



**BİLGİSAYAR GÖRÜNTÜSÜ VE YAPAY ZEKA KULLANARAK ALMA
VE YERLEŐTİRME GÖREVİ İÇİN GÖZ BAKIŐLI ROBOTİK KOL
KONTROLÜ**

Fahad ATA

**YÜKSEK LİSANS TEZİ
ELEKTRİK-ELEKTRONİK MÜHENDİSLİĐİ ANA BİLİM DALI**

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

ARALIK 2024

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Fahad ATA

06/12/2024

BİLGİSAYAR GÖRÜNTÜSÜ VE YAPAY ZEKA KULLANARAK ALMA VE YERLEŞTİRME GÖREVİ İÇİN GÖZ BAKIŞLI ROBOTİK KOL KONTROLÜ

(Yüksek Lisans Tezi)

Fahad ATA

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Aralık 2024

ÖZET

Bu tez, ciddi hareket bozukluğu olan bireyler için özerklik ve yaşam kalitesinde anlamlı iyileştirmeler sunmak amacıyla robotik bir kolu kontrol etmek için göz bakışı izlemenin gücünden yararlanma konusunda yenilikçi bir yolculuğa çıkıyor. Amacımız, en son teknoloji ürünü bilgisayar görüşü ve yapay zekayı kullanarak fiziksel engelli insanlar için anlamlı özerklik konusunda ilerlemeler sağlamaktır. Araştırma, kontrol algoritmalarının geliştirilmesi ve test edilmesi için sağlam bir çerçeve sağlayan Robot İşletim Sisteminin (ROS) simülasyon için Gazebo ile entegrasyonuna dayanmaktadır. Metodoloji, doğru algılama ve gerçek zamanlı işleme için evrişimli sinir ağlarından (CNN'ler) yararlanarak, hassas göz bakışı takibi için bir bilgisayarlı görme sisteminin ayrıntılı tasarımını kapsar. Ek olarak, robot kolunun ters kinematığı, göz bakışı koordinatlarını uygulanabilir komutlara dönüştürmek için hesaplanır. Deneysel doğrulama, farklı düzeylerde fiziksel engeli olan katılımcıların yer aldığı bir dizi kontrollü deneme aracılığıyla gerçekleştirildi. Sonuçlar, geleneksel kontrol yöntemleriyle karşılaştırıldığında görev tamamlama sürelerinde ve doğrulukta önemli gelişmeler olduğunu gösterdi. Kullanıcı geri bildirimleri, sistemin sezgisel doğasını ve gereken fiziksel çabanın önemli ölçüde azaldığını vurguladı. Sonuç olarak, bu çalışma, yardımcı teknoloji için geçerli bir çözüm sunmakta ve engelli topluluk için kapsayıcı araçlar oluşturmada yapay zeka ve bilgisayarlı görmenin potansiyelini vurgulamaktadır. Gelecekteki çalışmalar, sistemi daha geniş uygulama senaryoları için iyileştirmeye ve farklı koşullar altında bakış izleme algoritmasının sağlamlığını artırmaya odaklanacak.

Bilim Kodu : 93438
Anahtar Kelimeler : Arm, makine öğrenimi, nesne algılama, tensorflow, raspicam, sanal ağ bilişimi
Sayfa Adedi : 67
Danışman : Doç. Dr. Uğurhan KUTBAY

EYE GAZE ROBOTIC ARM CONTROL FOR PICK AND PLACE TASK USING
COMPUTER VISION AND ARTIFICIAL INTELLIGENCE

(M. Sc. Thesis)

Fahad ATA

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

December 2024

ABSTRACT

This thesis embarks on an innovative journey in harnessing the power of eye gaze tracking to control a robotic arm, to offer meaningful enhancements in autonomy and quality of life for individuals with severe mobility impairment. Our goal is to provide advances in meaningful autonomy for physically disabled people using state-of-the-art computer vision and artificial intelligence. The research is anchored in the integration of the Robot Operating System (ROS) with Gazebo for simulation, providing a robust framework for developing and testing the control algorithms. The methodology encompasses the detailed design of a computer vision system for precise eye gaze tracking, leveraging convolutional neural networks (CNNs) for accurate detection and real-time processing. Additionally, the inverse kinematics for the robotic arm is computed to translate eye gaze coordinates into actionable commands. Experimental validation was conducted through a series of controlled trials involving participants with varying levels of physical disability. The results demonstrated significant improvements in task completion times and accuracy compared to traditional control methods. User feedback highlighted the system's intuitive nature and the substantial reduction in physical effort required. In conclusion, this study presents a viable solution for assistive technology, emphasizing the potential of AI and computer vision in creating inclusive tools for the disabled community. Future work will focus on refining the system for broader application scenarios and enhancing the robustness of the gaze tracking algorithm under diverse conditions.

Science Code : 93438

Key Words : Arm, machine learning, object detection, tensorflow, raspicam, virtual network computing

Page Number : 67

Supervisor : Assoc. Prof. Uğurhan KUTBAY

TEŞEKKÜR

Duygularımı kelimelerle ifade edebilirim ama elimden geldiğince teşekkür etmeye çalışacağım. Öncelikle danışmanım Gazi Üniversitesi'ndeki araştırma yolculuğum boyunca sürekli destekleri, rehberlikleri ve paha biçilmez tavsiyeleri için Doç. Dr. Uğurhan KUTBAY'a teşekkür ederim. Onların içgöruları ve teşvikleri bu tezin başarılı bir şekilde tamamlanmasında etkili oldu. Komite üyelerim Prof. Dr. Fırat HARDALAÇ ve Prof. Dr. Kemal POLAT. Bu çalışmanın kalitesini önemli ölçüde artıran yapıcı geri bildirimleri ve önerileri için. Gazi Üniversitesi'ndeki meslektaşlarıma ve arkadaşlarıma da dostlukları ve araştırma deneyimimi zenginleştiren teşvik edici tartışmaları için minnettarım. Deneylerin teknik yönlerinde yardımcı olan Ali GÖZÜM'e ve zor zamanlarda manevi destek ve motivasyon sağlayan Mohamamad Rafay QADRI ve Zeeshan MALİK'e özel teşekkürler. Ayrıca Gazi Üniversitesi'ndeki idari personele de yardımlarından dolayı teşekkür etmek istiyorum. Çalışmalarımın prosedüre yönlerini yönlendirmede. Verimlilikleri ve yardımseverlikleri büyük ölçüde takdir edildi. Son olarak, sarsılmaz destekleri, sabırları ve teşvikleri için aileme en derin şükranlarımı sunmak isterim. Onların bana olan inancı benim itici gücüm oldu.

Bu çalışmaya doğrudan ya da dolaylı katkısı olan herkese teşekkür ediyorum. Desteğiniz çok değerliydi.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	x
SİMGELER VE KISALTMALAR.....	xi
1. GİRİŞ.....	1
2. LİTERATÜR İNCELEMESİ.....	15
2.1. Literatür Taraması	15
3. GÖZ ALGILAMA VE KIRIŞ SAYAÇ	25
3.1. Göz Algılama	25
3.2. Göz Algılama İçin Derin Öğrenme Modeli.....	25
3.2.1. Veri toplama.....	25
3.2.2. Görüntü ön işleme	25
3.2.3. Evrimsel sinir ağı	26
3.2.4. VGG-19.....	28
3.2.5. RestNet101V2.....	28
3.3. Mediapipe Kullanarak Göz Algılama	30
3.3.1. Gözlerin tespit edilmesindeki zorluklar	31
3.4. Göz Kırpma Algılama	32
3.4.1. Göz en-boy oranını anlamak	32
3.4.2. Yüzdeki yer işaretlerini kullanarak göz kırpma algılama	33
3.5. Mediapipe Mimarisi	34

4. NESNE ALGILAMA YOLOV8'İ KULLANARAK	39
4.1. Nesne Algılama	39
4.2. Mimari YOLOV8	39
4.3. Kayıplar	41
4.4. Eğitim	42
5. ROBOTİK İŞLETİM SİSTEMİ M(ROS) VE GAZEBO ORTAM KURULUMU	45
5.1. ROS	45
5.2. Gazebo	46
5.3. Kullanıcı Arayüzü (UI)	50
5.5. İleri Kinematik Ters Kinematik	55
5.5.1. İleri kinematik denklemler	55
5.5.2. Ters kinematik denklemler	55
6. SONUÇ	59
KAYNAKLAR	61
ÖZGEÇMİŞ	67

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 5.1. Projede kullanılan donanım bileşenleri	52



ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2.1. Wang ve diğerlerinin noktalarından göz açıklığı derecesi.	18
Şekil 2.2. Göz kapalı veya açıkken gözlerin açıklık derecesi.....	19
Şekil 2.3. Göz açık ve kapalıyken gözdeki yer işaretleri.....	19
Şekil 3.1. CNN Mimarisi	27
Şekil 3.2. RestNet101V2 Mimarisi.....	29
Şekil 3.3. CNN, VGG19 ve ResNet101v2 doğruluklarının karşılaştırılması	29
Şekil 3.4. İzolasyon sonrası maske üzerindeki şekillerin çizilmesi.....	36
Şekil 3.5. Göz kırpmalarını Tespit Etme ve Sayma.....	37
Şekil 3.6. Sol ve sağ göz hareketlerini algılama	37
Şekil 4.1. Ultralytics'ten YOLOV8 Mimarisi	40
Şekil 4.2. Eğitim sırasındaki metriklerin grafiği.....	43
Şekil 4.3. Eğitim kayıpları	43
Şekil 4.4. Doğrulama kayıpları	43
Şekil 5.1. ROS Ortamı	47
Şekil 5.2. Robotik kol alma ve yerleştirme hareketi.....	48
Şekil 5.3. Nesne tanımlama	51
Şekil 5.4. Nesne seçimi ve gereksinime göre görevin hareketi.	51
Şekil 5.5. Birleştirilmiş robotik kol.....	54
Şekil 5.6. Bir noktadan diğerine alma ve yerleştirme görevini gösterir	57
Şekil 6.1. Entegre görevin göz açıp kapayınca kadar yürütülmesi	60
Şekil 6.2. Nesneyi kavrama ve alma başlangıcı.....	60
Şekil 6.3. Son pozisyona ulaşmak ve nesneyi yerleştirmek.....	60

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılan semboller ve kısaltmalar, açıklamalarıyla birlikte aşağıda sunulmaktadır.

Simgeler

Açıklamaları

cm	Santimetre
dB	Desibel
Hz	Hertz
m²	Metrekare
m³	Metreküp

Kısaltmalar

Açıklamalar

ACC	Doğruluğu Yapay Zeka Yapay Zeka
AR	Artırılmış Gerçeklik
BCI	Beyin-Bilgisayar Arayüzü
CNN	Evrişimsel Sinir Ağı
DDS	Veri Dağıtım Hizmeti
DNN	Derin Sinir Ağı
EEG	Elektroensefalogram
EOG	Elektrookülografi
GAZEBO	Açık kaynaklı bir 3D robotik simülatörü
h-BCI	Hibrit Beyin-Bilgisayar Arayüzü
HCI	İnsan-Bilgisayar Etkileşimi
IMU	Atalet Ölçüm Birimi
KDL	Kinematik Dinamik Kitaplığı
LiDAR	Işık Algılama ve Aralığı
MS-COCO	Microsoft Ortak Nesnelere Bağlamda
PCA	Temel Bileşen Analizi
RESNET101V2	Artık Ağ sürüm 2, 101 katmanlı
ROS	Robot İşletim Sistemi
SLAM	Eşzamanlı Yerelleştirme ve Haritalama

Kısaltmalar**Açıklamalar****SSMI**

Şiddetli Konuşma ve Hareket Bozukluğu

SSVEP

Kararlı Durum Görsel Uyarılmış Potansiyeller

SVD

Tekil Değer Ayrışımı

TB6600

Spesifik bir step motor sürücüsü ve kontrol panosu

VGG-19

Evrışimli sinir ağı mimarisi VGG Group

VR

Sanal Gerçeklik

YOLO

Bir Kez Bakarsınız (Nesne Tespit Sistemi)



1. GİRİŞ

BT teknolojik gelişmeleri barındırmakta ve bilgisayarların söylediklerimizi ve onlara nasıl baktığımızı anladığı bir dünya sunmaktadır. İmkansız gibi görünse de gelecek yakın. Bilgisayar görüşü, makinelerin görsel girdileri görmesine, etkinleştirmesine ve yanıt vermesine olanak tanıyan yenilikçi bir teknolojidir. Bilgisayarın klavyede hangi tuşlara basacağını, el ve göz hareketlerini bildiğini hayal edin. Bilgisayar görüşünün büyüğü budur. Bu tez, Bilgisayarla Görme uygulamasının yüksek fikirli bir şekilde uygulanmasına yönelik çözümü, engelli insanlara yardım etmeyi araştıracaktır. Şimdi, felçli bir hastanın elleriyle değil gözleriyle hareket eden ve kontrol edilen bir robot kolunu hareket halindeyken hayal edelim. Bu bir rüya değil; bu uygulama doğru olabilir. Bilgisayarlı görme ve yapay zekanın yeteneklerinden yararlanarak, günlük yaşamlarında fiziksel zorluklarla karşılaşanlar için yeni fırsatların kilidini açmayı hedefliyoruz.

Peki ne olmuş yani? Ancak asıl önemli olan, engellilerin hayatları üzerindeki potansiyel dönüştürücü etkisindedir. Bir fincan almak gibi sıkıcı, günlük görevlerden daha karmaşık eylemlere kadar, bilgisayar görüşünün birleşimi, teknolojinin yetenek ve engellilik arasındaki köprü olduğu bir geleceğin altını çiziyor.

Bu tezin sayfalarını bilgisayarlı görüntünün çamaşır makinesinden geçireceğiz, teknolojinin gizemini çözeceğiz, durulayacağız ve tekrarlayacağız; böylece, kendileri için mümkün olduğu kadar çok bağımsızlık isteyenlere yardım etme potansiyelini büyük bir dikkatle ortaya çıkaracağız. O halde sıkı durun ve bu teknolojik hız trenine inerken keyfini çıkarın; Sadeliğin yenilikçilikle buluştuğu yerde, görme gücü tek aslına uygunluktur!

Arka plan bilgileri

Özellikle tetraplejik hastalar için, elektrookülografi (EOG) sinyalleriyle yönlendirilen üç boyutlu (3D) bir robot kolla nesnelere doğru bir şekilde kavrama sorunu önemlidir. EOG bakış tahmini her zaman nesne seçimi ve mekansal hassasiyet sorunları ile mücadele etmiştir, bu da tutarsız sonuçlara ve kavrama hatalarına neden olmuştur. Bu sorunu aşmak için yeni bir strateji, kamera görüntüsüne dayalı nesne tanımayı EOG bakış tahminiyle harmanlıyor. Bir EOG ölçüm analizörü, görüntüleri görselleştirmek için bir ekran, bir robot kolu ve üst ve yan kameraların tümü sisteme entegre edilmiştir. Kamera görüntüleri

sayesinde kullanıcılar robot kolunu kontrol edebiliyor ve EOG bakış tahmini, hedef nesnenin tanımlanmasına yardımcı oluyor. Deneyle, EOG bakış tahmininde önerilen yaklaşımın 1,8–3,0 cm'lik bir mesafe hata aralığına ulaştığını göstermektedir. Sistem, 3,0 cm'lik bir eşik uygulandığında 2,0 cm'lik bir eşik uygulandığında %27 daha yüksek bir kavrama hızı elde eder ve bu da uygun eşiklerin seçilmesiyle elde edilebilecek nesne kavrama performansında kayda değer bir iyileşme olduğunu gösterir. Sistemin faydaları arasında yer değiştirme görevleri için başarılı tutucu-hedef etkileşimi, temel görüntü işleme yoluyla artırılmış EOG bakış tahmini ve uzaktan robot kol kontrolü yer alır; bunların tümü, kullanıcı konumlandırma ve görüntü yakalama koşulları gibi sınırlamalara rağmen mümkündür. Bununla birlikte, zorluklar arasında, elde edilen görüntünün koşullarına duyarlılık ve kullanıcı sınırlamaları yer almaktadır (Amri bin Suhaimi, Matsushita, Kitamura, Laksono ve Sasaki, 2023).

Yardımcı manipülasyon sistemlerinde insan-robot işbirliğinin kritik sorunu, motor bozukluğu olan kişilerin bağımsızlığını artırmak için göz izlemeli yardımcı robot kontrol sisteminin geliştirilmesidir. Geliştirilmiş bir kontrol mimarisi ve grafiksel bir kullanıcı arayüzü kullanan proje, tekerlekli sandalyelerin ve tekerlekli sandalyeye monte edilen robot kollarının günlük yaşam aktiviteleri için etkin bir şekilde kontrol edilmesini sağlamayı amaçlamaktadır. Kullanıcı araştırmasındaki on sağlıklı gönüllü, görevleri %100 başarı oranıyla tamamlayarak sistemin etkinliğini vurguluyor. Kontrollü robotik kol, farklı yüksekliklerdeki nesnelere toplayarak uyarlanabilirliğini ortaya koyuyor. Yöntem şu anda sağlıklı gönüllülerle doğrulanıyor olsa da, çalışma, güçlü sonuçlar elde etmek için üst ekstremiteler hareket sorunları olan kişileri içeren ek araştırmaların gerekli olduğunu kabul ediyor. Olumlu katılımcı geri bildirimleri, göz bakışı arayüzünün ne kadar sezgisel olarak tasarlandığını vurguluyor ve daha iyi kullanım için nesne algılama özelliklerinin eklenmesi gibi gelecekteki iyileştirmeler planlanıyor. Kontrol mimarisi, motor fonksiyon bozukluğu olan ve günlük işleri yapmakta zorlanan kişiler için uygun bir çözüm sunuyor. İnsan-robot işbirliğinde başarıyla kanıtlanmıştır. Gelecekte, kontrol sistemi görevleri daha hızlı tamamlayacak şekilde iyileştirilecek ve üst hareket bozukluğu olan hedef kullanıcı grubu daha kapsamlı değerlendirilecektir (Ali, Hashim ve Al-Sakkal, 2022).

Ayrıca, beyin-bilgisayar arayüzü (BCI) sistemlerinin entegre edilmesi, hızla gelişen insan-robot etkileşimi alanında nörodejeneratif hastalıkları olan insanlara odaklanarak yardımcı robot kontrol yöntemlerini geliştirecektir. Bu çalışma, elektroensefalogram (EEG)

sinyallerindeki oküler artefaktın kasıtlı üretiminden ve yüksek sinyal-gürültü oranından yararlanan yeni bir BCI sunmaktadır. Bu çığır açan strateji, yan göz hareketlerini tanımlamak için F7 ve F8 kanallarında toplanan sinyallerdeki yapılandırılmış tepe ve vadi yapılarını kullanır. Bu teknik, hem zayıf hem de düzenli göz kırpmaları tespit eden, göz kırpması tespiti için çift eşik yaklaşımının kullanılması bakımından önceki yöntemlerden farklıdır. Algoritma gerçek zamanlı olarak uygulandığında, grafiksel bir kullanıcı arayüzü yardımcı robotu kontrol edebilir. Doğrulama çalışmalarına katılan beş katılımcı, BCI'nın robotu başarılı bir şekilde kontrol edebildiğini doğruladı. Çalışma, hem çevrimdışı hem de gerçek zamanlı senaryolardaki performansları değerlendiriyor, olumlu sonuçlar sunuyor ve BCI'nın engelli kişilerin yaşam kalitesini artırma potansiyelini vurguluyor. Oluşturulan algoritma, göz artefaktlarının tanımlanması ve sınıflandırılmasındaki zorlukları çözerek gerçek dünyadaki uygulamalar için potansiyel göstermektedir; yine de sinir ağları ve sinyal işleme teknikleri üzerine daha fazla araştırma yapılması, daha geniş çapta uygulanabilir iyileştirmelere yol açabilir. Bu çalışma, yardımcı teknolojiler alanını geliştiriyor ve çeşitli ortamlarda daha etkili ve kapsayıcı insan-robot işbirliğinin kapısını açıyor (Karas, Pozzi, Pedrocchi, Braghin ve Roveda, 2023).

Pratik kullanılabilirlik, esneklik, kullanıcının kendi kendini yönetmesi ve tek modlu Beyin-Bilgisayar Arayüzü (BCI) şemalarına ilişkin kılavuzlarda dezavantajlar ele alınmaktadır. Bu belgede, sabit durum görsel uyarılmış potansiyelleri (SSVEP) ve sanal gerçeklik (VR) arka planı içinde göz takibini sağlayan asenkron bir robotik kol düzenleme çerçevesi için hem eşzamanlı hem de seri kontroller ileri sürülmektedir. Hedef kategorizasyonu için elektroensefalografi (EEG) ve göz bakışı gerçekleri eş zamanlı olarak bir araya getirilerek insanların üçlü göz kırpması ve göz kapama kullanarak sistemi kendiliğinden başlatıp durdurmasına olanak sağlanır. Sıralı yöntem, herhangi bir kalibrasyona ihtiyaç duymadan, öncelikli olarak bakış alanının göz bakışıyla tanınması, daha sonra EEG ile hedeflerin sınıflandırılması ve son olarak göz kapatılarak yanlış komutların reddedilmesi yoluyla çalışmaktadır. Sistemin verimliliğini destekleyen ve ümit verici ortalama doğruluk (ACC) ve bilgelik aktarımı sergileyen siber ve çevrimdışı testlere yirmi kişi katıldı. Anket içeren çevrimiçi bir test, kullanıcı dostu olma, uygun komut tanıma, azaltılmış yorgunluk ve sistem manevra kabiliyeti ile ilgili iyimser sonuçları ortaya koyuyor. Seri tarzına yönelik tercih puanlarının yükselişte olması dikkat çekicidir ve bu da onun değerini ortaya koymaktadır. Gelecekte değerlendirmeler, sistemin engelli bireylere yönelik etkinliğini ölçmeyi, kapsamlı SSVEP provokasyon yaklaşımlarını harmanlamayı, kontrol tekniklerini

artırılmış gerçeklik (AR) ortamlarına doğru ilerletmeyi ve sistemin daha zorlu görevler için yeteneklerini artırmayı kapsıyor. Araştırma, VR ile çalışan robotik kol kullanımlarında kullanıcı memnuniyeti, kullanılabilirlik ve özgürlüğün altını çizen karlı bir hibrit BCI (h-BCI) sistemi göstergesini barındırıyor (Guo, Lin, Luo, Gao ve Zhang, 2023).

Ciddi konuşma ve hareket bozukluğu (SSMI) olan bireyler için insan robotlarına yönelik bir arayüz oluşturma gibi zorlu bir görev bu belgede tartışılmaktadır. Amaç, özellikle tekstil üretiminde kullanılmak üzere, bakışla kontrol edilebilen bir robot kol üretmektir. İki kullanıcı çalışmasında erişilebilirlik testleri ve alma ve bırakma görevleri kullanılmıştır. Göz bakışı kontrollü arayüzün uygulanabilirliği ve verimliliği, sonuçta konuşma ve hareket engeli olan kullanıcıların örnek alma ve bırakma görevlerini ortalama 15 saniyeden kısa sürede tamamladığında ortaya çıktı. İkinci bir testte, katılımcılar keyfi olarak belirlenen bir hedefe bir dakikadan kısa sürede ulaştılar ve gelecekteki araştırmalar için öneriler, kullanıcıların geçmiş etkileşimlerine dayalı olarak hareket genliklerinin değiştirilmesinden oluşuyordu. Hem uygun yeteneğe sahip kişiler hem de SSMI kullanıcıları sistemi değerlendirdi ve sonuçlar cesaret vericiydi. Gelecekteki ilerlemeler arasında robotik kol için bir tutucu oluşturulması, teknolojinin tekstil boyama görevlerine dahil edilmesi ve sürücüsüz araçlar veya robotik tekerlekli sandalyeler gibi diğer siber-fiziksel sistemlerdeki uygulamaların araştırılması yer alacak. Önerilen göz bakışı kontrollü robotik kol, potansiyel olarak konuşma ve motor yetenekleri ciddi şekilde kısıtlı olan kişilere rehabilitasyon süreçleri boyunca yardımcı olabilir (Sharma, Saluja, Mollyn ve Biswas, 2020).

Göz bakışı izleme teknolojisi aniden bazı ilerlemeler kaydetti ve yenilikçi makine arayüzleri, sanal gerçeklikler (VR) ve insan-bilgisayar etkileşimleri (HCI) gibi çeşitli alanlarda yenilikçi uygulamaların kapısını araladı. Elektro-okülografi (EOG), Sklera Arama Bobinleri, kızılötesi okülografi (IOG) ve video okülografi (VOG) gibi teknikleri kullanan süreç, odak noktasını belirlemek için kullanıcı davranışını analiz etmeyi içerir. Göz bakışı takibinin masaüstü ortamlarına dahil edilmesi, çeşitli uygulamalar arasında oldukça fazla ilgi gördü. Bu çalışma, bakış tahmini, kafa duruşu tanıma, göz bölgesi tespiti, göz özelliği tespiti, göz vektörü kalibrasyonu ve göz bakış takibi modüllerini içeren yeni çıkmış bir bakış izleme sisteminin tasarımını ve uygulanmasını içermektedir. HCI kullanım sorunlarının üstesinden gelmenin yanı sıra, önerilen yaklaşımın yardımcı teknolojiler, e-öğrenme, psikoloji araştırması, pilot eğitim desteği ve sanal ve artırılmış

gerçekliklerde olası uygulamaları vardır. Bu tez, gerçek zamanlı video işleme ve baş hareketlerine bakış vektörü duyarlılığına odaklanarak göz bakışı izleme teknolojisine katkıda bulunmaktadır. Hedefleri, geniş bir kullanıcı becerileri yelpazesi için teknolojinin doğruluğunu, dayanıklılığını ve kullanılabilirliğini geliştirmektir (Sorate ve Chhajer, 2017).

