

**BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİMDALİ
BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS
PROGRAMI**

**EĞİTİM ÇAĞINDA ÇOCUKLAR İÇİN EL HAREKETİ
TANIMA TABANLI GEOMETRİ ÖĞRENME UYGULAMASI**

YÜKSEK LİSANS TEZİ

HAZIRLAYAN

EVREN KEMAL UYSAL

ANKARA - 2020

**BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİMDALİ
BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS
PROGRAMI**

**EĞİTİM ÇAĞINDA ÇOCUKLAR İÇİN EL HAREKETİ
TANIMA TABANLI GEOMETRİ ÖĞRENME UYGULAMASI**

YÜKSEK LİSANS TEZİ

HAZIRLAYAN

EVREN KEMAL UYSAL

TEZ DANIŞMANI

DR. ÖĞR. ÜYESİ EMRE SÜMER

ANKARA - 2020

BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı çerçevesinde Evren Kemal Uysal tarafından hazırlanan bu çalışma, aşağıdaki jüri tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 03/02/2020

Tez Adı: Eğitim Çağında Çocuklar İçin El Hareketi Tanıma Tabanlı Geometri Öğrenme Uygulaması

Tez Jüri Üyeleri

İmza

Doç. Dr. Fikret ARI, Ankara Üniversitesi

Dr. Öğr. Üyesi Emre SÜMER, Başkent Üniversitesi

Dr. Öğr. Üyesi Tülin ERÇELEBİ AYYILDIZ, Başkent Üniversitesi

ONAY

Prof. Dr. Faruk ELALDI
Fen Bilimleri Enstitüsü Müdürü

Tarih: ... / ... / 2020

BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS TEZ ÇALIŞMASI ORJİNALLİK RAPORU

Tarih: 14/02/2020

Öğrencinin Adı, Soyadı : Evren Kemal UYSAL

Öğrencinin Numarası : 21310039

Anabilim Dalı : BİLGİSAYAR MÜHENDİSLİĞİ

Programı : BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS

Danışmanın Unvanı/Adı, Soyadı : Dr. Öğr. Üyesi Emre SÜMER

Tez Başlığı : Eğitim Çağında Çocuklar İçin El Hareketi Tanıma Tabanlı Geometri

Öğrenme Uygulaması

Yukarıda başlığı belirtilen Yüksek Lisans tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 53 sayfalık kısmına ilişkin, 14 / 02 / 2020 tarihinde tez danışmanım tarafından TURNITIN adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı %2'dir.

Uygulanan filtrelemeler:

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

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

Öğrenci İmzası:

Onay

.../.../2020

Dr. Öğr. Üyesi Emre Sümer

TEŞEKKÜR

Yazar, bu çalışmanın gerçekleşmesinde katkılarından dolayı, aşağıda adı geçen kişi ve kuruluşlara içtenlikle teşekkür eder.

Sayın Dr. Öğr. Üyesi Emre SÜMER'e (tez danışmanı), çalışmanın süresi boyunca yol göstermesi ve karşılaşılan sorunların çözülmesinde her zaman yardımcı olduğu için...

Önerilen sistemin geliştirilmesi için gerekli eğitim süreci ve test sürecinde veri setlerinin oluşturulması ve uygulamanın test edilmesinde emeği geçen herkese katkılarından dolayı...

Değerli arkadaşım Burak Kabasakal'a çalışmanın gerçekleştirilmesi süresince yardımları ve verdiği destek için...

Tüm bu süreçte destekleriyle yanımda olan aileme ve arkadaşlarıma desteklerinden dolayı...

ÖZET

Evren Kemal UYSAL

EĞİTİM ÇAĞINDA ÇOCUKLAR İÇİN EL HAREKETİ TANIMA TABANLI GEOMETRİ ÖĞRENME UYGULAMASI

Başkent Üniversitesi Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

2020

El hareketi tanıma, insan-bilgisayar etkileşimi arayüzleri geliştirilmesinde ve bu etkileşimin klavye ya da fare gibi araçlar gerektirmeden yapılabilmesinde büyük rol oynayabilecek, çok popüler bir araştırma konusudur. El hareketi tanıma, günlük bilgi ve haber elde etme, medya oynatıcısı yönetimi, sunum kontrol sistemleri, işitme engelliler için iletişim sistemleri, işaret dilleri ile insan-bilgisayar etkileşimi sağlayan uygulamalar, eğitim - matematik öğrenimi, yabancı dil ve kültür eğitimi, geometri öğrenimi - gibi alanlar üzerinde kullanılmaktadır. Bu tez çalışmasında, eğitim çağındaki çocuklar için el hareketi tanıma ile kullanılabilen bir geometri öğrenme uygulaması önerilmiştir. Bu uygulamada, standart bir ağ kamerası yardımıyla görüntü alınmakta, daha önceden sisteme tanıtılmış belirli el hareketleri ile eşleştirilerek, bu harekete karşılık gelen bir grafik aksiyonu çıktı olarak sunulmaktadır. Önerilen sistem iki ana aşamadan meydana gelmektedir. İlk olarak görüntünün elde edilmesi ve anlamlandırılması sağlanmaktadır. Bunun için OxfordNet'in VGG16 mimarisi, Evrişimsel Sinir Ağı (CNN) olarak seçilmiş ve gerekli değişiklikler yapılarak bu tez çalışmasına uyarlanmıştır. Modeli eğitmek için kullanılan veri seti bu tez çalışması kapsamında 6 farklı kişinin her el hareketi için fotoğraflarının çekilmesi suretiyle oluşturulmuştur. Önerilen sistemin ikinci adımında ise, algılanan el hareketleri ile geometrik şekillerin grafiksel gösterimleri ve çeşitli etkileşimler ilişkilendirilmiştir. Bu grafik aksiyonları önceki ve sonraki şekli çizmek, mevcut şekli büyötmek / küçölmek / döndörmek ve mevcut şekil için çevre, alan ve hacim formöllerini göstermek gibi işlemlerden oluşmaktadır. Sistemin kullanımı ve testleri hem yetişkinler hem de eğitim çağındaki çocuklar tarafından gerçekleştirilmiş ve tatmin edici sonuçlar elde edilmiştir.

ANAHTAR KELİMELELER: Bilgisayarla Görme, Geometrik Şekiller, El Hareketi Tanıma, Evrişimsel Sinir Ağları

ABSTRACT

Evren Kemal UYSAL

GEOMETRY LEARNING APPLICATION FOR SCHOOL-AGE CHILDREN USING HAND GESTURE RECOGNITION

Baskent University Institute of Science And Engineering

Computer Engineering Department

2020

Hand gesture recognition is a popular research topic which is a great way for implementing Human Computer Interaction interfaces for communicating with a computer without a need for having a mouse or a keyboard. Hand Gesture Recognition has been used on several areas such as daily information retrieval, media player controllers, presentation controllers, communication systems for the hearing impaired, mathematics education, language and culture education and geometry education. In this thesis a geometry learning application for school aged children using HGR is proposed. In this application a video stream is captured from a standart webcam and the received hand gestures are matched with previously defined hand gestures and the respective graphical action for the received hand gesture is given as an output. The system has two main stages. Firstly, the acquisition and interpretation of the image is employed. OxfordNet's VGG-16 architecture was selected as the CNN architecture and was modified in order to fit this study. The data set used to train the model was created by taking photos of 6 different people for each hand gesture. In the second step of the proposed system, the recognized hand gestures are mapped with graphical representations and various interactions. These graphical actions consist of drawing the next and previous shape, enlarging / shrinking / rotating the current shape and showing the perimeter, area and volume formulas for the current shape. The system was tested by adults and school aged children and satisfactory results were obtained.

KEYWORDS: Computer Vision, Geometric Shapes, Hand Gesture Recognition, Convolutional Neural Networks

İÇİNDEKİLER

	Sayfa
ÖZET.....	i
ABSTRACT.....	ii
İÇİNDEKİLER LİSTESİ	iii
TABLolar LİSTESİ.....	v
ŞEKİLLER LİSTESİ.....	vi
SİMGELER VE KISALTMALAR LİSTESİ.....	viii
1. GİRİŞ.....	1
1.1. Çalışmanın Amacı ve Kapsamı.....	1
1.2. Benzer Çalışmalar.....	2
1.2.1. El hareketi tanımlama kullanılarak geliştirilen genel kullanım uygulamaları	2
1.2.2. El hareketi tanımlama kullanılarak geliştirilen eğitim alanı uygulamaları	7
1.2.3. Geometri eğitimi uygulamaları ve el hareketi tanımlama kullanılarak geliştirilen geometri alanı uygulamaları	12
1.3. Tez Planı.....	15
2. SİSTEMİN GENEL YAPISI VE KULLANILAN YÖNTEMLER.....	17
2.1. Sistemin Yapısı.....	17
2.2. Evrişimsel Sinir Ağları ile El Hareketi Algılama.....	19
2.2.1. Evrişimsel sinir ağları.....	21
2.2.1.1. Evrişim katmanı (Convolution layer).....	22
2.2.1.2. ReLU (Rectified linear unit).....	23
2.2.1.3. Havuzlama katmanı (Pooling layer).....	23
2.2.1.4. Tam bağlı katman (Fully connected layer)	24
2.3. Geometrik Grafik İşlemleri.....	25
2.3.1. Bilgisayar grafikleri.....	25
2.3.2. Önerilen sistemin barındırdığı grafik işlemleri.....	28
2.3.2.1. Sonraki şekli çizme.....	29
2.3.2.2. Önceki şekli çizme.....	29
2.3.2.3. Mevcut şekle 2B-3B dönüşümü uygulama.....	30
2.3.2.4. Mevcut şekli büyütme.....	31

2.3.2.5. Mevcut şekli küçültme.....	31
2.3.2.6. Mevcut şekli X ekseninde d6ndürme.....	32
2.3.2.7. Mevcut şekli Y ekseninde d6ndürme.....	32
2.3.2.8. Mevcut şekil için çevre, alan ve hacim formüllerini gösterme.....	33
2.3.2.9. Aktive etme.....	34
2.3.2.10. Rotasyon yönünü deęiştirme.....	35
3. VERİ KÜMESİ.....	36
4. DENEYSEL SONUÇLAR.....	38
4.1. Geliştirme ve Test Ortamı Konfigürasyonları.....	38
4.2. Deęerlendirme Sonuçları ve Tartışma.....	38
5. SONUÇ.....	52
KAYNAKLAR.....	54

TABLolar LİSTESİ

	Sayfa
Tablo 2.1. Sistemdeki şekillerin 2 boyut 3 boyut dönüşümü için karşılıkları	31
Tablo 4.1. El Hareketine Karşılık Gelen 2 ve 3 Boyut Eşleşmeleri.....	40
Tablo 4.2. Sistemin eşik değeri uygulanmadan sınıflandırma yaptığında başarı yüzdeleri.....	42
Tablo 4.3. Sistemin eşik değeri uygulanarak sınıflandırma yaptığında başarı yüzdeleri.....	42
Tablo 4.4. %70'lik eşik uygulanmamış sınıflandırıcı için Cohen'in Kappa Katsayısı değeri.....	44
Tablo 4.5. %70'lik eşik uygulanmış sınıflandırıcı için Cohen'in Kappa Katsayısı değeri.....	44
Tablo 4.6. %70'lik eşik uygulanmamış sınıflandırıcı için F1 puanları	45
Tablo 4.7. %70'lik eşik uygulanmış sınıflandırıcı için F1 puanları.....	45

ŞEKİLLER LİSTESİ

Sayfa

Şekil 1.1.	El hareketi algılama tabanlı kontrollü oyun sistemi için genel akış [2] ...	3
Şekil 1.2.	Fare kontrol sistemi için belirlenen özellikler üzerinde yapılan Hesaplamalar [8].....	4
Şekil 1.3.	El hareketi algılama tabanlı televizyon kontrol sisteminin özelleştirme ekranı görseli [9].....	5
Şekil 1.4.	Microsoft Powerpoint kontrol sistemi için kullanılan el hareketleri görseli [11]	7
Şekil 1.5.	El hareketi algılama tabanlı animasyon oyunu genel akışı görseli [14].....	8
Şekil 1.6.	VECAR sisteminin akış şeması görseli [16].....	9
Şekil 1.7.	VECAR sisteminin nesne manipülasyonu için durum diyagramı [16].....	10
Şekil 1.8a.	Tüm insan vücudu [17].....	11
Şekil 1.8b.	Humerus kemiği ve cep telefonu uygulaması görseli [17].....	11
Şekil 1.9.	CNN modelinin mimarisinin görseli [17].....	11
Şekil 1.10.	El hareketi tanıma tabanlı cep telefonu uygulaması için genel akış [18]...	12
Şekil 1.11.	Geometri öğrenimi uygulamasında el hareketlerinin kullanımını gösteren görsel [23].....	15
Şekil 2.1.	Geliştirilen sistemin akış diyagramı	18
Şekil 2.2.	Sistem üzerinde tanımlanmış el hareketleri ve bu hareketlere karşılık gelen aksiyonlar	19
Şekil 2.3.	Video karesinde elin içerisinde olması beklenen ilgi bölgesi	20
Şekil 2.4.	VGG16 CNN modeli akışı.....	22
Şekil 2.5.	Evrişim işleminin gösterimi	23
Şekil 2.6.	ReLU işleminin girdisi ve çıktısı	23
Şekil 2.7.	Maksimum havuzlama işleminin girdisi ve çıktısı	24
Şekil 2.8.	Sonraki şekli çizme aksiyonu için el hareketi	29
Şekil 2.9.	Önceki şekli çizme aksiyonu için el hareketi	30
Şekil 2.10.	Mevcut Şekle 2B-3B dönüşümü uygulama aksiyonu için el hareketi	30
Şekil 2.11.	Mevcut şekli büyütme aksiyonu için el hareketi	31
Şekil 2.12.	Mevcut şekli küçültme aksiyonu için el hareketi	32
Şekil 2.13.	Mevcut şekli X ekseninde etrafında döndürme aksiyonu için el hareketi	32
Şekil 2.14.	Mevcut şekli Y ekseninde etrafında döndürme aksiyonu için el hareketi	33

Şekil 2.15. Mevcut şekil için çevre, alan ve hacim formüllerini gösterme aksiyonu için el hareketi	34
Şekil 2.16. Aktive etme aksiyonu için el hareketi	35
Şekil 2.17. Rotasyon yönünü değiştirme aksiyonu için el hareketi	35
Şekil 3.1. Eğitim işlemi için kullanılan orijinal görüntü ve çoğaltma işleminde kullanılan parametrelerin bireysel sonuçları	37
Şekil 4.1. VGG-16 mimarisi	39
Şekil 4.2. Sıradaki şekli çizme el hareketi ve karşılık gelen grafik operasyonu.....	46
Şekil 4.3. Mevcut şekil için 2 boyut – 3 boyut değiştirme el hareketi ve karşılık gelen grafik operasyonu.....	47
Şekil 4.4. Mevcut şekli küçültme el hareketi ve karşılık gelen grafik operasyonu.....	48
Şekil 4.5. Mevcut şekli büyütme el hareketi ve karşılık gelen grafik operasyonu	48
Şekil 4.6. Mevcut şekil için formülleri gösterme el hareketi ve karşılık gelen formülleri gösterme operasyonu	49
Şekil 4.7. Mevcut şekli Y ekseninde döndürme el hareketi ve karşılık gelen grafik operasyonu	50

SİMGELER VE KISALTMALAR LİSTESİ

AR	Augmented Reality
BSD	Berkeley Software Distribution
CAD	Computer Aided Design
CAM	Computer Aided Manufacturing
CGI	Computer Generated Imagery
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CV	Computer Vision
GPU	Graphics Processing Unit
HGR	Hand Gesture Recognition
HSV	Hue Saturation Value
ReLU	Rectified Linear Unit
RGB	Red Green Blue
RNN	Recurrent Neural Network
ROI	Region of Interest
VGG	Visual Geometry Group
VR	Virtual Reality

1. GİRİŞ

1.1. Çalışmanın Konusu, Amacı ve Kapsamı

Geometri konusu matematik bilimi içerisinde oldukça zor bir konu olup uzaysal düşünme, görselleştirme ve hayalde canlandırma gibi yetenekler gerektirir. Öğrenciler günümüzde de geometri öğrenimi konusunda zorluk çekmekte ve büyük çoğunluğu geometri çalışmaları sırasında kendilerine yeterince güvenmediklerini belirtmektedirler. Öğretmenler ve eğitmenler, geometri konusunda ideal öğretim tekniklerinin kullanılmasına gayret etmekte ve bunun için, geometri öğrenimini kolaylaştırmak ve öğrenilen bilgilerin kalıcı olmasını sağlamak için müfredat değişikliği yaparak, ders kitaplarını revize ederek ve eğitim materyallerini dijital ortama taşıyarak aktarılan bilgilerin işlevselliklerini ve kalitelerini yükseltme çabasındadırlar.

Bilgisayarların kullanımının yaygınlaşması ile bilimsel ve akademik alandaki gelişmelere paralel olarak eğitim ve öğretim alanında da hızlı bir ivmelenme dikkat çekmektedir. Artan teknolojik gelişmelere paralel olarak yeni araştırma ve uygulama alanları ortaya çıkmış, bu kapsamda özellikle son yirmi yıldaki teknolojik gelişmelerle birlikte çok sayıda araştırmacı örüntü tanıma ve bilgisayarlı görme alanları üzerine yoğunlaşmıştır.

Bilgisayarla görme, insan hareketlerini görüp algılayabilme yeteneğinden dolayı, bilgisayarlarla iletişim konusunda kullanılmaya çok elverişli bir alandır. Bilgisayarla görme kullanılarak, yüz tanıma, el hareketi tanıma, cinsiyet tespiti, hareket algılama gibi uygulamalar geliştirilebilir.

Geçtiğimiz 10 yıl içerisinde, bilgisayarla görme teknikleri kullanarak örüntü ve nesne tanıma konularında yapılan çalışmalarda ciddi bir artış olduğu görülmektedir. Özellikle son yıllarda el hareketi ve vücut hareketi tanıma yöntemleri kullanılarak doğrudan bilgisayar ile etkileşilebilen sistemler geliştirilmiştir. Bu sistemlerin geliştirilmesinde kullanılan en popüler yöntemlerden birisi Evrişimsel Sinir Ağları (Convolutional Neural Network-CNN)'dir [1]. CNN yaklaşımı, yüksek başarımlı oranı ve kullanım kolaylığı gibi avantajlar barındırmasının yanında yüksek hesaplama maliyeti ve büyük hacimli veriye ihtiyaç duyması gibi dezavantajlar da barındırmaktadır.

Bu tez çalışması kapsamında literatürde var olan bir CNN mimarisi, son katmanı 10 sınıf için sınıflandırma yapacak şekilde güncellenerek belirli el hareketlerini tanıyıp

anlamlandırabilen bir geometri eğitim aracına dönüştürülmüştür. İlköğretim çağındaki çocuklar için tasarlanan bu sistem, belirli el hareketlerini tanıyarak bazı geometrik şekillerin 2 ve 3 boyutlu olarak gösterimine, döndürülmesine, boyutlarının değiştirilmesine ve alan, çevre ve hacim formüllerinin gösterilmesine imkân tanımaktadır. Oldukça yüksek bir kullanım başarımı elde edilen bu sistem gerçek zamanlı olarak çalışabilmektedir. Bu çalışma ile, bazı kişiler için anlaşılması ve görsel olarak canlandırılması zor kabul edilen geometri eğitiminin daha çekici ve kolay hale getirilmesi amaçlanmıştır.

1.2. Benzer Çalışmalar

Son yıllarda bilgisayarla görme tabanlı birçok insan-bilgisayar etkileşimi uygulaması geliştirilmiştir. Özellikle el hareketi tanıma yoluyla yapılmış olan bu çalışmalar; genel kullanım uygulamaları, eğitim alanı uygulamaları ve geometri alanına yönelik uygulamalar olarak gruplandırılarak aşağıdaki bölümlerde detaylı bir biçimde aktarılmıştır.

