

**T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**KINECT RGB GÖRÜNTÜLERİNDE VE DERİNLİK HARİTALARINDA UZAM-
ZAMANSAL ÖZELLİKLERİ KULLANARAK İŞARET DİLİ TANIMA**

ABBAS MEMİŞ

**YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ PROGRAMI**

**DANIŞMAN
YRD. DOÇ. DR. SONGÜL ALBAYRAK**

İSTANBUL, 2013

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**KINECT RGB GÖRÜNTÜLERİNDE VE DERİNLİK HARİTALARINDA UZAM-
ZAMANSAL ÖZELLİKLERİ KULLANARAK İŞARET DİLİ TANIMA**

Abbas MEMİŞ tarafından hazırlanan tez çalışması 17.06.2013 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Yrd. Doç. Dr. Songül ALBAYRAK
Yıldız Teknik Üniversitesi

Jüri Üyeleri

Yrd. Doç. Dr. Songül ALBAYRAK
Yıldız Teknik Üniversitesi

Prof. Dr. Nizamettin AYDIN
Yıldız Teknik Üniversitesi

Yrd. Doç. Dr. Hatice KÖSE
İstanbul Teknik Üniversitesi

Bu alıřma, Yıldız Teknik Üniversitesi Bilimsel Arařtırma Projeleri Koordinatörlüğü' nün 2012-04-01-YL03 numaralı projesi ile desteklenmiřtir.

ÖNSÖZ

Lisans bitirme tezim ile yüksek lisans tezimde birlikte çalışma onuruna eriştiğim, bana çalışmalarım süresince her zaman destek olarak cesaretlendiren, bilgi ve tecrübelerini benimle paylaşarak olaylara bilimsel açıdan bakma pratiğini bana öğreten değerli hocam ve tez danışmanım Yrd. Doç. Dr. Songül Albayrak'a en içten teşekkürlerimle ve saygılarımla.

Çalışmalarımızın veri toplama aşamasında bizlere destek olan İşitme Engelliler Federasyonu başkanı sayın Muammer Ay'a, Hakan Kınık'a, Selman Karadoğan'a, Murat, Ahmet ve Zübeyir Bey'e çalışmamıza katkılarından dolayı teşekkür ederim.

Çalışmalarımıza destek olan Yıldız Teknik Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü'ne teşekkür ederim.

Tez çalışmalarım süresince benden desteklerini esirgemeyen küçük aileme teşekkür ediyorum. Uzunca bir süre aynı evi paylaştığım, değerli mühendis arkadaşlarım Anıl Gündoğdu'ya ve Ertunç Dayal'a teşekkür ederim.

Lisans ve yüksek lisans öğrenim hayatımı geçirdiğim Yıldız Teknik Üniversitesi'ne, Y.T.Ü. Bilgisayar Mühendisliği Bölümü'ne, Y.T.Ü. Bilgisayar Mühendisliği Bölümü'nün emektar ve değerli akademisyenlerine ayrıca teşekkürü bir borç bilirim. İstanbul Teknik Üniversitesi, Bilgisayar Mühendisliği Bölümü'nün değerli hocalarına ve bir süre beraber mesai yaptığımız araştırma görevlisi arkadaşlarıma bana destekleri için teşekkür ederim.

Mayıs, 2013

Abbas MEMİŞ

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ.....	vii
KISALTIMA LİSTESİ.....	viii
ŞEKİL LİSTESİ.....	ix
ÇİZELGE LİSTESİ	xii
ÖZET	xvi
ABSTRACT	xix
BÖLÜM 1	
GİRİŞ	1
1.1 Literatür Özeti	1
1.1.1 Türk İşaret Dili Üzerine Mevcut Çalışmalar.....	7
1.2 Tezin Amacı.....	8
1.3 Hipotez.....	9
BÖLÜM 2	
İŞARET DİLLERİ VE İŞARET DİLİ TANIMA SİSTEMLERİ	12
2.1 İşaret Dilleri ve Genel Özellikleri.....	12
2.2 İşaret Dili Tanıma Sistemleri	14
2.2.1 Elektronik Eldiven Tabanlı Sistemler	15
2.2.2 Görü Tabanlı Sistemler	16
2.2.2.1 Hareket Tabanlı Sistemler	16
2.2.2.2 Şekil Tabanlı Sistemler	17
BÖLÜM 3	
İŞARET DİLİ VERİSETLERİ	19
3.1 Amerikan İşaret Dili (ASL) Veri Seti	19
3.2 Türk İşaret Dili (TİD) Veri Seti.....	21

BÖLÜM 4

ÖNERİLEN İŞARET DİLİ TANIMA SİSTEMİ	27
4.1 Görüntü Ön-işleme	28
4.2 Öznitelik Çıkarımı	30
4.2.1 Zamansal Özniteliklerin Çıkarımı	30
4.2.2 Uzamsal Özniteliklerin Çıkarımı.....	36
4.3 Öznitelik Vektörlerinin Oluşturulması	37
4.4 İşaret Tanıma	39

BÖLÜM 5

DENEYSEL SONUÇLAR	40
5.1 ASL Veri Seti Üzerinde Sistem Performansı Değerlendirmeleri.....	42
5.2 TİD Veri Seti Üzerinde Sistem Performansı Değerlendirmeleri	48
5.2.1 Toplamsal İmge Yaklaşımında Çerçeve Ağırlıklandırma	78
5.2.2 Çerçeve Gruplama ve Gruplamalı Toplamsal İmge Yaklaşımı.....	80
5.2.3 Toplamsal İmgelerde Ortalama Eşik Katsayısı Yaklaşımı	85

BÖLÜM 6

SONUÇ VE ÖNERİLER	115
KAYNAKLAR	118
ÖZGEÇMİŞ	122

SİMGE LİSTESİ

α	Ayrık Kosinüs Dönüşümü sabit katsayısı
B	RGB renk uzayında mavi renk kanalı parlaklık değeri
C_{NZ}	Video görüntü çerçevesinde sıfırdan farklı piksel sayısı
$F(u, v)$	Ayrık Kosinüs Dönüşümü çıkış imgesi
g	Video çerçevesi gruplandırma grup sayısı
G	RGB renk uzayında yeşil renk kanalı parlaklık değeri
G_{sc}	Gri-ölçekte parlaklık değeri
$I(m, n)$	Ayrık Kosinüs Dönüşümü giriş imgesi
I_E	Eşikleme uygulanmış imge
I_F	Ardışıl iki video çerçevesi arasındaki fark imgesi
I_i	Video görüntü çerçevesi
M	XY koordinat düzleminde y doğrultusundaki imge çözünürlüğü
n	Video görüntüde yer alan toplam çerçeve sayısı
N	XY koordinat düzleminde x doğrultusundaki imge çözünürlüğü
R	RGB renk uzayında kırmızı renk kanalı parlaklık değeri
T	Eşikleme katsayısı
w	Ağırlıklandırma katsayısı
μ	Gauss dağılımı ortalama değeri
θ	Ayrık Kosinüs Dönüşümü radyan açı değeri
σ	Gauss dağılımı standart sapma değeri

KISALTMA LİSTESİ

ASL	American Sign Language
CCD	Charge Coupled Device
CRR	Correct Recognition Rate
DCT	Discrete Cosine Transform
HCI	Human Computer Interaction
HMM	Hidden Markov Models
K-NN	K-Nearest Neighbor
RGB	Red Green Blue
TİD	Türk İşaret Dili
TSL	Turkish Sign Language

ŞEKİL LİSTESİ

	Sayfa
Şekil 2. 1	İşaret dili tanıma sistemlerinin genel hiyerarşik yapısı15
Şekil 2. 2	Elektronik eldiven (cyberglove with 18 sensors) [42]16
Şekil 2. 3	Durağan olmayan (hareketli) işaret örneği (tid'de "dün" kelimesi) [2]17
Şekil 2. 4	Durağan işaret örneği (tid'de "a" harfi) [2]18
Şekil 3. 1	Asl veri setinde "computer" kelimesinin bir işaretçi tarafından ifade edilmesine ait seçilmiş ardışıl çerçeveler20
Şekil 3. 2	Kinect algılayıcı.....24
Şekil 3. 3	İşaret görüntülerinin kaydında sistem yapısı25
Şekil 3. 4	Tid veri setinde "günaydın" kelimesine ait örnek rgb ardışıl çerçeveleri ...25
Şekil 3. 5	Tid veri setinde "günaydın" kelimesine ait örnek derinlik haritası ardışıl çerçeveleri26
Şekil 4. 1	(a) Rgb renk uzayında yer alan örnek çerçeveler (b) gri-ölçeğe dönüştürülmüş çerçeveler.....29
Şekil 4. 2	Tid'de "günaydın" kelimesinin bir işaretçi tarafından uygulanışına ait seçilmiş gri-imge ardışıl çerçeveleri ve bu çerçevelere ait fark imgeleri. ...31
Şekil 4. 3	Tid'de "günaydın" kelimesinin bir işaretçi tarafından uygulanışına ait seçilmiş derinlik haritası ardışıl çerçeveleri ve bu çerçevelere ait fark imgeleri.....32
Şekil 4. 4	Gri-imgelere ait (a) gri-ölçek ve (b) renklendirilmiş toplamsal imgeler.....34
Şekil 4. 5	Derinlik haritalarına ait (a) gri-ölçek ve (b) renklendirilmiş toplamsal imgeler.....34
Şekil 4. 6	Gri-imgelere ait süzgeçlenmiş (a) gri-ölçek ve (b) renklendirilmiş toplamsal imgeler.....35
Şekil 4. 7	Derinlik haritalarına ait süzgeçlenmiş (a) gri-ölçek ve (b) renklendirilmiş toplamsal imgeler.35
Şekil 4. 8	Toplamsal imge yaklaşımı ve zamansal öznitelik çıkarımı.....36
Şekil 4. 9	Zigzag taramanın örnek bir imge üzerinde uygulanışı.38
Şekil 4. 10	Önerilen sistemde öznitelik çıkarımı ve öznitelik vektörü oluşturulması ...38
Şekil 5. 1	Asl veri setinde farklı mesafe ölçümlerini kullanan knn sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (4-katlı çapraz doğrulama).42
Şekil 5. 2	Asl veri setinde manhattan mesafesini kullanan knn sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı

	oranlarındaki doğru tanıma oranları (4-katlı çapraz doğrulama ve değişen knn k-parametrelerinde).	44
Şekil 5. 3	Asl veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (farklı k-katlı çapraz doğrulamalarda).	46
Şekil 5. 4	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve gri-imge + derinlik haritası özniteliklerinde).....	49
Şekil 5. 5	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve gri-imge özniteliklerinde).	52
Şekil 5. 6	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve derinlik haritası özniteliklerinde).	54
Şekil 5. 7	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime).	57
Şekil 5. 8	Tid veri setinde manhattan mesafesini kullanan knn (K=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1 ve gri-imge + derinlik haritası özniteliklerinde).....	59
Şekil 5. 9	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1).....	62
Şekil 5. 10	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2 ve gri-imge + derinlik haritası özniteliklerinde).....	64
Şekil 5. 11	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2).....	67
Şekil 5. 12	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3 ve gri-imge + derinlik haritası özniteliklerinde).....	69
Şekil 5. 13	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerde ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3).	72
Şekil 5. 14	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 3 kategoride yer alan 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (gri-imge + derinlik haritası öznitelikleri).	74

Şekil 5. 15	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime ve gri-imge + derinlik haritası özniteliklerinde).86	86
Şekil 5. 16	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime ve gri-imge özniteliklerinde).89	89
Şekil 5. 17	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime ve derinlik haritası özniteliklerinde).91	91
Şekil 5. 18	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime).94	94
Şekil 5. 19	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1 ve ortalama eşik katsayısı yaklaşımıyla gri-imge + derinlik haritası özniteliklerinde).96	96
Şekil 5. 20	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-1).99	99
Şekil 5. 21	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2 ve ortalama eşik katsayısı yaklaşımıyla gri-imge + derinlik haritası özniteliklerinde).101	101
Şekil 5. 22	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-2).104	104
Şekil 5. 23	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3 ve ortalama eşik katsayısı yaklaşımıyla gri-imge + derinlik haritası özniteliklerinde).106	106
Şekil 5. 24	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerde ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-3).109	109
Şekil 5. 25	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 3 kategoride yer alan 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında gri-imge + derinlik haritası öznitelikleri).111	111

ÇİZELGE LİSTESİ

	Sayfa
Çizelge 2. 1	Elektronik eldiven ve görü tabanlı sistemlerin temel karakteristikleri [26]14
Çizelge 3. 1	Asl veri setinde yer alan kelimeler ve türkçe karşılıkları20
Çizelge 3. 2	Tid veri setinde kelime kategorileri ve kategorilerdeki kelime sayıları...21
Çizelge 3. 3	Günlük konuşma terimleri kategorisinde yer alan kelimeler22
Çizelge 3. 4	Zaman terimleri kategorisinde yer alan kelimeler22
Çizelge 3. 5	Meslekler kategorisinde yer alan kelimeler23
Çizelge 5. 1	Asl veri setinde manhattan ve euclidean mesafesini kullanan knn sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (4-katlı çapraz doğrulama).....43
Çizelge 5. 2	Asl veri setinde manhattan mesafesini kullanan knn sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (4-katlı çapraz doğrulama ve değişen knn k-parametrelerinde).45
Çizelge 5. 3	Asl veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (farklı k-katlı çapraz doğrulamalarda).....47
Çizelge 5. 4	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve gri-imge + derinlik haritası özniteliklerinde).50
Çizelge 5. 5	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve gri-imge özniteliklerinde).....53
Çizelge 5. 6	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve derinlik haritası özniteliklerinde).55
Çizelge 5. 7	Tid veri setinde manhattan mesafesini kullanan knn (K=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime).....58

Çizelge 5. 8	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1 ve gri- imge + derinlik haritası özniteliklerinde).60	60
Çizelge 5. 9	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1).63	63
Çizelge 5. 10	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2 ve gri- imge + derinlik haritası özniteliklerinde).65	65
Çizelge 5. 11	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2).68	68
Çizelge 5. 12	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3 ve gri- imge + derinlik haritası özniteliklerinde).70	70
Çizelge 5. 13	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3).73	73
Çizelge 5. 14	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 3 kategoride yer alan 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (gri-imge + derinlik haritası öznitelikleri).75	75
Çizelge 5. 15	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki en başarılı doğru tanıma oranları (111 kelime ve gri-imge + derinlik haritası öznitelikleri).77	77
Çizelge 5. 16	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının gauss temelli ağırlıklandırma yaklaşımında farklı kategoriler için doğru tanıma oranları (dct katsayı oranı : % 0,52).79	79
Çizelge 5. 17	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının sabit katsayı ile ağırlıklandırma yaklaşımında farklı kategoriler için doğru tanıma oranları (dct katsayı oranı : % 0,52).79	79
Çizelge 5. 18	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırmaya dayalı toplamsal imge yaklaşımında farklı kategoriler için doğru tanıma oranları (toplam dct katsayı oranı: % 0,52, grup sayısı: 5).81	81
Çizelge 5. 19	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırmaya ve grup eksiltmeye dayalı toplamsal imge yaklaşımında farklı kategoriler için doğru tanıma oranları (toplam dct katsayı oranı: % 0,52, grup sayısı: 5).81	81
Çizelge 5. 20	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırmaya ve gruplardan dağılıma bağlı öznitelik seçemeye dayalı toplamsal imge yaklaşımında farklı kategoriler	

	için doğru tanıma oranları (toplam dct katsayı oranı: % 0,52, grup sayısı: 5).....	82
Çizelge 5. 21	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırmaya ve çerçeve eksiltmeye dayalı toplamsal imge yaklaşımında farklı kategoriler için doğru tanıma oranları(toplam dct katsayı oranı: % 0,52, grup sayısı: 5).	83
Çizelge 5. 22	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırmaya dayalı toplamsal imge ve genel gauss dağılımı temelli yaklaşımda farklı kategoriler için doğru tanıma oranları (toplam dct katsayı oranı: % 0,52, grup sayısı: 5).	84
Çizelge 5. 23	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırmaya dayalı toplamsal imge ve gruplamalı gauss dağılımı temelli yaklaşımda farklı kategoriler için doğru tanıma oranları (toplam dct katsayı oranı: % 0,52, grup sayısı: 5).	85
Çizelge 5. 24	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime ve gri-imge + derinlik haritası özniteliklerinde).	87
Çizelge 5. 25	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime ve gri-imge özniteliklerinde).	90
Çizelge 5. 26	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime ve derinlik haritası özniteliklerinde). ..	92
Çizelge 5. 27	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime).....	95
Çizelge 5. 28	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1 ve ortalama eşik katsayısı yaklaşımıyla gri-imge + derinlik haritası özniteliklerinde).	97
Çizelge 5. 29	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-1).	100
Çizelge 5. 30	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2 ve ortalama eşik katsayısı yaklaşımıyla gri-imge + derinlik haritası özniteliklerinde).	102

Çizelge 5. 31	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-2).	105
Çizelge 5. 32	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3 ve ortalama eşik katsayısı yaklaşımıyla gri-imge + derinlik haritası özniteliklerinde).	107
Çizelge 5. 33	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-3).	110
Çizelge 5. 34	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 3 kategoride yer alan 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında gri-imge + derinlik haritası öznitelikleri).	112
Çizelge 5. 35	Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki en başarılı doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında 111 kelime ve gri-imge + derinlik haritası öznitelikleri).	114

KINECT RGB GÖRÜNTÜLERİNDE VE DERİNLİK HARİTALARINDA UZAM-ZAMANSAL ÖZELLİKLERİ KULLANARAK İŞARET DİLİ TANIMA

Abbas MEMİŞ

Bilgisayar Mühendisliği Anabilim Dalı
Yüksek Lisans Tezi

Tez Danışmanı: Yrd. Doç. Dr. Songül ALBAYRAK

İnsan bilgisayar etkileşimi (HCI-Human Computer Interaction), 20. yy.'ın son çeyreği ile 21. yy.'ın ilk yıllarında bilgisayar sistemlerinin hızlı bir biçimde gelişmesine paralel olarak akademik dünyadaki çalışmalarda ve bu çalışmalara bağlı olarak gündelik yaşamda kendine yer edinmiştir. Özellikle makine öğrenmesi, robotik, görüntü işleme ve bilgisayarla görü alanlarındaki problem odaklı araştırma ve çalışmalar ile bu alan ivmeli bir biçimde genişlemektedir. Yeni problemlerin hayata girmesi ve yeni ihtiyaçların doğması da kuşkusuz insan bilgisayar etkileşimine dayalı çalışmalara ve çalışma alanlarına etki etmekte ve yön vermektedir.

İnsan bilgisayar etkileşimi dâhilinde temel olarak bilgisayar yazılımlarının, bilgisayar işletim sistemlerinin, robotların, uzaktan veya yakından kumanda edilmesi planlanan araçların kontrollerine yönelik araştırma ve uygulamalar yapılmaktadır. Bilgisayar sistemlerinin komuta edilebilmesinde ve bu sistemlerle etkileşimlerin sağlanmasında ses ve görüntü işaretleri kullanılmakta olup bu işaret türleri probleme ve ihtiyaca bağlı olarak tercih edilebilmektedir.

Son yıllarda özellikle görüntü işleme ve bilgisayarla görü alanlarındaki gelişmelerle beraber görü tabanlı sistemlerin insan bilgisayar etkileşimlerinde kullanımı yaygınlaşmaktadır. Görü tabanlı sistemler, kullanım kolaylığı ve sistem basitliği bakımından avantaj sağlamakla beraber ortam aydınlığındaki farklılıklar ve örtüşme gibi sebeplerle sorun çıkartabilmektedir. Bu tarz sistemlere yönelik sorunların azaltılması için problem odaklı çözümlere yönelinmektedir.

Elektronik cihaz ve bilgisayar yazılımına kumanda ederek etkileşim kurmak için görü tabanlı sistemler tercih edilmektedir. Etkileşimi sağlamak amacıyla, etkileşimde kullanılan öğelerin bilgisayar sistemleri tarafından yüksek başarımları ile tanınması da gerekmektedir. Bu amaçla makine öğrenmesi de bu tarz çalışma alanlarında sistemlerin eğitilmesi ve tanıma işlemleri gerçekleştirilerek karar verilmesinde önem arz etmektedir.

Sistemlerle etkileşim kurabilmek adına ses gibi elektriksel işaretlerin yanı sıra el, yüz, kafa ve vücut ile yapılan görsel işaretler de kullanılabilir. Son yıllarda özellikle el hareketleri kullanılarak kontrol edilen sistemlerin sayısında önemli artışlar görülmektedir. Bu sistemler arasında, işitme ve konuşma engellilerin iletişim sağlamak amacıyla kullandıkları işaret dillerinin tanınmasına yönelik araştırmalar da mevcuttur. Bu çalışmada işaret dillerinin tanınması amacıyla mevcut yöntemler incelenerek tanıma başarısını arttırmak hedeflenmiştir.

İşaret dili konuşma ve işitme engellilerin kendi aralarındaki ve başkaları ile iletişimlerinde kullandıkları; el, yüz, vücut hareketlerinden oluşan görsel bir dildir. İşaret dil, işitme ve konuşma engellilerin ana dili olarak tanımlanır ve iletişimlerinin çok önemli bir bölümünü oluşturur. Bu sebeple bu dillerin bilgisayar sistemleri tarafından tanınarak yorumlanması, her ne kadar teknolojik açıdan önemli ise sosyal açıdan da büyük önem arz etmektedir.

İşaret dillerine ait hareketler durağan (posture) ve durağan olmayan (gesture) hareketlerden oluşmaktadır. Bu tez çalışmasında değerlendirilen hareketler durağan olmayan video görüntüleridir. İşaretlere ait özniteliklerin çıkarımında iki aşamalı uzam-zamansal bir yapı kullanılmıştır. İlk aşamada işaretlerin zamansal öznitelikleri, ardışıl görüntü çerçeveleri arasındaki parlaklık farklarına dayalı toplamsal imge yaklaşımı kullanılarak çıkarılmış ve tek bir imge ile ifade edilmiştir. İkinci aşamada zaman uzayında çıkarılan bu öznitelikler Ayrık Kosinüs Dönüşümü (DCT - Discrete Cosine Transform) ile uzamsal özniteliklere dönüştürülmüştür. Kullanılan dönüşüm yöntemi ile elde edilen en yüksek enerjili katsayılar farklı oranlarda, zig-zag tarama yöntemi ile seçilerek tanıma aşamasında kullanılacak öznitelikler oluşturulmuştur.

Yaklaşımın başarımını sınamak amacıyla K-en yakın komşu (K-NN) sınıflandırıcısı kullanılmıştır. Sistemin eğitim ve test aşamalarında Amerikan İşaret Dili'ne (ASL) ait 20 kelimedenden, toplamda ise 800 işaretten oluşan bir veri seti ile bu tez çalışması kapsamında yeni oluşturulan Türk İşaret Dili'ne (TİD) ait 111 kelimedenden, toplamda ise 1002 işaretten oluşan bir veri seti kullanılmıştır. Test aşamalarında sistemin verimliliğini sınamak amacıyla çapraz doğrulama ile farklı test örnekleri seçilmiştir. Bu sayede her örneğin eğitim ve test kümelerinde kullanımı sağlanarak daha genel geçer bir sonuç alınması hedeflenmiştir.

Tez kapsamında Kinect algılayıcı kullanılarak oluşturulan Türk İşaret Dili'ne ait veri setinde RGB-D video görüntülerinde derinlik bilgisi de tanıma başarımını arttırmada etkin olarak kullanılmıştır. Ardışıl hareket farklarını ve dönüşüm yöntemlerini kullanarak uzam-zamansal öznitelikleri çıkaran sistemle ve K-en yakın komşu sınıflandırıcısı ile ASL veri setinde %95-99, TİD veri setinde %80-98 arasında değişen başarımlarla işaret diline ait durağan olmayan görüntüler tanınabilmiştir.

Anahtar Kelimeler: İnsan bilgisayar etkileşimi, işaret dili tanıma, uzam-zamansal özellikler, hareket farkları analizi, kinect algılayıcı

**SIGN LANGUAGE RECOGNITION USING SPATIO-TEMPORAL FEATURES ON
KINECT RGB VIDEO SEQUENCES AND DEPTH MAPS**

Abbas MEMİŞ

Department of Computer Engineering

MSc. Thesis

Advisor: Assist. Prof. Dr. Songül ALBAYRAK

Human Computer Interaction (HCI), has become widespread in academical researches and so in daily life with the growing up of computer systems in the last quarter of 20th century and first years of 21th century rapidly. Especially problem-focused researches and approaches in machine learning, robotic, image processing and computer vision improve HCI research area and make it grow up. New problems and requirements in life, also affect and direct the HCI based applications and research areas.

HCI mainly includes approaches and applications in controlling of computer softwares, computer operating systems, robots and control of devices that can be controlled near-fear or far-away. Speech and image signals have been used in order to control computer systems, interact with these kind of systems and these approaches are preferred considering the system requirements and problems.

Recent advances in computer vision enable vision based systems to be widespread in HCI systems. Although vision based systems have advantages on system simplicity and ease of use, can have disadvantages like illumination effects in environment and overlapping situations. So problem-focused approaches are proposed in researches in order to eliminate the effects of these kind of problems.

Vision based systems have been used for interaction in controlling electronical devices and computer softwares. Also components that used for interaction should be

recognized efficiently for interaction accuracy. So machine learning approaches carry weight in this kind of studies for system decisions that are based on recognition and training processes.

Not only electrical signals like speech can be used for interactions with computer systems but also visual signs which are carried out using hands, face, head and human body can be used. Especially systems that are controlled by hand signs, have increased significantly in recent years. These systems also works includes researches on recognition of sign languages which are used by deaf community for communication. This thesis aims to improve the recognition success rates for sign language recognition by analyzing the related and existing approaches.

Sign languages are visual languages consist of hand, face and body motions and which are used by deaf community for communication with each other and others. Sign languages are native languages of deaf people and constitute a important percent of their communication. So, not only recognition and interpretation of these languages via computer systems are important in technological advances, but also have quite importance in social perspective.

Sign languages consist of static signs (postures) and non-static signs (gestures). Proposed sign language recognition system is performed on videos of non-static (dynamic signs). In feature extraction process, a two stage spatio-temporal structure is performed. Firstly, temporal features of signs are extracted using accumulated motion image approach which is based on the intensity differences of sequential video frames and these features are presented in a single image. In second step, these features, which are in time domain, are transformed into spatial features via Discrete Cosine Transform (DCT). DCT provides coefficients that contain higher energy and feature vectors, which will be used for recognition, are obtained by selecting these coefficients in different ratios via zig-zag scanning.

K-Nearest Neighbor (K-NN) classifier is employed for performance evaluation. System performance is evaluated on a dataset which contains 20 words belong to American Sign Language (ASL) and totally 800 sign videos and evaluated on a second dataset, which is collected in scope of thesis works, contains 111 words belong to Turkish Sign Language (TSL) and totally 1002 sign videos. In system recognition process, different test samples are selected via K-fold cross validation for performance analysis on datasets. Thus, it is aimed to obtain more valid success rates by using every sample in train and test sets both.

Depth information of RGB-D video sequences are used effectively for improving the recognition rates of Turkish Sign Language dataset which have sign video samples captured by a Kinect sensor. Non-static signs are successfully recognized by proposed system that extract spatio-temporal features using sequential motion differences and transformation methods and that use K-NN classifier. Proposed system also has recognition rates between %95-99 on ASL dataset and %80-98 on TSL dataset.

Keywords: Human computer interaction, sign language recognition, spatio-temporal features, motion differences analysis, kinect sensor

1.1 Literatür Özeti

Bilgisayar sistemlerinin 1950'li yıllardan itibaren sürekli gelişim gösteren dünyasında, özellikle 20. yy.'ın son çeyreği ile içinde bulunduğumuz 21. yy'da akıllı sistemlerin gelişimi önemli yer tutmaktadır. Bilgisayar dünyasındaki gelişme, beraberinde bilgisayar ağları, yapay zekâ, programlama dilleri ile yazılım vb. gibi alt çalışma disiplinlerinin oluşmasını da beraberinde getirmiştir. Bu sayede farklı çalışma alanlarında problemlere odaklı çalışmalar ile yeniliklere dayalı çalışmalar yürütülmektedir. Akıllı sistemler de makine öğrenmesi, doğal dil işleme, bilgisayar görme ve görüntü işleme, makine çevirisi, robotik gibi çalışma disiplinlerinde yürütülen akademik ve sektörel çalışmalarla gelişimini sürdürmektedir.

Bilgisayar dünyasında, akıllı sistemler üzerine yapılan çalışmalarda insan bilgisayar etkileşimi önemli yer tutmaktadır. Nihayetinde geliştirilen ve geliştirilmesi planlanan sistemlerin birçoğunun içinde bir öge olarak insan yer almaktadır. İnsan bilgisayar etkileşimi temel alınarak çalışılan sistemler genellikle kumanda kontrolü sağlayan sistemler olmakla birlikte; biyometrik tanıma sistemlerinin, robotik sistemlerin de kapsamı içerisinde yer almaktadır.

Akıllı sistemlerin geliştirilmesinde, insanın zekâ yetisi makine öğrenmesi yöntemleri ile bilgisayar sistemlerine kazandırılmaktadır. Bu sayede bir veri kümesi kullanılarak eğitilen sistemler, eğitim setlerine bağlı olarak tanınması hedeflenen öğeleri birbirinden ayırarak tanıyabilmektedir.

İnsan bilgisayar etkileşimi temel alınarak geliştirilen sistemlerde ellerin, kolların şekil ve hareketleri etkileşimde sıklıkla kullanılmaktadır. El şekilleri ve hareketleri ile etkileşim kurulan sistemlerde genellikle bir sistemin kumanda edilmesi sağlanmaya çalışılmaktadır. Bu amaçla yapılan araştırma ve çalışmalar farklı alt problemlere yöneliktir. Aşağıda listelenen bazı insan bilgisayar etkileşimi temelli çalışma konularında ellerin hareket ve şekilleri kullanılarak araştırma faaliyetleri yürütülmektedir.

- Robot sistemlerin kontrolü
- Bilgisayar oyunlarının oynanması
- Elektronik cihaz (televizyon, bilgisayar, akıllı ev cihazları vb.) kontrolü
- Bilgisayar yazılımlarının kontrolü
- İşaret dillerinin tanınması

İşaret dilleri de, insan bilgisayar etkileşiminin bir alt çalışma alanı olarak üzerine çalışma yürütülen bir konudur. İşaret dilleri, işitme ve konuşma engellilerin iletişimde kullandıkları el, kol, yüz, baş hareketleri ile yüz ifadelerinden oluşan görsel bir dildir. Bu işaretlerin verimli bir şekilde bilgisayar sistemleri tarafından tanınabilmesi gerek bilgisayar sistemleri açısından gerekse de engelli kişilerin günlük hayattaki iletişim problemlerinin en aza indirilmesi açısından sosyal projeler kapsamında önem arz etmektedir.

İşaret dilleri, tıpkı ulusal diller gibi dünya üzerinde bir bütünlük arz etmemektedir; ülkelerin konuştukları diller gibi birbirinden ayrılabilirler. Bununla birlikte aynı dili konuşan ülkelerin işaret dilleri de birbirinden farklı olabilmektedir [1], [2]. Bu sebeple işaret dilleri üzerine yürütülen çalışmalar genelde ulusal olmakla birlikte, birçok dili kapsayacak biçimde de çalışmalar yürütülmektedir. Ülkelerde kullanılan işaret dillerinin farklı olması, işaret dili temel alınarak geliştirilen sistemlerin verimliliğini önemli ölçüde etkileyebilmektedir. Bu etki, dilin yapısına göre artan veya azalan yönde olabilmektedir. Bu sebeple işaret dili temelli bir sistem bir ülkenin işaret dili için yüksek verimlilikte çalışabilirken, başka bir ülkenin işaret dilinde aynı verimliliği gösteremeyebilir.

İşaret dillerini tanınmasına yönelik yapılan ilk çalışmalarda duyarlı eldivenler kullanılarak hareketçiden alınan veriler değerlendirilmiştir. Sistemi kullanan kişinin harekete duyarlı eldivenleri giyerek işaret diline ait ifadeleri icra etmesi ile duyarlılardan alınan bilgiler bilgisayara iletilmiş ve sistem tarafından değerlendirilmiştir [3]. Bu tarz sistemlerin veri güvenilirliği yüksek olmasına rağmen; yüksek derecede kablo karmaşıklığı içermesi ve kullanıcıyı belli bir derecede kısıtlaması sebebiyle kullanımları yaygın değildir. Bu sistemlere ek olarak, 1990'lı yılların ortalarından itibaren işaret dillerinin görü (kamera) tabanlı sistemler ile tanınmasına yönelik çalışmalar yapılmaya başlanmıştır [4]. Görü tabanlı sistemler; işaretleri bir kamera aracılığı ile bilgisayar ortamına iletip değerlendirdiği için oldukça pratiktir. Ancak görüntü işleme ve bilgisayarla görü çalışma alanlarındaki karşılaşılması muhtemel problemler ile verimlilikler, duyarlı sistemlere nazaran daha düşük olabilmektedir.

İşaret dillerinin görü tabanlı olarak tanınmasına yönelik ilk çalışmalar 90'lı yıllarında ortalarında yapılmıştır. Bu çalışmalardan birini gerçekleştiren Starner, basit bir kamera kullanarak işaret diline ait kelimeleri tanımaya çalışmıştır [4]. Hareketçiden alınan görüntülerde eller bölütlenmiş ve ellerin konumlarının da dâhil olduğu öznitelik vektörleri saklı Markov modelleri [5] ile modellenip işaret tanıma işlemi gerçekleştirilmiştir. Amerikan İşaret Dili (ASL)'ne ait 40 adet kelime üzerinde yapılan çalışmada %90'ın üzerinde başarımlar sağlanmıştır. Starner aynı zamanda işaret dili tanıması yapabilen bir giyilebilir bilgisayar sunmuştur [6]. Bu çalışmada bir şapkaya yerleştirilen kamera ile alınan görüntülerde, eller bölütlenerek ellere ait şekil, konum gibi özelliklerden tanıma yapılmıştır. Tanıma işlemlerini sınamak amacıyla yine ASL'ye ait 40 kelimedenden oluşan bir işaret veri seti üzerinde performans ölçümü yapılmış ve %99 oranlarında başarımlar sağlandığı belirtilmiştir.

İşaret dilleri üzerinde yapılan çalışmalarda hareketlerin yazıya çevrilmesine yönelik çalışmalar da mevcuttur. Hernandez – Rebollar tarafından [7]'te sunulan çalışmada ASL'ye ait birbirinden ayrılmış hareketleri yazıya ve sese dönüştüren bir yaklaşım sunulmuştur. 176 işaret ile gerçekleştirilen testlerde %95 oranında bir başarımlar ifade edilmiştir. Sistemde elektronik eldivenlerin ve kol cihazlarının kullanımı birleştirilmiştir. [8]'da ise Hernandez – Rebollar, yine eldiven benzeri cihazlar ile durağan işaretleri %78

ile %100 arasındaki oranlarda tanıyarak, sistemi bir ses sentezleyicisi ile birlikte önermiştir. Isaacs [9]'da iki boyutlu durağan el pozlarının, video tabanlı insan-bilgisayar arayüzlerde tanınmasına yönelik bir çalışma sunmuştur. Dalgacık dönüşümü tabanlı öznitelik vektörleri ileri beslemeli yapay sinir ağları ile sistem eğitiminde kullanılmış ve ASL alfabesindeki 24 durağan işaret %99,9 oranında tanınmıştır.