Pek çok bilimsel uygulaması nedeniyle, yüz özelliklerinin gelecekteki tespiti ve tanınması, modern araştırmalarda giderek daha önemli hale gelmiştir. Bu makaleler, ökaryotik mesafelere ve kenar tespitine bağlı olarak göz tespiti için yeni bir yöntem sağladı. İşleme, iki renk uzayındaki formülleri kullanarak cilt bölgelerini tanımlar. Seçilen cilt bölgesinde Prewitt ve Sobel maskelerinin sonuçları birleştirilerek yatay sınırlar tespit edilir. Erozyon ve dilatasyon gibi morfolojik süreçler, küçük kenarları ortadan kaldırır ve yakındakileri birleştirir. Yüzlerin bölgesi üç bölüme ayrılmıştır: sağ üst, sol üst ve alt yarılar. Her bölümdeki büyük ikili öğeler ağızları, sağ gözleri ve sol gözleri temsil eder. Doğrulama için sağ ve sol gözlerin ağız merkezleri arasındaki mesafelerin hesaplanması gereklidir. Önerilen yaklaşımlar, PIC'in çeşitli niteliklere sahip 100 görüntüden oluşan resim veritabanı üzerinde gerçekleştirilen testlerde %80 doğrulukla kenar yoğunluğu tabanlı yöntemden daha iyi performans göstermektedir. Önerilen yöntemler %93 doğruluk göstermektedir. Bu yöntemin gri tonlamalı görüntü yönelimi, loş ışıklı ortamlarda zorluk sunsa da, çeşitli yüz görüntüsü senaryolarında sürekli olarak etkinlik ve dayanıklılık gösterir. Strateji, Sobel ve Prewitts yöntemlerinin morfolojik işlemlerle birleştirilmesiyle geliştirildi. Bu, gerçek zamanlı uygulamalar için uygun olan göz tespitinde optimal %93 doğrulukla sonuçlanır (Azar ve Khalilzadeh, 2016).

En son eklenen makine öğrenimi, traktörler, nesnelere tanıma, görüntü ayırma ve özellikleri algılama gibi çeşitli bilgisayarla görme sorunlarını çözmek için sıklıkla kullanılan önemli bir teknolojidir. Robocop'lar ve diğer karmaşık sistemler, görsel sensörleri barındırarak bilgisayarlı görme ikilemlerini çözmelerine ve çevrelerinin durumunu kavramalarına olanak tanıdı. Gelecekteki muhtemel faaliyetlerle ilgili alınan kararlar büyük ölçüde bu işlerden alınan cevaplara bağlıdır. Çağdaş makine öğrenimi yöntemlerinden biri olan Takviyeli Öğrenme, kullanıcıların çevreleriyle etkileşime girmesini sağlayarak öğrenmeyi basitleştirir. Robotik bilgisayar görüşünde nesne algılama, görsel izleme, eylem tanımlama ve robotların gezinmesiyle bağlantılı sorun giderme sorunlarını içeren en son uygulamalardır. Bu makaleler, Güçlendirilmiş Öğrenme teknolojilerini ve bunların bilgisayarlı görme ve robot navigasyon problemlerine nasıl

uygulandığını kısaca tanıtmaktadır. Görüntü içeriğinin en önemli anlaşılması, bilgisayarla görme, nesnelere tutma ve sahne tanıma, olay algılama ve nesne gruplamanın amaçlarıdır. Pratik uygulamalarda bu bilgi, kendi kendine çalışan sistemlerde yol tanıma veya mobil robot navigasyonu gibi belirli sorunların düzeltilmesinde ilk adım olarak hizmet eder. Çevreleyen ortamların anlaşılmasına dayalı olarak oluşturulan kararlar, görüntü anlama teknolojilerinin etkinliği için ölçü görevi görür. Makine öğrenimi teknikleri, özellikle takviyeli öğrenme, görüntülerden otomatik örüntü tanımaya yardımcı olduklarından makine görüşü için çok önemlidir (Bernstein, Burnaev ve Kachan, 2018).

"Endüstri 4.0" olarak adlandırılan üretim ve yapay zekanın (AI) birleşimi, çağdaş çağda operasyonel ortamı değiştirdi. Robotik kollar üretim otomasyonunun temelini oluşturur ancak birinci sınıf performans için verimli yapay zeka eğitimi gerektirirler. Bu araştırma, yapay zekayı eğitmek için bir prototip olarak "Elektronik İkizler" in avangart alanını derinlemesine inceliyor. Özellikle, modern Oyun Motoru olarak Unity kullanılarak oluşturulan sanal bir atmosferin, 3D baskılı robotik kol replikasyonu ile tasvir edilen fiziksel bir muadili ile sorunsuz entegrasyonunu inceliyor. Araştırma, sanal alanda sağlam bir simülasyon sağlamak için Tensor akışından ve titiz hiperparametre ince ayarından ustaca yararlanıyor. Robotik kolun öğrenme eğrisini hızlandıran, müfredatı dayalı bir eğitim tekniğinin kabul edilmesi dikkate değerdir. Araştırma, Soft-Actor Critic (SAC) kuralları ve ters kinematik gibi yapay zeka modelinin tekrarlayan evrimini destekleyen çok sayıda eğitim tekniğini ortaya koyuyor. En önemlisi, sanal ikizden gelen ağ üretimleri, g-kodu direktifleri kullanılarak fiziksel ikize aktarılır ve iki dünya arasında somut bir bağlantı sağlanır. Bu vaka çalışması, Elektronik İkiz paradigmasındaki fiziksel ve sanal bileşenler arasındaki karmaşık etkileşimler hakkında aydınlatıcı bilgilerin yanı sıra, başarılı simülasyon için eğitim prosedürleri ve mimari müzakereler hakkında önemli tavsiyeler vermektedir. Daha hassas izleme sistemleri ve işbirlikçi robot iletişiminin araştırılması, gelecekteki girişimlere yol açtığı için gelecekteki çalışmalar için uygun bir yol gibi görünüyor (Matulis ve Harvey, 2021).

Ek olarak, robotların güvenli hareketinin sağlanması, özellikle çeşitli olayların meydana gelebileceği yapılandırılmamış bağlamlarda, insan-robot işbirliği için çok önemli bir husustur. Bu araştırma, endüstriyel robotların çalışma alanlarındaki nesnelere ve insanların etrafında manevra yapabilmelerini sağlamak için makine öğrenimi tabanlı bir yöntem önermektedir. Geliştirilen robot kontrol sistemi, hızlı engel tespiti için yapay sinir

ağlarını, kümeleme analizini ve yol planlama için en yakın komşu yöntemini içermektedir. Makalenin sonuç kısmı, insanın korunmasına yönelik uyarlanabilir yol planlamasını gösteren bir kavram kanıtını göstermektedir. Çalışmada insan-robot etkileşimine yönelik deneysel bir simülasyon platformunun (HIRIT) oluşturulması da anlatılıyor. Her ne kadar uyarlanabilir yol planlaması başarılı bir şekilde gösterilmiş olsa da, bazen yanlış kemik tespiti hatalı yol planlamasıyla sonuçlanıyordu. Bunu çözmek için simülasyon ortamında nesne tespiti ve yerelleştirmeye yönelik eğitilmiş bir Yapay Sinir Ağı (YSA) ve görüntü işleme program dizisi uygulamaya konmuştur. Ancak, özellikle işlem hızının ve tanıma doğruluğunun iyileştirilmesiyle ilgili hâlâ sorunlar var. Baskı altında gerçek zamanlı testleri kolaylaştırmak için yazarlar, veri işleme ve iletişim arayüzlerinde yanıt sürelerini en aza indirmenin önemini vurguluyorlar. Sürekli odak noktası, endüstriyel robot sistemlerinin ve insan-robot işbirliğinin uyarlanabilirliğini geliştirmek için makine öğrenimi algoritmalarının, özellikle de YSA'ların kullanılmasıdır. Bu algoritmaların üretimle ilgili uygulamalara pratik olarak yerleştirilmesi nihai amaçtır (Dröder, Bobka, German, Gabriel ve Dietriz, 2018).

İnsanlarla iletişimin zor olduğu veya düşünülemez olduğu çeşitli insanlık dışı koşullar nedeniyle mekanik kollara olan ihtiyaç artıyor. Bir bombayı dağıtmayı veya lav fırlatan çalışan bir kaynaktan okuma almayı içerebilirler. Burada ivmeölçerlerden elde edilen verilerle tipik insan kol gelişimiyle sınırlı bir mekanik kol inşa edilmesini öneriyoruz. İvmeölçerin verimini düzeltmek, uygun kontrol cihazını seçmek ve sensörlerden gelen gürültü miktarını azaltmak için uygun ortalama hesaplama kullanılır. Bu kolun ilerlemesi için Arduino ATmega328 gereklidir ve hepsi sıralı iletişim yoluyla birbirleriyle arayüz oluşturacaktır. Bu otomatik kol yapısı, 150 grama kadar ağırlık toplayabilir ve ileri ve geri dönüş, kolun açık ve kapalı olması ve 90 ve 180 derecelik eğimler dahil olmak üzere altı konuma sahiptir. 433 Hz'de çalışır. Son olarak, sorunu çözmek için gereken sonuçlar arasında yapılandırılmış kol modeli fotoğrafları ve tehlikeli veya tehlikesiz malların müşteriden uzağa yerleştirilmesi veya seçilmesi gibi ayrıntıların yer aldığı bir tablo yer alır. Bir kişinin koluna esnek ve jiroskop sensörüyle takılan bir Arduino ATmega328, bir parmağın düştüğünü veya kişinin kolunun döndüğünü algılayarak bilgiyi alıcı Arduino'ya iletir; o anda işaret oluşuyor ve 433 Hz'de çalışan ve 150 gram ağırlığı hareket ettirebilen robot kola takılı servo motoru ayarlıyor. Şekil 8'de gösterildiği gibi robotik kolun takma, sol, açık, kapalı ve 90 ve 180 derecelik eğimler dahil olmak üzere altı ayrı konumu vardır. Esnek ve jiroskop sensörüyle kontrol edilen mekanik bir kolun montajını ve

programlanmasını içeren bu çalışmanın hedeflerine ulaşıldı. Oluşan izlenim, gelişiminin kesin, kesin ve idare edilmesi ve kullanılması kolay olduğunu göstermektedir. Makale uzaktan aktarıma dayanmaktadır; İşareti 9 ila 10 metre içinde iletmek ve almak için küçük ölçekli ağırlık kaldırmak için tasarlanmış otomatik bir kol kullanıyoruz. Ayrıca altı aktivite gerçekleştiriyoruz: kolu sola ve sağa hareket ettirmek, parmakları açıp kapamak ve eli oraya buraya hareket ettirmek (Anughna, Ranjitha ve Tanuja, 2020).

Teknoloji, insanın sürekli artan ihtiyaçlarına ayak uydurabilmek için günümüzde de aynı doğrultuda ilerlemektedir. Bu ihtiyaçlara yönelik yapılan çalışmalarla hayat her geçen gün daha ulaşılabilir hale geliyor; bu çalışmalar robotik kol araştırmalarına odaklanıyor. Robotik kollar harici bir operatör tarafından veya önceden ayarlanmış komutlar takip edilerek çalıştırılabilir. Günümüzde endüstriyel ve medikal sektörler en gelişmiş robotik kol teknolojisine sahiptir. Proje için tasarlanıp hayata geçirilen robot kol, beş adet servo motor yardımıyla dört yönde hareket edebilmektedir. Tutucu, aldığı malzemeyi bir yerden başka bir yere taşımak için gerekli maddeyle karıştırmanıza olanak tanır. Robotik kollar için pek çok alan geliştirilebilir. Robotik kollar sayesinde birçok görev daha erişilebilir hale getirildi ve hata seviyesi önemli ölçüde azaldı. Örneğin, planlı bir robot kol ve eczanelere konuşlandırılan birkaç ilaç dağıtım robotu inşa edildi. Ek olarak robot kolunun hareket aralığı daha da artırıldı ve çok yönlülüğü, kamerayı parmak bölgesine yerleştirip hassasiyetini artırarak otomasyon sistemleri ve tıbbi alan dahil olmak üzere çeşitli ortamlarda kullanılmasına olanak tanıyor. Robotik kolların geliştirilmesine yönelik bu yaklaşım, tıp alanında hastalar arasında enfeksiyon olasılığını azaltır ve aynı zamanda cerrahi prosedürler sırasında insan hatası olasılığını da azaltır. Bu makalede, mobil bir uygulama aracılığıyla uzaktan çalıştırılabilen Arduino tabanlı bir robotik kolun modeli, tasarımı ve yapısı gösterilmektedir. Bu çalışma için altı serbestlik derecesine sahip bir robot kol geliştirilmiş ve kullanılmıştır. Kullanıcının mobil uygulaması, Bluetooth kablosuz kontrol sinyalleri aracılığıyla Arduino platformu kontrollü tasarıma komutlar verir. Servo motor, koldaki beş döner eklemden oluşan uç efektöre dönme hareketi sağlar. Her bağlantıyı üretmek için bir 3D yazıcı kullanılmadan önce Solid Works kullanıldı. Son kol prototipi, robotun bileşenlerinin ve motorun mekanik şekillerinin bir araya getirilmesiyle oluşturulur. Robotlar ve otonom sistemler dünya çapında birçok fonksiyonel süreçte kullanılıyor. Robotik kol, özellikle mesafe kontrollü özelliği sayesinde endüstriyel alanlardaki birçok insan kısıtlamasının üstesinden gelebilmektedir. Robotik kolların operasyonları ve fonksiyonları, prosedürleri otomatikleştirmek ve insan hatası oranlarını

azaltmak için işletmelerde ve araştırma laboratuvarlarında yaygın olarak kullanılmaktadır. Robotik kollar, montaj hattı çalışmaları ve kuvvet kontrolü ve kontrolöre giriş gerektiren hareketler dahil olmak üzere çeşitli faaliyetleri gerçekleştirebilir. Robotik sistemler, tüm endüstriyel sektörlerde genişleyen uygulanabilirliği ve hataları ve malzeme israfını azaltma kapasiteleri nedeniyle giderek daha popüler hale geldi. Müdahaleci teklifler için çok sayıda robotik sistem geliştirilmiştir (Ali, Hashim ve Al-Sakkal, 2022).

Yapay silahlar, insan bağımlı kurmanın ve korumanın en iyi ihtimalle zorlu olduğu çeşitli insanlık dışı durumlarda giderek daha fazla gerekli hale geliyor. Bu tür faaliyetler arasında örneğin bir bombayı etkisiz hale getirmek veya aktif bir yanardağdan ölçümler almak yer alıyor. Bu çalışmada, insan kolunun doğal hareketlerini taklit eden bir robotik kol oluşturmak için ivmeölçerler veya insan gözlemciler kullanmayı öneriyoruz. Bu kolun yapımında ATmega32 ve ATmega640 platformları, bir Arduino UNO veya MEGA kartı ve sinyal işleme için bir kişisel bilgisayar kullanıldı. Bu cihazlar seri bağlantı yoluyla bağlanacaktır. Son olarak bu prototip kol, güvenli bir şekilde yerleştirilemeyecek kadar uzaktaki nesnelere ulaşabiliyor. Bu proje robotik silahların geçmişini inceliyor. Proje, bir robot kolunu kontrol etmek için donanım ve yazılımı başarıyla oluşturdu. Temel amacımız nesnelere tutabilen robotik kol geliştirmek. Kapsamlı çalışmalara rağmen robotik kollar öncelikli olarak, asıl işlevi üretkenliği artırmak olan imalatla kullanılıyor. Bu uzuvlar artık olgun bir aşamaya ulaştıkları için hassas bir şekilde hareket edebilirler. Robotik kolların evsel ve evrensel birçok kullanım alanından daha fazla yararlanması gerekiyor. Robotik kolları yerleştirildiği her yerde bir "yardım eli" olarak kullanmak mümkün. İlk testlere göre cihazın hareketi ayarlanabilir, doğru, hassas ve keyiflidir (Bhadouria, Bhadouria, Patel ve Upasani, 2023).

Leap Motion ve Myo tarafından sağlanan el reaksiyon fonksiyonlarını inceledik ve karşılaştırdık. Süreç, her parmağa önceden belirlenmiş bir açı verilmesini ve sistemin tepkisinin izlenmesini içeriyordu. Kontrol programı kullanılarak pozisyonlar doğrudan parmaklarda kalibre edildi. Uzatılmış parmak duruşundan çeşitli dönme açıları ortaya çıkarıldı. Bu hareketleri yaparak parmakların, hareket tanıma ekipmanı tarafından girilen komutları nasıl kopyaladığını görebildik. Benzer açılarla iki hareket yakalama sistemi (Leap Motion ve Myo) kullanıldı. Çeşitli hareket yakalama teknolojilerini kullanarak el hareketlerini taklit edebilen 3 boyutlu baskılı bir robotik kol oluşturmak başarılı oldu. Projeye en uygun olanların seçilebilmesi için belirli hedeflere yönelik çeşitli "mocap"

cihazları kullanılarak bir çalışma yürütüldü. Robotik kol, insan elinin hareketlerini taklit edebildi. Kol, 22. MMVR (Tıp Sanal Gerçeklikle Buluşuyor) Los Angeles 2016 ve Oslo İnovasyon Haftası'nda sergilendi. Bu olaylar, prototipin uygulanabilir ve uygun fiyatlı bir çözüm olarak çekiciliğini ve potansiyelini ortaya koydu (Alpiste, Torner ve Brigos, diğerleri, 2017).

"Derin Öğrenmeyi Kullanarak Gerçek Zamanlı İris Yer İşaretlerinin Tespiti için Sağlam Bir Çerçeve" çalışması, etkili ve doğru bir iris yer işareti tespit sisteminin geliştirilmesine yönelik kapsamlı bir araştırma sunmaktadır. Araştırma, iris yer işaretlerinin yeni, açıklanmalı bir veri kümesini toplayarak ve iris yer işaretlerinin sağlam ve gerçek zamanlı tanımlanması için derin öğrenmeye dayalı bir model önererek alana önemli ölçüde katkıda bulunuyor. Yazarlar, çeşitli Evrişimli Sinir Ağı (CNN) modellerinin etkinliğini göstermekte ve iris yer işaretlerinin yerleştirilmesinin otomatik, hassas ve hızlı ölçümü için önerilen MobileNet v2 model tabanlı çerçevenin sağlamlığını doğrulamaktadır. Sonuçlar, Ortalama Mutlak Hata, model boyutu, yanıt süresi, doğrulama kaybı ve Birleşim Üzerindeki Ortalama Kesişme gibi olağanüstü performans parametrelerini sergileyerek çerçevenin sağlamlığını ve doğruluğunu vurguluyor. Çalışmanın bulguları, iris tabanlı imleç hareketleri yoluyla uzuv amputasyonu olan bireylerin desteklenmesinde potansiyel uygulamalarla birlikte, hassas iris lokalizasyonu için değerli bir kıyaslama veri seti sağlıyor. Sonuç olarak, önerilen çerçeve, kullanıcı kimlik doğrulaması, güvenlik uygulamaları ve engelli bireylere yönelik yardımcı teknolojiler için çıkarımlarla, gerçek zamanlı iris yer işaretlerinin tespiti için umut verici bir çözüm sunmaktadır (Kumar, Bajpai, Sinha ve Singh, 2023).

Makale, özellikle Endüstri 4.0 bağlamında verimli insan-makine etkileşimlerinin geliştirilmesinde el hareketi tanımanın önemine dair ilgi çekici bir örnek sunuyor. MediaPipe çerçevesini kullanarak el hareketi tanımanın uygulanması yoluyla çalışma, teknoloji endüstrisinde yaygın uygulama için umut verici bir potansiyel ortaya koyuyor. Sonuçlar, el hareketlerini tanımada %95'lik etkileyici bir doğruluk performansı sergiliyor ve bu yaklaşımın etkililiğini vurguluyor. Ayrıca yazarlar, diğer cihazlarla entegre olarak ve statik ve dinamik el hareketi tanıma sistemlerini keşfederek sistemin yeteneklerini genişletmeyi amaçlamaktadır. Bu araştırma, el hareketi tanımanın insan-bilgisayar etkileşimlerini geliştirmedeki önemli rolünü vurguluyor ve bu alanda daha fazla yenilik ve entegrasyon potansiyelinin altını çiziyor (Indriani, Harris ve Agoes, 2021).

Bu makale, ofis işlerinde stresi ve yorgunluğu azaltmaya yardımcı olabilecek, 3 boyutlu yer işaretlerini ve standart bir web kamerasını kullanarak göz kırpmalarını tespit etmeye yönelik yeni bir yöntem sunmaktadır. Algoritma, göz kırpma olasılığını belirlemek için göreceli bir eşik ve göz hareketi gradyanları kullanıyor ve bu da %90'ın üzerinde doğruluğa sahip güçlü bir algılama sistemi sağlıyor. Makalede ayrıca iş sağlığı açısından göz kırpma takibinin önemi ve bu teknolojinin gelecekteki potansiyel uygulamaları tartışılıyor. Yazarlar önerilen yöntemin performansını iki farklı veri kümesi üzerinde değerlendirdiler ve doğruluğunu belirlemek için F1 skorunu, kesinliği, hatırlamayı ve doğruluğu kullandılar. Sonuçlar, önerilen yöntemin göz kırpmalarını etkili bir şekilde algıladığını ve mevcut bilgisayar iş istasyonlarına kolayca entegre edilebileceğini gösterdi. Yazarlar, göz kırpma izlemenin daha ileri müdahale konseptleri için temel bir gereklilik olduğunu ve gelecekte kapsamlı sağlık desteğine olanak sağlayacağını ileri sürmektedir. Genel olarak bu makale, iş sağlığı yönetimine değerli bir katkı sağlamakta ve işyeri refahını iyileştirmek için 3 boyutlu işaretler kullanan kamera tabanlı göz kırpma algılamanın potansiyelini vurgulamaktadır (Kraft, Hartmann ve Bieber, 2022).

Google Araştırma tarafından geliştirilen MediaPipe, algılama hattı çerçevelerinde önemli bir ilerlemeyi temsil ediyor. "MediaPipe: Algılama İşlem Hatları Oluşturmak için Bir Çerçeve" makalesi, MediaPipe'in yeteneklerine ve uygulama geliştirme üzerindeki potansiyel etkisine kapsamlı bir genel bakış sunmaktadır. Çerçeve, bireysel modelleri soyutlayıp sürdürülebilir boru hatlarına bağlayarak algı modellerini çeşitli cihazlarda verimli bir şekilde çalıştırmanın zorluklarını ele alıyor. Bu, geliştiricilerin algoritma ve model geliştirmeye odaklanmasına olanak tanırken MediaPipe, duyuşal verilerden çıkarım yapmak ve algılanan sonuçları üretmek için gerekli adımları gerçekleştirir. Ayrıca MediaPipe'in GPU hızlandırma desteği ve farklı platformlarda tutarlı bir şekilde çalışabilme yeteneği, onu uygulama oluşturma ve dağıtma konusunda paha biçilmez bir araç haline getiriyor. Çerçevenin Google'da son altı yıldaki başarısı, etkinliğinin altını çizmektedir ve açık kaynak sürümü, topluluk desteği ve önerilen hesap makineleri ve grafiklerin geliştirilmesi yoluyla ekosistemini daha da geliştirmeye hazırdır (Lugaresi ve diğerleri, 2019).

MediaPipe ve OpenCV kullanılarak göz iris takibi ve yön tespiti, OpenCV ve MediaPipe kullanılarak bir göz izleme sisteminin geliştirilmesini araştırıyor. Sistem, tüm veri işlemlerinin akıllı bir kamerada gerçekleştirildiği, hem görünür ışık hem de karanlık

koşullarında göz irisini tespit edecek şekilde tasarlanmıştır. Yazarlar, iris koordinat tanımlaması için Circle Hough Transform'un kullanımının altını çiziyor ve sistemin çeşitli aydınlatma koşullarında göz irislerini tespit etme yeteneğini vurguluyor. Makale aynı zamanda yazarların dahili rehberinin ve meslektaşlarının katkılarına da teşekkür etmekte ve göz izleme teknolojisi alanındaki ilgili çalışmalara referanslar içermektedir (Doijad, Bhalerao ve Khan, 2022).

Sorun açıklaması

Omurilik yaralanmaları, nörolojik bozukluklar veya diğer sağlık durumlarından kaynaklanan felç, bireyin günlük görevleri yerine getirme yeteneğini önemli ölçüde sınırlar. Geleneksel yardımcı cihazların değerli olmasına rağmen çoğu zaman kullanıcılara kesintisiz ve sezgisel bir arayüz sağlaması gerekir. Felcin incelikli doğası ve etkilenen bireylerin farklı ihtiyaçları göz önüne alındığında, daha uyarlanabilir ve duyarlı bir teknolojik çözüme yönelik zorlayıcı bir ihtiyaç ortaya çıkmaktadır.

