



**İÇ ORTAMLAR İÇİN YAPAY ZEKA TABANLI DERİNLİK ALGILAYAN
OTONOM ARAÇ GELİŞTİRİLMESİ**

Utku ULUSOY

**YÜKSEK LİSANS TEZİ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANA BİLİM DALI**

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

ŞUBAT 2022

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.

Utku ULUSOY

15/02/2022

İÇ ORTAMLAR İÇİN YAPAY ZEKA TABANLI DERİNLİK ALGILAYAN OTONOM ARAÇ GELİŞTİRİLMESİ

(Yüksek Lisans Tezi)

Utku ULUSOY

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Şubat 2022

ÖZET

Engel tanıma, engelden kaçınma ve engel aşma görevleri bilgisayarlı görü alanındaki en önemli çalışma konularındandır. Son yıllarda yapay zeka, bilgisayarlı görü sistemleri ve robotik alanlarında gerçekleştirilen çalışmalar sayesinde pek çok gelişme kaydedilmiştir. Yapılan çalışmalar genel olarak incelendiğinde, otonom robotların engelden kaçınmalarını sağlamak amacıyla pek çok farklı algoritma ve sensör kullanıldığı görülmektedir. Durağan ortamlar için geliştirilen yaklaşımlarda robot daha önceden bildiği harita bilgisinden ya da konumunu belirleyebileceği sensör bilgisinden yararlanarak yönünü bulabilmektedir. Bu çalışmada dinamik yani yer değiştiren engellerin bulunduğu bir ortamda harita bilgisi gibi bilgilere bağlı olmaksızın, bilgisayarlı görü sistemlerini kullanarak engellerden kaçınabilen yapay zeka tabanlı otonom bir kara aracı geliştirilmiştir. Bu kapsamda görüntüleme teknolojisi olarak pasif derinlik ölçebilen stereo kamera sistemi kullanılmıştır. Stereo kameralar üzerinde gerçekleştirilen görüntü işleme algoritmaları sayesinde görüş alanındaki ortama ait derinlik haritası çıkarılmıştır. Tez çalışmasında daha önce eğitilmiş SegNet isimli yapay zeka tabanlı semantik segmentasyon (anlamsal bölütleme) modeli kullanılarak zemine ait piksellerin anlamsal çıkarımı gerçekleştirilmektedir. Böylelikle robotun hareketinin istendiği elverişli zemin bölgesi tespit edilebilmektedir. Elde edilen bu bilgisayarlı görü sistemi çıktıları robot üzerinde daha önce belirlenmiş engelden kaçınma algoritmasının çalışmasında kullanılmaktadır. Çalışmada açık kaynaklı ROS işletim sistemi kullanılmıştır. Sistem iç mekân ortamında gerçek zamanlı olarak test edilmiş ve engellerden başarılı olarak kaçınabildiği görülmüştür. Buna ek olarak hareket edilebilir zemine ait anlamsal bölütleme işlemleri başarıyla gerçekleştirilmiştir. Sistemin gerçek zamanlı olarak çalışması, düşük maliyetli cihazlara entegre edilebilmesi, hareketli engellerin bulunduğu dinamik ortamlarda çalışabilmesi ve otonom kontrol imkânı sunması gibi pek çok avantajı bulunmaktadır.

Bilim Kodu : 90541

Anahtar Kelimeler : Stereo Görüntüleme, Yapay Zeka, Segmentasyon, Robotik, Bilgisayarlı Görü, Derinlik Algılama

Sayfa Adedi : 93

Danışman : Dr. Öğr. Üyesi Ayşe DEMİRHAN

DEVELOPMENT OF AN ARTIFICIAL INTELLIGENCE BASED DEPTH
ESTIMATING AUTONOMOUS VEHICLE FOR INDOOR ENVIRONMENTS

(M. Sc. Thesis)

Utku ULUSOY

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

February 2022

ABSTRACT

Obstacle recognition, obstacle avoidance and overcoming tasks are the most important study topics in the field of computer vision. In recent years, many developments have been made thanks to the studies in the fields of artificial intelligence, computer vision systems and robotics. When the studies are examined in general, it is seen that many different algorithms and sensors are used to enable autonomous robots to avoid obstacles. In approaches developed for stationary environments, the robot can find its direction by using the map information it knows before or the sensor information that it can determine its position. In this study, an artificial intelligence-based autonomous land vehicle that can avoid obstacles by using computer vision systems, notwithstanding information including map information, has been developed in an environment where there are dynamic, in other words, moving obstacles. In this context, a stereo camera system that can measure depth passively is used as imaging technology. The depth map of the environment in the field of view was obtained thanks to the image processing algorithms performed on stereo cameras. In the thesis study, semantic extraction of ground pixels is performed by using the previously trained artificial intelligence-based semantic segmentation model called SegNet. Thus, the suitable ground area where the robot's movement is desired can be determined. These computer vision system outputs obtained are used in the study of the obstacle avoidance algorithm determined before on the robot. The open source ROS operating system has been used in the study. The system has been tested in real time in an indoor environment and it has been seen that it can successfully avoid obstacles. In addition, semantic segmentation of movable ground has been successfully performed. The system has many advantages such as real-time operation, integration with low-cost devices, operation in dynamic environments with moving obstacles, and autonomous control.