1.2.1. El hareketi tanımlama kullanılarak geliştirilen genel kullanım uygulamaları

S.S Rautaray ve A. Agrawal [2] tarafından önerilen çalışmada, C++ ve OpenCV kullanılarak, bir oyun kontrol sistemi geliştirilmiştir. Oyun; “tutma”, “yumruk atma”, “fırlatma” ve “hareket etme” fonksiyonlarını barındırır ve herbirine karşılık gelen bir el hareketi sayesinde bu işlemler gerçekleştirilir. Görüntüler kameradan alınır ve çeşitli aşamalarda işlemler uygulanır. İlk önce arkaplan çıkartma yöntemi uygulanır ve elin konumu tespit edilir. Bu işlemler bittikten sonra elin pozisyonunu bulmak ve hareketi sınıflandırmak için Haar Cascade Classifier (Haar Cascade Sınıflandırıcısı) [3] [4] yöntemi kullanılır. Haar-benzeri öznitelikler [4], gürültü ve aydınlanma değişimlerine karşı güçlü olduğu için bu yöntem seçilmiştir. Elin takibi, arka plan çıkartma ve hareket algılama ile yapılmıştır. Oluşan gürültü, aşındırma ve genişletme fonksiyonları kullanılarak giderilmiştir. Elin takibi için Camshift [5] ve Lukas Kanade Optical Flow (LKOF) [6] teknikleri kullanılmıştır. LKOF, Camshifti daha gürbüz bir hale getirmek için kullanılmıştır. El takip edilirken, karşılık gelen el ile bir kontur eşlenir ve bu da karşılık gelen dışbükey zarfı (convex hull) çıkartır. Sistem el etrafında çizilen dışbükeylik kusurlarını (convexity defect) takip eder ve önceden ayarlanmış bir komutla eşleştirir. 4 tane el hareketi için tanımlama yapılmıştır. Sanal oyun ile 15 kişinin test etmesi sonucu

sonuçlar elde edilmiştir. Işıklandırma değişimi ve kullanıcı sayısı gibi farklı parametrelerin değişimleri ile testler yapılmıştır. Sonuçlarda “tutma”, “yumruk atma”, “fırlatma” ve “hareket etme” hareketlerinin başarı oranları sırasıyla, %93.33, %86.67, %80.33, %93.33 olarak bulunmuştur. Sistemin akışı Şekil 1.1.’de görülmektedir.

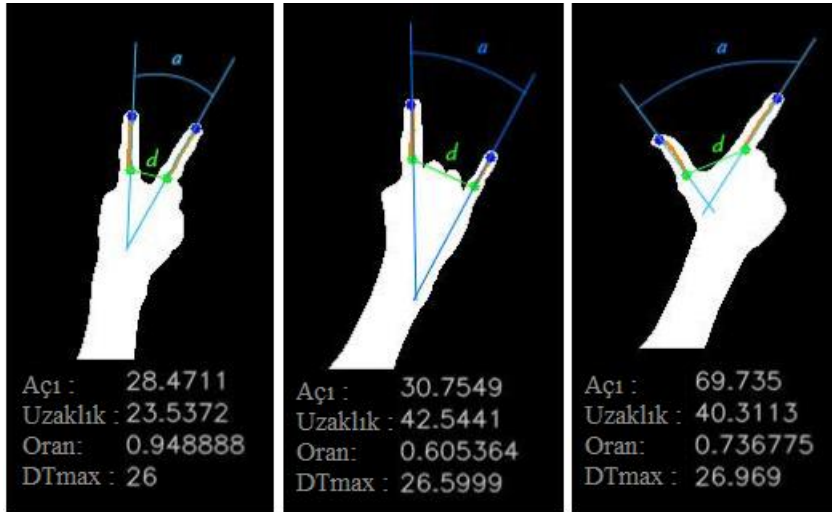


Şekil 1.1. El hareketi algılama tabanlı kontrollü oyun sistemi için genel akış [2]

H. Jalab ve H. Omer [7], HGR yardımıyla kullanılan bir medya oynatma sistemi geliştirmişlerdir. Medya oynatma sistemlerinde oldukça önemli olan “oynatma”,

“durdurma”, “ileri sarma” ve “geri sarma” işlemlerinin desteklendiği bu sistemde %95’lik bir başarı oranı sağlanmıştır.

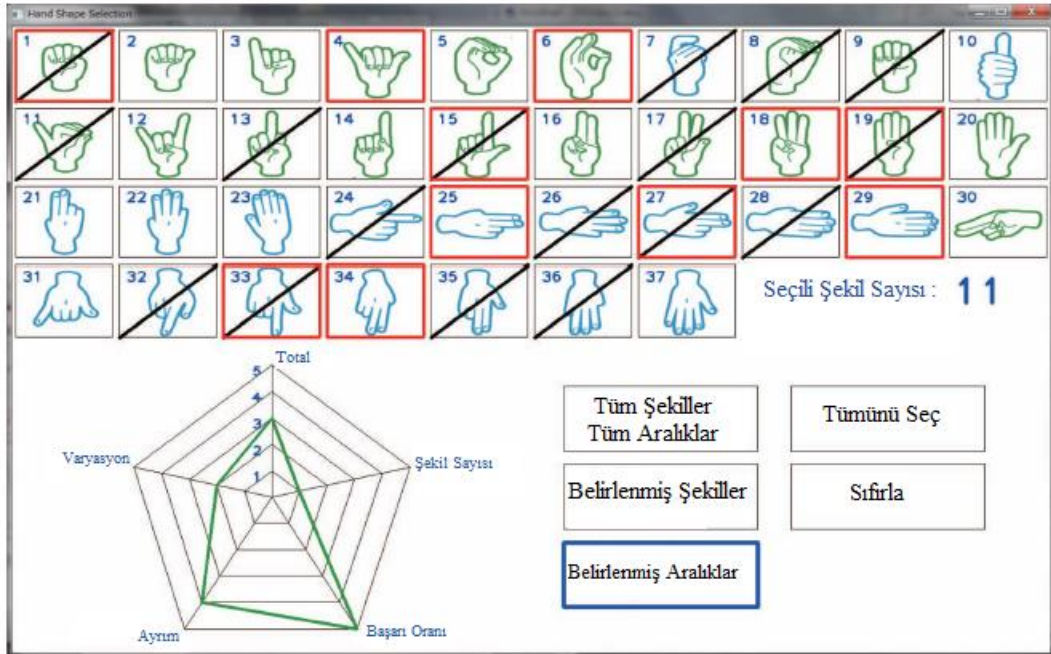
M. Czupryna ve M. Kawulok [8] görüntü tabanlı çalışan ve basit bir web kamerası ile görüntü almanın yeterli olduğu bir fare imleç kontrol uygulaması geliştirmişlerdir. Fare modu ve hareket modu olmak üzere 2 farklı modu vardır. Fare modu 4 hareketten (fare hareket kontrolü, sol tıklama, sağ tıklama ve kaydırma), hareket modu ise 8 hareketten oluşmaktadır. Görüntü alındıktan sonra HSV renk modeli kullanılarak cilt rengi algılaması yapılır. Cilt rengi algılayıcı elle konfigüre edilerek özelleştirilebilir ve uygulama önceden tanımlanmış standart değerler de sunabilmektedir (tungsten ışık için, güçlü güneş ışığı için gibi). El özniteliklerinin çıkartılması, ilk önce elin bölütlemesiyle başlar, daha sonra el merkezi, bilek noktası, avuç yönü ve her parmak için; parmak tabanı, parmak ucu, boyut, yön gibi özellikler çıkarılır. Hareketin sınıflandırılması parmakların görünümü ve yerlerine göre çıkarılan özelliklerle yapılır. Bu özellikler 2 parmak arasındaki açı, parmakların boylarının birbirine oranı ve parmak tabanı mesafesidir. Bu özellikler baz alınarak yapılan hesaplamalara ait görsel Şekil 1.2.’de görülebilir. Sistem kapalı ortamda ışıklandırılmış şartlarda ve el cilt rengine yüksek zıtlık oluşturan bir arka planda test edilmiştir. 50 görüntü üzerinde test yapılmış ve bu görüntülerin 48’inde başarılı sınıflandırma yapılarak %96 gibi yüksek bir başarı oranı elde edilmiştir.



Şekil 1.2. Fare kontrol sistemi için belirlenen özellikler üzerinde yapılan hesaplamalar [8]

Shimada vd. [9] kullanıcıların kendilerinin özelleştirebileceği bir televizyon kontrol arayüzü geliştirmişlerdir. TV üzerine yerleştirilmiş bir web kamerası ile video çerçeveleri okunur. El tespiti, cilt rengindeki pikselleri arayan bir modül tarafından yapılır. Yüzle

karıştırılmaması için el önceden tanımlanmış bir dikdörtgen içerisinde işleme sokulur. Bu dikdörtgen kullanıcı tarafından özelleştirilebilir. El takip edilir ve tespit edilen elin alanı el hareketinin tanımlanması için kullanılır. Dinamik el hareketleri için elin hareketini algılamada elin gezingesi kullanılır. Hangi el hareketinin hangi işi yapacağı da kullanıcı tarafından ayarlanabilir ve sınıflandırılan hareket gerekli işlemi yapmak için tanımlanmış komutu çalıştırır. Bu komut televizyona yönlendirilir ve önceden tanımlanan; “güç düğmesi açma”, “kanal artırma”, “kanal azaltma”, “ses artırma”, “ses azaltma”, “tamamen sessize alma” gibi işlemleri yapar. Hem kullanıcı hem de makine dostu olması hedeflenen bu sistemde, kullanıcı dostluğu kullanıcının kendi seçim yapabilmesi ve sistemi istediği gibi özelleştirebilmesi ile ilişkilendirilmiş olup, makine dostluğu da bilgisayar tarafından kolayca tanınacak el hareketlerinin tanımlanmasından kaynaklanmaktadır. Şekil 1.3.’te bu sistemin özelleştirme ekranı gösterilmiştir.

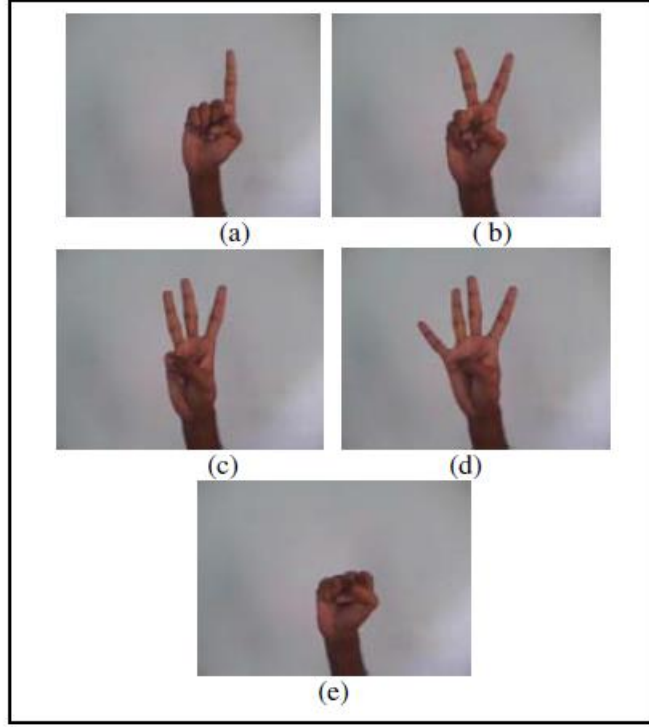


Şekil 1.3. El hareketi algılama tabanlı televizyon kontrol sisteminin özelleştirme ekranı görseli [9]

Peng vd. [10], kullanıcıların edinmek istediği günlük verileri (hava durumu, haberler, günlük fal) HGR yardımıyla alabilen bir sistem önermişlerdir. YCbCr renk uzayı ile ten rengi tespiti yapılarak el tespit edilir ve Camshift ile takip edilerek bölütleme gerçekleştirilir. Bu aşamadan sonra normalleştirme yapılır. Normalleştirme ile orijinal görüntüden, örnekleme azaltılmış (downsampled) bir görüntü elde edilip, ten bölgesi tespit

edilir. Ardından görüntü ikili hale getirilerek (binarization) en büyük ten bölgesi elde edilir. Bu bölge 100*100 çözünürlüğe getirilir. 6 hareket için her bir hareketin 5 farklı açıdan fotoğraflarıyla eğitilen model, Principal Component Analysis (PCA) kullanılarak sınıflandırma işlemini gerçekleştirir. Belirlenen el hareketleri için başarımları; “sıfır” hareketi için %93.5, “bir” hareketi için %96.9, “iki” hareketi için %86.1, “beş” hareketi için %94.1, “yedi” hareketi için %97.8 ve “kapalı avuç” hareketi için %90.3 olmak üzere, toplam başarımları %93.1 olarak ölçülmüştür.

Rajesh vd. [11] Microsoft Powerpoint sunumlarını kontrol etmek için bir sistem geliştirmişlerdir. Standart bir ağ kamerası ile çalışan bu uygulamada ilk önce kameradan görüntüler alınır. Bundan sonraki aşamada elin ayrıştırılması için arkaplan bölütlemesi yapılır ve RGB olarak alınan görüntü HSV renk modeline çevrilir. Daha sonra, bu görüntü ikili (binary) görüntüye dönüştürülür. Hareket sınıflandırması için aktif parmaklar sayılır. Bunun için elin ağırlık merkezi bulunur ve en uzun parmağın boyutu 0.7 ile çarpılarak yarıçap hesaplanır. Bu hesaplamalar sonucu oluşan merkez nokta ile yarıçap değerleriyle çizilen dairenin kesiştiği parmaklar aktif parmaklar olarak kabul edilir. Beyazdan siyaha geçişleri sayan bir graf oluşturulmuştur ve buradaki geçiş sayısı aktif parmak sayısını vermektedir. Bu yöntemde açık parmak sayısı aynı olan hareketler ayırt edilememektedir ve kullanılan 5 hareketin hepsinde aktif parmak sayısı farklıdır. Hareketler, işlevleri bakımından; “sonraki slayt”, “önceki slayt”, “slayt gösterisini başlat”, “slayt gösterisini bitir” ve “iki hareket arası geçiş yap” hareketleri olarak listenebilir. Başarı oranları sağ ve sol eller için ayrı ayrı hesaplanıp, “sonraki slayt” hareketi için %79.4, “önceki slayt” hareketi için %94.1, “slayt gösterisini başlat” hareketi için %96.1 ve “slayt gösterisini bitir” hareketi için %100 olarak ölçülmüş ve ortalama olarak %90’ın üzerinde bir başarımları sağlanmıştır. Bu sistemde kullanılan el hareketleri Şekil 1.4.’te görülebilir.



Şekil 1.4. Microsoft Powerpoint kontrol sistemi için kullanılan el hareketleri görseli [11]

El hareketi tanıma temelli uygulamalar, işitme engelliler için de çok büyük destek sağlamıştır. İşitme engelli kişilerin birincil iletişim mekanizması işaret dilleridir ve bu iletişim şekli HGR tabanlı bir uygulama geliştirmek için oldukça müsaittir.

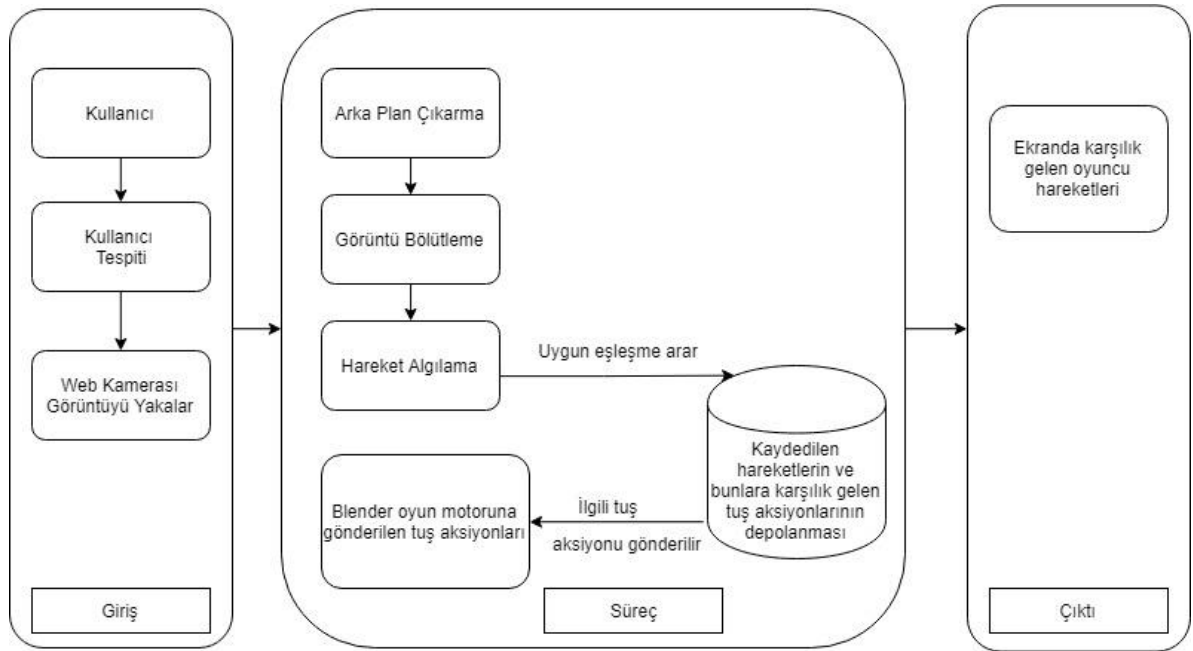
Tripathi vd. [12] tarafından gerçekleştirilen bir çalışmada, Hint işaret dili ile iletişim kurulması sağlanmıştır. Çalışmada, 2, 3 ya da 4 el hareketinin kombinasyonlarından oluşan toplam 10 farklı cümlelerin sınıflandırılması sağlanmıştır.

Benzer başka bir çalışma da Kim vd. [13] tarafından Kore işaret dilinin anlaşılması için ortaya konulmuştur. Kore işaret dilindeki 15 cümlelerin algılanmasını sağlayan bu uygulama %94 başarı oranıyla oldukça yüksek yüzdeli bir başarı ortaya koymuştur.

1.2.2. El hareketi tanımlama kullanılarak geliştirilen eğitim alanı uygulamaları

HGR tabanlı uygulamalar, eğitim konusunda da sıkça kullanılmakta ve özellikle eğitim çağındaki çocukların öğrenme süreçlerini daha kolay ve daha eğlenceli hale getirebilmektedir.