Araştırma soruları

Mevcut teknoloji araştırmaları, robotik kolu kontrol etmek için göz hareketini uygularken çeşitli teknik zorlukların altını çiziyor.

- Doğruluk ve Güvenilirlik
- Gecikme azaltma
- Kullanıcı dostu
- Robotik Kol dinamikleri

Mevcut teknoloji araştırmacılarının, robotik kol kontrol kontrolüyle birlikte hesaba kattığı dört zorluk bulunmaktadır: Tepki açısından sistem etkinliğini sağlayan kullanıcı dostu bir sistem oluşturmak - Robotik kol dinamiklerini, robot kollarını kontrol etmek için en son yazılım üzerinde çalışarak geliştirmek. Robotik kolun dinamikleri ve hareketleri sorunsuz bir şekilde sağlanır.

Çalışmanın önemi

Bu araştırmanın toplumsal etkisi teknolojik ilerlemenin ötesine uzanıyor. Felçli bireylerin kontrolü ve bağımsızlığı yeniden kazanmalarını sağlayan önerilen sistem, kapsayıcılık ve haysiyetin temel ilkeleriyle uyumludur. Ayrıca, kullanıcı aracılığını ve refahını ön planda tutan sorumlu inovasyon ihtiyacını vurgulayarak, yapay zekanın sağlık hizmetlerinde etik olarak konuşlandırılmasına ilişkin süregelen söylemlere katkıda bulunuyor.

Sonraki bölümleri incelerken okuyucu, araştırmacıya bu iddialı çabanın metodolojisi, uygulaması ve sonuçları boyunca yapacağı yolculukta eşlik edecek. Teknolojik yaratıcılık ve empatik tasarımın sentezi yoluyla bu tez, felç gibi karmaşık bir alanda yol alan kişilerin yaşamlarını iyileştirmede anlamlı adımlar atmaya amaçlıyor.

Tezin organizasyonu

6 bölümden oluşan tez;

- Bölüm 1'de araştırma hakkında bir giriş yapılır ve arka plan bilgisi, problem bildirimi, araştırma sorusu ve önemi özetlenir.
- Bölüm 2'de önerilen sistem ve boşluk hakkında gerekli arka plan açıklanmakta ve bazı örnekler verilmektedir.
- Bölüm 3, medya borusu ve derin öğrenme modellerini kullanan göz algılama yöntemlerini açıklamaktadır.
- Bölüm 4'te projede kullanılan nesne algılama parçaları ve veri kümesi sonuçları açıklanmaktadır.
- Bölüm 5'te Robotik parça ve onun uygulanması, hareketin kinetiği ve ayrıca kullanılan ortam, Donanım bileşeni, Kullanıcı Arayüzü açıklanmaktadır.
- Bölüm 6, tezin tüm göz algılama ve göz kırpma sayma, nesne algılama ve robotik seçme ve yerleştirme görevinin görevi gerçekleştirmek için entegre edildiği Entegrasyon ve sonuç kısmını açıklamaktadır.



2. LİTERATÜR İNCELEMESİ

2.1. Literatür Taraması

Bu bölümde göz kırpmaya algılama ve robotik kumanda kolu ile ilgili tüm çalışmaları ele alacağız.

Shafti, Orlov ve Faisal (2019)'ın araştırması, gerçek zamanlı niyet kodunun çözülmesini ve kullanıcı etkileşimini mümkün kılmak için bakış izlemeyi robotik sistemlerle birleştirerek yardımcı robotiğe yeni bir yaklaşım sunuyor. Robot İşletim Sisteminde yerleşik olarak bulunan RGB-D kameraları, göz izleme gözlüklerini ve nesne tanımlamayı kullanan çok modlu sistem, kullanıcı tarafından girilen komutları öngörerek 3 boyutlu bakış modellerini izler. Değerlendirme görevleri, 3D bakış entegrasyonu aracılığıyla etkinliklere ulaşmada %100 başarı oranı gösterdi ancak aynı zamanda mekanik tasarımda alınan kararlarla bağlantılı dökme görevlerinde sorunlar olduğunu da ortaya çıkardı. Yazarlar çalışmanın sınırlamalarını kabul ediyor ve sistemin performansını ve kullanılabilirliğini geliştirmek için gelecekteki çabalar öneriyor. Sonuç olarak bu çalışma, yardımcı robotikte insan-robot etkileşimi için çeşitli robotik sistemlerde kullanılacak umut verici bir yöntem sunuyor. Shafti ve diğerleri (2019)'ne göre gelecekteki çabalar, kısıtlamaların ele alınmasına ve kullanıcı takibi ve çevreyi kavramaya yönelik RGB-D kameranın yeni uygulamalarının araştırılmasına odaklanacak.

Allen, Timcenko, Yoshimi ve Michelman (1993), ufuk açıcı çalışmalarında robotik manipülasyon ve görüşe dayalı izlemenin temel ilkelerini araştırdılar. Gözlemlenen koordinatlardaki gürültünün etkilerini dikkate alan ve yükseklik farkı, yay uzunluğu ve yörünge bükülmesi gibi karmaşık özellikleri araştıran yörünge parametrelendirmesine yeni bir yaklaşım sunuyorlar. Çalışma, görüşe dayalı takip ve kavramayı inceliyor ve dinamik çevrenin sunduğu zorlukları vurguluyor. Aynı zamanda ölçüm gürültüsünü ve belirsizliği ele almak için sağlam sistemlere duyulan ihtiyacı da vurgulamaktadır. Önerilen çözüm, düzlemsel ve 3 boyutlu uzay eğrisi hareketlerini izlemeye yönelik olasılıksal yaklaşımların entegrasyonu ile karmaşık bir katman kazanıyor. Yazarlar, görüş tabanlı robotik alanındaki ilgili çalışmalar bağlamında yörünge inşası, robot görüşü, optik akış tespiti, gerçek zamanlı görüş geri bildirim ve erişim ile kavramanın zamansal etkileşimi üzerine araştırmaları kapsamaktadır. Yazarlar robotik, bilgisayar bilimi ve akıllı sistemlerdeki engin bilgilerini

kullanarak, gerçek zamanlı bilgisayarlı görme ve model tabanlı sensör planlamasının araştırılmasına önemli ölçüde katkıda bulunuyorlar. Literatür taraması ayrıca, incelenen araştırma alanındaki temel fikirlerin, yaklaşımların ve başarıların kapsamlı bir özetini sunarak, alanda önde gelen araştırmacıların daha önceki çalışmalarını da onurlandırmaktadır.

Paperno, Rupp, Maboudou-Tchao, Smither ve Behal (2016), fiziksel veya duyuşal sınırlamaları olan kişiler için robotik cihazlarla kullanıcı etkileşimini geliştirme ihtiyacını ele almaktadır. El becerisi, mekansal yetenekler ve çalışma belleği de dahil olmak üzere on temel insan faktörünü belirleyerek çalışma, seçme ve yerleştirme görevleri sırasında kullanıcı performansını tahmin etmeyi amaçlıyor. Sonuçlar, bilgi işleme hızının, uzamsal yeteneğin, el becerisinin ve çalışma belleğinin görev performansını önemli ölçüde etkilediğini göstermektedir. Çalışmada düşük hata oranlarıyla doğru tahminlere ulaşmak için doğrusal ve polinom modelleri kullanıldı. Genel popülasyonu dikkate almanın önemini vurgulayan araştırma, bireysel farklılıkların yardımcı teknolojilerle kullanıcı performansı üzerindeki etkisini vurguluyor. Araştırmacılar, belirli insan faktörlerini robotik manipülatörlerin tasarımına entegre ederek, farklı bir kullanıcı tabanı için kullanılabilirliği ve etkinliği artırabilir. Genel olarak bu çalışma, fiziksel ve duyuşal sınırlamaları olan bireyler için yardımcı teknolojileri optimize etme seçeneği konusunda değerli bilgiler sunarak daha kullanıcı merkezli ve verimli robotik sistemlerin önünü açıyor.

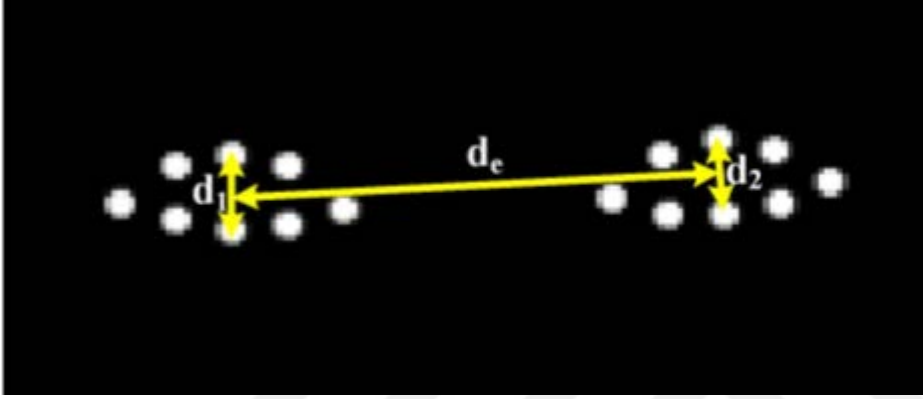
Qul'am, Oktarina, Dewi, Permatasari ve Risma (2019) tarafından yapılan çalışma, görüş güdümlü bir seç ve yerleştir robot sisteminde kenar algılama tekniklerinin kullanımını araştırıyor. Politeknik Negeri Sriwijaya, Endonezya, araştırma robotik teknolojisinin başta tarım olmak üzere çeşitli sektörlerdeki önemine odaklanıyor. Araştırmacılar, Sobel, Prewitt ve Canny gibi algoritmaları kullanarak gerçek zamanlı senaryolarda nesne algılama ve manipülasyon yeteneklerini geliştirmeyi amaçlıyor. Deney düzeneği, hassas hareket kontrolü için motorlu servolara sahip bir robot ve görüş rehberliği için göz-ele kamerayı içeriyor; sistemin domates, limon ve patlıcan gibi nesnelere ayırt etme ve kullanma konusundaki etkinliğini sergiliyor. Makale ayrıca tarımsal robotiklerle ilgili çalışmalara da atıfta bulunarak, tarımsal ortamlarda meyve hasadı, izleme ve haritalama gibi görevlerin iyileştirilmesinde kenar algılamanın potansiyelini vurguluyor. Genel olarak çalışma, hedef tanıma ve nesne manipülasyonu ile ilgili görevlerde görüş güdümlü robotik sistemlerin performansını ve verimliliğini artırmak için kenar algılama uygulamasına ilişkin değerli bilgiler sağlıyor.

Moriyama, Kanade, Cohn, Xiao, Ambadar, Gao ve Imamura (2002), hareketleri ve görünüm bilgilerini takip ederek göz kırpmaları tespit edebilen bir algoritma oluşturan ilk kişiler listesine girmiştir. Göz kırpmalarını güvenilir bir şekilde algılayan yöntemlerin geliştirilmesinde ortaya çıkan bazı zorlukları basıyorlar. "kafanın sert hareketi, önden olmayan duruş, kafa hareket gözlüklerinin tıkanması ve jestler, konuşma, yumuşak hareket birimleri, hızlı yüz ifadesi" gibi zorlukların sorunun daha da büyümesine neden olan unsurlar olduğunu öğreniyorlar ve tüm bunları yapıyorlar. Bir "Göz Kırpma Dedektörü" geliştirilirken göz önünde bulundurulması gerekenler, Wang, Ding, Fang, Liu ve Wang (2009), görüntülerin kalitesi veya parlaklığındaki farklılıklar gibi teknik istikrarsızlıkları da içeren ve aynı zamanda karşılaşılan zorluklar olan aynı zorlukları vurgulamaktadır. Moriyama ve diğerleri (2002), göz kırpmaları otomatik olarak tespit etmeye çalışırken, yöntemlerini Ekman ve Freisens'in, ilgili kaslara bakarak insan yüzünün çeşitli hareketlerini kategorize eden Yüz hareketi kodlama sistemi olarak bilinen 1978'e dayandırmaktadır. Bu hareketlerde sistem, eylemi (birimleri) AU'yu söz konusu kasın hareketi olarak tanımlar: kasılması veya gevşemesi.

Lalonde, Byrns, Gagnon, Teasdale ve Laurendeau (2007)'de ayrıca FACS'ı ve bunun kendi çalışmaları bağlamında uygunluğunu tartıştılar. Lalonde ve diğerleri (2007), bu yöntemin önceki yöntemlere göre daha stabil olduğunu iddia etmek için düşük kontrastlı görüntülerde gözleri tespit edip takip etmiştir. Çalışmalarında öncelikle bölgesel ilgiyi kullanarak katılımcıların gözlerini tespit edip buluyorlar. Aynı zamanda gözleri takip etmeyi ve katılımcının kameraya önden dönük olması gereken ilk karedeki baş pozisyonu ile sonraki kareler arasındaki dönüşümü tahmin etmeyi de sağlıyorlar. Göz kırpmaları tespit etmek için, hareketleri tespit eden ve bir formun ilgi alanı içindeki, diğeri veya sonraki form arasındaki farkları hafifleten bir algoritma kullanıyorlar.

Wang ve diğerleri (2009), göz kırpma tespitinde kullanılan iki ana yöntem türünü tanımlamaktadır; bilinen kontur şablonu tabanlı yöntemler ve görünüm tabanlı yöntemler. İlki, gözün şekline göre modelini oluşturur. Göz modeli moda girdikten sonra bir sonraki adım, gözlerin görüntüsüne uygun bir şablon oluşturmaktır. Bu yöntemlerin dezavantajının, hem açık hem de kapalı göz şablonlarının ayrı ayrı işlenmesi ihtiyacını ve parlaklık değişimine karşı hassasiyeti içerdiğini belirtiyorlar. İkinci yöntemde ikili sınıflandırma gözün açık mı yoksa kapalı mı olduğunu belirler. Bu tür bir yöntem büyük miktarda veri gerektirir. Bu yöntemin bir diğer dezavantajı ise göz sınıflandırıcının göz

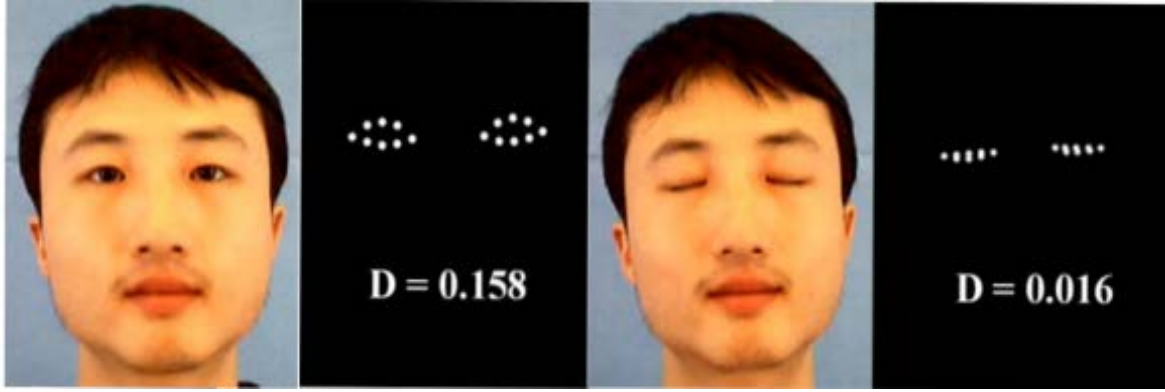
çevresini çıkarmamasıdır. Wang ve diğerleri (2009), her iki yöntemi de birleştirmeyi seçmiştir. Bu yöntemde, başlangıçta gözlerin konumunu tespit ederek başlıyorlar ve gözlerin dairesel bölgesine yakın 16 işaret noktası belirleyerek dış hatlarını çiziyorlar. Bu işaretler gözün açık mı kapalı mı olduğuna dair bilgi sağlar. İkinci adımda, yer işaretlerini tanımak üzere eğitilmiş bir sınıflandırıcı kullanırlar. Kameraya bakan, gözleri açık ve kapalı bir kişinin görüntüsünü kullanıyorlar. Gözlerin ne kadar açık ve kapalı olduğunu hesaplamayı içeren bir göz kırpma tahmin prosedürü geliştirdiler.



Şekil 2.1. Wang ve diğerlerinin (2009) noktalarından göz açıklığı derecesi

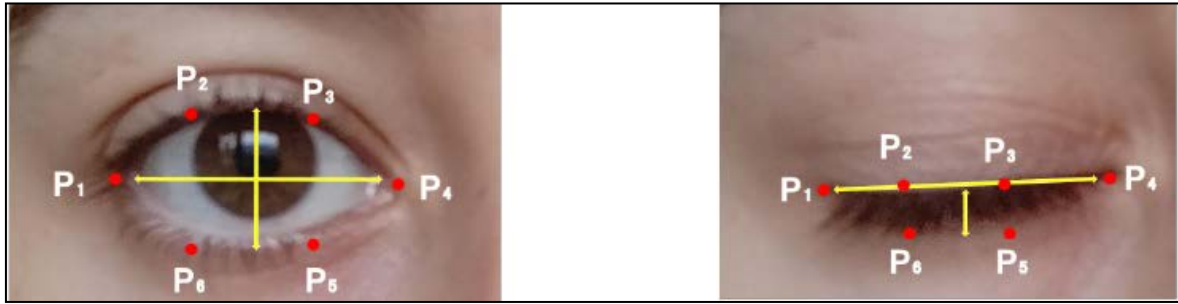
$$D_L = \frac{d_1}{d_e}, D_R = \frac{d_2}{d} \quad (2.1)$$

Choi, Han ve Kim (2011), ayrıca gözleri tespit etmeden önce görüntüler üzerinde bir insanın yüzünü tespit etti ve zamanla sonuçları cilaladı. Çalışmalarında katılımcıların yüzlerinin tespiti için AdaBoost Yüz dedektörünü, gözleri tespit etmek için ise Ad boost göz dedektörünü kullandılar . AdaBoost dedektör işlemi görüntüyü tamamlar ve sınıflandırıcılar topluluğu üzerine inşa edilmiş kademeli bir sınıflandırıcıyı kullanır. Bir sınıflandırıcının çıkardığı veriler diğerine verilir. Viola ve Jones (2001), kademeli sınıflandırıcı adı verilen ilk kademeli sınıflandırıcıyı yarattı. Choi ve diğerleri (2011) ayrıca Modifiye Sayım Transformatörü (MCT) özelliklerini Adaboost Ace ve Adaboost yüz dedektörü göz dedektörlerine girdi olarak kullanıyor. Daha sonra dedektörlerde yapılan çalışmalarda çoklu görüntü piramitleri kullanıldı. İkili sınıflandırma içeren MCT özellikleriyle Adaboost öğrenimine dayalı birleşik göz kırpma algılama algoritmasını kullandılar. Gözün açık veya kapalı olduğunu sınıflandırabilirler.



Şekil 2.2. Göz kapalı veya açıkken gözlerin açıklık derecesi Wang ve diğerlerinin (2009)

Soukupová ve Čech (2016), çalışmalarında kişinin gözlerinin ne kadar açık olduğunu ölçmek için Göz En-boy Oranı (EAR) olarak bilinen yeni bir yöntem sunmaktadır. Çalışmalarında bu oranı, bir kişinin gözlerinin etrafındaki farklı noktalar arasındaki mesafelere bakarak göz açıklığını tahmin etmenin bir yolu olarak yorumladılar. EAR'ı hesaplamak için üst ve alt göz kapağı noktaları arasındaki mesafeyi ve katılımcının göz köşeleri arasındaki mesafeyi ölçerler. Mesafe her iki göz için hesaplanır. Ayrıca, kişi göz kırptığında bunun genellikle her iki gözde aynı anda meydana geldiğini de buldular. Yani her iki gözden gelen kulağın ortalaması dikkate alındığında,



Şekil 2.3. Göz açık ve kapalıyken gözdeki yer işaretleri

Doğruluğu artırmak için, bir bilgisayar programını, göz kırpmaları analiz ederken tek seferde bir video karesi yerine birden fazla video karesine bakacak şekilde eğittiler. Göz kırpmasının daha güvenilir bir şekilde algılanmasına yardımcı olur. Ayrıca, katılımcıların baş hareketlerinin göz kırpmaları algılamada hatalara neden olduğu önceki çalışmalardan kaynaklanan bir sorunu da tartıştılar. Yatay kafa hareketlerinin kulak ölçümüne müdahale etmediğini buldular. Son olarak, göz kırpmaları algılamaya yönelik EAR ölçümleri için

eşik deęerini 0,2 olarak ayarladıktan sonra, bunların daha doęru ve daha doęru alıřtıklarını buldular.

Meena, Chowdhury, Cecotti, Wong-Lin, Dutta ve Prasad (2017)'in EMOHEX cihazı hakkında konuřmak, teknoloji alanında byk bir devrimdir ve kol ve bacaklarda engelli insanlara yardım etmek iin kullanılır. Bu cihaz, ucuz bir gz izleme cihazı ve tekerlekli sandalyeye baęlanan bir el dıř iskeleti ieriyor. Kullanıcıların herhangi bir zorluk yařamadan bir řeyleri tutmasına ve hızlı bir řekilde hareket etmesine yardımcı olacaktır. Bu cihazda engelli bir kiřinin gz yardımıyla hem tekerlekli sandalyeyi hem de dıř iskeleti sorunsuz bir řekilde kontrol edebileceęi iki kontrol paneli bulunmaktadır. eřitli testler yapıldı ve cihazın sorunsuz, iyi ve istikrarlı bir řekilde alıřtıęını gsterdi; bu da engelli insanlara uygun bir řekilde yardımcı olacaęı anlamına geliyor. Gz takibi hızının doęruluęu, basit ve karmařık grevler iin yeterince iyiydi ve bu cihazın iyi alıřtıęını gsteriyordu. Sonu olarak EMOHEX cihazı engelli bireylerin hayatında byk bir fark yaratacak ve engelliler gnlk yařamlarında gnlk grevlerini yerine getirirken kendilerini daha baęımsız hissedecekler.

Kragic, Bjrkman, Christensen ve Eklundh (2005), robotların evdeki nesnelere ynetmesine yardımcı olmak iin tasarlanmış bir grř sistemi kullanmıřtır. Bu sistem temel olarak  adıma odaklanıyor: Nesnelere tespit etmek, onlara yaklařmak ve onları yakalamak. Hem tek gzly hem de iki gzly grme ipularını kullanan sistem, nesnelere nasıl konumlandırıldıęını tanıma, takip etme ve anlama konusunda ok faydalı hale geliyor. Bu sistem, nesneyi tanıdıęında, nasıl grndęn řekilleriyle birleřtirerek, nesnenin nasıl yerleřtirildięini anlamaya sorunsuz bir řekilde geiř yapabiliyor. Bu sistemin bu avantajı, kullanıcıların her řeyin tutarlı olmadığı veya iyi bir aydınlatmanın olmadığı daęınık i mekanlarda bile iyi alıřmasına olanak tanır.

Bu alıřma aynı zamanda e-mentor grafikleri ve Kraliyet Teknoloji Enstits gibi birok řirketin bu teknolojiyi pratik ve deęerli kılmak iin bir araya geldięini de gsteriyor. Bu iřbirlięi, farklı mhendislik alanlarının gerek dnyadaki sorunları zme iin nasıl birlikte alıřabileceęini gsteriyor. Bu alıřma aynı zamanda robotik grř sistemi hakkında bildiklerimize de katkıda bulunuyor: Robotlar nesnelere anlıyor, nesnelere tanıyor ve insanlara gnlk yařamlarında yardımcı oluyor. Yeni fikirler, problemler ve ekip

çalışması, doğal dünyanın karmaşık görev ve sorunlarını değiştirmenin temel bileşenleridir.

Fuchs ve Belardinelli (2021), robotların, insanların toplama, kapma ve bir konumdan diğerine geçme gibi günlük görevler sırasındaki göz hareketlerini takip ederek ne yapma eğiliminde olduklarını nasıl anlayabildiklerini ayrıntılı olarak inceledi. İnsanlarla yapılan robotik görevlerin çok önemli olduğunu ve görevleri birlikte etkili bir şekilde planlayıp yürütmek için insanların niyetlerini hızlı ve doğru bir şekilde tahmin etmesi gerektiğini keşfettiler. Araştırmacılar, göz hareketlerinden elde edilen göz verilerini kullanarak, farklı durumlarda farklı konumlarda nesnelere manipüle etmek gibi niyetleri yüksek doğrulukla tanıyabilen modeller geliştirdiler. Bulguları, manuel görevler sırasında etkileşimi ve yardımcı geliştirmek için robotik sistemlerde göz izleme bilgilerinin önemini vurgulamaktadır. Bu çalışma, insan-robot etkileşimine yardımcı olmada güvenilir göstergeler görevi gören, bir sistemi daha verimli hale getirmenin yolunu açan, uzaktan kumanda gibi karmaşık ortamlarda bile insan-robot etkileşimi ile işbirliğine yardımcı olan göz hareketlerini kullanarak pek çok içgörü sunuyor. Sonuç olarak çalışma, göz izleme teknolojisinin entegre edilmesinin, yerleştirme gibi görev başarı şansını artırdığını ve bunun da etkili insan-robot etkileşimine yol açtığını öne sürüyor.

