

**BAŐKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĐİ ANABİLİM DALI
ELEKTRİK ELEKTRONİK MÜHENDİSLİĐİ TEZLİ YÜKSEK
LİSANS PROGRAMI**

**TÜRK İŐARET DİLİNDE KELİME TABANLI DERİN ÖĐRENME
UYGULAMASI**

HAZIRLAYAN

MEHMET ERKUŐ

YÜKSEK LİSANS TEZİ

ANKARA - 2020

**BAŐKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĐİ ANABİLİM DALI
ELEKTRİK ELEKTRONİK MÜHENDİSLİĐİ TEZLİ YÜKSEK
LİSANS PROGRAMI**

**TÜRK İŐARET DİLİNDE KELİME TABANLI DERİN ÖĐRENME
UYGULAMASI**

HAZIRLAYAN

MEHMET ERKUŐ

YÜKSEK LİSANS TEZİ

TEZ DANIŐMANI

DR.ÖĐR. Üyesi SELDA GÜNEY

ANKARA – 2020

BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Elektrik Elektronik Mühendisliği Anabilim Dalı Elektrik Elektronik Mühendisliği Tezli Yüksek Lisans Programı çerçevesinde Mehmet Erkuş tarafından hazırlanan bu çalışma, aşağıdaki jüri tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 27 / 05 / 2020

Tez Adı: Türk İşaret Dilinde Kelime Tabanlı Derin Öğrenme Uygulaması

Tez Jüri Üyeleri (Unvanı, Adı - Soyadı, Kurumu)

İmza

Prof. Dr. Mustafa DOĞAN(Başkan), İstanbul Teknik Üniversitesi

.....

Dr. Öğr. Üyesi Selda GÜNEY (Danışman), Başkent Üniversitesi

.....

Prof. Dr. Hamit ERDEM, Başkent Üniversitesi

.....

ONAY

Prof. Dr. Ömer Faruk ELALDI

Fen Bilimleri Enstitüsü Müdürü

Tarih: ... / ... /

BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS TEZ ÇALIŞMASI ORJİNALLİK RAPORU

Tarih: 10 / 06 / 2020

Öğrencinin Adı, Soyadı : Mehmet Erkuş

Öğrencinin Numarası : 21310041

Anabilim Dalı : Elektrik Elektronik Mühendisliği

Programı : Elektrik Elektronik Mühendisliği Tezli Yüksek Lisans

Danışmanın Unvanı/Adı, Soyadı : Dr. Öğr. Üyesi Selda Güney

Tez Başlığı : Türk İşaret Dilinde Kelime Tabanlı Derin Öğrenme Uygulaması

Yukarıda başlığı belirtilen Yüksek Lisans tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 58 sayfalık kısmına ilişkin, 09 / 06 / 2020 tarihinde tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 10'dır.

Uygulanan filtrelemeler:

1. Kaynakça hariç
2. Alıntılar hariç
3. Beş (5) kelimedenden daha az örtüşme içeren metin kısımları hariç

“Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını” inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası:.....

ONAY

Tarih: 10 / 06 / 2020

Öğrenci Danışmanı Unvan, Adı, Soyadı

Dr. Öğr. Üyesi Selda Güney

TEŐEKKÜR

Tez alıŐmalarım boyunca beni yÖnlendirdiĐi, bilgi ve birikimlerini bana aktardıĐı iin danıŐmanım Sayın Dr. ÖĐr. Üyesi Selda Güney'e teŐekkürlerimi sunarım.

Hayatım boyunca bana her konuda destek olan ve her türlü fedakârlıkta bulunan sevgili anneme ve sevgili babama teŐekkürlerimi bor bilirim.

Tez süresince karşılaŐılan güçlüklerin aŐılmasında her zaman yardımcı olan ve motive eden canım arkadaşım Sanem ElbaŐı'na, veri seti oluŐurmada yardımcı oldukları iin sevgili kuzenim Alihan Durhan'a ve dostum Emin BaŐıĐit' e teŐekkür ederim.

ÖZET

Mehmet ERKUŞ

TÜRK İŞARET DİLİNDE KELİME TABANLI DERİN ÖĞRENME UYGULAMASI

Başkent Üniversitesi Fen Bilimleri Enstitüsü

Elektrik Elektronik Mühendisliği Anabilim Dalı

2020

İşaret dili, işitme sorunu olan kişilerin kendilerini ifade etmek için kullandıkları bir görsel iletişim şeklidir. Bu çalışmanın temel amacı, işitme sorunu olan kişilerin yaşamını kolaylaştırmaktır.

Bu çalışma kapsamında, üç farklı kişiden alınan 32 sınıf için 3200 adet RGB görüntü kullanılarak bir veri seti elde edilmiştir. Veri kümelerine veri artırma yöntemleri uygulanmış, görüntü sayısı her bir sınıf için 600 adet olacak şekilde toplam 3200'den, 19200'e yükseltilmiştir. İşaretlerin sınıflandırılması için hem problemin çözümüne yönelik 10 katmanlı bir evrişimli derin ağ modeli oluşturulmuş hem de derin öğrenme yöntemlerinden biri olan aktarımlı öğrenme yöntemi kullanılarak hazır modellerden VGG166, Inception ve ResNet derin ağ mimarileri kullanılmıştır. Ayrıca derin öğrenmenin öznelik çıkarımı tekniğinden faydalanılarak elde edilen öznelik vektörü ile geleneksel makine öğrenme yöntemlerinden olan Destek Vektör Makineleri (Support Vector Machines, SVM) ve K- En Yakın Komşu(K- Nearest Neighbor, K-NN) yöntemleri ile sınıflandırılmıştır. Elde edilen sonuçlar zaman ve başarı oranlarına göre kıyaslanarak, en başarılı yöntem belirlenmiştir.

Bu çalışmada, görsel bir dil olan Türk İşaret Dili (TİD)'ne ait durağan kelimelerin yapılan analizler sonucunda başarılı bulunan derin öğrenme yöntemlerinden biri olan aktarımlı öğrenme kullanılarak gerçek zamanlı yazılı dile çevrilmesi sağlanmıştır. Ayrıca, tasarlanan bu gerçek zamanlı sistem ile TİD durağan kelimelere ait işaretleri tanıma ve tahminini bilgisayar ekranına yazdırmasındaki başarısı değerlendirilmiştir.

ANAHTAR KELİMELER: Görüntü Örüntüsü Tanıma, Derin Öğrenme, Aktarımlı Öğrenme, Türk İşaret Dili Tanıma, Evrişimli Sinir Ağları

ABSTRACT

Mehmet ERKUŞ

**TÜRK İŞARET DİLİNDE KELİME TABANLI DERİN ÖĞRENME
UYGULAMASI**

Başkent Üniversitesi Fen Bilimleri Enstitüsü

Elektrik Elektronik Mühendisliği Anabilim Dalı

2020

Sign language is a form of visual communication used by people with hearing problems to express themselves. The main purpose of this study is to make life easier for people with hearing problems.

In this study, a data set was obtained using 3200 RGB images for 32 classes taken from three different people. Data development methods were applied to the data sets and the number of images was increased from 3200 to 19200, 600 per class. For the classification of the signs, both a 10-layer convolutional deep network model was created for the solution of the problem, and VGG166, Inception and ResNet deep network architectures, which are one of the deep learning methods, were applied by using transfer learning method. In addition, the signs are classified using the Support Vector Machines (SVM) and K- Nearest Neighbor (K-Nearest Neighbor, K-NN) methods, which are the traditional machine learning methods, with the feature vector obtained by using the feature extraction technique of deep learning. The most successful method was determined by comparing the obtained results according to time and performance ratios.

In this study, stationary words belonging to Turkish Sign Language (TSL), which is a visual language, are translated into real time written language by using transfer learning with one of the deep learning methods, which is successful as a result of the analysis. In addition, with the real-time system designed, its success in recognizing the stationary words of TSL signs and printing its prediction on the computer screen were evaluated.

KEYWORDS: Image Pattern Recognition, Deep Learning, Transfer Learning, Turkish Sign Language Recognition, Convolutional Neural Network.

İÇİNDEKİLER

Sayfa

TEŞEKKÜR	i
ÖZET.....	ii
ABSTRACT	iii
İÇİNDEKİLER	iv
TABLolar LİSTESİ.....	vii
ŞEKİLLER LİSTESİ.....	viii
SİMGELER VE KISALTMALAR LİSTESİ.....	x
1. GİRİŞ	1
1.1. Konunun Tanımı	1
1.2. Konuya İlişkin Önceki Çalışmalar	3
1.3. Konunun Amacı.....	10
1.4. Yöntem ve Materyal.....	11
2. TÜRK İŞARET DİLİ.....	12
3. MAKİNE ÖĞRENMESİ.....	14
3.1. Makine Öğrenmesinde Büyük Verinin Önemi.....	14
3.2. Makine Öğrenmesine Yaklaşımlar	15
3.2.1. Danışmanlı öğrenme	15
3.2.2. Danışmansız öğrenme	16
3.2.3. Yarı danışmanlı öğrenme	16
3.3. Yapay Zeka ve Makine Öğrenmesinin Tarihçesi	16
3.4. Yapay Sinir Ağları.....	17
3.5. Yapay Sinir Ağları İle Durağan Kelimelerin Tanınmasına İlişkin Yapılan Önceki Çalışmalar	21
3.6. Yapay Sinir Ağları Tarihçesi.....	21
4. DERİN ÖĞRENME ve EVRİŞİMSEL SİNİR AĞLARI (ESA)	24

4.1. Derin Öğrenme ve Tarihçesi	24
4.1.1. Derin öğrenmenin gelişimi	25
4.2. Derin Öğrenmenin Güçlü Yönleri	25
4.3. Derin Öğrenmede Etkili Olan Kavramlar	27
4.3.1. Ağırlık	27
4.3.2. Sapma	27
4.3.3. Aktivasyon fonksiyonu	28
4.3.4. Öğrenme katsayısı ve eniyileme algoritması	29
4.3.5. Kayıp fonksiyonu	30
4.3.6. Geri yayılım	30
4.3.7. Yığın, epok ve tekrar sayısı	30
4.4. Derin Öğrenme Mimarileri	30
4.5. Evrimsel Sinir Ağlar (ESA)	31
4.6. 10 Katmanlı ESA YAPISI	34
4.7. Aktarımlı Öğrenme	34
4.7.1. VGG16	35
4.7.2. Inception	36
4.7.3. ResNet50	36
5. TİD VERİ SETİ OLUŞTURULMASI	38
5.1. Kullanılan Yöntem ve Materyal	38
5.2. Veri Setinin Oluşturulması	38
5.3. Veri Setinin Çeşitlendirme -1	39
5.4. Veri Setinin Çeşitlendirme -2	40
5.5. Veri Setinin Genel Özellikleri	42
6. ARAŞTIRMAYA İLİŞKİN BULGULAR	43
6.1. Araştırmaya İlişkin Bulgular	43
6.1.1. 10 Katmanlı ESA yapısını kullanarak başarımlı değerlendirme	43
6.1.2. Öğrenme aktarımı ile VGG16 kullanarak başarımlı değerlendirme ...	45
6.1.3. Öğrenme aktarımı ile Inceptionv3 kullanarak başarımlı değerlendirme	46

6.1.4. Öğrenme aktarımı ile ResNet50 kullanarak başarımlı değerlendirme	47
6.1.5. Öğrenme aktarımı kullanılarak öznelik çıkarımı ile birlikte destek vektör makineleri ve k- en yakın komşu sınıflandırması.....	48
6.1.6. Öznelik çıkarımı olmadan dvm ve k- en yakın komşu ile sınıflandırma	49
6.2. Tartışma	49
7. TÜRK İŞARET DİLİ DURAĞAN KELİMELEİN GERÇEK ZAMANLI TEST EDİLMESİ.....	53
7.1. Çalışma Prensipleri.....	53
7.2. Gerçek zamanlı programın özellikleri.....	54
7.3. Sistemin Farklı Kullanıcılar Üzerinde Test Edilmesi	55
8. SONUÇ VE ÖNERİLER.....	57
8.1. Sonuç.....	57
8.2. Öneriler	58
KAYNAKLAR	59

TABLULAR LİSTESİ

	Sayfa
Tablo 5.1. WebCam' in genel özellikleri	38
Tablo 5.2. Veri çeşitlendirme işlemleri	41
Tablo 5.3. Veri setinin genel özellikleri	42
Tablo 5.4. İki farklı veri setinin görüntü özellikleri	42
Tablo 6.1. 10 Katmanlı ESA (CNN) için farklı Epok sayılarında elde edilen sonuçlar.....	43
Tablo 6.2. VGG16' nın farklı epok sayılarında elde edilen sonuçları.....	45
Tablo 6.3. InceptionV3' ün farklı epok sayılarında elde edilen sonuçları	46
Tablo 6.4. ResNet50' ün farklı epok sayılarında elde edilen sonuçları.....	47
Tablo 6.5. Öznitelik çıkarımı yapılarına elde edilen DVM ve K en yakın komşu sonuçları .	48
Tablo 6.6. DVM ve K en yakın komşu sınıflandırma sonuçları	49
Tablo 6.7. Modellerin zaman ve başarımları karşılaştıran tablo	50
Tablo 7.1. Sistemin farklı ortamlar ve kişiler üzerinde test sonuçlarını gösterir.....	55

ŞEKİLLER LİSTESİ

	Sayfa
Şekil 1.1. Kelime tabanlı işaret dili tanıma	2
Şekil 1.2. Harf tabanlı (parmak alfabesi) işaret diline örnek	2
Şekil 2.1. Tez çalışmasında kullanılan durağan kelimelere ait görseller (TİD Sözlüğü).....	13
Şekil 3.1. Geleneksel makine öğrenmesini gösterimi.....	14
Şekil 3.2. Biyolojik beyin nöronu yapısı	18
Şekil 3.3. Çok katmanlı yapay sinir ağı nöronu çalışma fonksiyonu (F. Rosenblatt'ın 1958 yılında tanıttığı model)	19
Şekil 3.4. İleri beslemeli yapay sinir ağı katmanları.....	20
Şekil 3.5. Sigmoid fonksiyonu grafiği	22
Şekil 4.1. Derin sinir ağlarını genel yapısı	24
Şekil 4.2. Derin öğrenme ve geleneksel makine öğrenmesi yaklaşımının karşılaştırılması	26
Şekil 4.3. Katmanlar, kayıp fonksiyonu ve eniyileme algoritması arasındaki ilişki	27
Şekil 4.4. Evrişimli ağlarda bilgisayar mimarisi katmanları.....	32
Şekil 4.5. 4x4 boyutunda bir görselde 2x2 boyutunda bir filtre ile görüntü işleme biçimi .	33
Şekil 4.6. VGG16 mimari yapısı ve katmanları	35
Şekil 4.7. Inception mimarisine sahip iç içe ağların gösterimi.....	36
Şekil 4.8. Bir ResNet mimarisi için daha derin bir artık fonksiyon yapı taşı ve ResNet-50 için bir “darboğaz” yapı taşının gösterimi	37
Şekil 5.1. Veri kümesinde yer alan TİD’ e ait 32 sınıf için renkli görüntüler	39
Şekil 5.2. Beş işaretinin gerçek ve gürültü eklenmiş görüntüleri.....	40
Şekil 5.3. Dikdörtgen işaretinin gerçek ve karşıtlık eklenmiş görüntüleri.....	40
Şekil 5.4. Başarılı İşaretine gerçek ve bulanıklık eklenmiş görüntüleri.....	40
Şekil 5.5. Köprü işareti için çeşitlendirilmiş bazı görüntüler	41

Şekil 6.1. 10 katmanlı ESA'nın başarımlar grafiği.....	44
Şekil 6.2. 10 katmanlı ESA'nın kayıpları grafiği.....	44
Şekil 6.3. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen VGG16 mimarisinin başarımlar grafiği.....	45
Şekil 6.4. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen VGG16 mimarisinin kayıpları grafiği.....	45
Şekil 6.5. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen InceptionV3 mimarisinin başarımlar grafiği.....	46
Şekil 6.6. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen InceptionV3 mimarisinin kayıpları grafiği.....	46
Şekil 6.7. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen ResNet50 mimarisinin başarımlar grafiği.....	47
Şekil 6.8. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen ResNet50 mimarisinin kayıpları grafiği.....	47
Şekil 6.9. InceptionV3 karışıklık matrisi.....	51
Şekil 6.10. ResNet50 karışıklık matrisi.....	52
Şekil 7.1. TİD 'dokuz' işaretinin bilgisayar ekran görüntüsü.....	53
Şekil 7.2. TİD 'zor' işareti bilgisayar tahminini gösteren ekran görüntüsü.....	54
Şekil 7.3. TİD basit yapıda bir cümlenin yazıya çevrilmesini gösteren bilgisayar ekran görüntüsü.....	54
Şekil 7.4. Dur işareti ile tüm kelimelerin yazımını gösteren bilgisayar ekran görüntüsü....	55
Şekil 7.5. Taşduvar arka planında lazım kelimesinin metin alanına yazılan ekran görüntüsü.....	56
Şekil 7.6. Doku desenli duvar arka planında üçgen kelimesinin metin alanına yazılan ekran görüntüsü.....	56
Şekil 7.7. Dış çekim doğa görüntüsü arka planında dokuz kelimesinin metin alanına yazılan ekran görüntüsü.....	56

SİMGELER VE KISALTMALAR LİSTESİ

AİD	Amerikan İşaret Dili
A.G.E	Adı Geçen Eser
ARİD	Arap İşaret Dili
2B	İki Boyutlu
3B	Üç Boyutlu
CRT	Katot Işın Tüpü
dB	Desibel
DCT	Ayrık Kosinüs Dönüşümü
DVM	Destek Vektör Makinesi
EMG	Elektromiyografi
ESA	Evrişimli Sinir Ağı
ESAK	Evrişimli Katman
HİD	Hint İşaret Dili
HOİD	Hollanda İşaret Dili
İİD	İngiliz İşaret Dili
JİD	Japon İşaret Dili
KİD	Kore İşaret Dili
K –NN	K En Yakın Komşu Sınıflandırma Algoritması
MEB	Milli Eğitim Bakanlığı
MLP	Çok Katmanlı Algılayıcı
RDF	Rastgele Karar Ormanı
RGB	Kırmızı, Yeşil, Mavi ana renklerden oluşan renk uzayı
TDK	Türk Dil Kurumu
TİD	Türk İşaret Dili
TÜİK	Türkiye İstatistik Kurumu
YSA	Yapay Sinir Ağı
WEB	İnternet
WHO	Dünya Sağlık Örgütü

1. GİRİŞ

1.1. Konunun Tanımı

İnsan varoluşu gereği sosyal bir canlıdır. İnsanoğlu sosyalleşmek için yazılı yahut sözlü olarak gerek kendini ifade etme isteğinden gerekse kendini bir topluluğun parçası olarak hissetmenin, ata mirası olan korkularından kurtuluşun başlıca yöntemi olduğunu bildiğinden kısaca, içgüdüsel zekanın emri gereği diğer insanlarla iletişim kurmak zorundadır. İletişim sınırlı anlamı ile söz söyleme, dinleme ve anlama üzerine kurulu olan diyalog sistemidir. İletişimin sınırlı anlamında var olan herhangi bir kavramın aksaması, iletişimin aksaması olarak tanımlanır. İletişimin aksaması bireyin sosyal anlamda kısıtlanmasına yol açmaktadır.

Her iki kulakta, 0-20 dB kayıp normal kayıp olarak adlandırılırken, 35-40 dB yahut üzerinde olan kayıplar, duyma engeli olarak adlandırılır. Bu engel bireyin, dinleme buna bağlı olarak da, konuşma yetisindeki kaybın oluşması durumunu ortaya koymaktadır [1]. Dünya Sağlık Örgütü'nün 2018 yılında duyurduğu raporda, dünya genelinde dört yüz altmış altı milyona yakın duyma engeli yaşayan birey olduğu belirtilmiştir (2018, WHO Raporu). 2010 yılında Türkiye İstatistik Kurumu (TÜİK)'nin yaptığı sağlık taramasında ülkemizde üç milyona yakın kişinin duyma engelli olduğu belirtilmiştir (2010 TÜİK Raporu).

Toplum ile iletişim halinde olmaya ihtiyaç duyan ve işitme kaybına uğramış bireylerin, toplumsal ve sosyal iletişimi sürdürmesi, yazılı anlatım, dudak okuma yöntemi, işaret dili olarak kabul gören görsel iletişim çeşitlerine yönelmelerine neden olmaktadır.

Görsel iletişim biçimi; alıcı sinyalleri ile verici sinyallerinin aynı anda birlikte hareket ettiği bilgi aktarımını bu yolla sağlayan bir çeşit iletişim biçimidir [2]. Görsel iletişimin bir kolu olan, 16.yy'da İtalyan Dr. J. Cardan tarafınca geliştirilen işaret dillerinin, günlük kullanılan konuşma, okuma ve yazma dillerindeki imla kurallarına benzeyen, sabit ve bozulduğunda anlam karmaşasına yol açan kuralları vardır. İşaret dilleri; konuşma dillerinde olduğu gibi evrensel nitelikte olmayıp, bu dili konuşan kişilerin yaşadıkları bölge ile bağlıdır [3].

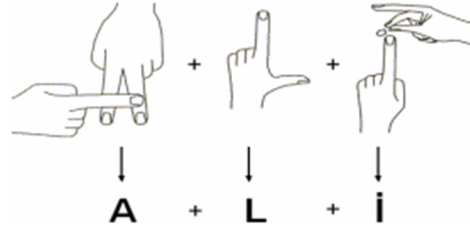
İşaret dili, yaşanılan bölgede konuşulan dilin etkisi altında ise de duyma yetisi yerinde olan bireylerin günlük kullandıkları dil bire bir benzerlik barındırmamakta ve bu görsel iletişim biçimi, toplumun çoğunluğu tarafından anlaşılmadığından bireyi kısıtlı bir sosyal çevre ile sınırlandırmaktadır. Örneğin; Ülkemizdeki günlük konuşma dili Türkçe ile Türk

İşaret Dili olan (TİD) arasında bire bir örtüşme olmadığı gibi, dünya genelinde ortak dil kabul edilen İngilizce ile İngiliz İşaret Dili olan (İİD) 'de birebir örtüşme sağlamamaktadır [4].

İşaret dili ifade ve anlam açısından kendi içinde ikiye ayrılmaktadır. Kelime tabanlı işaret dilinde ellerle birlikte, beden dili, yüz ifadesi ve yüzde ifade sırasında oluşan mimikler bir bütün olarak değerlendirilerek her hareketin bir kelimeyi anlattığı ifade biçimidir.



Şekil 1.1. Kelime tabanlı işaret dili tanıma [35]



Şekil 1.2. Harf tabanlı (parmak alfabesi) işaret diline örnek [35]

Buna karşın harf tabanlı işaret dilinde, her hareket bir harfi ifade ettiğinden harf tabanlı işaret diline, parmak alfabesi de denilmektedir. Kullanım açısından genel anlamda kelime tabanlı işaret dili tercih edilse de özel isimlerin ifadesinde, bilimsel terimlerde, kısaltmalarda yahut yabancı dillerden dilimize geçmiş olan kelimelerde, kelime tabanlı işaret dili ile tam olarak kelime anlamını karşılamak mümkün olmadığından parmak alfabesinin kullanımının tercih edilmektedir [5].

İşitme kaybı yaşayan bireyler iletişimde bulunurken, işaret dili kullanımının gerek yaşanan coğrafyada gerek ise dünya coğrafyasında yaygın olmaması hatta işitme kaybı yaşayan bireyler, yakınları ve bu görsel dile ait eğitim alan kişilerin dışında işaret dilini bilen kişi olmaması sebebi ile dudak okuma yöntemini kullanarak söyleneni anlayabilmekte ise de kendilerini ifade etme konusunda zorlanmaktadır.

İşitme kaybı olan bireylerin, kendilerini daha rahat ifade edebilmesini sağlamak amacı ile çeşitli işaret dili tanıma sistemleri tasarımları oluşturulmaktadır. İşaret dili tanıma

sistemleri ile ilgili çalışmalar, makine öğrenmesi sistemlerinde son 30 yıldır görülmeye başlanmış ve günümüze kadar pek çok çalışma yapılarak, çeşitli sistemler tasarlanmıştır.

Tasarlanan sistemler, görüntüyü işleme tabanlı sistemler ve elektronik eldiven tabanlı olarak ikiye ayrılmaktadır. Bu yöntemler kullanılarak tasarlanan ilk sistemler alıcıların fazlası ile mevcut olduğu elektronik eldivenler kullanılarak tasarlanan sistemlerdir. Elektronik eldivenler kullanılarak tasarlanan sistemler, pahalı sistemler olmasının yanı sıra, işaret dili hareketini yapan modelin eldiveni giymesi gerekli olduğundan ve eldivenin, hareketi yapan modelin hareketlerini kısıtlaması nedeni ile kullanımı kolay olmayan bir yöntem olarak görüldüğünden, son yıllarda görüntü tabanlı işaret dili tanıma ve sınıflandırma sistemleri kullanılmaya başlanmıştır [6,7].

İşaret dili tanıma ve sınıflandırma sistemlerinde üzerine en fazla araştırma yapılan dil, Amerikan İşaret Dili (AİD)'dir. İlk AİD tanıma sistemi tasarımları 1988 yılında Vogler ve Metaxas tarafınca saklı Markov modeli kullanarak geliştirdikleri sistem tasarımıdır. Bu sistem tasarımı üç boyutlu (3B) bilgi hareketlerini elde etmek için üç kamera kullandıkları tasarım sistemidir [7].

1.2. Konuya İlişkin Önceki Çalışmalar

Görüntüden işaret dili tanıma sistem tasarımları arasında, 1995 yılında Starner'ın AİD üzerine standart bir bilgisayar ve bağlı kamera ile gerçekleştirdiği Saklı Markov modeli ile tasarlanmış işaret dili tanıma sistemi ile başlamıştır. O tarihe dek tasarlanan sistemler arasında en fazla başarıya ulaşmış olan işaret dili tanıma sistemidir. İzleme, analiz etme ve tanıma işlemlerinin tamamının gerçek zamanlı olduğu bu sistem tasarımında Starner, elin duruşunun ayrıntılı şekilde incelenmesinin gerekli olmadığını düşünmüştür. Starner bunun yerine el şekli, el yönü ve el yörüngesi bilgisini kabaca kullanarak, elleri takip etmiştir. Bu takibin kolay olmasını sağlamak için ise işareti gerçekleştiren modelinin, sağ eline sarı, sol eline kavuniçi renkte elektronik özelliği olmayan basit eldivenler giydirmiş ve işareti yapmasını istemiştir. Starner, isim, fiil, zamir ve sıfattan oluşmuş 40 kelime ve bu kelimelerden türetilen 494 cümleden veri seti oluşturulmuş bir veri seti üzerinde çalışmıştır. Starner oluşturduğu veri setinin 395 tanesini eğitim, 99 tanesi test amaçlarıyla kullanmıştır. Starner, yaptığı testler sonucunda kelimelerin bir kısmı için, tek başına hareket bilgisinin yetersiz kaldığını ve parmak duruşlarının yapılan hareketlerde önemli olduğunu belirlemiştir. Testler tekrar edilip, kelimeler sınırlandırıldığında, örneğin, bir manzara resminin tarif edildiği bilindiğinde sistem başarısının, %99'a kadar ulaştığını doğrulamıştır.