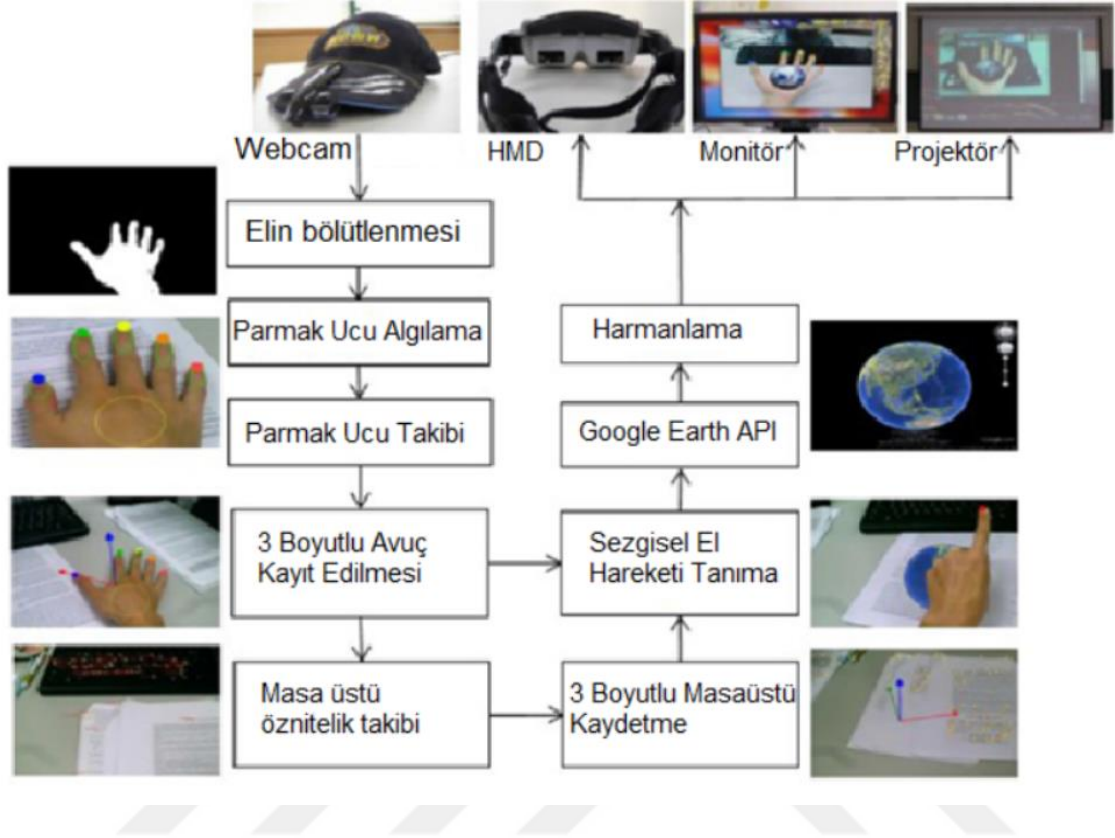
Mathmazing HGR tabanlı bir aritmetik eğitim uygulamasıdır. Sahasrabudhe vd. [14] tarafından geliştirilen uygulamada görüntü alma aracı olarak Kinect, oyun motoru olarak da Blender araçları kullanılmıştır. Bu uygulamada, ekranda bir aritmetik sorusu ve bir labirentin farklı uçlarında da cevap adayları yer alır. Kullanıcının el hareketleri ile labirente ilerleyerek doğru cevaba ulaşması beklenmektedir. Gosalia vd. [15] bu sistemde İngilizce dışında dil desteği olmaması, Kinect'e bağımlı olması, labirentin her seferinde aynı olması ve el hareketlerinin kısıtlılığın gidermek için çeşitli öneriler sunmuştur. Devanagari alfabesi için dil desteği, Kinect yerine web kamerası, yeni kullanıcı arayüzü ve her seviye için farklı labirentler ve yeni komutlar için yeni el hareketleri gibi eklentiler sunulan öneriler arasında yer almakta olup sistemin akış şeması Şekil 1.5.'te gösterilmiştir.



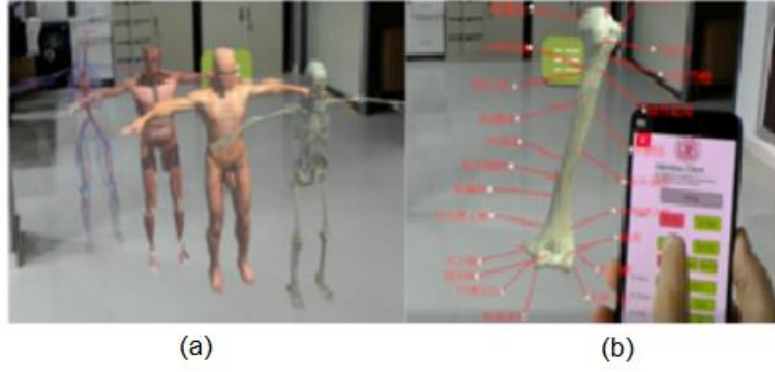
Şekil 1.5. El hareketi algılama tabanlı animasyon oyunu genel akışı görseli [14]

M.Yang ve W. Liao [16], VECAR adında, AR ve CV desteği olan bir sanal İngilizce sınıfı uygulaması geliştirmişlerdir. İngilizce ve İngiliz kültürü öğrenmek için tasarlanan bu uygulamada, öğrenciler fiziksel bir sınıfta ya da internet bağlantısı olan uzak bir konumda bulunabilmektedirler. Tüm katılımcılar HMC (Head Mounted Camera) kullanarak canlı görüntüler çekebilmekte ve bu görüntüler AR ve CV algoritmaları ile işlenmektedir. Çıktı görüntüleri, öğretmenler için projeksiyon ekranında, öğrenciler için HMD (head mounted display)'lerde gösterilebilmektedir. HGR ile avuç içerisinde ve masa üzerinde nesne manipülasyonu yapılabilmektedir. Sonuçlar göstermiştir ki VECAR, kültürel öğrenme ve öğrenci öğretmen arasındaki iletişimi geliştirmiştir. VECAR sisteminin akış şeması Şekil

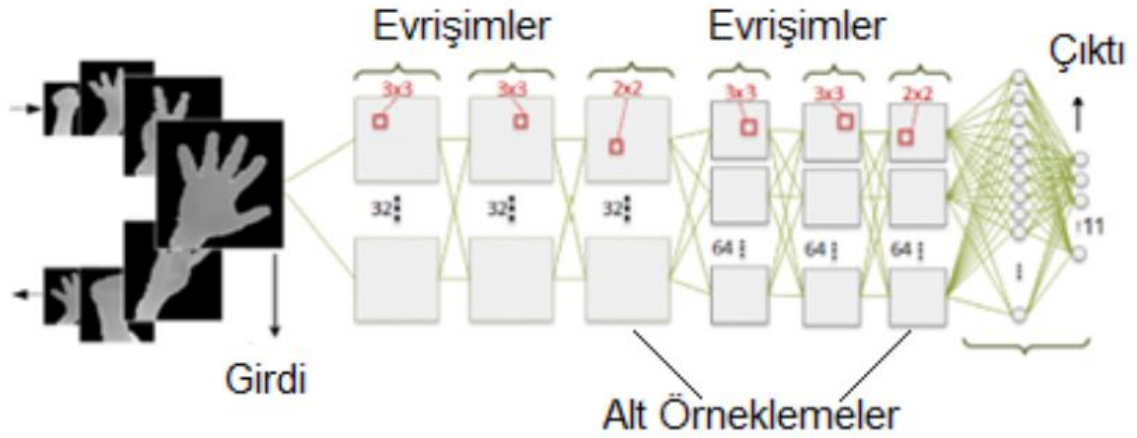
1.6.'da, 3 boyutlu nesne manipölasyonu için durum diyagramı ise Şekil 1.7.'de gösterilmiştir.



Şekil 1.6. VECAR sisteminin akış şeması görseli [16]



Şekil 1.8. (a) Tüm insan vücudu (b) Humerus kemiği ve cep telefonu uygulaması görseli [17]



Şekil 1.9. CNN modelinin mimarisinin görseli [17]

Yang vd. [18], erken yaşta matematik öğrenimi için oluşturulan ve HGR ile çalışan bir ürün tasarlamışlardır. Bu ürün sistem maliyeti açısından hafif bir ürün olup cep telefonlarında çalışabilmektedir. Elin bulunması için YCbCr renk modeli kullanılarak ten rengi tespiti yapılmıştır. Elin takip edilmesi ve ilgi alanının (ROI) sürekli güncellenmesi için CamShift yöntemi kullanılmıştır. Sınıflandırmadan önce morfolojik olarak gürültü giderme (3 daraltma, 5 genişletme) ve etiketleme yapılmıştır. Parmakların ardışık çevre çizgisi noktalarını bulmak için Snakes [19] modeli kullanılmıştır. Merkez ve çevre çizgileri arasındaki mesafenin grafikleri çizilmiştir. Grafikteki uç nokta sayıları sınıflandırma için kullanılmıştır. Bu yüzden parmak sayısı aynı olan hareketlerin algılanmasını desteklenmemektedir. Ayrıca, sistem herhangi bir öğrenme veri kümesine ya da eğitim sürecine ihtiyaç duymamaktadır. Bu sistemin genel akışı Şekil 1.10.'da görülebilir.

alıřan 3 farklı komut yer almakta olup, bu komutlar; “döndürme”, “zum” ve “hareket ettirme”’dir. Uygulamanın işlevselliğine dair bazı özellikler; 2 boyutlu şekillerden 3 boyutlu şekiller üretilmesi, yapısal uzay geometrisi (birleşim, kesişim, fark gibi işlemler) olarak sıralanabilir.

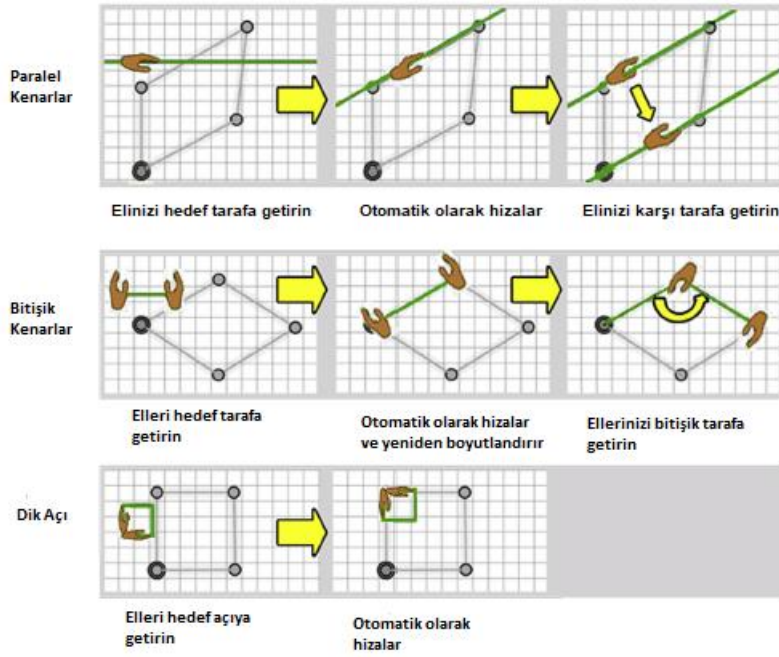
Sonuçlar ařağıdaki 4 metrik üzerinden kullanıcılara yapılan anketlere göre değerlendirilmiştir. Bu metrikler, sezgisellik (*3 boyutlu geometrik şekilleri gözünüzde canlandırabiliyor musunuz? Bu şekiller ve işlemlerini anlamak ve bunların yeni özellikleri keşfetmek sizin için kolay mı?*), kullanım kolaylığı (*Uygulamayı kullanırken nasıl hissediyorsunuz?, Tüm el hareketlerini hatırlamak ve kullanmak sizin için ne kadar kolay?*), anlama yeteneğı (*Geometrik şekil konseptini anlamak sizin için ne kadar kolay?, Bu uygulamayı kullanarak 3 boyutlu geometrik şekillerle ilgili gelişmiş özellikleri anlamak sizin için ne kadar kolay?*), dinamik değışiklik ve etkileşim (*Uygulama içerisindeki eylemleri gerçekleştirmek ve 3 boyutlu şekillerle etkileşmek sizin için kolay mı? Etkileşim sırasında 3 boyutlu geometrik şekillerin simüle edilmesi yeterince başarılı mı?*) olarak sıralanabilir. Bu soruların kullanıcılar tarafından 5 aralıklı bir cevaplandırma sisteminde (Çok zor, zor, normal, kolay, çok kolay) yanıtlanması istenmiştir. Kullanıcılar; sezgisellik metriğinde %82, kullanım kolaylığında %73, anlama yeteneğı metriğinde %75 ve dinamik değışiklik ve etkileşim metriğinde ise %80 oranlarında “kolay” ya da “çok kolay” seçeneğini seçmişlerdir.

Kaufmann vd. [21] VR tabanlı bir geometri uygulaması geliştirmişlerdir. Bunun için bir Kişisel Etkileşim Paneli (Personal Interaction Panel) [22], kullanılmıştır. Bu sistem pozisyonu takip edilen bir dijital bloknot ve kalemten oluşur ve uygulamayı kontrol etmek için kullanılır. Nokta, doğru, düzlem, küp, küre, koni ve silindir olmak üzere 7 tane şeklin oluşturulmasını sağlar. Uygulama, 6’sı kadın 8’i erkek olmak üzere, yaşları 22-34 arasında değışen 14 katılımcı ile test edilmiştir. Bu kullanıcıların 9’u matematik ve geometri öğrencileri olup tüm katılımcılar temel bilgisayar yeteneklerine sahiptirler. Bu kullanıcılara tamamlamaları için bir geometrik görev verilir. Bu görevi tamamlamaları 6 ila 13 dakika arasında sürmekte olup tüm kullanıcılar VR’ı tekrar deneyimlemek istediklerine ve çok güzel bir ortam olduğuna dair fikir beyan etmiştir. Bir öğrenim ortamı olarak VR kullanılması katılımcıların tamamı tarafından mantıklı bulunmuş ancak daha büyük bir öğrenci grubu ile kolaboratif bir çalışma yapılmasının önemli olacağına dair sorular ortaya çıkmıştır.

Vitale vd. [23] tarafından yapılan bir başka çalışmada, geometri öğreniminin sezgisel ve alışılmışın dışında konseptlerin bütünleştirilmesine ihtiyaç duyduğu önermesi üzerinden

yola çıkılarak, bu konseptlerin öğrenilmesinin gerçekleştirme tabanlı bir yöntemle daha başarılı olacağı ve daha derin bir kavrama sağlayacağı ortaya atılmıştır. Küçük çocukların bazı nesneleri (taksi, ada) algılamada ayırt edici özellikler yerine göze çarpan özellikleri kullanmaları gözlemlenmiştir. İlk öğretim çağındaki çocukları gerçekçi bütünleme (Grounding Integration – GI) ve sayısal bütünleme (Numerical Integration – NI) olarak iki farklı yöntem kullanarak, bilgisayar ortamında dinamik bir geometri öğrenme uygulaması üzerinde çalıştırılması sağlanmıştır. GI grubu, şekillerin özniteliklerini (dik açı, paralel kenar vb.) el hareketleriyle temsil ederek (elleriyle dik açı yaparak, ellerini birbirine paralel tutarak, vb.) belirtmişlerdir. NI grubunda ise, şekillerin öznitelikleri tamamen sayısal bir şekilde çıkartılmıştır. GI grubundaki öğrencilerin, çalışma sonrasında yapılan testlerde hedef şekilleri ve özniteliklerini daha başarılı bir şekilde tanımladıkları görülmüştür.

Bu uygulamada bir robotun bir parkuru, hedef nesneleri toplayarak ve engel nesnelerden kaçarak tamamlaması için bir kapalı yol çizilmesi amaçlanmaktadır. Bu yolların çizimi için kare, dikdörtgen, paralelkenar, yamuk ve üçgen gibi şekiller kullanılır. Ayrıca, el hareketi algılama da kullanılarak paralel olan kenarlar, eşleşik olan komşu kenarlar, dik açılar belirlenebilmektedir. Bu operasyon şekli GI olarak belirtilir. Sonuç olarak, aynı uygulamayı GI şeklinde kullanan çocukların, sadece sayısal temsillerle işlem yapan (NI) çocuklara oranla şekilleri tanımlamada daha başarılı olduğu görülmüştür. Bu çalışma sonunda geometri öğreniminin uzaysal gerçekçi (spatially grounded) olması gerektiği sonucu çıkmıştır. Bu uygulamanın el hareketi algılama kullanılan bir arayüzü Şekil 1.11.’de gösterilmiştir.



Şekil 1.11. Geometri öğrenimi uygulamasında el hareketlerinin kullanımını gösteren görsel [23]

1.3. Tez Planı

Eğitim çağındaki çocuklara yönelik el hareketi tanıma tabanlı geometri öğrenme uygulaması konulu bu çalışmanın birinci bölümünde çalışmanın konusu amacı ve kapsamı açıklanmıştır. Ayrıca, tez konusuna yakın literatür bilgilerine yer verilmiştir. Bu bağlamda; el hareketi tanıma kullanılarak geliştirilen genel kullanım uygulamalarına, el hareketi tanıma kullanılarak geliştirilen eğitim alanındaki uygulamalara ve geometri eğitimi alanında geliştirilen bazı uygulamalara ve yine geometri eğitimi için geliştirilen ve el hareketi tanıma tabanlı çalışan uygulamalara değinilmiştir.

Tezin ikinci bölümünde geliştirilen sistemin bileşenlerine ve kullanılan yöntemlere yer verilmiştir. İlgili bölümün alt başlıklarında sistemin el hareketi sınıflandırma için kullandığı CNN yapısından bahsedilmiş ve bu yöntem içerisinde kullanılan evrimsel katman, ReLU işlemi, havuzlama katmanı ve tam bağlı katman kavramları açıklanmıştır. Ayrıca, bilgisayar grafiklerinde ve sistemin kullanıcı arayüzünde görülen grafik çizme işlemlerinden bahsedilmiştir. Bilgisayar grafiklerinin ve grafik çizimleri hakkında tarihçe ve genel konseptlerden bahsedilerek sistemin gerçekleştirdiği grafik işlemlerine genel anlamda değinilmiştir.

Tezin üçüncü bölümünde veri kümesi anlatılmış olup tezde kullanılan veri kümesinin oluşturulması, sistemin eğitilmesi ve test edilmesi hakkında bilgiler verilmiştir.

Çalışmanın dördüncü bölümünde sistemin geliştirilmesine yönelik süreç ve değerlendirmeler üzerinde durulmuştur. Bu bağlamda, sistemin geliştirilme ve test süreçlerinin gerçekleştirildiği donanımlar ve kullanılan teknolojilerden bahsedilmiştir. Ayrıca, sistemin çalışma prensipleri, kullanım sonuçları, başarımları hakkında değerlendirmeler yapılmış ve sistemin farklı kullanım alanları ve hangi konularda geliştirilebileceğinin tartışması üzerinde durulmuştur.

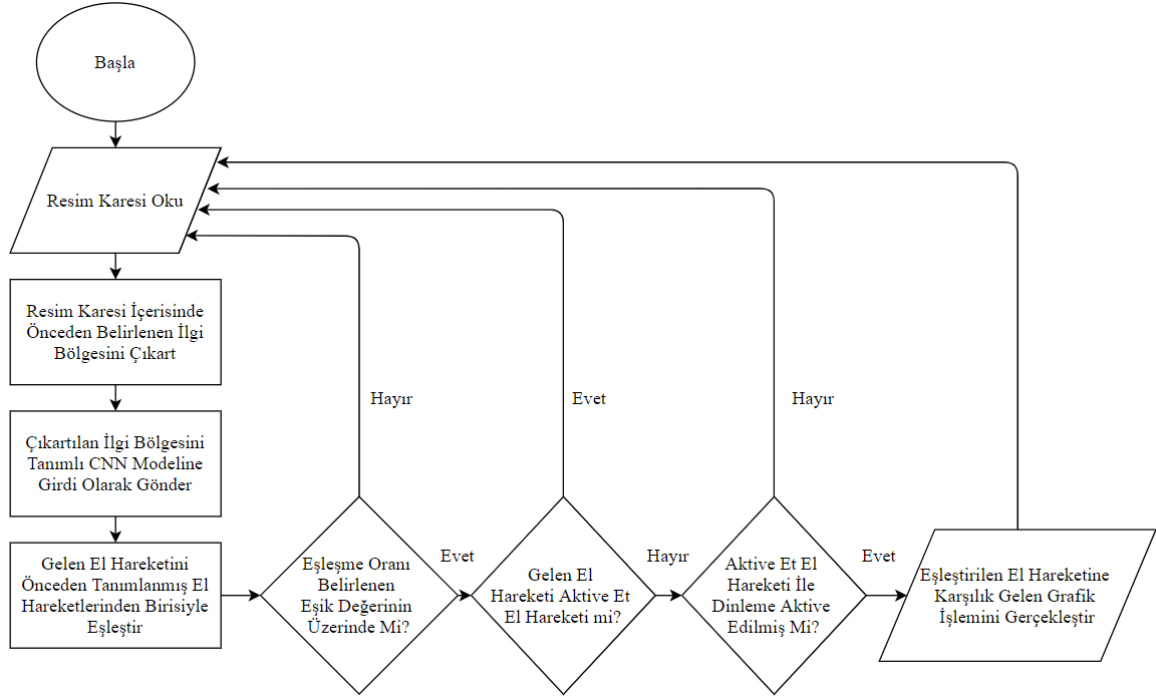
Son bölümde çalışmanın genel bir özetine yer verilmiş, elde edilen sonuçlara ve sonuçların ne şekilde iyileştirilebileceğine değinilmiş ve bu çalışmanın taban olarak kullanılıp ne gibi çalışmalar şeklinde genişletilebileceğinden bahsedilmiştir.