Levine, Pastor, Krizhevsky, Ibarz ve Quillen (2017), derin öğrenmeyi ve birçok veriyi kullanarak robotik kolun nesnelere daha iyi kavramasını öğretme konusunda harika bir çalışma yaptı. Sadece resimlere bakarak bir şeyleri yakalamanın en iyi yolunu öğrenmek için evrişimli sinir ağı adı verilen özel bir derin öğrenme modeli geliştirdiler. Çalışmaları, robotların nesnelere işleme biçimini iyileştirmeye yönelik yeni bir yaklaşım sunuyor. Yöntemlerinin gerçekten işe yaradığını kanıtlamak için farklı şeyleri yakalamak amacıyla yaklaşık 900.000 denemeden oluşan devasa bir veri kümesi topladılar.

Çalışmaları aynı zamanda bu farklı robot yaklaşımını da denedi ve onun büyük ölçekte ve çeşitli durumlarda iyi çalışabileceğini gösterdi. Bu, iyi eğitilmiş modeller ve veriye dayalı tekniklerin kullanılmasının, robotların nesnelere kullanma konusunda daha iyi olmalarına yardımcı olmanın ve onları herhangi bir çalışma alanında daha verimli ve esnek hale getirmenin harika bir yolu olduğunu gösteriyor. Gelecekteki araştırmacılar, kavramayı daha da iyi hale getirmek ve gerçek dünyadaki sorunları çözmek için pekiştirmeli öğrenmeyi deneyebilirler.

Robotik alanı her geçen gün daha da ilginçleşiyor, özellikle de dağınık yerlerdeki yeni öğelerin yönetimi söz konusu olduğunda. Zeng ve diğerleri (2019), akıllı kavrama tekniklerini yenilikçi nesne algılama ve tanıma ile birleştiren bir sistemi tanıttı. Her görev için özel eğitim verilerine ihtiyaç duymadan iyi sıkıştırma, emme, tripod ve sarma kavramalarını ortaya çıkaran, nesneden bağımsız bir kavrama çerçevesi oluşturdular.

Sistem, nesnelere tanımak için, çok fazla ekstra veriye veya yeniden eğitime ihtiyaç duymadan yeni öğelerle kolay kullanım için gözlemlenen görüntüleri çevrimiçi ürün resimleriyle eşleştiren alanlar arası görüntü sınıflandırmasını kullandı. Farklı kavrama görevlerini alanlar arası eşleştirmeyle birleştiren sistem, karmaşık alanlardaki nesnelere doğru bir şekilde tanıırken, maliyet oranlarına göre en iyi eylemi seçer. Sistem, 2017 Amazon Robotics yarışmasında da başarılı oldu.

Günümüz dünyasında, daha fazla araştırma desteği için çevrimiçi olarak tonlarca önceden eğitilmiş model ve kullanışlı veri mevcuttur. Sistem halihazırda güçlü olsa da gelecekteki çalışmaların temelleri kavramaya ve görüntü veritabanını daha fazla nesne türüyle oluşturmaya odaklanması gerekiyor.

Robotları görüşle kontrol etmek günümüzde robotikte, özellikle de yüksek doğruluk ve uyarlanabilirlik gerektiren görevler için çok önemlidir. Shih ve Lee (2019), iyi hizalama ve konumlandırma için basit modeller kullanarak eşyaları alma veya kavrama gibi robotik operasyonlarda kamera kontrolü için bir yöntem önerdiler.

Geleneksel yöntemler, hassas hareketler için 3 boyutlu ölçümlere ihtiyaç duyan PBVS'yi veya eylemleri yönlendirmek için 2 boyutlu görüntü koordinatlarına dayanan IBVS'yi içerir. Her iki sistemin bir karışımı, ilerledikçe sağlamlığı ve doğruluğu artıran hibrit çözümler oluşturur.

Görüş tabanlı robotik kontroller açık döngü veya kapalı döngü şeklinde de olabilir; açık döngü, sağlam kamera bilgisine ihtiyaç duyar ancak daha iyi hassasiyet sağlayan gerçek zamanlı hataları düzeltmek için kapalı döngü gibi geri bildirimlerden yoksundur. Zorluklar arasında kamera kalibrasyon hataları, teslim alma işleri gibi görevleri yavaşlatan karmaşık görüntü algoritmaları veya önceden bilgi olmadan 3 boyutlu nesne pozlarını ölçmede yaşanan zorluklar yer alıyor.

Ban ve diğlerleri (2023), engelli kişilere veya hassas görev desteğine ihtiyaç duyan kişilere, göz hareketleri yoluyla robotik kol kontrolü aracılığıyla olanak sağlayan, bakış algılama teknolojisini kullanan kalıcı İnsan-Makine Arayüzleri (HMI) üzerine araştırma başlattı - oldukça birinci sınıf şeyler! Çalışmaları, robotik kollar üzerinde gerçek zamanlı göz bakışı algılama ve yön kontrolü için makine öğrenimi tarafından yönlendirilen iki kameralı göz izleme sistemlerini (TCES) içeriyor.

TCES, tıp, endüstri ve güvenlik gibi çeşitli alanlara mükemmel bir şekilde hizmet eden basit göz hareketleri yoluyla hassas HMI kontrolünü hedefler ve diğler yöntemlerin kıyaslandığında başarısız olabileceği birden fazla eylemi süper doğru bir şekilde kontrol eden birinci sınıf doğrulukla mükemmel bir şekilde hizmet eder.

Bu çığır açıcı çalışma, pratik uygulamalarda genel olarak daha iyi performans için gelecekteki optimizasyonları ve genişletmeleri etkili bir şekilde teşvik ederek, gerçek dünyadaki zorlukların üstesinden gelmeye yönelik umut verici ilerlemeler göstererek depo sistemleri ve inşaat araçları arasındaki potansiyel faydaları ortaya koymaktadır.



3. GÖZ ALGILAMA VE KIRIŞ SAYAÇ

3.1. Göz Algılama

Göz tespiti, görüntü ve video akışlarında insan gözünün tanımlanmasına ve lokalizasyonuna olanak tanıyan bilgisayarlı görme teknolojilerinde çok önemli bir husustur. Kullanıcının bakışlarıyla hassas etkileşim gerektiren, örneğin göz izleme, göz kırpma algılama ve bakış tahmini gibi çeşitli uygulamaların temel bileşenini oluşturur. Etkili göz algılama, bakış kontrollü arayüzler, sürücü yorgunluğu tespiti, bazı erişilebilir teknolojiler ve ayrıca psikolojik araştırmalar gibi sonraki görevler için gereklidir. Bu bölümde göz kırpma algılamayı, modern dünyadaki uygulamasını ve farklı mühendislik alanlarına nasıl fayda sağladığını tartışacağız.

3.2. Göz Algılama İçin Derin Öğrenme Modeli

3.2.1. Veri toplama

Önerilen yöntemin ilk adımı, modelin eğitileceği gözlerin veri setini toplamaktır. Bu çalışma için seçilen birkaç videoyu ve bu videolardan çıkarılan ve web kaynaklı videolardan rastgele seçilen birkaç görüntüyü içeren bir veri kümesi. Görüntüler açık göz, kapalı göz, sol sağ ve sağ göz olmak üzere dört sınıfa ayrılmıştır. Bu veri seti, 2162'si eğitim, 322'si test ve 722'si doğrulama için olmak üzere 3206 görüntü içermektedir.

3.2.2. Görüntü ön işleme

Bilgisayarlı görmede uygulanması gereken çeşitli teknikler için bir verinin değerli hale getirilmesi amacıyla uygulanan bir dönüştürme işlemidir. Ön işlemede gürültü giderme gerçekleştirilir ve veri seti, saf bir veri seti sağlamak için görüntü işlemenin birkaç adımından geçer. Görüntü işleme teknikleri kullanılarak, görüntü içeriği anlaşılabilir analiz verimli bir şekilde gerçekleştirilebilir. Görüntü işlemenin temel amacı, görüntüdeki ilgi çekici noktayı işaret etmektir. Bu ilgi noktaları, özellik çıkarımı için kullanılır. Ön işleme, bozucu bozulmaları en aza indirmeyi ve ek işleme için gerekli olan unsurları geliştirmeyi amaçlar.

3.2.3. Evrimsel sinir ağı

CNN etkileyici başarılarla imza attı. Tam formu evrişimli bir Sinir Ağı olan CNN, derin öğrenme, en önemli sinir ağlarından biri olarak kabul edilir. Bilgisayar görüşüne dayanan evrişimli sinir ağı, insanların akıllı tıbbi bakım, yüz tanıma ve otonom araçlar gibi daha önce ulaşılamaz olduğu düşünülen görevleri gerçekleştirmesine olanak tanır (Sun, Li, Kang ve Shen, 2019; Alzubaidi ve diğerleri, 2021). Evrişimli bir sinir ağı, evrişim yapılarının yardımıyla verilerden özellikler çıkarabilen ileri beslemeli bir sinir ağı olarak kabul edilir. Özellikler, geleneksel özellik çıkarma prosedüründe manuel olarak çıkarılmıştır, ancak evrişimli sinir ağlarında, özelliklerin manuel olarak çıkarılmasına gerek yoktur. CNN'in mimarisi görsel algıdan ilham almaktadır (Aydogdu ve Demirci, 2017; Sun ve diğerleri, 2020). CNN çekirdekleri, birden fazla özelliğe tepki verebilen benzersiz reseptörlerdir. Yapay bir nöron biyolojik bir nörona eşdeğerdir. Aktivasyon fonksiyonları, nöronal elektrik sinyallerinin yalnızca belirli bir eşik üzerinde olduğunda bir sonraki nörona iletiildiği sürece benzer.

CNN'in mimarisi

Evrimsel katman

Bunlar temel blok CNN'lerdir. Bu bloğun her katmanı, çekirdekler veya evrişimli çekirdekler olarak da bilinen öğrenilebilir filtrelerden oluşur. Bu filtrenin işlevi, giriş görüntüsü boyunca kaydırmak ve mekansal özellikleri çıkarmak için evrişim işlemini gerçekleştirmektir. Son araştırmalar, bilgi işlem verimliliğini ve özellik öğrenme yeteneklerini geliştirmek için evrişim katmanlarının tasarımını optimize etmeye odaklanan sürekli çalışmayı göstermektedir (Zhang ve diğerleri, 2020).

Havuzlama katmanları

Havuzlama katmanları, önemli bilgileri korurken özellik haritasının uzamsal boyutunu azaltmak için evrişim katmanları arasında karıştırılarak önemli bir rol oynar. En yaygın havuzlama işlemi, özellik haritasının alt örneklemesini yapmak için kullanılan maksimum havuzlama ve ortalama havuzlamadır. Çeşitli çalışmalar, girdi özelliğine göre havuzlama

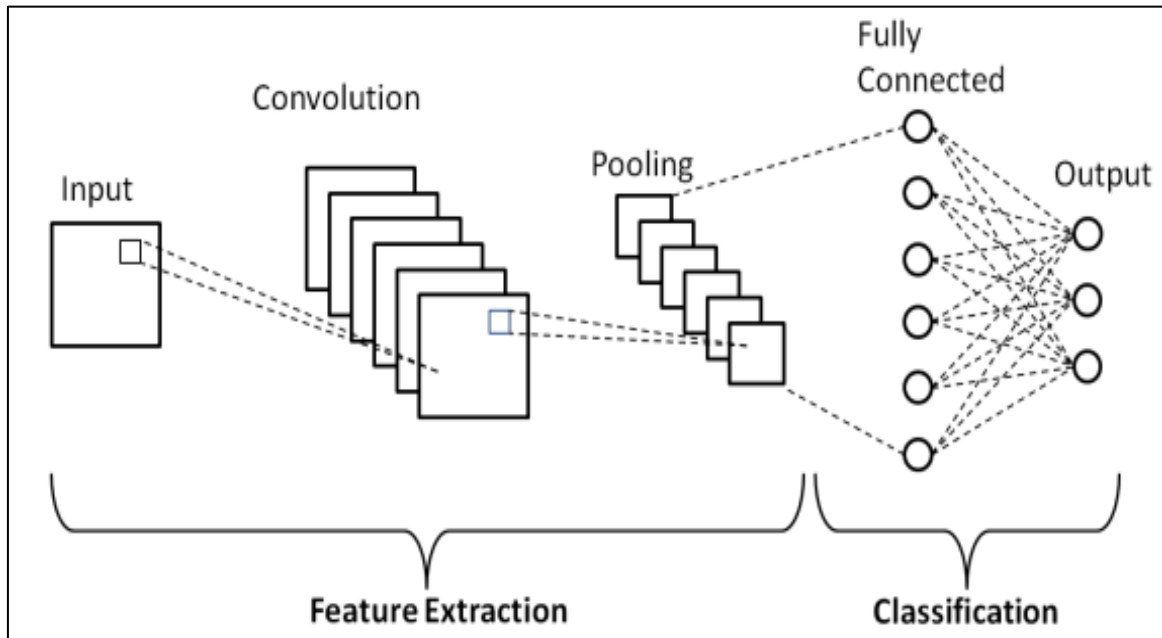
bölgelerini dinamik olarak ayarlamak için uyarlanabilir havuzlama tekniklerini araştırmıştır (He ve diğerleri, 2015).

Aktivasyon fonksiyonu

Etkinleştirme işlevi, ağın doğrusal olmamasına izin vererek verilerdeki karmaşık ilişkileri öğrenmesine olanak tanır. Yaygın olarak kullanılan ReLu (Düzeltilmiş Doğrusal Birim) gibi en yaygın kullanılan aktivasyon fonksiyonu. Son zamanlarda yapılan çeşitli araştırmalar, çeşitli görevlerde gelişmiş performans sergileyen Swish ve Mish gibi yeni aktivasyon fonksiyonları önermektedir (Ramachandran, Zoph ve Le, 2017).

Tamamen bağlı katmanlar

bir katmandaki her nöronu bir sonraki katmandaki her nörona bağlar. Çıkarılan özellikleri çıkış sınıflarıyla eşlemek için tipik olarak ağın sonuna bağlanırlar. Son gelişmeler, tamamen bağlı katmanların uzun menzilli bağımlılıkları yakalamadaki etkinliğidir (Vaswani ve diğerleri, 2017).



Şekil 3.1. CNN Mimarisi

3.2.4. VGG-19

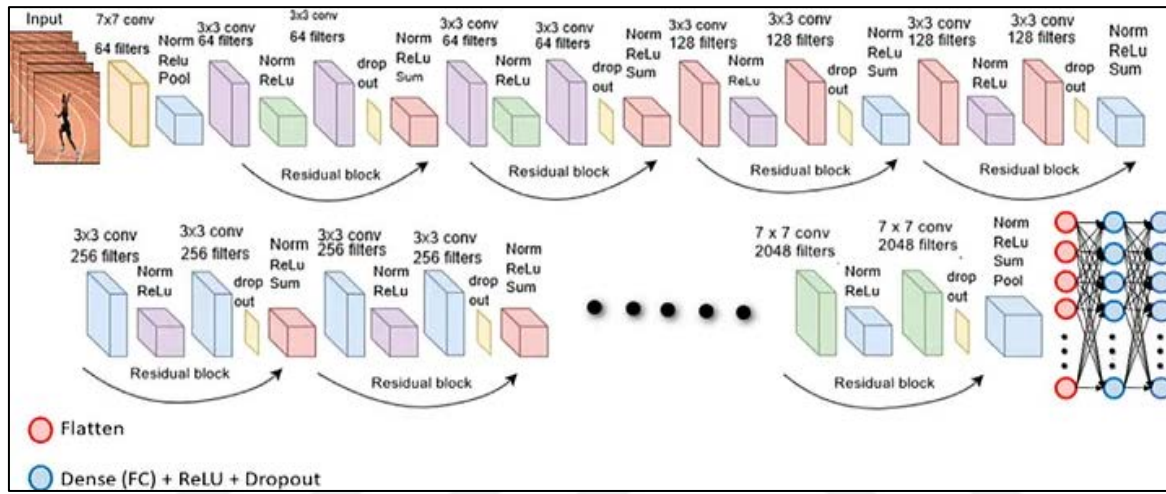
İlk olarak Oxford Üniversitesi'ndeki Görsel Geometri Grubu (VGG) tarafından önerilen derin Evrişimli Sinir Ağı mimarisidir. Basitliği ve görüntü tanıma görevlerindeki yüksek performansı nedeniyle en yaygın ve yaygın olarak kullanılan CNN mimarisidir. VGG-16 mimarisinin üç evrişimli katmana sahip bir uzantısıdır. VGG-19'daki evrişimli katmanlar, giriş görüntüsünden ince özelliklerin çıkarılmasına yardımcı olmak için kullanılan sol/sağ/aşağı ve merkez kavramını yakalayabilen küçük alıcı boyutlu alanlar kullanır. Ağın derinliği yüksektir çünkü ağın üst üste yığılmış birden fazla katman içermesi, daha karmaşık kalıpların öğrenilmesine olanak sağlar. Evrişimli katmanlar, non-lineerliği sağlamak için adım 1'i ve ardından ReLu aktivasyon fonksiyonunu kullanır. Maksimum havuzlama katmanlarında adım 2'ye sahip 2*2 filtre kullanılır ve özellik haritasının uzamsal boyutlarını azaltmak için kullanılır. VGGNet DNN giriş verileri, damar segmentli görüntülerden elde edildi. Eğitim aşaması sırasında özellik çıkarımı için evrişimli katmanlar kullanıldı ve özelliklerin boyutluluğunu azaltmak için evrişimli katmanların birçoğunun bağlantılı maksimum havuzlama katmanları kullanıldı. Görüntünün öznetelikleri, 3 x 3 filtre boyutunda 64 çekirdekten oluşan birinci evrişimsel katman yardımıyla çıkarılmaktadır. Özellik vektörü tamamen bağlı katmanların yardımıyla oluşturulur. Boyutluluğu azaltmak ve sınıflandırma sonuçlarını geliştirmek amacıyla görsel verilerden özellikler seçmek için elde edilen özellik vektörü, SVD ve PCA aracılığıyla ileri işlemlere tabi tutuldu (Sinha ve Le, 2021; Tammina, 2019). PCA ve SDA'yı kullanarak yüksek boyutlu verileri azaltmak önemli bir zorluktur. Önceki azaltma yaklaşımlarıyla karşılaştırıldığında PCA ve SVD daha hızlı ve istatistiksel olarak daha karardır, bu da onları daha değerli kılmaktadır. Son olarak softmax aktivasyon fonksiyonu yardımıyla 10-çapraz doğrulama kullanılarak sınıflandırma gerçekleştirilir.

3.2.5. RestNet101V2

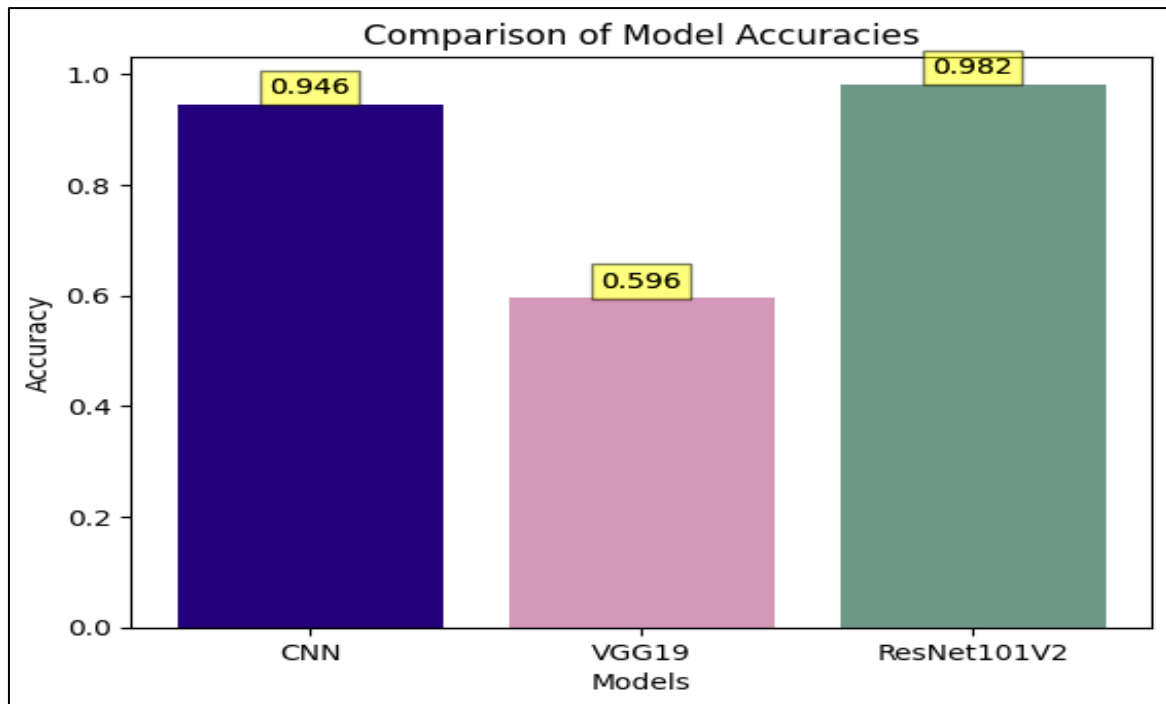
RestNet101V2, çeşitli bilgisayarlı görme görevlerinde sağlamlığa ve çok yönlülüğe sahiptir. Farklı mimari bileşenlerin, artık blokların ve ön aktivasyon birimlerinin madalya performansı üzerindeki etkisi, çeşitli görevlerde verimliliğini gösterir. RestNet101V2 mimarisi yüksek performanslı çoklu katmanlar içerir (Tammina, 2019). Sürüm (V2) ile sürüm (V1) arasındaki fark, sürüm (V2)'nin her ağırlık katmanından önce toplu

normalizasyon uygulamasıdır. RestNet, artık blokların yardımıyla kaybolan gradyan sorununu çözebilen evrişimli bir sinir ağıdır (He ve diğerleri, 2015).

Daha derin katmanlardaki darboğaz mimarisinin, modelin ölçeklenebilirliğine katkıda bulunan temsil gücü ile hesaplama verimliliğini etkili bir şekilde dengelediği gösterildi. ResNet101V2 gibi evrişimli sinir ağı mimarilerinin sürekli olarak iyileştirilmesi, bilgisayarlı görme alanında güçlü bir temel taşı olarak çalışır.



Şekil 3.2. RestNet101V2 Mimarisi



Şekil 3.3. CNN, VGG19 ve ResNet101v2 doğruluklarının karşılaştırılması

ResNet101v2, VGG-19 ve Evrişimli Sinir Ağı (CNN) mimarilerini kullanan göz kırpma algılamaya yönelik kapsamlı bir yaklaşımın ardından. Sonuçlar, bu modellerin çevrimiçi hareketli görüntülerden alınan veri kümeleri üzerinde ne kadar iyi performans gösterdiğini gösteriyor ve böylece görüntüler dört kategoriye ayrılıyor: açık, kapalı, sol ve sağ gözler. Ön işleme adımı, gürültünün engellenmesi ve özelliklerin geliştirilmesi yardımıyla modellerin performansını önemli ölçüde artırır ve bilgisayarlı görüntü uygulamaları için girdi verilerini optimize eder. Üstelik CNN mimarisi, evrişimli, havuzlamalı ve tamamen bağlantılı katmanlarıyla göz kırpmalarını tespit etmek için güçlü bir temel sağlar. Ayrıca, ResNet101v2'nin kalan blokları ve VGG-19'un uyarlanabilirliği, karmaşık görüntüleri analiz ederken her iki modelin neden daha iyi performans gösterdiğini açıklığa kavuşturuyor. %98,2 gibi oldukça etkileyici bir doğruluk elde eden ResNet101v2, karşılaştırmalı analizde hem VGG-19 hem de CNN'den daha iyi performans gösterdi. Doğruluğuna ve yakınsama oranına rağmen ResNet101v2, kaybolan degradeler gibi sorunları etkili bir şekilde çözebilir ve bu da onu karmaşık görüntü tanıma için güvenilir bir seçenek haline getirir. Araştırmacı, özellikle insan-bilgisayar arayüzleri veya sürücü yorgunluğu izleme gibi hassas göz kırpma algılamasının gerekli olduğu uygulamalar için pratik sonuçların olacağından emin. Avantajları ve dezavantajları hakkında doğru bilgi, çıkarların istedikleri görevler için en iyi derin öğrenme mimarisini seçmelerine yardımcı olacak ve onların iyi bilgiye dayalı kararlar almalarına olanak tanıyacaktır. Ayrıca, bilgisayarlı görme gelişmeye devam ettikçe, bu çalışma, CNN, VGG-19 ve ResNet101v2'nin göz kırpma tespiti için nasıl kullanıldığına dair kapsamlı bilgiler sunarak gerçek dünya uygulamalarının geliştirilmesine ve çeşitli alanlarda derin öğrenmenin genişletilmesine katkıda bulunuyor.