Cümle yapısı ve kelime sayısının bilinmediği testlerde ise %91 oranında başarılı sonuçlar alınmıştır. Bu sistem eldiven kullanılmadan ten rengi ile yeniden test edildiğinde, gerçek zamanlı değeri yakalayamamıştır. Sistemde eldivensiz uygulanan testlerde, doğru tahminleme başarı oranında %75 'e kadar düşüş tespit edilmiştir. Starner' ın tasarladığı işaret dili tanıma sistemi, sonraki yıllarda tasarlanan işaret dili tanıma sistemlerine yol göstermesi nedeni ile önemli bir sistem tasarımıdır [8].

1999 yılında Segen ve Kumar tarafından tek kamera ile gece, loş ışık, spot ışığı ortamlarında AİD'ne ait hareketleri yapan modellerin görüntüleri alınarak veri seti oluşturmuşlardır. Araştırmacılar, çalışan bir görüntü tanıma sistemi kurgulamayı hedeflemişlerdir [9].

2015 yılında Dong tarafından tasarlanan bir başka sistemde ise sistemin yapılan hareketleri algılayabilmesi için Microsoft Kinect hareket sensörü kullanılmıştır. Dong, sistem içerisinde görüntüleri arttırmak için segmentli el yapılandırmasını kullanmıştır. Dong'un kullandığı segmentli el yapılandırması, piksel başına sınıflandırma algoritmasına dayalı bir derinlik kontrast özelliğinin kullanılmasından ibarettir. Dong bu segment uygulamasından sonra, hareket kısıtlamaları ile el eklemi konumlarını lokalize etmek için bir sıralı bir mod bulma yöntemi geliştirmiş ve uygulamıştır. Dong, son olarak, AİD işaretlerini eklem açlarına göre tanımak için Rastgele Karar Ormanı (RDF) sınıflandırıcısı oluşturmuştur. Araştırmacı tasarladığı yöntemin tahmin performansını göstermek için, 75.000 örnekten aldığı 24 adet statik AİD alfabe işareti içeren bir veri kümesini kullanmıştır. Araştırmacı, tasarladığı sistemin doğruluk oranını %92 olarak tespit etmiştir. Araştırmacının, bu çoklu tasarımı, üzerinde araştırma yapılması ve kullandığı yöntemlerin değerlendirilmesi isteği ile Surrey Üniversitesi'ne gönderilmiştir. Dong'un, tasarımına ait veri seti, Surrey Üniversitesi'nin platformunda yayınlanarak kamuya açık bir veri setine dönüştürülmüştür [10]. Dong'un elde ettiği sonuçlar, önceki çalışmalarla kıyaslandığında, benimsediği yöntemin, AİD'ne ait hareketsiz kelimelere işaretlerinin çözümlenmesi konusunda, yüksek doğrulama performansına sahip bir tasarımıdır [11],[12].

2014 yılında Chuan, Regina ve Guardino [13] tarafınca geliştirilmiş sistem tasarımı ile kompakt ve uygun fiyatlı bir 3B hareket sensörü olan Leap Motion sensörü [14] kullanarak Amerikan İşaret Dili tanıma sistemi üzerinde oluşturdukları bu tasarımları ile Leap Motion sensörü' nün [15], mevcut tasarımlarda çoğunlukla tercih edilen Cyblerglove [16] veya Microsoft Kinect hareket sensörlerinden [17] çok daha ekonomik bir çözüm sunduğunu kanıtlamışlardır [13].

Cunnigham ve Delany 2007 yılında Amerikan İşaret Dili üzerinde yapılan diğer tasarım kurgularında ise Amerikan Alfabesinin 26 harfini duyuşal verilerden türetilmiş özellikleri kullanarak sınıflandırmak için K en yakın komşu (K-NN) ve destek vektör makineleri (DVM) algoritmaları kullanılmıştır [18]. Deneylerden alınan sonuçlar değerlendirildiğinde, sınıflandırmada en yüksek başarı performansı ortalaması sırasıyla %72.78 K en yakın komşu [19] ve % 79.83'ünün de DVM kullanımları ile elde edildiğini göstermiştir [19,20].

Wu vd. tarafınca 2016 yapılan çalışmada AİD tanınması amacıyla Atalet (Inertial) ve Yüzey Elektromiyografi (EMG)[22] sensörleri ile alınan bilgilerin kullanılması hedeflenmiştir. Bunun için giyilebilen bir sistem tasarım oluşturulmuştur. Bu sistem tasarımında, AİD'e ait günlük iletişimde diğer kelimelere nazaran daha fazla kullanılan 80 farklı işaret ile oluşturulan veri seti üzerinde Saf Bayes sınıflandırıcı (Naive Bayes) [23], K-NN [19] ve DVM [20] sınıflandırma yöntemleri kullanılmıştır. Bu yöntemlerin ardı sıra uygulanmasıyla %85.24 - %96.16 aralığında başarı elde edilmiştir[21].

Ameen ve Vedara tarafından 2017 yılında yapılan bir başka çalışmada Amerikan İşaret Dili parmak alfabesinin tanınması hedeflenmiştir. Araştırmacılar bunun için ESA yapısını kullanılmışlardır. Araştırmacılara ait sistem tasarımında, görüntü yoğunluğu ve derinlik verileri kullanılmıştır. Araştırmacılar, 60 binden daha fazla resim barındıran bir veri seti oluşturmuşlardır. Araştırmacılar daha sonra, fotoğraflar üzerinde renk ve derinlik özelliklerini çıkartmak için iki farklı ESA yapısı modellemişlerdir. Araştırmacılar elde ettikleri tüm özellikleri birleştirilerek sınıflandırma yapılmışlardır. Araştırmacıların, yaptıkları değerlendirmelerde %82 hassasiyet (precision) başarıları elde edilmiştir [24].

Quesada vd.[25]. tarafından tasarlanan bir diğer sistemde Leap Motion ve Intel RealSense [27] sensörleri kullanılarak, 50'den fazla birey ile DVM yöntemiyle sınıflandırma sağlanarak gerçekleştirilmiş ve bu tasarım ile %93 başarı elde etmişlerdir [25].

Hint İşaret Dili (HİD) üzerine Rao vd. tarafından 2018 yılında Hint İşaret Dili (HİD) kullanılarak işaret dili tanıma sistem tasarımı için, tek el ile yapılabilen 18 işaretin tanınması hedeflenerek videodan alınan her frame için segmentasyon işlemi uygulanmıştır. İleri beslemeli YSA ile sistem performansı test edilmiş ve tasarımın sistem performansının %90'a yakın oranda başarılı olduğu sonucuna varılmıştır [26].

Rao vd. tarafından 2018 yılında kurgulanmış diğer bir işaret dili tanıma sistemi tasarımında HİD'e ait 200 işaretin sınıflandırılması hedeflenerek ESA yapısı kurgulanmıştır. Araştırmacılar, bu sistem tasarımında kullanılmak üzere toplamda 300,000 RGB video

görüntüsünden oluşan veri seti oluşturmuşlardır. ESA yapısı 4'lü konvolüsyon katmanından oluşturulmuş ve sistemin performans testi sonucu %92.88 başarılı bulunmuştur [27].

Mapari ve Kharat'ın 2015 yılında yapmış oldukları işaret dili algılama sisteminde ise Leap Motion sensörünü kullanarak HİD'i tanıyan bir Hint İşaret Dili tanıma sistemi geliştirmişlerdir [28]. Leap Motion sensörü yapılan el hareketini yakalayarak 3B formatında parmak pozisyonu vermektedir (X, Y, Z eksenleri değerleri). Her iki el için avuç ortası ile birlikte beş parmak ucunun konumsal bilgileri, Öklid mesafesine ve Kosinüs benzerliğine dayalı işaret postürünü tanımak için kullanılır. Yapılan çalışmada sistemin test edilebilirliğini ölçmek için 10 farklı işitme ve konuşma engelli bireyler ile HİD işaretleri için test edilmiştir. HİD 'nin ortalama tanıma doğruluğu Öklid mesafe yöntemi için % 88.39 ve Kosinüs benzerliği için % 90.32'dir. İşaret yaparken, Leap Motion Camera, derinlik bilgilerinin doğru şekilde çıkarılması için yaklaşık 10 derece eğimli tutulur. Ayrıca, çalışmalarında Leap Motion sensörü ile her iki elin de doğru bir şekilde izlenmesine rağmen, diğer vücut kısımlarının ve yüz ifadelerinin yakalanamadığını belirtmişlerdir [28].

1996 yılında Yachida vd.[36] tarafından Japonca İşaret Dili (JİD) üzerinde saklı Markov modeli ile, tasarlanan işaret dili tanıma sistemi, çeşitli renklerde bantlar bulunan bir eldiven kullanılarak kurgulanmıştır. Sistem hem parmak alfabesi tabanlı hem de kelime tabanlı kelimeleri tanıyabilmek üzere eğitilmiş, doğrudan elin duruşu incelenmek yerine, elin şekline göre görünür bantların renkleri ve yoğunluğu incelenmiştir. Sistem tasarımında yanılma payının düşürülmesi hedeflenerek, eldivenlerin üzerine renkli bantlar yerleştirilirken, elin farklı duruşlarının aynı renk olmamasına özen gösterilmiştir [29].

1997 yılında Grobel ve Assan [30], tarafından CCD kamera kullanılarak tasarladıkları kişiye bağlı sistemde, Yachida vd. [36]'nın kullandığı renk kombinasyonundan etkilenmiş ancak tek el yerine her iki el içinde modelin sağ eline parmaklar, avuç içi ve elin arka yüzünü belli edecek şekilde yedi farklı renk bulunan eldiven, sol eline ise üzerinde sekiz farklı renk bulunan eldiven giydirilmiştir. Ellerin hareket sırasında izledikleri rota, elin şekli ve yönü gibi özelliklerden faydalanarak veri setini oluşturarak tanıma sistemi kurgulanmıştır [30]. Hollanda İşaret Dili (HOİD) üzerinde saklı Markov modeli kullanılarak tasarlanmış bu işaret dili tanıma sisteminde HOİD' e ait 262 hareket %94 başarı oranında sağlanmıştır [30].

2005 Mohandes vd. tarafından Arap İşaret Dili (ARİD) üzerinde yönlü gradyanların histogramı /HoG, kanonik korelasyon analizi ve rastgele Forest sınıflandırıcı, kullanarak oluşturduğu ARİD tanıma tasarımını, iki yıl sonra saklı Markov ve Kinect hareket sensörü

kullanarak aynı veri seti ile tekrarlamış, işaret dili tanıma tasarımını yeniden kurgulamıştır [31].

Literatüre geçmiş en ilginç işaret dili tanıma ve sınıflandırma sistemi tasarımı ise 2000 yılında İmagawa vd.'nin tasarladıkları sistem, ten renginden faydalanarak elleri ayırt etmekte ve iki aşamalı bir tanıma sistemi önermektedir. Öncelikle kişinin tanınmaya çalışılan hareketine, evrensel özelliklere göre çağrışım yapılan kelimeler listelenmekte ve devamında bu işaretlerden yerel özellikler kullanılarak işaret edilen kelime belirlenmektedir. Tasarlanan İşaret dili tanıma sisteminin başarısını ölçmek için postanedeki bir konuşmada geçebilecek Japon İşaret Dili'ne ait 160 kelime seçilmiştir [43]. Tasarlanan bu sistem belli bir hizmet almak amacı ile bulunulan, banka, postane, hastane gibi alanlarda belli bir amaca yönelik kullanılması muhtemel kelimeler düşünülerek, oluşturulan senaryolar üzerinden kelimeleri tanıyan sistemler tasarlamak, hem hizmeti sunan hem de hizmeti almak isteyen ancak bunu ifade edemeyen işitme ve buna bağlı olarak konuşma sorunu yaşayan bireylerin, engelsiz iletişim sağlanması açısından doğru bir yöntem olmasının yanı sıra ticari amaçlar ile uygulamaya yönelik başkaca katkıların sağlanabilmesi imkanı doğurması açısından da önem arz etmektedir. Bu sistem tasarımı üzerine yapılan testler sonucunda yakın evrensel özelliğe sahip kelimeleri yerel özelliklerinden faydalanarak tanıma başarısı %89,963 olarak gerçekleşmiştir [32].

Ülkemizde ise; Türkçe Dili ile tasarlanmış konuşma tanıma üzerine tasarlanmış pek çok sistem bulunmaktadır. Bunun nedenlerinden en önemlisi, kamu yahut özel sektör tarafından kurulmuş kendi başına dev bir sektör olan çağrı merkezi sektörüdür. Bu sektöre yönelik yapılan çalışmalar, bu çalışmalar üzerine tasarlanan sistemlerin çeşitli parametreleri değiştirilerek farklı kullanım alanlarına da sunulması, yan sektörleri de beraberinde getirmektedir. Bu alana sürekli yatırımların yapılması sonucu ticari faaliyet anlamında çok fazla olanak barındırmasıdır.

Literatürde Türk İşaret Dili üzerinde yer alan işaret dili tanıma tasarımları, Türkçe Dili ile tasarlanmış konuşma tanıma sistemlerine göre, çok kısıtlı bir kaynak sunmaktadır [33],[34]. Bu alanda ilk çalışmalardan biri olarak kabul edilen işaret dili tanıma tasarımı 2005 yılında Haberal ve Albayrak tarafından tarafınca kurgulanmış saklı Markov yöntemi kullanılan tasarımıdır. Bu tasarımda önceden oluşturulmuş bir veri seti bulunmaması nedeni ile 172 adet görüntü üzerinden çalışılmış ve başarı oranı %93 olarak tespit edilmiş test iki aşamalı olarak uygulanmış, veri setinde belirlenen sınıflar içerisinde K-En Yakın Komşu (K-NN) metodunda başarı sonucu %91 olarak bulunmuştur [35].

2010 yılında Altun vd. K-NN, SVM yöntemleriyle sırası ile %99.43 ve %98.83 başarılı [44], 2013 yılında Memiş ve Albayrak Kinect sensörü, Özellik çıkartma, Ayrık Kosinüs Dönüşümü (DCT) yöntemi ile ve farklı DCT katsayı oranları için sınıflandırma sonuçlarının karşılaştırılması amacı ile K-NN ile yapılan sınıflandırma sonucu %80-%90 arasında başarılı sonuçlar elde etmişlerdir[45],[46].

2018 yılında Ketenci ve Kayıkçıoğlu. Elektromiyografi (EMG) sinyalleri ile TİD'e ait rakamların tanınması çalışması Fourier dönüşümü ve Çapraz-Korelasyon fonksiyonunun standart sapması özellikleri kullanılarak ayrık lineer analiz sınıflandırıcısıyla Çapraz-Korelasyon kullanılmadığında %63 başarılı ; kullanıldığında ise %84 başarılı sonuçlar almışlardır [47].

2018 yılında Akış ve Akar Öznitelik tabanlı İDT sistemi el bölgesi bölütleme, öznitelik vektörlerini çıkarma, ve SVM kullanarak sınıflandırması ile ESA tabanlı İDT kullanılarak, Öznitelik tabanlı sistem ve ESA tabanlı sistem tasarlamışlardır. Araştırmacıların elde ettikleri başarı oranları sırasıyla %95.31 ve %93.12 olmuştur [48].

2018 yılında Ersoy ve Özdemir tarafından Gauss Karışım Modelinden türetilen Fisher Vektörler kullanılarak SVM yöntemi ile el betimleyicilerinin performansı üzerindeki etkisini analiz etmişlerdir. Araştırmacılar, Çok Katmanlı Algılayıcı (MLP) yöntemiyle %96.67 başarılı sonuçlar elde etmişlerdir [49].

2019 yılında Kın ve Erdem tarafından Evrişimli Yapay Sinir Ağ modeli kullanılarak Aktarımlı Öğrenme metodu ile yapılan çalışmada %90 başarılı bir çalışma yapmışlardır [37].

Hanecioğlu, 2019 yılında yaptığı Türk İşaret Dili için evrişimli sinir ağları ve tekrarlayan sinir ağları ile TV çevirmenliğinde kullanılmak üzere bir mimik tanıma çalışmasını yapmıştır. Araştırmacının, bu çalışmada, kullandığı videolar hem zamansal hem de uzamsal özellikleri içermektedir. Araştırmacı farklı mekansal özellikleri eğitmek için CNN modelini kullanmıştır. Araştırmacı, zamansal özellikleri eğitmek için de RNN modelini kullanmıştır. Araştırmacı, çalışmaları sonucunda, 3B-CNN'ye göre, daha derin CNN modeli olan Inception V3 ve RNN'nin ortaya koyduğu sorunun çözümünde daha uygun olduğu kararına ulaşmıştır. Araştırmacı ayrıca, derin öğrenme yöntemiyle çok sayıda ve çeşitli verilerle daha iyi sonuç elde edebileceğini ileri sürmüştür [38].

2017 yılında Beena, Agnisarman Namboodiri yaptıkları çalışmada Kinect sensöründen toplanan ASL'nin statik hareketlerinin tanınmasına odaklanarak, sistemin işaret dillerini tanıma için sınıflandırıcı tasarım kinect Depth, görüntülerinden CNN mimarisini kullanmışlardır. Sistem, CNN'leri 33000 görüntü kullanarak 24 alfabe ve 0-9 sayı sınıflandırması için eğitilmiştir. Önce çalışmaların karşılaştırıldığında, sınıflandırıcının %

94.6774'lük bir verimlilik elde etmişlerdir. Ayrıca sınıflandırıcı test etmek için basit bir java GUI uygulaması oluşturmuşlar ve eğitim sırasında farklı konulardan daha fazla veri eklediklerinde, doğruluğun arttığını da tespit etmişlerdir.[39] Liao, Xiong, Min, Lu arkadaşlarının 2019 yılında BLSTM-3D artık ağlarla video sırasına dayalı dinamik işaret dili tanıma çalışmasını yapmışlardır. Bu çalışma derin 3 boyutlu artık ConvNet ve BLSTM-3D artık ağ (3D ResNet) olarak adlandırılan çift yönlü LSTM ağlarına dayanan çok modlu bir dinamik işaret dili tanıma yöntemi sunmuşlardır. İlk el, ağ hesaplamasının zaman ve alan karmaşıklığını azaltmak için video karelerinde yerleştirilmekte, ardından, 3B ResNet, video dizilerinden uzamsal özellikleri otomatik olarak çıkarır ve özellik analizinden sonra video dizisindeki her eyleme karşılık gelen bir ara puan oluşturmuşlardır. Son olarak, video dizilerini sınıflandırarak dinamik işaret dili doğru bir şekilde tanımlamışlardır. Deney, DEVISIGN_D veri kümesi ve SLR_Dataset dahil olmak üzere test veri kümelerinde gerçekleştirildiği değinmişler, sonuçlar önerilen yöntemin en gelişmiş tanıma doğruluğunu elde edebileceğini belirtmişlerdir. (DEVISIGN_D veri kümesinde %89.8 ve SLR_Dataset'te % 86.9). Buna ek olarak, 3B ResNet, daha büyük video sekansı verileri sayesinde karmaşık el hareketlerini etkili bir şekilde tanıyabilir olacağını açıklamışlardır [40]. 2020 yılında Shukla, Agarwal, Gupta, Rathi arkadaşların yaptıkları çalışma "ResNet50 Derin Sinir Ağı Mimarisini Kullanarak İşaret Dili Tanıma" dır. Çalışmalarında otomatik işaret dili tanıma için derin öğrenme yaklaşımını kullanmaktadırlar. Parmak uçlu kelimeleri sınıflandırmak için yeni bir 2 seviyeli ResNet50 tabanlı derin sinir ağı mimarisi tasarlamışlar bunun için kullanılan veri kümesi standart Amerikan İşaret Dili el hareketi veri kümesi olduğunu belirtmişlerdir. Veri kümesi ilk önce çeşitli büyütme teknikleri kullanılarak büyütülmüşler, 2 seviyeli ResNet50 tabanlı yaklaşımlarındaki Seviye 1 modeli giriş görüntüsünü 4 setten birine sınıflandırmışlardır. Bir görüntü setlerden birine sınıflandırıldıktan sonra, görüntünün gerçek sınıfını tahmin etmek için karşılık gelen ikinci seviye modeline bir girdi olarak sağlamışlar ve yaklaşımları 12.048 test görüntüsünde % 99.03 doğruluk sağladığını da ayrıca değinmişlerdir [41].

Kim Bae vd. (2004), Kore vizyon dili (KSL) tanıma konusundaki çalışmalarımızın bazı renk sonuçlarını kullanarak bazı erken sonuçlarını bildirmişlerdir. İşaret dili gibi jest kelimelerini tanıyarak, kalıplar çok karmaşık ve çeşitlilik gösterdiğinden, sürekli bir işareti bireysel işaret kelimelerine ayırmanın zor olduğunu belirtmişlerdir. Araştırmacılar, bu sorunu çözmek için KSL'yi desenlerine göre 18 el hareketi sınıfına ayırmışlardır. Araştırmacılar, el hareketinin hızını değiştirerek, bulanık bölümler ve durum otomatlarını

kullanarak, hazırlık hareketi ve işaret sözcükleri arasındaki anlamsız hareket gibi istem dışı hareket hareketlerini reddetmişlerdir. Araştırmacılar, 18 el hareketi sınıfını tanımak için gizli Markov modelini benimsemişlerdir. Bu yöntemleri kullanarak 15 KSL cümlesini tanıyoruz ve % 94 tanıma oranı elde etmişlerdir [42].

2005 yılı itibari ile TİD üzerine yapılan çalışmalar sonucu halen ortak bir veri seti kütüphanesi eksikliği görülmektedir. Unutulmamalıdır ki Transfer Öğrenme metodu ile sınıflandırmada ilk katmanlarda görsellerden oluşan verileri algılamaktan ziyade görseli tüm olarak algılar bu sebeple el işaretleri ile yapılan hareketi tanımaz, katmanlarda üste doğru çıktıkça istediğimiz sınıflandırma düzeyini elde edebiliriz ancak farklı bir birey tarafından aynı işaret yapıldığında öğrenme ilk katmana döneceğinden farklı bireylerce elin yapısı, avuç içi, parmak yapısı, kadın yahut erkek bireyler, elde bulunan yara izi, deri hastalığı, parmak kaybı vb. gibi durumlar göz önünde bulundurularak ortak bir veri seti kütüphanesi oluşturmanın zorunluluk olduğu çıkmaktadır. Ortak kütüphanede veri seti oluşturulup milyonlarca görüntü alındıktan sonra sıradan bir bilgisayar sistemi ile veri seti işlenemeyeceğinden çok güçlü bir bilgisayar tercih edilerek, Transfer Öğrenme metodu ile hatasız ve kullanılabilir bir sınıflandırma yapmak mümkün olabilecek ve hedeflenen tasarıma kavuşulacaktır.

Bu çalışma ile işitme duyusunu kaybı yaşayan bireylerin iletişim sorunlarının en aza indirgenmesi hususunda, teknolojiden faydalanarak çözüme katkı sunabilmek hedeflenmiştir. Gerçekleştirilen çalışmada görsel bir dil olan TİD durağan kelimelere ait hareketlerin derin öğrenme yöntemi ile Yapay Sinir Ağlarının eğitilmesi, Transfer Öğrenmesi metodu ile görüntülerin en az hata ile sınıflandırma sağlanarak bu görüntülü iletişim biçimine ait anlatım biçimlerinin en az hata ile tanınması ve yazı diline çevrilmesi amaçlanmıştır.

1.3. Konunun Amacı

Bu çalışma kapsamı açısından görüntü işleme teknikleri ve yapay zeka ile ilişkilidir. Bilinen en geniş sınıflandırma ve algılama tekniği canlıların görme organı olan göze aittir. Göz, mükemmel algılama ve sınıflandırma kapasitesi ile görüntüyü işleme teknolojilerinin varmayı hedeflediği en mükemmel yapıdır. Bu mükemmel algılama ve sınıflandırma tekniği henüz aynı ile yapay zekâ teknolojilerine kopya edilememiş olsa da yapay zekâ teknolojilerinin, her geçen gün yapılan çalışmalar ile evrime zorlanması sonucu, mükemmel biraz daha yaklaşılmaktadır. Gözümüzün gördüğü canlı yahut cansız her varlığı

yapay zekâ görüntü havuzlarına atıp yapay zekanın, transfer öğrenmesi dediğimiz, insan beyni nöronlarını taklit eden yöntem ile bir model (bir nevi nöron) eğitimi yaparak, mükemmeli yakalamak teorik açıdan mümkün görünmekte ise de, pratikte, süregelen hareketleri tanımlama hususunda henüz varılması hedeflenen “yapay zekanın üstün ırkı” yaratılamamıştır. Bu nedenle çalışmamızda, öncelikle oluşturduğumuz görüntü havuzunda görüntü işleme teknikleri ile sınırlı sayıdaki TİD elemanlarını yapay zekaya tanıtırıp, yapay zekanın bu elemanları öğrenmesini sağlayarak (çalışmanın ilerleyen bölümlerinde ayrıntılı olarak yöntemlere değinilecektir), görüntülerden günlük yazı dili metnine uyarlanarak dönüştürülmesi hedeflenmiştir. İkincil hedef ise; gelecekte oluşturulması arzu edilen büyük görüntü havuzlarına katkı sunabilmektir.

1.4. Yöntem ve Materyal

Bu çalışmanın temelinde derin sinir ağları kullanılarak, TİD alfabesinde tanımlı olan durağan kelimeler ve işaretler (sayı, geometrik şekil ve harfler) için; 3 farklı kişi ile, 32 farklı figür üzerinde fotoğraflanmış, bir figür için 600 sayıda 2 boyutlu görüntü kümeleri kullanılmıştır. Bu tez çalışması kapsamında, sayı, geometrik şekil ve harften oluşan veri tabanı oluşturulmuştur. Veri tabanında toplam 19200 görüntüden elde edilmiş veri seti kullanılmıştır.