[10]'da önerilen sistemde, hareket yörüngeleri temel alınarak, video görüntü dizisindeki iki boyutlu hareketi çıkaran ve sınıflandıran bir algoritma sunulmuştur. Sistemde ardışıl görüntülerdeki eşleştirmeler ile hareket doğrultuları hesaplanarak harekete ait özniteliklerin çıkarımını takiben zaman gecikmeli yapay sinir ağı modeli ile ASL'ye ait hareketlerin eğitim işlemi gerçekleştirilmiştir. [11]'de ise görüntüyü alan, bölütleyen, süzgeçleyen, yeniden biçimleyen ve sınıflandırılan gürbüz olarak tanımlanan bir sistem önerilmiştir. ASL'ye ait hareketler siyah-beyaz imgelerde histogram eşiklemesi yolu ile bölütlenerek; bölgesel sınır dizileri arasındaki benzerlik doğrusal hizalama gerçekleştirilerek ölçülmüştür ve sistemin hatasız sınıflandırma yaptığı belirtilmiştir. Kevin tarafından yapılan çalışmada da [12] hareket yörüngeleri temel alınmıştır. Çalışmada el hareketlerine ait veriler, elektronik cihazlar yardımı ile elde edilmiştir. Hareket verilerini oluşturmak için elin üç boyutlu konumunu hesaplayan üç manyetik izleyici ile el açılarını hesaplayan bir çift elektronik eldiven kullanılmıştır. Hareketlerin olasılıksal modellerle tanındığı sistemin yüksek verimle çalıştığı ancak hareket geçişlerindeki etkilere karşı gürbüzlüğünün az olduğu vurgulanmıştır. Boston Üniversitesi, video işleme grubu tarafından ise ASL'de yer alan bazı hareketleri tanımak amacıyla hareketleri otomatik ayırabilen, izleyebilen ve tanıyabilen bir çalışma [13]'de sunulmuştur. Çalışmada efektif bir tanıma yapabilmek için ellerin, kafanın ve yüzün bulunarak izlenmesi için izlenen yöntemlere de yer verilmiştir. Ding tarafından tanımlanan yöntemde [14], önce dilbilimsel açıdan önemli parmakların üç boyutlu şeklini çıkaran daha sonra ise kameranın koordinat sisteminde elin üç boyutlu yörüngesini oluşturan bir algoritma tanımlanmıştır. Çalışmada üst üste gelme ve yer belirleme problemleri de ele alınarak, el özelliklerinin takibinde yer belirleme hatalarına karşı gürbüz ve basit bir algoritma sunulmuş. Tanımlanan yöntem ve teori ASL video görüntüleri üzerinde test edilmiştir.

İşaret dillerinde, hareketlerin başlangıç ve bitişlerinde uygulanışı farklı olabilmektedir. Bu durum işaretlerin ayrılmasında ve tanımlanmasında problemler oluşturabilmektedir. [15]'teki çalışmada bu problem ele alınarak, iki adımlı bir yaklaşım savunulmuştur. İlk aşamada işaretler bölütlenirken, uygulanışta farklı olan noktaları bulmak için şartlı rastgele alanlar kullanılmış ve bu yöntemle sistem eğitimi yapılarak saklı Markov modeline nazaran daha iyi başarımların gözlemlendiği ifade edilmiştir. Thangali [16]'daki yaklaşımında periyodik hareketlerin tespit ve tahmin edilmesinde uzay-zaman örnekleme ile yeni bir teknik önermiştir. Tekniğin önemli yanı olarak özellik takibi ve nesnelerin bölütleme gerektirmemesi vurgulanmıştır. Ardışık video çerçevelerdeki parlaklık bilgileri kullanılarak ASL üzerindeki yapay hareketlerde gerçekleştirilen analizlerde hareket yolu tahmin edilmeye çalışılmıştır.

İşaret dillerindeki ifadeler, iki el kullanılarak yapılan hareketlerin haricinde tek elle gerçekleştirilen hareketleri de kapsamaktadır. [17]'deki çalışma da ASL'deki tek elle yapılan hareketleri temel almaktadır. Elin şekli, yeri ve hareket bilgisi kullanılarak uzam-zamansal ağ işleme ile tanıma gerçekleştirilmiştir. Bu işlemler sırasında Markov zincirleri modellerinden de yararlanılmıştır. ASL'deki el hareketlerine ait alfabe imgelerinin değerlendirildiği bir çalışmada [18] ise, harf ifadesinin olduğu bölge normalleştirilerek ve Radon dönüşümü ile hizalanarak Gabor süzgeçlerinden geçirilmiştir. Temel bileşen analizi ile boyut olarak indirgenen veriler üzerinde %90 üzerinde tanıma gözlemlenmiştir. Yine ASL üzerindeki bir işaret tanıma yaklaşımında ise; el şekli, hareketi ve işaretlerin bağlantı yerlerinin bilinmesi ile hareketlerin tanınabileceği öne sürülmüştür [19]. Doğru sınıflandırmayı sağlamak adına şekil ve hareketin üç boyutlu bilgisinin kullanılması hedef alınmış ve bu amaçla video görüntülerden üç boyutlu model oluşturan yeni bir algoritma önerilmiştir.

İşaret dillerinin tanınmasına yönelik çalışmalardaki alt problemlerden biri de ellerin bulunması ve bölütlenmesidir. Bu bağlamda [20]'de yer alan çalışmada bu probleme odaklanılarak tanıma yapılmaya çalışılmıştır. Ellerin bulunması ile birlikte işaretlerin nerede başlayıp nerede bittiğini bulmak önemsenmiştir. Çalışmada tasarlanan sisteme düzlemsel bölütleme, zamansal bölütleme ve tanıma ile birlikte üç aşamalı bir yaklaşım getirilmiştir. [21]'de sunulan yaklaşımda da sürekli verilerden işaretlerin başlangıç ve bitiş noktalarının belirlenmesi hedeflenmiştir. İşaretlerin hareket ve şekil olarak

birbirinden farklı olması ile gerçekleştirilme süreleri problemler olarak incelenmiş ve değişken uzunluklu işaretlerde yarı otomatik şartlı Markov rastgele alanları yaklaşımı benimsenerek ASL ve Kore İşaret Dili üzerinde çalışmalar yürütülmüştür. Sürekli işaretlerin tanınmasına yönelik ve bölütleme problemini ortadan kaldıracak farklı yaklaşımlardan biri de [22]'deki çalışmada sergilenmiştir. Arap İşaret Dili'nin tanınması amacıyla yapılan bu çalışmada sıralı imgelerdeki hataların eşiklenmesi ile işaret bilgisinin çıkarımı sağlanmıştır. Hataların frekans uzayına aktarılması ile elde edilen özniteliklerin saklı Markov modelleri ile %94 oranında tanınabildiği gösterilmiştir. Çalışmanın eksik yönü olarak ise oluşturulan tanıma sisteminin kullanıcıya bağımlı olması belirtilmiştir. [22]'de yapılan çalışmanın yürütücüleri tarafından sunulan bir başka sistemde [23] kullanıcıdan bağımsız olarak işaretlerin tanınması hedeflenmiştir. İşaretlerin yapıldığı eller renkli eldivenlerin kullanımı ile bölütlenerek; ellerin alanı dâhilindeki yerlerde hata hesaplanması yoluyla hareket bilgisi çıkarılmıştır. Zaman uzayındaki bilgilerin frekans uzayına dönüşümü sağlanarak, polinom ağlarla tanınması sağlanmıştır.

El hareketlerinin tanınması amacıyla yapılan çalışmalarda robotların da kullanımı mevcuttur. Görü sistemi ile birleştirilen robotik sistemler, el hareketleri ile kumanda edilebilmekte ve tanıma yapabilmektedir. [24]'de Luo tarafından gerçekleştirilen çalışmada, hareket eden kamera sistemleri kullanılarak insan işaretlerinin tanınması hedeflenmiştir. Konumlandırılmış yüz ve vücut bölgelerinden elde edilen, el hareketlerinin yörüngelerinin çıkarılması için basit ve durağan bir yöntem olan insan izleyen bölgesel koordinat sistemi kullanılarak işaretlerin yörüngeleri elde edilmiştir. Sistemde yörünge bilgilerinin temel bileşen analizi ile dönüşümü sağlanarak öznitelikleri elde edilmiş ve bu öznitelikler tanıma probleminde kullanılmıştır.

İşaret tanıma üzerine Keskin tarafından yapılmış bir çalışmada [25] ise bilgisayar uygulamalarını kontrol etme amaçlı olan görü tabanlı ve gerçek zamanlı bir hareketsetel arayüz önerilmiştir. Yine Keskin tarafından yapılmış diğer çalışmalarda el hareketlerinin analizleri [26], ve başarılı olarak tanınmasına yönelik çalışmalar [27] mevcuttur.

1.1.1 Türk İşaret Dili Üzerine Mevcut Çalışmalar

İşaret dilleri, tüm dünyada genel bir standarda sahip olmayıp, ülkelere göre farklılıklar göstermektedir. Bu sebeple işaret dillerinin tanınması üzerine yapılan çalışmalar, genellikle belli bir işaret dilinin tanınmasına yönelik olmaktadır. Bu tez çalışmasında da öncelikle Türk İşaret Dili'ne ait kelimelerin tanınması hedeflenmiş ve bu amaçla tez çalışmaları yürütülmüştür.

Literatürde Türk İşaret Dili'ne yönelik yapılmış mevcut çalışmalar da yer almaktadır. Haberdar [28] yapmış olduğu çalışmada TİD'ye ait kelimeleri HMM kullanarak başarıyla tanıyabilmiştir. Bununla beraber Aran'a ait çalışmada ise [29] yine TİD örneklerinin görü tabanlı olarak tanınması problemi ele alınmış ve TİD tanınmıştır.

Türk İşaret Dili yapısı itibarı ile el ve kol hareketlerinin yanı sıra yüz ifadelerini de içermektedir. Arı [30] tarafından yürütülen çalışmada, yüz ifadelerinin takibi ve tanınması sağlanmıştır. Yüz ifadelerinin tanınması amacıyla yapılan bir çalışma [31] ise Güvensan ve Haberdar tarafından sunulmuştur. Ayrıca [32]'de, TİD'de kullanılan baş hareketlerini ve yüz ifadelerini içeren bir veritabanı tanıtılmıştır.

TİD'de parmak heceleme yoluyla ifade edilen kelimelerin tanınması amacıyla yapılmış çalışmalar mevcuttur. Altun ve Albayrak tarafından sunulan yaklaşımlar [33], [34] bu hedefe yöneliktir. Ayrıca Işıkdoğan ve Albayrak tarafından yapılan bir çalışmada [35] ise parmak heceleme görüntülerinin otomatik tanınması sağlanmıştır.

Tanımaya ek olarak, robotlar aracılığı ile işaretlerin gerçekleştirilmesi ve engellilere öğretilmesini hedefleyen Köse tarafından sunulan çalışmalarda da [36], [37] TİD kullanılmıştır. Yapılan bu çalışmalara ek olarak [38], [39], [40], [41] ve [42]'te yer alan farklı problemlere yoğunlaşmış ve yapılmış çalışmalarda da Türk İşaret Dili kullanılmıştır.

Literatürde yer alan çalışmalarla birlikte TİD için Türk Dil Kurumu [43] ve üniversite çalışma grupları [2], [44] sözlük oluşturma ve dili analiz etme çalışmaları yürütmektedir.

1.2 Tezin Amacı

İşaret dillerinin bilgisayar sistemleri tarafından tanınması, son yıllarda üzerine yürütülen araştırmalar sebebi ile önem kazanmıştır. Özellikle getirdiği kullanım kolaylıkları ve giderek yaygınlık kazanması sayesinde görü tabanlı sistemlerin üzerindeki akademik araştırmalar yoğunlaşmaktadır. Tez çalışmasında da bir görü tabanlı sistemle işaret dillerindeki kelimelere ait işaret örneklerinin tanınması hedeflenmiş ve bu bağlamda çalışmalar yürütülmüştür.

İşaret dilleri işitme engeller arasında ortak bir dil içermemektedir. Farklı ülkelerin işaret dilleri farklı olmakla birlikte aynı dili konuşan ülkelerin işaret dilleri de büyük oranlarda farklılık gösterebilmektedir. İşaret dilleri üzerine yapılan çalışmalar genellikle ulusal çalışmalardır. Çalışmalardaki yöntemler, verimler dile bağımlı olabilmektedir. İşaret dilleri işitme engellilerin birbirleriyle iletişimde kullandıkları görsel bir dil olmakla beraber, kendilerine özgü dil yapıları ve kuralları vardır. Bu yapılar ve kurallar dile göre değiştiği için bir işaret dili için kullanılan yöntem verimli olabilirken başka bir dil için aynı verimi gösteremeyebilir. Yapılan tez çalışmasında işaret dillerinin karmaşık olmayan ve verimli bir yöntemle tanınabilmesi amacıyla araştırmalar yürütülmesi ve farklı işaret dillerinin tanınabilmesi hedeflenmiştir.

İşaret dillerinin verimli bir biçimde tanınması ile başarılı bilgisayar çeviri sistemleri, akıllı kumanda sistemleri, robotik sistemlerin geliştirilebilmesi mümkün görünmektedir. Bu sebeplerle yürütülen tez çalışmasının amaçlarından biri de bu alanda yapılan çalışmalarla birlikte işaret dillerinin bilgisayar sistemlerinde kullanımının yaygınlığına katkı yapmaktır.

İşaret dili tanınmasının robotik ve kumanda sistemler başta olmak üzere diğer alanlarda bilimsel ve teknolojik olarak oluşturacağı yenilikler ve kolaylıklar ile birlikte sosyal-toplumsal olarak da önemi bulunmaktadır. Bu alanda yapılan çalışmalar ve geliştirilebilecek sistemler ile işitme engellilerin iletişimdeki ve iletişime bağlı olarak gündelik yaşamdaki problemlerinin en aza indirilmesi de sosyal anlamda önem arz etmektedir. Bu bağlamda yapılan diğer akademik çalışmalarla birlikte, üzerinde çalışma yürütülmüş olan tezin de sosyal anlamda bir katkı oluşturması hedeflenmiştir.

İşaret dili çalışmalarında teknik olarak öne çıkan çok önemli noktalardan biri ise sistemlerinin başarımlarının sınanması aşamasında değerlendirilecek olan veri setleridir. Bu tez çalışmasında özellikle Türk İşaret Dili'ne (TİD) ait kelimelerin tanınması hedeflenmiştir. Ancak TİD'ye ait genel, açık ve kapsamlı bir veri seti olmaması sebebiyle hedeflenen işaret dili tanıma yaklaşımının sınanması amacıyla, işitme engelli ve dili iyi bilen uzman işaretçilerin gerçekleştirdiği kelimelere ait işaretlerden oluşan yeni bir TİD veri setinin oluşturulması da tez kapsamında amaç olarak belirlenmiştir. Böylece bu alanda ileride yapılması muhtemel olan çalışmalarda ilgili veri setinin kullanımı mümkün olabilecek ve çalışmalar genişletilebilecektir.

İlgili TİD veri setindeki işaretlerin RGB görüntülerinin yanı sıra derinlik bilgilerini ifade eden görüntüler (derinlik haritaları) de kayda alınması planlanmış ve bu bağlamda gerekli işlemler yürütülmüştür. Çalışmada amaç olarak belirlenen bir diğer nokta ise derinlik bilgilerinin tanıma işleminde kullanılması ve işaret dillerinde derinlik bilgilerinin kullanılabilirliğinin sınanmasıdır.

Akademik olarak bu alanda yürütülen çalışmalar tezin ilerleme sürecinde tez çalışmalarına önemli katkılar sağlamıştır. Tez çalışması süresince yapılan araştırmalara, deneylere ve uygulamalara tezin ilerleyen bölümlerinde değinilmiştir. Yapılan çalışmaların tez kitabı olarak hazırlanması ve yayınlanabilecek bilimsel makaleler ile yeni çalışmalara ve fikirlere katkı sağlaması da bu tezin amaçlarından biridir.

1.3 Hipotez

İşaret dillerinin tanınabilmesi üzerine yürütülen tez çalışması sürecinde konu üzerine literatürde yer alan çalışmaların bir bölümü incelenmiştir. Çalışmalarda önerilen yöntemlerin birçoğu genelde belli bir işaret dilini tanımaya odaklı yaklaşımlardır. İşaret dillerini tanımak amacıyla yapılan çalışmalarda HMM gibi Markov modellerinin kullanımı oldukça yaygındır. Markov modellerinin işaretlerin zamandaki özellikleri oldukça iyi bir biçimde ifade edebilmesi bunun başlıca sebepleri arasındadır. Çünkü işaretin gerçekleştirilmesi süresince ellerin herhangi bir zaman anı ile bu andan önceki ve sonraki durumları birbirine olasılıksal olarak bağlıdır.

Bu tez çalışmasında ise alışlagelmiş yöntemlerin aksine hareketlere ait öznitelikleri ifade etmek amacıyla hareket farklarına dayalı toplamsal imge yaklaşımı tercih edilmektedir. Toplamsal imge yaklaşımında işaretlere ait görüntü çerçevelerinin farkları dikkate alınarak, bir işaretin bütün zamansal bilgilerini içeren tek bir toplamsal imge oluşturulur. Zaman uzayındaki öznitelikleri içeren bu toplamsal imgeler çeşitli dönüşüm yöntemleri ile uzamsal uzaya aktarılabilir ve işaret öznitelikleri basit bir şekilde elde edilebilir.

Genel olarak işaret dillerinin tanınabilmesi için görü tabanlı sistemlerde işareti gerçekleştiren el, kafa, yüz, kol gibi vücut öğelerinin imge dizilerinde bölütlenmesi ve takip edilmesi aşamaları vardır. Bu aşamalar tanıma probleminden önceki aşamalarda yer almakta ve imge dizileri gerçekleştirilmesi sistemlere zaman, karmaşıklık ve işlem açısından fazladan yük getirmektedir. Bilhassa gerçek zamanlı olarak tasarlanan sistemlerdeki işlem yükünü arttıran bu aşamalar ciddi manada zamanlama sorunu oluşturabilmektedir. Tez çalışmasında uygulanan yöntemle, imge dizileri bütün olarak işleme tabi tutularak herhangi bir nesne bölütleme ve nesne takip etme yaklaşımı kullanılmamıştır. Bu sayede tasarlanan sistemde ciddi anlamda bir zaman kazanımı ve basitlik sağlanmıştır.

Tez çalışması kapsamında oluşturulan TİD veri setindeki işaretlerin RGB görüntülerinin yanı sıra derinlik bilgilerini ifade eden görüntüler (derinlik haritaları) de kayda alınmıştır. Belirtilen öznitelik çıkarımı yönteminin derinlik haritalarındaki performansları ve RGB görüntülerle birlikte kullanımındaki performanslarının değerlendirilmesi tez çalışmasının amaçları arasında belirtilmiştir. Böylece derinlik bilgilerinin işaret dili tanımada etkin olup olmadığı da sınıanabilecektir.

Tanıma sistemlerinde özniteliklerin çıkarılması aşaması tanıma işlemi açısından önem arz etmektedir. Çıkarılan özniteliklerin makine öğrenmesi ve veri madenciliği yöntemleri kullanılarak verimli bir şekilde sınıflandırılabilmesi için özniteliklerin mümkün mertebede ayrılabilir olması gerekmektedir. Çalışmada kullanılan yöntem bu bağlamda zamansal öznitelikleri basit bir şekilde ve iyi bir biçimde ifade edebilmektedir. Bu sebeple çalışmada değerlendirilen veri setlerinde başarılı sonuçların gözlemlenebilmesi mümkün görünmektedir. Çalışmada uygulanan yöntemler gerek

basit olması gerekse de fazla işlem yükü gerektirmemesi sebebi ile bir işaret dili tanıma sistemi için oldukça kullanışlıdır.

Yukarıda verilmiş olan bilgiler ışığında tez çalışmaları yürütülmüştür.

İŞARET DİLLERİ VE İŞARET DİLİ TANIMA SİSTEMLERİ

2.1 İşaret Dilleri ve Genel Özellikleri

İşaret dilleri işitme engellilerin kendi aralarında iletişim kurmalarını sağlayan ve değişik görsel öğelerden oluşan dillerdir. İşaret dillerini oluşturan görsel öğeler ellerin hareketi, şekli, kafanın hareketi ve işaretlerin gerçekleştirilmesi esnasındaki yüz mimikleridir. Bu görsel öğelerin kullanımı; işaretlerin anlamlarına ve cümleler içerisindeki kullanımlarına göre değişiklik gösterebilmektedir. Öğelerin kullanımı aynı zamanda dilin biçimsel yapısına da bağlıdır.

İşaret dilleri tıpkı konuşma dilleri gibi kendine ait bir biçimsel yapıya sahiptir. İşaretler kelime anlamlarını ifade etmek için kullanılabilceği gibi cümleleri ifade etmek için de kullanılabilir. Bu bağlamda tek bir işaret bir kelimeyi ifade edebileceği gibi bir cümleyi de ifade edebilir. Bunun dışında bir anlama karşılık gelen işaret dili hareketi cümle içindeki kullanımına göre uygulama esnasında değişiklik gösterebilir. Bu durumlar genelde hareketlerin başlangıç ve bitiş noktalarında olmakla birlikte, dilin biçimsel yapısıyla da yakından ilişkilidir.

Kendine ait bir biçim yapısı ile işitme engellilerin ana dili olan işaret dilleri aynı zamanda bütün dünyadaki engelliler için aynı yapıda değildir [1], [2]. Farklı dilleri konuşan ülkelerin işaret dilleri de farklıdır. Hatta aynı dili konuşan ülkelerde bile işaret dilleri farklı olabilmektedir. Bu bağlamda işaret dillerinin ulusal bir özellik gösterdiği söylenebilir.

İşaret dillerinin bilimsel olarak belirlenmiş ana ilkeleri ise [2]'de belirtildiği üzere aşağıda belirtilmiştir.

- İşaret dilleri de sözlü diller gibi bir gramer yapısına sahiptir. Sanılanın aksine sözlü dillerden daha basit bir yapıda değildir.
- Her işaret dilinin kendine özgü gramer kuralları vardır ve her kavram için kullanılan işaretler de kullanıcılar arasında ortaktır. Bu özelliği ile işaret dili, konuşurken kullandığımız jestler ya da pantomimden çok farklıdır.
- Her ülkenin kendi işaret dili vardır. Örneğin Amerika'da kullanılan işaret dili (ASL) ile Almanya'da kullanılan işaret dili (DGS) birbirlerine benzemezler. Bu iki dil, İngilizce ile Almanca kadar farklıdır.
- Bir işaret dili çevrede kullanılan sözlü dilden etkilense de, farklı bir gramer yapısına sahiptir. Yani Türkçe ile Türk İşaret dili arasında mutlaka bir benzerlik olması gerekmez.
- Sözlü dillerde olduğu gibi işaret dili de erken yaşta öğrenilmelidir. İşitme engelliler 5-6 yaşına kadar işaret dili öğrenemezlerse daha sonra hem işaret dilini hem de başka dilleri öğrenmeleri zorlaşır.
- Beyin üzerindeki araştırmalara göre, sözel dillerle işaret dilleri aynı nörofizyolojik süreçlere ve aynı lokalizasyona (yani beynin sol yarımküresi) sahiptir.

İşaret dilleri işitme engellilerin iletişimde önemli yer tutar. Engelliler işaret dilinin yanı sıra yazarak da iletişim kurabilirler. Yazarak iletişim kurmak, konuşmayla veya işaret diliyle iletişim kurmaya nazaran oldukça yavaş kalmaktadır. Örneğin bir kişi bir dakikada konuşurken yaklaşık 150 kelime ifade edebilirken yazarken bu durum yaklaşık 15 kelime civarında olmaktadır. Bu durum ise engellilerin iletişimdeki motivasyonlarını olumsuz yönde etkileyebilmektedir. Yukarıda değinilen sebeplerden ötürü işaret dilleri, işaret dillerinin öğrenilmesi ve bunların bilgisayar sistemlerinde kullanımı önem arz etmektedir.

İlerleyen başlıklar ise altında işaret dili tanıma amacıyla geliştirilmekte olan sistemlerdeki teknik ve diller ile ilgili yapısal yaklaşımlara yer verilmiştir.

2.2 İşaret Dili Tanıma Sistemleri

İşaret dillerini tanımak amacıyla yapılan çalışmalar kullandıkları veri tiplerine göre; elektronik eldiven tabanlı ve görü tabanlı olmak üzere ikiye ayrılmaktadır. İşaret dilleri üzerine yapılan ilk çalışmalar elektronik eldiven tabanlı yaklaşımlardır. Elektronik eldiven tabanlı yaklaşımlarda genel olarak el ve parmaklara ait çeşitli konum ve açı gibi bilgiler değerlendirilmektedir. Görü tabanlı sistemlerde ise bir veya daha fazla kamera aracılığı ile alınan görüntüler üzerinde işaretlere ait değişik tipte öznitelikler çıkarılmakta ve işaretlerin tanınmasında kullanılmaktadır. Görü tabanlı sistemler üzerine son yıllarda önemli ölçüde çalışma ve araştırmalar yürütülmektedir. Özellikle görü tabanlı sistemleri kullanım kolaylığı, düşük karmaşıklık ve maliyetleri buna etki eden önemli faktörlerdendir. Görü tabanlı sistemler ve elektronik eldiven tabanlı sistemler birçok açıdan bunlara benzer farklılıklar sergilemektedir. Çizelge 2.1’de elektronik eldiven ve görü tabanlı sistemlere ait genel bilgilere karşılaştırmalı olarak yer verilmiştir.

Çizelge 2. 1 Elektronik eldiven ve görü tabanlı sistemlerin temel karakteristikleri [29]

	Elektronik Eldiven Tabanlı Sistemler	Görü Tabanlı Sistemler
Veri aygıtları	Sensörlü eldivenler, elektronik eldivenler	Kameralar
Veri tipi	Elektriksel sinyaller	Durağan ve ardışıl görüntüler
Veri saflığı	Gürültüden etkilenme oranı düşüktür	Ortama bağlı olarak düşük ve yüksek olabilmektedir
Kullanılabilirlik	Hareketleri kısıtlayabilir, her ortamda kullanılamayabilir	Hareketleri kısıtlamaz, görüntü alınan ortamlarda uygulanabilir
Kullanılan özellikler	Parmaklar ve ele ait dönme, hız, konum, hareket bilgileri	El ve parmakların şekli, hareketi, konumu
Veri ön işleme	Gürültü filtreleme	Nesne bölütleme, gürültü filtreleme, boyut indirgeme
Maliyet	Elektronik eldiven maliyeti (yüksek)	Görüntü alma aygıtı maliyeti (eldivene nazaran düşük)

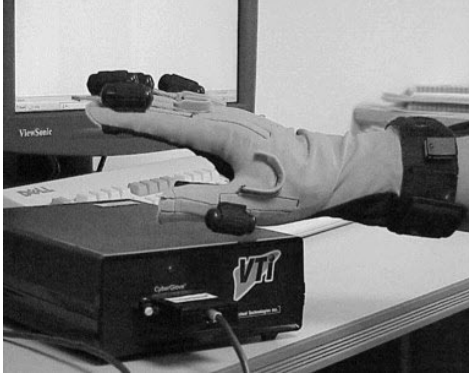
İşaret dillerinde yer alan işaretler durağan (postures) ve durağan olmayan (gestures) hareketlerden meydana gelmektedir. Durağan işaretler tek bir imgeyle temsil edilmektedir. Özellikle rakamlar ve harfler bu kategoriye dâhil olmaktadır. Durağan işaretler tek bir imge temsil edilirken, durağan olmayan işaretler ise bir zaman aralığında eller ile kolların konumlarındaki ve şekillerindeki değişimlerden meydana gelmektedir. Durağan olmayan işaretler, işaret dillerinde bir kelimeyi ifade edebileceği gibi bir cümleyi de ifade edebilir. Ayrıca ayrıık işaretlerin birleşimiyle cümleler oluşturulabilir. Bu sebeple hareket tabanlı olarak yapılan çalışmalarda kelime veya cümle tanıma yönelik yaklaşımlar mevcuttur. Bu bilgiler dâhilinde işaret dili tanıma sistemleri genel olarak şekil 2.1'deki gibi gruplandırılabilir.



Şekil 2. 1 İşaret dili tanıma sistemlerinin genel hiyerarşik yapısı

2.2.1 Elektronik Eldiven Tabanlı Sistemler

İşaret dillerinin tanınması amacıyla yapılan ilk çalışmalarda elektronik eldiven tabanlı sistemlerdir. Bu tarz sistemlerde işaretleri gerçekleştirecek olan işaretçi, işaretleri gerçekleştirdiği ellerine giydiği elektronik eldivenler ile hareketleri gerçekleştirmektedir. İşaretlerin gerçekleştirilmesi esnasında eldivenlerden alınan dönme, hız, açı gibi veriler kullanılarak tanıma işlemi gerçekleştirilmeye çalışılmaktadır. Şekil 2.2'de bir elektronik eldiven örneği gösterilmiştir.



Şekil 2. 2 Elektronik eldiven (cyberglove with 18 sensors) [45]

Elektronik eldiven tabanlı sistemlerde verinin güvenilirliği ve saflığı yani gürültüden etkilenmemesi oldukça yüksek olmaktadır. Ancak eldivenlerin sistemlere getirdiği bazı zorluklar da mevcuttur. Örneğin işaretleri gerçekleştirecek olan işaretçilerin bu eldivenleri giymesi gerekmekte ve kullanımına aşına olması gerekmektedir. Ayrıca eldivenlerin bir takım parçalarından dolayı işaretçileri kısıtlayarak hareketleri gerçekleştirmede zorluk çıkarması da mümkündür. Bu eksi yönlerle beraber bu tarz sistemler çok kullanışlı da değildir.

2.2.2 Görü Tabanlı Sistemler

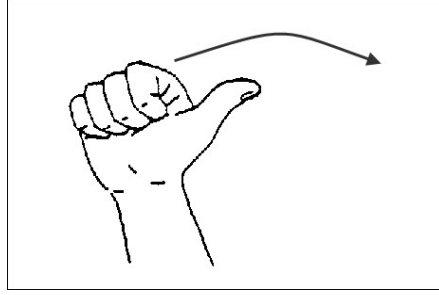
İşaret dili tanıma üzerine yapılan çalışmalarda 1990'lı yılların ortalarında itibaren görü tabanlı sistemler de kullanılmaya başlanmıştır. Görü tabanlı sistemler kullanım kolaylığı ve maliyet açısından diğer yaklaşımlara nazaran avantajlı olmakla birlikte veri alınması, işlenmesi gibi konularda problemler çıkarabilmektedir. Bu problemler verinin alındığı ortamın aydınlatma koşulları gibi bilgisayarla görmenin temel problemlerinin yanı sıra görüntülerde hareketi oluşturan öğelerin bölütlenmesi ve modellenmesi de olabilmektedir. Ancak görü tabanlı sistemlerin kullanılabilirlik ve uygulanabilirliğinin yüksek olması tercih oranını arttırmaktadır. İşaret dillerinin görü tabanlı sistemlerde kullanımı ve tanınması üzerine yapılan çalışmalar; hareket tabanlı sistemler ve şekil tabanlı sistemler olmak üzere iki alt ana başlıkta toplanabilir.

2.2.2.1 Hareket Tabanlı Sistemler

İşaret dillerinde bir kelimeyi veya kelimelerden bir cümleyi ifade etmek amacıyla durağan olmayan işaretler (gestures) kullanılır. İşaretler belli bir zaman aralığında

ellerin, kolların konum ve şekillerinde meydana gelen değişimlerle oluşur. Bu sebeple işaretin her bir zaman anındaki durumu bir imge olarak düşünülürse, hareketler bu imgelerin birleşiminden meydana gelen bir video çerçeve dizisi olarak tanımlanır.

İşaretlerin gerçekleştirilmesinde tek elin kullanıldığı ve iki elin kullanıldığı işaretler mevcuttur. Hareket tabanlı tanıma sistemlerinde kelimeleri veya cümleleri ifade eden durağan olmayan işaretler kullanılabilir. Şekil 2.3'te bir kelimeye ait durağan olmayan işaretin yapılışı ifade edilmiştir.

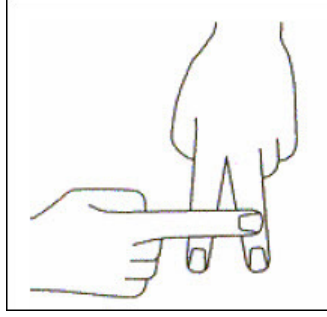


Şekil 2. 3 Durağan olmayan (hareketli) işaret örneği (tid'de "dün" kelimesi) [2]

İşaret dillerinde durağan olmayan işaretlerin yanı sıra durağan işaretler de kullanılmaktadır. İlerleyen başlık altında durağan işaretleri temel alan sistemlere ait bilgilere yer verilmiştir.

2.2.2.2 Şekil Tabanlı Sistemler

İşaret dillerinde durağan olmayan işaretlerin yanı sıra, durağan olmayan işaretlere (postures) nazaran az olmasına rağmen, durağan işaretler de yer almaktadır. Durağan işaretler durağan olmayan işaretlerin aksine tek bir imge ile ifade edilir. Özellikle işaret dillerinde harf ve sayılar durağan işaretler ile ifade edilir. Durağan işaretlerin kullanımı ile sayıların ifadesi söz konusu olabildiği gibi, işaret dillerinde karşılıkları olmayan bazı kişi, yer veya varlık isimlerini ifade etmek için de kullanılırlar. Parmak heceleme olarak tanımlanan bu işlemde, ifade edilecek kelimeye ait harfler sırasıyla durağan işaretler olarak gerçekleştirilir ve kelime harf harf hecelenerek ifade edilir. Şekil 2.4'te durağan bir işaret olan bir harfin ifade edilişi gösterilmiştir.



Şekil 2. 4 Durağan işaret örneği (tid'de "a" harfi) [2]

Farklı ülkelere ait işaret dillerinde, durağan olmayan işaretlerin farklı olması ile birlikte durağan işaretlerde de farklılıklar gözlenmektedir. Bu alanda yapılan çalışmalar ve tasarlanan sistemler özel olarak durağan işaretleri tanımayı hedeflemektedir.

İŞARET DİLİ VERİSETLERİ

Tez çalışması kapsamında hedeflenen işaret dili tanıma sisteminin başarımlarını sınamak ve sistem performansının değerlendirilmesi amacıyla 2 adet veri seti kullanılmıştır. Bu veri setlerinden ilki Amerikan İşaret Dili'ne (ASL) ait olmakla birlikte, tez çalışmasında kullanılan diğer işaret dili veri seti Türk İşaret Dili'ne aittir. Bu veri setlerine ait detaylı bilgiler bölümün ilerleyen alt başlıklarında verilmiştir.

3.1 Amerikan İşaret Dili (ASL) Veri Seti

İşaret dilleri tanıma problemlerinde üzerine en çok araştırma ve çalışma yapılan işaret dillerinden birisi de Amerikan İşaret Dili'dir. Görü tabanlı yaklaşımlar kullanılarak ilk yapılan çalışmalarda da yine ASL'ye ait kelime grupları kullanılmıştır. Yapılan tez çalışmasında, sistemin uygulanabilirliğini ve başarımlarını gözlemek adına Türk İşaret Dili dışında, ASL'ye ait kelimelerden oluşan bir veri seti [46] de kullanılmıştır. İşaretlere ait orijinal görüntü çerçeve boyutları 640x480 olup; görüntüler RGB renk uzayında bulunmaktadır. İlgili veri setinde ASL'ye ait 20 kelime bulunmaktadır. Her kelime 4 işaretçi tarafından 10'ar defa gerçekleştirilmiştir ve veri setinde toplamda 20 kelimeye karşılık gelen 800 işaret video görüntüsü yer almaktadır. Veri setinde yer alan kelimeler ve Türkçe karşılıkları çizelge 3.1'de gösterilmiştir.