Science Code : 90541

Key Words : Stereo Imaging, Artificial Intelligence, Segmentation, Robotic,
Computer Vision, Depth Estimation

Page Number : 93

Supervisor : Assist. Prof. Dr. Ayşe DEMİRHAN

TEŞEKKÜR

Tüm hayatım boyunca yanımda bulunarak beni koşulsuz destekleyen, beni cesaretlendiren ve bana güvenen canımdan çok sevdiğim sevgili annem Şahande ULUSOY'a sonsuz teşekkürlerimi sunarım. Hayatım boyunca desteklerini hiçbir zaman eksik etmeyen karizmatik babam Enver ULUSOY ve dünyalar tatlısı kardeşim Miyase ULUSOY'a minnet ve şükranlarımı sunarım. Çalışmalarımın yürütülmesi ve sonuçlandırılmasında desteğini eksik etmeyen çok değerli danışman hocam Dr. Öğr. Üyesi Ayşe DEMİRHAN'a maddi ve manevi tüm katkılarından dolayı en içten teşekkürlerimi sunarım. Yüksek lisans öğrenimimde kıymetli yönlendirmeleri ile bana yol gösteren Prof. Dr. Çetin ELMAS hocama tüm katkılarından dolayı teşekkürlerimi sunarım. Çalışmalarımı gözden geçirerek ikinci bir çift göz olan ailenin en zekisi ve tatlısı kuzenim Ayşegül EREN'e yardımları için teşekkür ederim. Yüksek lisans süresince bana desteklerini sunan TÜBİTAK BİLGEM İLTAREN ve K.K.K.lığına teşekkürlerimi sunarım.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	x
SİMGELER VE KISALTMALAR.....	xiii
1. GİRİŞ.....	1
1.1. Araştırmanın Tanımı ve Literatür Taraması	1
1.2. Araştırmanın Amacı	4
1.3. Araştırmanın Konusu ve Yöntem.....	4
2. MATERYAL VE METOTLAR.....	7
2.1. Görüntü İşleme.....	7
2.2. Bilgisayarlı Görü	11
2.3. Segmentasyon	13
2.4. Yapay Zeka	14
2.5. Yapay Sinir Ağları	16
2.6. SegNet Semantik Segmentasyonu.....	23
2.7. Derinlik Algılama.....	26
2.8. Stereo Görüntüleme.....	28
2.8.1. Sol ve sağ görüntünün eldesi.....	34
2.8.2. Kalibrasyon ve düzeltme (rectification) yapılması	34
2.8.3. Disparity (Farklılık) değerinin hesaplanması.....	38

	Sayfa
2.8.4. Derinlik değerin hesaplanması	41
2.8.5. Yakınsama ve filterelerin uygulanması.....	41
2.8.6. Reconstruction (üç boyutlu yeniden inşa) ve maskeleme işlemleri	42
2.9. Robot İşletim Sistemi	42
3. GELİŞTİRİLEN UYGULAMALAR	47
3.1. Gerçek Zamanlı Stereo Derinlik Tespiti	47
3.2. YZ Tabanlı Semantik Segmentasyon ile Gerçek Zamanlı Yol Tespiti	56
3.3. ROS Tabanlı Donanımın Oluşturulması	58
4. BULGULAR VE İRDELEME	71
4.1. Robotun Sahada Aldığı Görüntüler	71
4.2. Araç Yerleşim Planı ve Gerçek Zamanlı Hareket Sonuçları	75
5. SONUÇ VE ÖNERİLER	83
KAYNAKLAR	87
ÖZGEÇMİŞ	93

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 2.1. Aktivasyon fonksiyonları ve matematiksel ifadeleri	21
Çizelge 2.2. SegNet çalışması sonuçları	24
Çizelge 2.3. SUN RGB-D veri seti üzerinde gerçekleştirilen çalışma sonucu oluşturulan sınıflar ve doğruluk oranları	24
Çizelge 2.4. SegNet hafıza ve zaman performansı sonuçları.....	25
Çizelge 3.1. Hareket yönleri ve vektör bağlantısı.....	68

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2.1. Yaygın olarak kullanılan çözünürlük değerleri ve megapiksel değerleri	9
Şekil 2.2. Görüntü üzerinde piksel gösterimi.....	10
Şekil 2.3. a) Şehir görüntüsüne ait b) semantik, e) örnek, f) panoptik segmentasyon, c) görüntü sınıflandırma örneği ve d) nesne tanıma örneği gösterimi	13
Şekil 2.4. Yapay sinir hücresinin şematik gösterimi.....	19
Şekil 2.5. Yapay sinir ağı gösterimi.....	20
Şekil 2.6. 7 x 7'lik bir giriş matrisine piksel ekleme, konvolüsyon ve havuzlama işlemlerinin uygulanması.....	22
Şekil 2.7. SegNet semantik segmentasyonu kodlayıcı ve kod çözücü yapısı.....	23
Şekil 2.8. SegNet ve diğer ağ çıktılarının karşılaştırılması.....	25
Şekil 2.9. İnsan görme sistemi örnek görüntüleri	26
Şekil 2.10. Panum alanı, horopter, fiksasyon noktası gösterimi.....	27
Şekil 2.11. Stereo görüntüleme sistemi koordinat eksen ve değerleri	29
Şekil 2.12. Stereo görüntüleme belirsizlik	31
Şekil 2.13. Kameralardan derinlik görüntüsü eldesi şematik gösterimi	32
Şekil 2.14. Derinlik değeri elde etme adımları	33
Şekil 2.15. Radyal görüntü bozulmaları örnek gösterimleri	35
Şekil 2.16. Kalibrasyonda kullanılan satranç tahtası görüntüleri	36
Şekil 2.17. Kalibrasyon ve doğrultma adımları örnek gösterimi	37
Şekil 2.18. Sol ve sağ görüntünün epipolar düzlemde gösterimi.....	38
Şekil 2.19. Farklılık değeri hesaplanması adımları.....	39
Şekil 2.20. Sol ve sağ görüntü farklılık değeri maliyet hesaplaması.....	40
Şekil 2.21. Blok eşleme (BM) ve Yarı Global Blok Eşleme (SGBM) örnekleri.....	41
Şekil 2.22. ROS mimarisi basit gösterimi.....	43
Şekil 3.1. Stereo kamera modülü IMX29-83	48

