3(3), 47-64.
- Şenol, D., & Denizhan, B. (2023). Kripto Para Değerinin Yapay Sinir Ağları ile Tahmin. *Journal of Industrial Engineering*, 34(1), 42-69
- Tan, F. G., Yüksel, A. S., Aydemir, E., & Ersoy, M. (2021). Derin öğrenme teknikleri ile nesne tespiti ve takibi üzerine bir inceleme. *Avrupa Bilim ve Teknoloji Dergisi*, (25), 159-171.
- Tangsuksant, W., Adhan, S., & Pintavirooj, C. (2014). American Sign Language recognition by using 3D geometric invariant feature and ANN classification. In *The 7th 2014 Biomedical Engineering International Conference*, (pp. 1-5). IEEE.
- Tanışman, S., Karcıoğlu, A. A., Aybars, U. G. U. R., & Bulut, H. (2021). Lstm Sinir Ağı ve Arıma Zaman Serisi Modelleri Kullanılarak Bitcoin Fiyatının Tahminlenmesi Ve Yöntemlerin Karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (32), 514-520.
- Tasmere, D., & Ahmed, B.(2020). Hand Gesture Recognition for Bangla Sign Language Using Deep Convolution Neural Network. *IEEE 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI)*, Dhaka, Bangladesh, s. 1-5.
- Tutar, H., Yılmaz, M. K., Eroğlu, Ö. 2017., Genel ve Teknik İletişim, *Seçkin Yayıncılık*, Ankara,30-45
- Türk Dil Kurumu Sözlüğü [online], Erişim adresi : <https://sozluk.gov.tr/>, [Erişim tarihi:17 Kasım 2023].
- Türk İşaret Dili Sözlüğü [online], Erişim adresi: <https://orgm.meb.gov.tr/www/turk-isaret-dili-sozlugu-yayimlandi/icerik/541>, [Erişim tarihi:17 Ekim 2022].
- Türkiye Cumhuriyeti Aile, Çalışma ve Sosyal Hizmetler Bakanlığı, Engelli ve Yaşlı İstatistik Bülteni [online]. Erişim adresi: https://www.aile.gov.tr/media/120191/eyhgm_istatistik_bulteni_eylul2022.pdf, [Erişim tarihi:1 Ocak 2023].
- Uğurlu, M., DOGRU, İ., & Arslan, R. S. (2023). Karanlık ağ trafiğinin makine öğrenmesi yöntemleri kullanılarak tespiti ve sınıflandırılması. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 38(3), 1737-1746.
- Uysal, A. (2010). İşitme engellilerde Türkçe öğretimi, sorunlar, öneriler (Yüksek Lisans Tezi), *İstanbul Üniversitesi Sosyal Bilimler Enstitüsü*, İstanbul.20-23
- Walia, M.S., (2022). A Comprehensive Guide on Human Pose [online], Erişim adresi: <https://www.analyticsvidhya.com/blog/2022/01/a-comprehensive-guide-on-human-pose-estimation/>, [Erişim tarihi:15 Haziran 2023].

- Wang, H., Leu, M. C., & Öz, C. (2006). American Sign Language Recognition Using Multi-dimensional Hidden Markov Models. *Journal Of Information Science And Engineering*, 22, s.1109-1123.
- Wikipedia. (2023, Ocak 7). Yapay zekâ [online]. Erişim adresi: https://tr.wikipedia.org/wiki/Yapay_zek%C3%A2, [Erişim tarihi:20 Ekim 2023].
- Yasan, M. C. (2014). Türkçe metni Türk İşaret Diline dönüştürme (Yüksek Lisans Tezi), *Başkent Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, 30-35
- Yılmaz, M. C., & Orman, Z. (2021). LSTM Derin Öğrenme Yaklaşımı ile Covid-19 Pandemi Sürecinde Twitter Verilerinden Duygu Analizi. *Acta Infologica*, 5(2), 359-372.
- Zeshan, U. (2003). Aspects of Türk İşaret Dili (Turkish Sign Language). *Sign Language and Linguistics*, 6(1), 43-75.