Çizelge 3. 1 Asl veri setinde yer alan kelimeler ve türkçe karşılıkları

İNGİLİZCE KELİME	TÜRKÇE ANLAMI	İNGİLİZCE KELİME	TÜRKÇE ANLAMI
AGAINST	KARŞI	EXPLORE	KEŞFETMEK
AIM	AMAÇ	HUNT	AVLAMAK
BALLOON	BALON	KNIFE	BIÇAK
BANDIT	HAYDUT	RELAY	YERİNİ ALMAK
CAKE	KEK	REVERSE	TERS
CHAIR	SANDALYE	ROLE	ROL
COMPUTER	BİLGİSAYAR	SAD	ÜZGÜN
CONCENTRATE	KONSANTRE OLMAK	STUDENT	ÖĞRENCİ
CROSS	KARŞIYA GEÇMEK	TABLE	MASA
DEAF	SAĞIR	TELL	ANLATMAK

Şekil 3.1’de ASL veri setinde “bilgisayar” kelimesinin bir işaretçi tarafından ifade edilmesine ait seçilmiş ardışıl çerçeveler gösterilmiştir. Şekil 3.1’de ardışıl çerçeveler soldan sağa ve yukarıdan aşağıya doğru olacak biçimde sıralanarak ifade edilmiştir.



Şekil 3. 1 Asl veri setinde “computer” kelimesinin bir işaretçi tarafından ifade edilmesine ait seçilmiş ardışıl çerçeveler

3.2 Türk İşaret Dili (TİD) Veri Seti

Tez çalışması kapsamında diğer işaret dillerine ait mevcut veri setlerinin kullanımıyla beraber, Türk İşaret Dili'ne ait kelimelerin tanınması da hedeflenmiştir. Türk İşaret Dili üzerine, Türk Dil Kurumu ve ülkemizdeki çeşitli üniversitelerdeki çalışma grupları tarafından yapılmış çalışmalar vardır. Bu çalışmalarda Türk İşaret Dili'ne ait birtakım işaret sözlükleri oluşturulmuştur. İşaret sözlükleri genel olarak işaretlerin gerçekleştirilmesine yönelik görsel ve yazılı bilgileri içermektedir ancak Türk İşaret Dili'ne ait kelimeleri tanımak ve araştırma yapmak amacıyla büyük veya küçük çapta bir veri seti bulunmamaktadır. Bu amaçla tez sürecinde öncelikle Türk İşaret Dili'ne ait bir veri seti oluşturulması planlanmış ve bu bağlamda gerekli çalışmalar yürütülmüştür.

İşaret dilleri, konuşma dilleri gibi, kendine ait kuralları ve ifade biçimleri olan dillerdir. Bu sebeple işaretleri gerçekleştirecek olan kişilerin de bu konuda tecrübeli olmaları ve dillere hâkim olmaları gerekmektedir. Bu sebeple Türk İşitme Engelliler Federasyonu vasıtasıyla bu konuda uzman olan işaretçilerle görüşülerek, veri setinde yer alan kelimelere ait işaretleri gerçekleştirecek olan işitme ve konuşma engelli uzmanların tez çalışmalarına çok değerli katkıları sağlanmıştır.

Türk İşaret Dili'ne ait oluşturulan yeni veri setinde yer alan kelimeler 3 farklı kategoride toplanmıştır. Kelimeler 3 farklı işaretçi tarafından genelde 3 tekrar ile video görüntü olarak kayda alınmıştır ve çalışmada kullanılacak olan veri seti oluşturulmuştur. İlgili veri setinde kelime kategorilerini ve her bir kategoride yer alan kelime sayılarını gösteren bir tablo çizelge 3.2'de sunulmuştur.

Çizelge 3. 2 Tid veri setinde kelime kategorileri ve kategorilerdeki kelime sayıları

KATEGORİ ADI	KELİME SAYISI
GÜNLÜK KONUŞMA TERİMLERİ	22
ZAMAN TERİMLERİ	58
MESLEKLER	31

TİD veri setinde, çizelge 3.2'de ifade edilen kelime kategorilerinde, toplam 111 kelimeye ait işaret görüntüleri kaydedilmiştir ve tez çalışmasında öncelikli olarak kullanılmıştır. Çizelge 3.2'de belirtilen "günlük konuşma" terimleri, "zaman" terimleri

ve “meslekler” kategorilerinde yer alan kelimeler sırasıyla çizelge 3.3’te, çizelge 3.4’te ve çizelge 3.5’te sunulmuştur.

Çizelge 3. 3 Günlük konuşma terimleri kategorisinde yer alan kelimeler

AFFEDERSİNİZ	İYİ AKŞAMLAR	MERHABA	NE ZAMAN?
AFİYET OLSUN	İYİ GECELER	NASIL?	ÖZÜR DİLERİM
GÖRÜŞMEK ÜZERE	İYİ MİSİN?	NE?	SAĞOL
GÖRÜŞÜRÜZ	İYİYİM	NEDEN?	TEŞEKKÜRLER
GÜLE GÜLE	KİM?	NEREDE?	
GÜNAYDIN	LÜTFEN	NEREYE?	

Çizelge 3. 4 Zaman terimleri kategorisinde yer alan kelimeler

AĞUSTOS	ERKEN	HAFTASONU	SAAT
AKŞAM	EYLÜL	HAZİRAN	SABAH
ARALIK	GEÇ	HER GÜN	SALI
AY	GECE	HER ZAMAN	ŞİMDİ
BAZEN	GEÇEN AY	İLKBAHAR	SONBAHAR
BU AY	GEÇEN HAFTA	KIŞ	SONRA
BUGÜN	GEÇEN YIL	MART	ŞUBAT
BU HAFTA	GECE YARISI	MAYIS	TEMMUZ
BU YIL	GEÇMİŞ	NİSAN	YARIN
ÇARŞAMBA	GELECEK	OCAK	YAZ
CUMA	GELECEK AY	ÖĞLEN	YIL
CUMARTESİ	GELECEK HAFTA	ÖNCE	YÜZYIL
DAKİKA	GELECEK YIL	PAZAR	ZAMAN
DÜN	GÜN	PAZARTESİ	
EKİM	HAFTA	PERŞEMBE	

Çizelge 3.4'te yer alan kelimelerin bir bölümünün belli kısımları ifade edilişleri bakımından birbiriyle aynıdır veya çok büyük oranda birbirine benzemektedir. Bu durumun bu grupta yer alan kelimelerin tanıma aşamasında birbirleriyle karıştırılmasına sebep olması ve başarıma azalan yönde etki yapması olasıdır. Bu tarz durumlar tez çalışmasının deneysel sonuçlar bölümünde analiz edilmiş ve gözlemlenen sonuçlar yine deneysel sonuçlar bölümünde ifade edilmiştir.

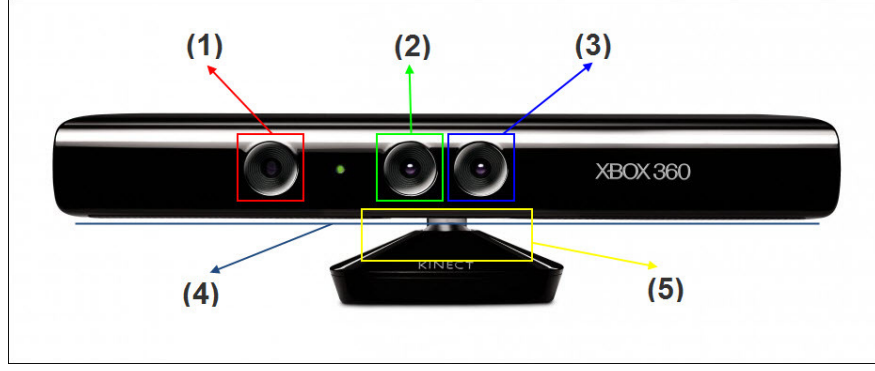
Çizelge 3.5'te ise "meslekler" kategorisinde yer alan kelimeler listelenmiştir.

Çizelge 3. 5 Meslekler kategorisinde yer alan kelimeler

BANKACI	FOTOĞRAFÇI	KUYUMCU	POSTACI
BAŞBAKAN	GAZETECİ	MATBAACI	PSİKOLOG
BERBER	HABER MUHABİRİ	MİLLETVEKİLİ	SAVCI
BOYACI	HİZMETLİ	MÜHENDİS	TAMİRCİ
CUMHURBAŞKANI	KAMERAMAN	NOTER	TERZİ
DİŞÇİ	KAPTAN	ÖĞRENCİ	VALİ
ELEKTRİKÇİ	KAYMAKAM	ÖĞRETMEN	VETERİNER
FIRINCI	KİMYAGER	POLİS	

İşaret dillerinin görü tabanlı yaklaşımla, gerçek zamanlı olarak veya gerçek zamanlı olmayarak tanınmasında genellikle CCD kameralar kullanılmaktadır. Bu kameralar farklı çözünürlüklerde çekim yapabilmekle beraber RGB olarak renkli görüntü sağlamaktadırlar.

Tez çalışmasında ise görüntüleri kayda almak amacıyla klasik CCD kameralar yerine Microsoft'un XBOX 360 oyun konsolu için tasarladığı ve ürettiği Kinect algılayıcı kullanılmıştır. Kinect algılayıcıya ait bir görüntü şekil 3.2'de gösterilmiştir.

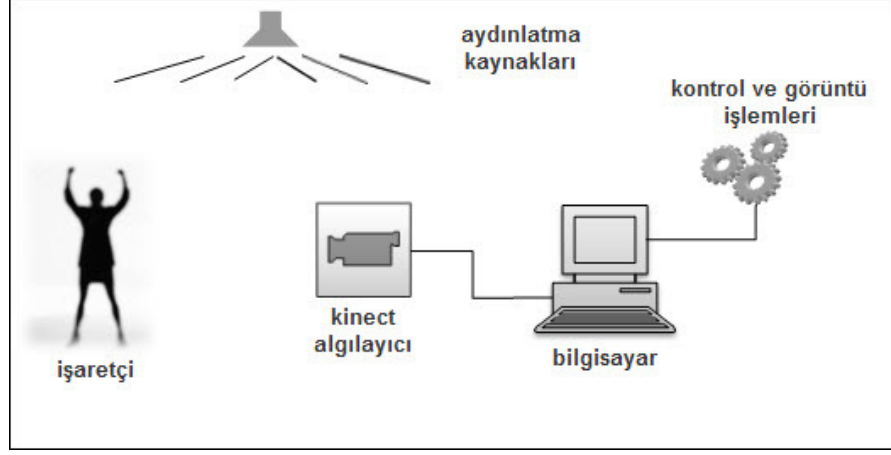


Şekil 3. 2 Kinect algılayıcı

Türk İşaret Dili'ne ait veri setindeki kelimelere ait işaret görüntülerini kayıt etmek amacıyla kullanılan Kinect algılayıcı üzerinde yer alan bir RGB kamera (2) ile renkli olarak kayıt yapabilmektedir. Bununla beraber şekil 3.2'de (1) ile gösterilen cihaz lazer yayılım yaparken, (3) ile gösterilen kızıl ötesi algılayıcı bu ışınların gidiş ve geliş sürelerini hesaplayarak kameranın görüş açısı dâhilindeki noktaların kameraya olan mesafelerini hesaplamaktadır. Bu mesafe bilgileri tek bir imgede toplanarak RGB görüntüye karşılık gelen bir derinlik haritası oluşturulmaktadır. Ayrıca Kinect algılayıcı üzerinde ses algılamayı sağlayan mikrofon dizisi (4) ile birlikte cihazın aşağı ve yukarı hareket etmesini sağlayan bir motor (5) da yer almaktadır [47].

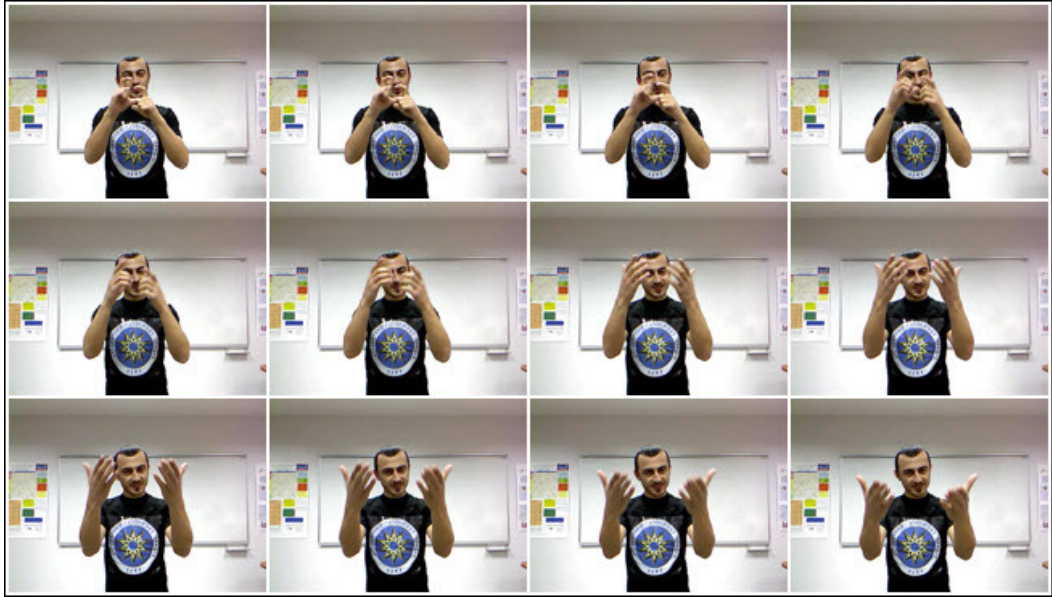
Çalışma kapsamında kullanılacak olan işaret görüntülerine ait RGB renk uzayındaki görüntülerle beraber bu görüntülere karşılık düşen derinlik haritaları da kaydedilmiştir. Bu sayede Türk İşaret Dili'ne ait kelimelerin tanınmasında derinlik bilgisinin de kullanılması hedeflenmiştir.

Veri setindeki görüntüler normal ışıklandırma altında çekilmiştir. Özellikle ışığa bağlı gürültü ve olumsuz etkileri azaltmak amacıyla ışığın çekim yapılan ortamdaki yayılımının eşit olması sağlanmıştır. İşaretçilerin Kinect algılayıcıya olan uzaklıkları sabitlenerek çekimler yapılmıştır. Kinect algılayıcı aracılığıyla yapılan çekimlerde görüntü verilerine erişim sağlamak ve bunları kayıt altına almak için ise Kinect'in bağlı olduğu donanımsal açıdan üst seviye bir bilgisayar ile bu bilgisayarda çalışan çeşitli yazılımlar kullanılmıştır. Şekil 3.3'te bu süreci betimleyen bir sistem şeması ifade edilmiştir.



Şekil 3.3 İşaret görüntülerinin kaydında sistem yapısı

Kinect algılayıcı kullanılarak, şekil 3.3'te ifade edilen sistemde, işaretçilere ait RGB görüntüler ve bu görüntülere karşılık düşen derinlik haritaları kaydedilmiştir. Şekil 3.4'te, oluşturulan Türk İşaret Dili'ne ait veri setindeki bir kelimeye ait örnek RGB sıralı çerçeveleri gösterilmiştir.



Şekil 3.4 Tid veri setinde "günaydın" kelimesine ait örnek rgb ardışıl çerçeveleri

Şekil 3.4'te ifade edilen ardışıl RGB çerçevelerine birebir karşılık gelen ardışıl derinlik haritaları ise şekil 3.5'te sunulmuştur.



Şekil 3. 5 Tid veri setinde “günaydın” kelimesine ait örnek derinlik haritası ardışıl çerçeveleri

Şekil 3.5’te yer alan derinlik haritası çerçevelerinde siyah bölgeler kameraya uzaklıkları yakın olan bölgeleri temsil ederken, grinin tonlarındaki bölgeler ise kameraya mesafeleri uzak olan bölgelerdir.

Bununla birlikte Kinect algılayıcı, kızıl ötesi algılayıcısı ile mesafesini hesaplayamadığı noktaların uzaklığını 0 (sıfır) olarak göstermektedir. Bu durum daha çok nesnelerin açılal konumlarına veya yapıldıkları maddelerin yansıtma özelliğinden kaynaklanmaktadır. Ayrıca çok fazla detay içeren nesnelere de bu durumun görülmesi söz konusudur. Bu durumlara özel olarak, kayıp derinlik bilgilerinin hesaplanması amacıyla literatürde hâlihazırda yapılmış çalışmalar mevcuttur. Tez çalışması kapsamında ise bu özel durum değerlendirmeye alınmamış olup, öncelikle Türk İşaret Dili’nde etkin tanıma ile derinlik bilgisinin kullanımının tanımadaki performansına odaklanılmıştır.

ÖNERİLEN İŞARET DİLİ TANIMA SİSTEMİ

İşaret dillerine ait kelimelerin bilgisayar sistemleri tarafından tanınması amacıyla yürütülen tez çalışmasında ilk aşamada literatürde yapılan çalışmalar incelenerek mevcut yaklaşımlar ve bu alandaki çalışmalar hakkında bilgiler toplanmıştır. Yapılmış olan çalışmalar hakkındaki bilgiler 1. bölümdeki “literatür özeti” alt başlığı altında detaylandırılmıştır.

İşaret dillerinin tanımak amacıyla yapılan çalışmalarda Markov modellerini temel alan çalışmalar çoğunluktadır. Bunun sebebi, Markov modellerinin işaretlerin zamandaki değişimini, yani işaretlerin herhangi bir zaman anından önceki ve sonraki durumlarını iyi bir şekilde ifade edebilmesi ve modelleyebilmesidir.

Tez kapsamında yapılan ve önerilen işaret dili tanıma sisteminde ise işaretlerin zamandaki değişimlerini yani zaman uzayındaki özelliklerini ifade etmek amacıyla basit ve maliyeti düşük bir yaklaşım olan toplamsal imge yaklaşımı kullanılmıştır. Toplamsal imge yaklaşımı, bir işaretin gerçekleştiği bütün çerçeveler arasındaki farkların toplamına dayanmaktadır. Bu sayede bir işaretin zaman uzayındaki değişimleri tek bir imgede temsil edilebilmiştir.

İşaretlere ait özniteliklerin çıkarımında iki aşamalı bir yapı kullanılmıştır. İlk aşama olan, toplamsal imgelerin elde edilmesiyle zamansal özniteliklerin çıkarılmasıdır. Toplamsal imgeler zaman özniteliklerini kuvvetli olarak içeren imgelerdir. İkinci aşamada ise zamansal öznitelikleri uzamsal uzaya aktarmak için DCT gibi imge dönüşüm yöntemlerinden faydalanılarak öznitelikler elde edilmiştir.

Önerilen işaret dili tanıma sistemi, genel yapısı itibariyle aşağıdaki başlıklar altında sıralanabilir;

- Görüntülerde ön-işleme uygulanması
- Özniteliklerin çıkarılması
 - Zamansal özniteliklerin çıkarılması
 - Uzamsal özniteliklerin çıkarılması
- Öznitelik vektörlerinin oluşturulması
- İşaret örneklerinin tanınması

İşaretlere ait özniteliklerin çıkarımı aşamasından önce video görüntüler üzerinde çeşitli ön-işleme işlemleri uygulanmıştır. Bu işlemleri takiben ise öznitelik çıkarımı aşamasına geçilmiş ve iki kademeli bir öznitelik çıkarım süreci yürütülerek öznitelikler elde edilmiştir. Öznitelik vektörlerinin oluşturulmasıyla beraber, eğitici makine öğrenmesi yöntemleriyle işaret örneklerinin tanınması sağlanmıştır. Sistemin sürecinde yer alan bu temel adımlara ait detaylı bilgiler ve işlemler bölümün ilerleyen başlıkları altında verilmiştir.

4.1 Görüntü Ön-işleme

İmge ve video görüntüler üzerinde yapılan ön-işleme adımları genellikle gürültülerin süzgeçlenerek verilerin temizlenmesi amacıyla yapılır. Ön-işlemler bu gibi sebeplerden dolayı görüntü işleme yaklaşımlarında önemli bir yer tutar ve sıklıkla tercih edilir.

İlgili tez çalışmasının verilerin toplanması aşamasından sonraki ve özniteliklerin çıkarımı aşamasından önceki aşamada imge ön-işleme işlemleri uygulanmıştır. Bu aşamada iki adımlı bir süreç yürütülmüştür;

- Video çerçeve boyutlarının ölçekleme ile yeniden boyutlandırılması
- Video görüntülerin renk uzayının değiştirilmesi

Öncelikle Türk İşaret Dili'ne ait veri setinde RGB görüntüler ve derinlik haritaları üzerinde öncelikle ölçekleme işlemi uygulanarak, boyutların küçültülmesi sağlanmıştır. Orijinal boyutları 640x480 olan video görüntülerine ait çerçeveler sırasıyla $\frac{1}{2}$ ve $\frac{1}{4}$

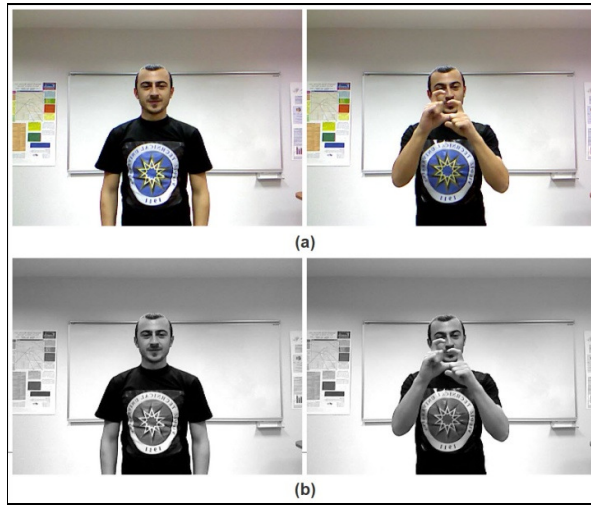
oranlarında ölçeklenerek ve belirli bir bilgi kaybı göze alınarak, yapılandırılmış boyutları 320x240 ile 160x120 olan yeni görüntüler oluşturulmuştur.

Ön-işleme adımlarında uygulanan çerçeve boyutlarının yeniden yapılandırılması işlemi ile farklı çerçeve boyutlarının tanımadaki performansları gözlemlenmek istenmiştir. Bu yaklaşımla bahsedilen tanıma başarımlarına ise deneysel sonuçlar bölümünde detaylıca yer verilmiştir.

İkinci aşamada ise uygulanan video görüntülerin renk uzayının değiştirilmesi işlemi gerçekleştirilmiştir. RGB renk uzayında yer alan görüntüler gri-ölçeğe geçirilmiştir ve görüntülerdeki 3 renk kanalı tek bir kanala indirgenmiştir. Dönüşümü sağlayan işlem eşitlik (4.1)'de gösterilmiştir.

$$G_{sc} = [R \quad G \quad B] \cdot \begin{bmatrix} 0.2989 \\ 0.5870 \\ 0.1140 \end{bmatrix} \quad (4.1)$$

Eşitlik (4.1) RGB değerlerini, gri-ölçeğe dönüştürürken R, G, B renk kanallarının ağırlıklı toplamını almaktadır. Şekil 4.1'de renk uzayı dönüşümleri yapılan örnek çerçeveler gösterilmiştir.



Şekil 4. 1 (a) Rgb renk uzayında yer alan örnek çerçeveler (b) gri-ölçeğe dönüştürülmüş çerçeveler.

Bu bölümde anlatılan ön-işleme adımlarının devamında öznitelik çıkarımı sürecine geçilmiştir. İşaretlere ait özniteliklerin çıkarımı amacıyla izlenen yol ve detaylı işlemler öznitelik çıkarımı başlığı altında etraflıca anlatılmıştır.

4.2 Öznitelik Çıkarımı

Yürütülen tez çalışmasındaki işaret dili tanıma sisteminde, özniteliklerin çıkarımı amacıyla iki aşamalı bir yol izlenmiştir. İlk aşamada işaretlere ait zamana bağlı öznitelikleri çıkarmak amacıyla hareket farkları analizine dayalı toplamsal imge yaklaşımı kullanılmıştır. Bu aşamayı takiben zamansal uzaydaki öznitelikler, DCT gibi çeşitli dönüşüm yöntemleri ile uzamsal uzaya aktarılmıştır ve öznitelikler elde edilmiştir. Yapılan bu işlemler RGB uzaydan gri-ölçeğe dönüştürülen video görüntüler ile derinlik haritalarında ayrı ayrı yapılmıştır. Öznitelik çıkarımında uygulanan bu yöntemlere ait detaylı bilgilere, sırasıyla zamansal özniteliklerin çıkarımı ve uzamsal özniteliklerin çıkarımı başlıkları altında yer verilmiştir.

4.2.1 Zamansal Özniteliklerin Çıkarımı

İşaret dillerinde yer alan işaretlerin büyük çoğunluğu ellerin ve kolların zamandaki yer ve şekil değişimleri sonucu meydana gelir. Bir zaman aralığı boyunca meydana gelen bütün değişimler aslında işaretleri tanımlar. Bu sebeple ellerin ve kolların zamandaki yer ve şekil değişimleri, işaretleri birbirinden ayıracak önemli bilgileri içerir.

Bu yaklaşım ışığında, bir işaret video görüntüsüne ait ardışıl video çerçeveleri arasındaki değişimler (farklar) bir hareketi temsil etmek amacıyla kullanılabilir. Tez çalışması kapsamında işaretlere ait zamansal öznitelikleri elde etmek için de bu temel yaklaşımdan faydalanılmıştır.

Ardışıl video çerçevelerindeki değişimlere dayalı hatalar, video sıkıştırmada kullanılan ara bir yöntemdir. Ardışıl video görüntülerde meydana gelen değişimler hesaplanarak bütün çerçeve bilgilerinin saklanması yerine, sadece gerekli değişimler saklanarak videoların bu temel yaklaşımla sıkıştırılması sağlanır.

Video sıkıştırmada bir ara yöntem olarak kullanılan bu yaklaşım, tez çalışmasında işaretlere ait zamansal öznitelikleri elde etmek amacıyla kullanılmıştır. İlk aşamada, bir işaret videosuna ait tüm ardışıl çerçevelerin farkları bulunmuştur. Fark imgeleri olarak da tanımlanabilecek olan bu çerçeve farkları eşitlik (4.2)'deki denklem yardımıyla ifade edilmiştir.

$$I_{Fi} = I_{i+1} - I_i ; \quad (i=1, \dots, n-1) \quad (4.2)$$

Eşitlik (4.2)'de yer alan denklemde I_i ve I_{i+1} ardışıl video çerçevelerini, I_{Fi} ise ardışıl çerçevelerin farklarını temsil etmektedir. Bu işlemle bir işaret videosunda yer alan tüm ardışıl çerçevelerin farkları hesaplanır. Şekil 4.2'de Türk İşaret Dili'nde "günaydın" kelimesine ait seçilmiş ardışıl gri-ölçek çerçevelerle ve bunlardan elde edilen fark imgeleri sıralı olarak gösterilmiştir.



Şekil 4. 2 Tid'de "günaydın" kelimesinin bir işaretçi tarafından uygulanışına ait seçilmiş gri-imege ardışıl çerçeveleri ve bu çerçevelere ait fark imgeleri.

Şekil 4.2'de 8 adet ardışıl çerçeveye karşılık düşen 7 adet fark imgesi gösterilmiştir. Fark imgelerini net bir biçimde gösterebilmek için, parlaklık değerlerinin mutlak değer karşılıkları alınmış ve bu değerler 0-255 renk aralığına normalleştirilmiştir. Aynı işlem şekil 4.3'te derinlik haritalarına ait fark imgeleri için de uygulanmıştır. Şekil 4.2'deki işarete ait derinlik haritalarındaki fark imgeleri de şekil 4.3'te sunulmuştur.



Şekil 4. 3 Tid’de “günaydın” kelimesinin bir işaretçi tarafından uygulanaşına ait seçilmiş derinlik haritası ardışıl çerçevesleri ve bu çerçeveslere ait fark imgeleri

Şekil 4.2’deki ve şekil 4.3’teki fark imgelerinin elde edilmesinden sonraki aşamada, bu fark imgeleri ikili eşikleme yöntemi ile ikili imgelere dönüştürülmüştür. Bu işleme ait denklem eşitlik 4.3’te ifade edilmiştir.

$$I_{Ei} = \begin{cases} 1 & |I_{Fi}| \geq T \\ 0 & \text{değilse} \end{cases} \quad (4.3)$$

Eşitlik (4.3)’te yer alan denklemde, I_{Fi} fark imgelerine, T eşikleme katsayısına, I_{Ei} ise eşiklenmiş ikili imgeye karşılık gelmektedir. İkili eşikleme ile seçilen T katsayısına bağlı olarak belirli bir değerin altında olan değışimler süzgeçlenmekte ve yüksek değışimler değeriendirilmeye alınmaktadır. Bu aşamada T eşikleme katsayısının alabileceğı değere bağlı olarak I_{Ei} ikili imgeleri değışebilir. Yapılan çalışmada bu katsayı deneysel gözlemlere bağlı olarak, gri-imgeler ve derinlik haritaları için ayrı ayrı belirlenerek kullanılmıştır. T katsayısı gri-imgelerde 60, derinlik haritalarında ise 10 olarak seçilmiştir.

Yine bu katsayının gözlemden bağımsız olarak belirlenmesini ve sistemin bu aşamadaki otomatikliğini sağlamak amacıyla fark değerleri ortalamasına dayalı bir yaklaşım da kullanılmıştır. Bu yaklaşımda, her bir çerçeve fark imgesine ait sıfırdan farklı değerlerin (değişim meydana gelen) mutlak değer toplamları alınarak, sıfırdan farklı piksel sayısına bölünmüştür. Böylece eşik katsayısı için gözleme dayalı sabit değerler yerine, değişime bağlı dinamik katsayılar da kullanılmıştır. Bu katsayılar bir video görüntüde yer alan her ardışıl iki çerçeve farkına özel olarak hesaplanır ve ilgili katsayı sadece ilgili çerçeve farkında kullanılır. Bu işlemi betimleyen bir ifade eşitlik (4.4)'te gösterilmiştir.

$$T_i = \frac{1}{C_{NZ}} \sum_{j=1}^M \sum_{k=1}^N |I_{Fi}(j, k)| \quad I_{Fi}(m, n) \neq 0 \quad (4.4)$$

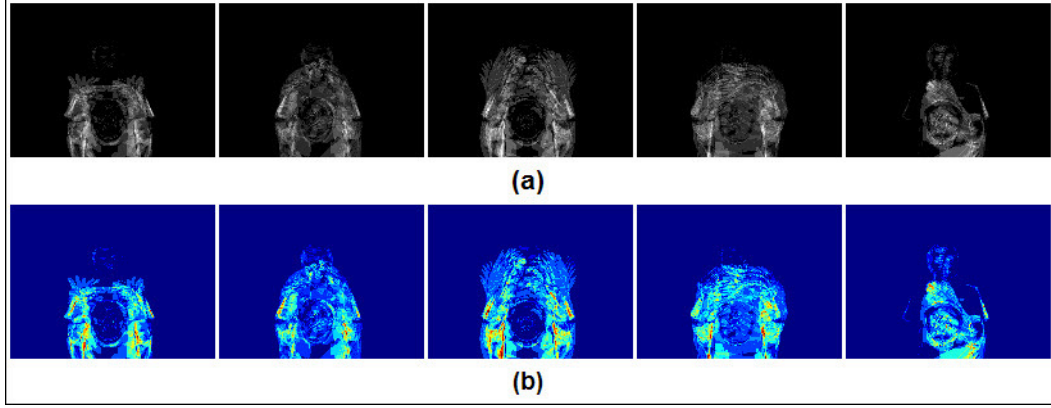
Eşitlik (4.3)'te M ve N çerçeve boyutları olmak üzere C_{NZ} sıfırdan farklı olan toplam piksel sayısıdır.

Eşikleme işlemlerini takiben, zamansal öznitelikleri ifade etmek amacıyla, eşikleme işlemi uygulanmış ikili fark imgeleri toplanmıştır. Bu işlem sonucunda hareket değişimlerinin bütününe içeren tek bir imge elde edilmiştir. Toplamsal imge olarak adlandırılan bu imgenin elde edilmesine ait denklem eşitlik (4.5)'te sunulmuştur.

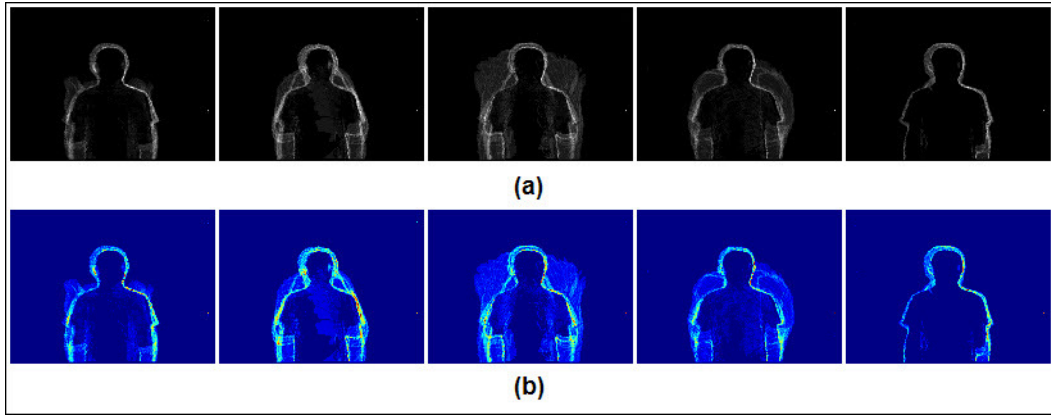
$$I_T = \sum_{i=1}^{n-1} I_{Ei} \cdot w_i \quad w_i = \frac{1}{n} \quad (4.5)$$

Eşitlik (4.5)'te n toplam çerçeve sayısı olmak üzere, w_i eşiklenmiş fark çerçevelerinin ağırlıkları ve I_T ise toplamsal imgedir. w_i ağırlık değerleri bütün eşiklenmiş imgeler için çerçeve sayısına bağlı olarak eşit olarak seçilmiştir. Ağırlıklandırma işlemi, işaret video görüntülerinin farklı sayıda çerçevelere sahip olması sebebiyle önem arz eder. Bu durum ise toplamsal imgelerin yapısını dolayısıyla tanımayı etkileyecek bir faktördür. Bu sebeple toplamsal imgeler elde edilirken eşiklenmiş imgeler ağırlıklandırılmıştır.

Şekil 4.4'te ve şekil 4.5'te Türk İşaret Dili'nde farklı kelimeler için tek bir işaretçiye ait gri-imgelerden ve derinlik haritalarından oluşturulan toplamsal imgeler gösterilmiştir. İlgili şekillerde gösterilen toplamsal imgeler sırasıyla soldan sağa olmak üzere; "affedersiniz", "görüşmek üzere", "günaydın", "iyi akşamlar" ve "lütfen" kelimelerine aittir.