Oluşturulan sistem transfer öğrenme metodu kullanılarak eğitilmiş, Evrişimli Sinir Ağları (ESA) tabanlı farklı yapılarla eğitilmiş ve sınıflandırılmıştır. Bu sistem **JetBrains PyCharm Community Edition 2019 1.1** ortamında, diğer programlama dillerine nazaran program kodunu en az çaba ile ve hızlıca yazmaya imkân tanması nedeni ile Python dili kullanılarak Intel® Core™ İ7 -8750h CPU @ 2.20 Ghz ve 16,00 GB RAM 'e sahip Nvidia GTX 1050 ti Ekran Kartı ve Microsoft Windows 10 Home x64 bit işletim sistemli bilgisayar tabanlı yazılım geliştirilmiştir. Elde edilen sonuçlar; geleneksel makine öğrenmesi yöntemleri ile elde edilen sonuçlarla karşılaştırılmış ve en başarılı yöntem tespit edilmiştir. Bunun yanı sıra elde edilen başarılı yöntemlerle TİD için gerçek zamanlı tanıma sistemi oluşturulmuştur.













2. TÜRK İŞARET DİLİ

İletişime ilişkin çeşitli tanımlamalar yapılabilir; Bunlardan biri de kişiden kişiye mesaj aktarımıdır. İletişimin; yazılı iletişim, sözlü iletişim, bilgi aktarımında bilgi aktaran taraf ile bilgi aktarımını alan tarafın aynı ortamda karşılıklı olarak bulunmasını zorunlu kılan görselliğin kullanıldığı görsel iletişim gibi çeşitleri vardır. İşitme engelli bireylerin kullandıkları iletişim biçimi görsel iletişimdir. Türkiye Cumhuriyeti sınırları içerisinde görsel iletişimin dili Türk İşaret Dili (TİD)'dir.

Türk İşaret Dili'nin dünyanın en eski işaret dillerinde biri olduğunu 2003 yılında Türk işaret dili hakkında araştırma yapan Alman asıllı araştırmacı yazar Ulrike Zessan tarafından ileri sürülmüş ve kabul görmüştür. Zessan'ın yazdığı Türk İşaret Dili adlı eser TİD üzerine yazılan neredeyse ilk eserdir. Bu eser haricinde TİD tarihi ile ilgili detaylı bir araştırma halen mevcut olmayıp dil bilgisi kurallarına yönelik çalışmalar ise neredeyse yoktur diyebiliriz. TİD ilk kullanılan işaret dili kabul edilmesine rağmen 1953 yılında sağır ve dilsizlere yönelik eğitim veren okullarda konuşmaya sevkini tembelliştirdiği gerekçesi ile yasaklanması nedeniyle çalışmaların ortaya çıkması, dilbilgisi kurallarının ve cümle örüntülerinin ortaya konulup çoğaltılması, ders kitaplarının hazırlanmasını engellemiştir ve işaret dilinin yaygın olarak kullanılmasını engellemiştir [50].

Bu anlayışın hali bölgeye bağlı olarak konuşma dillerindeki şive farklılığından kaynaklı olmaktan çok, ortak bir işaret dili kümesi, bu kümeyi kullanan eğitim kurumlarının olmayışı, kişinin aile bireyleri ile anlaşmak üzere kurguladığı hareketler sebebi ile gelişigüzelelikten kaynaklanmaktadır. Bu yasaklama 1980'li yılların ortalarına kadar sürmüş olduğundan yasağın kaldırılması ile birlikte sağır ve dilsiz okullarında yeniden bir ortak hareketleri barındıran işaret dili eğitimleri vermeye başlanmıştır. Bugün kullanılan TİD için ortaya konulan en kapsamlı çalışma 2003 yılında Zessan'ın yaptığı çalışma olmakla birlikte, Dilbilgisi kurallarını belirleyen ve cümle yapılarını kapsamlı bir şekilde ele alan sözlük çalışması 2015 yılında TDK tarafından Türk İşaret Dili Sözlüğü adlı eser yayımlanmıştır [51], [52].

Bunun yanı sıra Başak Uludağ ve Serdar Uludağ tarafından oluşturulan bu görsel sözlük isaretce.com adlı sitede arama motorunda kelimeye karşılık gelen hareketi belirtmektedir. Çalışmamızda bu iki kaynaktan yararlanılarak 32 durağan kelimeye karşılık gelen hareketsiz durağan kelimelerden bazıları Şekilde 2.1' de gösterilmektedir.

					
Baklava	Beş	Üçgen	Sıfır	Kare	Kurt
					
Köprü	Aferin	Artı	Başkent	Izci	Zor

Şekil 2.1. Tez çalışmasında kullanılan durağan kelimelere ait görseller (TİD Sözlüğü)

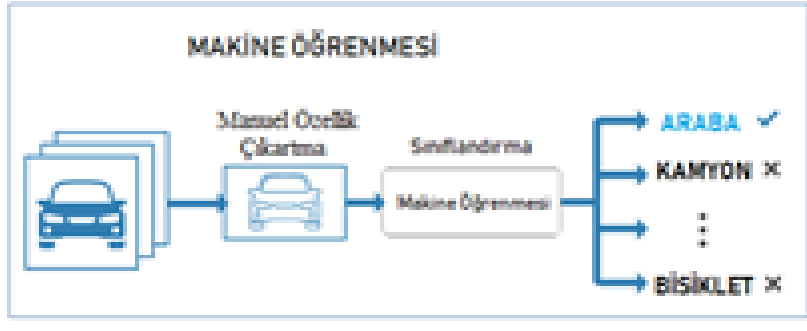
WEB sitesinde yer alan kelimeler bu sınıflardan birine ait olacak şekilde 32 tasnif edilmiş ve birçok örnek kısa cümleye yer verilmiştir [53].

Tez çalışması kapsamında kullanılması planlanan hareket içermeyen sabit kelimelerin gösterimi için bu bahse konu olan kaynaklardan yararlanılmış olup, sözlükte yer alan 32 adet görsel hareket tekrarlanarak fotoğraflanmış ve veri seti oluşturulmuştur.

3. MAKİNE ÖĞRENMESİ

Makine Öğrenmesi, belirli bir probleme yönelik model eğitimi yapılarak sorunun çözüme kavuşmasını hızlandıran, sistemi otomatik hale getiren, yazılımsal kodlama argümanlarının ortak adıdır. Makine öğrenmesini, performansı artırmak veya doğru tahminlerde bulunmak için deneyimi kullanan hesaplama yöntemleri olarak en geniş çerçevede tanımlayabiliriz [54].

Makine öğrenmesini, tanımlanan ve çözümü çıktı olarak vermesi amaç edinilen sorunu veri kümesindeki bilgilere göre tasarlanıp eğitilen model ile tarayarak çözüme yönelik bilgileri çıktı olarak sunan yapay zekanın bir alt birimidir. Makine öğrenmesi metodu, veri kümesindeki örüntüleri tespit ederek, akılcıl karar verme için istatistik, mantık ve hesaplamayı birlikte kullanan yapay zekâ disiplinler bütününe bağıl olarak çözümü istenen konuda eğitilip modellenen algoritmaya göre çözümü çıktı olarak sunan yöntemdir.



Şekil 3.1. Geleneksel makine öğrenmesini gösterimi [85]

3.1. Makine Öğrenmesinde Büyük Verinin Önemi

Büyük veri (big data) geleneksel veri işleme tekniği ile analizi tam ve sağlıklı bir şekilde yapılamayan doğrulamanın sınırlı olarak yapılabildiği durumlarda başvuru genel kümelerdir. Görüntü kavramında ise büyük data resimlerin yer aldığı ve sınıflandırmaların doğru olarak yapıldığı kavram olarak tanımlanabilir. Oluşturulan veri setlerinde görüntülerin sınıflandırılması ve girdi ve çıktıların doğru olarak isimlendirilmesi yönü ile makine öğrenmesinde, yapay zekâ teknolojilerinde ve derin öğrenme metotlarının uygulanması sırasında yapay sinir ağlarının daha kapsamlı bir şekilde eğitilmesinde önemli rol oynamaktadır. Pratikte oluşturulan ve bir amaca hizmet eden veri setlerinin bir araya getirilmesi aynı nesne yahut canlıya ait farklı açılardan alınan fotoğraflar, farklı tipte

canlılar, farklı şekle sahip canlı yahut cansız varlıkların doğru bir şekilde sınıflandırılmasını ve anlamını bize sunar. Çeşitli kareler sürekli olarak üretilip etiketlenerek depolanmakta ve veri kütüphanelerini oluşturmaktadır. Etiketlenen her bilgi veri deposunda yerini aldığından tekrar tekrar model eğitime gerek kalmaksızın sınıflandırmanın yapılmasına imkân sağlamaktadır. İncelenen her biçime ilişkin verinin çokluğu, makine öğrenimi ve YSA eğitiminin daha doğru yapılabilmesine imkân ve olanak sağlar. Bu sebeple verinin doğru sınıflandırılarak depolanması, erişime açık büyük veri kaynaklarına aktarılarak verilere ilişkin görsel kütüphanelerin oluşturulması ile üzerine katılarak ilerlenmesi yapay zekâ araçlarının kullanımının artırılması ve doğru tanımların alınabilmesi konusu önem arz eder. Tıbbi incelemelerde hastalıkların teşhisi konusunda yapay zekâ kullanımı ile hastalık tanısı nasıl ki çok kısa sürede yapılabiliyorsa, eğitim ve diğer alanlarda yaygınlaştırılması ile pek çok hizmetin çok daha az maliyet ile alınabilmesine katkı sunması açısından büyük veri setlerini barındıran kütüphanelerin oluşturulması bu sebeple önem arz etmektedir.

3.2. Makine Öğrenmesine Yaklaşımlar

Günümüzde sıkça duyduğumuz ve büyük veri setlerini oluşturmayı amaç edinen veri madenciliği çeşitli kaynaklardan toplanan, bilginin elde edilip depolanmasını, bilginin doğru kullanımını için analiz ve yönetilmesini, sağlama amacını güden bir bilimsel disiplindir. Veri madenciliğinden elde edilmiş olan tüm bilgi sistemleri, istatistik bilgisinin yanında, bilimsel çalışmalar, fotoğraflar, videolar, sensörler, sosyal medya siteleri, arama motorları, ziyaret edilen blog ve siteler gibi tüm kaynakları kullanarak bilgiyi toplayabilir, depolayabilir, ihtiyaca yönelik hizmet sunmak için kullanılabilir olmak amacı ile bilgiyi elde eden ve yönetmeyi amaçlayan hızlı ve karmaşık, deneyimli makine öğrenmesi ve yapay zekâ algoritmalarını kullanır. Burada deneyim, genellikle elektronik veri ile harmanlanan ve analizler için kullanılabilir halde saklı bulunan bilgilerdir. Makine öğrenmesi denetimli öğrenme, denetimsiz öğrenme ve yarı denetimli öğrenme olmak üzere üç sınıfa ayrılır.

3.2.1. Danışmanlı öğrenme

Bağıl yahut değişkenlerden oluşan bir veri setinde, yazılan her algoritmaya dair bir çıktının yer aldığı kısaca giriş yapılan bilgi ile sistem çıktısının belli olduğu öğrenme yöntemidir. Etiketlenmiş veri seti olarak adlandırılmış bu grup yüklenmiş algoritmalar ile sistem çıktısı arasında bir bağlantı kurmaya çalışarak karar verme yöntemini belirlemeyi

amaçlayarak öğrenir. Karar ağaçları, destek vektör makineleri, yapay sinir ağları, topluluk öğrenme yöntemleri denetimli makine öğrenmesine birer örnek teşkil eder [55].

3.2.2. Danışmansız öğrenme

Veri setinde etiketlemesi bulunmayan dolayısı ile girdilere yönelik sistem çıktısının alınmasının görece daha zor olduğu öğrenme biçimidir. Kümeleme işlemi bu tarz öğrenme biçimine örnektir [56].

3.2.3. Yarı danışmanlı öğrenme

Makine öğrenmesi grupları içerisinde insanın öğrenme sürecine en yakın olan öğrenme biçimidir. Etiketlenmiş verilerin, etiketlenmemiş verilere nazaran daha az olduğu bu öğrenme biçiminde etiketlenmemiş verileri, etiketli olan verilerden hareket ile etiketlemeye çalışır. Günümüz WEB sayfalarının arama motorları bu tarz öğrenmeye örnek gösterilebilir [57].

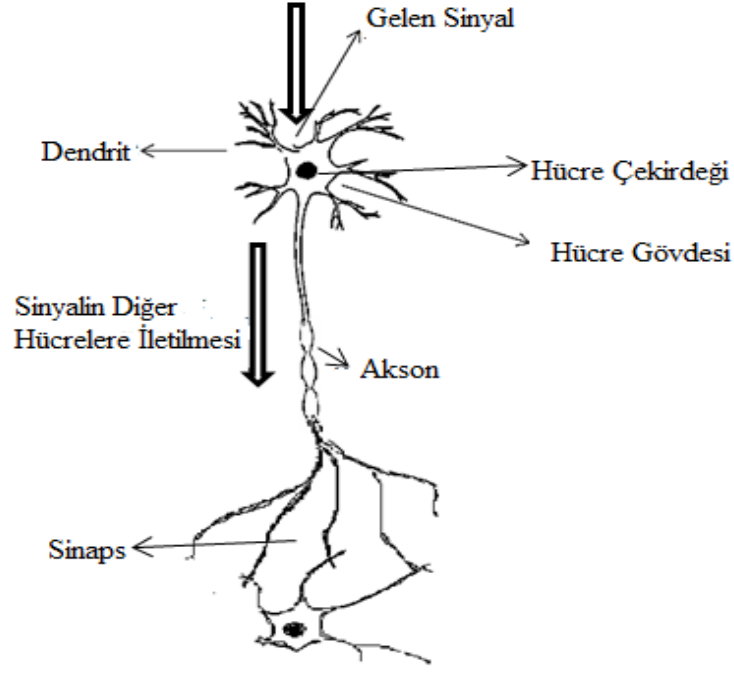
3.3. Yapay Zeka ve Makine Öğrenmesinin Tarihçesi

Cambridge Üniversitesi Profesörlerinden, Alfred North Whitehead ve Bertrand Russell tarafından 1925 yılında yazılan Principia Mathematica adlı üç ciltlik eser ile teorik mantığın temelleri atıldı. 1936 yılında ise Viyana Syntactic Structures'in yazarı Rudolf Canap mantıksal sözdizimi ile destek sundu. Yapay zekâ 1954 yılı soğuk savaş sırasında 1954'ten itibaren ABD ordusu Rus belgelerini ve akademik makalelerini otomatik olarak çevirmek için bir program istiyordu. Bunun için yapay zekâ önem kazandı ancak henüz istenilen düzeyde değildi. Sibernetiklerin kâşifi sayılan Walter Pitts ve Warren Mcculloch 1947 yılında yapay zeka ile kendini kontrol eden bir robot yapmışlardı ama 1951 yılında elektrik kontrolleri bilimi olan sibernetikler henüz ilgi uyandırmıştı. 1969 yılında Marvin Minsky ve Seymour Papert tarafından yazılan An Introduction to Computational Geometry adlı kitaplarında XOR sorunu yani eşdeğerliğin olumsuzlanması sorunu nedeniyle yapay zeka teknolojilerine olan ilgili azaldı. XOR sorunu kısaca modern sinir ağlarındaki gibi işleyişin henüz olmadığı zamanlarda tüm grafiklerin doğrusal olarak verilmesi sorunudur. İki boyutlu bir algılayıcı sisteminde sonucun doğrusal olması ve grafiğin herhangi bir eğri içermemesi sorunudur. XOR problemi doğrusal olmayan ve türevi alınabilen Sigmoid

fonksiyonunun kullanılması ile çözüme kavuş ve XOR sorununun çözülmesi ile birlikte çok katmanlı bir algılayıcı fikri ortaya çıkmıştır [58]. Destek vektör makinelerinin temeli (DVM) 1963 yılında Vladimir Vapnik ve Alexey Chervonenkis tarafından ortaya atılmış olup, istatistiksel öğrenmeyi esas alan, bir çeşit denetimli öğrenme algoritması olarak tanımlanabilir. Her ne kadar temelleri 60'lı yıllara dayansa da 1995 yılında Vladir Vapnik, Bernhard Boser ve Isabelle Guyon tarafından geliştirilmiştir. İlk olarak 1965' te Ivakhnenko ve Lapa tarafından geliştirilen ileri yayımlı perceptron çok ilgi görmemiştir. Yine aynı yıl Nilssons "Öğrenen Makineler" adlı bir kitap yayımlamıştır. Yayımlanan eserde tüm yapılan yapay zekâ ve makine öğrenmesi çalışmaları bir araya toplanmıştır [59] 1969 yılında ise XOR problemi nedeniyle ABD'de yapay zekâ araştırma çalışmalarını yürüten organizasyon olan DARPA yapay zeka ve makine öğrenmesi çalışmalarını desteklemeyeceğini açıkladı. Werbos tarafından 1974 yılında geliştirilen geriye yayılma tekniği ile çok katmanlı perseptron 1986 yılında Rumelhart'ın çalışmaları ile ünlenene dek yapay zekâ ve makine öğrenmesi çalışmaları için duraklama devri devam etmiş, ancak Rumelhart'ın çalışmaları tekrar hız kazanan yapay zekâ teknolojileri ve bu günkü derin sinir ağlarının temeli atılmıştır [60].

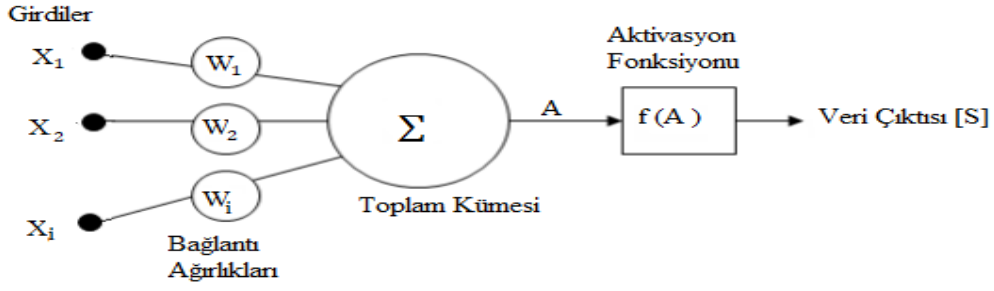
3.4. Yapay Sinir Ağları

İnsan beyni ve öğrenme işlevi dikkate alınarak kurgulanan bir çeşit makine öğrenmesidir. Yapay sinir ağları, biyolojik insan beyni nöronlarını taklit eder. Şekil 3.2 de insan beynindeki biyolojik nöron yapısı verilmiştir. Biyolojik nöronda bulunan dendritlerin görevi, gelen uyarı sinyallerini sinir hücresinin gövdesine iletir, hücredeki kararlı halin bozulmasına sinir hücresinde bir dizi kimyasal sürecin başlamasına neden olmaktadır. Sinir hücresinde vuku bulan bu kimyasal süreçte tüm girdiler bir sonraki sinir hücresine aksonlar aracılığı ile iletilir. Sinir hücresindeki akson ucu ile kendisinden sonra gelen dendrit arasındaki boşluk sinaptik boşluk olarak tanımlanır. Sinaptik boşluk içerisine bilgiyi ileten her dal iletken madde olarak kabul edilir ve bu bilgi iletimi sırasında dolarak bilginin geçişini sağlar. Sinaptik boşluktaki değişkenlik öğrenme süreci olarak tanımlanır.



Şekil 3.2. Biyolojik beyin nöronu yapısı

Yapısı gereği matematiksel algoritmalarından ibaret olan, yapay sinir ağları ise temelde üç ana bölümden oluşmaktadır. Giriş katmanı sisteme gelen verilerden oluşur. Veri seti bilgileri bu katmana işlenir. Özetle bu katman girdi vektörüdür. Orta katman olarak görülen gizli katman ise nöron görevini üstlenen algılayıcıların bulunduğu YSA'nın özüdür. Biyolojik nörondaki tüm görevleri burada gizli katmanda yer alan yapay sinir ağlarını çalıştıran algoritmalar yönetir. Gizli katman nöronları bir önceki katmandan alınan her veriyi bir dizi işlemde geçirerek bir sonraki katmana ileterek, anlamlı çıkışlar sunmaya çalışmaktadır. Giriş katmanındaki bilgiler bu katmanda işlenerek anlamlı çıktıların alınmasını sağlar. Giriş katmanındaki YSA nöronları $X_1, W_1; X_2, W_2; X_i, W_i$ gibi girdi verilerinin aktivasyon fonksiyonundan geçerek algoritmik kararı çıktı olarak sunar. X_1, X_2, \dots, X_i girdi kümesidir. W_1, W_2, \dots, W_i ise YSA nöronları ağırlık vektörü kümesinin elemanlarıdır. Algoritmik fonksiyon $X_i W_i$ girdilerinin çarpımlarının toplam kümesinden aktivasyon fonksiyonuna geçirilerek çıktı olarak sunulur.



Şekil 3.3. Çok katmanlı yapay sinir ağı nöronu çalışma fonksiyonu (F. Rosenblatt'ın 1958 yılında tanıttığı model)[68]

$$S = f(A) = f(X_1 \cdot W_1 + X_2 \cdot W_2 + \dots + X_i \cdot W_i) = f\left(\sum_{p=1}^i X_i W_i\right) \quad (3.1)$$

$$S = f(A) = f(X_1 \cdot W_1 + X_2 \cdot W_2 + \dots + X_i \cdot W_i) = f\left(\sum_{p=1}^i X_i W_i + b\right) \quad (3.2)$$

Burada

S = Verilerin Çıktısını,

A = Nöronun Aktivasyonunu,

f(A) = Aktivasyon Fonksiyonunu,

X_1, X_2, \dots, X_i = Sisteme Yüklenen Girdileri,

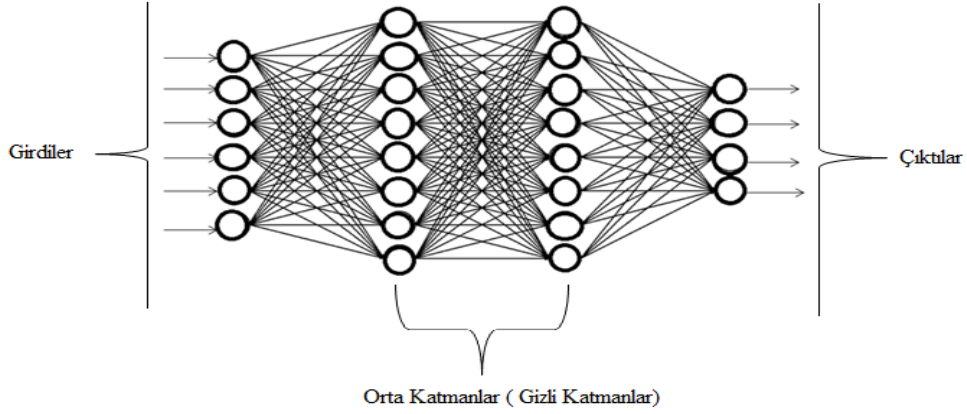
W_1, W_2, \dots, W_i = Bağlantının Vektörel Ağırlıklarını

$X_p = (p=1, \dots, i)$ Girdi Kümesi Olan X'in P Sırasında Bulunan Elemanını,

$W_p = (P=1, \dots, i)$ Ağırlık Vektörü W'nin P Sırasındaki Elemanını,

b = Bias önyargı değerini, ifade etmektedir.

1989 yılında K. Hornik Vd.[68] tarafından yazılan Neural Network adlı eserin 2. Cilt 5. Sayfasında ileri beslemeli yapay sinir ağını şekil 3.4'te gösterildiği gibi tanımlanmıştır. Şekilde insan beyni nöronlarına benzer şekilde görülen çizgiler (\ll) ağırlık vektörleridir. Ağırlık vektörünün atanması düzgün şekilde atanmamış ise bilgi diğer katmana geçmez yahut eksik geçer.



Şekil 3.4. İleri beslemeli yapay sinir ağı katmanları

Burada veri girişi ile giriş katmanında yer alan her yapay sinir ağı nöronu, giriş vektöründeki değerleri ayrı ayrı üstlenerek bir değer elde eder. Giriş vektörü değerleri; RGB bandındaki JPEG formatında bir resme ait piksel değeri olabileceği gibi, ses aralığındaki titreşim değeri gibi sayısallaştırılmış herhangi bir değer de olabilir. Giriş katmanı nöronları, kendilerine atanmış ağırlık değeri vektörleri ile bağlı gizli katman nöronlarına bilgiyi taşır. Gizli katman nöronları $S = f(A)$ başlangıcı ile (3.2) de bağlantısı verilen fonksiyonu tekrar tekrar işler ve her işleşte bias (b) adı verilen bir önyargı değeri ekler. Elde ettiği her değeri kendisinden sonra gelen diğer katmana taşır. Diğer gizli katmanlarda işlenen her değer bir sonraki katmana taşınır, değerlerin sürekli olarak sonraki katmana aktarılması şeklinde ilerlemesinden dolayı ileri beslemeli yapay sinir ağı adını alır. Sistemdeki her gizli katman aynı fonksiyonu işleyerek sonucunu sonraki katmana iletir. Son katman çıktı katmanıdır ki burada girdi katmanında yer alan veri sayısı ile farklılık göstermesi mümkündür. Bu giriş verilerinde değer kaybı yahut sayısallaştırmada sorun olması sebebi ile sınıflandırılmamış olması yahut aynı değer birden fazla kez yüklenmiş olması sebebi ile birlikte öznelik ataması yapıldığı algısını çağırır.

K.Hornik v.d. ; a.g.e.'de, evrensel yaklaşımlarla, standart çok katmanlı ileri beslemeli sinir ağlarının verileri, yakınlık derecesi niteliği ile sınıflandırılabilmesini, yahut derecesi fark etmeksizin bir fonksiyonun sonucunu kabul edilebilir bir sonuç ile tahmin edebileceğini ileri sürmektedir. Başarısız bir tahmin sonucunun oluşması halinde bu sorunun; girdi verilerinin eksik olmasında, gizli katmanların yetersiz olması ile öğrenme eksikliğinin olmasında yahut girdi verileri ile hedeflenen çıktı verileri arasında yeterli derecede ilişki kurulamamasına bağlı olduğunu vurgulamaktadır [61].

3.5. Yapay Sinir Ağları İle Durağan Kelimelerin Tanınmasına İlişkin Yapılan Önceki Çalışmalar

Literatürde, Türk İşaret Dilinde yer alan durağan kelimeler ile hareketli kelimelerin birlikte kullanıldığı bir adet çalışma bulunmaktadır. Bu çalışma 2018 yılında Karaca M., F., tarafından Üç Boyutlu Sanal Model İle Türk İşaret Dili Simülasyonu adlı doktora tez çalışmasıdır[51]. Bu çalışma kapsamında yapay sinir ağları kullanılmış ancak tek başına durağan kelimelere ilişkin bir çalışma değildir. Yapay sinir ağları veya derin öğrenme uygulanarak yapılan durağan kelimelere ilişkin başkaca bir çalışma bulunmamaktadır.