3.3. Mediapipe Kullanarak Göz Algılama

Mediapipe, Google tarafından geliştirilen ve gerçek zamanlı multimedya işleme hatları oluşturmak için kapsamlı bir çözüm sunan açık kaynaklı bir çerçevedir. Mediapipe'ta nesne algılama, yüz tanıma, el takibi, poz tahmini ve çok daha fazlası gibi görevlere yönelik çok çeşitli önceden oluşturulmuş bileşenler ve araçlar bulunur. Mediapipe'in tasarımı çok esnek, verimli ve kullanımı kolaydır; kullanıcı dostu olması onu artırılmış gerçeklik, robotik, sağlık ve eğlence gibi alanlardaki çeşitli uygulamalar için uygun bir seçenek haline getirir. Modüler mimarisi ve kapsamlı belgeleriyle Mediapipe, geliştiricilerin

çabalarını en aza indirerek son teknoloji multimedya uygulamaları oluşturmalarına olanak sağlar.

Mediapipe, gözleri tespit etmek için, göz izleme özelliklerini içeren yüz Mesh modeli de dahil olmak üzere çeşitli modeller ve araçlar kullanır. Yüz Mesh modeli, gözler ve yüzün çeşitli kısımları da dahil olmak üzere yüz yer işareti tespiti için mediapipe'in önemli bir parçasıdır. Makine öğrenimi algoritmalarına dayanan mediapipe modeli, gözler, burun ve ağız gibi yer işaretlerini doğru bir şekilde tanımlama yeteneğine sahiptir. Mediapipe, özellikle gözleri tespit etmek amacıyla, belirli bir görüntü veya video çerçevesinde bir insanın gözlerinin konumunu tespit etmek ve takip etmek için Face Mesh modelinin yeteneklerini kullanır. Model tarafından tespit edilen yüz işaretlerini inceleyerek. Mediapipe, gerçek zamanlı olarak gözlerin konumunu ve hareketini doğru bir şekilde belirleyebilir.

Ayrıca Mediapipe, uygulamaya yönelik özel çalışmamızın gerekliliğine bağlı olarak göz tespiti ve takibi için başka birçok araç ve model kullanabilir. Bu araçlar, özellikle göz algılama veya izleme için eğitilmiş özel makine öğrenimi modelini içerebilir veya mediapipe çerçevesindeki çeşitli görev sayısı için uygun modeller bulunabilir.

3.3.1. Gözlerin tespit edilmesindeki zorluklar

Yıldırım koşullarındaki değişiklik

Düşük ışık veya şiddetli yıldırım gibi yıldırım değişimleri, bir görüntü veya video karesinde insan gözünün görünürlüğünü etkileyebilir. Gözleri kısmak gözleri kısmen veya tamamen bloke edebileceği gibi yüz ifadesinin doğru bir şekilde algılanması ve takip edilmesinde zorluklara yol açabilir ve algılama algoritmasının bunları doğru bir şekilde tanımlamasını zorlaştırır.

Tıkanmalar

Nesneler ve diğer yüz özellikleri insanın gözünün görüşünü engellediğinde tıkanmalar ortaya çıkar. Örneğin gözlük takan, gözlerinin büyük bir kısmını kaplayan saçlar ya da bakmak gibi yüz ifadeleri gözleri kısmen ya da tamamen kapatabilen bir kişi, algoritmanın onları doğru bir şekilde tanımlamasını zorlaştıracak bir neden haline gelir.

Farklı göz şekilleri ve boyutları

Bu dünyaya her insan farklı göz şekilleri, boyutları ve iki göz arasındaki mesafeyle gelir. Göz özelliklerindeki bu çeşitlilik, farklı bireylerde doğru performans gösteren tüm göz algılama modellerine uyacak tek boyutlu bir model oluşturmayı zorlaştırabilir.

Baş hareketleri ve poz çeşitleri

Her birey başını hareket ettirir veya gözlerinin pozisyonlarını ve yönelimlerini değiştirir, bu da görüntülerdeki görünüşlerini etkiler. Bireylerin baş hareketlerinin ve göz hareketlerinin konumlarının bu dinamik doğası, göz tespit algoritmalarının baş hareketlerine karşı dayanıklı olmasını ve takip için gözlerin doğru ve kesin bir şekilde tespit edilmesini sağlamak üzere poz varyasyonlarını gerektirir.

Yansımalar ve parlama

Günümüzde pek çok kişi gözlük ve kontakt lens takıyor ve ışık kaynaklarında bir kişinin gözünde parlak vurgular yaratabilecek bir miktar parlama var. Göz üzerinde yapılan vurgular, gözleri tespit ederken rahatsızlık veren görüntülerdeki görünümünü etkiler. Bu yansımalar ve parlama, göz bölgesini bozarak göz algılama algoritmasını etkileyebilir.

Tüm bu zorlukların üstesinden geldikten sonra, farklı aydınlatma koşullarına, kafa hareketlerine ve tıkanmalara uyum sağlayabilen, göz görünümündeki değişiklikleri hesaba katabilen ve tüm bu koşullara rağmen gözleri sağlam bir şekilde takip edebilen algoritma ve tekniklerin geliştirilmesi gerekmektedir. Veri artırma, sağlam özellik çıkarma ve model optimizasyonu gibi çeşitli teknikler, her türlü zorlu koşulda göz algılama performansının iyileştirilmesine yardımcı olabilir.

3.4. Göz Kırpma Algılama

3.4.1. Göz en-boy oranını anlamak

Yüzdeki yer işaretlerinin tespiti, insan yüzündeki gözler, kaşlar, burun, kulaklar ve ağız gibi önemli özellikleri tanımlamak için kullanılan bir yaklaşımdır. Göz kırpma tespit çalışmasında sadece gözlere odaklandık. Bir kişinin yüzündeki her gözün haritası, gözün

sol köşesinden başlayıp onun etrafında saat yönünde hareket eden bir şekil oluşturan altı nokta kullanılarak haritalandırılabilir. Bu noktalara dikkat çektikten sonra gözün boyutunu, özellikle de genişlik ve yüksekliğini anlamamıza yardımcı olur.

Yanıp sönen izleme için Göz En-Boy oranı (EAR) adı verilen bir şey kullanırız. Bir yüzün gözündeki belirli noktalar arasındaki mesafeyi ölçen çok basit bir formüldür. Yatay mesafeleri ölçerek genişliği, dikey mesafeleri ölçerek ise yükseklikleri verir.

Göz en boy oranı, göz açıkken kabaca aynı kalır ancak göz kırıldığında önemli ölçüde değişir. Bu formülü kullanarak, karmaşık görüntü işleme adımlarına ihtiyaç duymadan herhangi bir kişinin göz kırıp kırılmadığını anlayabilirsiniz; yalnızca bu ölçümlerdeki değişikliklere bakın.

3.4.2. Yüzdeki yer işaretlerini kullanarak göz kırma algılama

Göz kırma, gözün milisaniyeler içinde hızla kapanıp açıldığı hızlı bir harekettir. Her insanın benzersiz bir göz hareketi ve açılıp kapanma düzeni vardır; bu düzen, hıza, gözlerin ne kadar sıkıldığına ve göz kırpmaların ne kadar süre kaldığına göre değişir. Tipik olarak, son göz kırpmasının süresi yaklaşık 100-400 milisaniyedir.

Göz kırpmaları tespit etmek için, gözleri ve göz kapaklarının hatlarını tam olarak tespit etmek amacıyla gelişmiş bir yüz yer işareti algoritması kullandık. Bundan sonra, bu yer işaretlerini kullanarak (EAR) kullanarak göz en-boy oranını hesaplayın; bu, gözlerin açık mı yoksa kapalı mı olduğunu analiz etmemize yardımcı olabilir. Ancak algılama sisteminin her birey için tamamen EAR'a yaslanması daha iyi değildir; çerçeve göz kırpmaları doğru şekilde algılayamıyor olabilir. Bunları ele alarak, göz kırpmaları daha iyi tanımlamak için daha büyük bir kare dizisini analiz eden bir sınıflandırıcıyı eğitiyoruz.

Her video karesinde yer işaretlerini tanımlayın ve ardından göz yüksekliğinin genişliğe oranını temsil eden Kulak değerini hesaplayın. Bu oran gözler açıkken sabit kalırken, göz kapatıldığında önemli ölçüde azalır.

3.5. Mediapipe Mimarisi

Mediapipe'da kullanıcı bir boru hattının aşamalı olarak prototiplenmesine izin verir. Boru hattı, her bileşenin bir hesap makinesi olduğu bileşenlerin yönlendirilmiş grafiği olarak tanımlanır. Grafik, bir graphconfig protokol arabelleği kullanılarak belirtilir ve ardından bir grafik nesnesi kullanılarak çalıştırılır.

Boru hattı yapısı

Mediapipe'da hesap makineleri adı verilen çeşitli bileşenlerden oluşan, grafik yönlendirmeli olarak tanımlanan bir Algılama boru hattı. Tüm bu farklı hesap makineleri tek bir şey yapmak için oradadır; ister bir model çıkarmak, ister bazı verileri dönüştürmek, ister bazı medyaları işlemek olsun. Grafik yapısı, veri akışınızı ve işlem hattındaki adım sürecini tanımlamanıza olanak tanır.

GraphsConfig protokol arabelleği

Mediapipe boru hattı grafikleri, hesap makinesi düzenlemesi ve hesap makinelerinin birbirine nasıl bağlandığı açısından grafik yapılandırmasıyla ilgili bilgileri saklayan bir graphconfig protokol arabelleği kullanılarak belirtilir. Protokol arabellekleri, boru hattının nasıl yapılandırılacağını ve yürütüleceğini tanımlar.

Veri akışları

Hesap makinesinde veri akışları, farklı bileşenler içindeki zaman serisi paketlerini temsil eden grafikler arasında bağlanır. Her akış, verilerin ardışık düzende senkronize ve sırayla hareket etmesini sağlamak için bir dizi veri paketini ve bu paketlerle ilişkili zaman damgalarını temsil eder.

Paket tabanlı veri işleme

Bu, Mediapipe'ın işleme sırasında zaman etiketleri ve C++'daki bazı değişmez türdeki yüklerle gönderdiği en küçük veri birimidir. Paketler değer sınıflarıdır ve dolayısıyla değer stilinde de işlenirler (yani gerçekten verimli bir şekilde kopyalanırlar). Bir cebin her

versiyonu, hesap makineleri arasında bilgi aktarımına izin vermek için diğerleriyle birlikte yük üzerinde kontrole sahiptir.

Paralel yürütme

Grafiklerdeki hesap makineleri eş zamanlı olarak çalışır ve her hesap makinesi belirli bir zamanda tek bir iş parçacığı üzerinde çalışır. Atılan adım iş parçacığı güvenliğini sağlar ve veri yarıışı koşulları riskini azaltır. Bu tasarımın seçilmesi, çok iş parçacıklı programlamadaki uzmanlık gereksinimini ortadan kaldırarak özel hesap makinelerinin geliştirilmesini kolaylaştırır.

Özel hesap makinesi geliştirme

Mediapipe'da geliştirici, özel hesap makinelerini kendi özel işlem gereksinimlerine göre tanımlama esnekliğine sahiptir. Geliştiriciler, grafik içine hesaplayıcılar ekleyerek veya bunları değiştirerek, işlem hattının genel yapısını bozmadan, aşamalı olarak hassaslaştırabilir ve yeni işlevler ekleyebilir.

Zaman akışı verilerinin akışı desteği

Mediapipe zaman serisi verilerini destekler; incelenen ses, video ve sensör verileri için uygun olmasını sağlar. Mediapipe çerçeve mimarileri, sürekli veri akışlarının verimli bir şekilde işlenmesini ve geçici veri serilerinin gerçek zamanlı analizini sağlar.

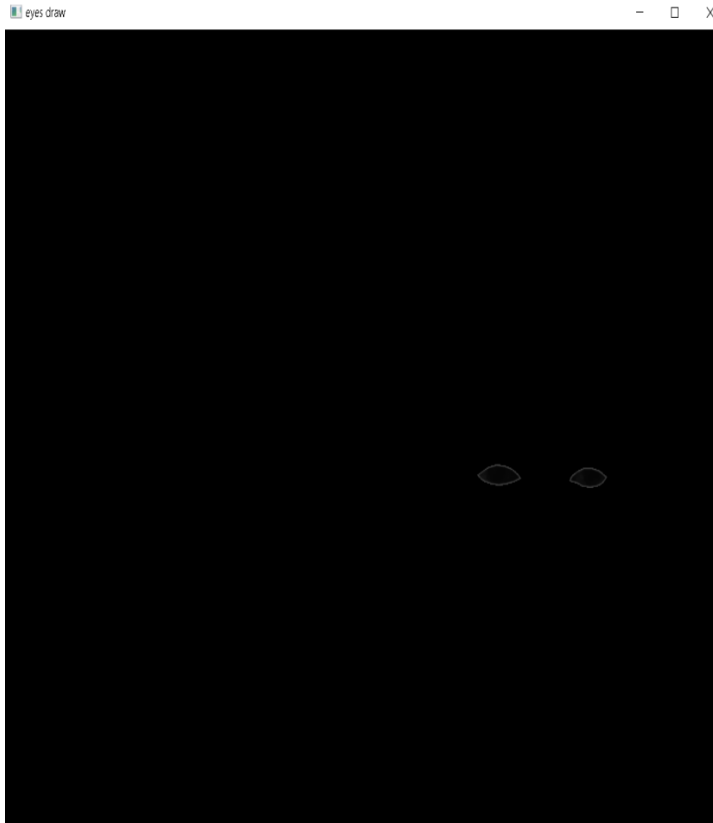
ResNet101V2 modelini kullanarak göz kırpması ve hareket

Bu çalışmada, işlemi gerçekleştirmek için yüz işaretlerini tespit etmek amacıyla Göz algılama konumu tahmincisi ve göz kırpması tespiti Mediapipe kütüphanesi tarafından gerçekleştirilmiştir. Adımlar aşağıda verilmiştir.

1. Yüzdeki yer işaretlerini algılama.
2. Göz yer işaretlerinin belirlenmesi.
3. Öklid uzaklığının bulunması.
4. ResNet101V2 modelinin başlatılması
4. Göz kırpması oranı.

5. Göz çıkarma.
6. Sayma yanıp sönüyor.
7. Göz pozisyonunun tahmin edilmesi.
8. Siyah piksel ile beyaz piksel arasında ayırım yapmak için eşik gözleri.
9. Her gözü üç parçaya ayırın.
10. Siyah pikselin sayılması ve konumunun belirlenmesi.

İnsan yüzündeki gözler, burun, ağız, kaşlar vb. gibi önemli noktaların konumlandırılmasını içeren yüz yer işaretlerini tespit ederken. Bu, esneme tespiti, yüz tespiti, duygu tespiti ve artırılmış gerçeklik gibi çeşitli uygulamalar için çok önemli bir süreçtir. Yüzdeki yer işaretlerini tam olarak konumlandırarak yüz ifadeleri ve yapıları hakkındaki önemli bilgileri kaldırabiliriz. Yer işaretleri algılama süreci aracılığıyla yüz yapısını ve geometrisini anladıktan sonra, bakış izleme, duygu tanıma ve esneme algılama gibi yüz ifadesi analizi gibi görevler için çok yararlı olan içgörüler kazanacaktır.



Şekil 3.4. İzolasyon sonrası maske üzerindeki şekillerin çizilmesi

Yüzdeki yer işareti tespiti alanında, gözlerle ilgili yer işaretlerinin belirli özelliklerini bulmak özellikle önemli hale geldi. Sol ve sağ gözler, derinlemesine bir bakışa olanak

tanıyan farklı yer işaretleri kümeleriyle sembolize edilmiştir. Göz kırpma oranlarını hesapladıktan sonra bakış yönlerini belirleyin ve oküler yer işaretlerini tam olarak konumlandırarak göz hareketini inceleyin. Bu bilgiler, kullanıcı katılımını anlamak, izleme sistemindeki sürücünün uykululuğunu belirlemek ve ayrıca bakışa dayalı etkileşimi tahmin etmek için çok faydalıdır. Göz işaretlerinin belirlenmesi, insan davranışını anlamak ve ayrıca İnsan bilgisayar etkileşimi, sürücü tabanlı izleme, güvenlik ve sağlık hizmetleri gibi çeşitli alan ve uygulamaların etkileşimini teşvik etmek için çok önemlidir.



Şekil 3.5. Göz Kırpmalarını Tespit Etme ve Sayma



Şekil 3.6. Sol ve sağ göz hareketlerini algılama

Öklid uzaklığı hesaplaması geometrik analizde, özellikle de gözlerin köşeleri ve kenarları gibi önemli yer işaretlerinin ölçülmesinde önemli bir rol oynar. Bu mesafeleri hesapladıktan hemen sonra göz genişliği, yüksekliği ve göz en-boy oranı gibi kullanışlı ölçümler elde ederek göz kırpma oranı hesaplamasını ve bakış tahminini kolaylaştırıyoruz. Bu çalışmada temel ölçüm, daha doğru özellik çıkarımına ve örüntü tanımaya olanak tanır ve ayrıca yüz hizalaması, yüz simetrisi değerlendirmesi ve bakış takibi gibi görevleri destekler. Öklid mesafe ölçümü, mekansal ilişkiyi ve nesne boyutunu anlama yeteneğini artırır, göz hareketlerinin yüz ifadesinin doğru analizi için kullanışlıdır. Göz algılama modelinde, RestNet101V2 medya borusu ile birlikte özellik çıkarıcı olarak görev yapar. RestNet101V2 modeli, kırılmış göz görüntülerini işler ve gözlerin şekiller, desenler ve dokular gibi yararlı özelliklerini yakalayan özellik haritaları oluşturur. Genelleştirme yeteneği, değişen aydınlatma ortamı koşullarında ve görüntü kalitesinde bile göz görüntülerinden ilgili özelliklerin hassas bir şekilde çıkarılmasına yardımcı olur. RestNet101V2 modeli gibi önceden eğitilmiş modellerin kullanılması, daha basit veya manuel olarak hazırlanmış özellikler kullanılarak yakalanması zor olabilecek göz hareketlerindeki ince farklılıkların tespit edilmesindeki hassasiyeti artırır.

4. NESNE ALGILAMA YOLOV8'İ KULLANARAK

4.1. Nesne Algılama

Nesne tespiti, bir görüntü veya videodaki nesneyi tanımlayan temel görevdir. Video veya görüntüdeki her bir nesne için, kullanılan ağ, nesnenin çevresine çok yakın bir şekilde uyan bir sınırlayıcı kutu ile nesnenin çıktısını vurgulamalıdır. Asıl amaç, nesnenin sınırlayıcı kutusunu tanımlamak ve onu sınıflandırmaktır. Önceden tanımlanmış kategoriler veya sınıflar.

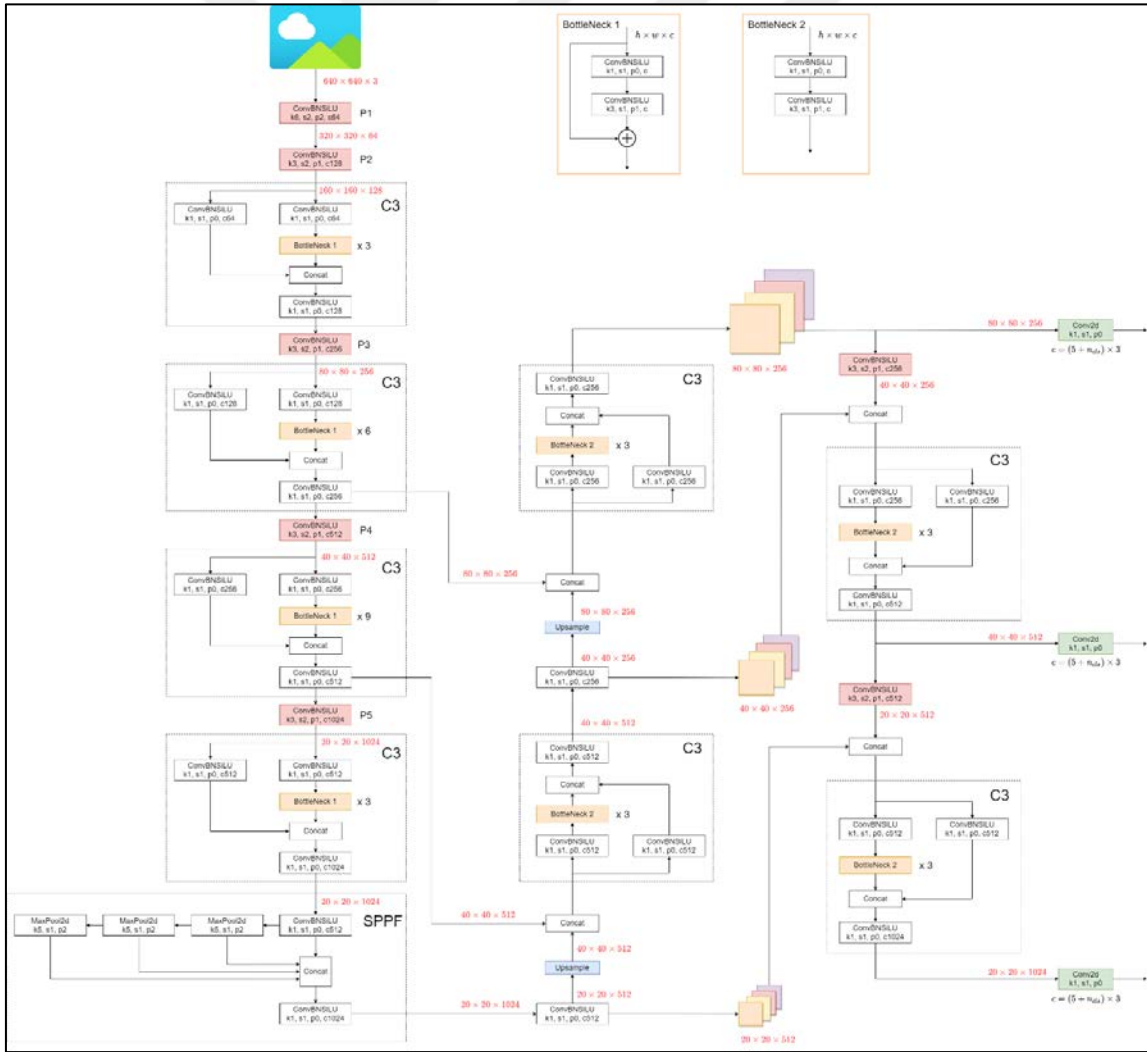
Nesne algılama görevi, araç algılama, sayı düzlemi algılama, robotik, şerit algılama vb. gibi bilgisayarlı görmenin çeşitli alanlarında çok önemlidir.

4.2. Mimari YOLOV8

Yolov8'de evrişimli blok, 2 boyutlu evrişimli bir katman, 2 boyutlu yatak normalizasyonu ve bir SeU aktivasyon fonksiyonundan oluşur; bunların hepsi tek bir evrişimli blok halinde birleştirilir. Yolov8'in omurgası, evrişimli bir dizilimden oluşan CSPDarkNet'in gelişmiş bir versiyonudur. iki evrişimli hızlı C2f Katmanına [Johnston ve Smith, 2020] sahip katman. Bir sonraki blok, evrişimli katmanları ve yeni tanımlanan C3f katmanlarını içeren C2f bloğudur. Buradaki evrişimli katmanlar standart 2 boyutlu özellik çıkarımını gerçekleştirir. C2f katmanları ELAN modülünden [Guo, Lin, Luo, Gao ve Zhang, 2023] ilham almıştır. C2f modeli, atlama bağlantısına sahip darboğazlara sahiptir; bu, C2f modülünü daha eğitilebilir hale getirecek, çok ölçekli özelliklerin öğrenilmesini geliştirecek ve aynı zamanda alıcı alanını genişletecektir. Omurganın sonunda bir sonraki blok SPPF'dir, SPF'nin değiştirilmiş versiyonu olan ancak daha yüksek bir hıza sahip olan Uzamsal piramit havuzlama hızlı anlamına gelir. SPPF bloğunun içinde, başlangıçta evrişimli blok, ardından 2d Max havuzlama katmanları ve evrişimli bir katmanla biten bir blok vardır. SPPF, daha hızlı çıkarım için uyarlanabilir olan ve doğruluğu artırmak için mekansal bilgiyi koruyan SPF ile aynı rolü üstlenir.

Boyun, üç farklı ölçekteki özellik haritasını omurgaya birleştirmede verimli bir şekilde çalışır. Daha küçük nesnelerin doğruluğunu ve algılama performansını artırmaya yardımcı olur. Yol entegrasyonunun yolunu takip eder ağları (PANet), yeni C2F modülünü [Fuchs ve Belardinelli, 2021] tanıttı.

Son blok, boyundan gelen özellik haritasından sınırlayıcı kutuyu nihayet algılayan kafadır. Yolov8'in tasarımı önceki versiyonlardan farklıdır. Yolov8 ankrajsız bir modeldir. Çapasız model ilk olarak Cornet Net [Kraft, Hartmann ve Bieber, 2022] tarafından tanıtılmıştır. Tespit bloğu iki parça içerir; bunlardan ilki sınırlayıcı kutu tahminleri içindir, diğeri ise sınıfları tahmin etmek içindir. Eğitim sürecinde sınırlayıcı kutuyu doğrudan tahmin etmek yerine, çapa tabanlı model bunların uzaklığını eğitim aşamasında önceden tanımlanmış bir dizi çapaya göre hesaplar. Sınırlayıcı kutuya güçlü bir öncelik verin ve bu modelin eğitilmesini kolaylaştıracaktır ancak. Çapasız tabanlı tasarım YOLOV8 modelini kullanarak daha iyi genelleştirmeye yol açar. Ayrıştırılmış algılama kafası ilk olarak Ayrıştırılmış iyileştirme (DCR) ağı tarafından tanıtıldı. Sınıflandırma ve regresyon görevini iki dala ayırmak ve ayrıca gradyan çatışmalarıyla ilgili sorunları ele almak ve nesne algılama hassasiyetini artırmak.