2. SİSTEMİN GENEL YAPISI VE KULLANILAN YÖNTEMLER

2.1. Sistemin Yapısı

Bu çalışmada geliştirilen el hareketi tanıma tabanlı geometri öğrenme uygulaması, kullanıcı ile el hareketleri aracılığıyla iletişim kurarak, geometri öğrenimini kolaylaştırmayı amaçlayan çeşitli özellikler sunar. Uygulama Python programlama dili kullanılarak geliştirilmiş olup işlem adımları şu şekilde özetlenebilir: standart bir ağ kamerası üzerinden görüntü akışı başlatılır, alınan görüntünün üzerinde ilgi bölgesi olarak tanımlanan bir bölge çizilir. Kullanıcının elinin bu bölgede olması beklenmekte olup el hareketinin tanınması için gerekli olan işlemler, sadece bu bölgedeki görüntü üzerinde yapılmaktadır. Görüntü akışı devam ederken, görüntü kareleri tek tek alınır ve ilgi bölgesi kırılarak tanımlanmış CNN modeline girdi olarak verilir. Model, gönderilen çıktıyı daha önceden tanımlanmış olan 10 adet el hareketiyle karşılaştırır ve bu 10 el hareketi sınıfının her birine ait olma ihtimalini hesaplar. Elde edilen bu 10 adet olasılık değerinin en yükseği seçilir ve önceden tanımlanmış bir eşik değeri ile karşılaştırılır. Eğer, seçilen değer bu eşik değerinden küçük ise sistem sınıflandırmayı yapmaz ve bir sonraki görüntü karesine geçer. Eğer seçilen değer, eşik değerinden büyük ise sınıflandırma başarılı olmuştur ve kullanıcının el hareketi, önceden tanımlı el hareketlerinden biri ile eşleştirilir. Her el hareketi algılandıktan sonra sistem, işlem yapmanın aktive edilip edilmediğini kontrol eder. Bunun sebebi sürekli görüntü akışı olması nedeniyle istenmeyen işlemlerin yapılmasının engellenmesidir. Aktivasyon işleminin gerçekleşmesi için özel olarak tanımlanmış bir el hareketi belirlenmiştir. Sistem bu hareketi algıladığında aktif hale geçer, diğer herhangi bir el hareketini görüp, o harekete karşılık gelen işlemi yaptıktan sonra tekrar kendini inaktif hale getirir. Eğer sistem aktif değilse ve algılanan hareket “aktive et” hareketiyse, sistem aktif duruma geçer. Eğer başka herhangi bir hareket algılandıysa işlem yapılmaz. Sistemin genel akış diyagramı Şekil 2.1.’de verilmiştir.

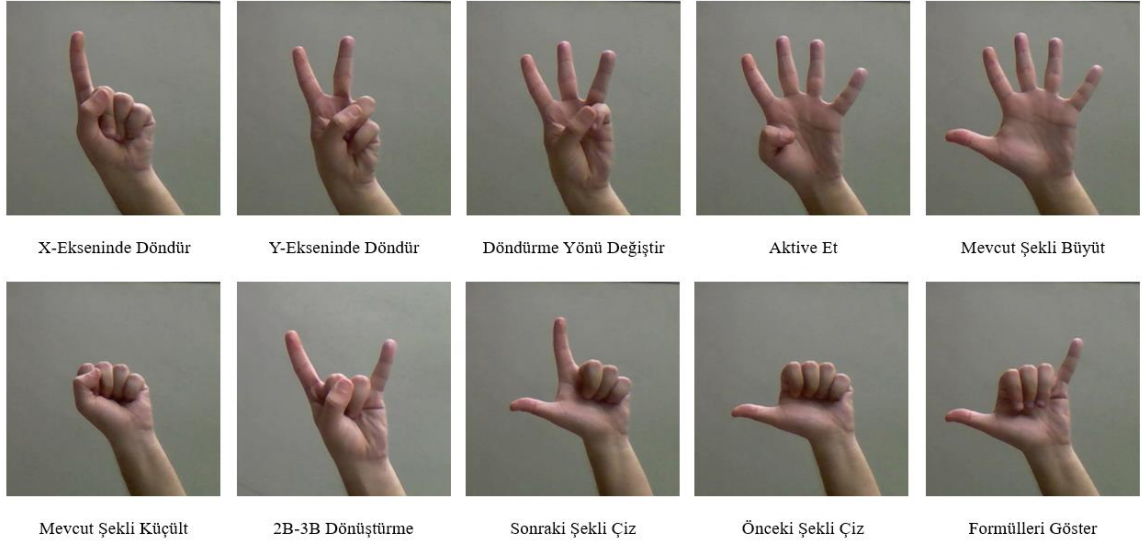


Şekil 2.1. Geliştirilen sistemin akış diyagramı

Sistem aktif haldeyken algılanan hareketler, kendilerine karşılık gelen grafik işlemlerini gerçekleştirir ve sistemi bir sonraki “aktive et” hareketini görene kadar pasif duruma getirir. Bu sistemde toplam 10 adet el hareketi tanımlıdır ve bunlara karşılık gelen grafik aksiyonları,

- Aktive Etme
- Sonraki Şekli Çizme
- Önceki Şekli Çizme
- Mevcut Şekli Büyütme
- Mevcut Şekli Küçültme
- Mevcut Şekli X Ekseninde Döndürme
- Mevcut Şekli Y Ekseninde Döndürme
- Dönüş Yönünü Değiştirme
- Mevcut Şekle 2 Boyut - 3 Boyut Dönüşümü Uygulama
- Mevcut Şekil İçin Çevre, Alan ve Hacim Formüllerini Gösterme

şeklinde tanımlanmıştır. Burada listelenen aksiyonlara karşılık gelen el hareketleri Şekil 2.2.’de gösterilmiştir.

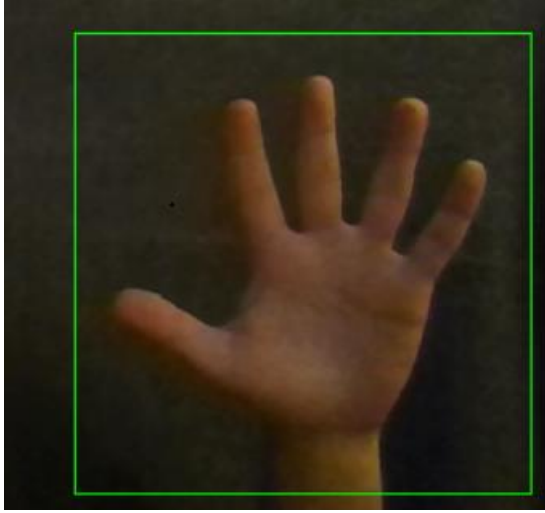


Şekil 2.2. Sistem üzerinde tanımlanmış el hareketleri ve bu hareketlere karşılık gelen aksiyonlar

El hareketi algılama gibi gerçek zamanlı görüntü üzerinde çalışan uygulamalarda donma, takılma ve görüntü kaybı gibi sorunlar yaşanabilir. Bu kapsamda, uygulamadaki CNN modeli GPU üzerinde sınıflandırma işlemini gerçekleştirecek şekilde konfigüre edilmiştir. Bu konfigürasyon, tespitin CPU üzerinde yapılmasına oranla daha hızlı bir şekilde çalışmasını sağlamaktadır.

2.2. Evrişimsel Sinir Ağları ile El Hareketi Algılama

Tez kapsamında geliştirilen bu sistemde, el hareketlerinin yüksek yüzde ile doğru bir şekilde sınıflandırılabilmesi için, ilk aşamada, sisteme girdi olarak verilen görüntülerde el bölgesinin net bir şekilde algılanabilmesi gerekir. Alınan görüntülerde, elin içerisinde bulunması gereken bir ilgi bölgesi ekrana çizilmiştir ve kullanıcının elinin bu bölgede olması beklenmektedir. Bu ilgi alanı bölgesi Şekil 2.3.'te gösterilmiştir.



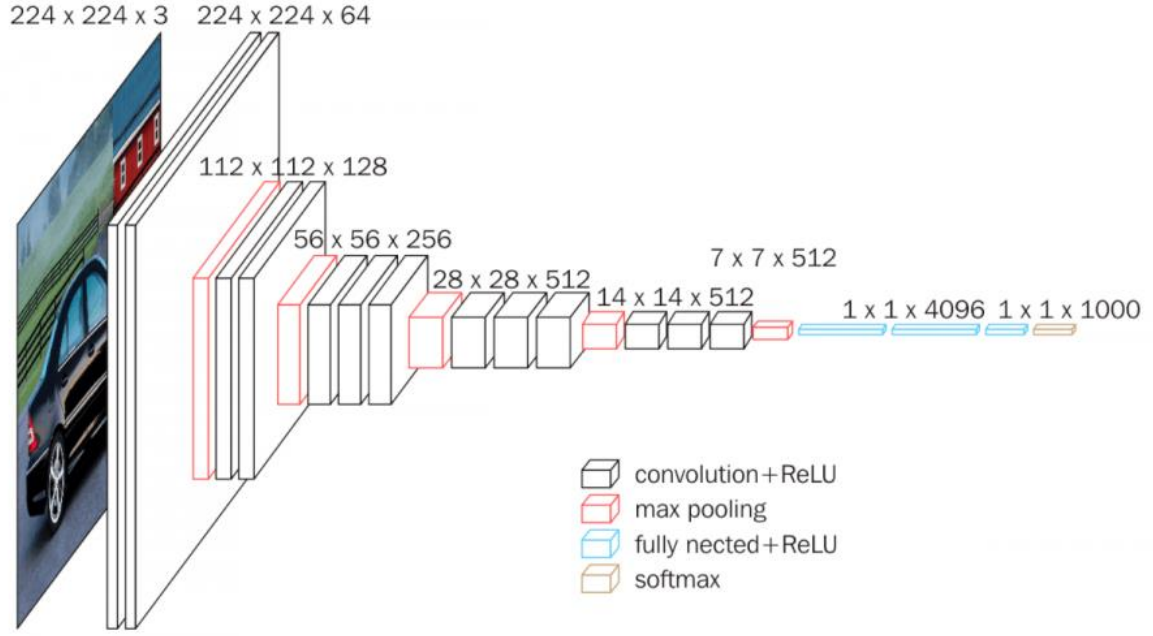
Şekil 2.3. Video karesinde elin içerisinde olması beklenen ilgi bölgesi

Bir görüntü karesi alındıktan sonra, görüntünün sadece bu ilgi bölgesi çıkartılmakta, geriye kalan kısımları kırpılmaktadır. Sistem, sade bir arka planda çalışabilmektedir. Arka plan karmaşıklaştığında başarı oranı düşmektedir. Işıklandırma değişimine oldukça dayanıklı olan sistem, çok şiddetli ışık değişimleri yaşanmadığı takdirde yüksek başarı oranı ile çalışmaya devam etmektedir. Sistem, her insanın el hareketlerini farklı bir şekilde yapacağını öngörerek, olabildiğince esneklik sağlamaktadır. En büyük gereksinim olarak, ilgi bölgesindeki elin, karşıdan net bir şekilde görünebilmesi gerekmektedir. Avuç içi doğrultusunun kameraya uzaklığı arttıkça, yani karşıdan görünmek yerine yandan görünmeye başladıkça, başarı oranı düşmektedir. Sistem, elin karşıdan görüldüğü durumlarda çevrinme (pan), eğilme (tilt) ve boyuttan bağımsız bir şekilde yüksek oranda sınıflandırma yapabilmektedir. Görüntü alınması ve karelere bölünmesinde OpenCV kütüphanesinin 3.4.0 versiyonu kullanılmıştır. OpenCV görüntü işleme ile ilgili çok sayıda bilgi ve fonksiyon içeren açık kaynak kodlu ve görüntü işleme, bilgisayarla görme gibi alanlarda sıkça kullanılan bir kütüphanedir [24]. El hareketlerinin sınıflandırılması için Keras kütüphanesinin 2.1.4 versiyonu kullanılmıştır. Keras yüksek seviyeli bir yapay sinir ağı uygulama geliştirme arayüzüdür. Python tabanlı geliştirilmiş olup, Tensorflow, CNTK ve Theano gibi platformlar üzerinde çalışabilir [25]. CNN, RNN ve her ikisinin kombinasyonlarını destekleyip hem CPU hem de GPU üzerinde sorunsuz bir şekilde çalışacak biçimde konfigüre edilebilir. Sınıflandırıcı olarak, Keras ile geliştirilen bir CNN modeli kullanılmıştır. Bu CNN modelinin mimarisi, AlexNet, GoogleNet, OxfordNet gibi mimari çeşitleri içersinden OxfordNet (VGG16) olarak seçilmiştir.

2.2.1. Evrişimsel Sinir Ağları

Bu ağ yapısı 1988’de Fukushima tarafından ortaya atılmıştır [26]. Eğitim için gerekli donanımın yetersizliğinden dolayı yaygın şekilde kullanılamamıştır. 1990’larda LeCun ve arkadaşları, CNN’ler için gradyan-tabanlı bir öğrenme algoritması geliştirmiş ve el yazımı rakam sınıflandırması problemi [27] üzerinde başarılı sonuçlar elde etmişlerdir. Daha sonrasında CNN’ler geliştirilmiş ve birçok sınıflandırma işleminde başarılı sonuçlar vermişlerdir. CNN’lerde kullanılan max pooling (maksimum havuzlama) katmanı, şekil varyasyonlarını kavramada oldukça etkilidir. Ayrıca, bağlı ağırlıklarla seyrek bağlantılı olmaları, benzer bir boyuttaki tam bağlantılı ağ’a (fully connected network) göre oldukça az parametreleri olmasını sağlar. Gradyan bazlı öğrenme algoritmaları ile eğitildikleri için, azalan gradyan probleminden daha az etkilenirler. CNN’ler ağırlık oluşturmada çok etkilidirler ve yüksek derecede optimize edilmiş ağırlıklar üretebilirler. CNN yapısında iki adet ana bölüm vardır. Bunlar, öznitelik çıkartıcılar ve sınıflandırıcıdır. Öznitelik çıkartıcılar, evrişim (convolution) ve havuzlama (pooling) katmanlarından oluşur. Bu katmanların verdikleri çıktılara öznitelik haritası (feature map) denir. Evrişim katmanındaki her düğüm, girdi düğümleri üzerinde evrişim işlemi yaparak, girdi görüntüsünün özniteliklerini çıkartır. Üst seviye öznitelikler, alt seviye katmanlardan yayılan özniteliklerden elde edilir. Öznitelikler en yüksek katmana veya seviyeye yayıldıkça, özniteliklerin boyutları evrişim ve maksimum havuzlama operasyonları için çekirdek boyutuna bağlı olarak azaltılır. CNN’in son katmanının çıktısı, tam bağlı bir ağa girdi olarak verilir, buraya sınıflandırma katmanı denir. Daha iyi performansları olduğundan bu katmanda genellikle ileri beslemeli binir ağları (feed-forward neural network) kullanılır.

VGG16, nesne algılama için geliştirilmiş bir evrişimsel sinir ağıdır. Oxford’un Görsel Geometri Grubu (Visual Geometry Group (VGG)) tarafından geliştirilmiş ve eğitilmiştir. Bu model, akademik araştırmacılar için çok büyük bir görüntü veritabanı olan ImageNet kullanılarak 2014 yılında düzenlenen ImageNet Large Scale Visual Recognition Challenge (ILSVRC) yarışmasında çok büyük bir başarı elde etmiştir. Yarışmada nesne yer tayini (image localization) görevinde 1’inci, nesne sınıflandırma (image classification) görevinde 2’nci olmuştur. Yarışmada kullanılan görüntüler 1000 farklı kategoriye bölünmüş ve yarışan modellerin, görüntüleri bu kategorilere yüksek bir yüzde ile ayırması beklenmiştir. VGG16 modelinin akışı Şekil 2.4.’te gösterilmiştir.

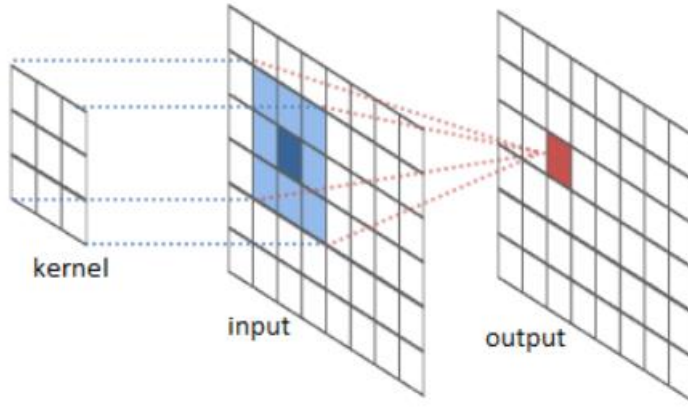


Şekil 2.4. VGG16 CNN modeli akışı

2.2.1.1. Evrişimsel katman (Convolution layer)

Bu katmanda, önceki katmanlardan gelen öznitelik haritaları, öğrenebilen çekirdeklerle evriştirilir. Alınan girdi görüntüsünden özniteliklerin çıkartılmasını sağlayan ilk katmandır. Evrişim, girdi verisini küçük karelere indirgeyerek, pikseller arasındaki bağlantıyı korur. $(h \times w \times d)$ boyutunda bir görüntü, $(fh \times fw \times d)$ boyutunda bir filtre ile evrişime sokularak, $(h - fh + 1) \times (w - fw + 1) \times 1$ büyüklüğünde bir çıktı oluşturur. Bu çekirdek çıktılar, doğrusal ya da doğrusal olmayan aktivasyon fonksiyonlarından geçerek yeni öznitelik haritalarını çıktı olarak verir.

Çeşitli filtrelerle evriştirilen görüntülerde, kenar bulma, keskinleştirme, bulandırma gibi işlemler uygulanabilir. Bu işlemler özniteliklerin çıkartılmasında fayda sağlar. Evrişim işleminin gösterimi Şekil 2.5.'te yapılmıştır.



Şekil 2.5. Evrişim işleminin gösterimi

2.2.1.2. **ReLU (Rectified linear unit)**

Bu fonksiyon doğrusal olmama durumunu sağlar. $f(x) = \max(0, x)$ şeklinde bir çıktı üretir. Evrişim sonunda oluşacak negatif piksellerin değerinin 0'a çekilmesini sağlar. ReLU'nun çalışma prensibini gösteren örnek bir görsel Şekil 2.6.'da verilmiştir.