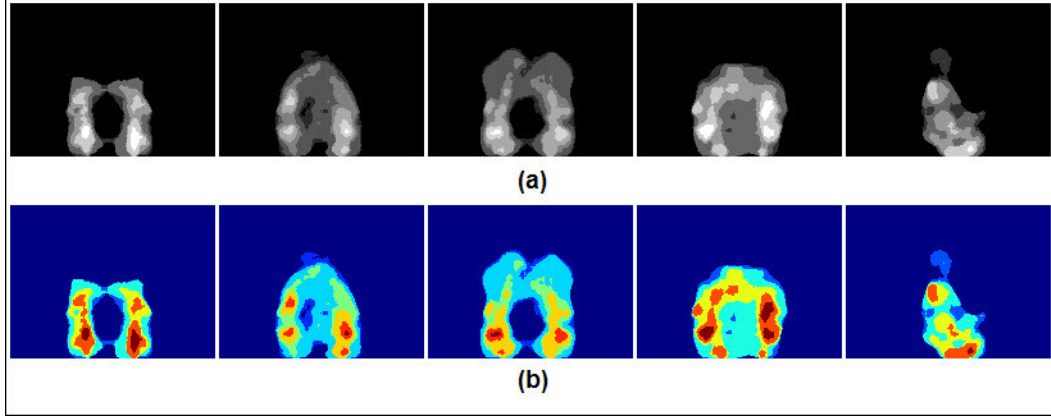
Şekil 4. 4 Gri-imgelere ait (a) gri-ölçek ve (b) renklendirilmiş toplamsal imgeler.



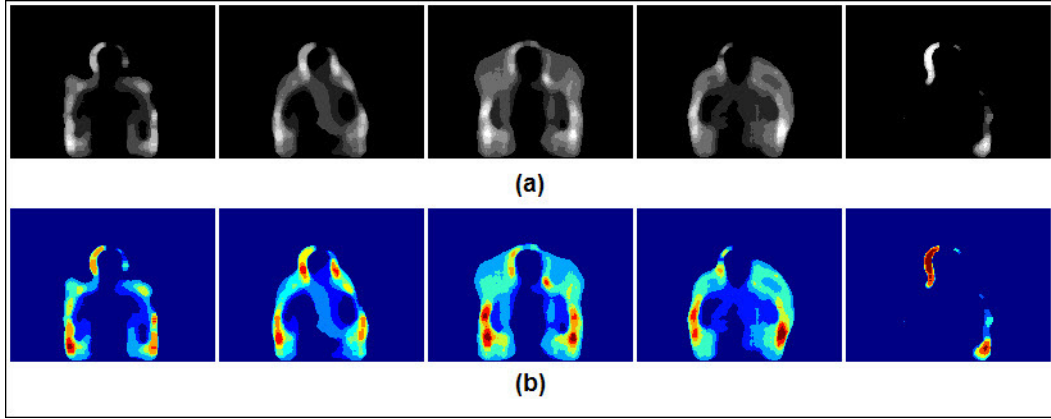
Şekil 4. 5 Derinlik haritalarına ait (a) gri-ölçek ve (b) renklendirilmiş toplamsal imgeler.

Şekil 4.4'te ve şekil 4.5'te ifade edilen toplamsal imgelerde ilk satırda yer alan imgeler gri-ölçekte (a) 0-255 renk aralığına normalleştirilmiş olarak gösterilirken, ikinci satırda yer alan imgeler (b) ise renklendirilmiş olarak gösterilmiştir. Renkli toplamsal imgelerde sıcak renkler (kırmızı ve tonları) yüksek sayısal değerleri, soğuk renkler ise (mavi tonları) düşük sayısal değerleri göstermektedir.

Toplamsal imgeler, genel görünümü itibari ile, parçalı ve gürültülü bir yapı sergilemektedir. Bu imgeleri parçalı yapıya nazaran daha genel bir yapıyla temsil etmek ve gürültü benzeri parçaları toplamsal imgelerden arındırmak amacıyla imgeler 9x9 pencere boyutlarındaki medyan filtresi ile süzgeçlenmiştir. Süzgeçlenen gri-ime ve derinlik haritası toplamsal imgeleri şekil 4.6'da ve 4.7'de sunulmuştur.

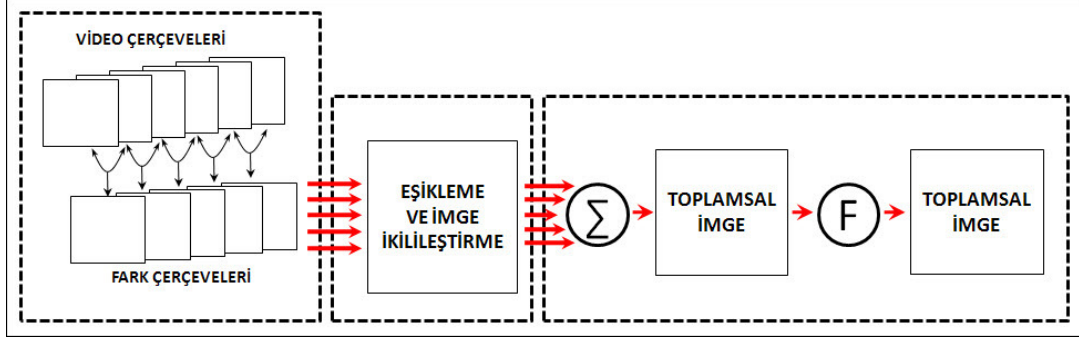


Şekil 4. 6 Gri-imgelere ait süzgeçlenmiş (a) gri-ölçek ve (b) renklendirilmiş toplamsal imgeler.



Şekil 4. 7 Derinlik haritalarına ait süzgeçlenmiş (a) gri-ölçek ve (b) renklendirilmiş toplamsal imgeler.

Şekil 4.6'da ve şekil 4.7'de gösterilen toplamsal imgeler işaretlere ait zamansal öznitelikleri içeren imgelerdir. Toplamsal imge yaklaşımı, temel olarak ardışıl çerçeveler arasındaki farkları dayandığı için ve zamansal öznitelikleri tek bir imgede ifade edebildiği için basit ve karmaşık olmayan bir yaklaşımdır. Toplamsal imgelerin elde edilmesi için sırayla; ardışıl çerçeveler arasındaki farklar bulunmuş, fark imgeleri eşiklenerek ikili imgeler elde edilmiş, ikili imgeler belli bir ağırlıklandırma katsayısı ile ağırlıklandırılarak toplanmış ve en son aşamada ise medyan filtresi ile süzgeçleme yapılmıştır. Bu süreci betimleyen bir akış şeması şekil 4.8'de gösterilmiştir. Şekil 4.8'de toplam (σ) sembolü imgelerin toplanmasını ve F sembolü ise süzgeçleme işlemine karşılık gelmektedir.



Şekil 4. 8 Toplamsal imge yaklaşımı ve zamansal öznitelik çıkarımı

Zamansal özniteliklerin elde edilmesinden sonraki aşamada, uzamsal özniteliklerin elde edilmesindeki süreç yürütülmüştür. Bu sürece ve süreçte yer alan işlemlere uzamsal özniteliklerin çıkarımı başlığı altında etraflıca değinilmiştir.

4.2.2 Uzamsal Özniteliklerin Çıkarımı

Önerilen işaret dili tanıma sisteminin öznitelik çıkarımı aşamasında, iki basamaklı bir süreç izlenmiştir. İlk aşamada işaretlere ait zamansal öznitelikleri elde etmek amacıyla, hareket farklarına dayalı toplamsal imge yaklaşımı kullanılmıştır. Bir işaretin zamandaki bütün değişimlerini içeren bu imgeler üzerinden, öznitelik vektörlerinde kullanılacak olan öznitelikleri çıkarmak için ikinci bir aşamaya geçilmiştir.

Toplamsal imgeler, zaman uzayında işaretlere ait zengin bilgiler içerir. Bu bilgileri uzamsal uzaya aktararak uzamsal öznitelikleri elde etmek için 2 boyutlu Ayırık Kosinüs Dönüşümü (2D DCT) [48] uygulanmıştır.

DCT, uygulandığı imgelerin yüksek enerjili olan önemli bilgilerini oluşturduğu katsayılar matrisinin sol üst bölümünde toplar. Bu katsayıların boyutu imgenin boyutundan oldukça küçüktür ve içerdiği bilgiler imgenin bütün enerjisinin çok büyük bir bölümünü saklar. Bu sebeple DCT, bir veriyi daha küçük boyutta temsil edebilmesi sayesinde, veri sıkıştırma yaklaşımlarında sıklıkla tercih edilir.

Çalışma kapsamında ise DCT, toplamsal imgelerin önem arz eden bilgilerini çıkarmak için kullanılmıştır. Bu sayede işaretlere ait önemli bilgiler öznitelikler olarak çıkarılmış ve işaret örneklerine ait öznitelik vektörlerinin çıkarımında kullanılmıştır. 2D DCT'ye ait denklem eşitlik (4.6)'da ifade edilmiştir.

$$F(u, v) = \alpha_u \alpha_v \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) \cdot \cos(\theta_m) \cdot \cos(\theta_n)$$

$$\theta_m = \left(\frac{\pi u (2m+1)}{2M} \right) ; \quad \theta_n = \left(\frac{\pi v (2n+1)}{2N} \right) \quad (4.6)$$

$$\alpha_u = \begin{cases} \frac{1}{\sqrt{M}} & , u=0 \\ \sqrt{\frac{2}{M}} & , 1 \leq u \leq M-1 \end{cases} \quad \alpha_v = \begin{cases} \frac{1}{\sqrt{N}} & , v=0 \\ \sqrt{\frac{2}{N}} & , 1 \leq v \leq N-1 \end{cases}$$

(4.6)'da yer alan 2D DCT denkleminde; α sabit katsayılar olmak üzere, I giriş imgesini, M ve N giriş imgesinin boyutlarını, F ise dönüşüm uygulanmış olan çıkış imgesini temsil etmektedir.

Toplamsal imgelere 2D DCT uygulanmasıyla elde edilen F çıkış imgelerinin yüksek enerjili katsayıları öznitelik vektörlerinin oluşturulmasında kullanılmıştır. Öznitelik vektörlerinin oluşturulmasında izlenen yöntemler başlık 4.3 altında detaylandırılmıştır.

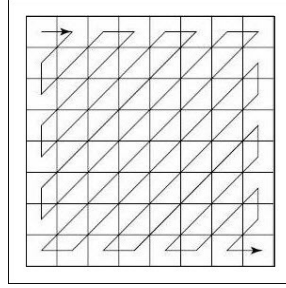
4.3 Öznitelik Vektörlerinin Oluşturulması

Önerilen işaret dili tanıma sisteminde öznitelik vektörlerini oluşturmak amacıyla, ilk aşamada işaretlere ait zamansal öznitelikler ve daha sonra da zamansal öznitelikleri içeren toplamsal imgelerden 2D DCT ile uzamsal öznitelikler çıkarılmıştır.

Türk İşaret Dili'ne ait kelimelerde, Kinect algılayıcı kullanılarak işaretlere ait RGB görüntülerin yanı sıra RGB görüntülere karşılık gelen derinlik haritaları da kaydedilmiştir. Özniteliklerin çıkarımı sürecinde, RGB renk uzayından gri-ölçeğe aktarılan görüntülere uygulanan işlemler aynı zamanda derinlik haritaları üzerinde de gerçekleştirilmiştir. Derinlik bilgilerinin işaret dili tanımda kullanılabilirliğinin sınanması da hedeflenerek bu doğrultuda işlemler yürütülmüştür.

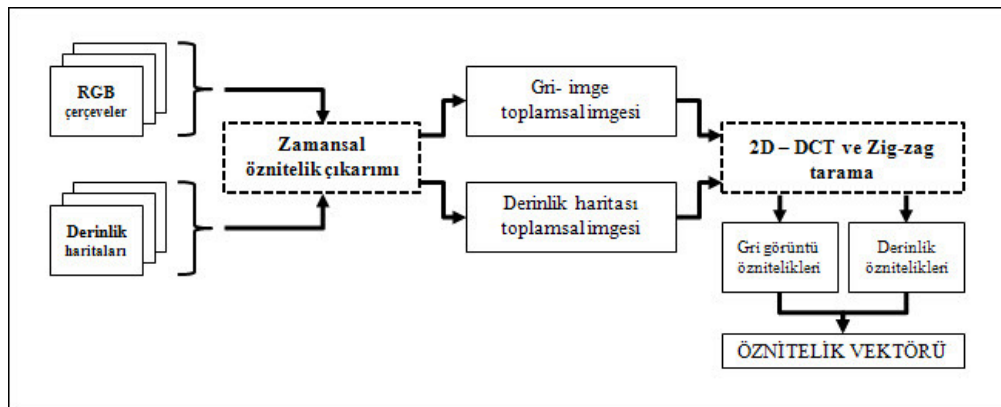
Çalışmada öznitelik vektörlerini oluşturmak amacıyla; RGB görüntülerden elde edilen gri-imgelere ve derinlik haritalarına ait özniteliklerin birlikte kullanımı önerilmiştir. Öncelikle ilgili görüntülere ait, öznitelik çıkarımı sürecinin son aşamasında 2D DCT ile elde edilen katsayıların seçilmesi sağlanmıştır. DCT, karakteristiği gereği uygulandığı

imgelerin yüksek enerjili bilgilerini katsayılar matrisi olarak da tanımlanan çıkış imgesinin sol üst köşesinde toplamaktadır. Bu katsayılar imgelere ait önemli derecede bilgi içermektedir. Bu katsayılardan seçim yapmak ve öznitelik vektörü oluşturmak amacıyla, zigzag tarama yöntemi kullanılmıştır. Zigzag tarama ile, sol üst köşeden başlanarak öznitelik katsayıları çapraz bir biçimde seçilmiş ve öznitelik vektörüne yerleştirilmiştir. Zigzag taramanın bir imge üzerinde uygulanışı şekil 4.9'da gösterilmiştir.



Şekil 4. 9 Zigzag taramanın örnek bir imge üzerinde uygulanışı.

DCT uygulanmış gri-imege ve derinlik haritalarına ait toplamsal imgelerden zigzag tarama ile seçilen katsayılar 1:1 oranında birleştirilerek işaretlere ait öznitelik vektörleri oluşturulmuştur. Öznitelik vektörlerini oluştururken kullanılan DCT katsayısının miktarının tanıma başarımına etkisini gözlemek için deneyler de yapılmış olup ilgili sonuçlar 5. Bölümdeki deneysel sonuçlar başlığı altında detaylıca verilmiştir. Şekil 4.10'da ise öznitelik vektörü oluşturma aşamasını ifade eden bir akış şeması verilmiştir.



Şekil 4. 10 Önerilen sistemde öznitelik çıkarımı ve öznitelik vektörü oluşturulması

4.4 İşaret Tanıma

İşaretlere ait öznitelik vektörlerinin oluşturulmasından sonraki aşamada ise öznitelik vektörleri kullanılarak sistemin eğitilmesi ve test edilmesi sürecinde çalışmalar yürütülmüştür. Tez çalışması kapsamında ASL ile TİD'ye ait ve tez kitabının 3. bölümünde detayları verilen iki farklı veri setinde işaretlerin tanınması amacıyla öncelikle sistem eğitilmiş ve bunun üzerinden testler yapılmıştır. Elde edilen test sonuçları 5. bölümde verilmiştir.

İşaret örneklerini tanımak amacıyla Manhattan mesafesini kullanan K-en yakın komşu sınıflandırıcısı kullanılmıştır. Veri setlerinde yapılan testlerde, eğitim ve test verilerinin ayrılması işleminde ise K-katlı çapraz doğrulama yapılarak daha genel başarı oranlarının hesaplanması hedeflenmiştir. ASL veri seti için çapraz doğrulama parametresi $K=4$ olarak seçilirken TİD veri seti için $K=3$ alınmıştır. Bu doğrultuda yürütülen testlere ve bu testlere ait sonuçlara 5. bölümde yer alan deneysel sonuçlar kısmında detaylı olarak yer verilmiştir.

BÖLÜM 5

DENEYSEL SONUÇLAR

Tez çalışması kapsamında geliştirilen işaret dili tanıma sisteminde Kinect algılayıcı tabanlı bir sistem yapısı önerilmiştir. Kinect algılayıcı aracılığıyla, işitme engelli uzman işaretçilere ait işaret görüntüleri kayda alınmıştır. Kinect'in donanımsal yapısı sayesinde işaret görüntülerine ait RGB görüntülerle birlikte bu görüntülere denk düşen derinlik haritaları da kayda alınmış olup; önerilen sistemde RGB görüntülerle bir bütünlük oluşturacak biçimde değerlendirmeye tabi tutulmuştur.

Önerilen sistemde, asıl tanıma işlemlerine ve sürecine geçilmeden önce işaretlere ait video görüntüler üzerinde video ön-işleme aşamaları gerçekleştirilmiştir. Daha çok imgelerde gürültü süzgeçleme amacıyla uygulanan ön-işlemler, önerilen sistemde renk uzayı dönüşümü ve video çerçeve boyutlarının nicemlenmesi amacıyla gerçekleştirilmiştir. Bu aşamada özellikle orijinal boyutları 640x480 olan video çerçeveleri farklı oranlarda ölçeklemeler ile yeniden yapılandırılmıştır. Video çerçevelerin yeniden boyutlandırılmasıyla, çerçeve boyutlarının tanıma aşamasındaki performanslarının değerlendirilmesi hedeflenmiştir.

İşaretlere ait özneliklerin çıkarımında ise hızlı ve basit bir yöntem olan toplamsal imge yaklaşımı kullanılmıştır. Bu yaklaşım, ardışıl video çerçevelerin parlaklık farklarına dayalı bir yöntemdir. Ardışıl çerçeve farklarında küçük hareket değişimleri öncelikle bir eşikleme katsayısına bağlı olarak süzgeçlenmekte ve hareket farklarını gösteren ikili imgeler elde edilmektedir. Bir işaret görüntüsüne ait bu ikili imgeler daha sonraki aşamada toplanarak, bir işareti temsil eden tek bir imge elde edilmektedir. Elde edilen bu imge toplamsal imge olarak adlandırılmakta olup işarete ait zamansal bilgileri

içermektedir. Daha sonraki aşamada ise zaman uzayındaki bu toplamsal imgeler Ayrık Kosinüs Dönüşümü (DCT) ile uzamsal uzaya dönüştürülmektedir. Ayrık Kosinüs Dönüşümü ile elde edilen dönüşüm katsayılarından yüksek enerjili olanlarının bir bölümü, imgelere ait önemli bilgileri içerdiği için bu katsayılar sistemde öznitelik olarak değerlendirilmeye alınmıştır. RGB görüntüler ve derinlik haritalarından bu şekilde ayrı ayrı elde edilen DCT katsayıları kullanılarak işaretlere karşılık gelen öznitelik vektörleri yapılandırılmıştır. Öznitelik vektörlerinin yapılandırılmasında, RGB görüntüler ve derinlik haritalarından elde edilen DCT katsayılarının birleştirilmesi yaklaşımı tercih edilmiştir. Bu yaklaşımla birlikte, bu görüntü tiplerinin tek başına kullanılmaları durumundaki performansı gözlemleyebilmek adına RGB görüntü ve derinlik haritaları öznitelik vektörleri tek başlarına da değerlendirilmeye alınmıştır.

DCT katsayılarının kullanım miktarına bağlı olarak farklı boyutlarda öznitelik vektörleri oluşturulması mümkündür. Bu durumun tanıma başarımında etki oluşturabilmesi muhtemeldir. Bu amaçla her işarete ait farklı boyutlarda öznitelik vektörleri de oluşturulup çalışma kapsamında değerlendirmeye tabi tutulmuştur.

Çalışma kapsamında, önerilen sistemin uygulanabilirliğini sınamak amacıyla iki farklı veri seti kullanılmıştır. Önceki bölümlerde detaylı olarak sunulan bu veri setleri; Amerikan İşaret Dili'ne (ASL) ve Türk İşaret Dili'ne ait kelimelere karşılık gelen video görüntülerden oluşmaktadır.

Tez çalışmasında önerilen işaret dili tanıma sisteminin tanıma aşamasında K-en yakın komşu (K-NN) sınıflandırıcısı yaklaşımıyla ile işaret örneklerinin tanınması sağlanmıştır. Sistemdeki başarımların ölçümünde doğru tanıma oranı (CRR – Correct Recognition Rate) değerleri hesaplanmıştır. Bu değerlerin hesaplanmasında başarılı olarak tanınan örneklerin sayısının tüm işaretlerin sayısına oranı alınmaktadır. CRR'ye ait denklem eşitlik (3.1)'de ifade edilmiştir.

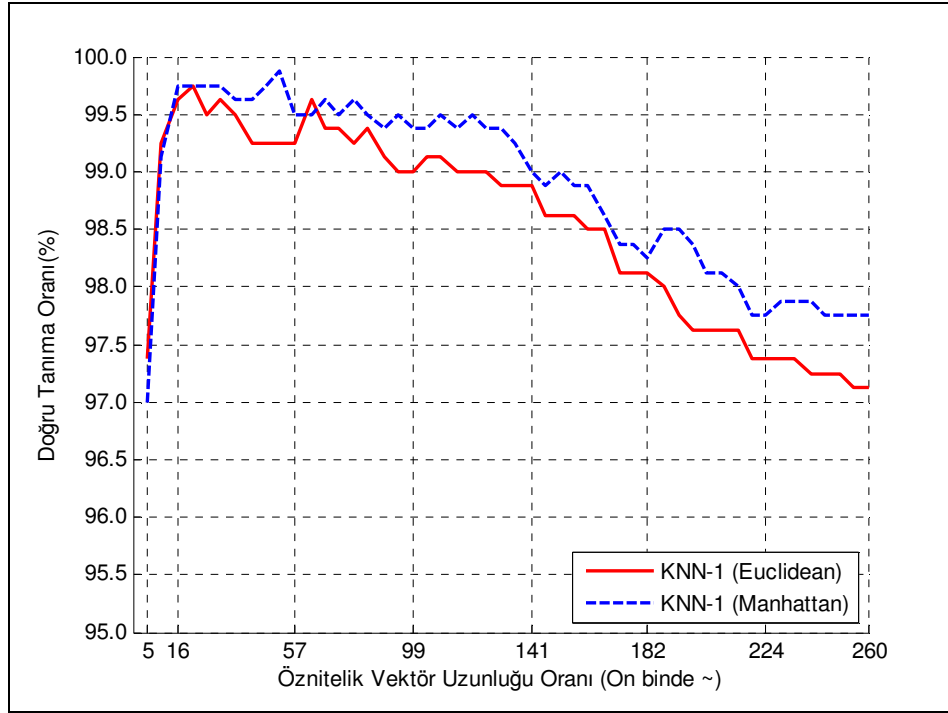
$$CRR = \frac{\# \text{doğru tanınan işaret örnekleri}}{\# \text{tüm işaret örnekleri}} \quad (5.1)$$

Sistemde yapılan kapsamlı başarımların sınamalarına ait sonuçlar ilerleyen kısımlarda detaylandırılmış ve yorumlanmıştır.

5.1 ASL Veri Seti Üzerinde Sistem Performansı Değerlendirmeleri

Tez çalışması kapsamında önerilen işaret dili tanıma sisteminin başarımının değerlendirilmesi amacıyla, Türk İşaret Dili ile birlikte, ASL'ye ait mevcut bir işaret dili veri seti de performans ölçümlerinde kullanılmıştır. Bu işaret dili veri setine ait detaylı bilgilere önceki bölümlerde etraflıca yer verilmiştir. Veri setindeki 20 kelime 4 farklı işaretçi tarafından 10'ar defa tekrar edilmiştir. Toplamda 800 işaret örneği video görüntüsünde yapılan performans değerlendirmelerinde mevcut yaklaşımla başarılı sonuçlar gözlemlenmiştir. Veri setindeki işaret videoları RGB renk uzayında yer almaktadır ve bu işaretlere ait derinlik haritaları bulunmamaktadır. Bu sebeple mevcut yaklaşım sadece RGB video görüntülerde uygulanmıştır.

Şekil 5.1'de ASL veri setinde farklı mesafe ölçümlerini kullanan KNN sınıflandırıcısının 160x120 çerçeve boyutlarındaki görüntülerde ve farklı DCT katsayı oranlarındaki doğru tanıma oranları verilmiştir.



Şekil 5. 1 Asl veri setinde farklı mesafe ölçümlerini kullanan knn sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (4-katlı çapraz doğrulama).

Şekil 5.1'deki doğru tanıma oranları analiz edildiğinde Manhattan mesafesini kullanan KNN sınıflandırıcısının Euclidean mesafesine göre daha başarılı olduğu görülmektedir.

Bununla birlikte öznitelik vektörlerinde kullanılan DCT katsayı miktarının artırılması genel olarak tanıma başarımını olumsuz etkilemektedir. Şekil 5.1’de yer alan tanıma başarımları çizelge 5.1’de detaylandırılmıştır.

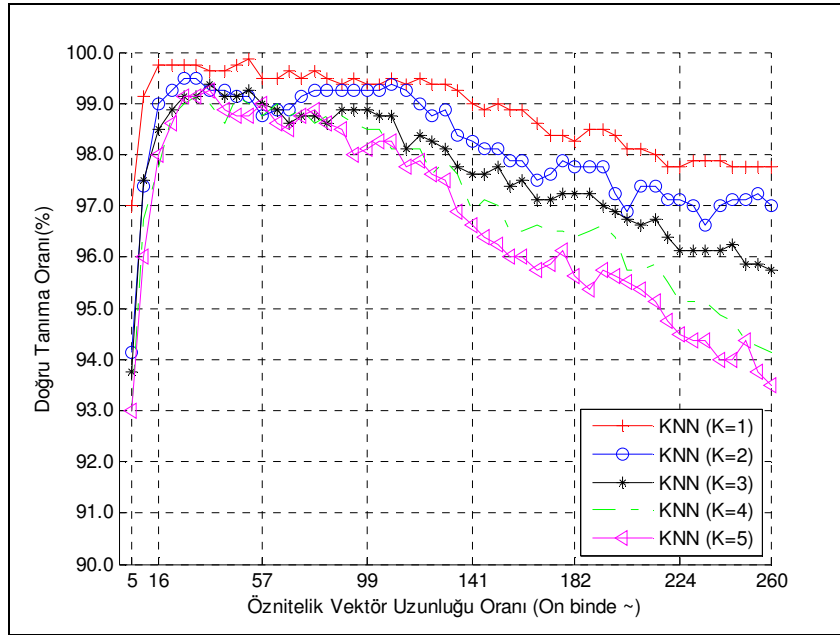
Çizelge 5. 1 Asl veri setinde manhattan ve euclidean mesafesini kullanan knn sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (4-katlı çapraz doğrulama).

DCT KATSAYI MİKTARI	DCT KATSAYI ORANI (~ ON BİNDE)	CRR (%) EUCLIDEAN MESAFESİ (K=1)	CRR (%) MANHATTAN MESAFESİ (K=1)
10	5	97,38	97,00
20	10	99,25	99,13
30	16	99,63	99,75
40	21	99,75	99,75
50	26	99,50	99,75
60	31	99,63	99,75
70	36	99,50	99,63
80	42	99,25	99,63
90	47	99,25	99,75
100	52	99,25	99,88
110	57	99,25	99,50
120	63	99,63	99,50
130	68	99,38	99,63
140	73	99,38	99,50
150	78	99,25	99,63
160	83	99,38	99,50
170	89	99,13	99,38
180	94	99,00	99,50
190	99	99,00	99,38

Çizelge 5. 1 (devam)

DCT KATSAYI MİKTARI	DCT KATSAYI ORANI (~ ON BİNDE)	CRR (%) EUCLIDEAN MESAFESİ (K=1)	CRR (%) MANHATTAN MESAFESİ (K=1)
200	104	99,13	99,38
210	109	99,13	99,50
220	115	99,00	99,38
230	120	99,00	99,50
240	125	99,00	99,38
250	130	98,88	99,38

Şekil 5.1'de ve çizelge 5.1'de detaylandırılan verilerde doğru tanıma oranlarının hesaplanmasında 4-katlı çapraz doğrulama uygulanmıştır. Aynı zamanda Manhattan mesafesini kullanan KNN sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerde ve farklı DCT katsayı oranlarındaki doğru tanıma oranları, değişen KNN K-parametreleri için şekil 5.2'de sunulmuştur.



Şekil 5. 2 Asl veri setinde manhattan mesafesini kullanan knn sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (4-katlı çapraz doğrulama ve değişen knn k-parametrelerinde).

Şekil 5.2’de KNN sınıflandırıcısında en yakın komşu sayısı K=1 alındığında tanıma başarımının diğer K değerlerindeki başarım değerlerine nazaran daha iyi olduğu görülmektedir. Genel olarak bütün K değerlerinde ise öznelilik vektör boyutu büyüdükçe başarım belirgin bir biçimde düşmektedir. Şekil 5.2’deki başarım değerlerinin detayları ise çizelge 5.2’de gösterilmiştir.

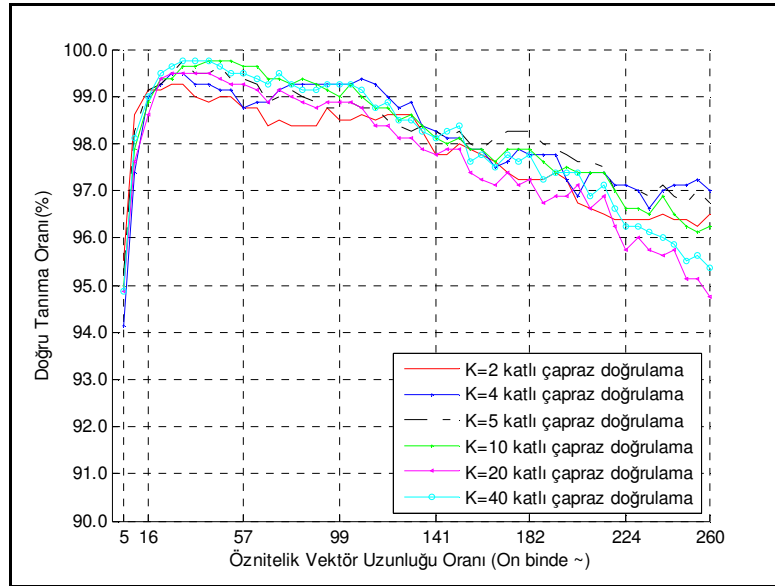
Çizelge 5. 2 Asl veri setinde manhattan mesafesini kullanan knn sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (4-katlı çapraz doğrulama ve değişen knn k-parametrelerinde).

DCT KATSAYI MİKTARI	DCT KATSAYI ORANI (~ ON BİNDE)	CRR (%) K=1	CRR (%) K=2	CRR (%) K=3	CRR (%) K=4	CRR (%) K=5
10	5	97,00	94,13	93,75	93,00	93,00
20	10	99,13	97,38	97,50	96,75	96,00
30	16	99,75	99,00	98,50	97,75	98,00
40	21	99,75	99,25	98,88	98,75	98,63
50	26	99,75	99,50	99,13	99,00	99,13
60	31	99,75	99,50	99,13	99,13	99,13
70	36	99,63	99,25	99,38	99,00	99,25
80	42	99,63	99,25	99,13	98,63	98,88
90	47	99,75	99,13	99,13	99,13	98,75
100	52	99,88	99,13	99,25	99,00	98,75
110	57	99,50	98,75	99,00	98,75	99,00
120	63	99,50	98,88	98,88	99,00	98,63
130	68	99,63	98,88	98,63	98,75	98,50
140	73	99,50	99,13	98,75	98,88	98,75
150	78	99,63	99,25	98,75	98,63	98,88
160	83	99,50	99,25	98,63	98,75	98,63
170	89	99,38	99,25	98,88	98,75	98,50

Çizelge 5. 2 (devam)

DCT KATSAYI MİKTARI	DCT KATSAYI ORANI (~ ON BİNDE)	CRR (%) K=1	CRR (%) K=2	CRR (%) K=3	CRR (%) K=4	CRR (%) K=5
180	94	99,50	99,25	98,88	98,63	98,00
190	99	99,38	99,25	98,88	98,50	98,13
200	104	99,38	99,25	98,75	98,50	98,25
210	109	99,50	99,38	98,75	98,13	98,25
220	115	99,38	99,25	98,13	98,13	97,75
230	120	99,50	99,00	98,38	98,13	97,88
240	125	99,38	98,75	98,25	97,63	97,63
250	130	99,38	98,88	98,13	97,88	97,50

İlgili veri setinde yapılan diğer başarımların sınamalarında farklı K-katlı çapraz doğrulama parametrelerinin de başarımların sonuçları gözlemlenmiştir. Şekil 5.3'te Manhattan mesafesindeki K=1 en yakın komşuyu kullanan KNN sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı DCT katsayı oranlarındaki doğru tanıma oranları, değişik K-katlı çapraz doğrulama parametreleri için verilmiştir.



Şekil 5. 3 Asl veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (farklı k-katlı çapraz doğrulamalarda).

Şekil 5.3'teki başarımların değerleri farklı çapraz doğrulama K-kat parametrelerinde gözlemlenmiştir. Grafikteki verilere göre K-kat değerinin arttırılmasında başarımların değerleri artmaktadır. K-kat değeri arttırılınca eğitim kümesindeki sınıflara ait örnek sayısı artmakta ve bu durumda işaretlerin tanınması kolaylaşmaktadır. Grafikteki başarımların değerleri yüksek K-kat değerleri için daha iyi olmakla beraber genel olarak birbirine yakındır. Öznitelik vektör boyutunun değişiminde başarımların değerlerinin değişimi de birbirlerine benzemektedir. Şekil 5.3'teki başarımların değerlerinin detayları ise çizelge 5.3'te gösterilmiştir.

Çizelge 5. 3 Asl veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 160x120 çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (farklı k-katlı çapraz doğrulamalarda).

DCT KATSAYI MİKTARI	DCT KATSAYI ORANI (~ ON BİNDE)	CRR (%) K=2	CRR (%) K=4	CRR (%) K=5	CRR (%) K=10	CRR (%) K=20	CRR (%) K=40
10	5	95,50	97,00	96,75	97,25	97,50	97,50
20	10	98,63	99,13	99,13	99,50	99,50	99,50
30	16	99,13	99,75	99,75	100,00	100,00	100,00
40	21	99,13	99,75	99,50	99,88	99,88	99,88
50	26	99,25	99,75	99,88	99,88	99,88	99,88
60	31	99,25	99,75	99,75	99,88	99,88	99,88
70	36	99,00	99,63	99,63	99,75	99,75	99,75
80	42	98,88	99,63	99,38	99,63	99,75	99,75
90	47	99,00	99,75	99,63	99,75	99,75	99,75
100	52	99,00	99,88	99,63	99,75	99,75	99,75
110	57	98,75	99,50	99,38	99,63	99,63	99,63
120	63	98,75	99,50	99,50	99,75	99,75	99,75
130	68	98,38	99,63	99,50	99,63	99,63	99,63
140	73	98,50	99,50	99,38	99,63	99,63	99,63

Çizelge 5. 3 (devam)

DCT KATSAYI MİKTARI	DCT KATSAYI ORANI (~ ON BİNDE)	CRR (%) K=2	CRR (%) K=4	CRR (%) K=5	CRR (%) K=10	CRR (%) K=20	CRR (%) K=40
150	78	98,38	99,63	99,38	99,75	99,75	99,75
160	83	98,38	99,50	99,38	99,63	99,63	99,63
170	89	98,38	99,38	99,25	99,63	99,63	99,63
180	94	98,75	99,50	99,00	99,50	99,63	99,63
190	99	98,50	99,38	99,00	99,50	99,50	99,50
200	104	98,50	99,38	99,13	99,63	99,63	99,63
210	109	98,63	99,50	99,00	99,63	99,63	99,63
220	115	98,50	99,38	98,88	99,50	99,50	99,50
230	120	98,63	99,50	98,88	99,38	99,38	99,38
240	125	98,63	99,38	98,88	99,38	99,50	99,50
250	130	98,63	99,38	98,75	99,50	99,50	99,50

Şekil 5.3'te ve çizelge 5.3'te K çapraz doğrulama parametresi arttırıldıkça eğitim kümesindeki sınıflara ait örnek sayıları da artmaktadır. Dolayısıyla sistemde test örnekleri daha iyi bir biçimde tanınabilmektedir ve başarımlar artmaktadır.