Şekil 4.1. Ultralytics'ten YOLOV8 Mimarisi

4.3. Kayıplar

Yerleştirme kaybı (Sınırlayıcı kutu regresyon kaybı)

Kayıp fonksiyonunun bu kısmı, tahmin edilen sınırlayıcı kutular ile temel doğruluk kutuları arasındaki hatayı ölçer. Genellikle yalnızca örtüşmeyi değil aynı zamanda kutuların merkez noktaları arasındaki mesafeyi ve en boy oranını da hesaba katan Birleşim Üzerinden Tam Kesişme (CIoU) kaybı kullanılarak hesaplanır.

$$Loss_{loc} = \sum_{i=0}^N CIoU(b_i, \hat{b}_i) \quad (4.1)$$

Burada b_i temel gerçek sınırlayıcı kutudur ve \hat{b}_i tahmin edilen sınırlayıcı kutudur.

Güven kaybı (Nesne kaybı)

Bu, modelin tahmin edilen sınırlayıcı kutudaki bir nesnenin varlığı konusunda ne kadar emin olduğunu ölçer. Genellikle İkili Çapraz Entropi (BCE) kaybı kullanılarak hesaplanır.

$$Loss_{conf} = \sum_{i=0}^N BCE(c_i, \hat{c}_i) \quad (4.2)$$

Temel gerçek nesnellik puanı nerede c_i ve \hat{c}_i tahmin edilen nesnellik puanı nerede?

Sınıflandırma kaybı

Bu, tespit edilen nesnelerin sınıf olasılıklarının tahmin edilmesindeki hatayı ölçer. Genellikle çok sınıflı sınıflandırma için Çapraz Entropi kaybı kullanılarak hesaplanır.

$$Loss_{cls} = \sum_{i=0}^N CE(p_i, \hat{p}_i) \quad (4.3)$$

Temel gerçek sınıf olasılığı ve \hat{p}_i tahmin edilen sınıf olasılığı nerededir? p_i

YOLOv8 için toplam kayıp fonksiyonu bu üç bileşenin ağırlıklı toplamıdır:

$$Total Loss = \lambda_{loc} \cdot Loss_{loc} + \lambda_{conf} \cdot Loss_{conf} + \lambda_{cls} \cdot Loss_{cls} \quad (4.5)$$

Yerelleştirme kaybı

Sınırlayıcı kutu tahminlerinin doğruluğunu ölçer (CIoU kaybını kullanarak).

Güven kaybı

Modelin nesnelerin varlığı konusunda ne kadar emin olduğunu ölçer (BCE kaybını kullanarak).

Sınıflandırma kaybı

Tahmin edilen sınıf olasılıklarının doğruluğunu ölçer (Çapraz Entropi kaybı kullanılarak).

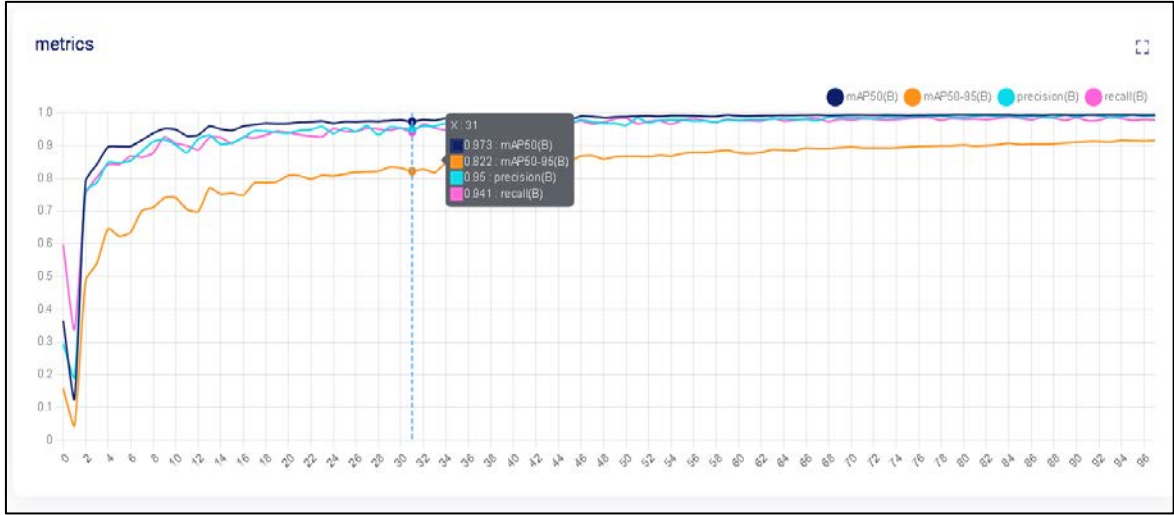
Toplam kayıp

Yerelleştirme, güven ve sınıflandırma kayıplarının ağırlıklı toplamı.

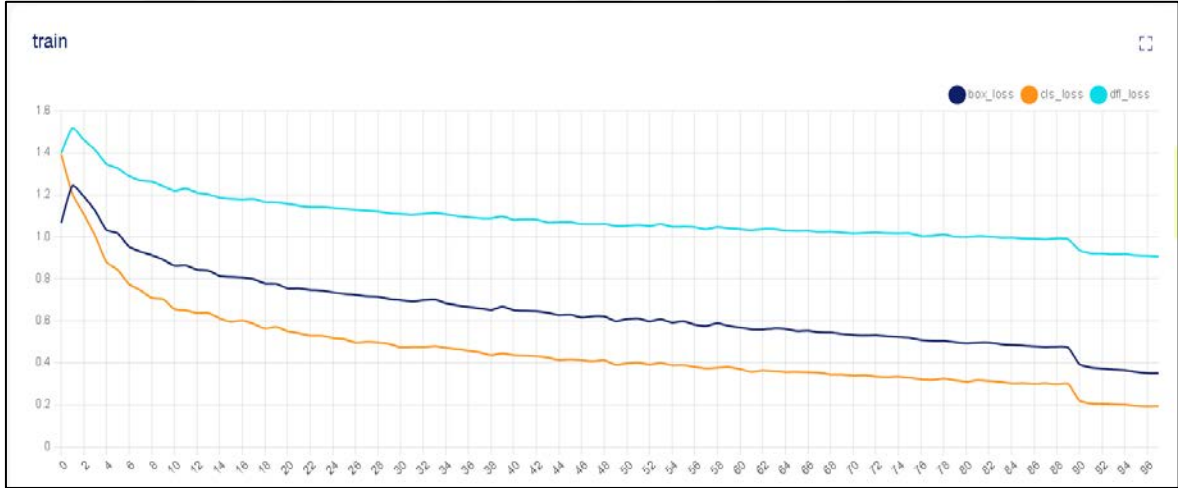
Bu kombinasyon, YOLOv8'in sınırlayıcı kutu yerelleştirmesinin hassasiyetini, nesne varlığına olan güveni ve sınıf tahminlerinin doğruluğunu dengeleyerek görüntüler içindeki nesnelere algılamayı ve sınıflandırmayı etkili bir şekilde öğrenmesini sağlar.

4.4. Eğitim

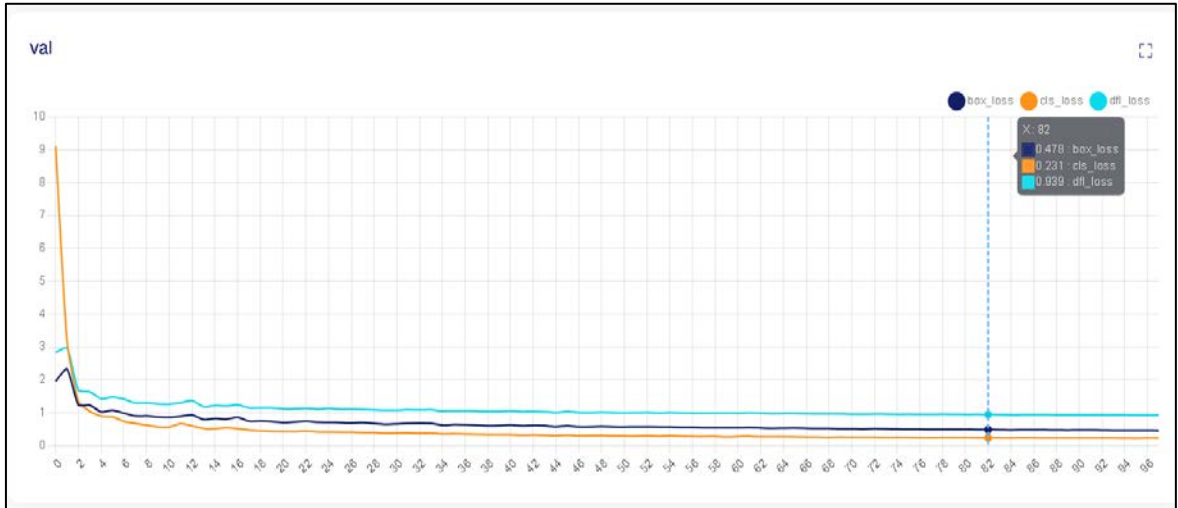
YOLOv8'deki eğitim süreci, yerleşik işlevlerle birlikte geldiğinden çok uygundur. Transfer öğrenimi verilerinin artırılmasının ayarlanması ve eğitim stratejilerinin değiştirilmesi, bu işlevlerin kullanılmasına yardımcı olur ve model performansının artırılmasına ve eğitim sürecinin hızlandırılmasına yol açar. YOLOv8'de modelin önceden eğitilmiş ağırlıklarla başlatılması, eğitimden önce transfer öğrenimiyle sonuçlanacaktır. Bu önceden eğitilmiş ağırlıklar, büyük veri kümeleri kullanılarak elde edilir. Önceden eğitilmiş ağırlıklar n YOLOv8, çok büyük ve çeşitli bir veri kümesi olan MS-COCO veri kümesi adı verilen büyük bir veri kümesi üzerinde eğitilir. Transfer öğrenimi, eğitim sürecini önemli ölçüde artırır, eğitim verisi ihtiyacını azaltır ve ayrıca model performansına fayda sağlar. Transfer öğrenimde özelliğin yeniden kullanılması yararlı avantajının nedenidir. Eğitim öncesi süreçte model, veri kümesiyle ilgili çeşitli özellikleri öğrenir ve veri kümesi büyük ve çeşitli olduğundan öğrenilen özellikler sağlamdır.



Şekil 4.2. Eğitim sırasındaki metriklerin grafiği



Şekil 4.3. Eğitim kayıpları



Şekil 4.4. Doğrulama kayıpları

4.5. Sonular

Eđitim veri seti, aıklamalı ve eđitim ve dođrulama setinde %84,8 ve %15,1 oranlarına blnmş 5524 grntden oluřmaktadır. Veri seti zerindeki eđitim, eřitli dikkate deđer modelleri ortaya koymaktadır. Eđitim ve dođrulamanın bařlangıcında, hem sınırlayıcı kutu (box_loss) hem de sınıflandırma (cls_loss) iin kayıplar, bazı dnemlerde azalma gsterir. Ancak kayıplarda, model istikrarı veya yakınsama konularındaki potansiyel zorluklara iřaret eden bazı ani artıřlar var. Model eđitimi, etkili bir řekilde alıřmayan optimizasyon sırasındaki zorluklara atıfta bulunduđunda ve performansının nasıl artırılacađını đrendiđinde yakınsama sorunları ortaya ıkar. Bu sorunlar, yavař yakınsama, plato kaybı, ařırı uyum ve bazı durumlarda ıraksama gibi eřitli řekillerde mdahale edebilir.

Tm bu dalgalanmalarla karřı karřıya kalan kayıpların genel gidiřatı, đrenmenin devam ettiđini gsteren bir azalma eđilimi gsteriyor. Tahmin edilen kayıpla (dfl_loss) temel geređe iliřkin diđer kayıp, bazı farklılıklarla aynı modeli izler. Metrikler hakkında, hassas ađrı ve mAP, belirli bir dnem boyunca tutarlı bir geliřme sergiler; bu, modelin nesnelere algılama yeteneđini yansıtır. Ortalama Keskinlik deđerleri istikrarlı bir byme gsterir; bu da farklı nesne rtřme seviyelerinde seviye performansında bir artıřa iřaret eder. Modelin genel eđitimi, model dođruluđunda ve veri kmesindeki eđitim sreci boyunca etkinliđinde olumlu bir geliřme olduđunu gstermektedir.

5. ROBOTİK İŞLETİM SİSTEMİ M(ROS) VE GAZEBO ORTAM KURULUMU

5.1. ROS

Robot İşletim Sistemi (ROS), robotik araştırma topluluğunun sağlam bir temel taşıdır; ROS, bir dizi araç, kitaplık ve sözleşmeyi içeren açık kaynaklı bir çerçeve sağlamak için tanıtıldı. Bu da, geniş bir robotik platform yelpazesinde karmaşık ve sağlam robotik davranışlar oluşturma görevini basitleştirir. Mimarisi, yazılım bileşenlerinin ölçeklenebilir, modüler ve yeniden kullanılabilir olmasını sağlamaya yardımcı olacak ve dolayısıyla araştırma ve ticari uygulamalarda uygulanabilecek şekildedir. Bunun nedeni, temeldeki birçok ana unsurun, özündeki mimari düzenlemenin gerçekleştirilmesini mümkün kılmasıdır. Düğümler, ROS mimarisindeki yürütme süreçlerini temsil eden en temel yapı taşlarıdır. Her düğüm modüler olacak şekilde tasarlanmıştır ve çeşitli programlama dillerinde uygulanabilir, ancak en sık kullanılanlar C++ veya Python'dur. Bu, geliştiricilerin birçok amaca özel düğüm oluşturmaya olanak tanıyan ve daha sonra birçok büyük sistemde belirli bir görevi gerçekleştirmek için de kullanılabilen olağanüstü bir tasarım felsefesidir (Johnston ve Smith, 2020). ROS Master, düğümlerin birbirlerini bulmaları ve eşler arası iletişim kanalları kurmaları için ad kaydı ve arama hizmetleri sağlar. Bu merkezi koordinatör, ROS tabanlı bir sistemde çalışan düğümlerin keşfedilmesini ve etkileşimini destekler.

Bu konfigürasyonda ROS, asenkron iletişim için konuları kullanarak çalışır; bu, düğümlerin bir konuya mesaj yayınlayabildiği veya bir konudan bildirimleri almak için abone olabileceği bir yayınlama-abone olma modelini takip eder. Böyle bir tasarım, düğümler arasında ayrıştırılmış iletişime olanak tanır ve böylece esnek ve ölçeklenebilir sistem mimarilerine olanak tanır. Dahası, istek/cevap etkileşimleri için hizmetler arasında senkronize bir iletişim modeline sahiptir ve bu, sonuçların daha sıkı bir şekilde veya belirli bir düzeyde onayla geri gelmesi gereken birçok uygulamayı gerçekleştirmesine yardımcı olur. Parametre Sunucusu, paylaşılan çok değişkenli bir sözlüktür; parametrelerin düğümler tarafından saklanmasına ve çalışma zamanında alınmasına olanak tanır. Merkezi konfigürasyon yönetimi için kullanılır.

ROS'u iyileştirmeye yönelik son çabalar, onu daha kullanıcı dostu, daha ölçeklenebilir ve günümüz teknolojisiyle daha iyi bağlantılı hale getirmeye yöneliktir. ROS 2, yeni nesil ROS olarak kabul edilir ve gerçek zamanlı sistemler için daha iyi destek, gelişmiş güvenlik özellikleri ve geliştirilmiş çoklu robot sistemi performansı dahil olmak üzere ROS 1'in dezavantajlarını kapsayan yeni özellikler içerir. ROS 2, robotikteki karmaşık uygulamalar için güçlü ve iletişim yeteneklerinde yüksek ölçeklenebilirlik sağlayan ara katman yazılımı olarak DDS'yi (Veri Dağıtım Hizmeti) benimser (Saéz-Pons ve diğerleri, 2014).

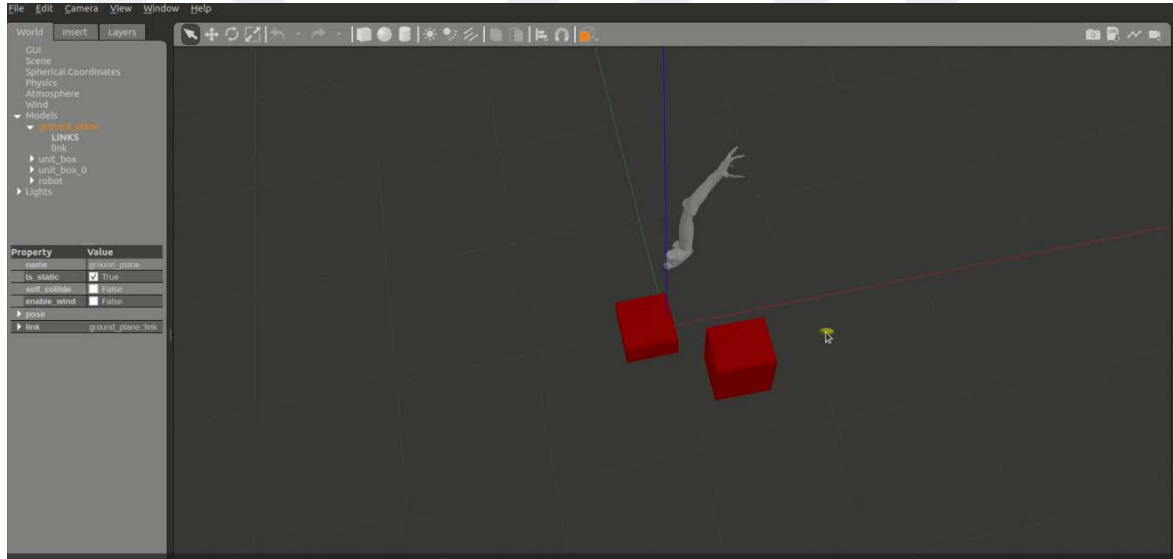
ROS kurulumuna yönelik adımlar, işletim sistemine göre birkaç alanda farklılık gösterir ancak en yaygın platform Ubuntu Linux'tur. ROS Noetic Ninjemys'i en son LTS sürümüne (Ubuntu 20.04) yüklemek için kullanıcı önce kaynak listesini ve anahtarları ayarlayacak, ardından paket dizinini güncelleyecek ve son olarak ROS paketlerini yükleyecektir. Tam bir kurulum için sonraki adımlar aynı zamanda rosdep'in başlatılmasını ve ortamın kurulumunu da içermektedir; bu, yalnızca temel parçaları değil aynı zamanda robotik uygulamaların geliştirilmesi için gerekli ek araçları da kuracaktır.

ROS'un esnekliği ve modülerliği, onun sayısız robotik sistem ve uygulamaya entegre edilmesine yol açmıştır. Örneğin, otonom araçlar alanına yapılan uygulamada ROS, LiDAR, kameralar ve GPS gibi birçok sensör türünü destekleyerek algılama, planlama ve kontrol için karmaşık algoritmalarla bunları uyumlu hale getirebilir (Johnston ve Smith, 2020). Robotik manipülasyonla ilgili olarak ROS, kavrama, manipülasyon ve kontrole ilişkin araçlar ve kütüphanelerle ve algoritmaları geliştirmek ve test etmek için MoveIt! kütüphanesiyle donatılmıştır. Hareket planlamadan manipülasyona, 3 boyutlu algılamadan kinematiğe kadar çeşitli işlevler içeren geniş destek sunan bir kütüphanedir ve bu ve diğer işlevler için uygulanabilir (Moll ve diğerleri, 2021). Öte yandan ROS, genellikle sağlık uygulamalarında insanlarla doğal ve sezgisel olarak etkileşime girebilen robotlar geliştirmek için yetkinliklerini kullanarak insan-robot etkileşim sistemlerinin tasarlanmasında da kullanılmıştır (Sharma ve diğerleri, 2020).

5.2. Gazebo

Gazebo, robotik algoritmaların test edilmesi ve geliştirilmesi için yüksek kaliteli ortamlar sağlamak üzere ROS ile entegre edilmiş güçlü, açık kaynaklı bir robot simülasyon aracıdır. Zengin işlevsellik, son teknoloji fizik simülasyonu, gelişmiş 3D grafikler ve esnek bir eklenti

sistemi sunar. Buna göre, Gazebo bugün robotik uygulamaların geliştirilmesi ve doğal ortamlarda konuşlandırılmadan önce test edilmesi için temel platformlardan biridir. Büyümesi ve başarısı için kritik bileşenler dünya dosyaları, modeller, sensörler ve eklentilerdir. Dünya dosyaları, Gazebo'da robotlar, sensörler vb. dahil olmak üzere simüle edilecek bir ortamı tanımlayan XML dosyalarını ifade eder. Bu dosyalar, gerçekçi bir simülasyon ortamının mevcut olmasını sağlayan yerçekimi, sürtünme ve aydınlatma gibi fiziksel özellikleri tanımlar. Modeller simülasyonda mevcut olan fiziksel varlıkları temsil eder: nesnelerin geometrisi, görünümü ve fiziksel özellikleri. Bu nesnelere özelliklerine göre farklılık gösterir. Gazebo'nun çok sayıda önceden oluşturulmuş modeli vardır, ancak kullanıcılar gereksinimlerine göre yeni modeller de tasarlayabilirler (Kehoe, Patil, Abbeel ve Goldberg, 2015).

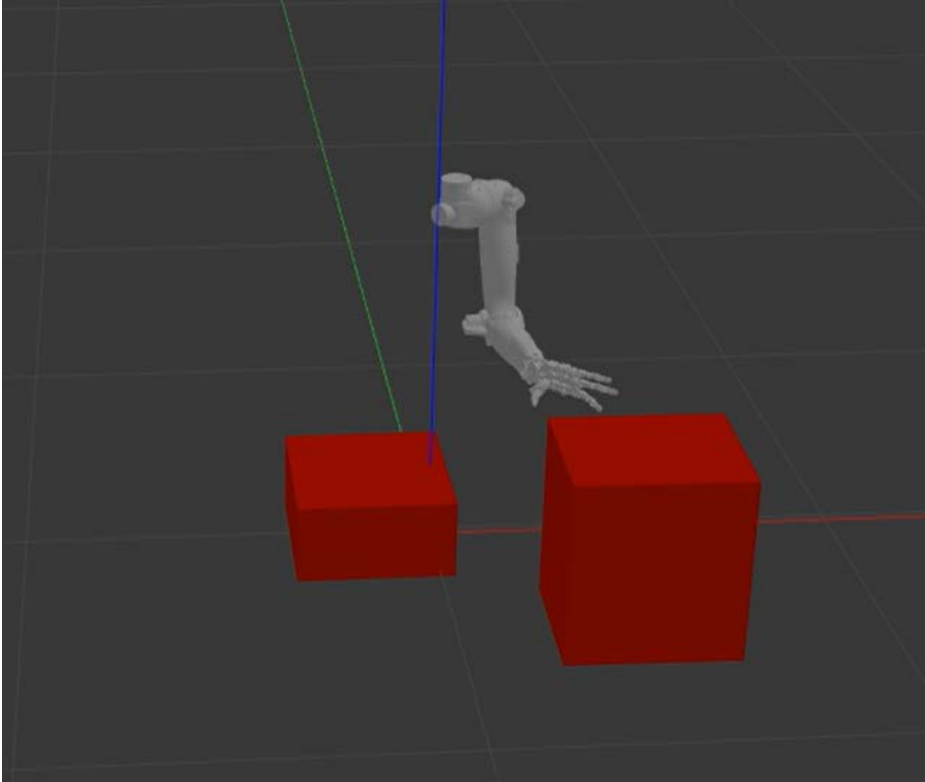


Şekil 5.1. ROS Ortamı

Gazebo, sensör girişlerine yönelik algoritmaların geliştirilmesine ve test edilmesine yardımcı olmak için kameralar, LiDAR'lar, IMU'lar ve GPS gibi bir dizi simüle edilmiş sensörden oluşur. Bu sensörler, mümkün olan en gerçekçi test için gerçek dünya donanımını taklit edecek bir seviyeye ayarlanabilir (Saéz-Pons ve diğerleri, 2014). Gazebo Eklentisi, modeli ve sensör davranışını veya simülasyon ortamıyla etkileşimi kontrol eden dinamik bir yükleme kitaplığıdır. Gazebo'nun işlevselliğini esnek bir şekilde genişletmenin ve onu diğer yazılım sistemleriyle entegre etmenin bir yoludur. Gazebo'nun geliştirilmesine yönelik son çalışmalar esas olarak performansına, kullanılabilirliğine ve ROS 2'deki entegrasyonuna odaklanmıştır. Gazebo sisteminin en yeni sürümleri, daha iyi gerçek zamanlı simülasyon mekanizmalarını, daha iyi grafik oluşturmayı ve aynı zamanda günümüz donanımıyla uyumluluğu bünyesinde barındırmaktadır. Bu gelişmeler Gazebo'yu

robotik araştırma ve geliştirme için daha sağlam ve çok yönlü bir araç haline getiriyor (Saéz-Pons ve diğerleri, 2014).