17	20	-30	35
18	-11	69	93
20	-45	25	-1
107	81	18	23

➔

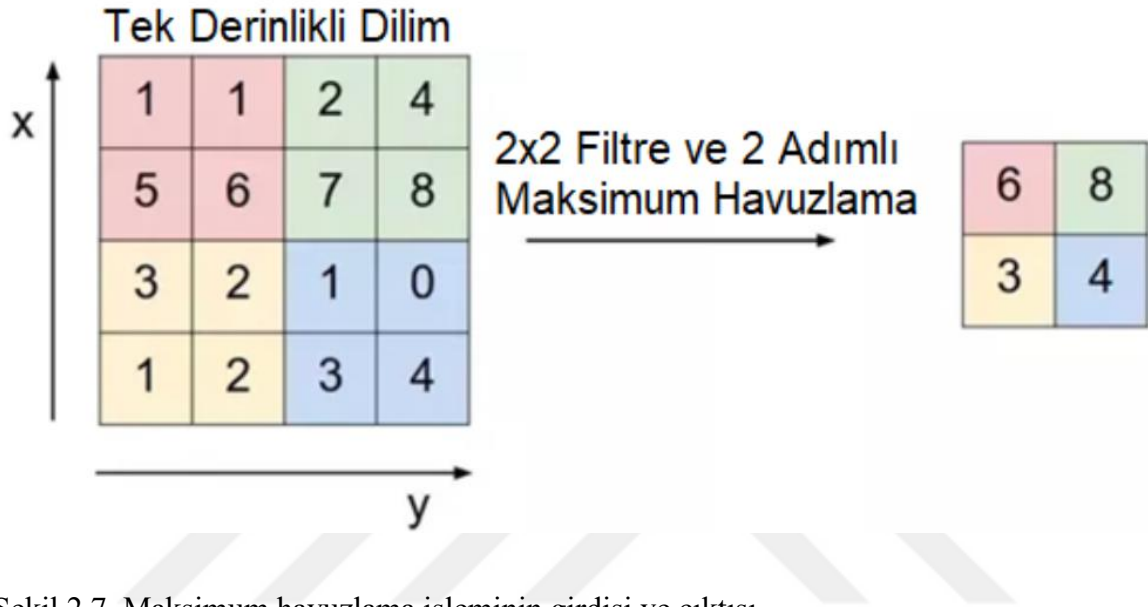
17	20	0	35
18	0	69	93
20	0	25	0
107	81	18	23

Şekil 2.6. ReLU işleminin girdisi ve çıktısı

2.2.1.3. **Havuzlama katmanı (Pooling layer)**

Bu katmanda, uzaysal havuzlama (spatial pooling) uygulanıp, alt örnekleme yapılarak görüntülerin boyutları azaltılır. Bu boyut azaltımı veri kaybı anlamına gelse de aslında ortaya çıkan alt örnekleme üzerinde öznitelikler hakkında varsayımlar yapılmasını sağlar. Aynı zamanda burada oluşan bilgi kaybı, sistemin öznitelik verilerini ezberlemesinin önüne geçmektedir. Boyut azaltımının bir diğer önemli avantajı da görüntüler daha küçük hale geldiği için ağırların hızının artmasını sağlamasıdır. Maksimum havuzlama (max pooling), ortalama havuzlama (average pooling) ve toplam havuzlama

(sum pooling) gibi çeşitleri vardır. VGG’de sadece maksimum havuzlama kullanılmıştır. Maksimum havuzlama, seçilen boyuta göre, görüntünün boyutunu azaltır. 2x2 filtre ile uygulanan bir maksimum havuzlama işlemi sırasında, 2x2’lik bölgeler seçilir ve o bölgedeki en büyük piksel değeri yeni değer için atanır. Maksimum havuzlama uygulanırken seçilen adım (stride) değeri, filtrenin ne kadar kaydırılacağını ifade eder. Şekil 2.7.’de maksimum havuzlama işleminin gösterimi verilmiştir.



Şekil 2.7. Maksimum havuzlama işleminin girdisi ve çıktısı.

2.2.1.4. Tam bağlı katman (Fully connected layer)

Önceki katmandan gelen çıkartılmış öz niteliklerden her sınıfa ait skoru hesaplayan tam bağlı katmandır. Son katmandan çıkartılan öz nitelik haritaları, skalar değerleri olan vektörler şeklinde ifade edilerek tam bağlı katmana gönderilir. Burada tam bağlı ağıdaki katman sayısı için katı kurallar olmasa da LeNet [27], AlexNet [28], ve VGG Net [9] gibi mimariler de dahil olmak üzere çoğu modelde 2 – 4 arası katman gözlemlenmektedir.

Evrişim ve havuzlama işlemlerinden sonra, çıktı bir sınıf olarak beklenir. Evrişim ve havuzlama katmanları, öz nitelik çıkarımını ve orijinal görüntülerin parametrelerinin azaltılmasında rol alırlar, ancak son çıktıyı alabilmek için bir tam bağlı katmana ihtiyaç duyulur. Evrişim katmanları 3 boyutlu aktivasyon haritaları (activation map) oluşturur ancak son çıktı olarak ihtiyaç duyulan veri, girdinin bir sınıfa ait olup olmadığıdır. Burada kullanılan tam bağlı katman, sınıflandırılan görüntünün sınıflara ait olma olasılıklarını hesaplar. Bu katmanda aktivasyon fonksiyonu olarak Softmax kullanılmıştır. Orijinal modelde, ILSVR için sınıflandırma yapılacağından 1000 çıktı sınıfı mevcuttur. Bu

çalışmada, softmax'ın çıktısı, çalışmada algılanması sağlanan 10 el hareketi sınıfına karşılık gelmesi için 10 olarak değiştirilmiştir. Softmax, tüm sınıflar için olasılık çıktılarını $[0,1]$ arasında normalize eder ve hepsinin toplamının 1 olmasını sağlar. Böylelikle her sınıfa ait olma olasılıkları elde edilmiş olur. Tam bağlı katman, tahmin hatalarını hesaplamak için kategorik çapraz entropi (categorical crossentropy) gibi bir kayıp fonksiyonuna (loss function) sahiptir. İleriye doğru akış bittikten sonra, hata ve kayıp değerlerinin azaltımına yönelik olarak ağırlık (weight) ve eğilim (bias) değerlerini güncellemek için geri yayılım (backpropagation) başlar.

Özet olarak CNN'ler öznitelik haritalarının oluşturulması için evrimsel katmana ihtiyaç duyarlar. Aktivasyon fonksiyonları, evrimsel katmanda oluşturulan öznitelik haritalarına doğrusal olmama durumunu sağlar ve bir nöronun ateşlenip ateşlenmemesi gerektiğine karar verir. VGG-16'da her bir evrimsel katmandan sonra ReLU aktivasyon fonksiyonu kullanılmıştır. Havuzlama katmanında, evrimden sonra elde edilen aktivasyon haritalarına alt örnekleme yapılarak öznitelikler hakkında çıkarım yapmak amaçlanır. Bu alt örnekleme işlemi verinin boyutunu azalttığından hem ağırlık performansını artırır hem de ağırlık verileri ezberlemesinin önüne geçer. VGG-16 mimarisinde havuzlama işlemi olarak maksimum havuzlama kullanılmıştır. Son olarak, çıktı katmanı olarak, tam bağlı bir katman yer alır. Tam bağlı katmanın amacı önceki katmanlardan elde edilen aktivasyon haritalarını kullanarak, sınıflandırılmak istenen girdinin tüm sınıflara ait olma olasılığını hesaplamaktır. VGG-16 mimarisinde tam bağlı katmanda aktivasyon fonksiyonu olarak Softmax kullanılmıştır. Softmax bir girdinin tüm sınıflara ait olma olasılığını hesaplayarak, 0 ile 1 arasında bir olasılık çıktısı sunar.

2.3. Geometrik Grafik İşlemleri

2.3.1. Bilgisayar grafikleri

Bilgisayar grafikleri, bir bilgisayar üzerinde, ilgili bir yazılımla, görüntülerin oluşturulması, gösterilmesi ve saklanması içerir. Bilgisayarlarla tamamen dijital görüntüler üretilebilir, ya da gerçek dünyadan elde edilen görüntüler üzerinde değişiklikler yapılabilir. Bilgisayarlar 2 tip grafik verisi saklama kapasitesine sahiptirler. Bunlar vektör grafikler ve raster (hücreli) grafiklerdir. Vektör grafikler, matematiksel formüllere dayanır ve doğruların birleştirilmesi ile oluşan geometrik şekiller oluştururlar. Hücreli grafikler ise, piksel denilen noktalardan oluşan bir ızgara (grid) yapısı üzerinde bireysel

piksellerin aydınlatılmasıyla çizilen grafikleri ifade eder. Bilgisayar grafikleri 2 boyutlu, 3 boyutlu ve animasyonlu (canlandırılmış) olmak üzere 3'e ayrılırlar.

Bilgisayar grafiklerinin, günlük hayatta çok sıklıkla karşılaşılan kavramlarda önemli bir yeri vardır. 3 boyutlu grafik ve animasyonlar, özellikle bilgisayar oyunlarının geliştirilmesinde çok önemli bir rol oynamaktadır. Bilgisayar oyunlarının devrimsel bir gelişim göstermesinin en büyük sebeplerinden bir tanesi grafik teknolojisinin ilerlemiş olması ve artık oyunlarda görüntünün çok gerçekçi bir halde sunulabiliyor olmasıdır. Aynı şekilde sinema ve film alanında çok sıkça kullanılan grafik ve animasyon teknolojileri, sinema endüstrisini de bambaşka bir seviyeye taşımıştır. Bunlar haricinde, bilgisayar grafiklerinin eğitim ve araştırma alanında da birçok kullanımı vardır. İnsan vücudunun ve dokularının sanal bir benzetimi oluşturularak, tıp alanında eğitim, klinik tanılarının konulması, operasyon benzetimleri gibi alanlarda kullanılabildiği gibi, eğitim alanında da, eğitimcilerin çeşitli kavramları ve prensipleri, 2 boyutlu görüntülere göre daha iyi açıklayabilecek ve öğrencilerin kavramasını daha rahat bir hale getirecek şekilde 3 boyutlu modellemeler yapmalarına imkan tanımaktadır. Mimari ve inşaat alanında, modellemeler yapılarak, sanal binalar ve evler üzerinde çeşitli denemeler yapılabilir ve inşaatın yapılmasından önce gerekli kararlar verilebilir.

Bilgisayar grafiklerinin tarihçesinde daha önce yapılan çalışmalar olsa da çok önemli bir kilometre taşı ve dönüm noktası olarak, Ivan Sutherland tarafından 1961 yılında doktora tezi olarak geliştirilen Sketchpad [29] adlı uygulama gösterilebilir. Bu çalışma insan bilgisayar etkileşimi konusunda çok özgün ve çığır açıcı bir uygulama olup grafik arayüzü kullanan ilk uygulamalardan birisidir. Çalışma, vektör grafiklerle çalışmakta olup monitor üzerindeki öğelerle etkileşim, fotosel barındıran bir ışık kalemı yardımıyla sağlanmaktaydı. Bu çalışma ile aynı zamanlarda General Motors'da yine tarihsel olarak çığır açıcı başka bir uygulama da temellenmekteydi. Bu uygulamanın adı DAC1 (Design Augmented by Computer) olup, bilgisayar destekli tasarım konusunda büyük bir ürüne dönüştü ve GM arabaları ve kamyonlarının tasarımında önemli bir öge haline geldi. Bu çalışmalardan sonra vektör grafikleri üzerinde çalışmalar devam etti. Bilgisayar grafikleri konusunun bu gelişimi üzerine büyük şirketler de konu ile ilgilenmeye başladı. IBM bu ilerleyen teknolojiye hızlı bir şekilde cevap vererek, IBM 2250 grafik terminali ismiyle, ilk ticari grafikli bilgisayarı piyasa sürdü. 1966'da yine Ivan Sutherland, ilk bilgisayarla kontrol edilen başa takılan göstergeyi geliştirdi. Bu cihaz her bir göz için birer adet wireframe (iskelet) görüntü oluştuyordu, böylelikle kullanıcı bilgisayardaki görüntüyü stereoskopik şekilde 3 boyutlu bir şekilde görebiliyordu. 1970'lerde doku eşleştirme

(texture mapping), saklı yüzey tayini (hidden surface determination) gibi kavramlar ortaya çıkartıldı. Aynı zamanda atari salonu konsepti ve atari salonu oyunları da bu dönemde ortaya çıkmaya başladı. 1970 yılında atari konsepti ortaya atılırken, bu konsept için geliştirilen oyunlarda 2 boyutlu hareketli (sprite) grafikler kullanıldı. 1972’de Pong, 1974’te Speed Race, 1978’de Space Invaders ortaya çıktı. 1980’lerde, bilgisayarların evlerde yaygınlaşmasıyla, daha öncesinde sadece akademik olarak görülen bilgisayar grafikleri konsepti, daha geniş bir çevre için erişilebilir hale geldi ve grafik programlama üzerinde çalışan insan sayısı oldukça arttı. 1982’de Japonya’daki Osaka Üniversitesi, LINKS-1 Computer Graphics System’ı geliştirdi. Bu sistemin geliştirilmesinin amacı gerçekçi 3 boyutlu grafikler üretebilmektir. LINKS-1, hızlı görüntü çizme için geliştirilen yeni bir metodoloji ile, hızlı bir şekilde yüksek derecede gerçekçi çizimler yapabiliyordu. LINKS-1, tamamen bilgisayar grafikleriyle oluşturulmuş, dünyanın ilk 3 boyutlu gökevi benzeri videosu ile evrenin bilinen kısmını çizmeyi başardı. Film ve müzik konusunda da CGI (Computer Generated Imagery) kullanımı çok ilerledi. Star Wars ve diğer bilim kurgu filmlerinin popülaritesi bilgisayar grafikleri kullanımıyla beraber daha da arttı. Başta Dire Straits’in “Money For Nothing” video klibi olmak üzere, CGI kullanılan klipler de ilgi görmeye başladı. 1988’de Pixar tarafından ilk gölgelendirici (shader) üretildi. 1980’ler video oyunlarının altın çağı olarak da adlandırıldı ve Atari, Nintendo ve Sega gibi şirketler, bilgisayar grafiklerini ürünlerinde kullanmaya başlayarak genç ve kolay etkilenebilir bir kitleye ürünlerini sundular. 1990’larda, eskiden sadece iş istasyonlarında (workstation) yapılan görüntü alma (rendering) işlemleri, ev bilgisayarlarında da yapılabilir hale geldiğinden CGI kalitesinde ciddi bir artış meydana geldi ve 3 boyutlu modelleme kavramının temelleri atılmış oldu. 3D Studio gibi evde görüntü almayı sağlayan yazılımların önemi ve bunlara olan rağbet arttı. 1999 yılında Nvidia tarafından GeForce256 olarak adlandırılan, GPU olarak bahsedilen ilk video kartı geliştirildi. 1990’ların sonunda, bilgisayarlar grafik işleme için kullanılan uygulama çatıları olan DirectX ve OpenGL’i kullanmaya başladı. Pixar, 1995 yılında ilk büyük filmi olan Toy Story’i ortaya çıkarttı. Wolfenstein 3D, Doom ve Quake gibi 3 boyutlu bilgisayar oyunları ortaya çıkmaya başladı. Sony Playstation, Sega Saturn ve Nintendo 64 gibi oyun konsolları ortaya çıktı ve yüksek bir popülariteye ulaştı. Bu cihazlar milyonlarca satıldı ve evde oyun oynama konsepti 3 boyutlu grafikleri popülerleştirdi. 2000’li yıllardan sonra bilgisayar grafikleri konusu daha da gelişirken, artık CAD (Computer Aided Design), CAM (Computer Aided Manufacturing) konseptleri birçok alanda yaygın bir şekilde kullanılmaya başlandı. Üretim, benzetim, inşaat, makine tasarımları, mimari gibi alanlarda sıkça kullanılır hale geldi [30].

Bilgisayar grafiklerinin ilerleyişi, bilgisayar kullanılarak, bir tasarımın oluşturulması, modifiye ya da optimize edilmesi olarak tanımlanabilecek CAD konseptinin de ilerlemesine ve üretim, mühendislik, mimari gibi alanlarda sıkça kullanılır hale gelmesine imkan tanımıştır.

Bilimsel hesaplamaların görselleştirilmesinde de bilgisayar grafikleri ciddi şekilde kullanılmaya başlamıştır. Sonlu eleman analizleri, akışkanlar dinamiği, hava durumu tahminleri, tıp gibi alanlarda bilgisayar grafiklerinin kullanımı artmıştır.

Ayrıca AR ve VR konseptlerinin son zamanlardaki gelişimi ve yayılımı ile, sınırlama tanımadan nerdeyse her alanda AR ve VR uygulamalarının artışı ile bilgisayar grafiklerinin önemi ve kullanımı büyük bir yükseliş göstermiştir.

2.3.2. Önerilen sistemin barındırdığı grafik işlemleri

Önerilen sistemde, el hareketlerinin tanımlanması ile, tanınan harekete karşılık gelen grafik işlemi yerine getirilmektedir. El hareketlerine karşılık gelen grafik işlemleri,

- Sonraki şekli çizme,
- Önceki şekli çizme,
- Mevcut şekle 2 boyut – 3 boyut dönüşümü uygulama
- Mevcut şekli büyütme,
- Mevcut şekli küçültme,
- Mevcut şekli X ekseninde döndürme,
- Mevcut şekli Y ekseninde döndürme,
- Mevcut şekil için çevre, alan ve hacim formüllerini gösterme

olarak sıralanabilir.

Doğrudan grafik işlemi yapmayan ancak grafik işlemlerinin yapılmasını etkileyen el hareketleri ise;

- Aktive etme
- Rotasyon yönünü değiştirme

el hareketleridir.

Söz konusu el hareketleri ve bunlara karşılık gelen grafik işlemleri, devam eden bölümlerde detaylı bir şekilde açıklanmıştır.

2.3.2.1. Sonraki şekli çizme

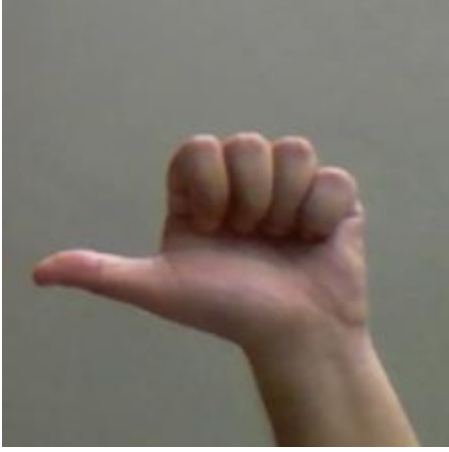
Önerilen sistemde çizilebilen 8 adet şekil olup, bunlar sırasıyla; kare, daire, dikdörtgen, üçgen, küp, küre, dikdörtgen prizma ve üçgen prizmadır. Bu el hareketine karşılık gelen grafik işlemi, yukarıdaki sıralamaya göre o anda ekranda bulunan şekilden bir sonraki şekli çizmektir. Şekillerin çizimi o anki boyuta bağlı olup şekillerin ekranın merkezinde kalması için, daire ve küre şekilleri hariç, boyuta göre köşelerin hesaplamaları yapılır ve bu noktalar birleştirilir. Daire ve küre için ise ekranın ortası merkez baz alınıp, boyuta göre yarıçap değeri verilerek çizdirilir. İlgili el hareketi Şekil 2.8.'de gösterilmiştir.



Şekil 2.8. Sonraki şekli çizme aksiyonu için el hareketi

2.3.2.2. Önceki şekli çizme

Bu el hareketine karşılık gelen grafik işlemi, bir önceki maddede belirtilen sıralamaya göre o anda ekranda bulunan şekilden bir önceki şekli çizmektir. Şekillerin çizilimi o anki boyuta bağlıdır, şekillerin ekranın merkezinde kalması için, daire ve küre şekilleri hariç, boyuta göre köşelerin hesaplamaları yapılır ve bu köşe noktaları birleştirilir. Daire ve küre için ise ekranın ortası merkez baz alınıp, boyuta göre yarıçap değeri verilerek çizdirilir. İlgili el hareketi Şekil 2.9.'da gösterilmiştir.



Şekil 2.9. Önceki şekli çizme aksiyonu için el hareketi

2.3.2.3. Mevcut şekle 2 boyut – 3 boyut dönüşümü uygulama

Bu el hareketine karşılık gelen grafik işlemi, o anda ekranda bulunan şekil 2 boyutlu ise, o şeklin 3 boyutlu karşılığını; eğer o anda ekranda bulunan şekil 3 boyutlu ise de, o şeklin 2 boyutlu karşılığını çizmektir. Sistemdeki şekillerin 2 boyut – 3 boyut karşılıkları Tablo 2.1’de gösterilmiştir. Bu işlem yapılırken o anki şeklin ebatları korunur. İlgili el hareketi Şekil 2.10.’da gösterilmiştir.



Şekil 2.10. Mevcut Şekle 2B-3B dönüşümü uygulama aksiyonu için el hareketi

Tablo 2.1. Sistemdeki şekillerin 2 boyut - 3 boyut dönüşümü için karşılıkları

2 Boyut	3 Boyut
Kare	Küp
Daire	Küre
Dikdögen	Dikdörtgen Prizma
Üçgen	Üçgen Prizma

2.3.2.4. Mevcut şekli büyütme

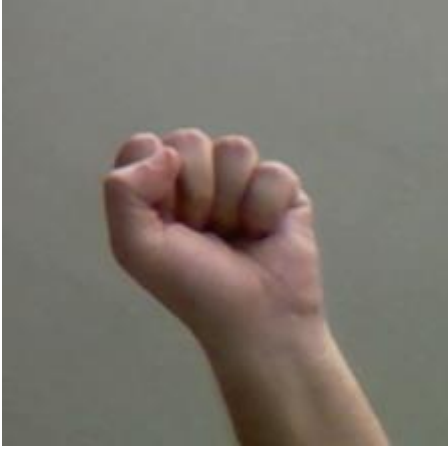
Bu el hareketine karşılık gelen grafik işlemi, ekranda o anda bulunan şekli büyütme işlemidir. Büyütme işlemi x birimden $2x$ birime kadar, $0.2x$ aralıklarla yapılacak şekilde, 6 farklı büyüklüğü desteklemektedir. Eğer şeklin büyüklüğü maksimum seviyede ise büyütme işlemi yapılmaz. İlgili el hareketi Şekil 2.11.'de gösterilmiştir.