5.2 TİD Veri Seti Üzerinde Sistem Performansı Değerlendirmeleri

Tez çalışması kapsamında önerilen işaret dili tanıma sisteminin performans değerlendirmeleri ASL'ye ait mevcut bir veri setinin yanı sıra, Türk İşaret Dili'ne ait, tez çalışması kapsamında oluşturulan ve detayları önceki bölümlerde verilen bir veri seti üzerinde de yapılmıştır. Bu kısımda ilgili veri seti üzerinde yapılan başarımların sınımlarına ait gözlemler ve sonuçlar verilmiştir.

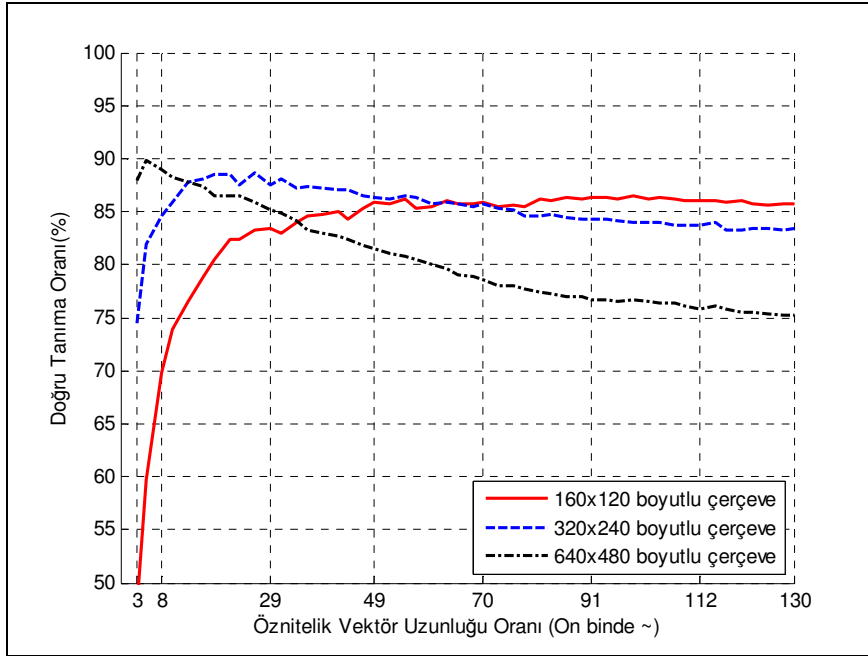
İlgili veri setinde 3 farklı işaretçi tarafından gerçekleştirilen 3 kategoride yer alan toplam 111 kelimeye ait 1002 adet işaret örneği değerlendirmeye alınmıştır. İşaretlere ait öznitelik vektörlerinin çıkarımında Kinect algılayıcı ile kaydedilen RGB görüntülerle

derinlik haritalarından faydalanılmıştır. Öznitelik vektörleri oluşturmada bu bilgiler birlikte kullanılırken, ayrı ayrı olarak tek başlarına da kullanılmıştır. Böylece sistemin tanıma başarımında RGB görüntüler ve derinlik haritalarının tek başlarına ve birlikte kullanım etkilerinin analizleri hedeflenmiştir.

Bununla beraber orijinal olarak 640x480 boyutlarından kaydedilen görüntüler, önerilen sistemin ön-işleme aşamasında belirtildiği üzere $\frac{1}{2}$ ve $\frac{1}{4}$ oranlarında ölçeklendirilip 320x240 ve 160x120 boyutlarına indirgenerek farklı çerçeve boyutlarında başarımların sınanması hedeflenmiştir.

Bu bağlamda Türk İşaret Dili'ne ait veri setinde performans değerlendirmeleri etrafıca yapılmış ve ilerleyen bölümlerde bu değerlendirmelere ait sonuçlara yer verilmiştir.

Şekil 5.4'te farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en (K=1) yakın komşu (KNN) sınıflandırıcısının veri setinde yer alan 111 kelime üzerindeki başarımları gösterilmiştir.



Şekil 5. 4 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve gri-imege + derinlik haritası özniteliklerinde).

Şekil 5.4'te farklı çerçeve boyutlarında gözlemlenen başarımlar sunulmuştur. İlgili testte en iyi başarımlar oranı 640x480 çerçeve boyutlarında gözlemlenmiştir. Ayrıca düşük

boyutlu öznitelik vektöründe en iyi performansı yine 640x480 çerçeve boyutlu görüntüler sergilemektedir. Öznitelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarımlar artmaktadır ve 320x240 boyutlu çerçevelerde ise başarımlar genel olarak düşmekle birlikte daha kararlı bir değişim sergilemektedir. Öznitelik vektör boyutunun arttırılması 640x480 çerçevelerin performansını ise oldukça olumsuz etkilemektedir. İlgili başarımlar grafiğinde 320x240 boyutlu çerçevelerin daha kararlı bir performans sergilemesi sebebiyle bazı testler özel olarak bu çerçeve boyutları için yapılmıştır. Şekil 5.4'teki başarımlar değerlerinin detayları ise çizelge 5.4'te gösterilmiştir.

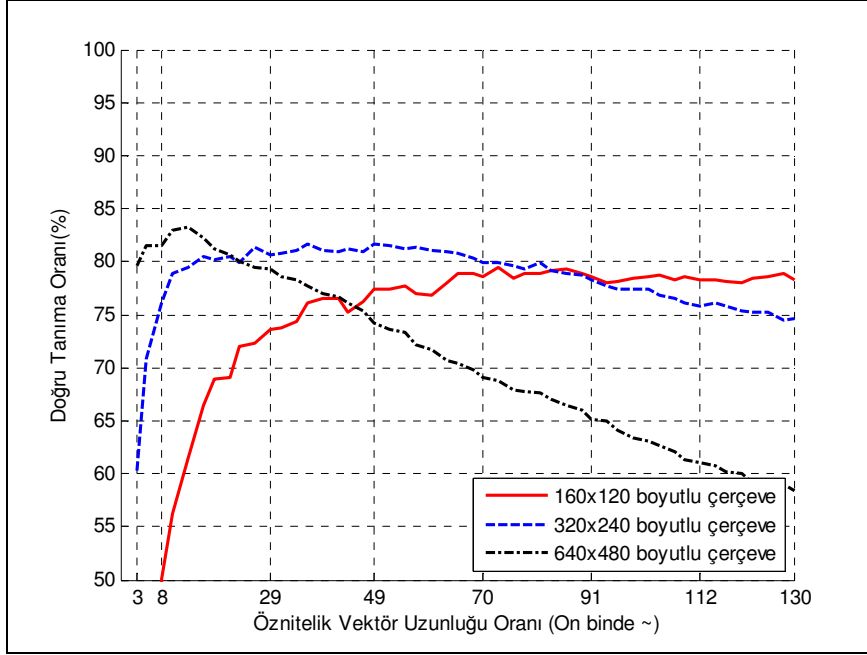
Çizelge 5. 4 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve gri-ime + derinlik haritası özniteliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	10	46,61	40	74,45	160	87,92
5	20	59,68	80	81,94	320	89,82
8	30	69,76	120	84,63	480	88,92
10	40	73,85	160	85,83	640	88,22
13	50	76,55	200	87,72	800	87,82
16	60	78,94	240	88,02	960	87,33
18	70	80,54	280	88,52	1120	86,43
21	80	82,34	320	88,52	1280	86,43
23	90	82,34	360	87,53	1440	86,43
26	100	83,23	400	88,62	1600	85,93
29	110	83,43	440	87,52	1760	85,13
31	120	82,93	480	88,02	1920	84,83
34	130	83,93	520	87,23	2080	84,13

Çizelge 5. 4 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
36	140	84,53	560	87,43	2240	83,33
39	150	84,73	600	87,23	2400	83,03
42	160	85,03	640	87,13	2560	82,63
44	170	84,33	680	87,03	2720	82,44
47	180	85,33	720	86,43	2880	81,74
49	190	85,93	760	86,33	3040	81,44
52	200	85,73	800	86,13	3200	81,14
55	210	86,23	840	86,53	3360	80,74
57	220	85,33	880	86,33	3520	80,44
60	230	85,43	920	85,73	3680	80,04
63	240	86,03	960	85,83	3840	79,54
65	250	85,73	1000	85,73	4000	79,04

Şekil 5.5'te farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının veri setinde yer alan 111 kelime üzerinde sadece grime öznitelikleri kullanıldığındaki başarımları gösterilmiştir.



Şekil 5. 5 Tid veri setinde manhattan mesafesini kullanan knn ($k=1$) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve gri-imege özneliklerinde).

Şekil 5.5'te farklı çerçeve boyutlarında sadece gri-imgelere ait özneliklerin kullanılmasıyla gözlemlenen başarımlar sunulmuştur. Elde edilen başarımların değerlerinin karakteristikleri şekil 5.4 ile büyük oranda benzerlik göstermektedir. Düşük boyutlu öznelik vektöründe en iyi performansı yine 640x480 çerçeve boyutlu görüntüler sergilemektedir. Öznelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarımlar artmaktadır ve 320x240 boyutlu çerçevelerde ise genel olarak başarımlar düşmekle birlikte daha kararlı bir değişim sergilemektedir. Öznelik vektör boyutunun arttırılması 640x480 çerçevelerin performansını ise oldukça olumsuz etkilemektedir. Şekil 5.5'teki başarımların değerlerinin detayları ise çizelge 5.5'te gösterilmiştir.

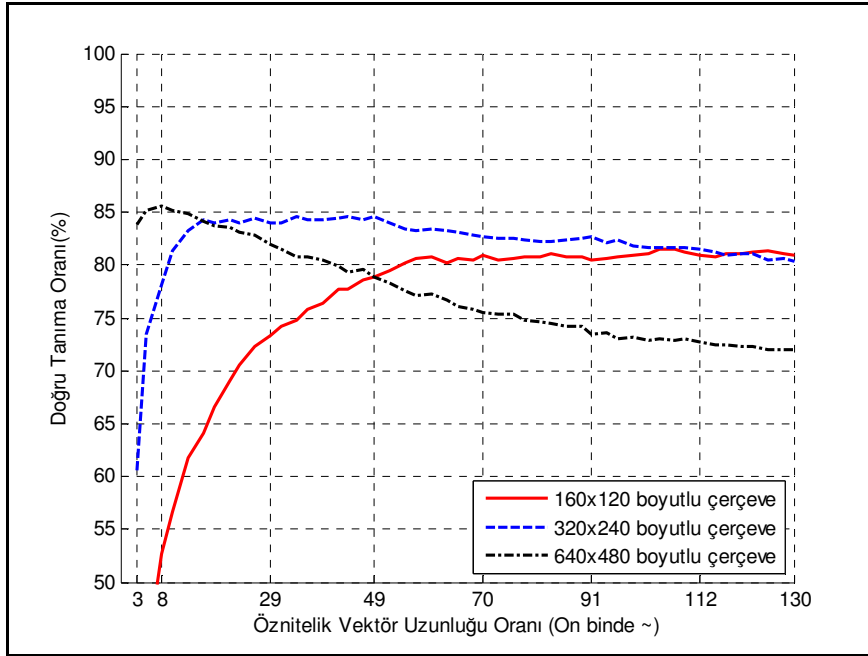
Çizelge 5. 5 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve gri-ime özniteliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	5	26,65	20	60,38	80	79,54
5	10	41,12	40	70,86	160	81,44
8	15	50,10	60	76,15	240	81,54
10	20	56,19	80	78,84	320	82,93
13	25	61,48	100	79,44	400	83,33
16	30	66,47	120	80,44	480	82,24
18	35	68,96	140	80,24	560	81,24
21	40	69,06	160	80,54	640	80,64
23	45	71,96	180	79,94	720	80,04
26	50	72,26	200	81,34	800	79,44
29	55	73,55	220	80,64	880	79,24
31	60	73,75	240	80,84	960	78,64
34	65	74,35	260	81,14	1040	78,24
36	70	76,05	280	81,64	1120	77,64
39	75	76,55	300	81,14	1200	76,95
42	80	76,55	320	80,94	1280	76,65
44	85	75,15	340	81,24	1360	76,15
47	90	76,25	360	80,94	1440	75,35
49	95	77,45	380	81,64	1520	74,25
52	100	77,45	400	81,44	1600	73,65

Çizelge 5. 5 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
55	105	77,64	420	81,24	1680	73,25
57	110	76,95	440	81,34	1760	72,16
60	115	76,85	460	81,04	1840	71,76
63	120	77,94	480	80,94	1920	70,76
65	125	78,84	500	80,84	2000	70,46

Şekil 5.6'da farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının veri setinde yer alan 111 kelime üzerinde sadece derinlik haritası öznitelikleri kullanıldığında başarımları gösterilmiştir.



Şekil 5. 6 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve derinlik haritası özniteliklerinde).

Şekil 5.6'da farklı çerçeve boyutlarında sadece derinlik haritalarına ait özniteliklerin kullanılmasıyla gözlemlenen başarımlar sunulmuştur. Elde edilen başarımların değerlerinin

karakteristikleri şekil 5.5 ile büyük oranda benzerlik göstermektedir. Düşük boyutlu öznitelik vektöründe en iyi performansı yine 640x480 çerçeve boyutlu görüntüler sergilemektedir. Öznitelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarımlar artmaktadır ve 320x240 boyutlu çerçevelerde ise genel olarak başarımlar düşmekle birlikte daha kararlı bir değişim sergilemektedir. Öznitelik vektör boyutunun arttırılması 640x480 çerçevelerin performansını ise oldukça olumsuz etkilemektedir. Şekil 5.5'teki gri-imgelere ait başarımlar ile şekil 5.6'daki derinlik haritalarına ait başarımlar karşılaştırıldığında, derinlik haritalarının işaretleri tanımda daha etkin olduğu görülmektedir. Şekil 5.6'daki başarımların detayları ise çizelge 5.6'da gösterilmiştir.

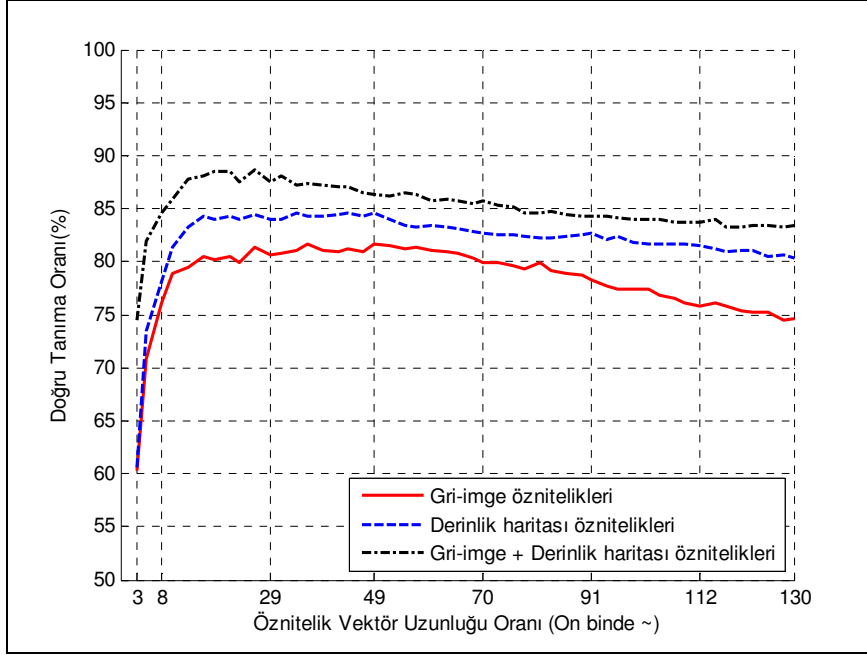
Çizelge 5. 6 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime ve derinlik haritası özniteliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	5	26,45	20	60,58	80	83,83
5	10	44,81	40	73,45	160	85,23
8	15	52,69	60	78,14	240	85,53
10	20	56,69	80	81,34	320	85,23
13	25	61,78	100	83,23	400	84,93
16	30	64,07	120	84,23	480	84,13
18	35	66,57	140	84,03	560	83,73
21	40	69,06	160	84,33	640	83,53
23	45	70,56	180	83,93	720	83,13
26	50	72,36	200	84,43	800	82,83
29	55	73,35	220	84,03	880	81,94

Çizelge 5. 6 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
31	60	74,15	240	84,03	960	81,44
34	65	74,85	260	84,53	1040	80,74
36	70	75,85	280	84,33	1120	80,84
39	75	76,35	300	84,23	1200	80,54
42	80	77,74	320	84,43	1280	79,84
44	85	77,74	340	84,53	1360	79,34
47	90	78,64	360	84,33	1440	79,54
49	95	78,84	380	84,53	1520	78,94
52	100	79,44	400	84,03	1600	78,24
55	105	80,24	420	83,43	1680	77,54
57	110	80,64	440	83,33	1760	77,15
60	115	80,74	460	83,43	1840	77,25
63	120	80,24	480	83,23	1920	76,75
65	125	80,64	500	83,13	2000	76,15

Şekil 5.7'de 320x240 çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının veri setinde yer alan 111 kelime üzerinde farklı özniteliklerin kullanımlarındaki başarımları gösterilmiştir.



Şekil 5. 7 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime).

Şekil 5.7, 320x240 çerçeve boyutlarındaki görüntülerde farklı özniteliklerin tanıma başarımlarını göstermektedir. 320x240 çerçeve boyutunda tanıma başarımının önceki testlerde daha kararlı bir performans sergilemesi sebebiyle, ilgili başarıım grafiği 320x240 çerçeve boyutu için gösterilmiştir. Şekil 5.7’de gri-imge, derinlik haritası özniteliklerinin ayrı ayrı değerlendirilmesiyle gözlemlenen tanıma başarımları ve bu özniteliklerin bire bir (1:1) oranında birleştirildiği durumda gözlemlenen tanıma başarımları yer almaktadır. İlgili başarıım grafiğinde en başarılı sonuçlara gri-imge+derinlik haritası özniteliklerinin kullanımında ulaşılmıştır. Bununla beraber derinlik haritası özniteliklerinin, gri-imge özniteliklerinden başarılı olduğu da görülmüştür. Şekil 5.7’deki başarıım değerlerinin detayları ise çizelge 5.7’de gösterilmiştir.

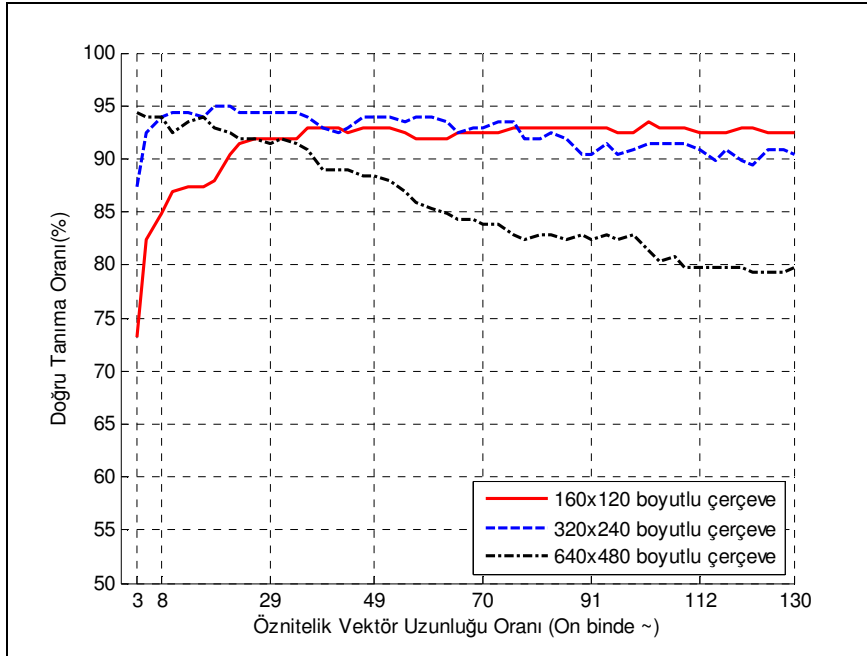
Çizelge 5. 7 Tid veri setinde manhattan mesafesini kullanan knn (K=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (111 kelime).

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNİTELİKLERİ		DERİNLİK HARİTASI ÖZİNİTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNİTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	20	60,38	20	60,58	40	74,45
5	40	70,86	40	73,45	80	81,94
8	60	76,15	60	78,14	120	84,63
10	80	78,84	80	81,34	160	85,83
13	100	79,44	100	83,23	200	87,72
16	120	80,44	120	84,23	240	88,02
18	140	80,24	140	84,03	280	88,52
21	160	80,54	160	84,33	320	88,52
23	180	79,94	180	83,93	360	87,53
26	200	81,34	200	84,43	400	88,62
29	220	80,64	220	84,03	440	87,53
31	240	80,84	240	84,03	480	88,02
34	260	81,14	260	84,53	520	87,23
36	280	81,64	280	84,33	560	87,43
39	300	81,14	300	84,23	600	87,23
42	320	80,94	320	84,43	640	87,13
44	340	81,24	340	84,53	680	87,03
47	360	80,94	360	84,33	720	86,43
49	380	81,64	380	84,53	760	86,33
52	400	81,44	400	84,03	800	86,13

Çizelge 5. 7 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNTELİKLERİ		DERİNLİK HARİTASI ÖZİNTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
55	420	81,24	420	83,43	840	86,53
57	440	81,34	440	83,33	880	86,33
60	460	81,04	460	83,43	920	85,73
63	480	80,94	480	83,23	960	85,83
65	500	80,84	500	83,13	1000	85,73

Şekil 5.8'de farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 1. kategoride yer alan 22 kelime üzerindeki başarımları gösterilmiştir.



Şekil 5. 8 Tid veri setinde manhattan mesafesini kullanan knn (K=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1 ve gri-imge + derinlik haritası öz niteliklerinde).

Şekil 5.8'de TİD veri setindeki 1. kategoride yer alan kelimelerin farklı çerçeve boyutlarındaki başarımları yer almaktadır. Başarım değerleri gri-imge + derinlik haritası

özniteliklerinin kullanımında hesaplanmış olup, en iyi başarımların değerlerine 320x240 çerçeve boyutlarında ulaşılmıştır. Öznitelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarımlar artmaktadır ve 320x240 boyutlu çerçevelerde ise genel olarak başarımlar düşmektedir. Öznitelik vektör boyutunun arttırılması 640x480 çerçevelere ait başarımların belirgin bir biçimde düşmesine sebep olmaktadır. Şekil 5.8’de yer alan başarımların detayları ise çizelge 5.8’de verilmiştir.

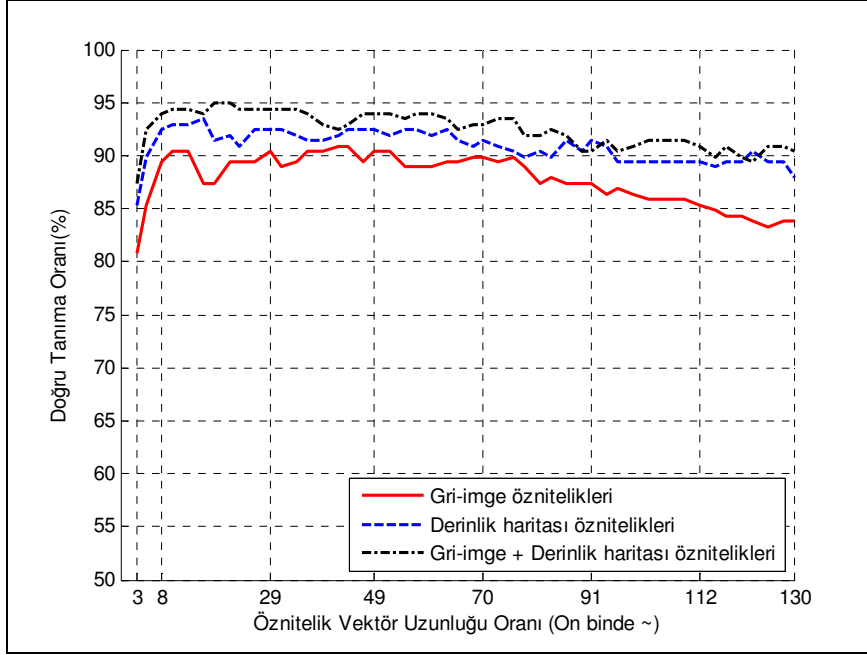
Çizelge 5. 8 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1 ve gri-imege + derinlik haritası özniteliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	10	73,23	40	87,37	160	94,44
5	20	82,32	80	92,42	320	93,94
8	30	84,85	120	93,94	480	93,94
10	40	86,87	160	94,44	640	92,42
13	50	87,37	200	94,44	800	93,43
16	60	87,37	240	93,94	960	93,94
18	70	87,88	280	94,95	1120	92,93
21	80	90,40	320	94,95	1280	92,42
23	90	91,41	360	94,44	1440	91,92
26	100	91,92	400	94,44	1600	91,92
29	110	91,92	440	94,44	1760	91,41
31	120	91,92	480	94,44	1920	91,92
34	130	91,92	520	94,44	2080	91,41

Çizelge 5. 8 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
36	140	92,93	560	93,94	2240	90,91
39	150	92,93	600	92,93	2400	88,89
42	160	92,93	640	92,42	2560	88,89
44	170	92,42	680	92,93	2720	88,89
47	180	92,93	720	93,94	2880	88,38
49	190	92,93	760	93,94	3040	88,38
52	200	92,93	800	93,94	3200	87,88
55	210	92,42	840	93,43	3360	86,87
57	220	91,92	880	93,94	3520	85,86
60	230	91,92	920	93,94	3680	85,35
63	240	91,92	960	93,43	3840	84,85
65	250	92,42	1000	92,42	4000	84,34

Şekil 5.9'da 320x240 çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 1. Kategoride yer alan 22 kelime üzerinde farklı özniteliklerin kullanımlarındaki başarımları gösterilmiştir.



Şekil 5. 9 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1).

Şekil 5.9'da gri-imge, derinlik haritası özniteliklerinin ayrı ayrı değerlendirilmesiyle gözlemlenen tanıma başarımları ve bu özniteliklerin bire bir (1:1) oranında birleştirildiği durumda gözlemlenen tanıma başarımları yer almaktadır. İlgili başarımlar grafiğinde en başarılı sonuçlara gri-imge+derinlik haritası özniteliklerinin kullanımında ulaşılmıştır. Bununla beraber derinlik haritası özniteliklerinin, gri-imge özniteliklerinden başarılı olduğu da görülmüştür. Şekil 5.9'daki başarımların değerlerinin detayları ise çizelge 5.9'da gösterilmiştir.

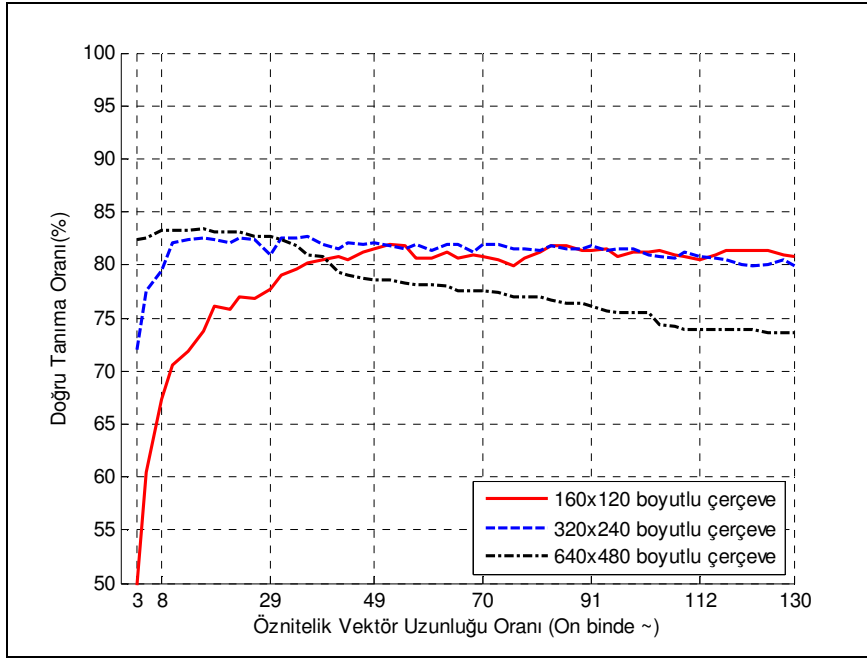
Çizelge 5. 9 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1).

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNİTELİKLERİ		DERİNLİK HARİTASI ÖZİNİTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNİTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	20	80,81	20	85,35	40	87,37
5	40	85,35	40	89,90	80	92,42
8	60	89,39	60	92,42	120	93,94
10	80	90,40	80	92,93	160	94,44
13	100	90,40	100	92,93	200	94,44
16	120	87,37	120	93,43	240	93,94
18	140	87,37	140	91,41	280	94,95
21	160	89,39	160	91,92	320	94,95
23	180	89,39	180	90,91	360	94,44
26	200	89,39	200	92,42	400	94,44
29	220	90,40	220	92,42	440	94,44
31	240	88,89	240	92,42	480	94,44
34	260	89,39	260	91,92	520	94,44
36	280	90,40	280	91,41	560	93,94
39	300	90,40	300	91,41	600	92,93
42	320	90,91	320	91,92	640	92,42
44	340	90,91	340	92,42	680	92,93
47	360	89,39	360	92,42	720	93,94
49	380	90,40	380	92,42	760	93,94
52	400	90,40	400	91,92	800	93,94

Çizelge 5. 9 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNTELİKLERİ		DERİNLİK HARİTASI ÖZİNTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
55	420	88,89	420	92,42	840	93,43
57	440	88,89	440	92,42	880	93,94
60	460	88,89	460	91,92	920	93,94
63	480	89,39	480	92,42	960	93,43
65	500	89,39	500	91,41	1000	92,42

Şekil 5.10'da farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 2. Kategoride yer alan 58 kelime üzerindeki başarımları gösterilmiştir.



Şekil 5. 10 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2 ve gri-imge + derinlik haritası özneliklerinde).

Şekil 5.10'da TİD veri setindeki 2. kategoride yer alan kelimelerin farklı çerçeve boyutlarındaki başarımları yer almaktadır. Başarım değerleri gri-imge + derinlik haritası

özniteliklerinin kullanımında hesaplanmış olup, en iyi başarımların değerlerine 640x480 çerçeve boyutlarında ulaşılmıştır. Öznitelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarımlar artmaktadır ve 320x240 boyutlu çerçevelerde ise genel olarak başarımlar düşmektedir. Öznitelik vektör boyutunun arttırılması 640x480 çerçevelere ait tanıma performansını olumsuz etkilemektedir. Şekil 5.10'da yer alan başarımların detayları ise çizelge 5.10'da verilmiştir.

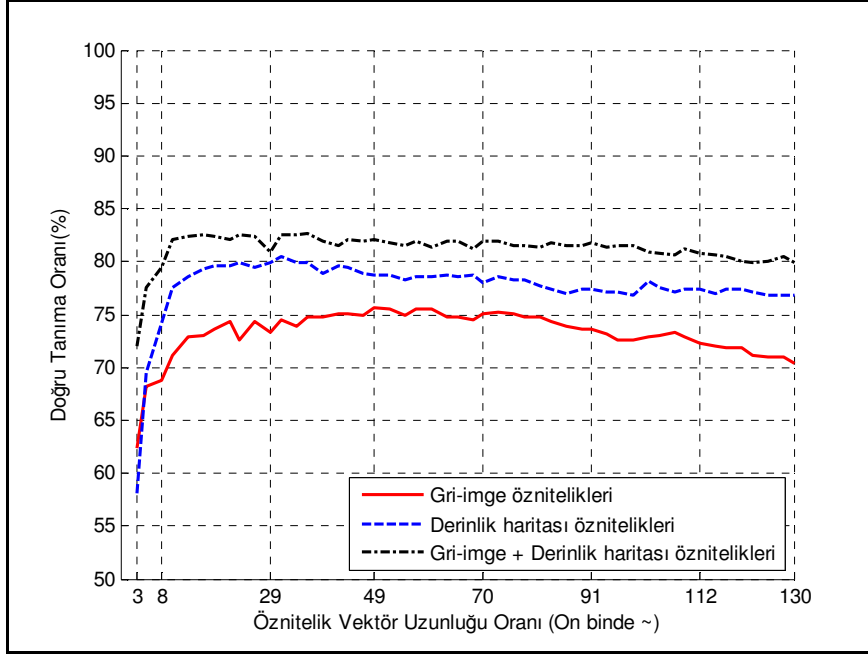
Çizelge 5. 10 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2 ve gri-imege + derinlik haritası özniteliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	10	48,29	40	72,05	160	82,32
5	20	60,46	80	77,57	320	82,51
8	30	67,30	120	79,47	480	83,27
10	40	70,53	160	82,13	640	83,27
13	50	71,86	200	82,32	800	83,27
16	60	73,76	240	82,51	960	83,46
18	70	76,05	280	82,32	1120	83,08
21	80	75,86	320	82,13	1280	83,08
23	90	77,00	360	82,51	1440	83,08
26	100	76,81	400	82,32	1600	82,70
29	110	77,76	440	80,99	1760	82,70
31	120	79,09	480	82,51	1920	82,32
34	130	79,66	520	82,51	2080	81,75
36	140	80,23	560	82,70	2240	80,99

Çizelge 5. 10 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
39	150	80,42	600	81,94	2400	80,80
42	160	80,80	640	81,56	2560	79,28
44	170	80,42	680	82,13	2720	79,09
47	180	81,18	720	81,94	2880	78,71
49	190	81,56	760	82,13	3040	78,52
52	200	81,94	800	81,75	3200	78,52
55	210	81,75	840	81,56	3360	78,33
57	220	80,61	880	81,94	3520	78,14
60	230	80,61	920	81,37	3680	78,14
63	240	81,18	960	81,94	3840	77,95
65	250	80,61	1000	81,94	4000	77,57

Şekil 5.11’de 320x240 çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 2. Kategoride yer alan 58 kelime üzerinde farklı özniteliklerin kullanımlarındaki başarımları gösterilmiştir.



Şekil 5. 11 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2).

Şekil 5.11’de gri-ime, derinlik haritası özneliklerinin ayrı ayrı değerlendirilmesiyle gözlemlenen tanıma başarımları ve bu özneliklerin bire bir (1:1) oranında birleştirildiği durumda gözlemlenen tanıma başarımları yer almaktadır. İlgili başarımlar grafiğinde en başarılı sonuçlara gri-ime+derinlik haritası özneliklerinin kullanımında ulaşılmıştır. Bununla beraber derinlik haritası özneliklerinin, gri-ime özneliklerinden başarılı olduğu da görülmüştür. Şekil 5.11’deki başarımların detayları ise çizelge 5.11’de gösterilmiştir.