Şekil	Sayfa
Şekil 3.2. Kırmızı renkli görüntü bozuklukları.....	48
Şekil 3.3. Kalibrasyonda kullanılan satranç tahtası görüntüsü örneği.....	49
Şekil 3.4. OpenCV SGBM parametreleri değişikliği için hazırlanan arayüz.....	50
Şekil 3.5. Sol ve sağ görüntü örneği.....	51
Şekil 3.6. Doğrultulmuş ve kalibre edilmiş gri tonlamalı sol ve sağ görüntü.....	52
Şekil 3.7. Sol eşleme kullanılarak elde edilen görüntü.....	52
Şekil 3.8. Sağ eşleme kullanılarak elde edilen görüntü.....	53
Şekil 3.9. WLS filtre çıktısı derinlik haritalanması.....	54
Şekil 3.10. Nihai farklılık değerleri gri tonlamalı gösterimi.....	55
Şekil 3.11. Farklılık değerleri renklendirilmiş gösterimi.....	55
Şekil 3.12. Sistem algoritması şematik gösterimi.....	56
Şekil 3.13. Test alanında semantik segmentasyon sonuçları.....	57
Şekil 3.14. Semantik segmentason işlem süreleri.....	57
Şekil 3.15. a) Manyetik ve b) optik enkoder iç yapısı.....	59
Şekil 3.16. a) Enkoderli DC motor, b) redüktör ve c) enkoder yapısı.....	59
Şekil 3.17. 12 V DC (doğru akım) motora ait 64 CPR enkoderin A ve B kanalı osiloskop görüntüsü.....	60
Şekil 3.18. Görüntü piksel ayrıştırması şeması.....	63
Şekil 3.19. Gerçek uzaklık ve filtre çıktısı derinlik sonucu karşılaştırması.....	64
Şekil 3.20. Çalışmada tasarlanan ROS sistemi şematik gösterimi.....	65
Şekil 3.21. Kaçınma algoritması.....	66
Şekil 3.22. Diferansiyel sürüş ROS şematik gösterimi.....	67
Şekil 3.23. İki tekerlekli araç görüntüsü ve üst görünüşüne ait koordinat sisteminde değerlerinin gösterilmesi.....	68
Şekil 4.1. Test alanında engellerle alınan sol ve sağ görüntü örneği.....	71
Şekil 4.2. Test alanında engellerle alınan sol ve sağ görüntünün kalibre edilip doğrultulmuş örneği.....	71

Şekil	Sayfa
Şekil 4.3. Test alanına ait derinlik haritası gri tonlamalı gösterimi	72
Şekil 4.4. Colormap Jet renk haritası gösterimi	72
Şekil 4.5. Test alanına ait derinlik haritası Colormap Jet tonlamalı gösterimi	73
Şekil 4.6. Test alanında semantik segmentasyon çıktısı üretimi	73
Şekil 4.7. Test alanında semantik segmentasyon çıktısının maskelenmesi sonucu	74
Şekil 4.8. Otonom robota ait odyometrik verilerin RVIZ programı aracılığıyla görselleştirilmesi.....	75
Şekil 4.9. İlk kat araç yerleşim üst görüntüsü.....	76
Şekil 4.10. İlk kat araç yerleşim alt görüntüsü.....	77
Şekil 4.11. İkinci kat araç yerleşim üst görüntüsü	78
Şekil 4.12. Araç ölçüleri teknik resim görüntüsü	79
Şekil 4.13. Gerçek zamanlı harekete ait kaydedilen video görüntülerinin çerçeve görüntülerinden örnekleri	80

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler	Açıklamalar
A	Amper
Ah	Amper-saat
Cm	Santimetre
GB	Gigabayt
M	Metre
Mah	Miliamper-saat
Mm	Milimetre
MP	Megapiksel
Sn	Saniye
V	Volt
W	Watt
°	Derece
°C	Santigrat derece

Kısaltmalar	Açıklamalar
3B	3 Boyut
AMCL	Adaptive Monte Carlo Localization (Adaptif Monte Carlo Lokalizasyonu)
ANN	Artificial Neural Network (Yapay Sinir Ağları)
BM	Block Matching (Blok Eşleme)
CCD	Charge Coupled Device (Yük Bağlısımlı Aygıt)
CMOS	Complementary Metal Oxide Semiconductor (Bütünleyici Metal Oksit Yarı İletken)
CNN	Convolutional Neural Network (Konvolüsyonel Sinir Ağları)
CPR	Counts Per Revolution (Devir Başına Sayım)
CSI	Camera Serial Interface (Kamera Seri Arayüzü)

Kısaltmalar**Açıklamalar**

DNA	Deoksiribonükleik Asit
DPI	Dot Per Inch (İnç Başına Düşen Nokta)
DSP	Digital Signal Processors (Sayısal İşaret İşlemcisi)
EKF	Extended Kalman Filter (Genişletilmiş Kalman Filtresi)
FCL	Fully Connected Layer (Tam Bağlantı Katmanı)
FCN	Fully Convolutional Networks (Tam Konvolüsyonel Ağlar)
FOV	Field of View (Görüş Alanı)
FPGA	Field Programmable Gate Arrays (Alanda Programlanabilir Kapı Dizileri)
GPU	Graphics Processing Unit (Grafik İşlemci Birimi)
HDMI	High Definition Multimedia Interface (Yüksek Çözünürlüklü Çokluortam Arayüzü)
İHA	İnsansız Hava Aracı
LIDAR	Light Detection and Ranging (Işık Tespiti ve Uzaklık Tayini)
MIT	Massachusetts Institute of Technology (Massachusetts Teknoloji Enstitüsü)
OODA	Observe, Orient, Decide, Act (Gözlemle, Uyumlandır, Karar Ver, Eyleme Geç)
PPI	Pixels per inch (Piksel Hassasiyeti)
RAM	Random Access Memory (Rastgele Erişimli Bellek)
RGB	Red, Green, Blue (kırmızı, yeşil, mavi)
RNN	Recurrent Neural Networks (Tekrarlayan Sinir Ağları)
ROS	Robot Operating System (Robot İşletim Sistemi)
RVIZ	3B ROS Görselleştirme Aracı
SGBM	Semi-Global Block Matching (Yarı-Global Blok Eşleme)
SSH	Secure Shell (Güvenli Veri İletimi için Kriptografik Ağ Protokolü)
SVM	Support Vector Machines (Destek vektör makinesi)
USB	Universal Serial Bus (Evrensel Seri Veriyolu)
XML	Extensible Markup Language (Genişletilebilir İşaretleme Dili)