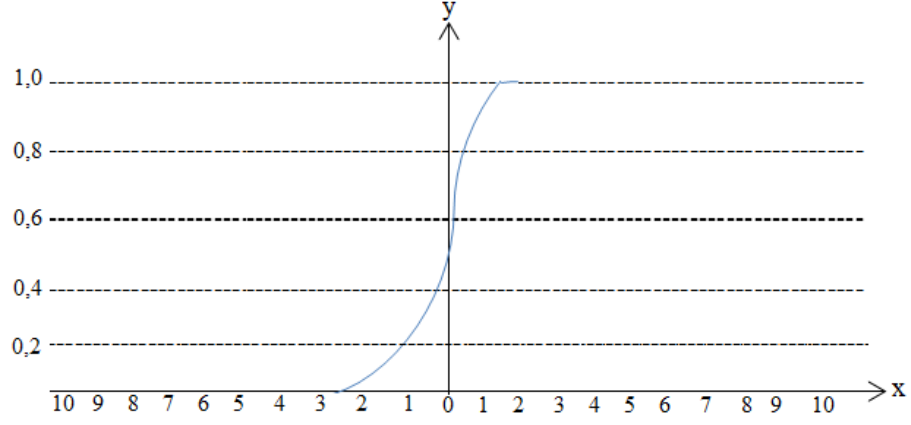
Dünyada işaret dillerinde durağan hareketsiz kelimelere yönelik yapılan çalışmalar ise kısıtlı sayıda mevcuttur. Bu çalışmalarda tek başına durağan kelimelerin kullanıldığı çalışmalar olmayıp hareketli ve durağan kelimelerin bütünleşik kullanıldığı bir çalışma bulunmaktadır. Bu çalışmalardan biri; Ludtke tarafından 2015 yılında yapılan çalışmada Alman İşaret Dili'nde yer alan durağan kelimeleri tanımada yapay sinir ağları ile daha kolay bir çalışma prensibinin geliştirilebilirliğinden bahsetmektedir [62]. Bu çalışma haricinde işaret dilinde durağan kelimelere yönelik yapılan başkaca çalışmaya rastlanılamamıştır.

3.6. Yapay Sinir Ağları Tarihçesi

Yapay sinir ağlarına(YSA) yönelik bir tarih yazılacak olursa 1980 öncesi ve sonrası olarak ikiye ayırma konusunun hiçbir yazar tarafından göz ardı edilemeyecek kadar büyük bir ayırım barındırmaktadır. YSA kullanımının, yapay zeka tarihçesinde kısaca bahsetmiş olduğumuz gibi doğrusal olmayan Sigmoid fonksiyonunu kullanan çok katmanlı algılayıcı modelin tanınması ile birlikte hız kazandığı görülmektedir. Sigmoid fonksiyonu, temelde bulanık mantık kümesi olan $[0,1]$ aralığındaki değerlerden oluşan kümeyi kullanır. Yapay sinir ağları oluşturulurken sıklıkla başvurulan aktivasyon fonksiyonu olarak bilinir. Doğrusal ve doğrusal olmayan davranışlar arasında denge sağlayan ve sürekli artan bir fonksiyon olarak tanımlanmaktadır[63] [73].

$$y = f(x) = \sigma(x) = \frac{1}{1+e^{-x}} \quad , \quad 0 \leq f(x) \leq 1 \quad (4.1)$$

Sigmoid fonksiyonuna ilişkin grafik ise Şekil 3.5 'de yer aldığı gibidir:



Şekil 3.5. Sigmoid fonksiyonu grafiği

Sigmoid Fonksiyonu, nöral kompresyonda genellikle sadece ağırlıklandırılmış toplamların katmanlı ağları ve ardından harici programlamanın kısmen öğrenme ile sübvansede edildiği bir karar fonksiyonu olarak, 1980'lerde bağlantıcılığın güçlü bir canlanmasını sağlayan fonksiyon olarak kullanılması ile karşımıza çıkar. Bu vizyon, sonunda takip problemlerini ortaya koymaktadır [74].

Sigmoid fonksiyonunu kullanan aktivasyon kodlarına ilişkin problemleri şu şekilde özetleyebiliriz[75]:

1. Gerçek nöronların bildiklerini açıklamak için, açıkça yetersiz olan hesaplama modüllerinin kullanımı, yapay nöronların biyolojik gerçeklikten geri dönüşü olmayan bir açıyla uzaklaşmasına yol açmaktadır.
2. Sentez süreçlerinde metodoloji eksikliği ve aşırı ampirizm vardır.
3. Kendini, yapay zekanın sembolik bakış açısıyla bütünleştirilemeyen bir alternatif olarak sunar.

Yapay Sinir Ağlarına ilişkin problem her ne kadar çözülmüşse de 2000'li yılların başlarına kadar hayatın içinde ve hemen her kademesinde var olduğunu söylememiz de gerçeği tam olarak yansıtmayacaktır [76]. 1980 yılında çalışmaları desteklenmeyen ve sıkıntılı sürece karşın çalışmalarına devam eden çok az bilim adamı çalışmalarının karşılığını almaya başlamışlardır [64].

Yukarıda yer alan problemlere ilişkin yapay zekâ ve YSA'lar için farklı aktivasyon kodları kullanılmaya çalışılmıştır. En bilinen aktivasyon kodları: Hiperbolik Tanjant

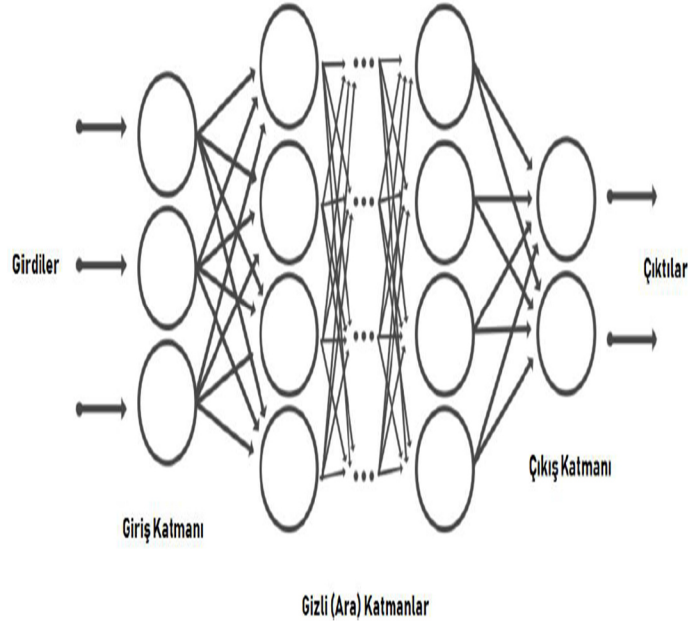
Fonksiyonu, ReLU, Softmax fonksiyonunu kullanan aktivasyon kodlarıdır. Bu aktivasyon kodlarına ilişkin bilgilere derin öğrenme konusunda değinilecektir.

Rumelhard V.d. tarafından 1986 ve 1988 tarihlerinde yayımlanan, “*Parallel distributed processing*” ve “*Explorations in the microstructure of cognition*” adlı eserleri ile bugünkü yapay sinir ağlarının temelleri atılmıştır [65]. 1988 yılında Broomhead ve Lowe tarafından “Radyal Tabanlı Fonksiyonlar” tanıtılmıştır. 1997 yılında Hochreiter ve Schmidhuber tarafından Tekrarlayan Sinir Ağları (TSA) çalışması tanımlanmış ve tanıtılmıştır [66]. Ancak 2000 yılına dek YSA algoritmaları makine öğrenmesine göre çok daha fazla avantaj sunmasına rağmen maliyeti fazla görüldüğünden, model eğitim süresinin uzun olması sebepleri ile tercih edilmeyerek bunun yerine probleme özel tasarlanan destek vektör makineleri tercih edilmiştir. Yapay sinir ağlarının tercih edilebilirliği bilgisayar hız kapasitelerinin artması, grafik işlemcileri (GPU)’un hesaplamalarda kullanılabilir olması ile modellerin daha hızlı eğitilebilmesine imkân tanınması ile olmuştur. [67] 2000 yılında YSA için “Derin Öğrenme” metodu Igor Aizenberg tarafından tanıtılmış ve bu doğru öğrenme metodu ilgi uyandırmıştır.

4. DERİN ÖĞRENME ve EVRİŞİMSEL SİNİR AĞLARI (ESA)

4.1. Derin Öğrenme ve Tarihçesi

İlk kez 2006 yılında Hinton vd. tarafından çok katmanlı yapay sinir ağlarının daha verimli eğitebileceğinin öne sürülmesiyle ortaya çıkmış [79], olsa da derin öğrenme terimi ilk olarak 1987 yılında Rina Dechter tarafında kullanılmış [80] 2000 yılında Igor Aizenberg ve arkadaşları yaptıkları çalışmalarda bir yapay sinir ağı modeli olarak derin öğrenme kavramını ortaya koymuştur. [81] Burada ki derin kelimesi ağ da ki katman sayısını ifade eder. Katman sayısı arttıkça ağ yapısı derinleşir. Yapay sinir ağları genelde bir iki katmandan oluşurken, derin ağlar yüzlerce katmanlardan oluşabilmektedir. Ancak Derin öğrenme 1990 ile 2000'ler arası veri yetersizliği, ana darboğaz ve donanım yetersizliğinden dolayı ışıldayamamıştır. Çünkü sinir ağ modeli derinleştikçe işlem yükü artar ve donanımsal yetersizlikler ortaya çıkmaktadır. Ayrıca ağın geriye doğru eğitilmesi sırasında katman sayısı arttıkça geri bildirim sinyali yok olup gidebilir. Veri yetersizliği de derin öğrenmenin en büyük problemlerinden olan aşırı uydurmaya sebep olmaktadır.



Şekil 4.1. Derin sinir ağlarının genel yapısı [85]

4.1.1. Derin öğrenmenin gelişimi

Danışmansız öğrenme, veri setinde etiketlemesi bulunmayan dolayısı ile girdilere yönelik sistem çıktısının alınmasının görece daha zor olduğu öğrenme biçimidir. Kümeleme işlemi bu tarz öğrenme biçimine örnektir [56].

1- İnternetin yaygınlaşması ile makine öğrenmesinin ihtiyaç duyduğu büyük veri setlerinin toplanması ve dağıtımının mümkün kılınması.

2- 2009-2010 civarında basit ama önemli birkaç algoritmanın keşfi ile hatta sinyalinin geriye daha iyi iletilmesi mümkün hale gelmiştir. Bu algoritmalar:

- a. Katmanlar için daha uygun aktivasyon fonksiyonları (Relu, Softmax gibi),
- b. Daha uygun ilk ağırlık başlatma yöntemleri,
- c. Daha iyi eniyileme (optimizasyon) yöntemleri (RMSProp ve Adam gibi).

3- 2014,2015, 2016 yıllarında yığın normalizasyonu (batch normalization), artık bağlantılar (residual connections) ve derinlemesine evrişim (depthwise separable convolution) gibi hata sinyalinin geriye daha iyi yayılımını sağlayan bu yöntemler ile günümüzde, yüzlerce katmandan oluşan ağları sıfırdan eğitebilir durum geldi [84].

4- Bilgisayar oyunları pazarı için geliştirilen ekran kartları ile daha verimli paralel hesaplamaları gerçekleştirebilmesidir.

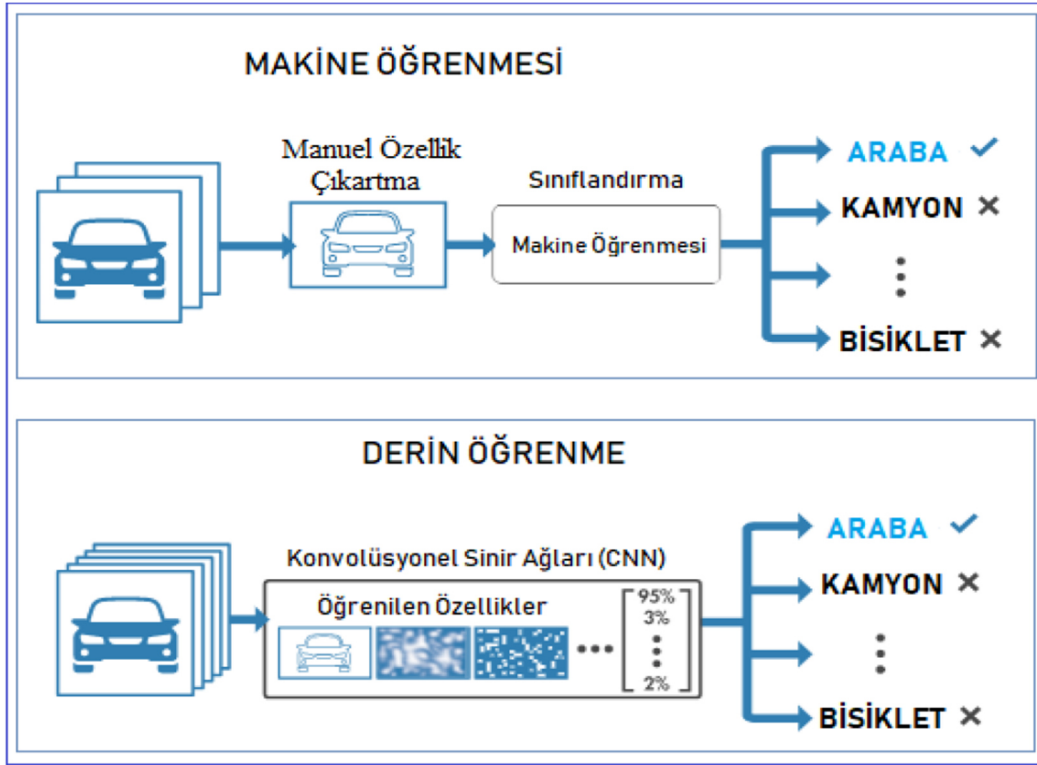
Sonuç olarak derin öğrenme bugün sağlık, eğitim, iletişim gibi sektörlerde doğal dil işleme, örüntü tanıma, konuşma tanıma gibi birçok alanlarda yaygın bir biçimde kullanılmaktadır. Bu çalışmada da derin öğrenmeyi TİD durağın kelimeleri tanımada kullanılmıştır.

4.2. Derin Öğrenmenin Güçlü Yönleri

Derin öğrenme, tüm katmanlarının gösterimlerini sıralı olarak öğrenmek yerine bütünleşik olarak öğrenir. Her şey tek bir geri bildirim sinyali tarafından kontrol edilir ve modeldeki her değişim hedefe ulaşmaya hizmet eder. Böylece sıg modelleri üst üste sıralama yerine gösterimleri ara katmanlardan birbirinden soyut olarak öğrenebilmek mümkün olmaktadır [84].

Derin öğrenmeyi farklı kılan en önemli özelliklerinden biri; Klasik makine öğrenme tekniklerinde veriyi kullanmadan önce ön işlem den geçirerek veriyi model için anlamlı kılmak gerekmektedir. Bunun içinde problemin çözümünü kolaylaştırmak için girdilerin daha ayırt edici olmasını sağlayan öz nitelikler çıkarılır. Derin öğrenme, makine

öğrenmesinin bu kritik adımını tamamen otomatikleştirmiş olarak kendi içinde yapması nedeniyle problemlerin çözümünü kolaylaştırmıştır. Şekil 4.2’ de görüntüde bulunan girdilerin tahmin edilmesinde hem klasik makine öğrenmesi hem de derin öğrenme modellerinin nasıl bir yol izlediği gösterilmiştir [85].



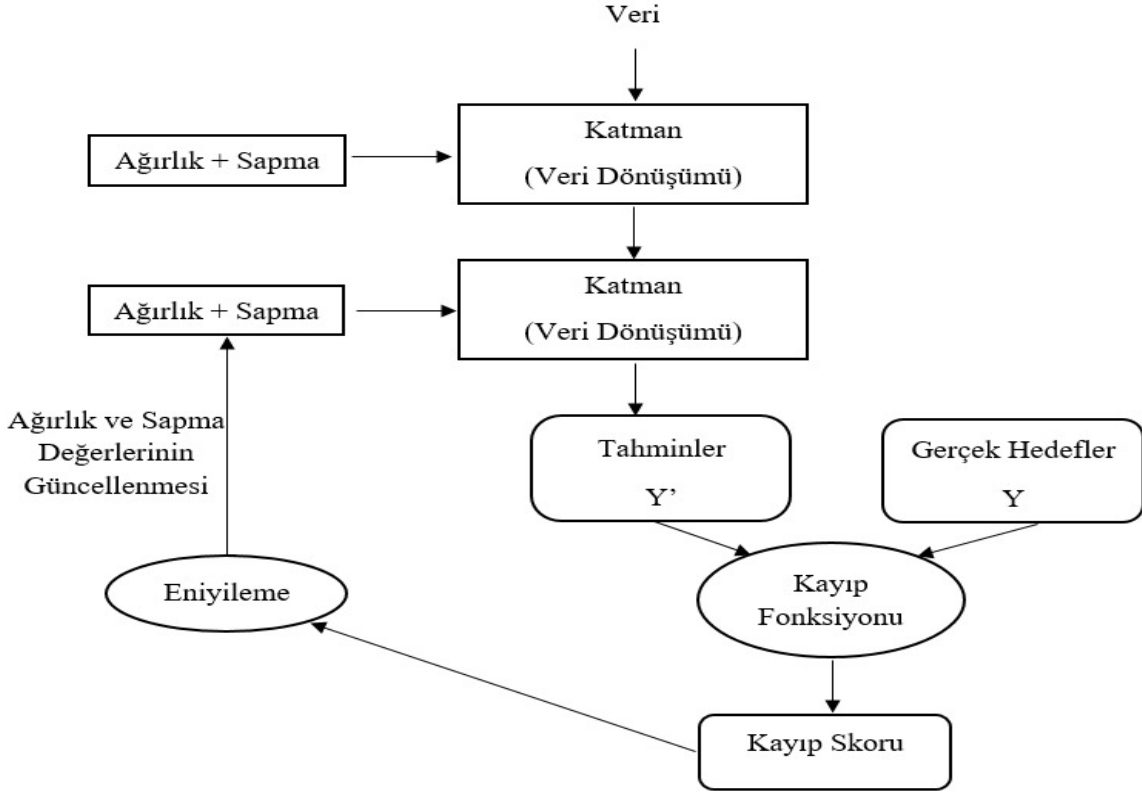
Şekil 4.2. Derin öğrenme ve geleneksel makine öğrenmesi yaklaşımının karşılaştırılması[85]

Geleneksel makine öğrenmesi tekniklerinde eğitim verilerinin artırılması her zaman başarıyı olumlu etkilemezken; derin öğrenmede eğitim verisi çeşitlenip artıca başarı oranı da artar [85]. Bu durumda derin öğrenme için çok miktarda eğitim verisine ihtiyaç vardır. Geleneksel makine öğrenmesinde başarılı bir sonuç için az miktarda eğitim verisi yeterli olabilir.

Derin sinir ağlarıyla uygulama yapılırken genelde 3 farklı yöntem uygulanır. Birinci olarak problemin çözümüne uygun derin ağ modeli oluşturularak eğitimin gerçekleştirilmesi, ikinci olarak önceden eğitilmiş hazır bir ağ yapısını yeni verilerle eğitilmesi ve son olarak da derin sinir ağı yapısından öznelik çıkarımı yapılarak diğer makine öğrenme yöntemleri ile kullanılmasıdır. Gerçekleştirilen bu çalışmada TİD durağın kelimelerine ait işaretleri tanıyabilmek için; Her üç yöntemde uygulanmıştır.

4.3. Derin Öğrenmede Etkili Olan Kavramlar

Derin ağ yapısı, zincirleme eklenen katmanlardan oluşur ve girdi verisini tahminlere eşler. Sonrasında kayıp fonksiyonu vasıtasıyla tahminler ile hedefler arasındaki fark olan kayıp değeri bulunarak ağırlık tahminleri ile hedeflerin ne kadar eşleştiği ölçülür. Eniyileme algoritması kayıp değerini kullanarak ağırlık ve sapma değerleri güncellenir.



Şekil 4.3. Katmanlar, kayıp fonksiyonu ve eniyileme algoritması arasındaki ilişki

4.3.1. Ağırlık

Eğitilen parametrelerden bir tanesi İngilizce de weight anlamına geldiği için genellikle W harfi ile gösterilir. Siner ağı tarafından eğitim esnasından daha önemli olduğu düşünülen girdilere daha yüksek ağırlık değeri verilmektedir.

4.3.2. Sapma

Eğitilen parametrelerden bir tanesi İngilizce de bias anlamına geldiği için genellikle b harfi ile gösterilir. Ağırlıktan farklı ek olarak uygulanan doğrusal bileşene sapma denir. Bu değer; ağırlıklı katsayılı girişin aralığını temel olarak değiştirmek için eklenmektedir.

4.3.3. Aktivasyon fonksiyonu

Bu fonksiyonlar tensörlere etki eden doğrusal olmayan operasyonlardır. Amaç ağırlık ve sapma değerlerin ayarlanmasıdır. Bunu neronların tetiklenmesi olarak düşünebiliriz. Yüksek bir değer geldiği zaman neron daha yüksek bir sinyal verecektir.

Yapay zekâ ve YSA'da kullanılması ile neredeyse devrim yaratan Sigmoid Fonksiyonu hakkında bilgiler çalışma kapsamında ilgili bilgiler verildiğinden diğer en sık kullanılan fonksiyonlara ilişkin bilgilere değinilecektir. Bunlar sırası ile ; Hiperbolik Tanjant Fonksiyonu, ReLU Fonksiyonu ve Softmax Fonksiyonlarıdır.

1. Hiperbolik Tanjant Fonksiyonu (Tanh); Sigmoid Fonksiyonunun [-1 ve 1] aralığını kapsayan fonksiyondur. Eşikleme 0,5 te değil 0 da sağlanır. Bu sebeple 0 dan küçük olan sayılar 0 (sıfır) olarak kabul edilmeyerek -1 'e yaklaşan değerler alır ve fonksiyon denklemi şu şekilde formüle edilir:

$$, -1 \leq f(x) \leq 1 \quad (4.1.)$$

2. ReLU Fonksiyonu; Düzeltilem doğrusal birim aktivasyon fonksiyonudur. Fonksiyon aynı Sigmoid fonksiyonunda yer alan sorunu yansıtır. Bu sorun 0'dan küçük değerleri 0 değeri gibi kabul etmektedir. Pozitif değerler x değerini alırken negatif değerler 0 olarak kabul edilir ve şu şekilde formülize edilir:

$$y = f(x) = \max(0, x), \quad 0 \leq f(x) \quad (4.2.)$$

ReLU fonksiyonu, Sigmoid ve Hiperbolik Tanjant fonksiyonlarını kullanan aktivasyon kodlarına nazaran daha YSA eğitimi sırasında daha hızlı çalıştığına tecrübe edilmiş olması nedeniyle, ağı öğrenmesi sürecinde sıkça kullanılan bir fonksiyon olmasına rağmen tıpkı Sigmoid Fonksiyonu aktivasyon kodunda olduğu gibi dezavantajlara sahiptir. Bu dezavantaj negatif değerlerin 0 olarak kabul edilmesidir. Bu durum 0'dan küçük ağırlık değerlerini alan perseptronların diğer değışle nöronların yok olmasını ve bir daha eğitim sürecinde kullanılamaması sonucunu beraberinde getirmektedir. Bu dezavantajın önüne geçilmek için sızdırmalı (leaky) ReLU fonksiyonu ileri sürülmüştür. Bu fonksiyonun formülasyonu şu şekildedir:

$$y = f(x) = 1(x < 0)(\alpha x) + 1(x \geq 0)(x), \quad (4.3.)$$

Fonksiyondaki α sabiti genellikle küçük bir sayı olarak eklenir ve 0'a yakın bir değer elde edilir (0,1) gibi ve bu değer negatif ağırlık değeri ile çarpılmaktadır.

3. Softmax Fonksiyonu;

Çoklu sınıflandırma işlemleri için kullanımına en çok başvuru fonksiyon olmakla birlikte, girdiler için bir sınıflandırma olasılığına bağlı kayıp değeri üreten fonksiyondur. Sınıflandırıcı, girdilerin gerçek sınıfa ait dağılımı ile ait olma olasılığı bulunan kümenin

dağılımı (en çok olabilirlik fonksiyonunu) normalize eder ve iki değer arasındaki çarpaz entropi değerinin en aza indirgenmesi mantığı ile hareket ederek çıktıları sunar.

Formülü ise şu şekildedir:

$$\sigma: \mathbb{R}^n \rightarrow \text{int}(\Delta^{n-1}), \quad (4.4.)$$

$$, \lambda > 0 \quad (4.5.)$$

softmax işlevi $\sigma: \mathbb{R}^n \rightarrow$ ile verilir

$\Delta = 1$ olduğunda, (1) 'e standart softmax işlevinde ise $\Delta \rightarrow 0$ olarak, σ çıkışı noktası merkezinde akıllıca, yani tekdüze bir olasılık dağılımı uygular. Öte yandan, $\lambda \rightarrow \infty$ olarak, σ çıkışı noktasal olarak $e, j \in \mathbb{R}^n$ 'ye yakınsar, burada $j = \text{argmax}_{1 \leq i \leq n} e_i$, z , değerleri ile ifade edilir. İki veya daha fazla bileşen arasındaki fark z 'nin çok küçük olmaması gerekir [69].

Bu fonksiyonlara ek olarak farklı Derin Öğrenmede kullanılan farklı aktivasyon kodları da bulunmaktadır. Derin öğrenmede en sık başvurulan aktivasyon fonksiyonlarına yer verilmiş olup, bu çalışma kapsamında, sınıflandırma için önerilen ve kullanacağımız aktivasyon kodu ise softargmax, argmax ya da yanlış olmasına rağmen en bilinen adı ile softmax fonksiyonu olacaktır.

4.3.4. Öğrenme katsayısı ve eniyileme algoritması

Öğrenme katsayısı, başkaca bir söylem ile hata düzeltme katsayısıdır. YSA'da aksion görevi gören taşıyıcı sinaptikleri eğitim düşümü algoritmalarında kullanılan bir katsayıdır. Öğrenme katsayısı, eğitim düşümüne göre tanımlanmış algoritmalarının yaklaşmasını sağlamaktadır. Öğrenme katsayısı büyüdükçe sinaptiklerin bir sonraki katmana bilgiyi taşıyamamasına neden olur. Bu katsayı için çok küçük bir değer atandığında ise öğrenme sürecinin uzamasına etki eder.

Öğrenme katsayıları için çeşitli eniyileme algoritmalar kullanılır amaç Eniyileme algoritması, öğrenmenin nasıl gerçekleşeceğini belirlemektir. En sık kullanılan algoritmalar ise, eğimine göre, Adagrad, Adamax, Nadam, Adadelata, RMSprop ve Momentum algoritmalarıdır.