Şekil 2.11. Mevcut şekli büyütme aksiyonu için el hareketi

2.3.2.5. Mevcut şekli küçültme

Bu el hareketine karşılık gelen grafik işlemi, ekranda o anda bulunan şekli küçültme işlemidir. Küçültme işlemi $2x$ birimden x birime kadar, $0.2x$ aralıklarla yapılacak şekilde, 6 farklı büyüklüğü desteklemektedir. Eğer şeklin büyüklüğü minimum seviyede ise küçültme işlemi yapılmaz. İlgili el hareketi Şekil 2.12.'de gösterilmiştir.



Şekil 2.12. Mevcut şekli küçültme aksiyonu için el hareketi

2.3.2.6. Mevcut şekli X eksen etrafında döndürme

Bu el hareketine karşılık gelen grafik işlemi, ekranda o anda bulunan şekli, X eksen etrafında, “Rotasyon Yönünü Değiştirme” el hareketine göre alınan değere göre pozitif ya da negatif yönde döndürmektir. Döndürme işlemi herhangi bir kısıtlama olmadan yapılabilmektedir. Şekil eksen etrafında sonsuza kadar dönebilir. İlgili el hareketi Şekil 2.13.’te gösterilmiştir.



Şekil 2.13. Mevcut şekli X eksen etrafında döndürme aksiyonu için el hareketi

2.3.2.7. Mevcut şekli Y eksen etrafında döndürme

Bu el hareketine karşılık gelen grafik işlemi, ekranda o anda bulunan şekli, Y eksen etrafında, “Rotasyon Yönünü Değiştirme” el hareketine göre alınan değere göre pozitif ya

da negatif yönde döndürmektir. Döndürme işlemi herhangi bir kısıtlama olmadan yapılabilir. Şekil eksen etrafında sonsuza kadar dönebilir. İlgili el hareketi Şekil 2.14.'te gösterilmiştir.



Şekil 2.14. Mevcut şekli Y eksen etrafında döndürme aksiyonu için el hareketi

2.3.2.8. Mevcut şekil için çevre, alan ve hacim formüllerini gösterme

Bu el hareketine karşılık gelen grafik işlemi, ekranda o anda bulunan şekil için çevre, alan ve hacim hesaplama formüllerini göstermektir. 2 boyutlu şekiller için sadece çevre ve alan formülleri gösterilirken, 3 boyutlu şekiller için bu formüllerin yanısıra hacim formülleri de kullanıcıya sunulmaktadır. Formüllerin gösterim şekli ana pencerenin yanında bir yan pencere açarak, bu pencerede yazı bazlı, ya da şeklin kenarları, yüksekliği, yarıçapı gibi değerlerin gösterildiği sembolik bir şekil üzerinde gösterilebilir. İlgili el hareketi Şekil 2.15.'te gösterilmiştir.



Şekil 2.15. Mevcut şekil için çevre, alan ve hacim formüllerini gösterme aksiyonu için el hareketi

2.3.2.9. Aktive etme

Bu el hareketinin doğrudan grafik çizimi işleminde bir etkisi olmasa da, grafik çizimi işlemlerini yapacak diğer hareketlere önemli bir etkisi bulunmaktadır. Ağ kamerasından sürekli görüntü okunduğu için, herhangi bir hareket algılanmaması istenilen durumlarda veya hareketler arası geçişlerde istenilenden farklı bir hareketin algılanmaması için sistem, aktif olmadığı durumlarda sınıflandırma yapmayacak şekilde tasarlanmıştır. Bu el hareketi sistemi aktif duruma geçirmek için kullanılır. Bu hareket algılanmadan başka bir hareket sınıflandırılmaz. Sistem bu hareketi algıladıktan sonra dinlemeye geçer ve diğer hareketlerden birisini algıladığında o hareket için gerekli işlemi yapıp kendini tekrar pasif hale getirir. Her hareketten sonra bu el hareketiyle uygulamayı tekrar aktive etmek gerekmektedir. İlgili el hareketi Şekil 2.16.'da gösterilmiştir.



Şekil 2.16. Aktive etme aksiyonu için el hareketi

2.3.2.10. Rotasyon yönünü değiştirme

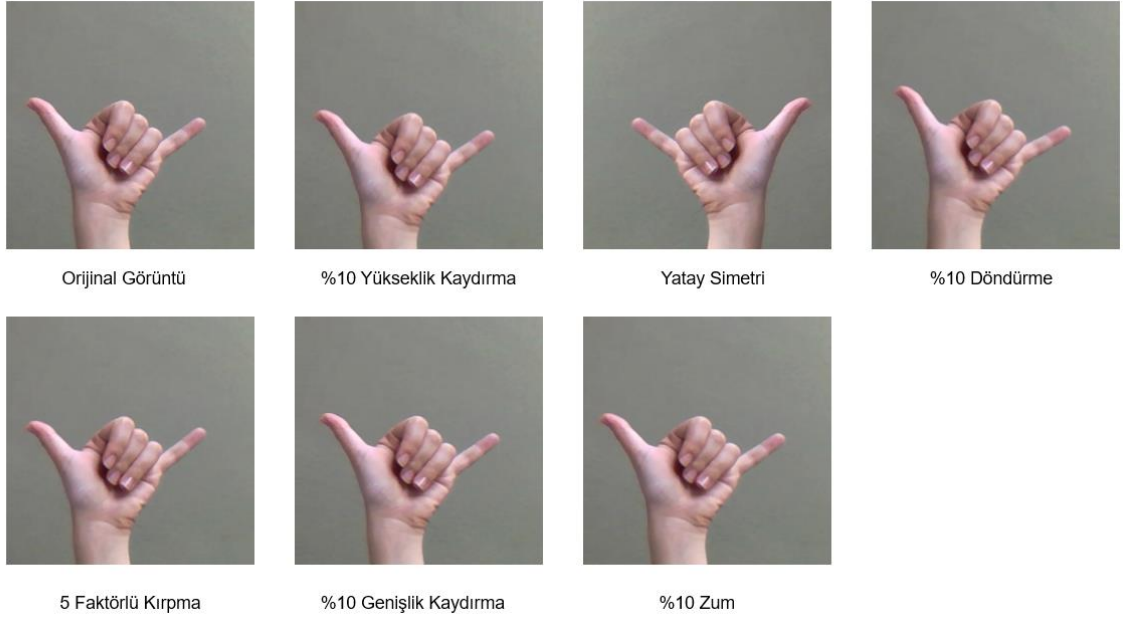
Bu el hareketinin doğrudan grafik çizimi işleminde bir etkisi olmasa da, grafik çizimi işlemlerini yapacak diğer hareketlere etki etmektedir. X ve Y eksenleri etrafında döndürme fonksiyonlarına doğrudan etki eden bu hareket varsayılan olarak pozitif başlayan bir değeri, her algılandığı durumda terse çevirir. Bu hareket en son tanındığında değer pozitif ise, döndürme işlemleri pozitif yönde, değer negatif ise, döndürme işlemleri negatif yönde yapılır. İlgili el hareketi Şekil 2.17.'de gösterilmiştir.



Şekil 2.17. Rotasyon yönünü değiştirme aksiyonu için el hareketi

3. VERİ KÜMESİ

Çalışmada kullanılan veri kümesinin oluşturulmasında, yaşları 20 ila 30 arasında değişen 2 kadın ve 4 erkek katılımcının el görüntüleri kullanılmıştır. Veri kümesinin oluşturulması için Python dilinde bir uygulama geliştirilmiştir. Bu uygulama ağ kamerasından bir akış başlatmakta ve kullanıcının komutuyla o anki kareyi kaydetmektedir. Bu şekilde veri kümesinde yer alan kullanıcılar bu uygulamayı kullanarak; (i) ellerini sağa, sola, yukarı ve aşağı kaydırarak, (ii) ellerini bilekten öne, geriye, yukarı ve aşağı doğru bükerek, (iii) ellerini kameraya yaklaştırıp, uzaklaştırarak her el hareketinin olabildiğince fazla şekilde görüntüsünü kaydetme imkânı bulmuşlardır. Aynı uygulama sayesinde her katılımcının, her el hareketi için 50'şer adet görüntü çekilmiş ve her görüntü saat yönünde ve ters saat yönünde 10° döndürme aralığı, %10 genişlik ve yükseklik kaydırma aralığı, %10 faktörlü yakınlaştırma ve uzaklaştırmayla zum işlemi, 5 faktörlü kırpma işlemi ve yatay simetri parametrelerinin rastgele seçilmesi ile 10 katına çıkartılmıştır. Orijinal görüntülerin bu parametrelerle çoğaltılmış şekillerinden bazı örnekler Şekil 3.1.'de gösterilmiştir. Bu çoğaltma işleminden sonra her el hareketi için kullanıcı başına 500 adet görüntü elde edilmiştir. Elde edilen bu 500 görüntüden rastgele seçilen 200 tanesi eğitim, 20 tanesi doğrulama (validation) için ayrılmıştır. Modelin eğitim sürecine yönelik her el hareketi için her katılımcıdan alınan 200 görüntünün toplamı olan 1200 (6×200) adet görüntü eğitim için, her katılımcıdan alınan 20 görüntünün toplamı olan 120 (6×20) görüntü ise doğrulama amaçlı kullanılmıştır. Kalan görüntülerden, her el hareketi için rastgele 1000'er adet seçilmiş ve bu görüntüler test süreci için kullanılmıştır.



Şekil 3.1. Eğitim sürecine yönelik veri çoğaltma işleminde kullanılan parametrelerin sonuçlarına yönelik birer örnek

4. DENEYSEL SONUÇLAR

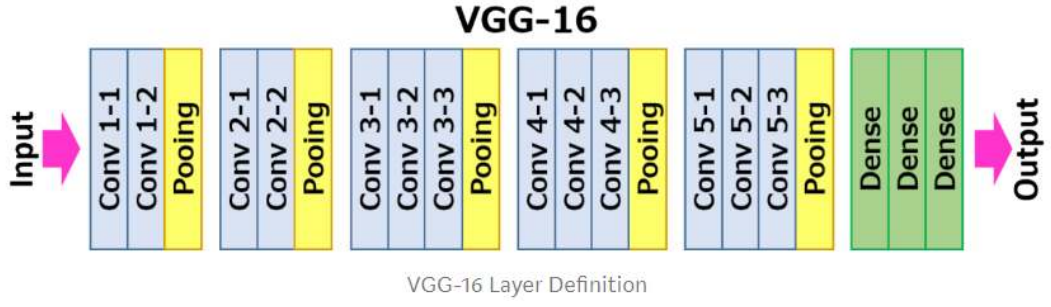
4.1. Geliştirme ve Test Ortamı Konfigürasyonları

Eğitim çağında çocuklar için el hareketi tanıma tabanlı geometri öğrenme uygulaması, Windows 10 işletim sistemi üzerinde çalışan, NVidia GeForce 750M ekran kartlı, 16GB rastgele erişimli bellekli (RAM) ve 256 GB katı hal sürücülü (SSD) donanımları kullanan bir dizüstü bilgisayarda geliştirilmiş ve test edilmiştir.

Uygulama Python 3.6.3 versiyonu kullanılarak geliştirilmiştir. Bilgisayarla görme işlemleri için OpenCV kütüphanesinin 3.4.0 versiyonu kullanılmıştır. OpenCV seçiminin yapılması, BSD (Berkeley Software Distribution) lisansı altında tamamen ücretsiz olması ve tüm büyük işletim sistemlerini desteklemesinin yanı sıra Python ile olan uyumu ve geniş bir topluluk tarafından kullanılır olması gibi nedenlere dayanmaktadır. El hareketi tanımlama sürecinin sonunda yer alan el hareketlerinin sınıflandırılması için kullanılan CNN modelinin gerçekleştirilmesi için Keras kütüphanesinin 2.1.4 versiyonu kullanılmıştır. Diğer derin öğrenme çatılarının yanında (Tensorflow, Theano, Caffe, Torch) Keras'ın seçilmesinin sebebi Keras'ın Tensorflow, CNTK, Theano gibi düşük seviye derin öğrenme dillerini desteklemesi ve tek bir arka uca bağımlı kalmaması, kullanıcı kolaylığı sağlaması, mobil, web, bulut gibi birçok platform üzerinde çalışabilme desteği ve GPU üzerinde çalışabilmesi gibi nedenlerdir. Çalışmadaki grafik işlemlerinin gerçekleştirilmesinde bir Python kütüphanesi olan Pyglet'in 1.3.1 versiyonu kullanılmıştır. Pyglet seçiminin yapılmasında temel unsurlar Windows, OS X ve Linux gibi işletim sistemlerinin hepsinde çalışabilir olması, görüntü ve video formatlarının çok büyük bir bölümünü destekliyor olması ve başka dış bağımlılıkları olmamasıdır.

4.2. Değerlendirme Sonuçları ve Tartışma

Bu çalışma kapsamında geliştirilmiş olan sistemde el hareketi algılama tabanlı çalışan bir geometri öğrenme uygulaması geliştirilmiştir. El hareketlerinin algılanması ve sınıflandırılması işleminde derin öğrenme çeşitlerinden CNN kullanılmıştır. CNN modeli olarak Oxford'un VGG16 mimarisi seçilmiştir. VGG16 modelinin görseli Şekil 4.1.'de gösterilmiştir.



Şekil 4.1. VGG-16 mimarisi

Modelin eğitimi GPU üzerinde gerçekleştirilmiştir. 50 devirde (epoch) eğitilen modelde devir süreleri en az 110 saniye, en fazla 118 saniye sürmüştür, 50 devrin ortalaması 114 saniye olarak hesaplanmıştır. Toplam eğitim işlemi 95 dakika sürmüştür. Eğitim işleminde kayıp fonksiyonu olarak kategorik çapraz entropi kullanılmıştır. Bu fonksiyon çoklu sınıflandırma işlemlerinde kullanılabilir. Çoklu sınıflandırma işlemlerinde sınıflar kategorik veri şeklindedir. Bilinen bir örneğin bir sınıfa ait olma olasılığı 1, geriye kalan sınıflara ait olma olasılığı 0'dır. Çapraz entropi, gerçek sınıf ve tahmin edilen sınıf arasındaki ortalama farkı hesaplar. Kategorik veri gereksiniminden ötürü çıktı katmanının her sınıf için bir tane olmak üzere n düğümden oluşması ve aktivasyon fonksiyonu olarak Softmax kullanmasını bekler. Eğitim işleminde çapraz entropi fonksiyonunun minimize edilmesi beklenir ve kusursuz değer 0'dır. 50 devirde yapılan eğitim işleminin eğitim ve doğrulama işlemleri için başarı oranları ve kayıp değerleri Tablo 4.1.'de gösterilmiştir. Eğitim işleminde 45'inci devirden sonra başarı oranlarının hemen hemen sabitlendiği ve çok küçük değişimler gösterdiği gözlemlenmiştir.

Tablo 4.1. CNN modelinin eğitim ve doğrulama başarı oranları

Devir (Epoch)	Kayıp (Loss)	Başarı (Accuracy)	Doğrulama Kayıp (Validation Loss)	Doğrulama Başarı (Validation Accuracy)
1	1.9570	0.3383	1.2924	0.6167
2	1.0110	0.7300	0.7527	0.8583
10	0.2760	0.9383	0.2362	0.9583
20	0.1520	0.9683	0.1468	0.9583
30	0.0960	0.9817	0.1115	0.9750
40	0.0810	0.9875	0.0597	1.0000
49	0.0550	0.9925	0.0561	0.9917
50	0.0560	0.9900	0.0546	1.0000

Klasik makine öğrenme temelli sınıflandırıcılar yüksek başarı yüzdesine erişebilmek için, eğitimi sırasında, oldukça maliyetli bir işlem olan öznitelik vektör çıkarımı yapılmasına ihtiyaç duyar. Derin öğrenme kullanan sınıflandırıcılarda ise modelin içerisindeki katmanlarda gerekli öznitelikler çıkartılır. Bu yüzden klasik makine öğrenme tabanlı sınıflandırıcılar yüksek yüzdelere erişebilmek için ham veri üzerinde bir takım ön işlemler (preprocessing) yapma gereği duyarlar [31]. Bu çalışmadaki CNN modelinin eğitiminde de ham veriler kullanılmış olup oldukça yüksek seviyelerde başarı oranı elde edilmiştir.

Derin öğrenme kullanılan bu çalışmada, modelin eğitim ve test süreçleri hem CPU hem de GPU üzerinde denenmiştir. GPU üzerinde daha hızlı sonuç alınmasına rağmen geliştirme ve testlerin yapıldığı cihaz üzerindeki ekran kartının çok yüksek performanslı bir ekran kartı olmamasından dolayı çok büyük bir fark gözlenmemiştir.

Çalışma sonucunda el hareketi algılama ile kontrol edilebilen bir geometri öğrenme uygulaması geliştirilmiş ve bu uygulamada kullanıcının; şekil çizme, şekilleri büyütme, küçültme, döndürme, şekilleri 2 boyuttan 3 boyuta ve 3 boyuttan 2 boyuta çevirme ve şekiller için çevre, alan ve hacim formüllerini gösterme gibi grafiksel işlemler yapması sağlanmıştır. Sistemin çalışması gerçek zamanlıya çok yakın olmakla beraber 0.2-0.3 saniye civarında küçük bir gecikme göstermektedir. Burada da yine sistemin test edildiği cihazda orta seviyeli bir ekran kartı bulunmasından dolayı oluşan bu gecikme, daha üst seviye bir ekran kartı kullanan bir sistemde çalıştırıldığında bu gecikmenin ortadan kalkması beklenmektedir.

Sistemin test edilmesi, her el hareketi için 1000'er adet görüntü kullanılarak yapılmıştır. Görüntüler sistemin çalışmasındaki gibi akış halinde değil, statik şekilde girdi

olarak verilmiş ve sistemden bir sınıflandırma çıktısı alınmıştır. Sisteme 1000'er adet görüntü verilip doğrudan sınıflandırma sonuçları alındığında sistemin bireysel olarak sırasıyla Fist (Yumruk), Four (Dört), Horn (Boynuz), L, One (Bir), Palm (Avuç içi), Shaka, Three (Üç), ThumbLeft (Baş parmak sola), Two (İki) el hareketleri için sırasıyla %100, %96.5, %85.1, %88.3, %86.6, %96.4, %62.9, %87.4, %71, %92.2 olmak üzere toplam %86.6 başarı yüzdesi ürettiği gözlemlenmiştir. Sistemin çalışma prensibi içerisinde bir el hareketini sınıflandırırken, benzerlik yüzdesinin %70'lik bir eşik değerini geçmesi beklenmektedir. Eğer sınıflandırma yapıldığında minimum benzerliği %70'in üzerine çıkan hiçbir sonuç yoksa, sistem grafik aksiyonu almaz. Bu eşik değeri seçilirken %50 - %80 değerleri arasında %5'lik adım sayısı kullanılarak yapılan testler sonucu en yüksek başarı oranının %70 değerinde elde edildiği görülmüştür. Bu nedenle optimal eşik değeri olarak %70 değerinin kullanılmasına karar verilmiştir. Bu eşik değeri dikkate alınarak yapılan testlerde yukarıda belirtilen sırayla %100, %98.1, %93.5, %94.5, %90.5, %98.3, %77.2, %93.7, %73.7, %97.5 olmak üzere, toplamda %92.6 oranında bir başarı oranı gözlemlenmiştir. Sabit görüntülerle yapılan eşik uygulanmamış ve eşik uygulanmış şekilde çıktı veren testlerin hata matrisleri Tablo 4.2. ve Tablo 4.3.'te gösterilmiştir. Tablolarda el hareketleri Fist (Yumruk), Four (Dört), Horn (Boynuz), L, One (Bir), Palm (Avuç içi), Shaka, Three (Üç), ThumbLeft (Baş parmak sola), Two (İki) sırasıyla 0-9 arası rakamlarla gösterilmiştir.