Kısaltmalar**Açıklamalar****YSA**

Yapay Sinir Ağları

YZ

Yapay Zeka

1. GİRİŞ

Doğadaki modellerin taklit edilerek, bu modellere benzer görevlerde kullanılacak araçlar, mekanizmalar veya sistemler tasarlanması biyomimetik olarak adlandırılmaktadır. Böylelikle doğanın sahip olduğu bu problemlerin üstesinden gelme yeteneği insan için kullanılabilir. Doğadan alınan ilham birçok örneği kapsar ve insanlar da bu örneklerle dahildir. İnsan DNA'sından bilgisayar hafıza sistemine, insan kaslarından robotiğe, kulak zarından kulaklığa gibi insan biyolojisinden ilham alarak tasarlanmış birçok biyomimetik örnek vardır [1]. İnsan gözünün çevreyi algılama modeli de en çok ilgi duyulan ve taklit edilmeye çalışılan sistemlerdendir. Yapay görme sistemleri oluşturmak amacıyla pek çok çalışma gerçekleştirilmiştir. Yine benzer şekilde bu görme sistemlerinden alınan yapay görüntüler üzerinde de insanın karar mekanizmalarının benzerleri taklit edilmeye çalışılmıştır.

Araştırmanın Tanımı ve Literatür Taraması

Bilgisayarlı görü sistemleri insana ait olan bu süreci taklit etmeye çalışırken yaygın olarak nesnelerin tanınması, nesnelerin tespiti, nesne hareketinin tahmini vb. görevleri gerçekleştirmeye çalışmaktadır. Bahse konu görevleri gerçekleştiren sistemler pek çok alanda başarıyla kullanılmaktadır. Bunlara birkaç örnek vermek gerekirse günümüz dünyasında kendi başına karar verebilen medikal sistemler, dost ya da düşman tespiti yapabilen askeri teknolojiler ve insanların yerine sanayide çalışan dev robotlar verilebilir. Bu anlamda bilgisayarlı görü sistemleri son yıllarda teknolojiye meydana gelen büyüme, yarı iletken teknolojisinde meydana gelen gelişmeler ve geliştirilen algoritmalar sebebiyle en çok araştırma gerçekleştirilen alanlardan olmuştur [2].

Görüntülemeli sistemler pek çok farklı bant aralığında kameraları kullanarak görüntüleme yapabilmektedir. Bunlardan en yaygın olanı ise insan görme sistemiyle aynı aralığa sahip olan görünür bant kameralardır. Bu kameralardan alınan üç kanal (RGB) görüntüler birleştirilerek görüş alanındaki sahne sanal olarak üretilebilmektedir. Bu alanda gerçekleştirilen en önemli keşiflerden biri de stereo kameralardır. Stereo kameralar sayesinde çevreye ait derinlik değerlerine ulaşılabilir. Stereo kameralar pasif derinlik tespiti gerçekleştiren sistemlerdir. Çevreden yansıyan ışığı sensörleriyle iki boyutlu olarak algılayarak ikili kamera sistemi sayesinde derinlik tahmininde bulunmaktadırlar. Aktif

sistemler bir yayın yapar ve yansımanın dönüşündeki bilgiden yararlanarak derinlik değerini elde edebilirler [3]. Pasif sistemler ise çevreden topladığı bilgileri işleyerek derinlik kestirimi yapmaya çalışırlar [4]. Elde edilen bu derinlik bilgisi sayesinde otonom araçlar yollarına çıkan engelleri tespit edebilirler, otonom askeri cihazlar hedefleriyle olan mesafelerinin tahminini yapabilirler ve robotlar nesnelerle temasını bu üçüncü boyut bilgisi sayesinde sağlayabilirler [5].

Stereo kameralar yardımıyla derinlik tahmininin aktif sistemlere karşı avantajlı ve dezavantajlı olduğu durumlar bulunmaktadır. Örneğin derinliği tahmin edilen hedef tarafından tespit edilmek istenmediği durumda pasif sistemler avantajlı olmaktadır. İyi aydınlatılmamış, puslu ya da sisli ortamlarda da stereo kameraların alacağı görüntülerin bozulması sebebiyle LIDAR gibi aktif sistemler çok daha iyi performans verebilmektedir. Amaca yönelik seçimler yapılarak bu sistemler pek çok projede kullanılabilmektedir. Otonom araç teknolojilerinde yapay zeka ile LIDAR performansı birleştirilerek çalışmalar gerçekleştirilmektedir [6]. Bunun yanı sıra stereo kameralar kullanılarak da otonom araçlar üzerinde çalışmalar sürdürülmektedir [7]. Stereo görüntüleme sistemleri gerçek dünyayı insan gözüne en benzer şekilde irdeleyen sistemlerdir. Görünür bant bölgesinde alınan görüntüler üzerinde görüntü işleme yöntemleri kullanarak çevreye ait anlamsal çıkarımlar yapılmasına da izin vermektedir. Aktif sensörlerin aksine görüntü alanındaki cisimlerde özellik çıkarımları (renk bilgisi, şekil, sınıf vb.) daha hassas gerçekleştirilebilmektedir. Bu tezin ana konusu olan stereo görüntüleme alanı ve semantik segmentasyon alanında pek çok çalışma gerçekleştirilmiştir.

1989 yılın D. J. Kriegman vd. ve 1998 yılında K.Onoguchi vd. tarafından gerçekleştirilen çalışmalarda mobil araçlar için navigasyon ve yol kestirimi çalışmalarının ilk örnekleri gerçekleştirilmiştir [8, 9].