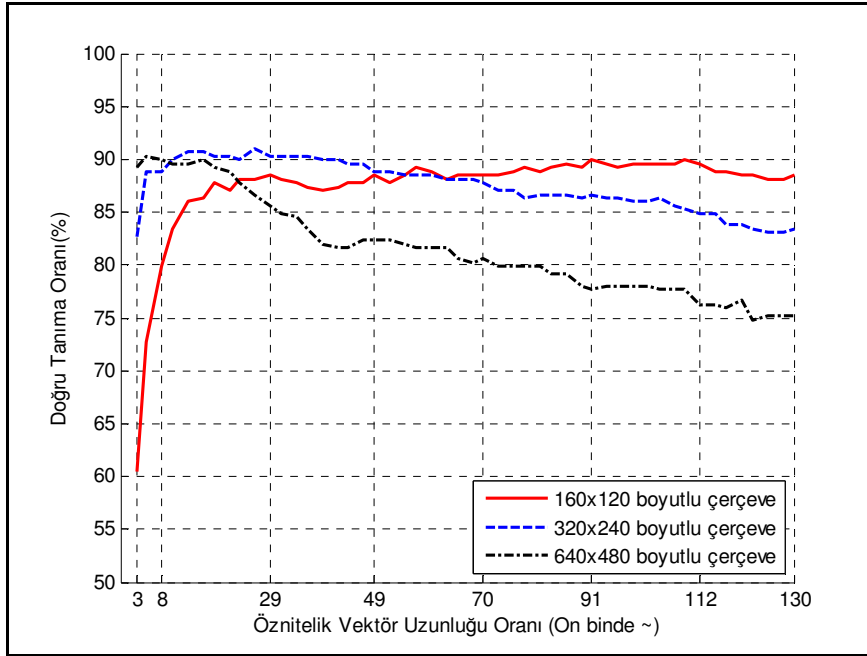
Çizelge 5. 11 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2).

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNİTELİKLERİ		DERİNLİK HARİTASI ÖZİNİTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNİTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	20	62,36	20	58,17	40	72,05
5	40	68,25	40	69,58	80	77,57
8	60	68,82	60	74,14	120	79,47
10	80	71,10	80	77,57	160	82,13
13	100	72,81	100	78,52	200	82,32
16	120	73,00	120	79,28	240	82,51
18	140	73,57	140	79,66	280	82,32
21	160	74,33	160	79,66	320	82,13
23	180	72,62	180	79,85	360	82,51
26	200	74,33	200	79,47	400	82,32
29	220	73,38	220	79,85	440	80,99
31	240	74,52	240	80,42	480	82,51
34	260	73,95	260	79,85	520	82,51
36	280	74,71	280	79,85	560	82,70
39	300	74,71	300	78,90	600	81,94
42	320	75,10	320	79,66	640	81,56
44	340	75,10	340	79,47	680	82,13
47	360	74,90	360	78,90	720	81,94
49	380	75,67	380	78,71	760	82,13
52	400	75,48	400	78,71	800	81,75

Çizelge 5. 11 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNTELİKLERİ		DERİNLİK HARİTASI ÖZİNTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
55	420	74,90	420	78,33	840	81,56
57	440	75,48	440	78,52	880	81,94
60	460	75,48	460	78,52	920	81,37
63	480	74,71	480	78,71	960	81,94
65	500	74,71	500	78,52	1000	81,94

Şekil 5.12'de farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 3. Kategoride yer alan 31 kelime üzerindeki başarımları gösterilmiştir.



Şekil 5. 12 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3 ve gri-imge + derinlik haritası özneliklerinde).

Şekil 5.12'de TİD veri setindeki 3. kategoride yer alan kelimelerin farklı çerçeve boyutlarındaki başarımları yer almaktadır. Başarım değerleri gri-imge + derinlik haritası

özniteliklerinin kullanımında hesaplanmış olup, en iyi başarımların değerlerine 320x240 çerçeve boyutlarında ulaşılmıştır. Öznitelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarımlar artmaktadır ve 320x240 boyutlu çerçevelerde ise genel olarak başarımlar düşmektedir. Öznitelik vektör boyutunun arttırılması 640x480 çerçevelere ait tanıma başarımlarının belirgin biçimde düşmesine sebep olmaktadır. Şekil 5.12’de yer alan başarımların oranlarının detayları ise çizelge 5.12’de verilmiştir.

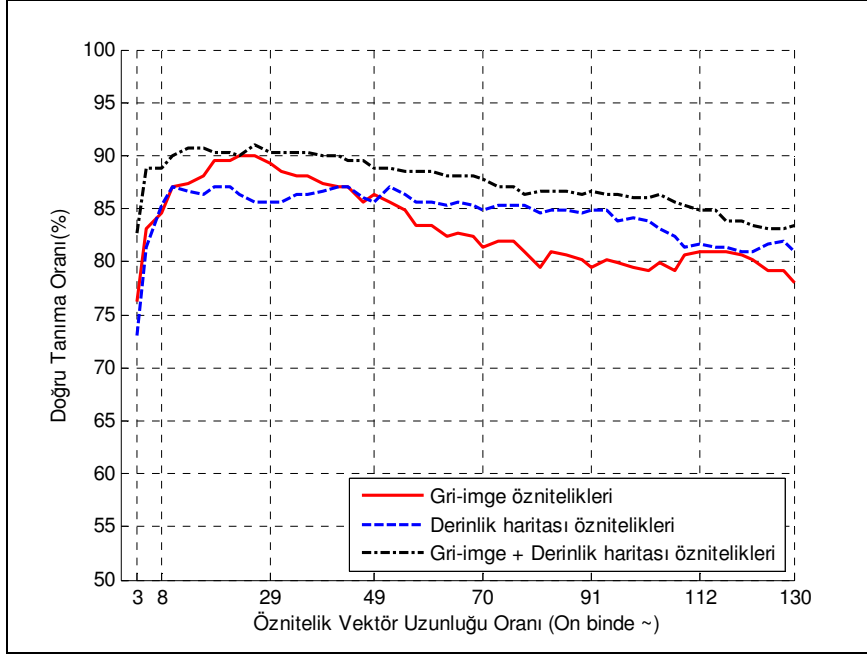
Çizelge 5. 12 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3 ve gri-imge + derinlik haritası özniteliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	10	60,43	40	82,73	160	89,21
5	20	72,66	80	88,85	320	90,29
8	30	79,86	120	88,85	480	89,93
10	40	83,45	160	89,93	640	89,57
13	50	85,97	200	90,65	800	89,57
16	60	86,33	240	90,65	960	89,93
18	70	87,77	280	90,29	1120	89,21
21	80	87,05	320	90,29	1280	88,85
23	90	88,13	360	89,93	1440	87,77
26	100	88,13	400	91,01	1600	86,69
29	110	88,49	440	90,29	1760	85,61
31	120	88,13	480	90,29	1920	84,89
34	130	87,77	520	90,29	2080	84,53
36	140	87,41	560	90,29	2240	83,45

Çizelge 5. 12 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
39	150	87,05	600	89,93	2400	82,01
42	160	87,41	640	89,93	2560	81,65
44	170	87,77	680	89,57	2720	81,65
47	180	87,77	720	89,57	2880	82,37
49	190	88,49	760	88,85	3040	82,37
52	200	87,77	800	88,85	3200	82,37
55	210	88,49	840	88,49	3360	82,01
57	220	89,21	880	88,49	3520	81,65
60	230	88,85	920	88,49	3680	81,65
63	240	88,13	960	88,13	3840	81,65
65	250	88,49	1000	88,13	4000	80,58

Şekil 5.13'te 320x240 çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 3. Kategoride yer alan 31 kelime üzerinde farklı özniteliklerin kullanımlarındaki başarımları gösterilmiştir.



Şekil 5. 13 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerde ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3).

Şekil 5.13'te gri-imege, derinlik haritası özniteliklerinin ayrı ayrı değerlendirilmesiyle gözlemlenen tanıma başarımları ve bu özniteliklerin bire bir (1:1) oranında birleştirildiği durumda gözlemlenen tanıma başarımları yer almaktadır. İlgili başarımlar grafiğinde en başarılı sonuçlara gri-imege+derinlik haritası özniteliklerinin kullanımında ulaşılmıştır. Bununla beraber derinlik haritası özniteliklerinin, gri-imege özniteliklerinden bazı öznitelik vektör uzunluklarında başarılı olduğu görülmüştür. Şekil 5.9'da 1. kategoriye ve şekil 5.11'de 2. kategoriye ait başarımlar grafiklerinde derinlik haritası öznitelikleri hemen her vektör boyutunda gri-imege özniteliklerine nazaran daha verimlidir. 3. kategoride ise düşük vektör boyutlarında gri-imege öznitelikleri verimliyken, vektör boyutunun büyümesiyle derinlik haritası öznitelikleri daha iyi verim sağlamaktadır. Şekil 5.13'teki başarımlar değerlerinin detayları ise çizelge 5.13'de gösterilmiştir.

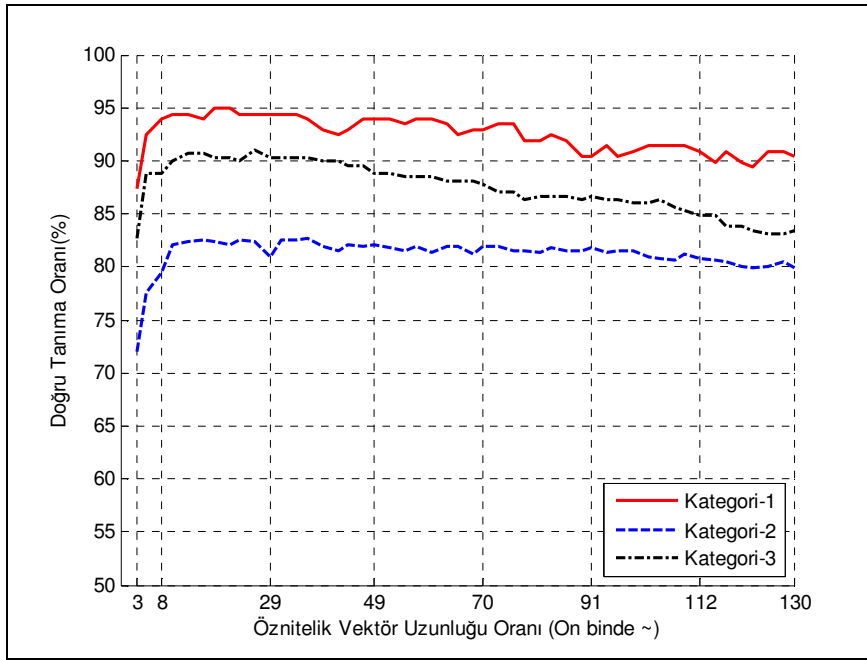
Çizelge 5. 13 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3).

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNİTELİKLERİ		DERİNLİK HARİTASI ÖZİNİTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNİTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	20	76,26	20	73,02	40	82,73
5	40	83,09	40	81,30	80	88,85
8	60	84,53	60	85,25	120	88,85
10	80	87,05	80	87,05	160	89,93
13	100	87,41	100	86,69	200	90,65
16	120	88,13	120	86,33	240	90,65
18	140	89,57	140	87,05	280	90,29
21	160	89,57	160	87,05	320	90,29
23	180	89,93	180	86,33	360	89,93
26	200	89,93	200	85,61	400	91,01
29	220	89,21	220	85,61	440	90,29
31	240	88,49	240	85,61	480	90,29
34	260	88,13	260	86,33	520	90,29
36	280	88,13	280	86,33	560	90,29
39	300	87,41	300	86,69	600	89,93
42	320	87,05	320	87,05	640	89,93
44	340	87,05	340	87,05	680	89,57
47	360	85,61	360	85,97	720	89,57
49	380	86,33	380	85,61	760	88,85
52	400	85,61	400	87,05	800	88,85

Çizelge 5. 13 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNİTELİKLERİ		DERİNLİK HARİTASI ÖZİNİTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNİTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
55	420	84,89	420	86,33	840	88,49
57	440	83,45	440	85,61	880	88,49
60	460	83,45	460	85,61	920	88,49
63	480	82,37	480	85,25	960	88,13
65	500	82,73	500	85,61	1000	88,13

Şekil 5.14'te 320x240 çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının her bir kategorideki başarımları gösterilmiştir.



Şekil 5. 14 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 3 kategoride yer alan 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (gri-imge + derinlik haritası öznitelikleri).

Şekil 5.14'te TİD veri setindeki her kategorinin 320x240 çerçeve boyutlarındaki başarımları yer almaktadır. Başarım değerleri gri-imge + derinlik haritası özniteliklerinin kullanımında hesaplanmıştır. Sırasıyla en başarılı sonuçlar 1. kategori, 3. kategori ve 2.

kategoride gözlemlenmiştir. 2. kategoriye ait başarımların düşük olmasının sebebi; bu kategorinin diğer kategorilere nazaran daha fazla kelime içermesi ve bazı kelimelerin birbirlerine benzemeleridir. Şekil 5.14'te yer alan başarımların oranlarının detayları ise çizelge 5.14'te verilmiştir.

Çizelge 5. 14 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 3 kategoride yer alan 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (gri-imege + derinlik haritası öznitelikleri).

DCT KATSAYI ORANI (~ ON BİNDE)	KATEGORİ - 1		KATEGORİ - 2		KATEGORİ - 3	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	40	87,37	40	72,05	40	82,73
5	80	92,42	80	77,57	80	88,85
8	120	93,94	120	79,47	120	88,85
10	160	94,44	160	82,13	160	89,93
13	200	94,44	200	82,32	200	90,65
16	240	93,94	240	82,51	240	90,65
18	280	94,95	280	82,32	280	90,29
21	320	94,95	320	82,13	320	90,29
23	360	94,44	360	82,51	360	89,93
26	400	94,44	400	82,32	400	91,01
29	440	94,44	440	80,99	440	90,29
31	480	94,44	480	82,51	480	90,29
34	520	94,44	520	82,51	520	90,29
36	560	93,94	560	82,70	560	90,29
39	600	92,93	600	81,94	600	89,93

Çizelge 5. 14 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	KATEGORİ - 1		KATEGORİ - 2		KATEGORİ - 3	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
42	640	92,42	640	81,56	640	89,93
44	680	92,93	680	82,13	680	89,57
47	720	93,94	720	81,94	720	89,57
49	760	93,94	760	82,13	760	88,85
52	800	93,94	800	81,75	800	88,85
55	840	93,43	840	81,56	840	88,49
57	880	93,94	880	81,94	880	88,49
60	920	93,94	920	81,37	920	88,49
63	960	93,43	960	81,94	960	88,13
65	1000	92,42	1000	81,94	1000	88,13

Çizelge 5.15'te ise farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının her bir kategoride farklı özneliklerin kullanımlarındaki en iyi doğru tanıma oranları ve bu değerlere karşılık gelen DCT katsayı oranları gösterilmiştir.

Çizelge 5. 15 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki en başarılı doğru tanıma oranları (111 kelime ve gri-imge + derinlik haritası öznitelikleri).

ÖZİNİTELİK VEKTÖRÜNDE YER ALAN ÖZİNİTELİKLER (DCT KATSAYI ORANI ~ ON BİNDE)	DOĞRU TANIMA ORANLARI (%)					
	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT K.O.	CRR	DCT K.O.	CRR	DCT K.O.	CRR
GRI-İMGE	73	79,44	36	81,64	13	83,33
DERİNLİK	104	81,44	24	84,53	8	85,53
GRI-İMGE + DERİNLİK	99	86,43	26	88,62	5	89,82

Tez çalışmasında önerilen işaret dili tanıma sisteminde hareket farklarına dayalı toplamsal imge yaklaşımıyla çıkarılan özniteliklerin 2D DCT ile uzamsal uzaya aktarılmasıyla oluşturulan öznitelik vektörleriyle, çalışma kapsamında oluşturulan Türk İşaret Dili veri setinde başarılı sonuçlar alınmıştır. Toplam 3 kategoride toplanmış 111 kelimedeki uygulanan yöntem ile, farklı çerçeve boyutlarında ve farklı DCT katsayı oranlarında performans ölçümleri yapılmıştır. Mevcut veri setinde bütün kelimelerin birlikte değerlendirilmesinin yanı sıra kategoriler ayrı ayrı da değerlendirmeye tabi tutulmuştur.

Tüm veri seti üzerinde yapılan testlerde genellikle DCT katsayı oranının değişimine göre 320x240 boyutlu çerçevelerde daha dengeli başarımlar gözlemlenmiştir. Bununla birlikte gri-imge toplamsal imgeleri üzerinden çıkarılan özniteliklerin başarılı olduğu, derinlik haritaları üzerinden elde edilen özniteliklerin ise gri-imgelere nazaran daha başarılı olduğu görülmüştür. Ayrıca gri-imge ve derinlik haritalarına ait özniteliklerin birlikte kullanımının genelde en başarılı sonuçları verdiği gözlemlenmiştir. Kategorilerin ayrı ayrı değerlendirilmesinde ise en başarılı sonuçlara 1. ve 3. kategoride ulaşılmıştır. 2. kategoride diğer kategorilere nazaran gözlemlenen düşük başarımlar, bu kategoride yer alan kelime sayısının diğer kategorilerden fazla olması ve kategoride yer alan kelime işaretlerinin bir kısmının birbirine çok benzemesine bağlanmıştır.

Tez çalışmasında önerilen sistemde gri-imege ve derinlik haritalarında toplamsal imge yaklaşımı ile birlikte 2D DCT kullanılmıştır. Bu yaklaşıma ek olarak mevcut sistem yapısının bazı kısımlarında farklı yaklaşımlar da uygulanarak sistem performansının değişiminin gözlemlenmesi ve sistemin iyileştirilmesi amaçlanmıştır. İlerleyen kısımlarda sistemde denenen diğer yaklaşımlar gözlemlenen sonuçlarla birlikte detaylı olarak verilmiştir.

5.2.1 Toplamsal İmge Yaklaşımında Çerçeve Ağırlıklandırma

Önerilen sistemde, zamansal özniteliklerin çıkarımı aşamasında çerçeve farklarından elde edilen ikili imgeler toplandıktan sonra, işlenen görüntüye ait toplam çerçeve sayısına bölünerek, normalleştirme uygulanmıştır. Başka bir ifadeyle toplamsal imgelerin parçaları sabit bir katsayı ile ağırlıklandırılmıştır. Bu işlem eşitlik (5.2)'deki şekilde ifade edilmiştir.

$$I_T = \sum_{i=1}^{n-1} I_{Ei} \cdot w_i \quad w_i = \frac{1}{n} \quad (5.2)$$

Eşitlik (5.2)'te n toplam çerçeve sayısı olmak üzere, w_i eşiklenmiş fark çerçevelerinin ağırlıkları ve I_T ise toplamsal imgedir. w_i ağırlık değerleri bütün eşiklenmiş imgeler için çerçeve sayısına bağlı olarak eşit olarak seçilmiştir ve I_{Ei} eşiklenmiş ikili imgelerine uygulanmıştır.

Tez çalışmasında, bu yaklaşımla birlikte farklı bir çerçeve ağırlıklandırma yöntemi daha izlenmiştir. Bir boyutlu Gauss fonksiyonunu temel alan bu yaklaşımda ikili imgelerin ağırlıkları 1D Gauss fonksiyonuna göre belirlenmiştir.

$$w(i) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-(i-\mu)^2/2\sigma^2} \quad (i = 1, \dots, n-1) \quad (5.2)$$

Eşitlik (5.2)'de yer alan 1D Gauss fonksiyonu normal olarak dağılmış ağırlık katsayıları sağlamaktadır. Fonksiyon aracılığı ile üretilen ağırlıklar sırasıyla işaret görüntülerine ait ikili imgelere uygulanarak toplamsal imge üzerinde, işaretlerin başlangıç ve bitişlerindeki çerçevelerin etkilerinin azaltılarak orta kısımlarda kalan çerçevelerin

etkilerinin arttırılması hedeflenmiştir. (5.2)'deki eşitlikte $\mu = n/2$ olarak alınırken $\sigma = n/4$ olarak alınmıştır.

İlgili yaklaşımında devamında toplamsal imgelerin elde edilmesi sağlanmış ve 2D DCT ile uzamsal öznitelikler çıkarılmıştır. Toplamsal imge çıkarımında 1D Gauss fonksiyonuna bağlı ağırlıklandırma yöntemi ile elde edilen sonuçlar çizelge 5.16'da gösterilmiştir.

Çizelge 5. 16 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının gauss temelli ağırlıklandırma yaklaşımında farklı kategoriler için doğru tanıma oranları (dct katsayı oranı : % 0,52).

	160x120 ÇERÇEVE			320x240 ÇERÇEVE			640x480 ÇERÇEVE		
	K-1	K-2	K-3	K-1	K-2	K-3	K-1	K-2	K-3
G	85,86	68,25	82,37	86,36	71,10	82,01	80,30	63,31	66,55
D	89,39	72,43	80,58	89,90	75,67	82,73	82,32	70,91	71,58
G+D	89,90	77,00	87,05	91,92	78,14	84,17	84,85	73,95	72,66

Çizelge 5.16'daki verilere göre, genel olarak her çerçeve boyutunda ve her kategoride derinlik haritası öznitelikleri (D) gri-imge özniteliklerinden (G) daha başarılı sonuçlar vermektedir. Aynı zamanda bu özniteliklerin birlikte kullanımı (G+D) tek başına kullanımlarından daha iyi başarımlar vermektedir. Önerilen sistemde belirtilen sabit ağırlıklandırmaya dayalı yöntemde, 2D DCT ile çıkarılan özniteliklerin % 0,52'si değerlendirmeye alındığında gözlemlenen sonuçlar ise çizelge 5.17'de sunulmuştur.

Çizelge 5. 17 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının sabit katsayı ile ağırlıklandırma yaklaşımında farklı kategoriler için doğru tanıma oranları (dct katsayı oranı : % 0,52).

	160x120 ÇERÇEVE			320x240 ÇERÇEVE			640x480 ÇERÇEVE		
	K-1	K-2	K-3	K-1	K-2	K-3	K-1	K-2	K-3
G	87,88	69,58	85,97	90,40	75,48	85,61	84,34	72,62	74,82
D	92,42	75,85	85,25	91,92	78,71	87,05	82,32	77,19	77,69
G+D	92,93	81,94	87,77	93,94	81,75	88,85	87,88	78,52	82,37

Çizelge 5.17'deki verilere göre, genel olarak her çerçeve boyutunda ve her kategoride derinlik haritası öznitelikleri (D) gri-imege özniteliklerinden (G) daha başarılı sonuçlar vermektedir. Aynı zamanda bu özniteliklerin birlikte kullanımı (G+D) tek başarısına kullanımlarından daha iyi başarımlar değerleri vermektedir. Çizelge 5.17 ile çizelge 5.16 karşılaştırılınca, sabit katsayı ile ağırlıklandırma yaklaşımının Gauss temelli ağırlıklandırma yaklaşımına nazaran daha performanslı olduğu görülmektedir.

5.2.2 Çerçeve Gruplama ve Gruplamalı Toplamsal İmge Yaklaşımı

Zamansal özniteliklerin çıkarılması amacıyla toplamsal imgelerin yapılandırılması aşamasında uygulanan diğer bir yaklaşım ise çerçevelerin önceden belirlenen bir grup sayısı parametresine göre gruplandırılması olmuştur. Bu yaklaşımda işaret görüntülerine ait çerçeveler, her grupta eşit oranda çerçeve olacak biçimde bölümlenmiştir. Daha sonraki aşamada ise her grupta yer alan çerçeveler üzerinden sabit ağırlıklandırma ile toplamsal imgeler oluşturulmuştur. Bu bağlamda görüntü çerçevelerinin ayrıldığı grup sayısı kadar toplamsal imge elde edilmiş ve her bir toplamsal imgeye 2D DCT uygulanarak toplamsal imgelere karşılık gelen uzamsal öznitelikler çıkarılmıştır. Öznitelik vektörlerinin oluşturulması aşamasında ise her bir toplamsal imgeden elde edilen DCT katsayıları eşit oranda birleştirilmiştir. Eşitlik (5.3) her çerçeve gruplama yaklaşımında toplamsal imgelerin elde edilmesine ait ifadeyi yansıtmaktadır. Eşitlikte yer alan g sembolü grup sayısına denk düşmektedir.

$$I_{Ti} = \sum_{j=1}^{n_i-1} I_{Ej} \cdot w_j \quad w_j = \frac{1}{n_i} \quad (i = 1, 2, \dots, g) \quad (5.3)$$

Bu yaklaşımla, işaretlere ait zamansal bilgiler parçalanıp gruplanarak, zamansal özniteliklerin daha detaylı ve belirgin olarak çıkarılması hedeflenmiştir. Çizelge 5.18'de bu yaklaşımla gözlemlenen sonuçlar verilmiştir.

Çizelge 5. 18 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırmaya dayalı toplamsal imge yaklaşımında farklı kategoriler için doğru tanıma oranları (toplam dct katsayı oranı: % 0,52, grup sayısı: 5).

	160x120 ÇERÇEVE			320x240 ÇERÇEVE			640x480 ÇERÇEVE		
	K-1	K-2	K-3	K-1	K-2	K-3	K-1	K-2	K-3
G	69,70	47,53	56,83	75,25	54,75	62,59	78,79	57,60	60,07
D	63,64	44,30	48,56	71,21	57,22	60,43	76,77	67,68	64,03
G+D	73,74	50,57	60,79	77,78	59,51	65,11	80,81	68,63	67,27

Çizelge 5.18'deki verilere göre, genel olarak her çerçeve boyutunda ve her kategoride gri-imge öznitelikleri (G) derinlik haritası özniteliklerinden (D) daha başarılı sonuçlar vermektedir. Aynı zamanda bu özniteliklerin birlikte kullanımı (G+D) tek başına kullanımlarından daha iyi başarımlar vermektedir. Sonuçları çizelge 5.18'de verilen çerçeve gruplandırmaya dayalı yaklaşım, normal yaklaşıma göre (çizelge 5.17) düşük performans sergilemektedir.

Sonuçları çizelge 5.18'de gösterilen yaklaşımda, ilk ve son gruplardaki toplamsal imgelere ait özniteliklerin öznitelik vektörlerinde kullanılmaması sonucu elde edilen doğru tanıma oranları ise çizelge 5.19'da verilmiştir. İlk ve son grupların değerlendirilmeye katılmaması ile grup sayısının iki eksiği kadar toplamsal imge kullanılmıştır ve bu toplamsal imgelerden elde edilen özniteliklerden, öznitelik vektörüne eşit oranda katılım yapılmıştır.

Çizelge 5. 19 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırmaya ve grup eksiltmeye dayalı toplamsal imge yaklaşımında farklı kategoriler için doğru tanıma oranları (toplam dct katsayı oranı: % 0,52, grup sayısı: 5).

	160x120 ÇERÇEVE			320x240 ÇERÇEVE			640x480 ÇERÇEVE		
	K-1	K-2	K-3	K-1	K-2	K-3	K-1	K-2	K-3
G	78,28	55,51	65,47	87,37	65,59	72,66	83,84	62,36	65,47
D	68,69	55,89	63,31	73,23	73,38	70,14	79,80	72,24	68,71
G+D	83,33	63,50	71,58	85,86	75,48	75,18	85,35	74,71	70,50

Çizelge 5.19'daki verilere göre, genel olarak her çerçeve boyutunda ve her kategoride gri-imege öznitelikleri (G) derinlik haritası özniteliklerinden (D) daha başarılı sonuçlar vermektedir. Aynı zamanda bu özniteliklerin birlikte kullanımı (G+D) tek başarına kullanımlarından daha iyi başarımlar değerleri vermektedir. Sonuçları çizelge 5.19'da verilen çerçeve gruplandırmaya dayalı yaklaşım, normal yaklaşıma göre (çizelge 5.17) düşük performans sergilemektedir.

İşaret görüntülerine ait çerçeve gruplama yönteminde, birden çok toplamsal imgenin oluşturulduğu yaklaşımlarda öznitelik vektörü oluşturulurken, her toplamsal imgeye ait özniteliklerden eşit oranlarda katılım yapılmıştır. Fakat her gruba ait özniteliklerden alınacak olan öznitelik miktarının tanıma performansına etkileri söz konusu olabilmektedir. Bu bağlamda gruplardan seçilecek özniteliklerinin yani DCT katsayılarının sayısını belirlemek amacıyla yine eşitlik (5.2)'de ifade edilen 1D Gauss fonksiyonundan faydalanılmıştır. Fonksiyonda toplam çerçeve sayısı parametresi yerine grup sayısı parametresi konularak her grup için dağılımdaki olasılıkları hesaplanmıştır ve gruplardaki öznitelikler bu olasılık değerlerinin doğru oranındaki sayıda öznitelik vektörüne aktarılmışlardır. Diğer bir ifade ile ilk ve son gruplardan öznitelik vektörüne seçilen öznitelikler az iken, ortada yer alan gruplardan seçilen öznitelikler daha fazladır. Bu yaklaşımla elde edilen sonuçlar ise çizelge 5.20'de sunulmuştur.

Çizelge 5. 20 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırmaya ve gruplardan dağılıma bağlı öznitelik seçmeye dayalı toplamsal imge yaklaşımında farklı kategoriler için doğru tanıma oranları (toplam dct katsayı oranı: % 0,52, grup sayısı: 5).

	160x120 ÇERÇEVE			320x240 ÇERÇEVE			640x480 ÇERÇEVE		
	K-1	K-2	K-3	K-1	K-2	K-3	K-1	K-2	K-3
G	74,75	54,37	61,87	84,85	64,45	71,22	83,84	61,22	65,47
D	71,72	55,70	60,07	75,76	70,91	68,35	82,32	72,24	66,19
G+D	78,28	59,13	70,50	83,33	69,39	71,94	85,86	73,38	69,78

Çizelge 5.20'deki verilere göre, genel olarak her çerçeve boyutunda ve her kategoride gri-imege öznitelikleri (G) derinlik haritası özniteliklerinden (D) daha başarılı sonuçlar vermektedir. Aynı zamanda bu özniteliklerin birlikte kullanımı (G+D) tek başarına

kullanımlarından daha iyi başarımlar değerleri vermektedir. Sonuçları çizelge 5.20’de verilen çerçeve gruplandırmaya dayalı yaklaşım, normal yaklaşıma göre (çizelge 5.17) düşük performans sergilemektedir.

Gruplamalı olarak toplamsal imge oluşturulmasına dayalı olarak uygulanan diğer bir yöntem ise, işaret görüntülerinde hareketin başladığı ve tamamlandığı zaman aralıklarındaki görüntü çerçevelerinin öznitelik çıkarımı aşamasına dâhil edilmemesi olmuştur. Bu yaklaşımda ise, yine görüntülere ait toplam çerçeve sayısı belli sayıda gruba ayrılmıştır. Ancak öznitelik çıkarımı aşamasında hareketin başlangıç ve bitiş anlarındaki çerçeveleri içermesi dolayısıyla, ilk ve son gruplarda yer alan çerçeveler öznitelik çıkarımına dâhil edilmemiştir. Diğer gruplarda bulunan çerçeveler ise birlikte değerlendirilerek, bu çerçevelerden tek bir toplamsal imge edilmiş ve bu tek toplamsal imgeden çıkarılan öznitelikler değerlendirilmeye alınmıştır. Bu bağlamda yapılan deneylere ait sonuçlar ise çizelge 5.21’de belirtilmiştir.

Çizelge 5. 21 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırmaya ve çerçeve eksiltmeye dayalı toplamsal imge yaklaşımında farklı kategoriler için doğru tanıma oranları (toplam dct katsayı oranı: % 0,52, grup sayısı: 5).

	160x120 ÇERÇEVE			320x240 ÇERÇEVE			640x480 ÇERÇEVE		
	K-1	K-2	K-3	K-1	K-2	K-3	K-1	K-2	K-3
G	85,35	65,97	82,73	87,37	71,67	82,01	80,81	66,16	69,06
D	86,87	72,81	81,29	87,88	75,86	82,01	84,34	75,86	78,42
G+D	91,41	77,76	86,69	91,41	79,28	86,69	88,89	77,76	79,50

Çizelge 5.21’deki verilere göre, genel olarak her çerçeve boyutunda ve her kategoride derinlik haritası öznitelikleri (D) gri-imge özniteliklerinden (G) daha başarılı sonuçlar vermektedir. Aynı zamanda bu özniteliklerin birlikte kullanımı (G+D) tek başarımlar kullanımlarından daha iyi başarımlar değerleri vermektedir. Sonuçları çizelge 5.21’de verilen çerçeve gruplandırmaya dayalı yaklaşım, normal yaklaşıma (çizelge 5.17) yakın bir performans sergilemektedir ve çizelge 5.18’deki, çizelge 5.19’daki ve çizelge 5.20’deki çerçeve gruplandırmaya dayalı yaklaşımlardan başarılıdır.

Gruplamalı toplamsal imge yaklaşımında ikili imge ağırlıklandırmanın performanslarını gözlemlemek amacıyla da farklı yaklaşımlar sergilenmiş ve bu bağlamda deneyler yapılmıştır. Klasik gruplama yaklaşımıyla beraber, gruplarda yer alan ikili imgeleri ağırlıklandırmak için eşitlik (5.2)'de yer alan 1D Gauss fonksiyonundan yararlanılmıştır. Ancak ikili imgelerin ağırlıklandırılmasında her grupta yer alan çerçeve sayısı değil de, görüntüde yer alan toplam çerçeve sayısı baz alınmıştır ve dağılım katsayıları tüm çerçeveleri kapsayacak biçimde oluşturulmuştur. Böylece her grupta yer alan ikili imgeler genel dağılımın bir bölümüyle ağırlıklandırılmıştır. İşlemlerin devamında her gruptan elde edilen toplamsal imgelere 2D DCT uygulanmış ve öznitelikler çıkarılmıştır ve öznitelik vektörleri gruplara ait özniteliklerden eşit oranda seçilerek oluşturulmuştur.

Bu yaklaşımla elde edilen sonuçlar çizelge 5.22'de sunulmuştur.

Çizelge 5. 22 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırma dayalı toplamsal imge ve genel gauss dağılımı temelli yaklaşımda farklı kategoriler için doğru tanıma oranları (toplam dct katsayı oranı: % 0,52, grup sayısı: 5).

	160x120 ÇERÇEVE			320x240 ÇERÇEVE			640x480 ÇERÇEVE		
	K-1	K-2	K-3	K-1	K-2	K-3	K-1	K-2	K-3
G	67,17	44,30	58,99	74,75	53,04	60,79	75,25	53,04	53,24
D	62,63	41,83	51,08	73,74	56,27	59,71	71,72	61,41	60,79
G+D	73,23	47,15	59,35	74,24	57,22	65,83	78,28	63,31	62,23

Çizelge 5.22'deki verilere göre, genel olarak her çerçeve boyutunda ve her kategoride gri-imge öznitelikleri (G) derinlik haritası özniteliklerinden (D) daha başarılı sonuçlar vermektedir. Aynı zamanda bu özniteliklerin birlikte kullanımı (G+D) tek başına kullanımlarından daha iyi başarımlar değerleri vermektedir. Sonuçları çizelge 5.22'de verilen çerçeve gruplandırma dayalı yaklaşım, normal yaklaşıma göre (çizelge 5.17) ve sonuçları çizelge 5.21'de gösterilen çerçeve gruplandırma dayalı yaklaşıma göre düşük performans sergilemektedir.