Tablo 4.2. Sistemin eşik değeri uygulanmadan sınıflandırma yaptığında başarı yüzdeleri (Yatay Eksen: Gerçek El Hareketleri, Dikey Eksen: Sınıflandırma Çıktıları)

	0	1	2	3	4	5	6	7	8	9
0	1000	8	3	1	87	0	13	4	281	8
1	0	965	7	0	13	36	48	101	0	20
2	0	0	851	2	0	0	11	0	2	0
3	0	0	5	883	0	0	14	0	0	0
4	0	0	12	7	866	0	13	0	3	1
5	0	19	2	10	1	964	23	4	0	1
6	0	0	5	6	0	0	629	0	4	0
7	0	8	22	0	0	0	6	874	0	44
8	0	0	2	84	1	0	207	0	710	4
9	0	0	91	7	32	0	36	17	0	922

Tablo 4.3. Sistemin eşik değeri uygulanarak sınıflandırma yaptığında başarı yüzdeleri (Yatay Eksen: Gerçek El Hareketleri, Dikey Eksen: Sınıflandırma Çıktıları)

	0	1	2	3	4	5	6	7	8	9
0	999	5	1	1	68	0	1	1	214	2
1	0	923	0	0	2	16	6	46	0	2
2	0	0	685	0	0	0	4	0	0	0
3	0	0	1	785	0	0	3	0	0	0
4	0	0	1	0	784	0	3	0	1	1
5	0	9	1	0	1	927	3	0	0	0
6	0	0	0	5	0	0	414	0	0	0
7	0	3	6	0	0	0	1	767	0	15
8	0	0	0	38	0	0	96	0	604	1
9	0	0	38	2	11	0	5	5	0	809
Eşik Altı	1	60	267	169	134	57	464	181	181	170
Eşik Üstü	999	940	733	831	866	943	536	819	819	830
Sınıflandırılma Yüzdesi	0.999	0.940	0.733	0.831	0.866	0.943	0.536	0.819	0.819	0.830

Sınıflandırma sonuçları incelendiğinde “boynuz” ve “shaka” el hareketlerinin belirlenen eşik değeri geçme konusunda başarısız oldukları gözlemlenmiştir. Bu durumun, ilgili hareketlerin karmaşıklığı ve birbirlerine benzerliklerinden kaynaklanmakta olduğu düşünülmektedir. Özellikle shaka hareketinin %70 eşiğini geçerek sınıflandırılma oranının oldukça düşük olduğu gözlemlenmektedir. Örneğin; bu hareketin eşik değeri

uygulanmadan sınıflandırılması %62.9 başarı oranına sahip iken eşik uygulandığında bu yüzde %77.2'ye yükselmektedir. Bu sonuçlar da bu hareketin sınıflandırılması esnasında çıkan benzerlik oranlarının birbirine çok yakın olduğunu göstermektedir. Bu iki hareket haricinde düşük sınıflandırılma oranı veren iki hareket de “üç” ve “başparmak sola” hareketleridir. Özellikle “başparmak sola” hareketinin eşik uygulanmış halde başarı yüzde artışı çok düşük bulunmuştur, “başparmak sola” el hareketinin, “yumruk” hareketi olarak yanlış sınıflandırılma oranının da çok yüksek olduğu görülmektedir. Bunun sebebi “yumruk” hareketi yapılırken de başparmağın kapalı olması durumunda da sola doğru bir uzantı şeklinde uzamasıdır. İki el hareketinin ayırt edici özelliği sadece sola doğru olan uzantının boyutudur ve farklı el büyüklüğü ve hareketin gerçekleştiriliş şekline göre değişir. Bu nedenle farklı el büyüklükleri ve hareketin gerçekleştiriliş şekline göre farklı başarı oranları yüksek olduğu düşünülmektedir.

Eşik değeri uygulandığında bile, “shaka” hareketinin “başparmak sola” hareketi olarak sınıflandırıldığı ve “başparmak sola” hareketinin “yumruk” hareketi olarak sınıflandırıldığı durumlar hata matrislerinden görülebileceği üzere yüksek orandadır. Bu durum, başparmak ve serçe parmağının insan elindeki en kısa parmaklar olduğundan dolayı, çeşitli açılar altında daha da kısa görülebilmelerinden kaynaklı, uzantılarının doğru algılanamamasıyla açıklanabilir.

Sınıflandırma sonuçları incelendiğinde, “yumruk”, “dört” ve “avuç içi” el hareketlerinin oldukça yüksek bir yüzdeyle sınıflandırıldığı görülmektedir. Bu hareketlerin basit yapısı ve diğer hareketlerle benzerliğinin az olması bu yüksek oranın elde edilmesinde etkilidir.

Sınıflandırıcı başarımının değerlendirilmesinde kullanılan başka bir yöntem olan Cohen'in Kappa Katsayısı iki değerleyici arasındaki karşılaştırmalı uyuşmanın güvenilirliğini ölçen bir yöntemdir [32]. Bu çalışmada iki değerleyici olarak sınıflandırıcı ve kesin referans (ground truth) kullanılmıştır. Tablo 4.4. ve Tablo 4.5.'te %70'lik eşik değeri uygulanmamış ve uygulanmış sınıflandırma sonuçları için Cohen'in Kappa Katsayısı değerleri görülmektedir.

Tablo 4.4. %70’lik eşik uygulanmamış sınıflandırıcı için Cohen’in Kappa Katsayısı değeri

	Gözlemlenen Uyuşma (Observed Agreement)	Şans Eseri Uyuşma (Chance Agreement)	Kappa
0	0.8664	0.0141	0.6664
1		0.0119	
2		0.0087	
3		0.0090	
4		0.0090	
5		0.0102	
6		0.0064	
7		0.0095	
8		0.0101	
9		0.0111	

Tablo 4.5. %70’lik eşik uygulanmış sınıflandırıcı için Cohen’in Kappa Katsayısı değeri

	Gözlemlenen Uyuşma (Observed Agreement)	Şans Eseri Uyuşma (Chance Agreement)	Kappa
0	0.9256	0.0155	0.7256
1		0.0120	
2		0.0083	
3		0.0095	
4		0.0095	
5		0.0113	
6		0.0050	
7		0.0095	
8		0.0089	
9		0.0105	

Landis ve Koch [33] elde edilen Kappa değerlerini yorumlamak için 6 aralıklı bir ölçek kullanmıştır. Buna göre; değer sıfırdan küçükse “hiçbir uyuşma yok”, 0-0.20 arasında “hafif uyuşma”, 0.20 – 0.40 arasında “vasat uyuşma”, 0.40 – 0.60 arasında “orta derecede uyuşma”, 0.60 – 0.80 arasında “önemli derecede uyuşma” ve 0.80 – 1 arasında “neredeyse mükemmel uyuşma” değerleri olduğunu savunmuşlardır. Bu çalışmada hem eşik uygulanmamış hem de eşik uygulanmış şekilde sınıflandırılma yapıldığında, sırasıyla 0.67 ve 0.73’lük değerlerle önemli derecede uyuşma olduğu görülmüştür.

Sınıflandırma başarısı ölçümü yapılırken dikkate alınan bir diğer metrik ise F1 puanıdır [34]. F1 puanı duyarlılık ve kesinlik değerlerinin harmonik ortalamasıdır. Bu çalışmada yanlış negatif ya da yanlış pozitif sınıflandırmalar arasında önem farkı yoktur. Her iki durumda da istenilen grafik aksiyonu alınamaz. Bu yüzden yanlış pozitifte bağlı

olan kesinlik değeri ve yanlış negatife bağlı olan duyarlılık değeri karşılaştırıldığında, modelin performansı açısından biri diğerine oranlı daha önemli değildir. Modelin ortalama F1 puanları %70'lik eşik değeri uygulanmadığında 0.87, bahsedilen eşik değeri uygulandığında 0.92'dir. Bu da başarımlar oranlarıyla doğru orantılı bir sonuca işaret etmektedir. Sistemin sınıflandırma sonuçları üzerinden hesaplanan duyarlılık, kesinlik ve F1 puanı değerleri, eşik değeri uygulanmamış ve eşik değeri uygulanmış sınıflandırıcılar için sırasıyla Tablo 4.6. ve Tablo 4.7.'de gösterilmiştir.

Tablo 4.6. %70'lik eşik uygulanmamış sınıflandırıcı için F1 puanları

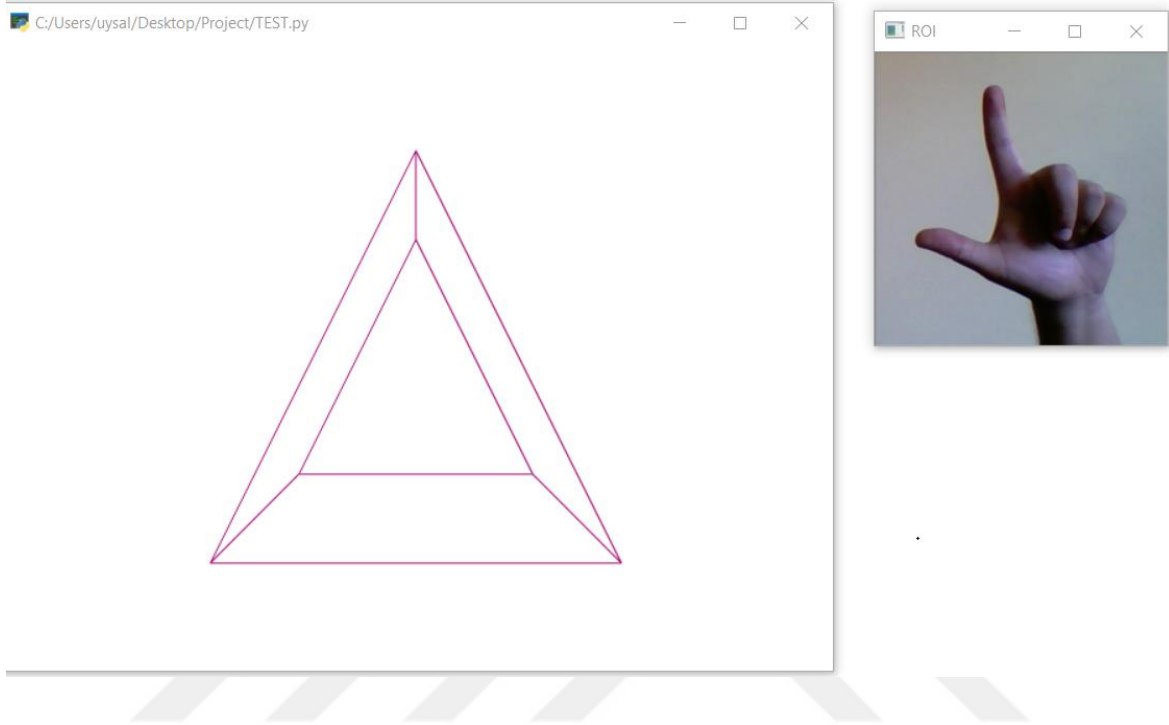
	Duyarlılık (Recall)	Kesinlik (Precision)	F1 Puanı (F1 Score)
0	1.000	0.712	0.832
1	0.965	0.811	0.881
2	0.851	0.983	0.912
3	0.883	0.979	0.928
4	0.869	0.961	0.912
5	0.964	0.941	0.953
6	0.629	0.977	0.765
7	0.874	0.916	0.895
8	0.710	0.704	0.707
9	0.922	0.834	0.876

Tablo 4.7. %70'lik eşik uygulanmış sınıflandırıcı için F1 puanları

	Duyarlılık (Recall)	Kesinlik (Precision)	F1 Puanı (F1 Score)
0	1.000	0.773	0.872
1	0.982	0.928	0.954
2	0.935	0.994	0.963
3	0.945	0.995	0.969
4	0.905	0.992	0.947
5	0.983	0.985	0.984
6	0.772	0.988	0.867
7	0.937	0.968	0.952
8	0.737	0.817	0.775
9	0.975	0.930	0.952

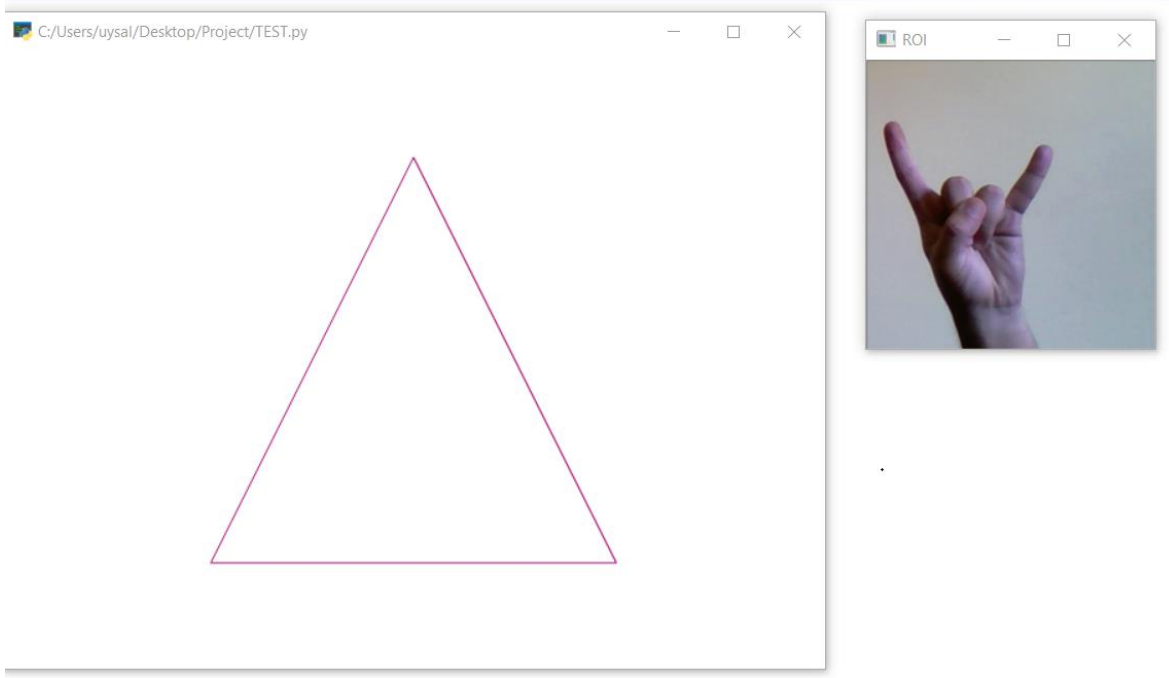
Geliştirilen sistemin çalışma prensibi, girdi olarak aldığı bir el hareketi sonucu, bu el hareketine karşılık gelen bir grafiksel operasyon gerçekleştirmektir. Sistemin desteklediği bazı el hareketleri ve bunlara karşılık gelen grafiksel işlemlerin ekran görüntüleri Şekil 4.2., 4.3., 4.4., 4.5., 4.6. ve 4.7.'de gösterilmiştir.

Şekil 4.2.'de gösterildiği üzere, sıradaki şekli çizme aksiyonu alınmıştır ve sistem, gösterilen el hareketini her algıladığında yazılım geliştirilirken belirtilen sırada olmak koşuluyla bir sonraki şekli ekrana çizdirir. Sistem aynı zamanda önceki şekli çizdirmek için de farklı bir el hareketini destekler.



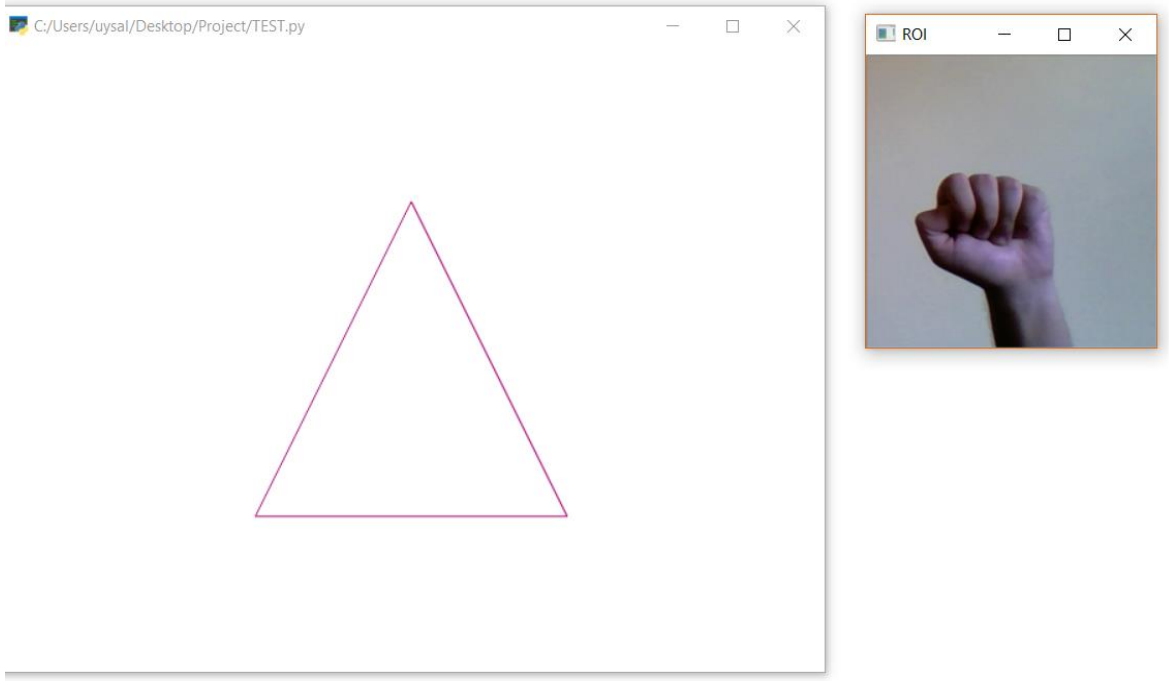
Şekil 4.2. Sıradaki Şekli Çizme el hareketi ve karşılık gelen grafik operasyonu

Şekil 4.3.'te gösterildiği üzere, sistem 2 boyut - 3 boyut dönüşümü hareketini algıladığında, o esnada ekranda bulunan şekil 2 boyutluysa o şeklin 3 boyutlu karşılığını, ekranda bulunan şekil 3 boyutluysa o şeklin 2 boyutlu karşılığını ekrana çizdirir. Bu çizdirme işlemi yapılırken ilk baştaki şeklin boyutu ve dönüş açıları dikkate alınır ve çizdirilen şekil bu değerlere göre çizdirilir.