Tsukuba Üniversitesi'nde, akıllı robot geliştirme laboratuvarında, 2000 yılında Masako Kumano vd. tarafından gerçekleştirilen projede, otonom robotlar için engel aşma problemi stereo kameralar yardımıyla çözülmüştür. Çalışma kapsamında parlaklık değerlerinden faydalanılarak iki kamera yardımıyla derinlik değerine ulaşılmakta ve engel tespiti yapılabilmektedir. Çevresel koşullardan çok fazla etkilenmesine rağmen bu sistemin hızlı çalışabildiği gösterilmiştir [10].

2002 yılında Johns Hopkins üniversitesinde gerçekleştirilen bir çalışmada ise engelden kaçınabilen stereo görüntülemeli bir mobil araç çalışması gerçekleştirilmiştir. Sistem iç mekân testlerinde yolundaki engelleri gerçek zamanlı olarak tespit edebilmektedir ve diğer çalışmalardan hem pozitif hem de negatif engelleri (merdivenler vb.) tespit edebilmesi yönüyle ayrılmaktadır [11].

2004 yılında ise insansı robot QRIO için engelden kaçınma ve yol çıkarımı alanında stereo görüntüler kullanılmış ve bu problemlere hızlı çalışan iyileştirmeler gerçekleştirilmiştir [12].

Stereo görüntüleme sistemleri ile bulanık mantık denetleyicileri kullanarak hedef takibi gerçekleştiren ve engelden kaçınabilen mobil araç sistemleri için de çalışmalar son yıllarda giderek artmıştır [13, 14]. Böylelikle yapay zeka ve stereo algoritmaları bir arada kullanılarak faydalı modeller üretilmeye çalışılmıştır.

2009 yılında bilgisayarlı görü temelli araç belirleme ve takip sistemi gerçekleştirilmiştir. Bu sistem sayesinde sürücülere tehdit oluşturabilecek araçlar tespit edilebilecek ve güvenli sürüş mesafesi otonom olarak korunabilecektir [15].

2012 yılında iç mekanlarda görüntüye dayalı bir engel tanıma sistemi geliştirilmiştir. Bu sistem donanımsal olarak gerçekleştirilmiş ve kamera aracılığıyla alınan görüntülerde engel tanıma işlemi uygulanmıştır. Zemin ve engel arasında ayırım yapılarak engelden kaçınma gerçekleştirilmektedir [16].

Bu sistemler sadece otonom kara araçları için değil hava araçları için de uygulanmaya başlanmıştır. 2018 yılında ise MIT’de gerçekleştirilen bir çalışmada başarılı bir yüksek hızda otonom engelden kaçınma çalışması gerçekleştirilmiştir [17].

2018 yılında gerçek zamanlı stereo kamera sisteminin İHA (İnsansız Hava Aracı)’lar için de başarılı bir şekilde uygulanabileceği gösterilmiştir [18].

Ülkemizde ise bu alanda çalışmalar son yıllarda artmıştır. 2020 yılında görüntü temelli çalışan mobil robot kontrolü ve yol çıkarımı (path planning) algoritmaları üzerine bir çalışma gerçekleştirilmiştir [19].

Yine aynı yıl harita bilgisi olmadan bilgisayarlı görü yardımıyla engellere çarpmadan hareket edebilen otonom gezgin bir robot benzetimi gerçekleştirilmiştir [20].

Bu sisteme benzer olarak ROS yazılımı kullanılarak iç mekanlar için özelleştirilmiş engel tanıyabilen gezgin robot tasarımları yapılmıştır [21, 22].

2021 yılında ise bu tez çalışmasından türetilen Stereo Kamera ve Yapay Zeka Tabanlı Derinlik Algılama Sistemi İle İnsan Mesafe Tespiti çalışması gerçekleştirilmiştir [23].

Araştırmanın Amacı

Bu araştırma iç ortamlar için yapay zeka tabanlı çalışan ve engelden kaçınabilen otonom bir sistem üretimini gerçekleştirmeyi hedeflemektedir. Son yıllarda çok fazla ilerleme kaydedilmiş olan yapay zeka alanının getirdiği avantajlar ile bilgisayarlı görü sistemlerinin özellikleri bu çalışmada birleştirilmiştir. Engelden kaçınabilen otonom araçta görüntülemeli sistemlerden olan stereo kamera kullanılmıştır. Son dönemde sürücüsüz araçlar, keşif robotları, sanayi robotları ve derinlik tespit sistemleri gibi pek çok görevde kullanılan bu sistemin engelden kaçınabilen otonom araçlar geliştirilmesinde de kullanılması hedeflenmiştir. Bu tez kapsamında tamamıyla pasif derinlik algılaması yapılmıştır. Buna ek olarak yapay zeka alanında gerçekleşen gelişmelerden faydalanarak gerçek zamanlı anlamsal bölütleme ile otonom sürüş için ek bir faydalı veri üretilmesi hedeflenmiştir. Gerçekleştirilen literatür taramalarında bu çalışmadaki gibi yapay zeka tabanlı anlamsal bölütleme ile stereo derinlik algılama sistemlerinin gerçek zamanlı olarak bir arada kullanımına benzer bir çalışma bulunamamıştır. Bu alanda bulunan araştırma eksikliğinin de bu çalışmayla giderilmesi ve literatüre katkıda bulunulması amaçlanmaktadır.