Gazebo'nun özellikle ROS ile birlikte kurulumu birkaç adımdan oluşur. Öncelikle kurulu ve doğru yapılandırılmış bir ROS'a ihtiyaç vardır. Daha sonra Gazebo kurulur; kurulum, paket indeksinin güncellenmesini ve gerekli Gazebo paketlerinin kurulmasını içerir. Bu iki sistem arasındaki entegrasyon ROS-Gazebo paketlerinin kurulumunu gerektirir. Aracın başlatılmasıyla kurulumun başarılı olduğu ve Gazebo'nun kullanıma hazır olduğu doğrulanabilir. ROS, Gazebo ile birlikte çalışabilir ve bu da onu, robotik sistemlerin gerçek dünya uygulamalarında konuşlandırılmasına gerek kalmadan yapılmasına yardımcı olan paha biçilmez bir geliştirme ve test aracı haline getirir. Bu entegrasyon, bir ROS paketi oluşturmayı, bir Gazebo dünya dosyası oluşturmayı ve simülasyonu hem ROS'u hem de Gazebo'yu başlatan bir başlatma dosyasıyla başlatmayı içerir. Bu temel adımlar, ROS ve Gazebo'yu entegre etmek için bir çerçeve oluşturarak araştırmacılar ve geliştiriciler tarafından araştırma ve geliştirme süreçlerinde robot sistemi simülasyonunun verimliliğini sağlar.



Şekil 5.2. Robotik kol alma ve yerleştirme hareketi

Gazebo, otonom navigasyon algoritmalarının geliştirilmesinde ve test edilmesinde yaygın olarak kullanılmaktadır. İçerisindeki farklı ortamlar, depo otomasyonu ve dış mekan araştırmaları gibi uygulamalar için büyük önem taşıyan çeşitli senaryolarda navigasyon algoritmalarının performansının test edilmesine yardımcı olur. ROS'un Gazebo ile entegre edilmesiyle, gelişmiş algılama algoritmaları üzerinde geliştirmeler yapılabilen, karmaşık senaryoların simülasyonu sağlanmakta ve geliştirilen algı sistemlerinin performansı, simüle edilen ortamdaki sensör verilerine dayalı olarak değerlendirilebilmektedir. Bu, nesne tanıma, SLAM ve engellerden kaçınma gibi farklı uygulamalar için hayati öneme sahiptir. Gazebo'nun paylaşılan bir ortamda çoklu robotları simüle etme yeteneği, onu çoklu robotlu sistemlerin geliştirilmesi ve test edilmesi için de ideal bir araç haline getiriyor. Çeşitli koordinasyon ve iletişim stratejilerinin emülasyonu, çok robotlu ekiplerin arama ve kurtarma, çevresel izleme ve işbirlikçi manipülasyon gibi görevler için optimum performans göstermesine yardımcı olacaktır (Saéz-Pons ve diğerleri, 2014).

Robotikte seçme ve yerleştirme görevi, nesneyi nesneyi kavramak ve onu istenen konuma koymak için görevi yürütmek üzere güçlü bir mekanizma kazanır. Bu proje görevinde, robotun nesneyi doğru bir şekilde taşımasına olanak tanıyan mekanizmayı çalıştırmak için ROS hizmetlerini kullandık. Tutucu, nesneye yapışmak için vakum oluşturma sürecini başlatan belirli ROS hizmetini çağırarak çalışır. Emme mekanizması başlamadan önce bu hizmetler proxy'ler oluşturur, bekler, kullanılabilirlik sağlar ve ardından emiş işlemini başlatır. Emme mekanizmasının kullanılması görevi ortaya çıktı çünkü Robotik kolumuzun fiziksel modelinde, bilekte tüm parmaklara bağlı yalnızca 1 eklem var, eğer eklem o eklem hareketinin motoruysa, bu da ona bağlı olan tüm parmakların hareketine yol açar. MoveIt genellikle robotların hareket ve hareket etmelerinde, bu aktivasyonun robotlarla senkronize edilmesinde kullanılır. Öncelikle hedefe ulaşmak için istenen konumu belirledi. Emme tutucusu, nesneyi yerine yerleştirildiğinde almak için sistemle birleşir. Nesne tutucu tarafından tamamen kavrandıktan sonra robot kol, nesneyi oraya yerleştirmek için istenen noktaya doğru ilerler. Daha sonra ilgili servis kullanılarak vakum devre dışı bırakılır.

Sonuç ROS ve Gazebo'nun entegrasyonu, hem robotik hem de bilgisayarlı görme alanını önemli ölçüde ileriye taşımış, dolayısıyla karmaşık robotik sistemlerin geliştirilmesi, test edilmesi ve konuşlandırılması için sağlam araçlar sağlamıştır. Açık kaynak yapıları, modüler mimarileri ve ölçeklenebilir platformları, gelişmiş robotik geliştirmeye erişimi

demokratikleştirerek dünya çapında inovasyonu ve araştırmayı teşvik ediyor. Sonuç olarak bu, biz bu alanda devam ederken ROS ve Gazebo'nun entegrasyonunun kritik olacağı anlamına geliyor.

5.3. Kullanıcı Arayüzü (UI)

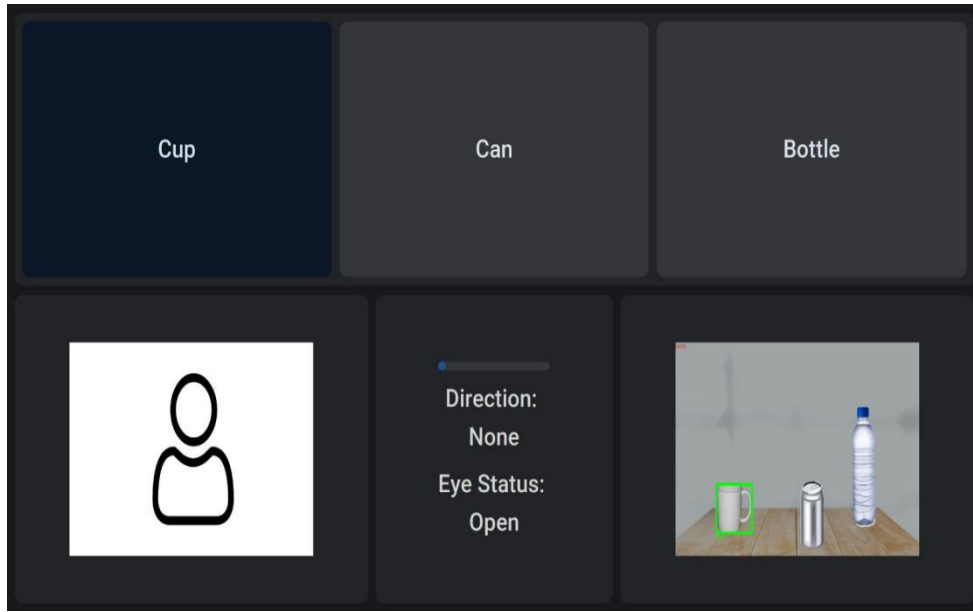
Kullanıcı Arayüzü (UI), esas olarak insanları bilgisayar sistemlerine ve uygulamalarına bağlayan ve teknolojiyle zahmetsizce etkileşim kurmamızı mümkün kılan bir köprüdür. Favori uygulamanızı veya web sitenizi kullandığınızı hayal edin. Uygulamada gezinme, aradığınızı bulma ve görevleri hayal kırıklığı yaşamadan gerçekleştirme kolaylığı, iyi tasarlanmış bir kullanıcı arayüzü sayesinde.

Bir kullanıcı arayüzü tasarlarken birkaç temel bileşen devreye girer. Görsel tasarım tamamen arayüzün görünümü ve hissi ile ilgilidir. Bu, her şeyi sadece güzelleştirmekle kalmayıp aynı zamanda okunmasını ve anlaşılmasını da kolaylaştıran doğru renkleri, yazı tiplerini ve düzenleri seçmek gibidir (Nielsen, 2021). Etkileşim tasarımı tıklattığımız, kaydığımız veya yazdığımız öğelere odaklanır. Doğal ve sezgisel hissettiren düğmeler, simgeler ve hareketler oluşturmakla ilgilidir. Kullanılabilirlik, arayüzün kullanıcı dostu olmasını sağlar; bu, onu nasıl kullanacağımızı anlamak için bir kılavuza ihtiyacınız olmadığı anlamına gelir (Shneiderman ve diğerleri, 2018).

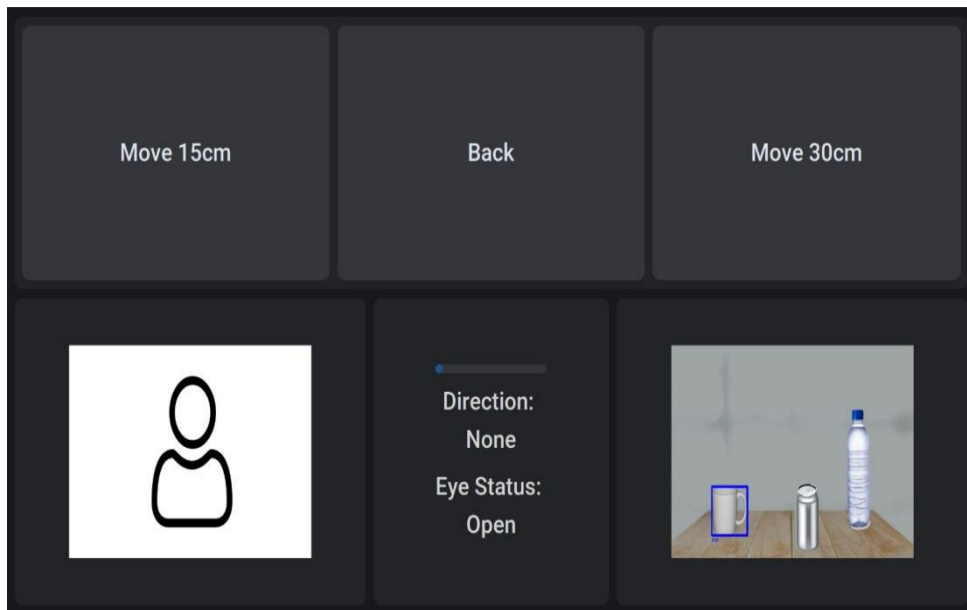
Duyarlılık, arayüzün telefon, tablet veya masaüstü olsun her cihazda iyi görünmesini ve iyi çalışmasını sağlayan bir diğer önemli husustur. Erişilebilirlik, engelli kişiler de dahil olmak üzere herkesin arayüzü herhangi bir sorun yaşamadan kullanabilmesini sağlamakla ilgilidir. Bu, ekran okuyucular gibi özelliklerin içeriği etkili bir şekilde yorumlayabilmesini sağlamak için aşağıdaki yönergeleri içerir (W3C, 2022).

Harika bir kullanıcı arayüzü oluşturmak, kullanıcıların neye ihtiyaç duyduğunu ve istediğini anlamakla başlar. Bu, kullanıcıları daha iyi tanımak için anketler ve röportajlar gibi birçok araştırmayı içerir. Tasarımcılar daha sonra arayüz planları gibi olan tel çerçeveler ve prototipler oluşturur. Bu prototipler, gerçek hayattaki senaryolarda iyi çalıştıklarından emin olmak için birçok kez test edilir ve ince ayarlar yapılır (Cooper, Reimann, Cronin ve Noessel, 2014).

İyi bir kullanıcı arayüzünün en önemli yönlerinden biri geri bildirimdir. Bir düğmeye tıkladığınızda ve renginin değiştiğini gördüğünüzde veya bir onay mesajı aldığınızda, bu size eyleminizin başarılı olduğunu bildiren bir geri bildirimdir. Bu, kullanıcıların neler olduğunu anlamalarına yardımcı olur ve hataları azaltır (Norman, 2013). Tutarlılık, basitlik ve öngörülebilirlik de temel ilkelerdir. Arayüzün bir bölümünü kullanmayı öğrendikten sonra diğer bölümlerde kafa karışıklığı yaşamadan kolayca gezinebilmenizi sağlarlar (Tidwell, Brewer ve Valencia, 2020).



Şekil 5.3. Nesne tanımlama



Şekil 5.4. Nesne seçimi ve gereksinime göre görevin hareketi

Sonuçta iyi hazırlanmış bir kullanıcı arayüzü, teknolojiyle etkileşimi keyifli ve verimli bir deneyime dönüştürür. Kullanıcı ihtiyaçları ve teknolojik gelişmelerle birlikte gelişir, daha iyi kullanılabilirlik ve memnuniyet sağlamak için sürekli olarak gelişir.

5.4. Kullanılan Donanım Bileşenleri

Çizelge 5.1. Projede kullanılan donanım bileşenleri

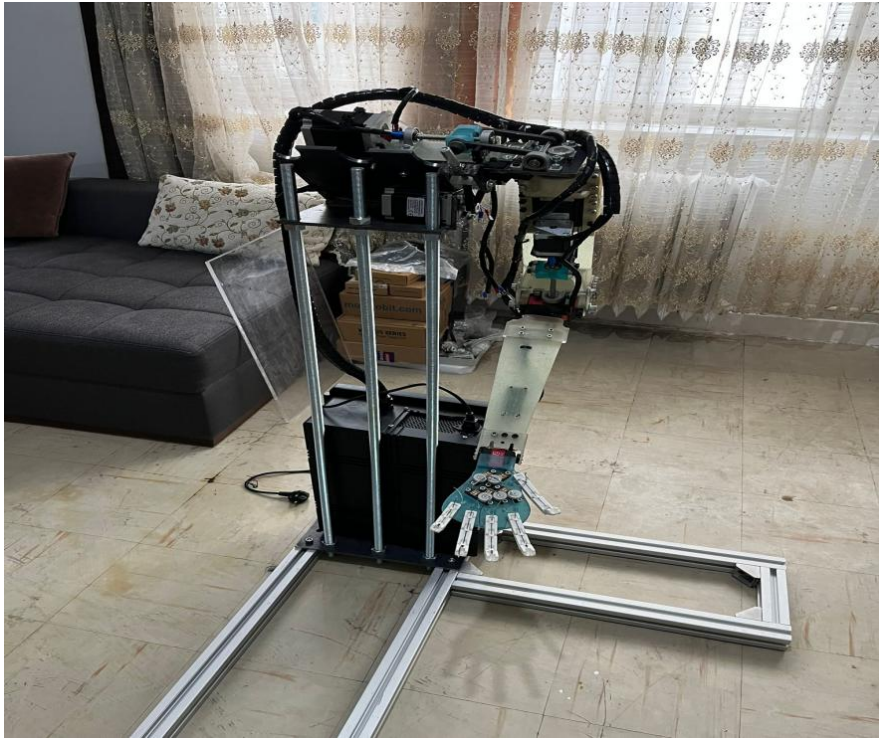
1. Fladen Maxximus Örgü Pe 0,55 mm Olta 100 m.	Yaygın olarak bilinen sağlamlığı ve dayanıklılığı ile bilinen örgülü bir oltadır. Tipik olarak balıkçılık için kullanılır ve ağır yükleri kırılmadan taşıyabilir.
2. Meon Robotik Otomasyon Nema 23 Adım Motoru 2,2 Nm	Robotik ve CNC makinelerinde yaygın olarak kullanılan yüksek torklu bir step motordur. NEMA standartları, motorun boyutlarını ve motorun montaj düzenini tanımlar.
3. JK42HS48 NEMA17 Adım Motoru.	NEMA 23'e kıyasla daha küçük bir adım motoru olan NEMA 17 adım motoru, 3D yazıcılarda ve diğer hassas ekipmanlarda yaygın olarak kullanılır.
4. Adım Motoru Sürücü Kartı DRV8825.	Genellikle yazıcılar ve diğer ekipmanlar gibi küçük makinelerde kullanılan NEMA 23'e kıyasla küçük bir step motordur.
5. 4A Step Motor Sürücü ve Kontrol Kartı TB6600.	· 4A'e kadar akıma sahip motorları sürebilen step motor sürücüsü. Step motorların hassas kontrolünü gerektiren uygulamalarda kullanılır.
6. 3D Yazıcı A4988 Step Motor Sürücü Genişletme Kartı.	· 3D yazıcılarda ve ince motor kontrolü gerektiren diğer uygulamalarda kullanılan, A4988 step motor sürücüsünü içeren bir genişletme kartı.
7. Emax Es3054 Metal Dişli Servo Motor 3,5 Kg	· Yüksek tork ve dayanıklılık sunan, metal dişlilere sahip bir servo motor. Robotikte, RC arabalarında ve diğer hassas kontrol uygulamalarında kullanılır.
8. LD-220MG Çift Eksenli Servo Motor	Robotik kollar veya pan-tilt kamera sistemleri gibi iki yönde hareket gerektiren uygulamalar için kullanılan çift eksenli bir servo motor.
9. 1 mm Esnek Olta 100 Metre Şeffaf	Balıkçılık ve güçlü, neredeyse görünmez bir misininin gerekli olduğu diğer kullanımlar da dahil olmak üzere çeşitli uygulamalar için kullanılan ince, esnek bir olta.
10. Güçlü Cıvata Somunu	Çeşitli mekanik düzeneklerde kullanılan, güvenli bir sabitleme çözümü sağlayan sağlam bir somun.
11. Güçlü Cıvata – Gijon M16	Güçlü, dayanıklı bağlantı elemanları gerektiren endüstriyel ve mekanik uygulamalarda kullanılan, 16 mm çapında ağır hizmet tipi bir cıvata.
12. 6006 2RS Rulman	Sürtünmeyi azaltmak ve radyal ve eksenel yükleri desteklemek için kullanılan, her iki tarafında lastik conta bulunan sabit bilyalı rulman.
13. Hobi Mekatronik Çapı 16 Krom Kaplama İndüksiyon Mili 300MM	300 mm uzunluğunda, 16 mm çapında şaft, krom kaplamalı ve daha fazla dayanıklılık için indüksiyonla sertleştirilmiş. Çeşitli mekanik uygulamalarda kullanılır.

Çizelge 5.1. (devam) Projede kullanılan donanım bileşenleri

14. En İyi Kale Galvanizli Çelik Somun M16 (10 Adet)	· Ağır hizmet tipi mekanik aksamlarda kullanılan, korozyon direnci için galvanizli çelikten yapılmış on adet M16 somun paketi.
15. M16 Gijon 1 Metre	Güçlü ve ayarlanabilir bağlantılar için inşaat ve mekanik uygulamalarda kullanılan, 16 mm çapında, 1 metre uzunluğunda dişli çubuk.
16. 50x50 Zamak Menteşe Burcu Krom Eksen: 30*30	Mobilya ve dolaplarda pürüzsüz ve dayanıklı dönme hareketi için kullanılan, krom kaplı zamak gövdeli bir menteşe.
17. Mikro Anahtar SS-5GL	Çeşitli elektronik uygulamalarda mekanik hareketi algılamak ve anahtarlama sinyali sağlamak için kullanılan küçük, hassas bir anahtar.
18. AeroCool VX PLUS AE-VXP750 750W 58A ATX Aktif PFC	Gelişmiş enerji verimliliği için Aktif PFC (Güç Faktörü Düzeltme) özelliğine sahip, masaüstü bilgisayarlar için 750 W güç kaynağı ünitesi.
19. Robolink Nema17 Step Motor Tutucu	· NEMA 17 step motorlar için, motoru çeşitli robotik ve mekanik düzeneklere güvenli bir şekilde monte etmek için kullanılan bir tutucu.
20. Somun Perçin	Cıvatalar için güvenli bir bağlantı noktası sağlayan, iş parçasına dişli bir parça takmak için kullanılan bir perçin türü.
21. Fiber Kablo	Esneklik ve dayanıklılık gerektiren çeşitli uygulamalarda kullanılan, fiber malzemelerden yapılmış yüksek mukavemetli bir kablo.
22. Robiz Teknoloji – Muz Soketi	Test ekipmanında ve hoparlörlerde yaygın olarak kullanılan, kabloları ekipmana bağlamak için kullanılan bir tür elektrik konektörü.
23. Robolink – Yüksük 10 adet	Çok telli kabloyu sonlandırmak için kullanılan, güvenli ve güvenilir bir bağlantı noktası sağlayan 10 halkadan oluşan bir set.
24. Robolink – Anahtar	Bir devredeki elektrik akışını kontrol etmek için kullanılan bir elektrik anahtarı.
25. Konya Caddesi – Direnç	Çeşitli elektronik devrelerde kullanılan, elektrik akımının akışına direnç sağlayan pasif bir elektrik bileşeni.
26. Konya Caddesi – Isıyla Daralan Boru	Isıtıldığında büzüşen, kabloları yalıtım ve koruma ve gerilim azaltma sağlamak için kullanılan boru.
27. Aydınlar Elektrik – Kablo Spiral	Kabloları toplayıp düzenlemek için kullanılan, koruma sağlayan ve dağınıklığı azaltan spiral sargı.
28. Aydınlar Elektrik – PG21 Kablo Rakoru / Siyah	Bir kablunun ucunu sabitlemek ve korumak için kullanılan, gerilim azaltma ve çevre koruma sağlayan bir kablo rakoru.
29. Mat Siyah Sprey Boya	Çeşitli uygulamalarda kullanılan, dayanıklı ve estetik bir görünüm sağlayan mat siyah sprej boya.
30. 5 Metre Kablo Toplama Spiral Siyah (8mm)	· Kabloları düzenlemek ve korumak için kullanılan, 8 mm çapında, kabloları demetlemek için 5 metre uzunluğunda spiral sarım.

Çizelge 5.1. (devam) Projede kullanılan donanım bileşenleri

31. Lojik Seviye Dönüştürücü 4 Kanal.	Farklı voltajlarda çalışan bileşenler arasındaki mantık seviyelerini güvenli bir şekilde düşürmek veya yükseltmek için kullanılan bir cihaz.
32. IRFZ44 - 49A 55V MOSFET - TO220 Mosfet.	Çeşitli elektronik cihazlarda elektronik sinyalleri değiştirmek ve yükseltmek için kullanılan bir tür transistör.
33. RobitShop - TB6600 4A Step Motor Sürücü ve Kontrol Kartı.	Hassas uygulamalarda kullanılan, 4A'e kadar akımla motorları sürebilen step motor sürücüsü ve kontrol panosu.
34. Logitech C310 HD Web Kamerası.	Video konferans, akış ve kayıt için kullanılan, yerleşik mikrofona ve net video kalitesine sahip bir HD web kamerası.
35. Meon Robotik Otomasyon Nema 23 Adım Motoru 2,2 Nm.	2,2 Nm torklu Meon Robotik Otomasyon Nema 23 Adım Motoru için yinelenen bir giriş.
36. NVIDIA Jetson Nano Geliştirici Kiti (4GB).	Yapay zeka ve makine öğrenimi uygulamaları için kullanılan, dört çekirdekli ARM işlemci ve NVIDIA GPU içeren güçlü bir geliştirici kiti.
37. DS5160 HV 60KG Dijital Servo Motor.	Ağır hizmet tipi robotik ve RC uygulamalarında kullanılan, 60 kg-cm'ye kadar torku idare edebilen yüksek torklu bir dijital servo motor.



Şekil 5.5. Birleştirilmiş robotik kol

5.5. İleri Kinematik Ters Kinematik

5.5.1. İleri kinematik denklemler

İleri kinematik, verilen eklem parametrelerine dayalı olarak uç efektörün konumunun ve yöneliminin hesaplanmasını içerir. N eklemlili bir robot kolu için bu, Denavit-Hartenberg (DH) parametreleri kullanılarak temsil edilebilir. Her eklem bir dönüşüm matrisine katkıda bulunur ve tabandan uç efektöre kadar genel dönüşüm, bu matrislerin çarpılmasıyla elde edilir.

Dönüşüm matrisindeki T_i bir eklem i , şu şekilde temsil edilebilir:

$$T_i = \begin{bmatrix} \cos \theta_i & -\sin \theta_i \cos \alpha_i & \sin \theta_i \sin \alpha_i & \alpha_i \cos \theta_i \\ \sin \theta_i & \cos \theta_i \cos \alpha_i & -\cos \theta_i \sin \alpha_i & \alpha_i \sin \theta_i \\ 0 & \sin \alpha_i & \cos \theta_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \dots \dots \dots \quad (5.1)$$

Nerede;

- θ_i burada eklem açısı, önceki z eksenini boyunca sapmadır
- d_i önceki z eksenini boyunca uzaklıktır
- α_i bağlantı uzunluğunu temsil eder
- α_i bağlantı bükümünü temsil eder.

Tabandan son efektöre kadar genel dönüşüm matrisi T, bireysel dönüşüm matrislerinin çarpımıdır:

$$T = T_1 T_2 T_3 \dots T_n \quad (5.2)$$

5.5.2. Ters kinematik denklemler

Ters kinematik, istenen uç efektör konumu ve yönelimini sağlayan eklem parametrelerinin bulunmasını içerir. Bu genellikle daha karmaşıktır ve analitik çözümler, sayısal yöntemler veya optimizasyon teknikleri gibi çeşitli yöntemler kullanılarak çözülebilir.