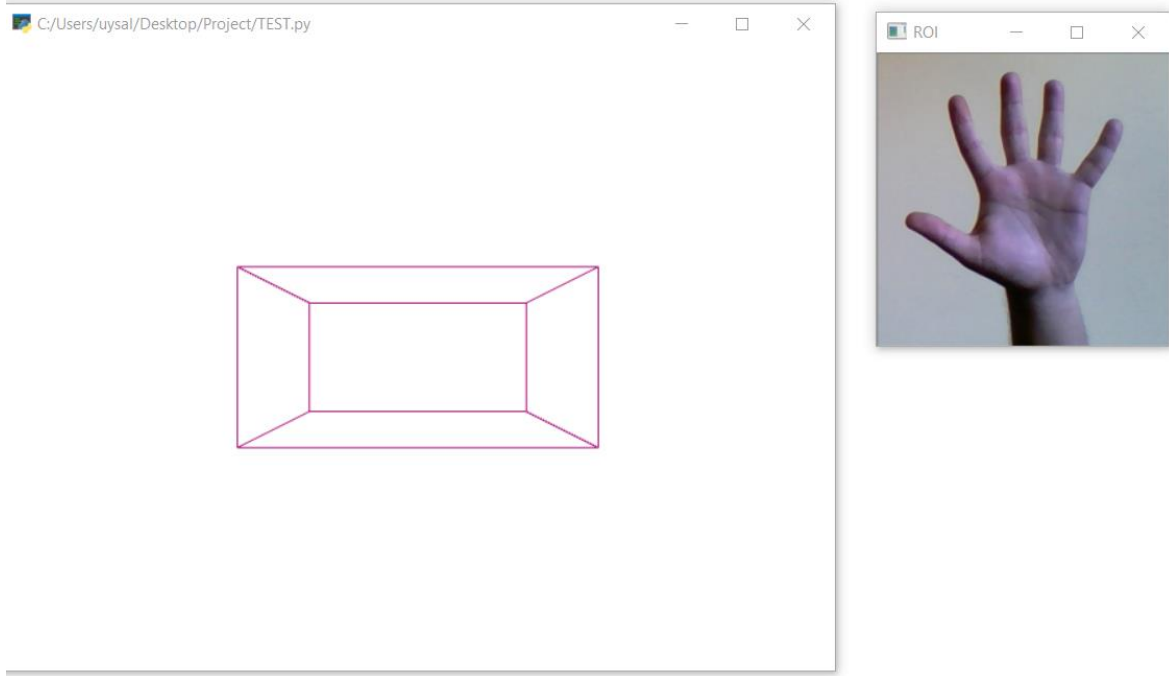


Şekil 4.3. 2 Boyut – 3 Boyut Değişirme el hareketi ve karşılık gelen grafik operasyonu

Şekil 4.4. ve Şekil 4.5.'te gösterildiği üzere, mevcut şekli küçültme/büyütme aksiyonu alınmıştır. Bu işlem yapılırken 5 aralıklı bir ölçek dikkate alınır ve küçültme/büyütme işlemi bu ölçek üzerinde birer kademe azaltılarak/artırılarak yapılır. Eğer o esnada ekranda bulunan şekil hali hazırda ölçeğin en küçük/büyük değerindeyse, sistem daha fazla küçültme/büyütme gerçekleştirmez ve herhangi bir aksiyon almaz.



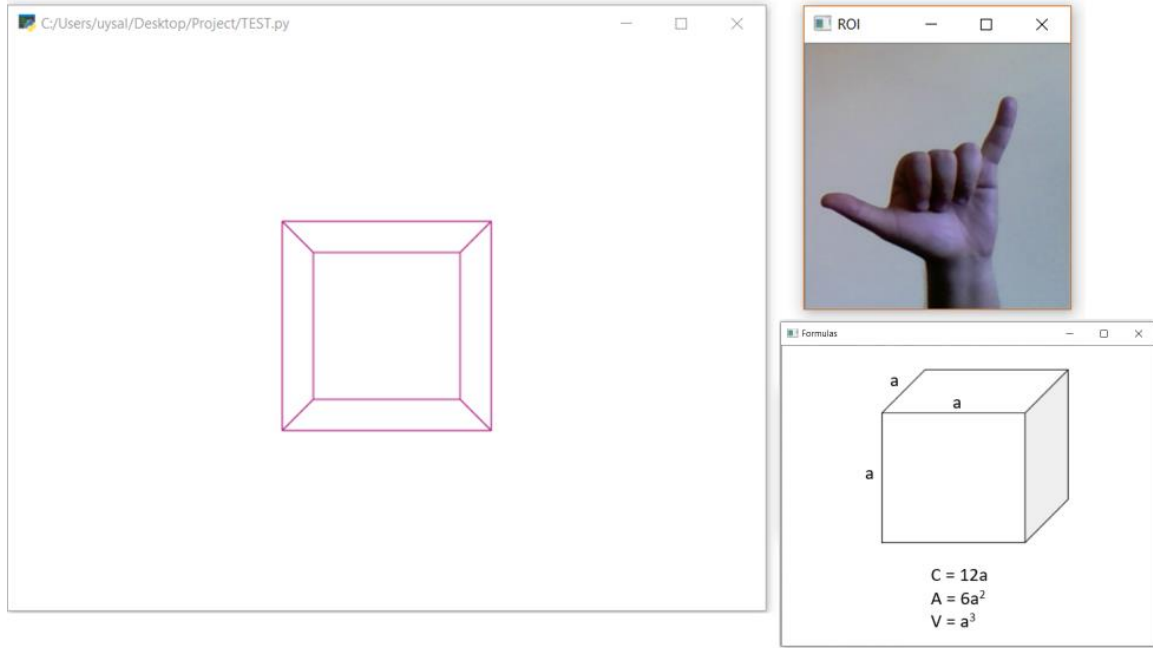
Şekil 4.4. Mevcut şekli küçültme el hareketi ve karşılık gelen grafik operasyonu



Şekil 4.5. Mevcut şekli büyütme el hareketi ve karşılık gelen grafik operasyonu

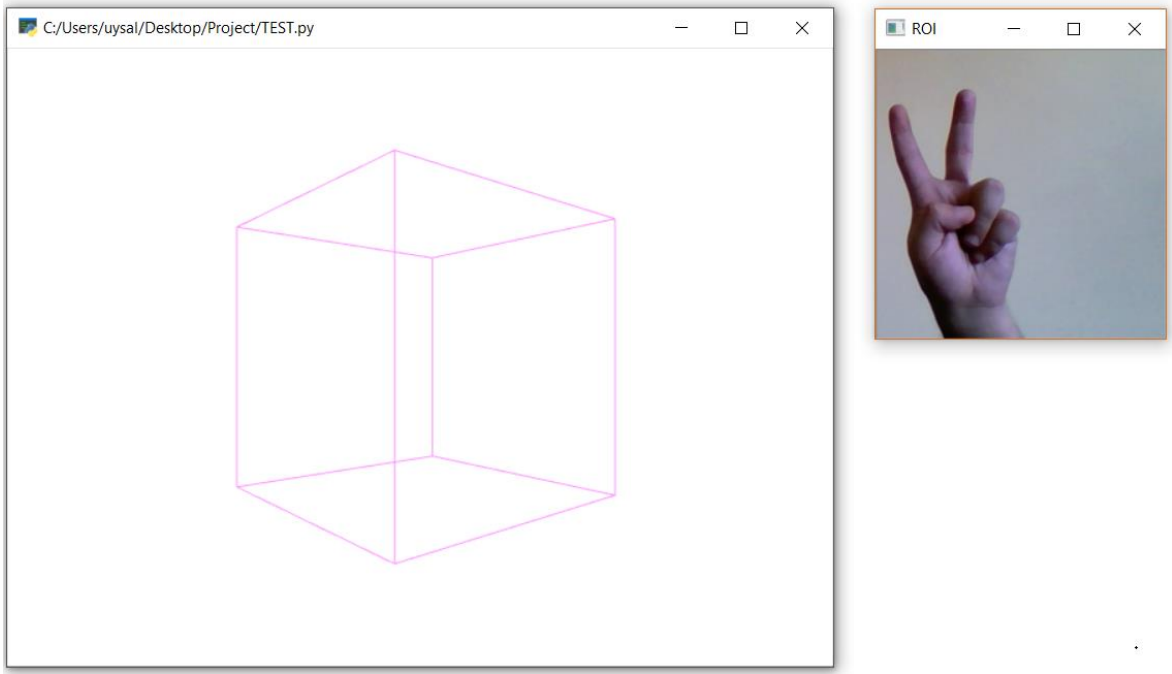
Şekil 4.6.'da gösterildiği üzere, mevcut şekil için çevre, alan ve hacim formülleri gösterme aksiyonu alınmıştır. Bu el hareketi algılandığı zaman o esnada ekrandaki şekil için çevre, alan ve varsa hacim formüllerini gösteren bir pencere daha açılır. Açılan

pencerede, anlamayı kolaylaştırmak için, şeklin bir statik temsili görüntüsü ve kenar, yükseklik, yarıçap gibi özellikleri etiketlenmiştir ve formüller bu etiketlere göre gösterilir.



Şekil 4.6. Mevcut şekil için formülleri gösterme el hareketi ve karşılık gelen formülleri gösterme operasyonu

Şekil 4.7.'de gösterildiği üzere, mevcut şekli Y-ekseni etrafında döndürme aksiyonu alınmıştır. Bu el hareketi algılandığı zaman o esnada ekrandaki şekil Y-ekseni etrafında döndürülür. Sistem bu el hareketi dışında, mevcut şekli X-ekseni etrafında döndürmek için de ayrıca bir el hareketini daha desteklemektedir. Aynı zamanda döndürme işlemleri için sistemin desteklediği başka bir el hareketi de döndürme yönünü değiştirmeye yönelik el hareketidir. X-ekseni ve Y-ekseninde döndürme el hareketleri, döndürme yönü değişkenine bağlı bir şekilde çalışır ve bu değişkene göre pozitif veya negatif yönde döndürme işlemi yaparlar. Döndürme yönünü değiştir el hareketi bu bahsedilen değişkenin değerini değiştirmek için kullanılır.



Şekil 4.7. Mevcut Şekli Y ekseninde döndürme el hareketi ve karşılık gelen grafik operasyonu

Sistem genel anlamda değerlendirildiğinde, geometri eğitimi konusunda HGR tabanlı çalışan bir uygulama olmasıyla öne çıkmaktadır. Küçük bir gecikmeyle de olsa gerçek zamanlı çalışabilmesiyle de öne çıkan bu uygulama, ortalama bir donanım ile kullanılabilir. Ayrıca, sistemin eğitim çağındaki çocuklar için ulaşılabilir ve düşük maliyetli bir sistem üzerinde çalışabilir olması hedeflenmektedir. Yüksek fonksiyonellik ve görsellik sunması da sistemin avantajlarından bir diğeridir. El hareketlerinin tanınması konusunda hareketlerin yapılış şeklinde esneklik tanır, ölçekleme, döndürme gibi yapısal farklılıklardan etkilenmez, standart aydınlanma ve aydınlanma üzerinde çok büyük bir değişiklik yapılmadığı sürece yüksek başarı oranını sürdürür.

Sistemin zayıf yanlarından bahsedecek olursak, düz bir arka plana gereksinim duyar ve karmaşık arka planlarda kullanıldığı zaman başarı yüzdesi düşmektedir. Ortalama donanımlarda elde edilen performans küçük bir gecikmeye sebep olur. Bu gecikme gerçek zamanlı kullanımı etkilemeyecek boyutta olmasına rağmen çok iyi bir kullanıcı deneyimi oluşturamamaktadır. Standart aydınlanma altında yüksek başarı yüzdesiyle çalışan uygulamanın doğrudan görüntü kaynağına ya da ele yüksek şiddette ışık miktarının düşmesi, arka planda belirgin gölge oluşması durumunda sınıflandırma performansı düşmektedir. Sistemin başka bir zayıf yönü, birbirine oldukça benzeyen el hareketlerinin

ayrımını yaparken ayrışabilen hareketlere oranla daha düşük bir başarı yüzdesi göstermesidir.



5. SONUÇ

Bu çalışma kapsamında öğrenme çağındaki çocuklar için el hareketleri tanıma yoluyla kullanılabilen bir geometri öğrenme uygulaması geliştirilmiştir. Bu uygulama geometri öğrenimini görselleştirecek ve kullanıcıların standart öğrenme yöntemlerine göre daha başarılı bir biçimde kavramasına imkân tanıyacaktır.

Geliştirilen uygulamada el hareketi tanıma modülü ve grafiksel operasyonlar modülü olmak üzere iki farklı modül bulunmaktadır. El hareketi tanıma modülü Python programlama diliyle geliştirilmiş olup Keras derin öğrenme kütüphanesi kullanılmıştır. El hareketlerinin sınıflandırılması için bir CNN mimarisi kullanılmıştır ve bu mimari OxfordNet'in VGG16 modeli olarak seçilmiştir. Modelin eğitilmesi için kullanılan veri kümesi 6 kişinin, her bir el hareketi için görüntülerinin alınması ve bu görüntülerin çeşitli veri arttırma işlemleri ile çoğaltılması şeklinde oluşturulmuştur. Eğitim işleminde, her bir el hareketi için 1200 görüntü kullanılmış, eğitim ve doğrulama başarı oranı %93 civarında gözlenmiştir. Derin öğrenmenin verdiği esneklikten dolayı sistem, alınan görüntüler üzerinde el bölgesinin kırılması haricinde hiçbir ön işleme gerektirmeden ham veri üzerinde başarılı bir biçimde çalışabilmektedir. Uygulama gerçek zamanlı çalışabilecek şekilde geliştirilmiştir ancak test edilen sistemin orta sınıf ve eski bir ekran kartına sahip olması nedeniyle, 0.2 saniye civarında bir gecikme yaşanabilmektedir.

Sistemde bulunan ikinci modül, grafiksel operasyonlar ile ilgilidir. Bu modül el hareketi algılama modülünün verdiği çıktıyı girdi olarak alıp el hareketine uygun bir grafik işlemi gerçekleştirmektedir. Bu modül de Python programlama dilinde geliştirilmiş olup Pyglet isimli OpenGL tabanlı bir grafik kütüphanesi kullanmaktadır. Alınan girdiye göre, şekil çizme, mevcut şekli büyütme, küçültme, döndürme, boyut dönüştürme ve şekil için çevre, alan ve hacim formülleri gösterme gibi işlemleri yerine getirmektedir.

Gelecek dönemlerde yapılacak çalışmalarda, sisteme yeni şekiller eklenerek açılış işlemi, şekillerin bölümlendirilmesi ve aynı anda birden fazla şekil üzerinde işlem yapılması gibi yetenekler kazandırılabilir. Sisteme grafiksel estetik kazandırmak adına ışıklandırma eklenebilir, doku ekleme işlemleri yapılarak iskelet görüntü yerine dokuyla kaplanmış şekiller gösterilebilir. Bunların dışında uygulamanın daha erişilebilir olması ve daha alt seviye donanımlarda daha yüksek performansla çalışabilir hale gelmesi için CNN modeli üzerinde performans optimizasyonu yapılması düşünülebilir. Ayrıca bu sistemin, el

hareketi tabanlı alıřacak eđitim uygulamaları iin nemli bir altlık alıřma olabileceđi dřnlmektedir.



KAYNAKLAR

- [1] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp. 1097-1105, 2012
- [2] S. S. Rautaray and A. Agrawal, "Interaction with virtual game through hand gesture recognition," in *International Conference on Multimedia, Signal Processing and Communication Technologies*, pp. 244-247, 2011
- [3] Q. Chen, N. D. Georganas, E. M. Petriu, "Hand gesture recognition using haar-like features and a stochastic context-free grammar," *IEEE Transactions on Instrumentation and Measurement*, vol 57(8), pp 1562-1571, 2008
- [4] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR 2001, Vol 1, pp. 511-518, 2001
- [5] G. R. Bradski, "Computer vision face tracking as a component of a perceptual user interface," in *Workshop on Applications of Computer Vision*, Princeton, NJ, pp 214-219, 1998
- [6] B. Jean-Yves "Pyramidal implementation of the Lucas Kanade feature tracker description of the algorithm," *Intel Corporation, OpenCV Documents*.
- [7] H. A. Jalab & H. K. Omer, "Human computer interface using hand gesture recognition based on neural network," in *2015 5th National Symposium on Information Technology: Towards New Smart World (NSITNSW)*, pp. 1-6, 2015
- [8] M. Czupryna, & M. Kawulok, "Real-time vision pointer interface," in *Proceedings ELMAR-2012*, pp. 49-52, 2012

- [9] A. Shimada, T. Yamashita & R. I. Taniguchi, "Hand gesture based TV control system - towards both user & machine - friendly gesture applications," *The 19th Korea-Japan Joint Workshop on Frontiers of Computer Vision*, pp. 121-126, 2013
- [10] S. Y. Peng, K. Wattanachote, H. J. Lin, & K. C. Li, "A real-time hand gesture recognition system for daily information retrieval from internet," *Fourth International Conference on Ubi-Media Computing*, pp. 146-151, 2011
- [11] J. Ram Rajesh, R. Sudharshan, D. Nagarjunan & R. Aarthi, "Remotely controlled PowerPoint presentation navigation using hand gestures," *Proc. of the Intl. Conf. on Advances in Computer, Electronics and Electrical Engineering*, 2012
- [12] K. Tripathi & N. B. G. Nandi, "Continuous Indian sign language gesture recognition and sentence formation," in *Procedia Computer Science*, vol 54, pp 523-531, 2015
- [13] J. B. Kim, K. H. Park, W. C. Bang & Z. Z. Bien, "Continuous gesture recognition system for Korean sign language based on fuzzy logic and hidden Markov model," in *2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. Proceedings*, Vol. 2, pp. 1574-1579, 2002
- [14] S. Sahasrabudhe, S. H. A. H. Adeet, M. Thakkar, V. Thakkar & I. Y. E. R. Sridhar, "MathMazing: 3D gesture recognition exergame for arithmetic skills," in *Proceedings of the 20th International Conference on Computers in Education*, 2012
- [15] N. Gosalia, P. Jain, I. Shah, A. R. Joshi, N. Katre & S. Sahasrabudhe, "3D Gesture-recognition based animation game," in *Procedia Computer Science*, vol 45, pp 712-717, 2015
- [16] M. T. Yang & W. C. Liao, "Computer-assisted culture learning in an online augmented reality environment based on free-hand gesture interaction," in *IEEE Transactions on Learning Technologies*, vol 7(2), pp 107-117, 2014

- [17] A. Karambakhsh, A. Kamel, B. Sheng, P. Li, P. Yang & D. D. Feng, "Deep gesture interaction for augmented anatomy learning" in *International Journal of Information Management*, vol 45, pp 328-336, 2019
- [18] D. Yang, J. K. Lim & Y. Choi, "Early childhood education by hand gesture recognition using a smartphone based robot," in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pp 987-992, 2014
- [19] M. Kass, A. Witkin & D. Terzopoulos, "Snakes: Active contour models," in *International Journal of Computer Vision*, vol 1(4), pp 321-331, 1988
- [20] H. Q. Le, & J. I. Kim, "An augmented reality application with hand gestures for learning 3D geometry," in *2017 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 34-41, 2017
- [21] H. Kaufmann, D. Schmalstieg & M. Wagner, "Construct3D: a virtual reality application for mathematics and geometry education," in *Education and Information Technologies*, vol 5(4), pp 263-276, 2000
- [22] Z. Szalavári & M. Gervautz, "The personal interaction panel - a two-handed interface for augmented reality," in *Computer Graphics Forum*, Vol. 16(3), pp. C335-C346, 1997
- [23] J. M. Vitale, M. I. Swart & J. B. Black, "Integrating intuitive and novel grounded concepts in a dynamic geometry learning environment," in *Computers & Education*, vol 72, pp 231-248, 2014
- [24] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000
- [25] F. Chollet, "keras, GitHub," *GitHub repository*, <https://github.com/fchollet/keras>, 2015

- [26] K. Fukushima, "Neocognitron: A hierarchical neural network capable of visual pattern recognition," in *Neural Networks*, vol 1(2), pp 119-130, 1988
- [27] Y. LeCun, L. Bottou, Y. Bengio & P. Haffner, "Gradient-based learning applied to document recognition," in *Proceedings of the IEEE*, vol 86(11), pp 2278-2324, 1998
- [28] K. Simonyan & A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint*, pp 1409-1556, 2014
- [29] I. E. Sutherland, "Sketchpad: A man-machine graphic communication system," in *AFIPS Conference Proceedings, 1963 Spring Joint Computer*, pp 323–328, 1963
- [30] H. Lv & Z. Wu, "A survey of computer graphics and graphics image processing technology," in *2011 2nd International Symposium on Intelligence Information Processing and Trusted Computing*, pp. 196-198, 2011
- [31] B. Kabasakal, "Uzaklık ve cinsiyet tabanlı akıllı reklam görüntüleme sistemi," *Master's thesis, Başkent Üniversitesi Fen Bilimleri Enstitüsü*, 2018
- [32] J. Cohen, "A coefficient of agreement for nominal scales," in *Educational and Psychological Measurement*, Vol.20, No.1, pp.37-46, 1960
- [33] J. R. Landis & G. G. Koch, "The measurement of observer agreement for categorical data," in *Biometrics*, Vol.33, pp.159-174
- [34] C. J. Van Rijsbergen, "Information retrieval/CJ van Rijsbergen," London, Butterworths, 1979