Araştırmanın Konusu ve Yöntem

Bu tez projesinde, iç mekanlar için otonom hareket edebilen stereo görüntülemeli bir araç geliştirilmiştir. Tez çalışmasına başlamadan önce gerçekleştirilen mali ve teknik fizibilite çalışmalarında ihtiyaç duyulan yazılımın tüm detayları ana hatlarıyla ortaya koyulmuştur. Teknik anlamda gereksinimler belirlenmiş ve ilgili algoritmalarda çalışmanın yol haritası çıkarımı için analizler gerçekleştirilmiştir. Tasarım, kodlama ve test aşamasında ihtiyaç duyulacak yazılım ve donanım netleştirilerek bu hedefte materyal ve metot seçimi

gerçekleştirilmiştir. Maddi kısıtlar göz önünde bulunularak sistemin gerçekleştirileceği minimum donanım gereksinimleri sağlanması amacıyla mali bir fizibilite çalışması da yapılmıştır. Tez sürecinde hazırlanan bu taslak plan zaman zaman adımlarda yenilemeler yapılarak takip edilmiş ve başarılı bir sistem oluşturulmuştur. Sistem OODA döngüsünü kullanarak engelden kaçınma görevini gerçekleştirmektedir. Engelden kaçınma görevini başarıyla gerçekleştiren bu sistem aynı zamanda yapay zeka yöntemlerini kullanarak hareket edilebilir zemine ait anlamsal bir ayırım gerçekleştirebilmektedir. Engelden kaçınabilen bu teknolojik çalışma sayesinde otonom araçlar bilgisayarlı görü sistemlerini kullanarak çevreye ait değerlendirmelerde bulunabileceklerdir. Bu araç gerek duyulması halinde istenilen ortama ait anlamsal bölütleme gerçekleştirebilecektir. Araç üzerinde aktif derinlik algılama sistemi kullanılmaması sebebiyle çevreye herhangi bir yayın yapılmamaktadır. Lazer, ultrasonik ya da kızılötesi yayınlar yaparak derinlik algılayan sistemlerin bu bantlarda tespiti gerçekleştirilebilmektedir. Bu tez kapsamında gerçekleştirilen araç sadece pasif derinlik tespiti gerçekleştirmesi yönüyle de benzerlerinden ayrılmaktadır.

Çalışmanın 2'nci bölümünde tezde kullanılan materyal ve metotlar hakkında genel bilgilendirmelerde bulunulmuştur. Görüntü işleme, bilgisayarlı görü, segmentasyon, yapay zeka kavramları hakkında genel bilgiler verilmiştir. Tezin temel konularından olan yapay sinir ağları konusu hakkında bilgilendirme ve çalışmada kullanılan SegNet semantik segmentasyonu incelemesi, derinlik algılama konusu ve stereo görüntüleme sistemlerinin temel prensipleri de 2'nci bölümde verilmiştir. Buna ek olarak Robot İşletim Sistemi (ROS) hakkında genel bilgilendirmeler de yine 2'nci bölümde bulunmaktadır. 3'üncü bölümde geliştirilen uygulamalarla ilgili bilgiler verilmiştir. Gerçekleştirilen sisteme ait bulgular ve bulguların irdelenmesi ise 4'üncü bölümde verilmiştir. Son bölüm olan 5'inci bölümde ise sonuçlar ve değerlendirme sunulmuştur. Gelecekte yapılması hedeflenen çalışmalar ve öneriler de 5'inci bölümde verilmiştir.

2. MATERYAL VE METOTLAR

Bu bölümde tez kapsamında kullanılan konular hakkında açıklamalara yer verilmiştir. Görüntü işleme, bilgisayarlı görü, yapay zeka, yapay sinir ağları, derinlik algılama sistemleri, stereo derinlik algılama yöntemi ve robot işletim sistemi konuları hakkında sırasıyla bilgilendirmeler yapılmıştır.

2.1. Görüntü İşleme

Görüntü işleme, görüntü üzerinde hedeflenen farklılıkların (kontrast, parlaklık vb.) bulunması ya da görüntüden istenilen bilgilerin çıkarılması için görüntü üzerinde uygulanan işlemlerin toplamıdır. Görüntü işleme terimi çok geniş bir alana hitap etmekle birlikte hedeflenen iş için çok farklı metotlar gerçekleştirilebilmektedir. Görüntü işlemeyi tam anlamıyla anlayabilmek için görüntü olarak tanımlanan kavramı bilmek gerekmektedir.

Sağlıklı bir insanın gündelik yaşamının sıradan bir parçası olan görme eyleminin arkasında tabii (güneş) ya da suni (led) bir kaynaktan gelen ışınların cisimlerden yansması ve bu ışınların görme organında bazı işlemlerden geçmesi vardır. Göz üzerinde bulunan bazı bölgelerden (saydam tabaka, iris, göz merceği, ağ tabaka vb.) kırılarak geçen bu ışınlar gözün içerisinde bulunan ve sarı leke adı verilen bölgede terslenerek cismin görüntüsünü oluşturur. Bu görüntü sinir hücreleriyle beyne aktarılır ve beyinde tekrar terslenerek cisim algılanmış olur.

Bu karmaşık görüntüleme modelinin becerisi taklit edilerek dış dünyaya aktarılacak istenmiştir. İbn-i Heysem'in gümüş nitratin karardığını keşfetmesiyle ilk adımı atılan ve ondan asırlar sonra Leonardo Da Vinci'nin karanlık odada cisim görüntülerini elde etmeye çalışmasıyla birlikte başlayan fotoğrafçılık alanındaki ilerlemeler ilk başlarda sadece kimyasal yöntemlere dayanmaktaydı. Gümüş nitrat, cıva, iyot, bakır vb. kimyasalların ışık ve kullanılan malzemeler ile etkileşimi kullanılarak cisimlerden gelen görüntüler mercekler yardımıyla bu malzemeler üzerine düşürülmekte ve bazı kimyasal işlemlerden geçirilerek cisimlerin kararmış silüetleri çıkarılmaya çalışılmaktaydı. 1838'de Fransız sanatçı ve kimyacı Louis Daguerre, gerçek anlamdaki ilk fotoğrafı Paris'te çekmiştir. İlk fotoğrafın çekim işleminin gerçekleşmesi 10 dakikadan uzun sürmüştür. Bu fotoğrafta ayaklarını çekim