4.3.5. Kayıp fonksiyonu

Ağın doğruluk değeri, kayıp fonksiyonu kullanılarak hesaplanmaktadır. Bu fonksiyon, ağ, gerçek değerlerinin tahmininde hata oluşursa cezalandırmaya çalışmaktadır. Sınıflandırma problemlerinde genelde kategorik çapraz entropi fonksiyonu kullanılır. Bu tez çalışması bir sınıflandırma problemi olduğu için kullanılan derin öğrenme modellerinde bu fonksiyon kullanılmıştır.

4.3.6. Geri yayılım

Tek bir iterasyon ile elde edilen çıktı için ağın hatası hesaplanabilmektedir. Elde edilen hata değeri, maliyet fonksiyonun eğimi ile birlikte ağın ağırlıklarının, güncellenebilmesi için ağa geri beslenmektedir. Sonraki iterasyonlar da hatanın minimize edilmesi için bu ağırlıklar güncellenir. Kayıp fonksiyonun eğimi kullanılarak ağırlıkların güncellenmesi geri yayılım olarak ifade edilmektedir.

4.3.7. Yığın, epok ve tekrar sayısı

Yığın, model tarafından işlenmesi için bir defada alınan miktar. Tüm girdinin tek seferde gönderilmesi yerine girdiler rastgele seçilmiş eşit büyüklükte birbirinden bağımsız yığınlara bölünerek ağda eğitilir. Her bir yığın kümesi işlendiğinde bir tekrar, tüm girdi yığınlarının bir defada işlenmesine de epok denir.

4.4. Derin Öğrenme Mimarileri

Farklı derin öğrenme mimarileri vardır. En çok kullanılanlar;

- Tekrarlayan Sinir Ağı -Recurrent Neural Networks (RNN):

Ara katman çıkışı hem bir sonraki katmana hem de tekrar aynı katmana giriş olarak vermesinden dolayı tekrarlayan sinir ağları denir. Sınıflandırma problemlerinde kullanılabilen bu yapı ses tanımadaki çalışmalarında başarılı olduğu görülmüştür.

- Kısıtlanmış Boltzmann Makineleri (Restricted Boltzmann Machines) (RBM):

Sınıflandırma, bağlanım ve özellik öğrenimi işlemlerinde kullanılan boltzman makinesi giriş veri seti üzerinde olasılık dağılımlarını öğrenebilen bir sinir ağıdır.

- Uzun Kısa Dönemli Bellek (Long / Short Term Memory) (LSTM):

RNN'lerin dezavantajı olarak bilinen bağlam boşluklarının tahmin edilmesi ihtiyacından dolayı Hochreiter ve Schmidhuber tarafından 1997 yılında LSTM'ler ortaya çıkmıştır [ref ver]. Gizli durumları hesaplamak için LSTM ağlarında bir yapı kullanılır. LSTM içerisinde hafıza hücreleri yer alır. Bu hafıza hücreleri önceki durum ile girdi bilgisini tutan hücrelerdir. Ağ mimarisi içerisinde yer alan bu hücreler hangi verinin tutulacağına ya da hangi verinin sileceğine karar verirler. Sonraki aşamada ise önceki durumu mevcut bellek ile giriş verisini birleştirirler. Böyle bir yaklaşımla uzun vadeli bağımlılıkların ortadan kaldırılarak veri dizilerinin devam ettirilmesi mümkün kılınır.

- Derin Oto Kodlayıcılar (Denoising Autoencoders) (DAE):

Danışmansız öğrenme yaklaşımının sahip özel bir yapay sinir ağıdır. DAE, girdi katmanındaki değerleri çıktı katmanına kopyalayan bir sinir ağıdır. Yani, sinir ağına girdi olarak verilen veriyi, çıktı katmanında tekrar oluşturur. AE, kısaca girdi verisinin sıkıştırılmış gösteriminden en iyi özelliklerin öğrenilmesini hedefleyen bir ileri beslemeli sinir ağıdır.

- Evrişimsel Sinir Ağları (Convolutional Neural Networks) (CNN):

Evrişimli sinir ağları oluşturulurken görsel korteksin nörobiyolojik modelinden esinlenilerek geliştirilmiştir. Bu anlamda bu ağ yapısı, görsel nesne tanımda daha etkindir. Çok sayıdaki evrişim ve havuzlama katmanına sahip ileri beslemeli bir sinir ağıdır.

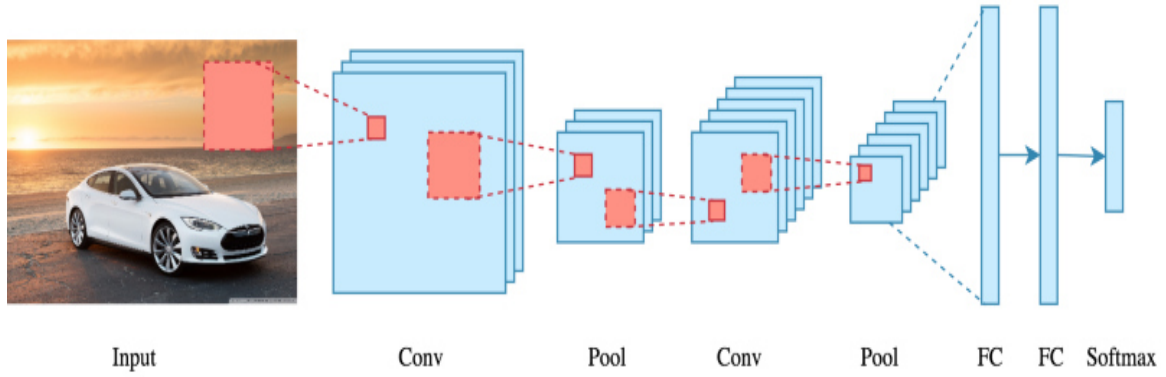
Bu tezde görüntü tabanlı bir sınıflandırma işlemi yapıldığı için yukarıda belirtilen modellerden görüntü sınıflandırmadaki başarısıyla diğerlerinden ayrılan evrişimsel sinir ağları tercih edilerek, TİD durağın kelime tanıma sistemi gerçekleştirilmiştir.

4.5. Evrişimsel Sinir Ağlar (ESA)

İsmi evrişim katmanında alan bu ağ yapısı ilk olarak el ile yazılmış rakamları tanımak için LeCun tarafından tanıtılarak literatüre kazandırılmıştır. ABD Posta Servisi tarafından zarflarının üzerinde yazan posta kodlarını okumak için kullanılmıştır [70]. Evrişimsel katmanlar daha çok özelleştirilmiş doğrusal işlemleri barındırır. Doğrusal fonksiyonlardan seçilerek belirlenen fonksiyon kullanılır ve bir katmandan az olmayacak halde matris çarpımları yerine evrişimsel işlem uygulaması yapan sinir ağları olarak tanımlanabilir [71].

Evrişimsel sinir ağları, görüntü gibi 2 boyutlu veriler için kullanımının uygun olduğu genel kanısını taşımaktadır. Bu genel kanıya evrişimsel katmanda yer alan her bir gizli evrişim filtresinin, girdisini nöron aktivasyonlarının 3 boyutlu bir çıktısına dönüştürme

becerisinden varılmaktadır. Evrişimli sinir ağı oluşturulurken görsel korteksin nörobiyolojik modelinden esinlenilerek geliştirilmiştir. Bu anlamda bu ağ yapısı, görsel nesne tanımada daha etkindir [72]. Şekil 4.4 'de evrişimli ağlarının katmanları gösterilmiştir.



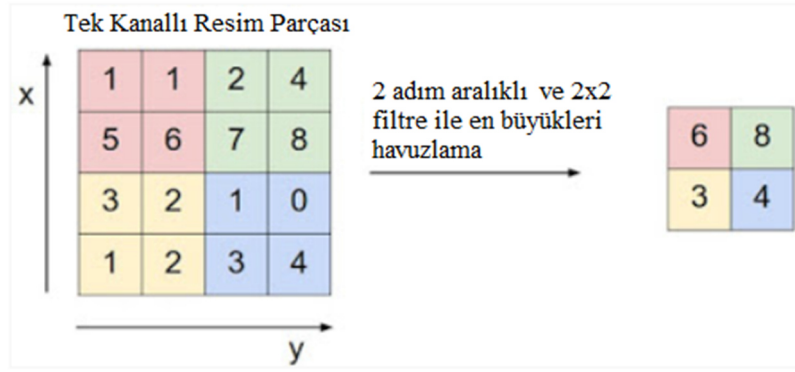
Şekil 4.4. Evrişimli ağlarda bilgisayar mimarisi katmanları [86]

ESA mimarileri temelde beş katmandan oluşan modellemelerden ibarettir. Bu katmanlar sırası ile, giriş, evrişimsel katman, havuzlama, tam bağlı katman ve çıktı katmanlarıdır.

Giriş Katmanı: İsminden anlaşılacağı üzere ESA 'nın ilk katmanını oluşturulur. Girdi herhangi bir işlemde geçmeden ağına verilir.

Evrişim Katmanı: Izgara benzeri bir topolojide girdilerin evrişim adı verilen matematiksel işlemde geçirildiği katmandır. Doğrusal olan evrişim işlemi ile girdilerin bölgesel örüntüleri öğrenilir. Bu sayede bir resmin sol alt köşesinde öğrendiği bir örüntüyü farklı bir yerde konumlandığı zaman yeniden eğitilmeden tanıyabilir.

Havuzlama Katmanı: Bu katmanda kanal sayısı sabit tutulur. Belirlenen filtreler içerisindeki piksellerin maksimum değerleri veya ortalaması alınarak yükseklik ve genişlikte boyut küçülmesi yapılır. Bu işlem bilgi kaybına yol açar ancak az katsayıyla işlem yapmak hesaplama yükünü azaltır ve evrişim katmanlarını her seferinde daha büyük pencerelere bakması sağlanmış olur.



Şekil 4.5. 4x4 boyutunda bir görselde 2x2 boyutunda bir filtre ile görüntü işleme biçimi [77]

Tam Bağlı Katman: Bu katmanda her girişin tüm nöronlara bağlı olduğu için tam bağlı katman denir. Genelde modelin sonlarına doğru konularak sınıf tahminlerini optimize etmek için kullanılır.

Çıktı Katmanı: Bu katman sınıflandırmanın olduğu katmandır. Genellikle sınıflandırmada başarılı olan softmax aktivasyon fonksiyonu tercih edilir. Kaç sınıfsa varsa o kadar çıkış değeri vardır.

Evrişimsel sinir ağları için tasarlanabilecek birçok farklı modelden bahsetmek mümkündür. Model seçimi, problemin çözülmesinde önemli bir etken olarak kabul edilmektedir. 2 Boyutlu görüntülerin ayırt edilmesi amacı ile tasarlanmış ve başarısı kanıtlanan pek çok model bulunmaktadır. Temelde yer alan katmanların farklı kombinasyonlarla dizilimi, girdi ve çıktı katmanları hariç tutularak, farklı mimari modellerin tasarlanmasına imkan tanımaktadır. İnternet ortamında bulunan resimlerin toplanıp etiketlenmesiyle oluşan 1000 kategoride yaklaşık 1.4 milyon görüntüyü barındıran ImageNet veri seti üzerinde 2012 yılında düzenlenen Büyük Ölçekli Görsel Tanıma Yarışmasında (ILSVRC) birinci olan AlexNet[86], derin öğrenmeyi daha da yaygınlaştırmış aynı zaman esa yapılarının görüntü tanımda ki başarısını kanıtlamıştır[87].

İşaret dilindeki durağan kelimelerin tanınmasında, derin öğrenmenin yaygın olan üç yöntemi de kullanılmıştır. Bunlar sırasıyla problemin çözümüne yönelik bir modelin oluşturularak eğitilmesi, önceden eğitilmiş modelin aktarımlı öğrenme yolu ile yeni verilerle eğitilmesi ve derin öğrenme modelinin yalnızca öznetelik çıkartılması sağlanarak geleneksel makine öğrenme yöntemleri ile sınıflandırılması.

4.6. 10 Katmanlı ESA YAPISI

Oluşturulan ESA(CNN) yapısının giriş katmanına 50x50 piksel boyutunda rgb görüntüler verilmiştir. Birinci evrişim katmanında 2x2 boyutunda 16 adet filtre kullanılmıştır. İkinci evrişim katmanında 5x5 boyutunda 32 adet filtre kullanılmıştır. Üçüncü evrişim katmanında 5x5 boyutunda 64 adet filtre kullanılmış ve her evrişim katmanından sonra en büyükleri havuzlama katmanları koyulmuştur ve düzleştirme işlemi yapıldıktan sonra iki tam bağlantılı katman eklenerek model oluşturmuştur. Son katmanda aktivasyon fonksiyonu olarak sınıflandırma için uygun olan softmax fonksiyonu seçilmiştir. Diğer katmanların aktivasyon fonksiyonları ReLu fonksiyonudur. Eniyileme algoritması olarak Stochastic gradient descent (SGD) kullanılmıştır. Aşırı uydurmayı önlemek için bırakma(dropout) tekniği uygulanmıştır. Bu teknikten kısaca değinirsek;

Bırakma, Sinir ağlarında ezberlemeyi engelleyen en etkin ve yaygın tekniktir. Eğitim esnasında o katmanın öğrendiği bilginin bir bölümü rastgele sıfır yapılır. Amaç eğitim esnasında nöronlara bağımsızlık kazandırılarak ağırlıkların daha iyi eğitilmesini sağlamaktır.

4.7. Aktarımlı Öğrenme

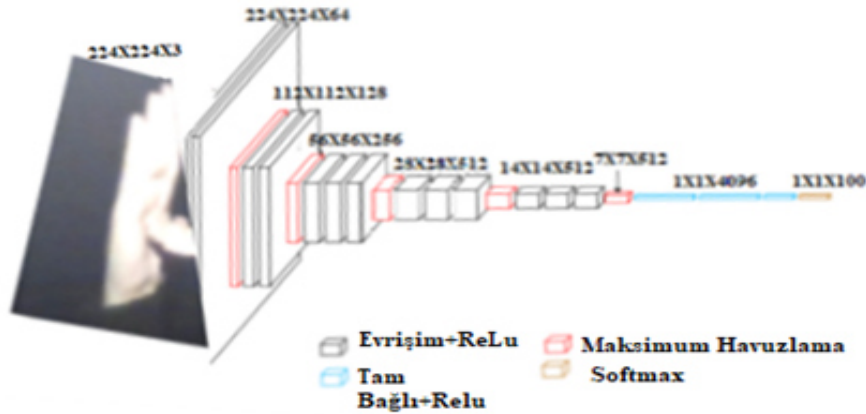
İnsanlar bir alanda öğrendikleri tecrübeleri, benzer özelliklerdeki başka bir problemi çözümünde kullanırlar. Örneğin araba sürmeyi öğrenmiş bir birey tecrübelerinden faydalanarak kamyon sürmeyi daha kolay öğrenir. Bu kapsamda yapay sinir ağlarında bilgi ya da tecrübe eğitilen modelin ağırlık ve sapma değerlerinde saklanmaktadır. Belli bir alanda eğitilmiş bir modelin o alana benzer problemde tamama men ya da kısmen yeniden eğitilmesine aktarımlı öğrenme (transfer learning) denir. Sonuçta daha önce eğitilmiş modelin, eğitilmemiş bir modele göre daha hızlı öğrenmesi ya da daha yüksek başarı el etmesi beklenir. Bu sayede görüntü dünyası için anlamlı özellikler sunarak eğitilen sınıflardan tamamen farklı bir problem daha az veri ile çözülebilir. Bu çalışmada da bu yöntemden faydalanılarak sınıfların çoğu hayvanlar ve günlük hayatta kullanılan nesnelere oluşan ImageNet veri setini sınıflandırama da başarılı olmuş ve birbirlerine göre farklı mimarilere sahip olan VGGNET, Inception ve ResNet ESA modelleri kullanılmıştır.

4.7.1. VGG16

VGG16, 2014 yılında ILSVR (Imagenet) yarışmasını kazanmak için kullanılan bir evrişimli sinir ağı (ESA) mimarisidir. 2014 yılındaki yarışmada başarılı olan VGGNet evrişimli sinir ağı yapısı Simonyan ve Zisserman tarafından geliştirilmiştir. 16 katmandan oluşan çok düzgün bir yapısı vardır. Mimari olarak AlexNet benzemektedir Bugüne kadar üretilmiş mimariler arasında mükemmel vizyon modeli mimarisinden biri olarak kabul edilmektedir. Ağı yaklaşık olarak 138 milyon parametreden oluşmaktadır.

VGG16 ile ilgili en eşsiz şey, çok sayıda hiper parametreye sahip olmak yerine, tek adımda 3x3 filtrenin konveksiyon katmanlarına sahip olmaya odaklanmasıdır. Diğer taraftan VGG16 yapısında, evrişim ve maksimum havuz katmanları tüm mimari boyunca tutarlı bir şekilde 2x2 filtrelerde 2 adımlı olmak aynı dolgu ve en büyükleri havuzlama katmanının kullanılmış olmasıdır. Son katmanda ise tamamen bağlı katmanlar olan 2 tam bağlı katmanı ve ardından çıktı için bir Softmax kullanılmıştır. VGG16, ağırlıkları olan 16 katmanı olduğunu ifade eder. Bu ağı oldukça büyük bir ağıdır ve yaklaşık 138 milyon (yaklaşık) parametreye sahiptir.

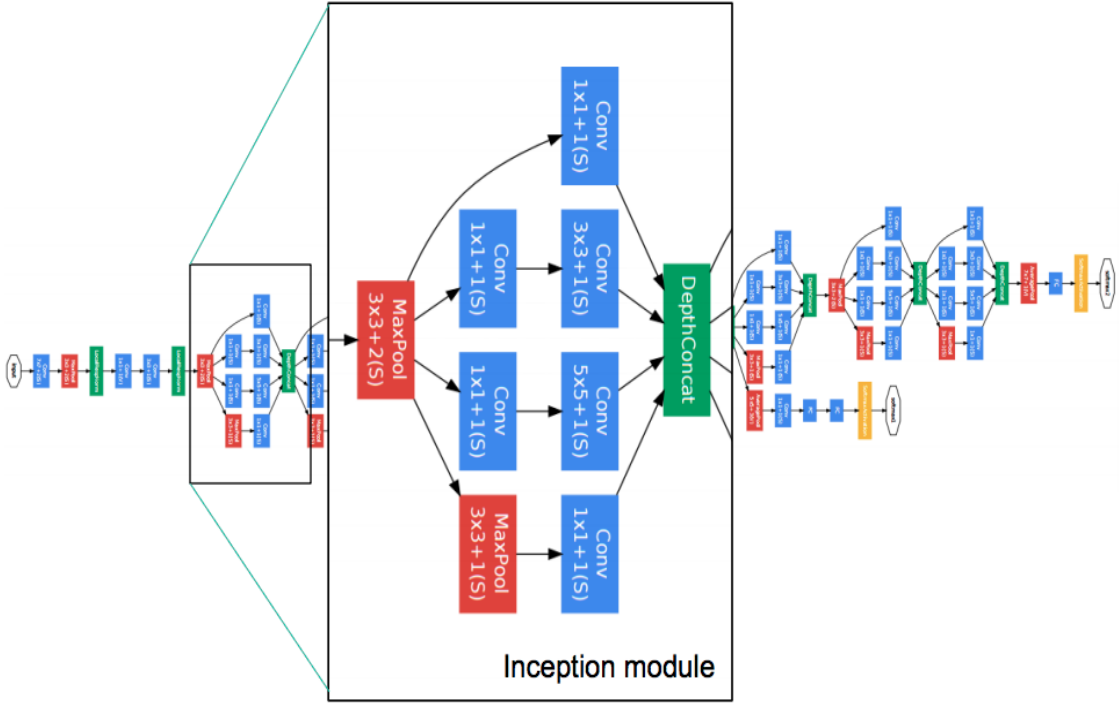
Şekil 4.6.'dan de anlaşılacağı üzere VGG 16 mimarisinde 224x224 ebatlarında bir görüntü tanıma ve sınıflandırma öğrenimi pixel bölümlenme ile sağlanır [78] ve ağı eğitimi yapılır. Böylelikle ağı sınıflandırmayı öğrenmiş olur ve görüntü çözümlemesinde ağı öğrenmede doğruluk payı artmaktadır [82].



Şekil 4.6. VGG16 mimari yapısı ve katmanları

4.7.2. Inception

Google ekibinden Szegedy ve ekibi tarafından geliştirilen yapı 2014' de ki ImageNet yarışmasının kazananı olmuştur. Temelinde LeNet mimarisi olan ancak yapıya ağ içinde ağ mimarisini uygulayarak başlangıç(inception) adı verdikleri yeni modül oluşturmuşlardır. Inception filminden de hatırlanacağı üzere rüya içinde rüya gibi katmanlar arasında dibe inip yukarı çıkma işlemidir. Inception mimarisinde, paralel olarak birbirine bağlı katmanlar kullanılmıştır. Bu da parametre sayısını oldukça düşürmüştür (yaklaşık 4 milyon) [80]. Parametre sayısındaki düşüğe bağlı olarak, bellek kullanımı ve performans arttırmada avantaj sağlamıştır. Modelin yapısı şekil 4.7'de gösterilmiştir [83].

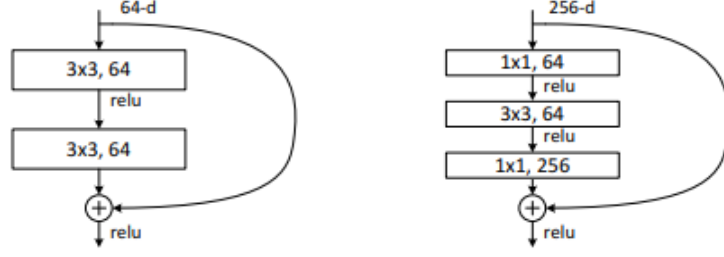


Şekil 4.7. Inception mimarisine sahip iç içe ağların gösterimi

4.7.3. ResNet50

Microsoft ekibinden Kaiming He ve arkadaşları tarafından geliştirilen artık sinir ağ mimarisi 2015 yılındaki ImageNet yarışmasında hata oranını %3.57 ye düşürerek başarılı olmuştur. 152 katmandan oluşan yapı yaklaşık 23 milyon parametreye sahiptir. ResNet50 görüntü tanıma ve sınıflandırmalarında sıkça kullanılan görüntü bir mimaridir. ResNet50 mimarisi de diğer öğrenme ağı mimarilerinde olduğu gibi derin ağ öğrenmelerindeki tıkanıklığı iyileştirme amacı gütmektedir. Bu bilgisayar mimarisi, yenilenen sinir ağlarına

benzeyen yapısı ile, girişin “*artık*” bir değer olarak sonraki iki evrişimli katmanın çıkışına beslenmesi prensibi ile çalışmaktadır. Şekil. 4.8’te ResNET mimarisi için hizmet eden “*artık*” fonksiyon yapı taşı ve ResNET50 mimarisinde darboğaz oluşumu yer almaktadır [88],[89]. Artık değer fonksiyonunun kullanılması ile ImageNET yarışmasında başarısını kanıtlamıştır.



Şekil 4.8. Bir ResNet mimarisi için daha derin bir *artık* fonksiyon yapı taşı ve ResNet-50 için bir “darboğaz” yapı taşının gösterimi

5. TİD VERİ SETİ OLUŞTURULMASI

5.1. Kullanılan Yöntem ve Materyal

Gerçekleştirilen tez çalışması üç temel bölümden oluşmaktadır. Birinci kısımda bilgisayara bağlı webcam ile veri setinin oluşturulması ikinci olarak sistemin işaret dilindeki durağan kelimeleri tanıyabilmesi için farklı modellerin geliştirilip eğitilmesi sınıflandırması ve başarısının değerlendirilmesidir. Son aşamada bunu bir program yazılarak gösterilen el işaretinin ekrana gerçek zamanlı metin olarak yazdırılması amaçlanmaktadır. Bu kapsamda kullanılan materyal bilgileri şu şekildedir:

Oluşturulan sistem, Intel® Core™ İ7 -8750h CPU @ 2.20 Ghz ve 16,00 GB RAM 'e sahip Nvidia GTX 1050 ti Ekran Kartı ve Microsoft Windows 10 Home x64 bit işletim sistemli bilgisayar üzerinde geliştirilmiştir. Diğer programlama dillerine nazaran program kodunu en az çaba ile ve hızlıca yazmaya imkân tanınması nedeni ile Python dili seçilmiş ve JetBrains PyCharm Community Edition 2019 1.1 ortamında TensorFlow arkauç motorunu kullanarak çalışan keras kütüphanesi kullanılarak yazılım gerçekleştirilmiştir.

Bu çalışma bir kişisel dizüstü bilgisayarında yazılan kodlar ve dizüstüne tümleşik bağlı olan bir web kamerasından oluşmaktadır. Kameranın özellikleri tablo 5.1.'de yer verilmiştir. Gerçekleştirilen çalışmada kamera karşısında duran kişinin yaptığı her bir hareketin Türk İşaret Dili'nde karşılığını doğru olarak tanıyıp; metin haline dönüştürmek için bir Python programı yazılmıştır.

Tablo 5.1. WebCam' in genel özellikleri

Kamera Model	Front- Facing 720p kamera
Ön kamera çözünürlüğü	1 MP
Ön kamera sinyal formatı	720p
Ön kamera HD türü	HD

5.2. Veri Setinin Oluşturulması

TİD durağan kelimelerine ait hareketleri içeren veri seti üç farklı bireyden alınan görüntülerle oluşturulmuştur. Veri setinde bulunan görüntüler, bilgisayara bağlı web kamerası yardımıyla iki farklı zamanda (gece-gündüz) elde edilmiştir. Çekim esansında arka fonda lacivert perde kullanılmıştır.

Birinci aşamada üç farklı bireyden alınan her bir sınıf için 100 adet görüntü kullanılarak toplam 32 sınıftan oluşan 3200 resimden oluşan bir veri seti oluşturulmuştur. Bunlar 10 tane rakam ve 22 tane sabit kelimededen oluşmaktadır. Sabit kelimededen kastedilen hareket içermeden yapılan işaret türüdür. Bu kelimeler: aferin, başarılı, üçgen, dikdörtgen, kare, artı, bölme, başkent, izci, ev, zor, baklava, kurt, köprü, mola, simit, kale, ada, mantar, çapraz, lazım ve dur. 32 işaretin görüntüsü aşağıdaki Şekil 5.1. de gösterilmiştir.