Çizelge 5.22'de sonuçları verilen yaklaşıma ek olarak, Gauss dağılımı, görüntüde yer alan toplam çerçeve sayısını temel alan genel dağılım yerine grup sayısı bazında

oluşturulmuştur. Çerçevelerin bölümlendiği grup sayısına göre oluşturulan dağılım katsayıları, her gruptan sabit ağırlıklandırma ile oluşturulmuş toplamsal imgelerle çarpılarak, gruplamalı toplamsal imgeler normleştirilmiştir. Sürecin devamında ise her gruptan elde edilen toplamsal imgelere 2D DCT uygulanmış ve öznitelikler çıkarılmıştır. Öznitelik vektörleri gruplara ait özniteliklerden eşit oranda seçilerek oluşturulmuştur. Bu yaklaşımla elde edilen sonuçlar ise çizelge 5.23’de ifade edilmiştir.

Çizelge 5. 23 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının çerçeve gruplandırmaya dayalı toplamsal imge ve gruplamalı gauss dağılımı temelli yaklaşımda farklı kategoriler için doğru tanıma oranları (toplam dct katsayı oranı: % 0,52, grup sayısı: 5).

	160x120 ÇERÇEVE			320x240 ÇERÇEVE			640x480 ÇERÇEVE		
	K-1	K-2	K-3	K-1	K-2	K-3	K-1	K-2	K-3
G	69,70	47,53	56,83	75,25	54,75	62,59	78,79	57,60	60,07
D	63,64	44,30	48,56	71,21	57,22	60,43	76,77	67,68	64,03
G+D	73,74	50,57	60,79	77,78	59,51	65,11	80,81	68,63	67,27

Çizelge 5.23’teki verilere göre, genel olarak her çerçeve boyutunda ve her kategoride gri-imge öznitelikleri (G) derinlik haritası özniteliklerinden (D) daha başarılı sonuçlar vermektedir. Aynı zamanda bu özniteliklerin birlikte kullanımı (G+D) tek başına kullanımlarından daha iyi başarımlar değerleri vermektedir. Sonuçları çizelge 5.23’te verilen çerçeve gruplandırmaya dayalı yaklaşım, normal yaklaşıma göre (çizelge 5.17) ve sonuçları çizelge 5.21’de gösterilen çerçeve gruplandırmaya dayalı yaklaşıma göre düşük performans sergilemektedir.

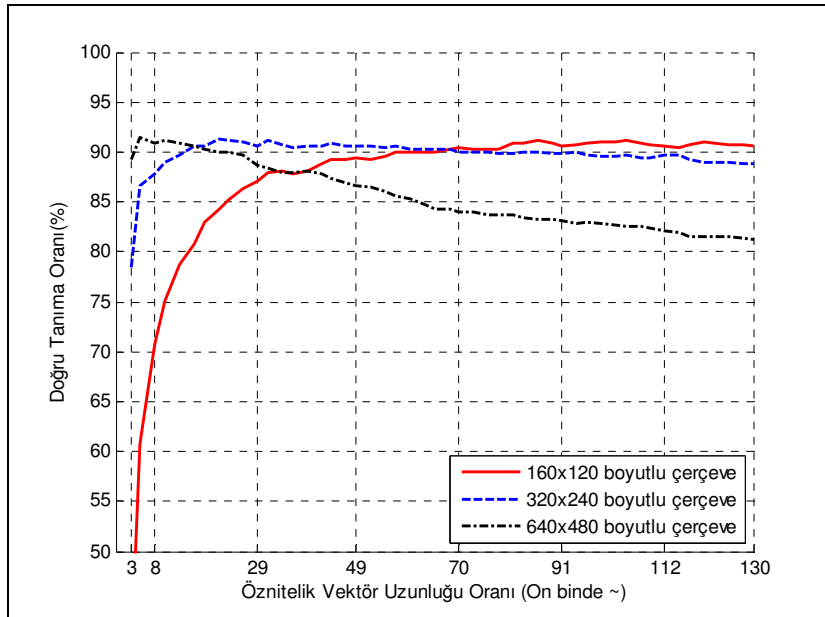
5.2.3 Toplamsal İmgelerde Ortalama Eşik Katsayısı Yaklaşımı

Önerilen işaret dili tanıma sisteminde toplamsal imgeler oluşturulurken öncelikle bir işaret video görüntüsüne ait ardışıl video çerçeveleri arasındaki farklar bir eşik katsayısı kullanılarak eşiklenmiş ve ikili imgelere dönüştürülmüştür. Sistemin bu aşamasında kullanılan sabit eşik katsayıları gri-imgeler ve derinlik haritaları için deneysel gözlemlere bağlı olarak ayrı ayrı belirlenmiştir.

Bu aşamada, eşik katsayılarını sabit olarak belirlemek yerine çerçeve farklarına bağlı olarak otomatik olarak belirlenmesi amacıyla ortalama eşik katsayısı yaklaşımı tercih edilmiştir. Bu yaklaşımla, ardışıl iki video çerçevesi arasındaki fark alındıktan sonra mutlak parlaklık değeri sıfırdan farklı olan piksellerin ortalama parlaklık değeri bulunur ve fark imgesini ikilileştirmek için kullanılacak olan eşik katsayısı olarak bu ortalama parlaklık değeri kullanılır. Bu işlem bütün ardışıl video çerçeveleri boyunca uygulanır ve her fark imgesi için otomatik olarak bir eşik katsayısı hesaplanır.

Bu başlık altında yer verilen başarımların sınamaları bu yaklaşımın ilgili veri setinde uygulanmasıyla gözlemlenen sonuçları kapsamaktadır. Bu yaklaşım uygulanırken işaretlere ait ardışıl video çerçeve farklarından eşikleme ile ikili imgeler elde edilmiştir. İkili imgelere sabit katsayı ile ağırlıklandırma uygulanarak tek bir toplamsal imge elde edilmiştir. Daha sonraki aşamalarda ise medyan filtre ve DCT işlemleri ile öznitelikler elde edilmiştir.

Şekil 5.15'te farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en (K=1) yakın komşu (KNN) sınıflandırıcısının veri setinde yer alan 111 kelime üzerindeki başarımları gösterilmiştir.



Şekil 5. 15 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime ve gri-imge + derinlik haritası özniteliklerinde).

Şekil 5.15'te farklı çerçeve boyutlarında gözlemlenen başarımlar sunulmuştur. İlgili testte bütün çerçeve boyutlarında başarılı sonuçlar gözlemlenmiştir. Ayrıca düşük boyutlu öznitelik vektöründe en iyi performansı 640x480 çerçeve boyutlu görüntüler sergilemektedir. Öznitelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarımlar artmaktadır ve 320x240 boyutlu çerçevelerde ise daha kararlı yapı sergilemektedir. Öznitelik vektör boyutunun arttırılması 640x480 çerçevelerin performansını ise oldukça olumsuz etkilemektedir. İlgili başarımlar grafiğinde 320x240 boyutlu çerçevelerin daha kararlı bir performans sergilemesi sebebiyle bazı testler özel olarak bu çerçeve boyutları için yapılmıştır. Şekil 5.15'teki başarımlar değerlerinin detayları ise çizelge 5.24'te gösterilmiştir.

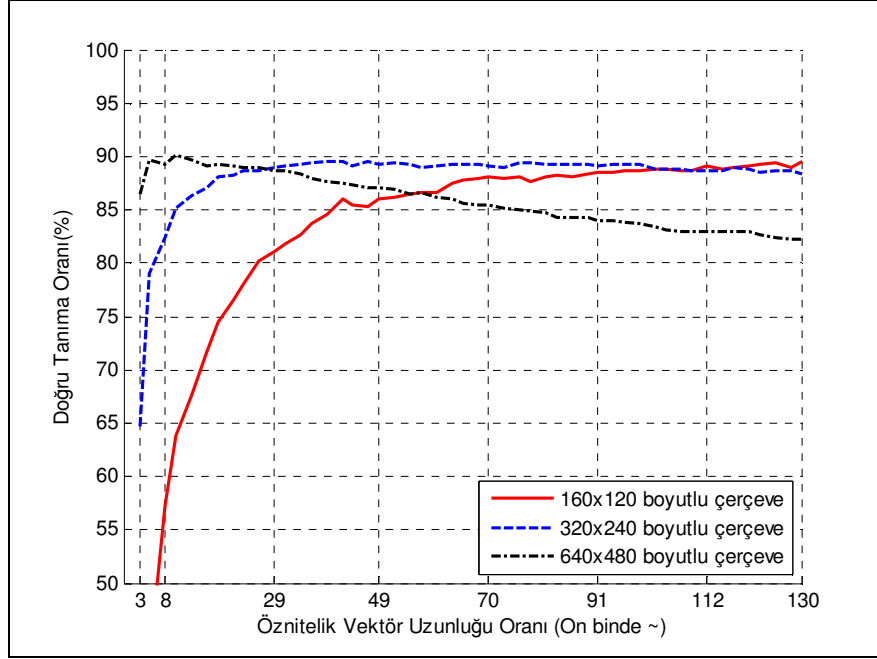
Çizelge 5. 24 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime ve gri-imege + derinlik haritası özniteliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	10	40,82	40	78,44	160	89,32
5	20	60,78	80	86,63	320	91,52
8	30	70,66	120	87,72	480	90,92
10	40	75,05	160	88,92	640	91,22
13	50	78,74	200	89,72	800	90,92
16	60	80,84	240	90,62	960	90,62
18	70	83,03	280	90,62	1120	90,32
21	80	84,33	320	91,32	1280	89,92
23	90	85,23	360	91,12	1440	89,92
26	100	86,33	400	91,02	1600	89,72
29	110	87,03	440	90,62	1760	88,72

Çizelge 5. 24 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
31	120	87,92	480	91,12	1920	88,32
34	130	88,02	520	90,72	2080	87,92
36	140	87,72	560	90,42	2240	87,92
39	150	88,12	600	90,52	2400	88,02
42	160	88,82	640	90,52	2560	87,72
44	170	89,22	680	90,82	2720	87,43
47	180	89,32	720	90,52	2880	86,93
49	190	89,42	760	90,62	3040	86,63
52	200	89,32	800	90,62	3200	86,53
55	210	89,52	840	90,42	3360	86,03
57	220	90,02	880	90,62	3520	85,63
60	230	89,92	920	90,32	3680	85,33
63	240	89,92	960	90,32	3840	84,73
65	250	90,02	1000	90,22	4000	84,33

Şekil 5.16'da farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının veri setinde yer alan 111 kelime üzerinde sadece gri-imge öznitelikleri kullanıldığındaki başarımları gösterilmiştir.



Şekil 5. 16 Tid veri setinde manhattan mesafesini kullanan knn ($k=1$) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşıkyle 111 kelime ve gri-imege özniteliklerinde).

Şekil 5.16'da farklı çerçeve boyutlarında sadece gri-imgelere ait özniteliklerin kullanılmasıyla gözlemlenen başarımlar sunulmuştur. Elde edilen başarımların değerlerinin karakteristikleri şekil 5.15 ile büyük oranda benzerlik göstermektedir. Düşük boyutlu öznitelik vektöründe en iyi performansı yine 640x480 çerçeve boyutlu görüntüler sergilemektedir. Öznitelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarımlar artmaktadır ve 320x240 boyutlu çerçevelerde ise başarımların artışıyla beraber bir noktadan itibaren başarımlar genel olarak sabit kalmaktadır. Öznitelik vektör boyutunun arttırılması 640x480 çerçevelerin performansını ise oldukça olumsuz etkilemektedir. Şekil 5.16'daki başarımların detayları ise çizelge 5.25'te gösterilmiştir.

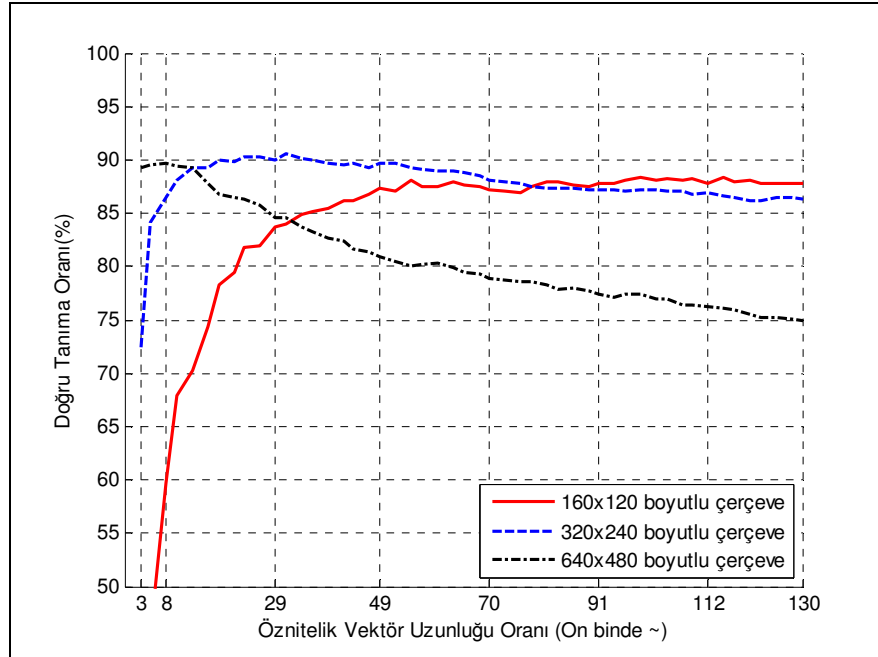
Çizelge 5. 25 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımla 111 kelime ve gri-imege özneliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	5	25,65	20	64,67	80	86,43
5	10	43,81	40	79,04	160	89,62
8	15	57,19	60	82,34	240	89,22
10	20	63,87	80	85,13	320	90,12
13	25	67,66	100	86,33	400	89,72
16	30	71,66	120	87,13	480	89,12
18	35	74,55	140	88,12	560	89,22
21	40	76,55	160	88,22	640	89,12
23	45	77,94	180	88,72	720	89,02
26	50	80,24	200	88,72	800	89,02
29	55	81,14	220	89,02	880	88,72
31	60	81,74	240	89,12	960	88,62
34	65	82,63	260	89,32	1040	88,42
36	70	83,63	280	89,42	1120	87,92
39	75	84,53	300	89,52	1200	87,62
42	80	86,03	320	89,52	1280	87,53
44	85	85,43	340	89,12	1360	87,33
47	90	85,33	360	89,52	1440	87,13
49	95	86,03	380	89,22	1520	87,03

Çizelge 5. 25 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
52	100	86,23	400	89,42	1600	86,93
55	105	86,53	420	89,22	1680	86,53
57	110	86,63	440	89,02	1760	86,63
60	115	86,63	460	89,12	1840	86,13
63	120	87,53	480	89,22	1920	86,03
65	125	87,72	500	89,32	2000	85,63

Şekil 5.17'de farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının veri setinde yer alan 111 kelime üzerinde sadece derinlik haritası öznitelikleri kullanıldığındaki başarımları gösterilmiştir.



Şekil 5. 17 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime ve derinlik haritası özniteliklerinde).

Şekil 5.17’de farklı çerçeve boyutlarında sadece derinlik haritalarına ait özniteliklerin kullanılmasıyla gözlemlenen başarımlar sunulmuştur. Düşük boyutlu öznitelik vektöründe en iyi performansı yine 640x480 çerçeve boyutlu görüntüler sergilemektedir. Öznitelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarımlar artmaktadır ve 320x240 boyutlu çerçevelerde ise genel olarak başarımlar düşmekle birlikte daha kararlı bir değişim sergilemektedir. Öznitelik vektör boyutunun arttırılması 640x480 çerçevelerin performansını ise oldukça olumsuz etkilemektedir. Şekil 5.16’daki gri-imgelere ait başarımlar ile şekil 5.17’deki derinlik haritalarına ait başarımlar karşılaştırıldığında, derinlik haritalarının işaretleri tanımda mevcut yaklaşım ile performanslı olduğu görülmektedir. Şekil 5.17’deki başarımların detayları ise çizelge 5.26’da gösterilmiştir.

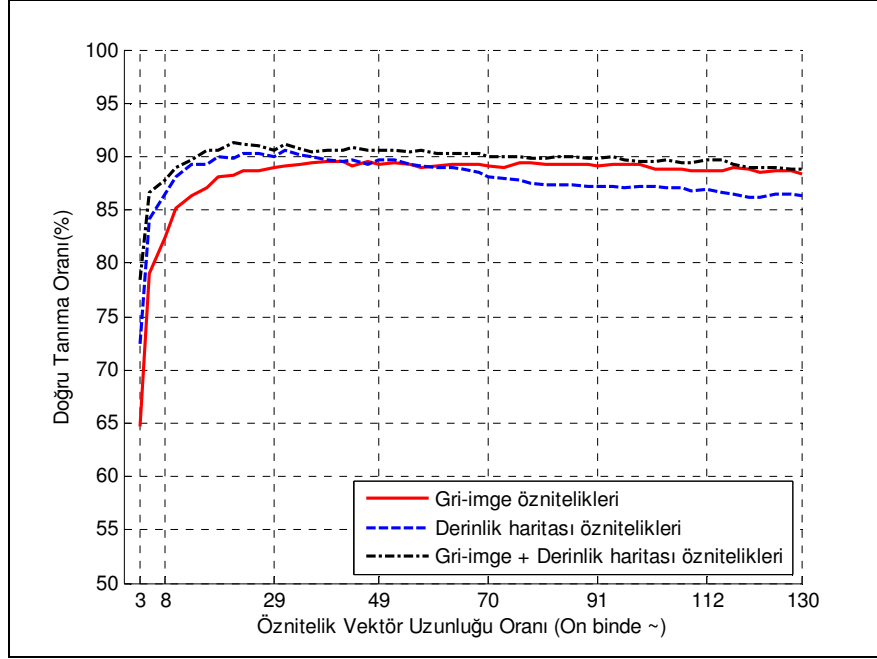
Çizelge 5. 26 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime ve derinlik haritası özniteliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	5	25,15	20	72,46	80	89,22
5	10	45,51	40	84,13	160	89,52
8	15	59,78	60	86,53	240	89,72
10	20	67,96	80	88,02	320	89,42
13	25	70,26	100	89,22	400	89,32
16	30	74,35	120	89,22	480	87,82
18	35	78,34	140	89,92	560	86,83
21	40	79,44	160	89,82	640	86,43
23	45	81,74	180	90,22	720	86,33
26	50	81,94	200	90,32	800	85,73

Çizelge 5. 26 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
29	55	83,73	220	90,02	880	84,63
31	60	83,93	240	90,52	960	84,53
34	65	84,83	260	90,12	1040	83,73
36	70	85,13	280	90,02	1120	83,23
39	75	85,43	300	89,72	1200	82,73
42	80	86,23	320	89,52	1280	82,34
44	85	86,23	340	89,62	1360	81,64
47	90	86,73	360	89,32	1440	81,34
49	95	87,33	380	89,62	1520	80,94
52	100	87,03	400	89,62	1600	80,54
55	105	88,02	420	89,22	1680	80,04
57	110	87,53	440	89,12	1760	80,24
60	115	87,53	460	89,02	1840	80,34
63	120	87,92	480	89,02	1920	79,84
65	125	87,62	500	88,82	2000	79,44

Şekil 5.18’de 320x240 çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının veri setinde yer alan 111 kelime üzerinde farklı özniteliklerin kullanımlarındaki başarımları gösterilmiştir.



Şekil 5. 18 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımla 111 kelime).

Şekil 5.18, 320x240 çerçeve boyutlarındaki görüntülerde farklı özniteliklerin tanıma başarımlarını göstermektedir. 320x240 çerçeve boyutunda tanıma başarımının önceki testlerde daha kararlı bir performans sergilemesi sebebiyle, ilgili başarıım grafiği 320x240 çerçeve boyutu için gösterilmiştir. Şekil 5.18’de gri-imege, derinlik haritası özniteliklerinin ayrı ayrı değerlendirilmesiyle gözlemlenen tanıma başarımları ve bu özniteliklerin bire bir (1:1) oranında birleştirildiği durumda gözlemlenen tanıma başarımları yer almaktadır. İlgili başarıım grafiğinde en başarılı sonuçlara gri-imege+derinlik haritası özniteliklerinin kullanımında ulaşılmıştır. Bununla beraber gri-imege özniteliklerinin, derinlik haritası özniteliklerinden başarılı olduğu da görülmüştür. Şekil 5.18’deki başarıım değerlerinin detayları ise çizelge 5.27’de gösterilmiştir.

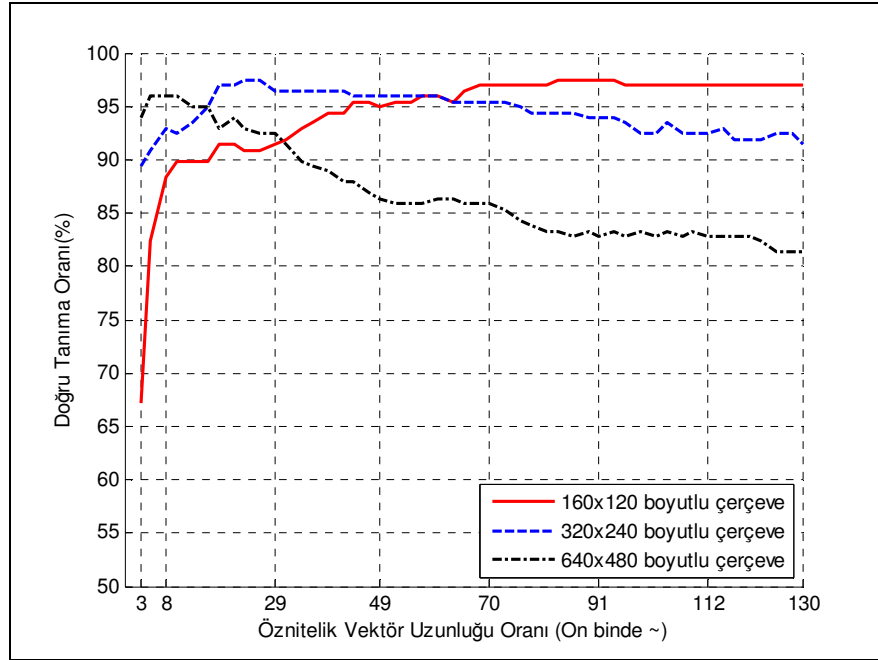
Çizelge 5. 27 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımıyla 111 kelime).

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNİTELİKLERİ		DERİNLİK HARİTASI ÖZİNİTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNİTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	20	64,67	20	72,46	40	78,44
5	40	79,04	40	84,13	80	86,63
8	60	82,34	60	86,53	120	87,72
10	80	85,13	80	88,02	160	88,92
13	100	86,33	100	89,22	200	89,72
16	120	87,13	120	89,22	240	90,62
18	140	88,12	140	89,92	280	90,62
21	160	88,22	160	89,82	320	91,32
23	180	88,72	180	90,22	360	91,12
26	200	88,72	200	90,32	400	91,02
29	220	89,02	220	90,02	440	90,62
31	240	89,12	240	90,52	480	91,12
34	260	89,32	260	90,12	520	90,72
36	280	89,42	280	90,02	560	90,42
39	300	89,52	300	89,72	600	90,52
42	320	89,52	320	89,52	640	90,52
44	340	89,12	340	89,62	680	90,82
47	360	89,52	360	89,32	720	90,52
49	380	89,22	380	89,62	760	90,62
52	400	89,42	400	89,62	800	90,62

Çizelge 5. 27 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNTELİKLERİ		DERİNLİK HARİTASI ÖZİNTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
55	420	89,22	420	89,22	840	90,42
57	440	89,02	440	89,12	880	90,62
60	460	89,12	460	89,02	920	90,32
63	480	89,22	480	89,02	960	90,32
65	500	89,32	500	88,82	1000	90,22

Şekil 5.19'da farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 1. kategoride yer alan 22 kelime üzerindeki başarımları gösterilmiştir.



Şekil 5. 19 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1 ve ortalama eşik katsayısı yaklaşımıyla gri-imge + derinlik haritası öz niteliklerinde).

Şekil 5.19'da TİD veri setindeki 1. kategoride yer alan kelimelerin farklı çerçeve boyutlarındaki başarımları yer almaktadır. Başarım değerleri gri-imge + derinlik haritası özniteliklerinin kullanımında hesaplanmış olup, en iyi başarım değerlerine 160x120 çerçeve boyutlarında ulaşılmıştır. Öznitelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarım artmaktadır ve 320x240 boyutlu çerçevelerde ise genel olarak başarım düşmektedir. Öznitelik vektör boyutunun arttırılması 640x480 çerçevelere ait başarımların belirgin bir biçimde düşmesine sebep olmaktadır. Şekil 5.19'da yer alan başarım oranlarının detayları ise çizelge 5.28'de verilmiştir.

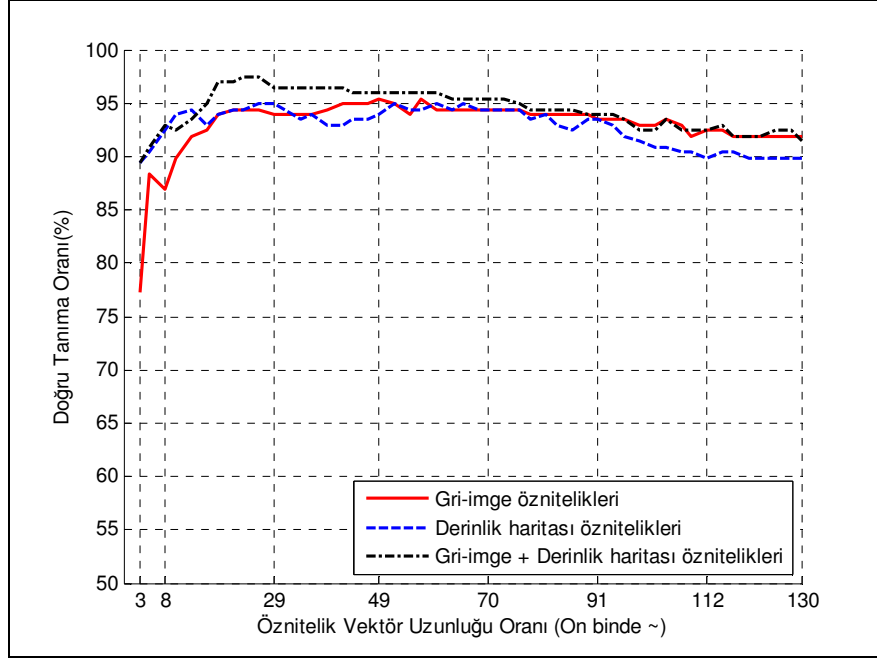
Çizelge 5. 28 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-1 ve ortalama eşik katsayısı yaklaşımlıyla gri-imge + derinlik haritası özniteliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	10	67,17	40	89,39	160	93,94
5	20	82,32	80	90,91	320	95,96
8	30	88,38	120	92,93	480	95,96
10	40	89,90	160	92,42	640	95,96
13	50	89,90	200	93,43	800	94,95
16	60	89,90	240	94,95	960	94,95
18	70	91,41	280	96,97	1120	92,93
21	80	91,41	320	96,97	1280	93,94
23	90	90,91	360	97,47	1440	92,93
26	100	90,91	400	97,47	1600	92,42
29	110	91,41	440	96,46	1760	92,42
31	120	91,92	480	96,46	1920	91,41

Çizelge 5. 28 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
34	130	92,93	520	96,46	2080	89,90
36	140	93,43	560	96,46	2240	89,39
39	150	94,44	600	96,46	2400	88,89
42	160	94,44	640	96,46	2560	87,88
44	170	95,45	680	95,96	2720	87,88
47	180	95,45	720	95,96	2880	86,87
49	190	94,95	760	95,96	3040	86,36
52	200	95,45	800	95,96	3200	85,86
55	210	95,45	840	95,96	3360	85,86
57	220	95,96	880	95,96	3520	85,86
60	230	95,96	920	95,96	3680	86,36
63	240	95,45	960	95,45	3840	86,36
65	250	96,46	1000	95,45	4000	85,86

Şekil 5.20’de 320x240 çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 1. Kategoride yer alan 22 kelime üzerinde farklı özniteliklerin kullanımlarındaki başarımları gösterilmiştir.



Şekil 5. 20 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-1).

Şekil 5.20'de gri-imge, derinlik haritası özniteliklerinin ayrı ayrı değerlendirilmesiyle gözlemlenen tanıma başarımları ve bu özniteliklerin bire bir (1:1) oranında birleştirildiği durumda gözlemlenen tanıma başarımları yer almaktadır. İlgili başarımlar grafiğinde en başarılı sonuçlara gri-imge+derinlik haritası özniteliklerinin kullanımında ulaşılmıştır. Bununla beraber derinlik haritası öznitelikleri ile gri-imge özniteliklerinden çoğu zaman birbirlerine yakın performanslar sergilediği görülmüştür. Şekil 5.20'deki başarımların değerlerinin detayları ise çizelge 5.29'da gösterilmiştir.

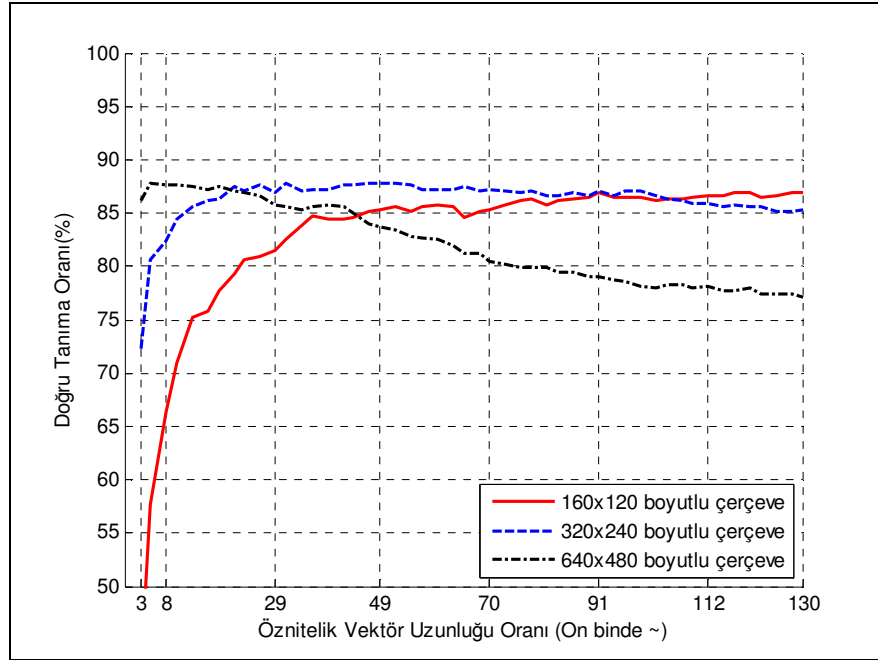
Çizelge 5. 29 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-1).

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNİTELİKLERİ		DERİNLİK HARİTASI ÖZİNİTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNİTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	20	77,27	20	89,39	40	89,39
5	40	88,38	40	90,40	80	90,91
8	60	86,87	60	92,42	120	92,93
10	80	89,90	80	93,94	160	92,42
13	100	91,92	100	94,44	200	93,43
16	120	92,42	120	92,93	240	94,95
18	140	93,94	140	93,94	280	96,97
21	160	94,44	160	94,44	320	96,97
23	180	94,44	180	94,44	360	97,47
26	200	94,44	200	94,95	400	97,47
29	220	93,94	220	94,95	440	96,46
31	240	93,94	240	94,44	480	96,46
34	260	93,94	260	93,43	520	96,46
36	280	93,94	280	93,94	560	96,46
39	300	94,44	300	92,93	600	96,46
42	320	94,95	320	92,93	640	96,46
44	340	94,95	340	93,43	680	95,96
47	360	94,95	360	93,43	720	95,96
49	380	95,45	380	93,94	760	95,96
52	400	94,95	400	94,95	800	95,96

Çizelge 5. 29 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNTELİKLERİ		DERİNLİK HARİTASI ÖZİNTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
55	420	93,94	420	94,44	840	95,96
57	440	95,45	440	94,44	880	95,96
60	460	94,44	460	94,95	920	95,96
63	480	94,44	480	94,44	960	95,45
65	500	94,44	500	94,95	1000	95,45

Şekil 5.21’de farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 2. Kategoride yer alan 58 kelime üzerindeki başarımları gösterilmiştir.



Şekil 5. 21 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2 ve ortalama eşik katsayısı yaklaşımıyla gri-imge + derinlik haritası özniteliklerinde).

Şekil 5.21’de TİD veri setindeki 2. kategoride yer alan kelimelerin farklı çerçeve boyutlarındaki başarımları yer almaktadır. Başarım değerleri gri-imge + derinlik haritası özniteliklerinin kullanımında hesaplanmış olup, tüm çerçeve boyutlarında başarılı sonuçlara ulaşılmıştır. Öznitelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarım artmaktadır ve 320x240 boyutlu çerçevelerde ise genel olarak sabit kalarak bir yapı göstermektedir. Öznitelik vektör boyutunun arttırılması 640x480 çerçevelere ait tanıma performansını olumsuz etkilemektedir. Şekil 5.21’de yer alan başarım oranlarının detayları ise çizelge 5.30’da verilmiştir.

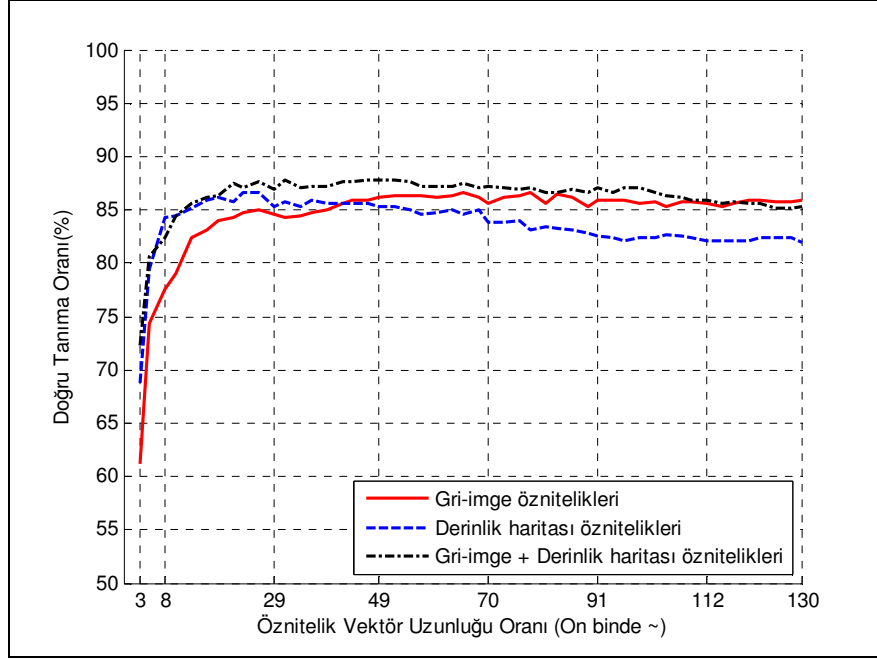
Çizelge 5. 30 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-2 ve ortalama eşik katsayısı yaklaşımla gri-imge + derinlik haritası özniteliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	10	41,25	40	72,24	160	86,12
5	20	57,60	80	80,61	320	87,83
8	30	66,35	120	82,32	480	87,64
10	40	70,91	160	84,41	640	87,64
13	50	75,29	200	85,55	800	87,45
16	60	75,86	240	86,12	960	87,26
18	70	77,76	280	86,31	1120	87,45
21	80	79,28	320	87,45	1280	87,07
23	90	80,61	360	87,07	1440	86,88
26	100	80,99	400	87,64	1600	86,69
29	110	81,56	440	86,88	1760	85,74
31	120	82,51	480	87,83	1920	85,55

Çizelge 5. 30 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
34	130	83,84	520	87,07	2080	85,36
36	140	84,79	560	87,26	2240	85,55
39	150	84,41	600	87,26	2400	85,74
42	160	84,41	640	87,64	2560	85,55
44	170	84,60	680	87,64	2720	84,98
47	180	85,17	720	87,83	2880	84,03
49	190	85,36	760	87,83	3040	83,65
52	200	85,55	800	87,83	3200	83,46
55	210	85,17	840	87,64	3360	82,89
57	220	85,55	880	87,26	3520	82,70
60	230	85,74	920	87,26	3680	82,51
63	240	85,55	960	87,26	3840	81,94
65	250	84,60	1000	87,45	4000	81,18

Şekil 5.22’de 320x240 çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 2. Kategoride yer alan 58 kelime üzerinde farklı özniteliklerin kullanımlarındaki başarımları gösterilmiştir.