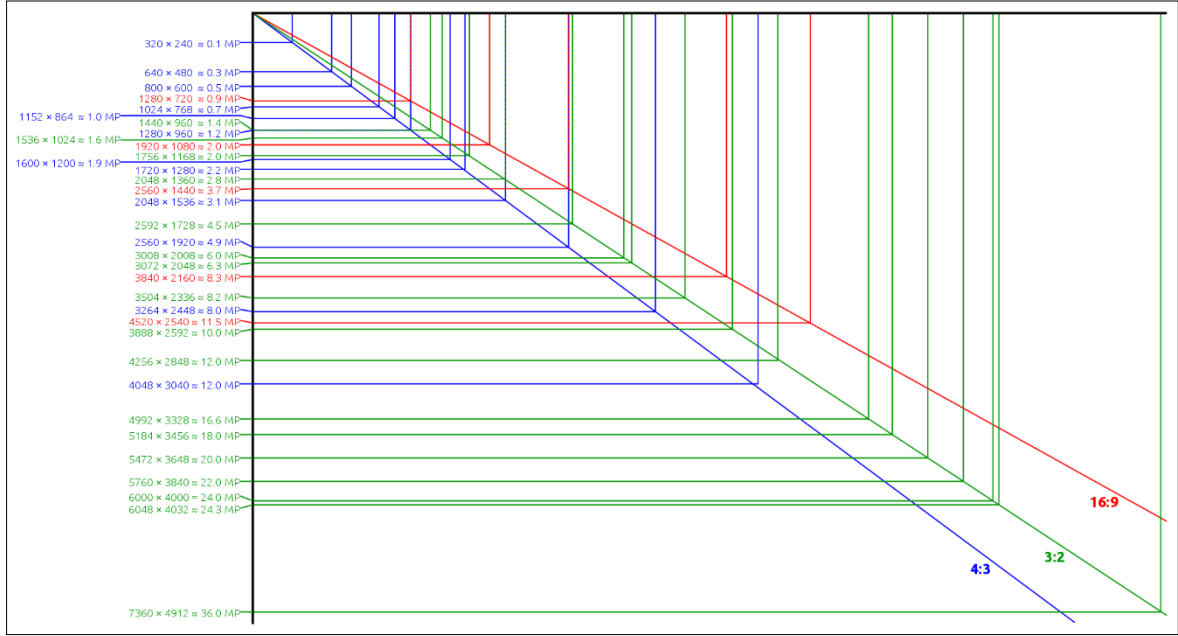
süresince cilalanan bir kişi görüntülenmektedir. Bu fotoğraf bir insana ait çekilmiş ilk fotoğraftır.

Daguerre'in buluşu çığır açıcı bir değer taşıyor olsa da görüntüyü gümüş ve bakır karışımı bir levha üzerine düşürmek zorunda olması sebebiyle ilerleme kaydedememiştir. Fakat daha sonrasında Henry Fox Talbot'un kâğıt üzerine görüntüleri taşıma çalışmaları ve bunların negatiflerinin pozitive çevrilmesi ile görüntüleme tekrar hız kazanarak olumlu bir yola girmiştir. Talbot ayrıca 'fotoğraf' kelimesini de ilk defa kullanarak literatüre kazandıran kişidir. Talbot'tan sonra birçok kişi bu alanda çalışmalar yaparak fotoğrafçılığın gelişimine katkıda bulunmuştur. Örneğin, 1888'de George Eastman'ın 10 poz çekebilen jelatin ruloyu keşfetmiştir. 1870'lerde ise Herman Vogel emülsiyonları spektrumun geri kalanına duyarlı hale getirmenin bir yolu olarak keşfetmiş ve bununla birlikte renkli fotoğrafçılığın önü açılmıştır. 1880'de kullanılmaya başlanan renk çarkının bütün ışınlarına karşı muhteşem duyarlılık sergileyen pankromatik filmler ile fotoğrafçılık daha da yaygınlaşabilmiştir. 19. ve 20. yy.'da merceklerdeki gelişmeler, selüloz filmler ve optikte yaşanan teknolojik gelişmeler ile birlikte günümüzdeki analog fotoğrafçılık seviyesine ulaşılabilmiştir.

Fotoğrafçılık alanında yapılan birçok çalışma ile bilim insanları geleneksel fotoğraf makinelerinde olduğu gibi fotoğraf filmlerini kullanmak yerine fotoğrafları sayısal olarak saklamanın bir yolunu keşfedebildiler. Işığın, filmler üzerine değil de sensörlerin üzerine düşürülmesiyle oluşturulan sayısal veriler işlemci üzerindeki RAM ya da ROM bellekler üzerinde saklanabilir hale geldi. Bu çalışmaların sonucunda 1975 yılında Kodak'ta mühendis olarak çalışan Steven Sasson ve çalışma arkadaşları tarafından ilk dijital fotoğraf makinesi icat edilmiş oldu. Dijital fotoğraf makinelerinde cisimlerden gelen yansımalar kırmızı, yeşil ve mavi renklere duyarlı minik hücrelerden geçerek cisme ait görüntü bilgileri elektriksel değerler olarak elde edilmekte ve işlemci birimi tarafından bu farklı değerler kaydedilmektedir. Görüntü tekrar elde edilmek istendiğinde bu değerlere ait pikseller benzer şekilde çıkış biriminde renklendirilerek kullanıcıya sunulabilmektedir. Devrim niteliğindeki bu gelişme sayesinde sayısal görüntü işleme alanındaki çalışmalar hızlanmıştır.

Dijital görüntü de aslında bir çeşit sinyaldir. Sayısal sinyallerin analog sinyallere göre analizleri daha kolaydır, gösterimleri ayrık zamanlı olmak zorundadır, kolaylıkla depolanabilirler ve gürültü etkisi daha azdır. Ancak bu avantajlarının yanında kayıt esnasında sinyal örneklendirilerek elde edildiği için daha az doğruluk değerleri göstermeleri

olağandır. Dolayısıyla sayısal bir görüntüde örnekleme artıran piksel sayısı görüntünün büyüklüğünü etkilediği gibi detaylarını da etkilemektedir. Ancak tek başına piksel sayısının kaliteyi artıran bir faktör olduğunu düşünmek de yanıltıcıdır. Dijital bir görüntü iki boyutlu bir düzleme oturtulduğunda görüntüyü meydana getiren her bir elemana piksel ismi verilmektedir. Dijital bir görüntünün piksel sayıları için yaygın kullanılan çözünürlük değerleri ve megapiksel (MP) ölçeğinde içerdiği piksel sayılarına ait diyagram Şekil 2.1’de verilmiştir.