Şekil 5.1. Veri kümesinde yer alan TİD' e ait 32 sınıf için renkli görüntüler

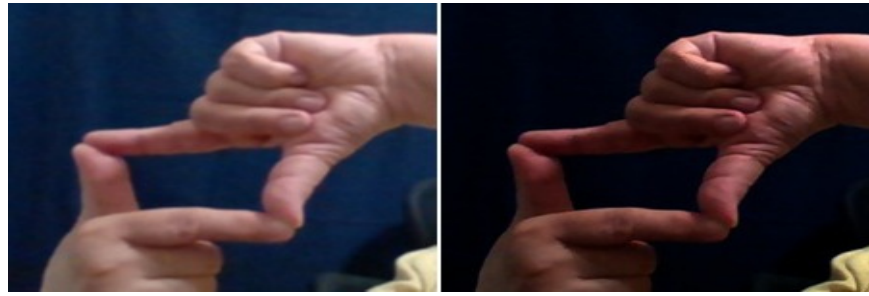
5.3. Veri Setinin Çeşitlendirme -1

Aşırı uydurmayı önlemek için yapılan veri çeşitlendirme yönteminde amaç eldeki verilerden birtakım dönüşümler ile gerçeğe yakın daha fazla eğitim verisi oluşturmak böylece eğitim esnasında modelimiz aynı resmi ikinci kez görmemesini sağlamaktır. Bu da modelimizin daha çok örüntü aramasını ve daha iyi genelleştirmesini sağlar.

Çalışmada öncelikle her sınıf görüntü için 100 adet fotoğrafa gürültü(noise), karışıklık(contrast) ve bulanıklık(blur) işlemleri yapılmış ve her sınıf için 100 adet olan görüntü sayısı 300'e çıkarılmıştır. Çalışma sırasında bulanıklık (blur) işlemi için, dönme eksenini kullanılmıştır. Şekil 5.2' de beş işarete gürültü eklenmiş ve Şekil 5.3' de dikdörtgen işarete karışıklık işlemi Şekil 5.4'te ise başarılı işarete bulanıklık işlemi yapılmıştır.



Şekil 5.2. Beş işaretinin gerçek ve gürültü eklenmiş görüntüleri



Şekil 5.3. Dikdörtgen işaretinin gerçek ve karışıklık eklenmiş görüntüleri



Şekil 5.4. Başarılı İşaretine gerçek ve bulanıklık eklenmiş görüntüleri

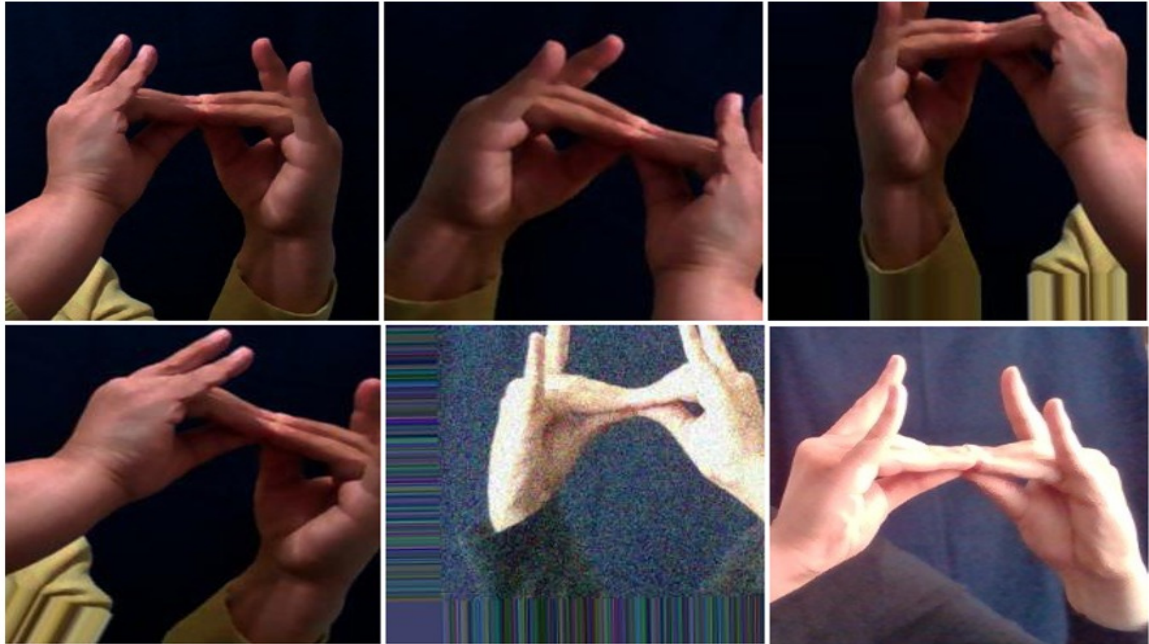
5.4. Veri Setinin Çeşitlendirme -2

İkinci aşamada sistemin örüntüyü farklı açılardan ve konumlardan çıkarabilmesi için keras kütüphanesinin sağladığı ImageDataGenerator sınıfını kullanarak rastgele dönüştürme parametreleri ile veriyi çeşitlendirdik. Her sınıf için 300 olan görüntü sayısı 600 çıkarılmıştır. Uygulanan işlemler tablo 5.2' de yer verilmiştir.

Tablo 5.2. Veri çeşitlendirme işlemleri

İşlem Adı	Değer	Açıklama
Görüntü Döndürme Aralığı	15	Görüntüyü en fazla 15 olacak şekilde sağa sola döndürme işlemidir.
Görüntüyü uzay düzlemde X ekseninde Kaydırma Aralığı	0.1	Yatay eksende kaydırma
Görüntüyü uzay düzlemde Y ekseninde kaydırma aralığı	0.1	Dikey eksende kaydırma
Kesişme Noktası Aralığı	0.2	Burkma işlemi
Yakınlaştırma Aralığı	0.1	Yakınlaşma işlemi
Görüntünün Dikey Ekseninde Yatay Simetriğinin Alınması	True	Resmin dikey eksende simetriği

Yukarıdaki değerleri seçerken örüntünün tamamen bozulmaması için uygun değer seçilmeye çalışılmıştır. Örneğin döndürmede işleminde 15 dereceden fazlası örüntüyü farklı bir işaret anlamına getirebilir ya da yatay eksende simetrik alınması işlemi yapılıyorsa farklı bir hareket ortaya çıkarırdı. Şekil 5.5 'da köprü işareti üzerinde yapılmış rastlantısal işlemler gösterilmiştir.



Şekil 5.5. Köprü işareti için çeşitlendirilmiş bazı görüntüler

5.5. Veri Setinin Genel Özellikleri

Veri çeşitlendirme işlemlerinden sonra her bir sınıf için 600 örüntü, toplamda da her biri 32 sınıftan oluşan veri setimizde $32 \times 600 = 19200$ örüntü bulunmaktadır. Veri setinin 5/6 eğitim 1/12' si doğrulama ve 1/12' si de test için ayrılmıştır. Veri setinin genel özellikleri Tablo 5.3' de belirtilmektedir.

Tablo 5.3. Veri setinin genel özellikleri

Toplam Sınıf Sayısı	32
Her Sınıf İçin Kaydedilen Görüntü Sayısı	600
Veri Setindeki Toplam Görüntü Sayısı	19200
Değişkenler	Gürültü, Işık, Yön, Uzaklık, Karşıtlık
Sınıfların Ortalama Dosya Büyüklüğü	16mb-22mb
Veri Seti Dosya Büyüklüğü	290 MB
Toplam Veri Seti Oluşturma Süresi	13-15 saat

Son olarak resimler yeniden boyutlandırılarak 50x50 formatta ayrı bir veri seti oluşturulmuştur. Tablo 5.4' te iki farklı görüntünün de özelliklerine yer verilmektedir.

Tablo 5.4. İki farklı veri setinin görüntü özellikleri

Boyutlar	50 x 50	224 x 224
Genişlik	50 piksel	224 piksel
Yükseklik	50 piksel	224 piksel
Yatay Çözünürlük	96dpi	96dpi
Dikey Çözünürlük	96dpi	96dpi
Bit Derinliği	24	24

6. ARAŞTIRMAYA İLİŞKİN BULGULAR

6.1. Araştırmaya İlişkin Bulgular

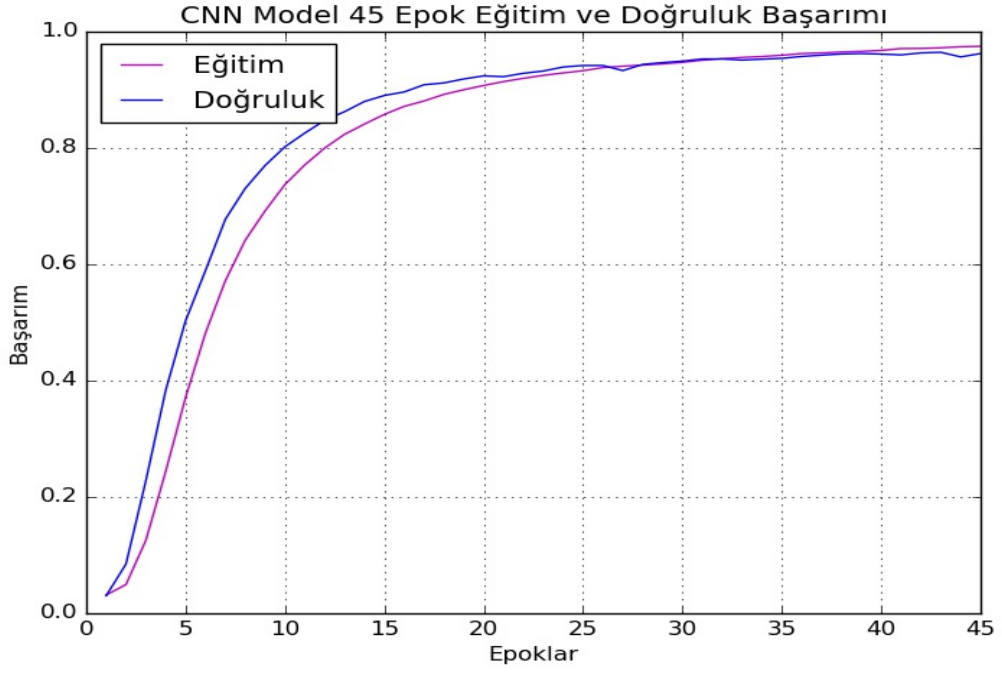
Modeller 16 Gb ram, Intel I7 işlemci dizüstü bilgisayarda NVIDIA GeForce GTX 1050-Ti ekran kartı üzerinde test edilmiştir. Donanım kısıtlarına bağlı olarak yığın miktarında 100 den daha fazla üzerine çıkılamadığından tüm derin öğrenme modellerinde yığın miktarı (batch size) 100 seçilmiştir. Uygulanan derin öğrenme yöntemlerinde çok fazla sayıda parametre bulunmaktadır ve bu nedenle tüm parametreler için eniyileme yönteminin uygulanması çok zaman almaktadır. Literatür incelendiğinde ve birkaç elle deneme sonucunda epok sayısının başarıyı bir hayli etkilediği gözlenmiştir. Bundan dolayı bu tez çalışması kapsamında diğer parametreler sabit tutulduğunda en uygun epok sayısının tüm yöntemlerde belirlenmesi için bir analiz yapılmıştır. En uygun epok sayısı literatürde sıklıkla kullanılan 15, 30 ve 45 sayıları arasında aranmıştır.

6.1.1. 10 Katmanlı ESA yapısını kullanarak başarımların değerlendirme

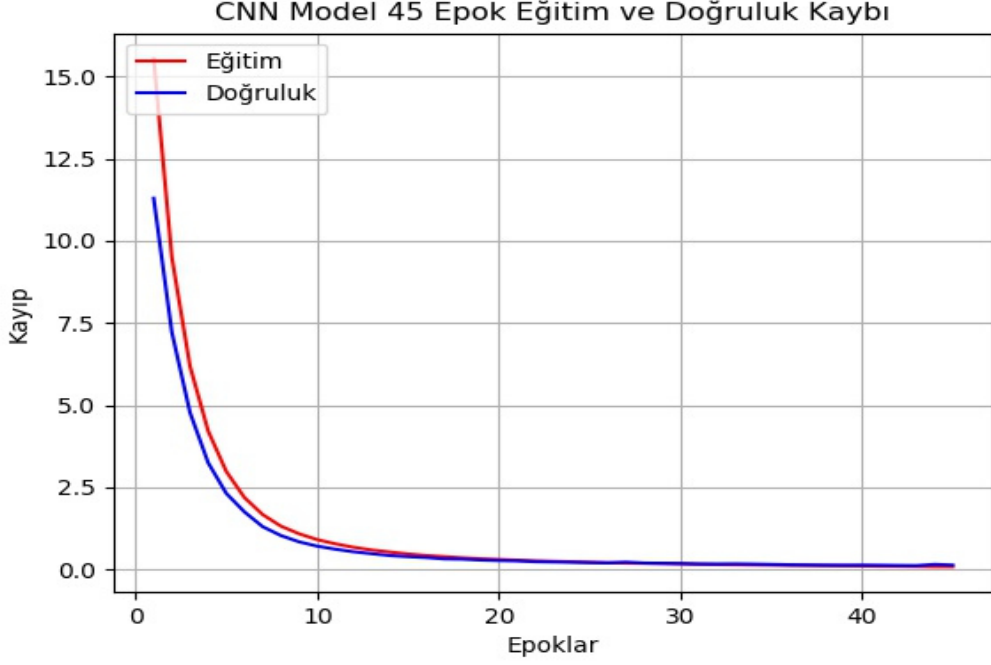
Bu tez çalışması kapsamında oluşturulan ESA mimarisinde epok sayısının değişimine göre kayıp ve başarımın değişimi incelenmiş ve uygun bir epok sayısı seçilmeye çalışılmıştır. Tablo 6.1. 'e göre en başarılı sonuç 45 epokta elde edilen %98' lik başarıdır. Epok sayısının 45 seçilmesi durumunda elde edilen kayıp ve başarımların grafiği Şekil 6.1 ve Şekil 6.2'de verilmektedir.

Tablo 6.1. 10 Katmanlı ESA (CNN) için farklı Epok sayılarında elde edilen sonuçlar

Epok	Resim Boyutu	Yığın Boyutu	Zaman	Kayıp ve Başarımlar
15	50x50	100	0:00:43	Kayıp:0.3111-Başarımlar: 91.7500%
30	50x50	100	0:01:08	Kayıp:0.0773-Başarımlar: 97.8125%
45	50x50	100	0:01:41	Kayıp:0.0430-Başarımlar: 98.6250%



Şekil 6.1. 10 katmanlı ESA'nın başarımları grafiği



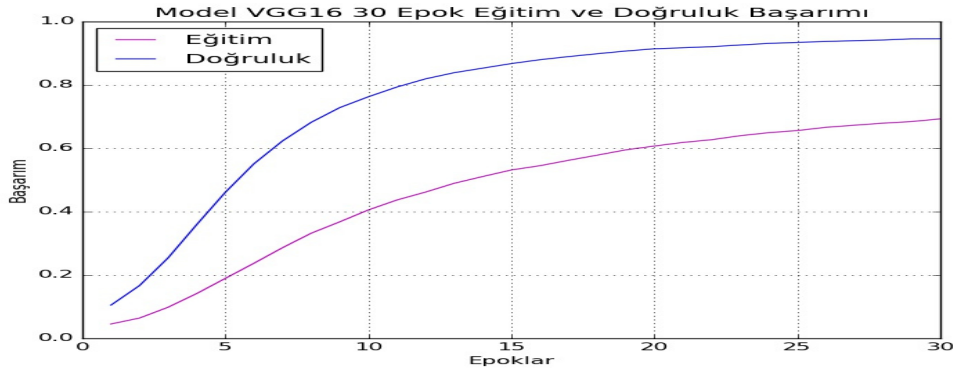
Şekil 6.2. 10 katmanlı ESA'nın kayıpları grafiği

6.1.2. Öğrenme aktarımı ile VGG16 kullanarak başarımların değerlendirilmesi

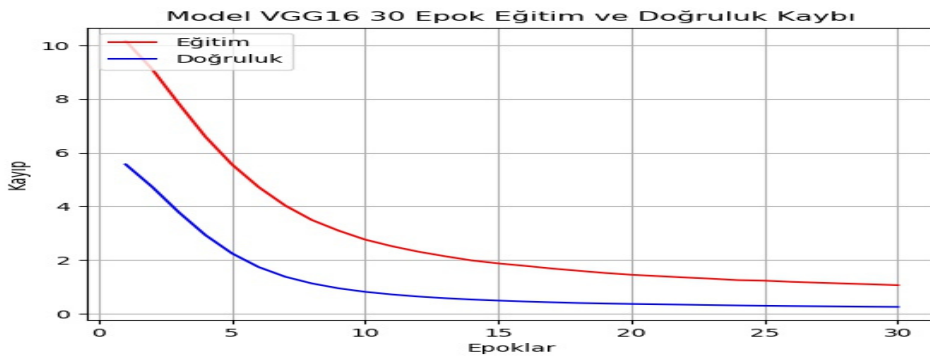
Öğrenme aktarımı ile katsayıları hazır alınmış VGG16 mimarisine sahip ESA’da en uygun epok sayısını bulmak için incelemeler yapılmıştır. Tablo 6.2’ye göre en yüksek başarımlar oranı 45 epokta elde edilen %96’lık oran olsa da 30 epokta elde edilen sonuca göre yaklaşık 45 dk. daha yavaş olduğundan ve iki sonuç arasından yaklaşık %2’lik az bir fark olduğundan en uygun epok sayısı 30 olarak kabul edilmiştir. Bu sonucun kayıp ve başarımlar grafiği Şekil 6.3 ve Şekil 6.4’de verilmiştir.

Tablo 6.2. VGG16’ nın farklı epok sayılarında elde edilen sonuçları

Epok	Resim Boyutu	Zaman	Kayıp ve Başarımlar
15	224x224	0:46:37	Kayıp:0.4652-Başarımlar: 87.7500%
30	224x224	1:32:22	Kayıp:0.2502-Başarımlar: 94.6250%
45	224x224	2:17:34	Kayıp:0.1871-Başarımlar: 96.8125%



Şekil 6.3. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen VGG16 mimarisinin başarımlar grafiği



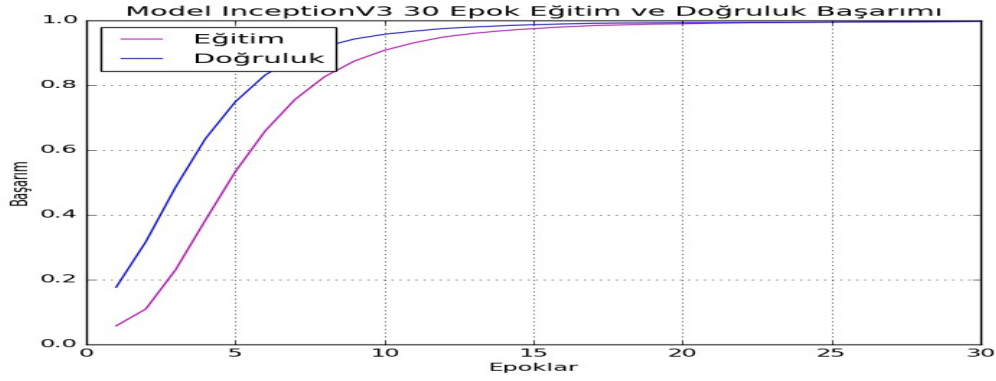
Şekil 6.4. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen VGG16 mimarisinin kayıp grafiği

6.1.3. Öğrenme aktarımı ile Inceptionv3 kullanarak başarımların değerlendirilmesi

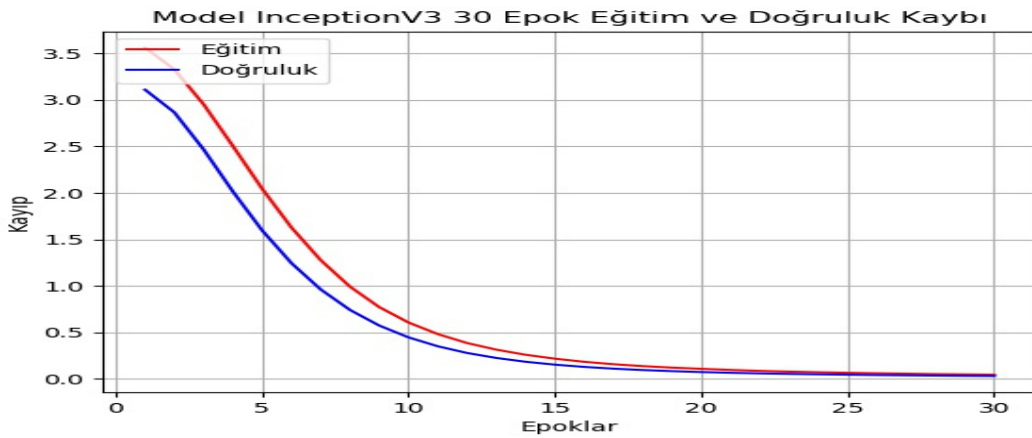
Öğrenme aktarımı ile ESA’da InceptionV3 mimarisi kullanıldığında en uygun epok sayısı için incelemeler yapılmıştır. Tablo 6.3’e göre başarımların oranları aynı olan 30 ve 45 epok sayılarına göre daha kısa süreye sahip olan 30 epokta elde edilen sonuç en başarılı kabul edilmiştir. Bu sonucun kayıp ve başarımların grafiği Şekil 6.5 ve Şekil 6.6’de verilmiştir.

Tablo 6.3. InceptionV3’ ün farklı epok sayılarında elde edilen sonuçları

Epok	Resim Size	Zaman	Kayıp ve Başarımların
15	224x224	0:34:23	Kayıp:0.1240-Başarımların: 98.8125%
30	224x224	1:07:40	Kayıp:0.0294-Başarımların: 99.7500%
45	224x224	1:37:58	Kayıp:0.0238-Başarımların: 99.7500%



Şekil 6.5. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen InceptionV3 mimarisinin başarımların grafiği



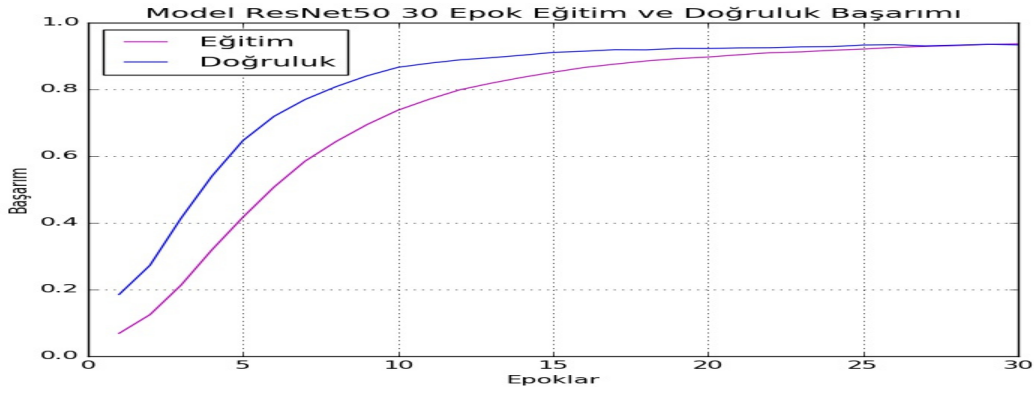
Şekil 6.6. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen InceptionV3 mimarisinin kayıp grafiği

6.1.4. Öğrenme aktarımı ile ResNet50 kullanarak başarımların değerlendirilmesi

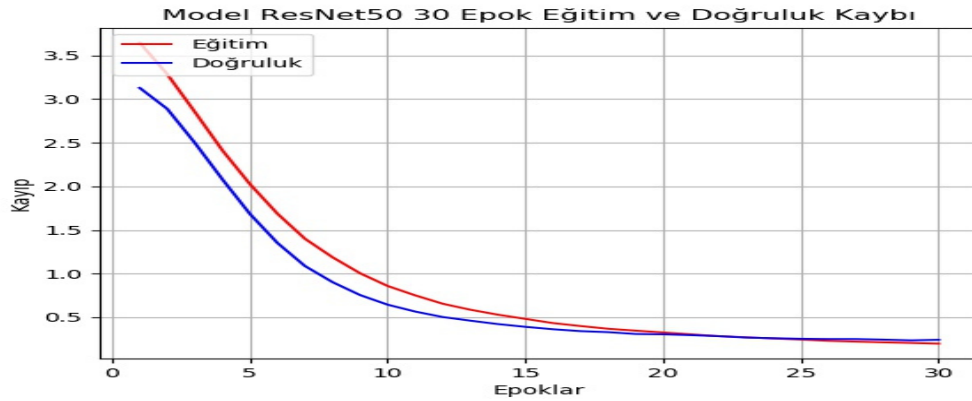
Tablo 6.4'e göre en yüksek başarımlar oranı 45 epokta elde edilen %95,5'lik oran olsa da 30 epokta elde edilen sonuca göre yaklaşık 42 dk. daha yavaş olduğundan ve iki sonuç arasından yaklaşık %2'lik az bir fark önemsenmeyip en uygun epok sayısı 30 olarak belirlenmiştir. Bu sonucun kayıp ve başarımlar grafikleri Şekil 6.7 ve 6.8 'de verilmektedir.

Tablo 6.4. ResNet50'ün farklı epok sayılarında elde edilen sonuçları

Epok	Resim Boyutu	Zaman	Kayıp ve Başarımlar
15	224x224	0:34:30	Kayıp:0.3658-Başarımlar: 90.8125%
30	224x224	1:24:16	Kayıp:0.2480-Başarımlar: 93.1875%
45	224x224	2:06:09	Kayıp:0.1725-Başarımlar:95.4375%



Şekil 6.7. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen ResNet50 mimarisinin başarımlar grafiği



Şekil 6.8. Öğrenme Aktarımı Yöntemi ile son katmanı eğitilen ResNet50 mimarisinin kayıp grafiği

6.1.5. Öğrenme aktarımı kullanılarak öznitelik çıkarımı ile birlikte destek vektör makineleri ve k- en yakın komşu sınıflandırması

Vapnik Destek Vektör Makinesi algoritmalarını 1963 yılında tanıtmıştır [58]. Destek vektör makineleri (DVM), denetimli öğrenme modellerinden biridir. DVM'lerde girdi olarak verilen örneklere göre çıktı olarak neyin beklendiği etiketlenir. DVM' nin amacı, bir veri setini iki sınıfa ayıran bir hiper düzlem veya bir karar yüzeyi bulmaktır [59]. Hiper düzlemi belirleyen veri noktalarına "destek vektörleri" denir [60],[61]. Veri noktalarını iki sınıfa ayıran birçok olası hiper düzlemin olması mümkündür. Yani, asıl amaç n-boyutlu bir alanda "azami verim" e sahip bir köprü bulmaktır. verim, iki sınıf arasındaki mesafe için hiper düzlemin boyutuna bağlıdır, bu da giriş özelliklerinin sayısı ile ilişkilidir. DVM, doğrusal olmayan modellerde iki sınıf arasındaki mesafeyi hesaplamak için "çekirdek" adı verilen bir mekanizma kullanır. Çekirdek seçimi, yerleşik DVM modelinin performansını etkiler. Doğru çekirdeğe karar vermek için en iyi yol, veri üzerindeki çekirdek işlevlerinin ayarlanarak denetlenmesini sağlamaktır. Doğrusal destek vektör makinesi, karar düzlemi sınıfları mükemmel bir şekilde +1 ve -1 düzleminde ayırır.