Şekil 5. 22 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-2).

Şekil 5.22’de gri-imge, derinlik haritası özniteliklerinin ayrı ayrı değerlendirilmesiyle gözlemlenen tanıma başarımları ve bu özniteliklerin bire bir (1:1) oranında birleştirildiği durumda gözlemlenen tanıma başarımları yer almaktadır. İlgili başarımlar grafiğinde en başarılı sonuçlara gri-imge+derinlik haritası özniteliklerinin kullanımında ulaşılmıştır. Bununla beraber gri-imge özniteliklerinin genel olarak derinlik haritası özniteliklerinden başarılı olduğu da görülmüştür. Şekil 5.22’deki başarımların değerlerinin detayları ise çizelge 5.31’de gösterilmiştir.

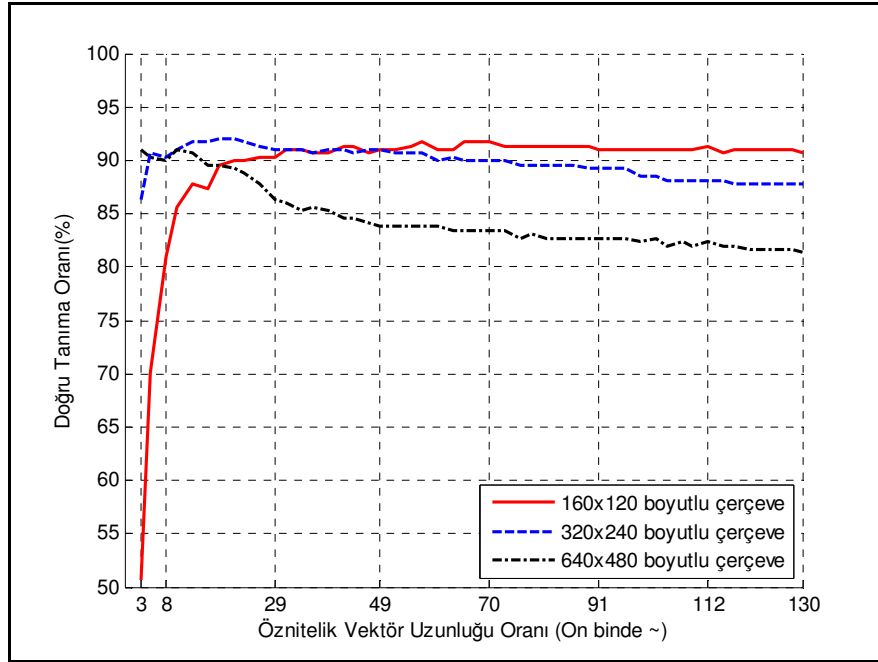
Çizelge 5. 31 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-2).

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNİTELİKLERİ		DERİNLİK HARİTASI ÖZİNİTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNİTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	20	61,22	20	68,82	40	72,24
5	40	74,33	40	79,47	80	80,61
8	60	77,57	60	84,22	120	82,32
10	80	79,09	80	84,41	160	84,41
13	100	82,32	100	85,17	200	85,55
16	120	83,08	120	85,93	240	86,12
18	140	84,03	140	86,12	280	86,31
21	160	84,22	160	85,74	320	87,45
23	180	84,79	180	86,69	360	87,07
26	200	84,98	200	86,69	400	87,64
29	220	84,60	220	85,36	440	86,88
31	240	84,22	240	85,74	480	87,83
34	260	84,41	260	85,36	520	87,07
36	280	84,79	280	85,93	560	87,26
39	300	84,98	300	85,55	600	87,26
42	320	85,55	320	85,55	640	87,64
44	340	85,93	340	85,55	680	87,64
47	360	85,93	360	85,55	720	87,83
49	380	86,12	380	85,36	760	87,83
52	400	86,31	400	85,36	800	87,83

Çizelge 5. 31 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNTELİKLERİ		DERİNLİK HARİTASI ÖZİNTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
55	420	86,31	420	84,98	840	87,64
57	440	86,31	440	84,60	880	87,26
60	460	86,12	460	84,79	920	87,26
63	480	86,31	480	84,98	960	87,26
65	500	86,69	500	84,60	1000	87,45

Şekil 5.23'te farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 3. Kategoride yer alan 31 kelime üzerindeki başarımları gösterilmiştir.



Şekil 5. 23 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3 ve ortalama eşik katsayısı yaklaşımıyla gri-imge + derinlik haritası özniteliklerinde).

Şekil 5.23’de TİD veri setindeki 3. kategoride yer alan kelimelerin farklı çerçeve boyutlarındaki başarımları yer almaktadır. Başarım değerleri gri-imge + derinlik haritası özniteliklerinin kullanımında hesaplanmış olup, en iyi başarım değerlerine 320x240 çerçeve boyutlarında ulaşılmıştır. Öznitelik vektör boyutunun arttırılmasıyla 160x120 boyutlu çerçevelerde başarım bir noktadan sonra genel olarak sabit kalmakta ve 320x240 boyutlu çerçevelerde ise genel olarak başarım düşmektedir. Öznitelik vektör boyutunun arttırılması 640x480 çerçevelere ait tanıma başarımının belirgin biçimde düşmesine sebep olmaktadır. Şekil 5.23’te yer alan başarım oranlarının detayları ise çizelge 5.32’de verilmiştir.

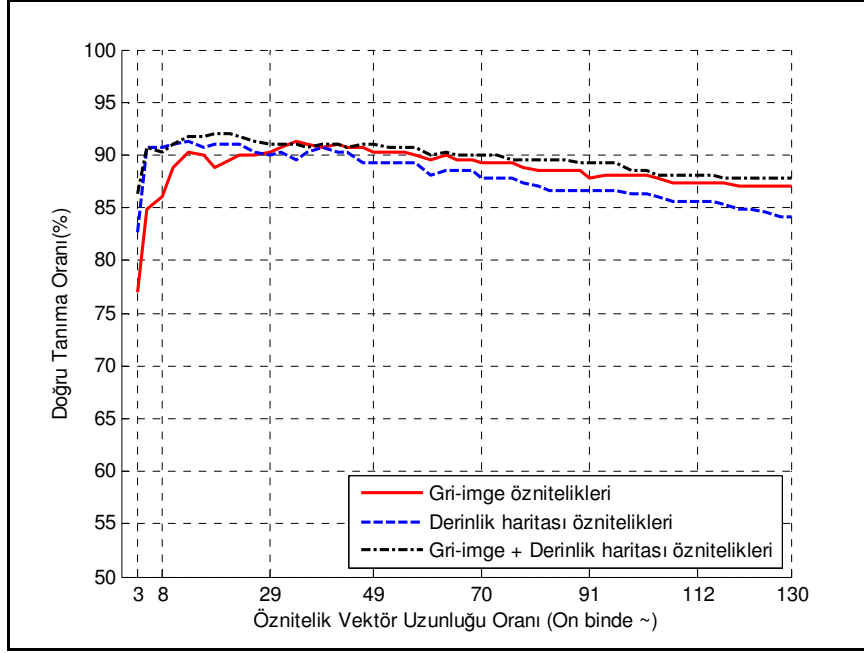
Çizelge 5. 32 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (kategori-3 ve ortalama eşik katsayısı yaklaşımıyla gri-imge + derinlik haritası özniteliklerinde).

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	10	50,72	40	86,33	160	91,01
5	20	70,14	80	90,65	320	90,29
8	30	80,94	120	90,29	480	89,93
10	40	85,61	160	91,01	640	91,01
13	50	87,77	200	91,73	800	90,65
16	60	87,41	240	91,73	960	89,57
18	70	89,57	280	92,09	1120	89,57
21	80	89,93	320	92,09	1280	89,21
23	90	89,93	360	91,73	1440	88,85
26	100	90,29	400	91,37	1600	87,77
29	110	90,29	440	91,01	1760	86,33

Çizelge 5. 32 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
31	120	91,01	480	91,01	1920	85,97
34	130	91,01	520	91,01	2080	85,25
36	140	90,65	560	90,65	2240	85,61
39	150	90,65	600	91,01	2400	85,25
42	160	91,37	640	91,01	2560	84,53
44	170	91,37	680	90,65	2720	84,53
47	180	90,65	720	91,01	2880	84,17
49	190	91,01	760	91,01	3040	83,81
52	200	91,01	800	90,65	3200	83,81
55	210	91,37	840	90,65	3360	83,81
57	220	91,73	880	90,65	3520	83,81
60	230	91,01	920	89,93	3680	83,81
63	240	91,01	960	90,29	3840	83,45
65	250	91,73	1000	89,93	4000	83,45

Şekil 5.24'te 320x240 çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının 3. Kategoride yer alan 31 kelime üzerinde farklı özniteliklerin kullanımlarındaki başarımları gösterilmiştir.



Şekil 5. 24 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerde ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-3).

Şekil 5.24'te gri-imge, derinlik haritası özniteliklerinin ayrı ayrı değerlendirilmesiyle gözlemlenen tanıma başarımları ve bu özniteliklerin bire bir (1:1) oranında birleştirildiği durumda gözlemlenen tanıma başarımları yer almaktadır. İlgili başarımlar grafiğinde en başarılı sonuçlara gri-imge+derinlik haritası özniteliklerinin kullanımında ulaşılmıştır. Bununla beraber derinlik haritası özniteliklerinin, gri-imge özniteliklerinden bazı öznitelik vektör uzunluklarında başarılı olduğu görülmüştür ancak grafiğin tamamında gri-imge öznitelikleri daha iyi performans sergilemektedir. Şekil 5.24'teki başarımların değerlerinin detayları ise çizelge 5.33'te gösterilmiştir.

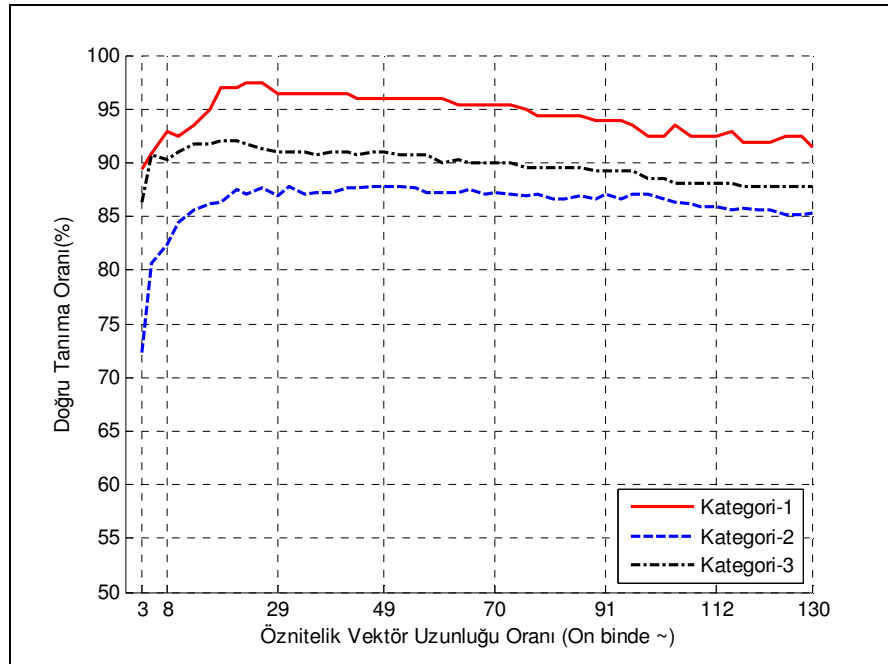
Çizelge 5. 33 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında kategori-3).

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZİNİTELİKLERİ		DERİNLİK HARİTASI ÖZİNİTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZİNİTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	20	76,98	20	82,73	40	86,33
5	40	84,89	40	90,65	80	90,65
8	60	85,97	60	90,65	120	90,29
10	80	88,85	80	91,01	160	91,01
13	100	90,29	100	91,37	200	91,73
16	120	89,93	120	90,65	240	91,73
18	140	88,85	140	91,01	280	92,09
21	160	89,57	160	91,01	320	92,09
23	180	89,93	180	91,01	360	91,73
26	200	89,93	200	90,29	400	91,37
29	220	90,29	220	89,93	440	91,01
31	240	90,65	240	90,29	480	91,01
34	260	91,37	260	89,57	520	91,01
36	280	91,01	280	90,29	560	90,65
39	300	90,65	300	90,65	600	91,01
42	320	91,01	320	90,29	640	91,01
44	340	90,65	340	90,29	680	90,65
47	360	90,65	360	89,21	720	91,01
49	380	90,29	380	89,21	760	91,01
52	400	90,29	400	89,21	800	90,65

Çizelge 5. 33 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	GRİ-İMGE ÖZNİTELİKLERİ		DERİNLİK HARİTASI ÖZNİTELİKLERİ		GRİ-İMGE + DERİNLİK HARİTASI ÖZNİTELİKLERİ	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
55	420	90,29	420	89,21	840	90,65
57	440	89,93	440	89,21	880	90,65
60	460	89,57	460	88,13	920	89,93
63	480	89,93	480	88,49	960	90,29
65	500	89,57	500	88,49	1000	89,93

Şekil 5.25'te 320x240 çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının her bir kategorideki başarımları gösterilmiştir.



Şekil 5. 25 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 3 kategoride yer alan 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında gri-imege + derinlik haritası öznitelikleri).

Şekil 5.25'te TİD veri setindeki her kategorinin 320x240 çerçeve boyutlarındaki başarımları yer almaktadır. Başarım değerleri gri-imege + derinlik haritası özniteliklerinin

kullanımında hesaplanmıştır. Sırasıyla en başarılı sonuçlar 1. kategori, 3. kategori ve 2. kategoride gözlemlenmiştir. 2. kategoriye ait başarımların düşük olmasının sebebi; bu kategorinin diğer kategorilere nazaran daha fazla kelime içermesi ve bazı kelimelerin birbirlerine benzemeleridir. Şekil 5.25'te yer alan başarımların oranlarının detayları ise çizelge 5.34'te verilmiştir.

Çizelge 5. 34 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının 3 kategoride yer alan 320x240 boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında gri-imge + derinlik haritası öznitelikleri).

DCT KATSAYI ORANI (~ ON BİNDE)	KATEGORİ - 1		KATEGORİ - 2		KATEGORİ - 3	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
3	40	89,39	40	72,24	40	86,33
5	80	90,91	80	80,61	80	90,65
8	120	92,93	120	82,32	120	90,29
10	160	92,42	160	84,41	160	91,01
13	200	93,43	200	85,55	200	91,73
16	240	94,95	240	86,12	240	91,73
18	280	96,97	280	86,31	280	92,09
21	320	96,97	320	87,45	320	92,09
23	360	97,47	360	87,07	360	91,73
26	400	97,47	400	87,64	400	91,37
29	440	96,46	440	86,88	440	91,01
31	480	96,46	480	87,83	480	91,01
34	520	96,46	520	87,07	520	91,01
36	560	96,46	560	87,26	560	90,65
39	600	96,46	600	87,26	600	91,01

Çizelge 5. 34 (devam)

DCT KATSAYI ORANI (~ ON BİNDE)	KATEGORİ - 1		KATEGORİ - 2		KATEGORİ - 3	
	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)	DCT KATSAYI MİKTARI	CRR (%)
42	640	96,46	640	87,64	640	91,01
44	680	95,96	680	87,64	680	90,65
47	720	95,96	720	87,83	720	91,01
49	760	95,96	760	87,83	760	91,01
52	800	95,96	800	87,83	800	90,65
55	840	95,96	840	87,64	840	90,65
57	880	95,96	880	87,26	880	90,65
60	920	95,96	920	87,26	920	89,93
63	960	95,45	960	87,26	960	90,29
65	1000	95,45	1000	87,45	1000	89,93

Çizelge 5.35'te ise farklı çerçeve boyutlarında Manhattan mesafesini kullanan K-en yakın komşu (KNN) sınıflandırıcısının her bir kategoride farklı özneliklerin kullanımlarındaki en iyi doğru tanıma oranları ve bu değerlere karşılık gelen DCT katsayı oranları gösterilmiştir.

Çizelge 5. 35 Tid veri setinde manhattan mesafesini kullanan knn (k=1) sınıflandırıcısının farklı çerçeve boyutlarında olan görüntülerdeki ve farklı dct katsayı oranlarındaki en başarılı doğru tanıma oranları (ortalama eşik katsayısı yaklaşımında 111 kelime ve gri-imege + derinlik haritası öznitelikleri).

ÖZNİTELİK VEKTÖRÜNDE YER ALAN ÖZNİTELİKLER (DCT KATSAYI ORANI ~ ON BİNDE)	DOĞRU TANIMA ORANLARI (%)					
	160x120 ÇERÇEVE BOYUTU		320x240 ÇERÇEVE BOYUTU		640x480 ÇERÇEVE BOYUTU	
	DCT K.O.	CRR	DCT K.O.	CRR	DCT K.O.	CRR
GRI-İMGE	130	89,52	39	89,52	10	90,12
DERİNLİK	99	88,42	31	90,52	8	89,72
GRI-İMGE + DERİNLİK	86	91,22	21	91,32	5	91,52

SONUÇ VE ÖNERİLER

Bu tez çalışmasında izlenen yaklaşım ve yöntemlerle işaret dillerine ait durağan olmayan (gesture) kelimelerin tanınması hedeflenmiştir. Çalışmada kullanılan ve ardışıl görüntü çerçevelerinin farklarına dayalı toplamsal imge yaklaşımıyla Amerikan İşaret Dili (ASL) ve Türk İşaret Dili'ne (TİD) ait işaret örnekleri başarılı olarak tanınabilmektedir.

Çalışmada öznitelik çıkarımı aşamasında kullanılan iki kademeli yöntemin ilk aşamasında hareketlere ait toplamsal imgeler oluşturulmuş; ikinci aşamada ise bu toplamsal imgelere 2-B DCT uygulanarak zamansal imgeler uzamsal uzaya aktarılmış ve karmaşık olmayan bu iki kademeli yapı ile öznitelik vektörleri elde edilmiştir. İşaret örneklerinin sınıflandırılması aşamasında Manhattan mesafesini kullanan K-en yakın komşu (K-NN) sınıflandırıcısı kullanılmıştır. Sistemin performansının değerlendirilmesi aşamasında çalışmada kullanılan ilgili veri setlerinin eğitim ve test seti olarak bölünmesi için çapraz doğrulama uygulanmıştır.

Bu tez çalışması kapsamında oluşturulan TİD veri setinde yer alan görüntülerin tamamı işitme engelli uzman işaretçiler tarafından gerçekleştirilmiştir. Başarım sınamalarının yapılacağı sistemlerde veri setleri bu açıdan önemlidir. Önerilen tez çalışmasında da veri seti oluşturulurken bu durum göz önünde bulundurulmuş ve bu özelliğe sahip TİD veri setinde başarılı sonuçlar gözlemlenmiştir. İlgili veri setinde yer alan görüntüler Kinect duyarga kullanılarak oluşturulmuştur. Cihazın derinlik algılayabilmesi özelliği de kullanılarak RGB görüntülerle birlikte derinlik bilgilerini içeren derinlik haritaları da kaydedilmiştir. Derinlik bilgilerinin işaret dili tanımada belirlenen yöntemle kullanılabilir olup olmadığı tez çalışmasının amaçlarından biri olarak belirlenmiştir.

Deneyisel çalışmalar bölümünde kapsamlı olarak değinilen testler sonucunda derinlik bilgisinin tanımada etkin olduğu gözlemlenmiştir. RGB görüntülerde, derinlik haritalarında ve bu görüntülerin birlikte kullanımı işaretlerin tanınmasında değerlendirilmiştir. Genel olarak derinlik haritalarında sağlanan başarımlar RGB görüntülerde sağlanan başarımlardan üstündür. En iyi başarımlar ise RGB görüntüler ve derinlik haritalarının birlikte kullanımında görülmüştür. Bu bağlamda tez çalışmasında, derinlik bilgilerinin kullanımının mevcut yöntemde oldukça başarılı sonuçlar verdiği ve işaret tanımada etkin olduğu sonucuna varılmıştır.

Genel olarak işaret dillerinin tanınabilmesi için görü tabanlı sistemlerde işareti gerçekleştiren el, kol, kafa, yüz gibi vücut öğelerinin imge dizilerinde bölütlenmesi ve takip edilmesi aşamaları vardır. Bu aşamalar tanıma probleminden önceki aşamalarda yer almakta ve gerçekleştirilmesi sistemlere zaman, karmaşıklık ve işlem açısından fazladan yük getirmektedir. Bilhassa gerçek zamanlı olarak tasarlanan sistemlerdeki işlem yükünü arttıran bu aşamalar ciddi manada zamanlama sorunu oluşturabilmektedir. Tez çalışmasında uygulanan yöntemle, imge dizileri bütün olarak işleme tabi tutularak herhangi bir nesne bölütleme ve nesne takip etme yaklaşımı kullanılmamıştır. Bu sayede tasarlanan sistemde ciddi anlamda bir zaman kazanımı ve basitlik sağlanmıştır.

Önerilen sistemde farklı çerçeve boyutlarında ve farklı DCT katsayı miktarlarının tanıma başarımlarını gözlemlemek amacıyla testler yapılmış ve sonuçlar tez kitabının deneysel sonuçlar bölümünde etraflıca anlatılmıştır. Farklı çerçeve boyutlarında ve öznitelik vektörlerinde düşük miktarda DCT katsayısı kullanımında oldukça başarılı sonuçlar gözlemlenmiştir.

Yapılan çalışmada zamansal öznitelikleri ifade etmek amacıyla toplamsal imge yaklaşımı kullanılmış ve uzamsal öznitelikleri elde etmek için ise 2-B DCT'den faydalanılmıştır. Uzamsal öznitelikleri çıkarmak amacıyla farklı dönüşüm yöntemleri ve yaklaşımları kullanılabilir. Bu bağlamda ileride yapılması muhtemel çalışmalarda uzamsal öznitelikleri elde etmek amacıyla farklı yöntemler kullanılabilir ve sistemin iyileştirilmesi için bu yönde çalışmalar yürütülebilir.

Tez kapsamında Kinect algılayıcı kullanılarak oluşturulan Türk İşaret Dili'ne ait veri setinde RGB-D video görüntülerinde derinlik bilgisi de tanıma başarısını arttırmada etkin olarak kullanılmıştır. Ardışıl hareket farklarını ve dönüşüm yöntemlerini kullanarak uzam-zamansal öznitelikleri çıkaran sistemle ve K-en yakın komşu sınıflandırıcısı ile ASL veri setinde %95-99, TİD veri setinde %80-98 arasında değişen başarımlarla işaret diline ait durağan olmayan görüntüler tanınabilmektedir.

KAYNAKLAR

- [1] Lucas, C., (2001). The Sociolinguistics of Sign Languages, Cambridge University Press, Cambridge.
- [2] Koç Üniversitesi, Türk İşaret Dili, <http://www.turkisaretdili.ku.edu.tr>, 21 Mayıs 2013.
- [3] Murakami, K. ve Taguchi, H., (1991). "Gesture Recognition Using Recurrent Neural Networks", Proceeding of the SIGCHI Conference on Human Factors in Computing Systems, 27 April-2 May 1991, New Orleans, 237-242.
- [4] Starner, T. ve Pentland, A., (1995). "Real-time American Sign Language Recognition from Video using Hidden Markov Models", Proceedings of International Symposium on Computer Vision, 21-23 November 1995, Coral Gables, 265-270.
- [5] Alpaydın, E., (2011). Yapay Öğrenme, Boğaziçi Üniversitesi Yayınları, İstanbul.
- [6] Starner, T., Weaver, J. ve Pentland, A., (1997). "A Wearable Computer Based American Sign Language Recognizer", First International Symposium on Wearable Computers, 13-14 October 1997, Cambridge, 130-137.
- [7] Hernandez-Rebollar, Jose L., Kyriakopoulos, N. ve Lindeman, Robert W., (2004). "A New Instrumented Approach for Translating American Sign Language Into Sound and Text", Proceedings of Sixth International Conference on Automatic Face and Gesture Recognition, 17-19 May 2004, Seoul, 547-552.
- [8] Hernandez-Rebollar, Jose L., Lindeman, Robert W. ve Kyriakopoulos, N. (2002). "A Multi-Class Pattern Recognition System for Practical Finger Spelling Translation", Proceedings of Fourth International Conference on Multimodal Interfaces, 14-16 October 2002, Pittsburgh, 185-190.
- [9] Isaacs, J., ve Foo, S., (2004). "Hand Pose Estimation for American Sign Language", Proceedings of the Thirty-Sixth Southeastern Symposium on System Theory, 14-16 March 2004, Atlanta, 132-136.
- [10] Yang, M., Ahuja, N., ve Tabb, M., (2002). "Extraction of 2D Motion Trajectories and Its Application to Hand Gesture Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, 24 (8):1061-1074.

- [11] Gupta, L., ve Ma, S., (2001). "Gesture Based Interaction and Communication: Automated Classification of Gesture Contours", IEEE Transactions on Systems, Man, and Cybernetics, 31 (1):114-120.
- [12] Kevin, N.Y.Y., Ranganath, S., ve Ghosh, D., (2004). "Trajectory Modeling in Gesture Recognition Using CyberGloves and Magnetic Trackers", IEEE TENCON Region 10 Conference-Analog and Digital Techniques in Electrical Engineering, 21-24 November 2004, Chiang Mei, 571-574.
- [13] Sclaroff, S., Betke, M., Kollios, G., Alon, J., Athitsos, V., Li, R., Magee, J., ve Tian, T., (2005). "Tracking, Analysis and Recognition of Human Gestures in Video", Proceedings of Eighth International Conference on Document Analysis and Recognition, 29 August-1 September 2005, Seoul, 806-810.
- [14] Ding, L. ve Martinez, A.M., (2006). "Three Dimensional Shape and Motion Reconstruction for the Analysis of ASL", Conference on Computer Vision and Pattern Recognition Workshop, 17-22 June 2006, New York, 146.
- [15] Yang, R., ve Sarkar, S., (2006). "Detecting Coarticulation in Sign Language Using Conditional Random Fields", 18th International Conference on Pattern Recognition, 20-24 August 2006, Hong Kong, 108-112.
- [16] Thangali, A. ve Sclaroff, S., (2005). "Periodic Motion Detection and Estimation via Space-Time Sampling", Seventh IEE Workshops on Application of Computer Vision, 5-7 January 2005, Breckenridge, 176-182.
- [17] Saman, G., Sajjad, A., Hameed, A.O. ve Shah, A.R., (2006). "Visual Sign Language Interpretation Using Spatial Temporal Neural Processing", IEEE International Multitopic Conference, 23-24 December 2006, Islamabad, 128-133.
- [18] Amin, M.A. ve Yan, H., (2007). "Sign Language Finger Alphabet Recognition From Gabor-PCA Representation of Hand Gestures", International Conference on Machine Learning and Cybernetics, 19-22 August 2007, Hong Kong, 2218-2223
- [19] Ding, L. ve Martinez, A.M., (2007). "Recovering the Linguistic Components of the Manual Signs in American Sign Language", IEEE Conference on Advanced Video and Signal Based Surveillance, 5-7 September 2007, London, 447-452.
- [20] Alon, J., Athitsos, V., Yuan, Q. ve Sclaroff, S., (2009). "A Unified Framework for Gesture Recognition and Spatiotemporal Gesture Segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, 31 (9):1685-1699.
- [21] Cho, S.S., Yang, H.D. ve Lee, S.W., (2009). "Sign Language Spotting Based on Semi-Markov Conditional Random Field", Workshop on Applications of Computer Vision, 7-8 December 2009, Snowbird, 1-6.
- [22] Assaleh, K., Shanableh, T., Fanaswala, M., Bajaj, H. ve Amin, F., (2008). "Vision Based System for Continuous Arabic Sign Language Recognition in User Depended Mode", Fifth International Symposium on Mechatronics and Its Applications, 27-29 May 2008, Amman, 1-5.

- [23] Shanableh, T. ve Assaleh, K., (2007). "Arabic Sign Language Recognition in User-independent Mode", International Conference on Intelligent and Advanced Systems, 25-28 November 2007, Kuala Lumpur, 597-600.
- [24] Luo, D. ve Ohya, J., (2010). "Study on Human Gesture Recognition from Moving Camera Images", International Conference on Multimedia and Expo, 19-23 July 2010, Suntec City, 274-279.
- [25] Keskin, C. ve Akarun, L., (2009). "STARS: Sign Tracking and Recognition System Using Input-output HMMs", Pattern Recognition Letters, 30 (12):1086-1095.
- [26] Keskin, C., Aran, O. ve Akarun, L., (2011). "Hand Gesture Analysis", Computer Analysis and Human Behavior, Springer London, 125-149.
- [27] Keskin, C., Cemgil, A.T. ve Akarun, L., (2011). "DTW Based Clustering to Improve Hand Gesture Recognition", Human Behavior Understanding Lecture Notes in Computer Science, Springer Berlin Heidelberg, 7065:72-81.
- [28] Haberdar, H., (2005). Saklı Markov Model Kullanılarak Görüntüden Gerçek Zamanlı İşaret Dili Tanıma Sistemi, Yüksek Lisans Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul.
- [29] Aran, O., (2008). Vision Based Sign Language Recognition: Modelling and Recognizing Isolated Signs with Manual and Non-manual Components, Doktora Tezi, Boğaziçi Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
- [30] Arı, İ., (2008). Facial Feature Tracking and Expression Recognition for Sign Language, Yüksek Lisans Tezi, Boğaziçi Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
- [31] Guvensan, M.A. ve Haberdar, H., (2007). "Detection and Tracking of Face and Facial Features for Recognition of Turkish Sign Language", 15th IEEE Signal Processing and Communications Applications, 11-13 June 2007, Eskisehir, 1-4.
- [32] Aran, O., Ari, I., Guvensan, A., Haberdar, H., Kurt, Z., Turkmen, I., Uyar, A. ve Akarun L., (2007). "A Database of Non-Manual Signs in Turkish Sign Language", 15th IEEE Signal Processing and Communications Applications, 11-13 June 2007, Eskisehir, 1-4.
- [33] Altun, O. ve Albayrak, S., (2011). "Turkish Fingerspelling Recognition System Using Generalized Hough Transform, Interest Regions, and Local Descriptors", Pattern Recognition Letters, 32 (13):1626-1632.
- [34] Altun, O. ve Albayrak, S., (2012). "An Evaluation of Local Interest Regions for Non-rigid Object Class Recognition", Expert Systems with Applications, 39 (3):2335-2340.
- [35] Isikdogan, F. ve Albayrak, S., (2011). "Automatic Recognition of Turkish Fingerspelling", International Symposium on Innovations in Intelligent Systems and Applications, 15-18 June 2011, Istanbul, 264-267.
- [36] Kose, H., Yorganci, R. ve Itama, I.I., (2011). "Humanoid Robot Assisted Interactive Sign Language Tutoring Game", IEEE International Conference on Robotics and Biomimetics, 7-11 December 2011, Phuket, 2247-2248.

- [37] Kose, H., ve Yorganci, R., (2011). "Tale of A Robot: Humanoid Robot Assisted Sign Language Tutoring", 11th IEEE-RAS International Conference on Humanoid Robots, 26-28 October 2011, Bled, 105-111.
- [38] Selda, B. ve Vasif, N.F., (2006). "Recognition of Vowels Letters of Turkish Sign Language by Artificial Neural Networks", 14th IEEE Signal Processing and Communications Applications, 17-19 April 2006, Antalya, 1-4.
- [39] Dibeklioglu, H., Dikici, E., Santemiz, P., Balci, K., ve Akarun, L., (2007). "Sign Language Motion Tracking and Generating 3D Motion Pieces Using 2D Features", 15th IEEE Signal Processing and Communications Applications, 11-13 June 2007, Eskisehir, 1-4.
- [40] Santemiz, P., Aran, O., Saraclar, M. ve Akarun, L., (2009). "Alignment Based Extraction of Isolated Signs from Sign Language Videos", 17th IEEE Signal Processing and Communications Applications, 9-11 April 2009, Antalya, 532-535.
- [41] Santemiz, P., Aran, O., Saraclar, M. ve Akarun, L., (2009). "Automatic Sign Segmentation from Continuous Signing via Multiple Sequence Alignment", 12th International Conference on Computer Vision Workshops, 27 September-4 October 2009, Kyoto, 2001-2008.
- [42] Aran, O. ve Akarun, L., (2010). "A Multi-class Classification Strategy for Fisher Scores: Application to Signer Independent Sign Language Recognition", Pattern Recognition, 43 (5):1776-1788.
- [43] T.C. Başbakanlık Atatürk Kültür, Dil ve Tarih Yüksek Kurumu, Türk Dil Kurumu, Türk İşaret Dili Sözlüğü, http://www.tdk.gov.tr/index.php?option=com_content&view=article&id=264, 21 Mayıs 2013.
- [44] T.C. Boğaziçi Üniversitesi, Bilgisayar Mühendisliği Bölümü, Türk İşaret Dili Kaynak Sitesi, <http://www.cmpe.boun.edu.tr/tid/>, 21 Mayıs 2013.
- [45] Oz, C. ve Ming, C.L., (2011). "American Sign Language Word Recognition with A Sensory Glove Using Artificial Neural Networks", Engineering Applications of Artificial Intelligence, 24 (7):1204-1213.
- [46] Stylianos, T., Gesture Database, <http://cgi.di.uoa.gr/~stzikop/myPage.php#anchorGesture>, 21 Mayıs 2013.
- [47] Wikipedia The Free Encyclopedia, Kinect, <http://en.wikipedia.org/wiki/Kinect>, 21 Mayıs 2013.
- [48] Jain, A.K., (1989). Fundamentals of Digital Image Processing, Prentice Hall, Englewood Cliffs NJ.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Abbas MEMİŞ
Doğum Tarihi ve Yeri : 23.03.1987 – Arguvan / Malatya
Yabancı Dili : İngilizce
E-posta : abbasmemis@gmail.com

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Lisans	Bilgisayar Mühendisliği	Yıldız Teknik Üniversitesi	2010
Lise	Matematik-Fen	Gemlik Lisesi	2005

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2009-2012	Naryaz Bilgisayar Yazılım ve Tic. Ltd. Şti.	Analist Programcı
2012-2013	İstanbul Teknik Üniversitesi	Araştırma Görevlisi

YAYINLARI

Bildiri

1. Memiř, A., Albayrak, S., "Aısal Radyal Paralamaya ve İstatistiksel Doku Analizine Dayalı İris Tanıma Sistemi", Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu (ASYU),2010.
2. Memiř, A., Albayrak, S., "Kinect RGB Görüntülerde ve Derinlik Haritalarında Uzun-zamansal Özellikleri Kullanarak Türk İşaret Dili Tanıma", Sinyal İşleme ve İletişim Uygulamaları Kurultayı (SIU),2013.