6-DOF'lu (Serbestlik Derecesi) bir robotik kol için ters kinematik problemi, her bir eklem açısı için çözüme ayrıştırılabilir. Konsepti açıklamak için düzlemsel 2-DOF koluna yönelik basitleştirilmiş bir örneği ele alalım:

Verilen;

- x ve y uç efektörün koordinatlarıdır.
- l_1 and l_2 iki bağlantının uzunluklarıdır.

de eklem açısı θ_1 and θ_2 denklemleri trigonometri kullanılarak türetilebilir.

$$\cos\theta_2 = \frac{x^2 + y^2 - l_1^2 - l_2^2}{2l_1l_2} \quad (5.3)$$

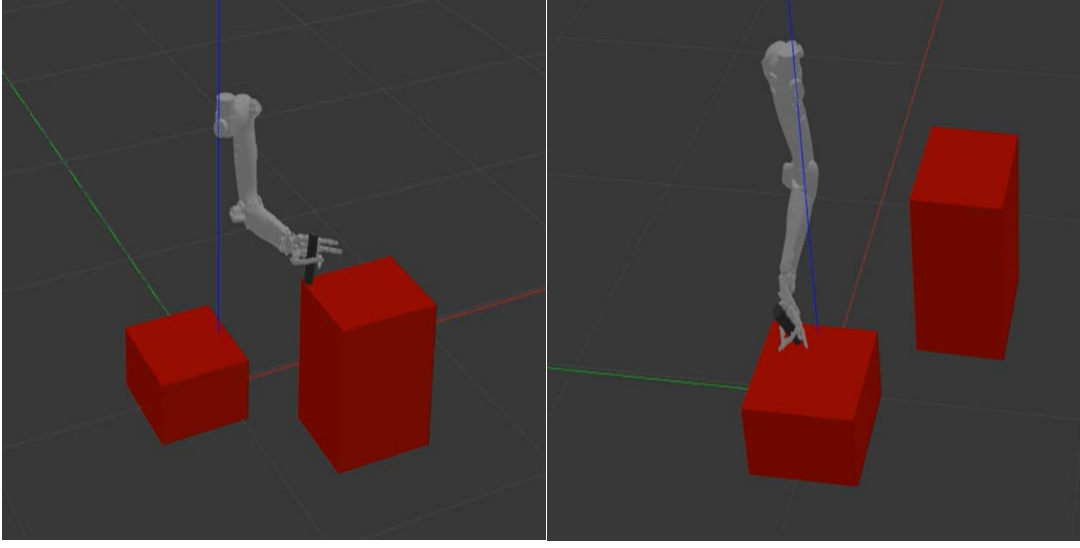
$$\sin\theta_2 = \sqrt{1 - \cos^2\theta_2} \quad (5.4)$$

Böylece θ_2 şu şekilde hesaplayabiliriz:

$$\theta_2 = \arctan2(\sin\theta_2, \cos\theta_2) \quad (5.5)$$

$$\theta_1 = \arctan2(y, x) - \arctan2(l_2 \sin\theta_2, l_1 + l_2 \cos\theta_2) \quad (5.6)$$

Robotik kolun hareketi ve yürütülmesi, kinematiğin soyut karmaşıklıklarını çözmek için entegre çözümleriyle iyi bilinen MoveIt tarafından kolaylaştırılır. Mevcut eklem durumlarından yararlanmak için MoveIt, robotun tabanından uç efektörüne kadar ileri kinematiği kullanır. Bu süreç, Denavit-Hartenberg parametrelerinde tanımlandığı gibi her bir bağlantı için dönüşüm matrisleri T'nin sırayla çarpılması yoluyla genel dönüşüm matrisi T'nin hesaplanmasını içerir. Her alma ve yerleştirme görevinde önemli olan görev, o MoveIt için ters kinematik yürütülürken istenen konumun elde edilmesidir. Belirlenen uç efektör pozunu elde etmek için eklem açıları gereksinimini belirlemek için çeşitli algoritmalar kullanır. KDL (kinematik dinamik kütüphane), mevcut konum ile istenen konum arasındaki hatayı en aza indirecek şekilde eklem açılarını yinelemeli olarak ayarlar ve aynı zamanda hassas hareketin elde edilmesini sağlar.



Şekil 5.6. Bir noktadan diğerine alma ve yerleştirme görevini gösterir



6. SONUÇ

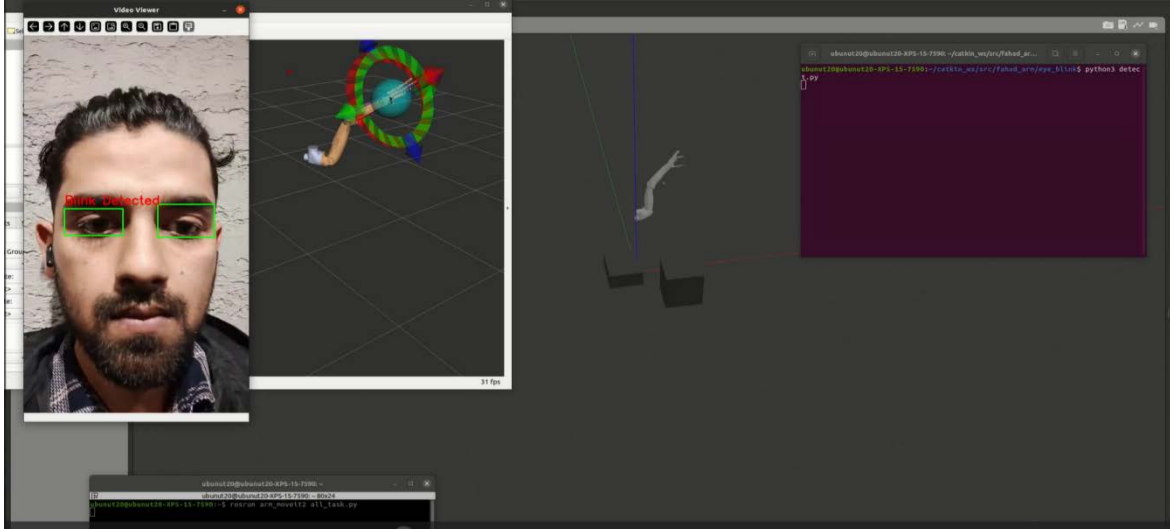
Proje, simüle edilmiş bir ortamda göz izleme teknolojisinin robotik bir kolla etkili bir şekilde entegre edilmesini başarıyla gösterdi ve göz kırpmalarının algılanmasından yararlanarak robotik hareketleri kontrol etti. Bu teknoloji, ciddi motor bozuklukları olan bireylerin özerkliğini artırmayı amaçlayan yardımcı uygulamalar için hayati önem taşıyan nesnelere hassas bir şekilde manipüle edilmesini sağlar.

Yapılan deneyler, sistemin göz kırpmalarını güvenilir bir şekilde algılayıp bunları robotik bir kolun doğru hareketlerine dönüştürebildiğini açıkça gösterdi. Sistemin, ortalama 0,45 metrelik bir yer değiştirme doğruluğu ve yer değiştirme ölçümlerinde $\pm 0,02$ metrelik bir standart sapma ile alma ve yerleştirme görevlerini gerçekleştirme yeteneği, hassasiyetini ve güvenilirliğini vurgular. Bu tür bir doğruluk, robotik kolun kullanıcı komutlarına yanıt verecek şekilde hassas bir şekilde ayarlanabilmesini sağlayarak gerçek dünya uygulamaları için olmazsa olmaz olan düzgün ve kontrollü hareketleri kolaylaştırır.

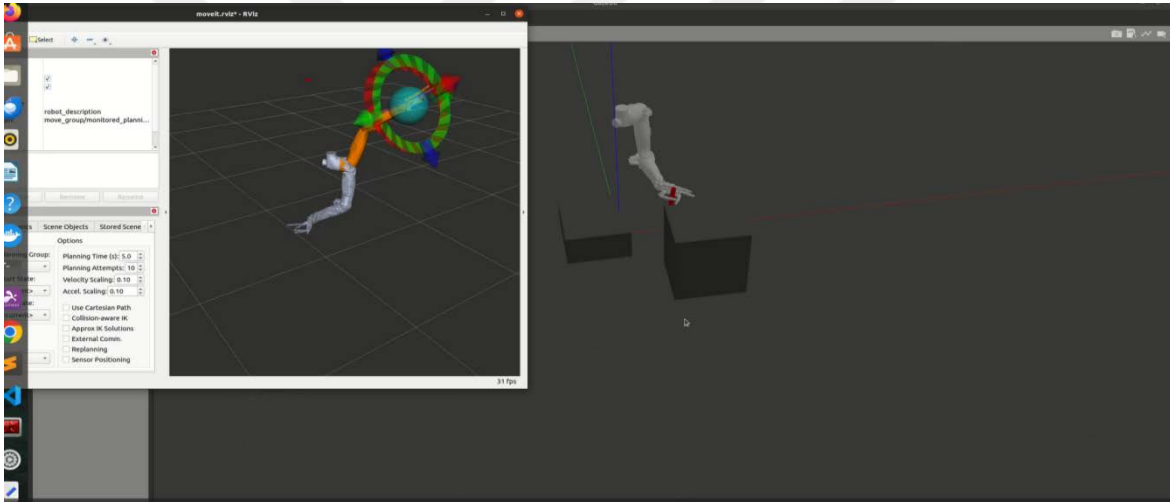
Dahası, sistemin göz kırpması algılamasından eylem başlatmaya kadar olan tepkisi, gerçek zamanlı uygulamalar için kritik öneme sahiptir ve kullanıcılara sorunsuz bir etkileşim deneyimi sunar. Göz izleme arayüzünün sağlamlığı, robotik kolun yol planlaması ve yürütmesinin hassasiyetiyle birleştiğinde, bu teknolojinin gelişmiş yardımcı cihazlar geliştirmedeki potansiyelini destekler.

Gelecekteki çalışmalar, daha karmaşık görevleri dahil ederek ve kullanıcı arayüzünü daha kolay kontrol için iyileştirerek sistemin etkileşim yeteneklerini geliştirmeye odaklanacaktır. Ek olarak, gerçek kullanıcılarla gerçek dünya testleri, sistemin daha fazla doğrulanması ve iyileştirilmesi için önemli olacak ve yardımcı olmak üzere tasarlandığı kişilerin ihtiyaçlarını karşıladığından emin olunacaktır.

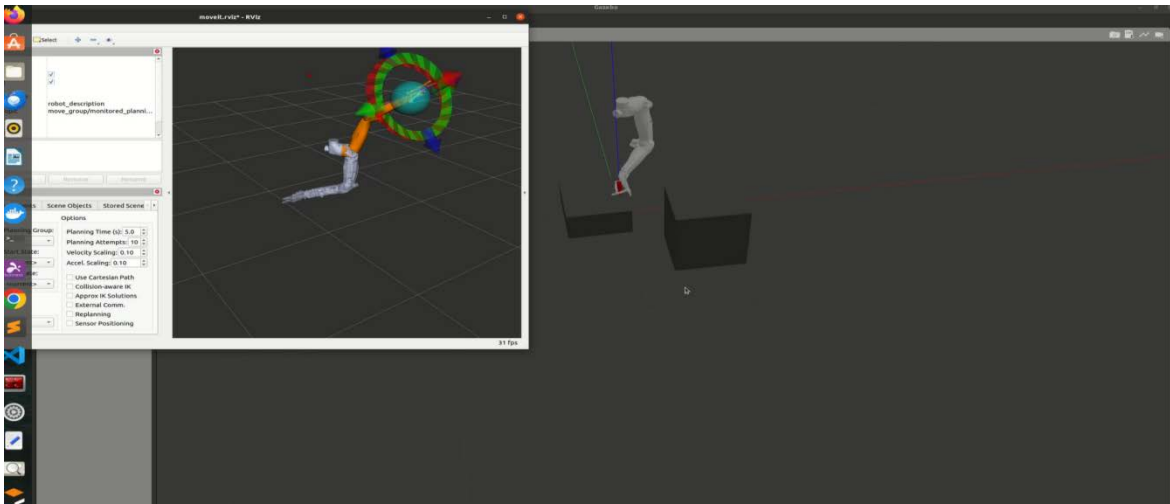
Bu proje yalnızca yardımcı robotik alanını ilerletmekle kalmıyor, aynı zamanda hassas kontrol ve kullanıcı tepkisinin kritik olduğu diğer alanlarda göz izleme teknolojisinin uygulanması için yeni olanaklar da sunuyor.



Şekil 6.1. Entegre görevin göz açıp kapayıncaya kadar yürütülmesi



Şekil 6.2. Nesneyi kavrama ve alma başlangıcı



Şekil 6.3. Son pozisyona ulaşmak ve nesneyi yerleştirmek

KAYNAKLAR

- Allen, P. K., Timcenko, A., Yoshimi, B., and Michelman, P. (1993). Automated tracking and grasping of a moving object with a robotic hand-eye system. *IEEE Transactions on Robotics and Automation*, 9(2), 152-165.
- Alpiste, F., Torner, J., and Brigos, M. (2017). Robotic hand controlled through vision and biomechanical sensors. *DYNA New Technologies Journal*, 4(1), 1-14.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaria, J., Fadhel, M. A., Al-Amidie, M., and Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, 53.
- Amri bin Suhaimi, M. S., Matsushita, K., Kitamura, T., Laksono, P. W., & Sasaki, M. (2023). Object grasp control of a 3D robot arm by combining EOG gaze estimation and camera-based object recognition. *Biomimetics*, 8(2), 208. <https://doi.org/10.3390/biomimetics8020208>
- Ali, H. M., Hashim, Y., and Al-Sakkal, G. A. (2022). Design and implementation of Arduino based robotic arm. *International Journal of Electrical and Computer Engineering*, 12(2), 1411-1418.
- Amri bin Suhaimi, M. S., Matsushita, K., Kitamura, T., Laksono, P. W., and Sasaki, M. (2023). Object grasp control of a 3D robot arm by combining eog gaze estimation and camera-based object recognition. *Biomimetics*, 8(2), 208.
- Anughna, N., Ranjitha, V., and Tanuja, G. (2020). Design and implementation of wireless robotic arm model using flex and gyro sensor. *International Journal of Recent Technology and Engineering*, 8, 2978-2983.
- Aydogdu, M. F., and Demirci, M. F. (2017). Age classification using an optimized CNN architecture. In *Proceedings of the International Conference on Compute and Data Analysis*, 19 May, University of Florida, USA, 233-239.
- Azar, A. R., and Khalilzadeh, F. (5-6 Kasım 2016). Real time eye detection using edge detection and euclidean distance. *2nd International Conference on Knowledge Based Engineering and Innovation (KBEI)*, Tehran, İnan, 43-48.
- Ban, S., Lee, Y. J., Yu, K. J., Chang, J. W., Kim, J. H., and Yeo, W. H. (2023). Persistent human-machine interfaces for robotic arm control via gaze and eye direction tracking. *Advanced Intelligent Systems*, 5(7), 1-9.
- Bernstein, A. V., Burnaev, E. V., and Kachan, O. N. (2018). Reinforcement learning for computer vision and robot navigation. In Perner, P. (Ed.), *Machine learning and data mining in pattern recognition*. Champaign: Springer International Publishing, 258-272.
- Bhadouria, A. S., Bhadouria, I. S., Patel, V., and Upasani, A. (2023). Robotic arm and Arduino controlled smart machine. *International Journal of Creative Research Thoughts*, 11(6), 2320-2882.

- Choi, I., Han, S., and Kim, D. (2011). Eye detection and eye blink detection using AdaBoost learning and grouping. *2011 Proceedings of 20th International Conference on Computer Communications and Networks (ICCCN)*, 1-4.
- Cooper, A., Reimann, R., Cronin, D., and Noessel, C. (2014). *About face: the essentials of interaction design*. New York: John Wiley & Sons, 118.
- Doijad, S., Bhalerao, A., and Khan, S. (2022). Impact factor: 2.205 eye irises tracking using MediaPipe ve OpenCV. *International Journal of Progressive Research in Engineering Management and Science*, 2(5), 319-322.
- Dröder, K., Bobka, P., German, T., Gabriel, F., and Dietriz, F. (2018). A machine learning-enhanced digital twin approach for human-robot-collaboration. *Procedia CIRP*, 76, 187-192.
- Fuchs, S., and Belardinelli, A. (2021). Gaze-based intention estimation for shared autonomy in pick-and-place tasks. *Frontiers in Neurorobotic*, 15, 647930.
- Guo, R., Lin, Y., Luo, X., Gao, X., and Zhang, S. (2023). A robotic arm control system with simultaneous and sequential modes combining eye-tracking with steady-state visual evoked potential in virtual reality environment. *Frontiers in Neurorobotics*, 17, 1146415.
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37, 1904-1916.
- Indriani, I., Harris, M., and Agoes, A. S. (2021). Applying hand gesture recognition for user guide application using MediaPipe. In *Proceedings of the 2nd International Seminar of Science and Applied Technology*. India: Atlantis Press, 101-108.
- İnternet: W3C. (2022). WCAG 2 overview. Web: <https://www.w3.org/WAI/standards-guidelines/wcag/>, Son Erişim Tarihi: 22.10.2024.
- Johnston, M., and Smith, J. (2020). Autonomous Vehicles and ROS. *Robotik Araştırma Dergisi*, 39(6), 789-810.
- Karas, K., Pozzi, L., Pedrocchi, A., Braghin, F., and Roveda, L. (2023). Brain-computer interface for robot control with eye artifacts for assistive applications. *Scientific Reports*, 13, 17512.
- Kehoe, B., Patil, S., Abbeel, P., and Goldberg, K. (2015). A survey of research on cloud robotics and automation. *IEEE Transactions on Robotics and Automation*, 12(2), 398-409.
- Kraft, D., Hartmann, F., and Bieber, G. (2022). Camera-based blink detection using 3D-landmarks. In *Proceedings of the 7th International Workshop on Sensor-Based Activity Recognition and Artificial Intelligence*. Association for Computing Machinery, New York, USA, Article 13, 1-7.

- Kragic, D., Björkman, M., Christensen, H. I., and Eklundh, J.-O. (2005). Vision for robotic object manipulation in domestic settings. *Robotics and Autonomous Systems*, 52(1), 1-16.
- Kumar, R., Bajpai, A., Sinha, A., and Singh, S. K. (2023). Mediapipe and CNNs for real-time ASL gesture recognition. *arXiv*, arXiv:2305.05296v3.
- Lalonde, M., Byrns, D., Gagnon, L., Teasdale, N., and Laurendeau, D. (2007). Real-time eye blink detection with GPU-based SIFT tracking. In *4th Canadian Conference on Computer and Robot Vision (CRV'07)*, 28-30 May, Montreal, Quebec, Canada, 481-487.
- Levine, S., Pastor, P., Krizhevsky, A., Ibarz, J., and Quillen, D. (2017). Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *International Journal of Robotics Research*, 37(4-5), 426-435.
- Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C. L., Yong, M. G., Lee, J., Chang, W. T., Hua, W., Georg, M., and Grundmann, M. (2019). MediaPipe: A framework for building perception pipelines. *arXiv*, arXiv:1906.08172.
- Matulis, M., and Harvey, C. (2021). A robot arm digital twin utilising reinforcement learning. *Computers & Graphics*, 95(1), 106-114.
- Meena, Y. K., Chowdhury, A., Cecotti, H., Wong-Lin, K., Dutta, A., and Prasad, G. (2017). EMOHEX: An eye tracker based mobility and hand exoskeleton device for assisting disabled people. *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary*, 2122-2127.
- Moll, M. ve ark. (2021). Hareket Et! 2: Robotik Manipülasyonda Yeni Bir Dönem. *IEEE Robotic and Automation Journal*, 28(2), 123-134.
- Moriyama, T., Kanade, T., Cohn, J. F., Xiao, J., Ambadar, Z., Gao, J., and Imamura, H. (2002). Automatic recognition of eye blinking in spontaneously occurring behavior. *Proceedings of the 16th International Conference on Pattern Recognition (ICPR' 2002)*, 4, 78-81.
- Nielsen, J. (2021). *Usability engineering*. Burlington, MA: Morgan Kaufmann, 56.
- Norman, D. A. (2013). *The design of everyday things* (Revised ed.). New York: Basic Books, 102.
- Paperno, N., Rupp, M., Maboudou-Tchao, E. M., Smither, J. A. A., and Behal, A. (2016). A predictive model for use of an assistive robotic manipulator: Human factors versus performance in pick-and-place/retrieval tasks. *IEEE Transactions on Human-Machine Systems*, 46(6), 846-858.
- Qul'am, H. M., Dewi, T., Risma, P., Oktarina, Y., and Permatasari, D. (2019). Edge detection for online image processing of a vision guide pick and place robot. In *International Conference on Electrical Engineering and Computer Science (ICECOS)*, 2-3 October, Batam, Indonesia, 102-106.

- Ramachandran, P., Zoph, B., and Le, Q. V. (2017). Searching for activation functions. *arXiv*, arXiv:1710.05941v2.
- Saéz-Pons, J., ve diğerleri. (2014). Multi-Robot Systems: A Survey of Current Approaches. *IEEE Robotics and Automation Journal*, 21(2), 83-92.
- Shafti, A., Orlov, P. and Faisal, A. A. (2019). Gaze-based, context-aware robotic system for assisted reaching and grasping. In *2019 International Conference on Robotics and Automation (ICRA)*, 20-24 May, Montreal, Canada, 863-869.
- Sharma, A., Pillai, D.R., Lu, M., Doolan, C., Leal, J., Kim, J., Hollis, A. (2020). Impact of isolation precautions on quality of life: A Meta-Analysis. *Journal of Hospital Infection*, 1-32.
- Sharma, R., ve diğerleri. (2019). Human-Robot Interaction in Healthcare: Challenges and Opportunities. *International Journal of Social Robotics*, 11(3), 495-509.
- Sharma, V. K., Saluja, K. S., Mollyn, V., and Biswas, P. (2020). Eye gaze controlled robotic arm for persons with severe speech and motor impairment. In *ACM Symposium on Eye Tracking Research and Applications*, 2-5 June, Stuttgart, Germany, 1-9.
- Shih, C. L., and Lee, Y. (2019). A simple robotic eye-in-hand camera positioning and alignment control method based on parallelogram features. *Robotics*, 7(2), 31.
- Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S., Elmqvist, N., and Diakopoulos, N. (2018). *Designing the user interface: Strategies for effective human-computer interaction*. New York: Pearson, 98.
- Sinha, S., and Le, D. V. (2021). Completely automated CNN architecture design based on VGG blocks for fingerprinting localisation. *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 29 November - 2 December, Lloret de Mar, Spain, 1-8.
- Sorate, P. P., and Chhajed, G. J. (2017). Computer vision system for eye gaze tracking. *International Journal of Modern Computer Science*, 5(3), 31-36.
- Soukupová, T., and Čech, J. (2016). Eye-blink detection using facial landmarks. *21st Computer Vision Winter Workshop*, February 3-5, Rimske Toplice, Slovenia, 1-8.
- Sun, K., Yu, J., Zhang, L., and Dong, Z. (2020). A convolutional neural network model based on improved softplus activation function. In Atiquzzaman, M., Choo, K. K. R., Zhang, X., Abawajy, J., and Islam, R. (Eds.), *International Conference on Applications and Techniques in Cyber Intelligence*. Champaign: Springer, 1326-1335.
- Sun, X., Li, Y., Kang, H., and Shen, Y. (2019). Automatic document classification using convolutional neural network. *Journal of Physics Conference Series*, 1176, 1-6.
- Tamina, S. (2019). Transfer learning using VGG-16 with deep convolutional neural network for classifying images. *International Journal of Scientific and Research Publications*, 9(10), 9420.

- Tidwell, J., Brewer, C., and Valencia, A. (2020). *Designing interfaces: Patterns for effective interaction design* (3rd ed.). New York: O'Reilly Media, 98.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., M., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). *Attention is all you need*. *arXiv*, arXiv:1706.03762v7.
- Viola, P., and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 8-14 December, Kauai, 511-518.
- Wang, L., Ding, X., Fang, C., Liu, C., and Wang, K. (2009). Eye blink detection based on eye contour extraction. *Image Processing: Algorithms and Systems*, 7245, 222-228.
- Zeng, A., Song, S., Yu, K. T., Donlon, E., Hogan, F. R., Bauza, M., Ma, D., Taylor, O., Liu, M., Romo, E., Fazeli, N., Alet, F., Chavan-Dafle, N., Holladay, R., Morona, I., Nair, P. Q., Green, D., Taylor, I., Liu, W., Funkhouser, T., and Rodriguez, A. (2019). Robotic pick-and-place of novel objects in clutter with multi-affordance grasping and cross-domain image matching. *International Journal of Robotics Research*, 41(7), 690-705.
- Zheng, X. ve Chen, T. (2020, Ekim). *Segmentation of high spatial resolution remote sensing image based on U-Net Convolutional Networks*. 2020 IEEE International Geoscience and Remote Sensing Symposium, HI, ABD.





Gazili olmak ayrıcalıktır