K-NN (*K-Nearest Neighbor*) algoritması en basit ve en sık kullanılan sınıflandırma algoritmasından biridir. K-NN eğitim aşaması olmayan bir yöntemdir. Yani tembel tabir edilen bir öğrenme algoritmasıdır. Eğitim verilerini öğrenmez, bunun yerine eğitim veri kümesini "ezberler". Bir tahmin yapmak istediğimizde, tüm veri setinde en yakın komşuları arar. Dezavantajı uzaklık hesaplanırken bütün uzaklıklar saklandığı için çok büyük veride yüksek bellek alanına ihtiyaç duyar. Algoritmanın çalışmasında bir *K* değeri belirlenir. Bu değer verilen noktaya en yakın komşu sayısıdır. Uzaklık fonksiyonları ile yeni noktanın mevcut noktalara göre tek tek uzaklığı hesaplanır en yakın k-komşuların özniteliklerine göre komşuların sınıfına yerleştirilir. Uzaklık hesaplama işleminde genelde Öklid fonksiyonu kullanılır. Ancak Öklid fonksiyonuna alternatif olarak Hamming, Manhattan ve Minkowski fonksiyonları da vardır.

Tablo 6.5. Öznitelik çıkarımı yapılarak elde edilen DVM ve K en yakın komşu sonuçları

Model	Resim Boyutu	Zaman	Kayıp ve Başarım
InceptionV3-DVM	224x224	0:01:07	99.5%
InceptionV3-K En Yakın Komşu (K-NN)	224x224	0:00:34	97.5%

Derin öğrenme modellerinde en başarılı modelin en düşük kayıp ve en yüksek başarı oranına ait Inception mimarisine sahip model olduğu önceki sonuçlarda görülmüş olup bu nedenle bu model ile öznelik çıkarımı yapılmıştır. Yaklaşık 15 dakikalık bir öznelik çıkarımı yapıldıktan sonra değerler diske kayıt edilerek. SVM ile K-NN makine öğrenme algoritmalarıyla sınıflandırma yapılmıştır ve başarı oranı SVM’ de 99.5, K-NN de ise 97.5 elde edilmiştir. Buradan çıkarılacak sonuçlardan biri, InceptionV3 mimarisinin, uygulanan resimlerden ayırt edici öznelikleri iyi tespit ettiğiidir. Bu sayede klasik makine öğrenme yöntemlerinden de başarılı sonuçlar alınmıştır.

6.1.6. Öznelik çıkarımı olmadan dvm ve k- en yakın komşu ile sınıflandırma

Bu tez kapsamında derin öğrenme yöntemlerinin başarıyı ne kadar arttırdığını anlamak için klasik makine öğrenmesi yöntemleri de kullanılmıştır. Öznelik çıkarımı olmadan ham veri ile yapılan K-NN ve SVM sınıflandırmada sınıflandırma doğruluk oranlarının sırasıyla tablo 6.6 da gösterildiği gibi %94.6 ve %96.5 elde edilmiştir. Ancak bu yöntemlere yüksek boyutlu veri uygulanması durumunda donanım yetersizliği yaşandığı için veri boyutu 224x224 yerine 50x50 seçilmiştir.

Tablo 6.6. DVM ve K en yakın komşu sınıflandırma sonuçları

Model	Resim Boyut	Zaman	Kayıp ve Başarım
DVM	50x50	0:12:49	96.5%
K En Yakın Komşu (K-NN)	50x50	0:01:47	94.6%

6.2. Tartışma

Tablo 6.7’de tüm uygulanan yöntemler hızları, kayıp ve başarı oranları verilmektedir. Genel olarak tüm sonuçlar değerlendirildiğinde 50x50 boyuttaki verilerde klasik makine öğrenme yöntemlerine göre kendi oluşturduğumuz CNN modeli ile zaman ve başarı açısından daha iyi sonuçlar elde edilmiştir. Veri sayısı arttıkça ve gürültü arttıkça derin öğrenmenin SVM ve k-NN göre daha da başarılı olacağı öngörülmektedir. Ayrıca her problem için ona özgü derin ağ modelini oluşturmak daha hızlı ve başarılı sonuçlara ulaşılacağı bu tez kapsamında anlaşılmaktadır.

30 epok eğitim sonunda öğrenme aktarımı ile hassas ayar yapılan modeller içerisinde hem başarı oranı hem de zaman olarak InceptionV3 mimariye sahip ESA modelinin daha başarılı olduğu tespit edilmiştir. Çok küçük boyutta ve çok miktarda evrişim katmanı kullanıldığı için bu ağ, parametre sayısını büyük ölçüde azaltmayı başarmıştır. Alexnet 60 milyon parametre sayısı varken burada 4 milyona düşürülmüştür, Resnet50 23milyon ve VGG16 138 milyon parametreye sahiptir. Bu değerler göz önüne alındığında InceptionV3'ün diğer modellere göre daha hızlı olması normal kabul edilmiştir.

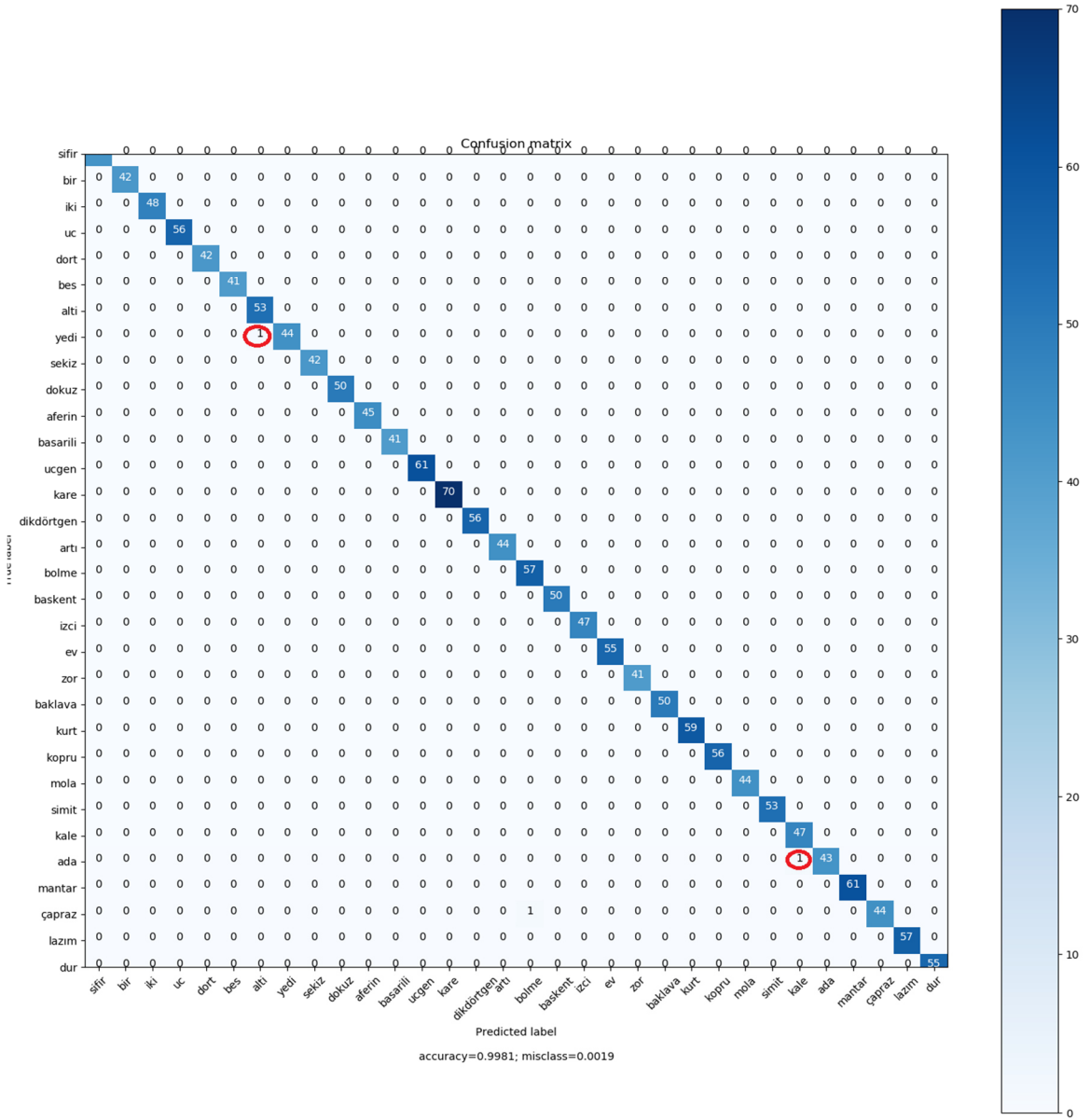
Bu tez kapsamında derin öğrenme yöntemleri ile klasik makine öğrenme yöntemleri hibrit şekilde birleştirilerek bu problem kapsamında analiz edilmiştir. Yine yüksek başarı gösteren InceptionV3 modeli öznelik çıkarmada kullanılmış ve klasik makine öğrenmesi yöntemlerinden seçilen SVM ve k-NN yöntemleri ile sınıflandırılması ile de yüksek başarı elde edilmiştir. Bunun nedeni olarak InceptionV3 mimarisinin bu problemin ayırt edici özneliklerini çıkarmada etkili olduğu görülmektedir. InceptionV3 mimarisinin doğrudan kullanmak yerine sınıflandırmayı klasik makine öğrenmesi yöntemleri ile yapmak işlemin hızını arttıracığı tespit edilmiştir.

Derin öğrenme modellerinde InceptionV3 ve 10 katmanlı bu problem kapsamında oluşturulan ESA modelinin zaman ve başarımlar açısından en başarılı gelen modellerdir. InceptionV3 modeli Imagenet yarışmasından daha fazla veri ile eğitilmiş olduğundan farklı arka planlarda başarılı olabileceği öngörülerek 30 epokta eğitilen InceptionV3 modelinin, gerçek zamanlı işaret dilini tanıma sisteminde kullanılmasına karar verilmiştir.

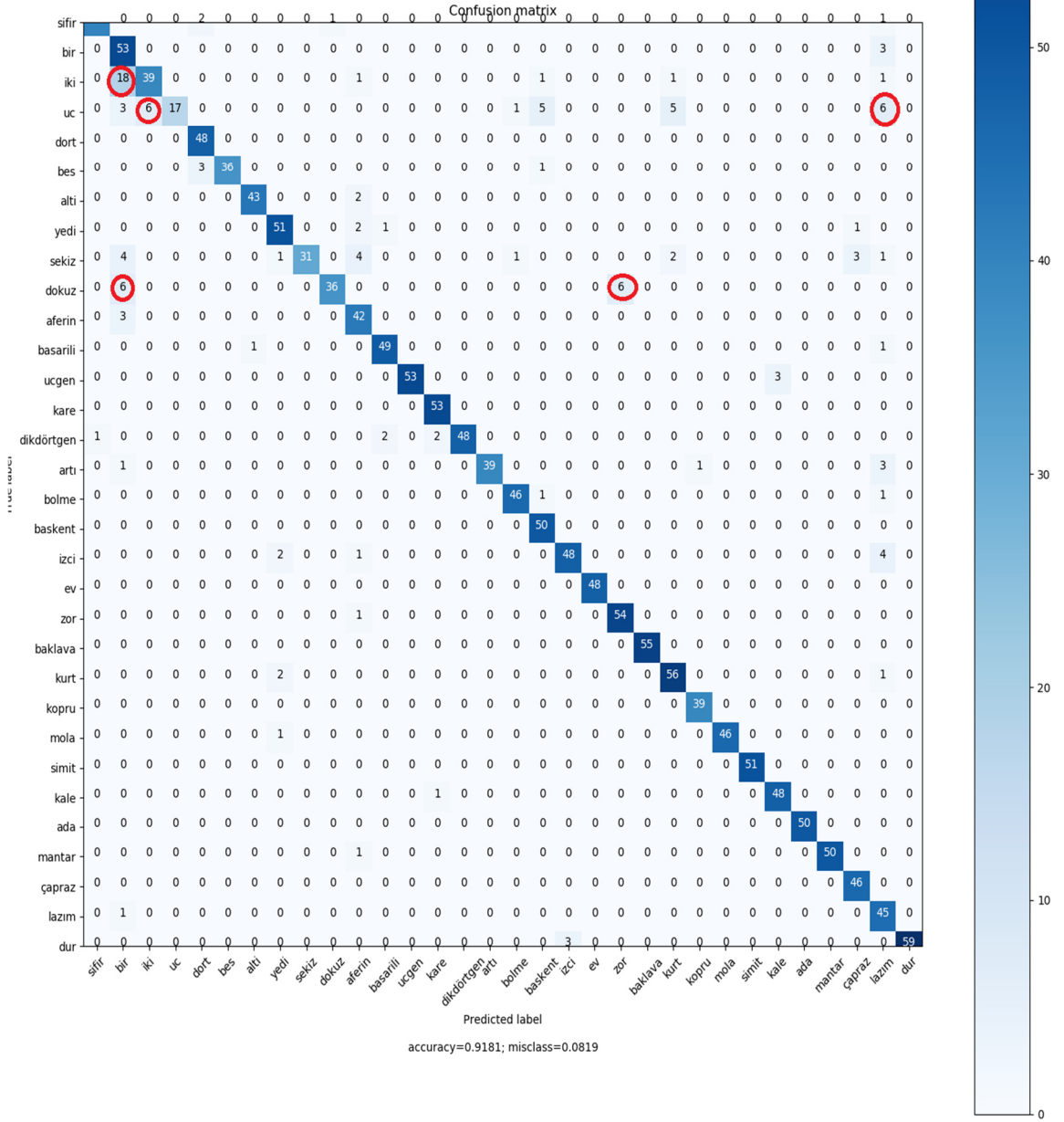
Tablo 6.7. Modellerin zaman ve başarımlar oranları karşılaştıran tablo

Model	Resim Boyutu	Epok	Zaman	Başarımlar Oranı
DVM	50x50	-	0:12:49	96.5%
10-Katmanlı ESA(CNN)	50x50	45	0:01:08	98.6250%
K -NN	50x50	-	0:01:47	94.6%
InceptionV3 – K-NN	224x224	-	0:00:34	97.5%
InceptionV3 – DVM	224x224	-	0:01:07	99.5%
VGG16	224x224	30	1:32:22	94.6250%
InceptionV3	224x224	30	1:07:40	99.7500%
Resnet50	224x224	30	1:24:16	93.1875%

En yüksek ve en düşük başarımlarına sahip Inception ve ResNet mimarilerinin, tahmin edilen değerler ile gerçek değerleri gösteren karışıklık matrisleri Şekil 6.9 ve 6.10’ da gösterilmiştir. Modelin hatalarını görmek ve incelemek için sıklıkla kullanılır. Inception yüzde yüze yakın bir oranla sadece yedi rakamını altı ve ada kelimesini kale olarak yanlış tahmin etmiştir. Resnet ise en çok on sekiz kez doğrusu iki iken bir olarak tahmin etmiştir. Ayrıca altı kez üç iken iki ve lazım, altı kez dokuz iken zor ve bir olarak yanlış tahmin etmiştir. Karıştırılan kelimeler birbirine yakın işaretlere sahiptir.



Şekil 6.9. InceptionV3 karışıklık matrisi

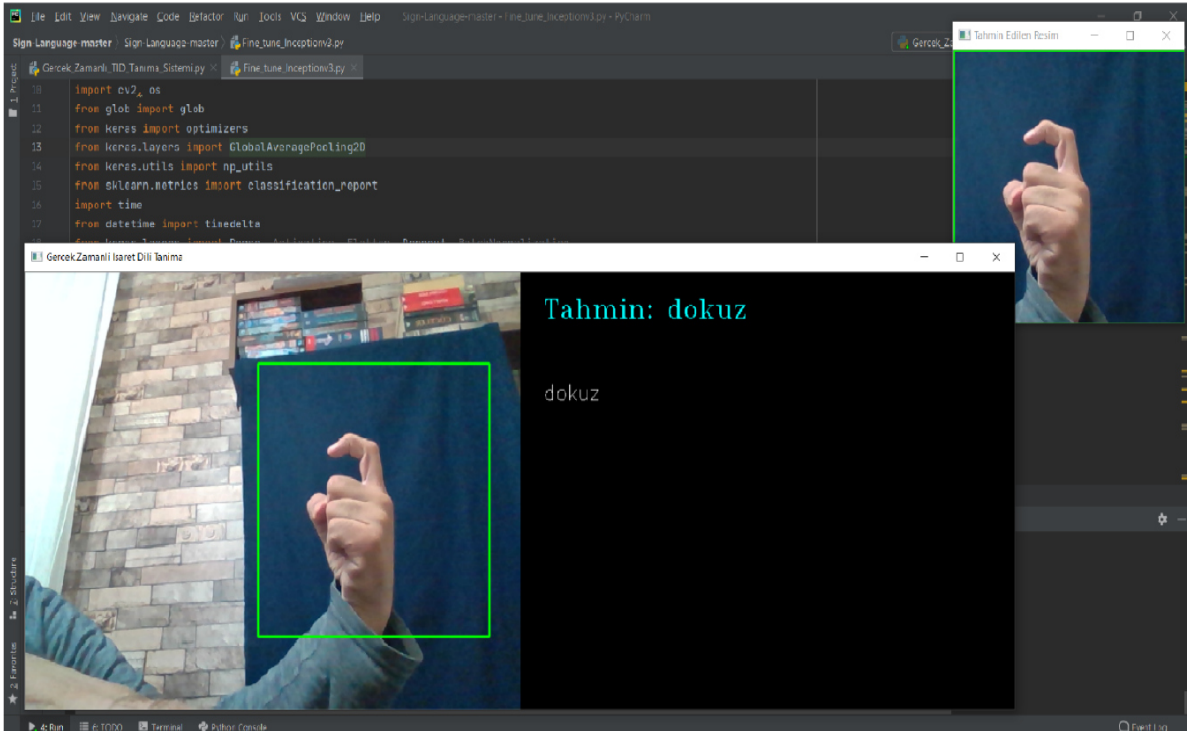


Şekil 6.10. ResNet50 karışıklık matrisi

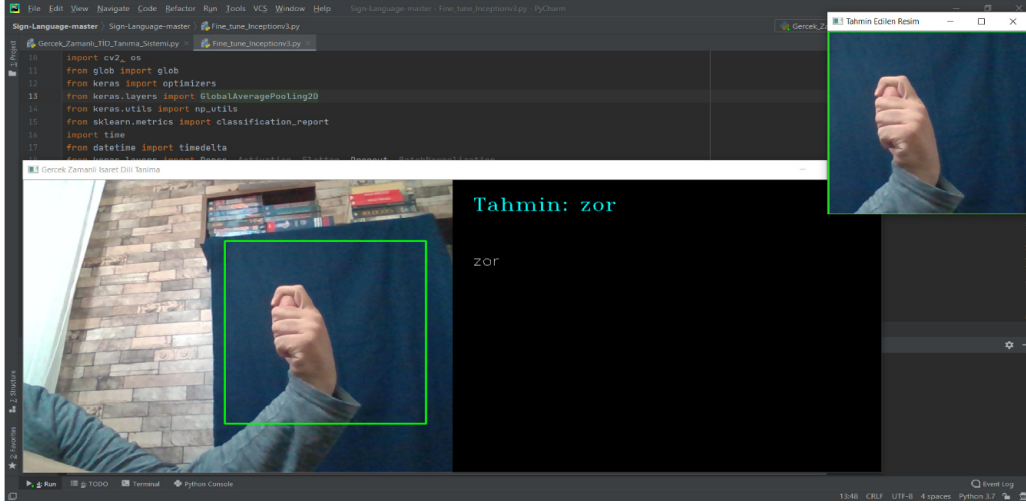
7. TÜRK İŞARET DİLİ DURAĞAN KELİMELERİN GERÇEK ZAMANLI TEST EDİLMESİ

7.1. Çalışma Prensipleri

Şekil 7.1 ve Şekil 7.2’de gösterildiği gibi sistem daha önce eğitilmiş olan Inceptionv3 modelinden beslenerek tahmin yapmaktadır. Kamera önünde yeşil çerçevenin içerisinde yapılan örüntü daha önceden eğitilmiş model tarafından tahmin edilir ve güvenilirlik sınırı %70 ve üstü olarak belirtilmiştir. Tahmin kısmına tahmin edilen kelime yazılır. Eğer bu tahmin 20 çerçeve(frame) boyunca aynı ise tahmin edilen işareti aşağıdaki metin alanına yazar. Burada karşılaşılan engel TİD yer alan durağan kelimelerin birbirine benzer hareketler ile tanımlanmış olmasıdır. TİD içerisinde yer alan durağan kelimelerden dokuz ile zor, üçgen ile kale, aferin ile başarı kelimeleri yaklaşık olarak birbirine benzerdir. Bu benzer kelimelerin yanı sıra durağan kelime ile parmak tabanlı işaret diline ait harfler benzerlik göstermektedir. Bu benzerlik örneğin “L” sesi ile “Lazım” kelimesi aynı işarete sahiptir. Çalışmamız kapsamında durağan kelimeleri tercih ettiğimizden “L” sesi çalışma kapsamına alınmamıştır.



Şekil 7.1. TİD ‘dokuz’ işaretinin bilgisayar ekran görüntüsü



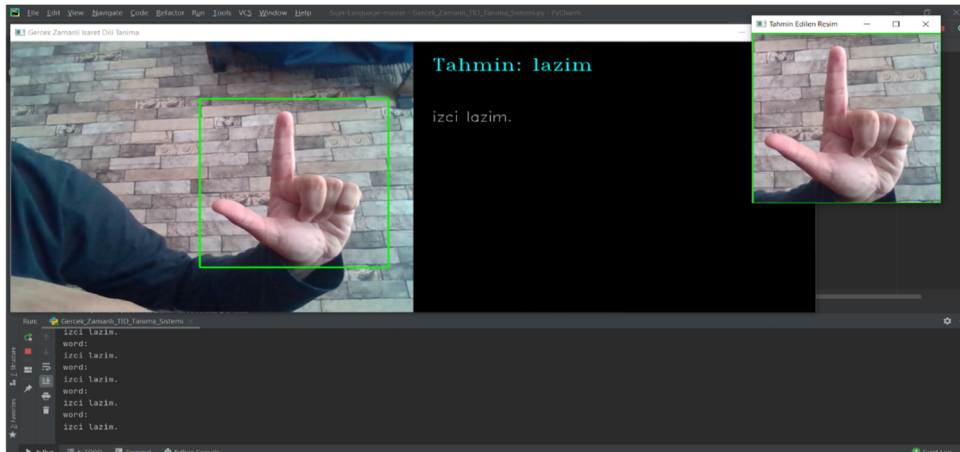
Şekil 7.2. TİD ‘zor’ işareti bilgisayar tahminini gösteren ekran görüntüsü

7.2. Gerçek zamanlı programın özellikleri

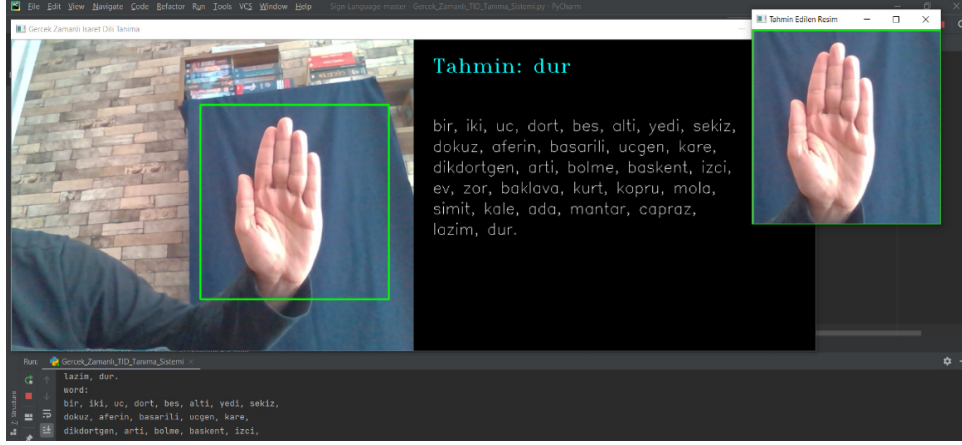
-Her kelime tahmininden sonra sitem otomatik olarak boşluk koyar, klavyeden virgül ve noktalama işaretleri koyula bilir.

- Klavyeden R tuşuna basıldığında bütün yazılanlar silinir.
- Backspace tuşu ile son yerleştirilen kelime silinir.
- Klavyeden C tuşuna basıldığında sistem kapanır.

Programın özellikleri kullanılarak şekil 7.3’te bir basit cümle kurulmaya çalışılmıştır ve şekil 7.4’de programın özelliklerinden faydalanarak kullanılan tüm işaretler ekranın metin kısmına yazılmıştır.



Şekil 7.3. TİD basit yapıda bir cümlenin yazıya çevrilmesini gösteren bilgisayar ekran görüntüsü



Şekil 7.4. Dur işareti ile tüm kelimelerin yazımını gösteren bilgisayar ekran görüntüsü

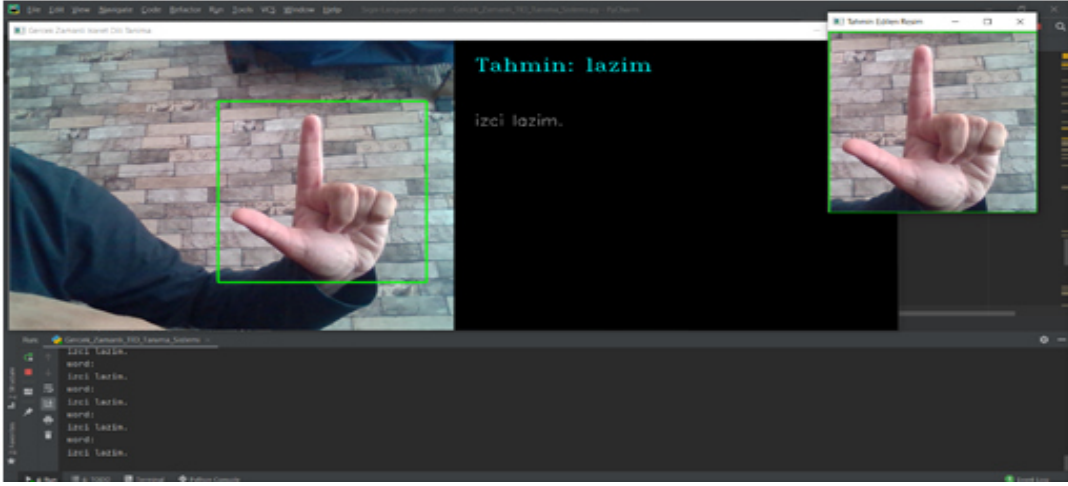
7.3. Sistemin Farklı Kullanıcılar Üzerinde Test Edilmesi

Eğitim veri setinde olmayan dört farklı kullanıcı üzerinde beş farklı kelime için program test edilmiştir. Kullanıcılar Türk İşaret Dili'ni bilmeyen ve sistemi daha önce kullanmamış kişilerdir. Buna ilaveten, eğitim setinde kullanılmayan arka planlarda program denenmiştir. Kelimeler birbirine yakın işareteler olan kale ve üçgen sonrasında zor ve dokuz en son olarak aferin kelimesi ile test edilmiştir. Sonuçlar Tablo 7.1' de belirtilmiştir.

Tablo 7.1. Sistemin farklı ortamlar ve kişiler üzerinde test sonuçlarını gösterir

Kişiler / Ortam	Dokuz	Zor	Üçgen	Kale	Aferin
1.Kişi / Ortam 1	X	✓	✓	✓	✓
1.Kişi / Ortam 2	X	✓	✓	✓	✓
1.Kişi / Ortam3	✓	X	✓	✓	✓
2.Kişi / Ortam 1	✓	X	✓	✓	✓
2.Kişi / Ortam 2	✓	✓	✓	✓	✓
2.Kişi / Ortam 3	X	✓	✓	✓	✓
3.Kişi / Ortam 1	✓	✓	✓	✓	✓
3.Kişi / Ortam 2	✓	X	✓	✓	✓
3.Kişi / Ortam 3	✓	✓	✓	✓	✓
4.Kişi / Ortam 1	✓	✓	✓	✓	✓
4.Kişi / Ortam 2	✓	✓	✓	✓	✓
4.Kişi / Ortam 3	✓	✓	✓	✓	✓

Buradan elde edilen bulgu ise, sistemin dokuz ve zor kelimelerini ayırt etmekte zorlandığı tespit edilmiştir. 60 denemde 54 doğru sonuçla %90 oranında sistem başarısı elde edilmiştir. Farklı arka planlarda ve kişilerle yapılmış testlere ait ekran görüntülerine ise Şekil 7.5, Şekil 7.6 ve Şekil 7.7'de yer verilmiştir.



Şekil 7.5. Taşduvar arka planında لازم kelimesinin metin alanına yazılan ekran görüntüsü



Şekil 7.6. Doku desenli duvar arka planında üçgen kelimesinin metin alanına yazılan ekran görüntüsü



Şekil 7.7. Dış çekim doğa görüntüsü arka planında dokuz kelimesinin metin alanına yazılan ekran görüntüsü

8. SONUÇ VE ÖNERİLER

8.1. Sonuç

Bu tez çalışmasında temel amaç TİD' e ait durağan hareketlerin başarılı bir şekilde tanınmasını sağlamak için derin öğrenme yöntemleri uygulanarak Türk işaret dilindeki bazı kelimelerin gerçek zamanlı tanınmasını gerçekleştirmektir. Bunu gerçekleştirmek için tanınacak kelimeleri sınıflandırmada derin öğrenme yöntemleri irdelendi. Tasarlanan sistem görüntü tanıma yaptığı için de bu alanda en başarılı sonuçları veren Evrişimsel Sinir Ağları incelendi. Bu kapsamda ilk olarak bu probleme özgü tasarlanan Evrişimsel Sinir Ağıyla %98 başarı elde edilerek klasik makine öğrenme yöntemlerinden daha başarılı sonuç elde edildiği tespit edildi. Öğrenme aktarımı yoluyla farklı mimariye sahip Evrişimsel Sinir Ağları karşılaştırıldı. Bunların içinden InceptionV3 mimarisine sahip yapının %99 ile en başarılı sonucu verdiği tespit edildi. Bunların yanısıra yine derin ağlar ile öznitelik çıkartılarak Destek Vektör Makineleri ile sınıflandırdığında %99 başarı elde edilmiştir. Tüm yöntemler zaman ve sınıflandırma başarı oranı açısından kıyasandığında, bu problem için en iyi yöntemin hibrit yöntem olan derin ağlarla öznitelik çıkardıktan sonra sınıflandırmak için klasik makine öğrenmesini kullanan yöntem olduğu tespit edilmiştir.

Ülkemizde TİD çeviri sistemlerine ait çalışmaların az olması, ekstra bir cihaz kullanılmadığından maliyetsiz olması, renkli eldiven gibi araçlar kullanılmadan tanıma işlemi gerçekleştirerek kullanıcı dostu olması, veri çeşitlendirme yöntemlerinden contrast, noise gibi işlemler yapılarak ve gray scale yerine rgb görüntü kullanılarak ortamın aydınlatma şartlarından bağımsız çalışabilmesini sağlamış olması, eğitim verilerinin doğrudan herhangi bir ön işlemde geçmeden yada bir öznitelik çıkarımı yöntemi uygulanmadan sınıflandırması bu çalışmayı özgün kılan özelliklerdir.

Sonuç olarak, bu tez çalışması kapsamında elde edilen kendi özgün verilerimiz ile oluşturulan TİD'deki durağan kelimelerin tanınması ve gerçek zamanlı olarak yazılı metne dönüştürülmesinde başarılı sonuçlar alınmıştır.

8.2. Öneriler

1-Bu tez kapsamında yapılan webcamdan fotoğraf çeken kod sayesinde makine öğrenmesi modellerine uygun olacak şekilde yeni veriler eklenebilir. Kişi sayısı arttırılabilir. Farklı ortamlarda fotoğraf çekim gerçekleştirilebilir.

2-Mevcut çalışmada öğrenme transferi ile başka başarılı olmuş önceden eğitilmiş modellere hassas ayar yapılarak modellerin başarısı kıyaslanabilir.

3-Durağan kelimelerin daha da arttırılması ya da harf tabanlı işaretlerin de eklenmesi ile sınıf sayısı arttırılarak modellerin başarısı değerlendirilmeye devam edilebilir.

4-TİD kullanılan hareketli kelimeleri tanımak için video bazlı bir tanıma sistemi yapılabilir. Ayrıca gerçek zamanlı işaret dili tanıma sisteminin yazılan işareti sesli olarak söylemesi duyabilen ve duyamayan kişiler arasından kullanılabilirliği kolaylaştırabilir.

5-Bu tez çalışmasında tasarlanan sistem geliştirilerek kamuya açık alanlarda, okullarda ve eğitim kuruluşlarında kullanılabilir. Hatta hem harf hem de kelime tabanlı olarak çalışabilecek, yalnızca ellerin değil vücut hareketlerinin, yüz ifadesinin ve dudakların da incelenebileceği kapsamlı bir çalışma yapılarak akıllı telefonlarda çalışabilen bir uygulama ile günlük hayatta da kullanılması sağlanabilir. Böyle bir çalışma ile işitme sorunu olan kişilerin daha kolay iletişim kurmalarına ve topluma kazandırılmasına önemli katkı sağlayacaktır.

KAYNAKLAR

- [1] Ç. Batman, “İşitme Kayıpları,” <https://www.caglarbatman.com/isitme-kayıplari>, (Erişim Tarihi: 17 Aralık 2019)
- [2] Ş. H. Akalın ve S. Cavkaytar, İşaret Dilinin Kökleri, T.C. Anadolu Üniversitesi Yayın No:3140, Açıköğretim Fakültesi Yayınları, No: 2045 s: 11
- [3] Marmara İşitme Engelliler Derneği İnternet Sitesi, <http://www.marmaraied.org/detail/35/isaret-dili>,(Erişim Tarihi: 17 Aralık 2019)
- [4] Ş. H. Akalın ve S. Cavkaytar, İşaret Dilinin Kökleri, T.C. Anadolu Üniversitesi Yayın No:3140, Açıköğretim Fakültesi Yayınları, No: 2045, s: 4, 12, 17
- [5] M.F. Karaca ve S. Görgünoğlu, “Türkçeden Türk İşaret Dili’ne,” *İleri Teknoloji Bilimleri Dergisi*, ISSN:2147-3455, Sayı 5, C:1, 2015, s:22
- [6] B. Oktekin ve N. Çavuş, “İşitme ve Konuşma Engelli Bireyler için İşaret Tanıma Sistemi Geliştirme,” *Folklor-Edebiyat Dergisi*, Sayı: 97-1, C: 25, 2019, s: 598
- [7] C. Vogler and D. Metaxas, “ASL recognition based on a coupling between HMMs and 3D motion analysis,” International Conference on Computer Vision, 1998, s: 363-369 IEEE.
- [8] T. Starner, S. Mann, B. Rhodes, J. Healey, J. Levine and A. Pentland, “Cambridge Wearable Computing and Augmented Reality,” to appear 1997 Presence issue on Augmented Reality, MIT Press, Cambridge, MA; also Vision and Modeling Technical Report #355, MIT Media Laboratory, 1995.
- [9] J. Segen and S. Kumar, Shadow gestures: 3D hand pose estimation using a single camera. IEEE International Conference on Computer Vision and Pattern Recognition, 1999, (s: 1479- 1485) IEEE: Fort Collins, CO, USA.

- [10] Dong, Cao, “American Sign Language alphabet recognition using Microsoft Kinect Masters Theses 7392,” https://scholarsmine.mst.edu/masters_theses/7392 (Eriřim Tarihi: 20.12.2019)
- [11] J. A. Udoff (2015). Mouthings in American sign language: Biomechanical and representational foundations. Dissertation Abstracts International: Section B: The Sciences and Engineering. US: ProQuest Information & Learning.
- [12] Dong, Cao, https://scholarsmine.mst.edu/cgi/viewcontent.cgi?article=8391&context=masters_theses (Eriřim Tarihi: 20.12.2019)
- [13] C.H. Chuan, E. Regina and C. Guardiano, Amerikan, “Sign Language Recognition Using Leap Motion Sensor, 2014 13th,” International Conference On Machine Learning And Applications (ICMLA), IEEE December, 2014, s: 541-544
- [14] VR Technologies İn Cultural Heritage First İnternational Conference, VRTCH 2018, Brasov, Romaia, May.2018, Communications in Computer and İnformation Science, ISSN 1865-0929 ISBN 978-3-030-05818-0, e book ISBN 978-3-030-05819-7, <https://doi.org/10.1007/978-3-030-05819-7>, s:115-126
- [15] LeapMotion Sensor İncelemeleri <https://www.hackster.io/leapmotion/projects> (Eriřim Tarihi: 24.12.2019)
- [16] Cyblerglove Sensör İncelemeleri <http://www.cyberglovesystems.com/cyberglove-ii> (Eriřim Tarihi: 24.12.2019)
- [17] Microsoft Kinect Sensör İncelemeleri <https://www.instructables.com/id/Kinect-Projects/> (Eriřim Tarihi: 24.12.2019)
- [18] H. Sever ve A. Uçan, “Otomatik Duygu Sözlüğü Çevirimi ve Duygu Analizinde Kullanımı,” Hacettepe Üniversitesi, Yüksek Lisans Tezi, 2014
- [19] P. Cunningham and S.J. Delany, “K-Nearest Neighbour Classifiers”, Mult. Classif. Syst., 2007, s: 8

- [20] S.A. Özel, ve F. Güven, “Using Text Representation And Deep Learning Methods For Turkish Text Classification,” Çukurova Üniversitesi, Bilgisayar Mühendisliği Anabilim Dalı, Lisans Tezi, 2019
- [21] Ö.K. Baykan ve E. Ezel, “Derin Öğrenme Yöntemi Kullanılarak Görüntü-Tabanlı Türk İşaret Dili Tanıma,” Selçuk Üniversitesi, Yüksek Lisans Tezi, 2018
- [22] Y.H. Usta, M. Ghosheh ve Y. İşler, “El Protezlerinde Kullanılan Yüzey EMG Elektrotları ile Kuvvet Sensörlerinin Karşılaştırılması,” *Biyomedikal ve Klinik Mühendisliği Derneği, 29 Ek.2016 Tıp Teknolojileri Kongresi Dergisi*, 2016, s: 430-432
- [23] S.A. Arpacı ve O. Kalıpsız, “Yazılım Hata Sınıflandırmasında Farklı Naive Bayes Tekniklerin Kıyaslanması,” *Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, Sayı 1, C: 7, s: 1-13 ISSN: 2564-6605, 2018
- [24] S. Ameen & S.Vadera, A convolutional neural network to classify American Sign Language fingerspelling from depth and colour images. *Journal Expert Systems* 2017 Vol.34, Is. 3., pp. 97-121.
- [25] L. Quesada, G. Opez & L. A. Guerrero, Sign Language Recognition Using Leap Motion 2015 (pp. 277–288).
- [26] G. A. Rao & P. V. V. Kishore, Selfie video based continuous Indian sign language recognition system. *Ain Shams Engineering Journal* 2018a, 9(4), 1929–1939.
- [27] G. A. Rao, K. Syamala, P. V. V. Kishore & A. S. C. S. Sastry, (2018b). Deep convolutional neural networks for sign language recognition. In 2018 Conference on Signal Processing And Communication Engineering Systems, SPACES 2018 (Vol. 2018-January, pp. 194–197). Institute of Electrical and Electronics Engineers Inc.
- [28] R.B. Mapari and G.U. Kharat, Real Time Human Pose Recognition Using Leap Motion Sensor, *Computer Science Published in IEEE International Conference*, 2015, doi:10.1109 / icrcicn.2015.7434258

- [29] P. Backhaus, ALPHABET Ante Portas: How English Text Invades Japanese Public Space. *Visible Language*, 41(1), 2007, 70–87. Retrieved from
- [30] M. Assan & K. Grobel, Video-based sign language recognition using hidden markov models. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 1371, pp. 97–109). Springer Verlag.1998
- [31] M. Mohandes & M. Deriche, Image based Arabic Sign Language recognition. In *Proceedings-8th International Symposium on Signal Processing and its Applications, ISSPA 2005* (Vol. 1, pp. 86–89)
- [32] M. Uslu, <http://veribilimci.org/oklid-uzakligi/>, (Eriřim Tarihi: 29.12.2019)
- [33] G.ř. Ögüdücü and A. Madylova, “Kosinüs Benzerliđini Kullanarak Belgeler Arası Anlamsal Benzerliđi Kavramsal Sözlüđe Dayalı Hesaplama Yöntemi,” İTÜ, Yüksek Lisans Tezi, 2009
- [34] C. Öz, B. Eken, B. řahin, E. Akbulut & F. Akbulut, Saklı Markov Model Kullanarak Türkçe Konuşmayı ve Yazıyı İşaret Diline Çevirme, TMMOB Elektrik Mühendisleri Odası, Elektrik, Elektronik ve Bilgisayar Sempozyumu 2011, Fırat Üniversitesi, Elazığ.
- [35] S. Albayrak ve H. Haberdar, “Saklı Markov Model Kullanılarak Görüntüden Gerçek Zamanlı TİD Tanıma Sistemi,” YTÜ, Yüksek Lisans Tezi, 2005.
- [36] Y. Iwai, K. Watanabe, Y. Yagi and Yachida, “Information Intelligence and Systems (Cat. No.96CH35929),” *International Conference on Systems, Man and Cybernetics*, 2002, IEEE, ISBN: 0-7803-3280-6
- [37] H. Erdem ve Z.B. Kın, “Türk İşaret Dili Alfabetesinin Derin Öğrenme Yöntemi İle Sınıflandırılması,” Başkent Üniversitesi, Yüksek Lisans Tezi, 2019

- [38] N. Haneciođlu, “Türkçe İşaret Dili İçin Evrişimli Sinir Ağları ve Tekrarlayan Sinir Ağları İle Mimik Tanıma,” Ankara Yıldırım Beyazıt Üniversitesi, Yüksek Lisans Tezi, 2019
- [39] M. V. Beena & M. N. A. Namboodiri, Automatic Sign Language Finger Spelling Using Convolution Neural Network: Analysis. *International Journal of Pure and Applied Mathematics*, 117(20), 9–15, 2017
- [40] Y. Liao, P. Xiong, W. Min, & J. Lu, Dynamic Sign Language Recognition Based on Video Sequence with BLSTM-3D Residual Networks. *IEEE Access*, 7, 38044–38054, 2019.
- [41] P. Rathi, R. Kuwar Gupta, S. Agarwal, & A. Shukla, Sign Language Recognition Using ResNet50 Deep Neural Network Architecture. *SSRN Electronic Journal* 2020.
- [42] J. Kim, Bae, K.H. Park, W.C. Bang and Z.Z. Bien, “Continuous Gesture Recognition System for Korean Sign Language based on Fuzzy Logic and Hidden Markov Model,” *IEEE International Conference on Fuzzy Systems, 2004, FUZZ-IEEE’04, Proceedings of the 2004*, vol. 2, s: 1574-1579
- [43] K. Imagawa, S. Lu & S. Igi, Color-based hands tracking system for sign language recognition. In *Proceedings-3rd IEEE International Conference on Automatic Face and Gesture Recognition, FG 1998* (pp. 462–467). IEEE Computer Society 1998.
- [44] S. Albayrak ve O. Altun, “İlgi bölgeleri ve yerel tanımlayıcılar ile geliştirilmiş Hough dönüşümü ve en az eylemsizlik eksenli tabanlı hizalama yaklaşımı ile Türk işaret dili tanıma sistemi,” *Bilgisayar Mühendisliği Anabilim Dalı, YTÜ, Doktora Tezi*, 2010
- [45] S. Albayrak ve A. Memiş, “Kinect RGB Görüntülerinde Ve Derinlik Haritalarında Uzam-Zamansal Özellikleri Kullanarak İşaret Dili Tanıma,” *YTÜ, Yüksek Lisans Tezi*, 2013a

- [46] A. Memiş, & S. Albayrak, A Kinect based sign language recognition system using spatio-temporal features. In Sixth International Conference on Machine Vision (ICMV 2013b) (Vol. 9067, p. 90670X). SPIE.
- [47] T. Kayıkçıoğlu ve S. Ketenci, “Aktif ve Düşünsel Parmak Hareketlerine İlişkin EEG Sinyallerinin Sınıflandırılması,” KTÜ, Doktora Tezi, 2018
- [48] G.B. Akar ve A.H. Akış, “Düşük hesap karmaşıklığına sahip öğrenme tabanlı işaret dili tanıma sistemi,” ODTÜ, Yüksek Lisans Tezi, 2018
- [49] L.A. Ersoy ve O. Özdemir, “El Betimleyicileri ve Gezinge Tabanlı Yöntemler Kullanarak Yalıtılmış İşaret Dili Sınıflandırma,” Boğaziçi Üniversitesi, Yüksek Lisans Tezi, 2018
- [50] Z. Korkmaz ve Ş.H. Akalın, “Türkiye Türkçesinin Temeli Oğuz Türkçesinin Gelişimi,” TDK Yayınları, Ankara, 2013, ISBN : 9789751625595 S:55-70
- [51] G. Eryiğit, C. Eryiğit, S. Karabüklü, M. Kelepir, A. Özkul, T. Pamay, H. Köse, Building the first comprehensive machine-readable Turkish sign language resource: methods, challenges and solutions. *Language Resources and Evaluation*, 54(1), 97–121, 2020.
- [52] Türk İşaret Dili Sözlüğü, MEB Yayınları, Ankara 2015 http://orgm.meb.gov.tr/alt_sayfalar/duyurular/1.pdf (Erişim Tarihi: 26.01.2020)
- [53] “İşaret dili tercümanları derneği” <http://www.tidted.org/turk-isaret-dilinin-gelisim-sureci/>, (Erişim Tarihi: 01.03.2020)
- [54] A.R. Weeks, *Fundamentals of Electronic Image Processing*, Spie Press Book ISBN: 9780819491602 , volume: PM32, 1996, pp: 37-40
- [55] C. Samut and G.I. Webb, “Encyclopedia of Machine Learning,” Volume:8. pp:11-25-29 ISBN:978-0-387-30768-8, Springer Science+Business Media, LLC, 2010

- [56] Ç. Özarı, & Ö. Eren, K-Ortalamlar Yönteminin Başlangıç Merkez Seçim Sorunsalı Üzerine Bir Çalışma. *Business & Management Studies: An International Journal* 7(2), 1117–1135. 2019
- [57] Ş. Sağıroğlu, E. Beşdok ve M. Erler, Mühendislikte yapay zekâ uygulamaları-1: yapaysinir ağları, *Ufuk Yayınevi*, 2003, ISBN: 9789759594855, s: 24-47
- [58] V. Vapnik, *Pattern Recognition Using Generalized Portrait Method*, 1963, pp.774–780.
- [59] J.E.T. Akinsola, Supervised Machine Learning Algorithms: Classification and Comparison, *International Journal of Computer Trends and Technology (IJCTT)*, 48, 2017, 128 – 138
- [60] S. Huang, N. Cai, P.P. Pacheco, S. Narrandes, Y. Wang and W. Xu, Applications of Support Vector Machine (SVM) Learning in Cancer Genomics, *Cancer Genomics Proteomics*, 2017, 15, 41 – 51.
- [61] D. K. and B. Srivastava Lekha, Data Classification using Support Vector Machine, *Journal of Theoretical and Applied Information Technology*, 2010
- [62] M. Bramer, *Principles of Data Mining*, Springer, London, 2016
- [63] M. Chih-Min, Y. Wei-Shui and C. Bor-Wen, How the Parameters of K-nearest Neighbor Algorithm Impact on the Best Classification Accuracy: In Case of Parkinson Dataset, *Journal of Applied Sciences*, 2014, 171–176.
- [64] B. Lantz, *Machine learning with R*, Packt Publishing, Birmingham, 2015
- [65] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, ISBN: 978-0387310732. 2., Vol.: 2, 2006, pp:12-15.

- [66] L. Y. Pratt, Experiments on the transfer of knowledge between neural networks. In S. Hanson, G. Drastal, and R. Rivest, editors, Computational Learning Theory and Natural Learning Systems, Constraints and Prospects, volume 1, chapter 19, pages MIT Press, franklinite.mines.colorado.edu: pub/prattpapers/pratt-clnlbook.ps.Z. (Eriřim Tarihi: 22.01.2020)
- [67] D. E. Rumelhart, J. L. McClelland and the PDP Research Group, Parallel distributed processing Explorations in the microstructure of cognition, Vol. 1. Foundations, Cambridge, Massachusetts, and MIT Press and McClelland, J. L., Rumelhart, D. E., & the PDP Research Group, Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 2. Psychological and biological models. Cambridge, Massachusetts: MIT Pres, 1986
- [68] K. Hornik, M. Stinchcombe and H. White, “Multilayer feedforward networks are universal approximators,” Neural networks, vol. 2, no. 5, 1989, pp. 359-366
- [69] U. M. Ludtke, Emotion in Language: Theory – research – application, John Benjamins Publishing Company, ISBN: 9789027267658, 2015, pp:106-109
- [70] S.Çetiner, Yapay Sinir Ađları, Elektronik Kitap https://books.google.com.tr/books/about/YAPAY_S%C4%B0N%C4%B0R_A%C4%9ELARI.html?id=1q6oDwAAQBAJ&printsec=frontcover&source=kp_read_button&redir_esc=y#v=onepage&q&f=false, (Eriřim: 25.01.2020)
- [71] J. Mira, Neural Computation Lecture Notes İn Computer, International Workshop on Artificial Neural Networks Malaga- Torremolinos, Spaine June 1995, Proceeding, IWANN '95, Sprenger ISBN 3 1741 102 204 973, 1995, pp. 4-15.
- [72] D. E. Rumelhart, J. L. McClelland adn the PDP Research Group, Parallel distributed processing Explorations in the microstructure of cognition, Vol. 1. Foundations, Cambridge, Massachusetts, and MIT Press and McClelland, J. L., Rumelhart, D. E., & the PDP Research Group, Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 2. Psychological and biological models. Cambridge, Massachusetts: MIT Pres, 1986

- [73] M. F. Keskenler, E.F. Keskenler, “Geçmişten Günümüze Yapay Sinir Ağları ve Tarihçesi,” *Takvim-i Vekayi Dergisi*, ISSN: 2148-0087, C.5, S.2, 2017, S.8-18.
- [74] S. Roweis, L. Saul and G. Hinton, Global Coordination Of Local Linear Models, *Advances in Neural Information Processing Systems 14*, MIT Press, Cambridge, MA., 2002, pp:491.
- [75] S. University, “CS231n: Convolutional Neural Networks for Visual Recognition,” <http://cs231n.github.io/neural-networks-2>, (Erişim Tarihi: 27.01.2020)
- [76] S. Ruder, “An overview of gradient descent optimization algorithms,” *ArXiv e-prints*, vol. 1609, <http://adsabs.harvard.edu/abs/2016arXiv160904747R> (Erişim Tarihi: 29.01.2020)
- [77] A. Krizhevsky, I. Sutskever and ve G. E. Hinton, “Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*,” 2012, 1097-1105
- [78] I. Goodfellow, Y. Bengio and A. Courville, *Deep learning (Adaptive computation and machine learning)*. Cambridge, Massachusetts: The MIT Press, 2016, pp. 22.
- [79] H. Wang and B. Raj, *On the Origin of Deep Learning*, Language Technologies Institute School of Computer Science Carnegie Mellon University, 2017
- [80] Al. Aggarwal, *The Resurgence of Artificial Intelligence During 1983–2010*. March 22, Datafloq, 2018
- [81] N. Ganatra and A. Patel, *A Comprehensive Study of Deep Learning Architectures, Applications and Tools*, *International Journal of Computer Sciences and Engineering*, Vol. 6, No:12, 2018, pp.701-705.
- [82] M. Jaderberg, K. Simonyan, A. Zisserman and K. Kavukcuoglu, *Spatial Transformer Networks*, *Neural Information Processing Systems (NIPS) Papers* published at the *Neural Information Processing Systems Conference*, 2015

- [83] C. Szegedy, V. Vanhoucke, S. offe, J. Shlens and Z. Wojna, Rethinking the inception architecture for computer vision. Arşiv Preprint Arşiv: 1512.00567, 2015
- [84] F. Chollet, *Deep Learning with Python*, Manning Publications, ISBN- 13: 978-1617294433, 2017
- [85] MathWorks introduces Release 2017a (R2017a) with a range of new capabilities in MATLAB and Simulink.
- [86] N. K. Ure and U. Genc, CNN to RL how AI is changing the ADAS scene, Eaton Technologies, 2019
- [87] Z. C. Lipton, J. Berkowitz and C. Elkan, A Critical Review of Recurrent Neural Networks for Sequence Learning, arXiv:1506.00019v4 [cs.LG] 17 Oct 2015.
- [88] K. He, X. Zhang, S. Ren and J. Sun, Deep Residual Learning for Image Recognition, Microsoft Research. arXiv:1512.03385v1 [cs.CV] 10 Dec 2015.
- [89] S. Targ, D. Almeida, K. Lyman, Resnet in Resnet: Generalizing Residual Architectures, Workshop track - ICLR 2016.