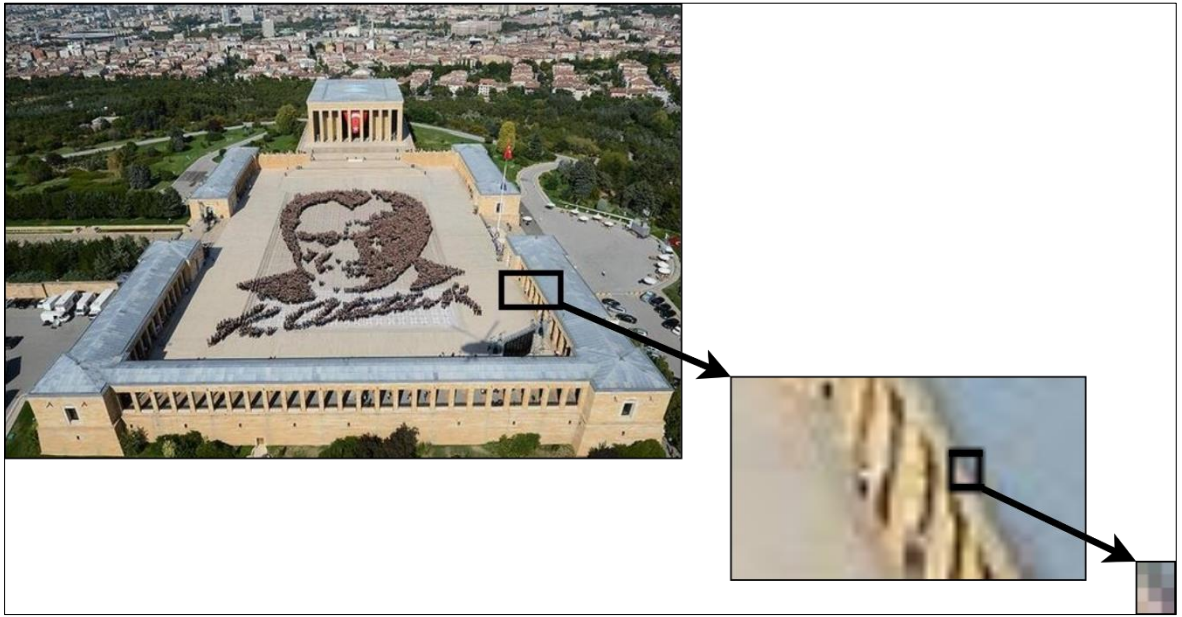


Şekil 2.1. Yaygın olarak kullanılan çözünürlük değerleri ve megapiksel değerleri [24]

Piksel sayıları gösterilirken yatayda ve dikeydeki piksel sayılarının çarpımı kullanılır. Örnek vermek gerekirse 2048 x 1536 piksellik bir görüntü içerisinde 3 145 728 adet piksel bulunur ve çoğunlukla 3,2 megapiksel şeklinde ifade edilir. Burada 2048 yatay eksenindeki dizili piksel sayısını, 1536 ise yukarıdan aşağıya dizili olan sütundaki piksel sayısını göstermektedir. Pikseller çözünürlüğün de bir ifadesi olarak kabul edildiği için yaygın olarak isimlendirmelerde kullanılmaktadır. Çözünürlükler arttıkça pikseller küçülecekleri için bu piksel hassasiyetini artırır.

Günümüzde görüntüleme cihazlarında kullanılmakta olan DPI (inç başına düşen nokta sayısı) ya da PPI (inç başına düşen piksel sayısı) gibi terimlerle 1 inch üzerinde kaç adet piksel bulunduğu bilgisi kullanıcıya verilerek bilgilendirme sağlanabilir. Piksellere ait görüntü alma teknolojisinde daha çok CCD ve CMOS gibi teknolojiler kullanılmaktadır. Her

iki teknoloji de foto duyarlı teknolojiler olmasına rağmen CCD sensörler uzun yıllardır daha yaygın bir şekilde kullanılmaktadır. CCD sensörlerin, CMOS teknolojisinden daha yaygın kullanılmasının asıl sebepleri arasında daha iyi çözünürlükte sonuçlar verebilmesini sayabiliriz. Fakat son yıllarda CMOS teknolojisinin kaydettiği ilerlemeler sayesinde CCD teknolojisini yakalamasına ramak kalmıştır. Genel değerlendirme yapıldığında ise her ikisinin de avantajlı ve dezavantajlı olduğu durumlar bulunmaktadır. Her iki teknoloji de sayısal görüntüyü RGB duyarlı kanallar aracılığıyla alarak son görüntüye ait piksel değerlerini oluşturmaktadır. Şekil 2.2’de bir görüntüye ait piksel gösterilmiştir.



Şekil 2.2. Görüntü üzerinde piksel gösterimi

Belirli piksel değerlerinin bir araya gelmesi ile oluşturulan görüntülerde bu iki boyutlu (renkli görsellerde 3’üncü boyut olarak RGB matrisleri eklenebilir) matrisler üzerinde belirli bir amaç için işlemler gerçekleştirilir. Bu amaçlar görüntünün kullanım alanına göre değişiklik göstermektedir. Görüntü işlemenin en çok kullanıldığı alanlar aşağıdaki gibi sıralanabilir:

- Görüntü zenginleştirme
- Gürültü azaltma ve görüntü restorasyonu
- Medikal görüntüleme
- Uzaktan algılama sistemleri
- Mikroskopik görüntüleme

- Örüntü (Pattern) tanıma
- Video görüntülerinin işlenmesi
- Bilgisayarlı görü sistemleri

Bilgisayarlı görü sistemleri ile görüntülerin anlamlandırılması amacıyla görüntü işleme teknikleri kullanılır. Tez kapsamında yapılan çalışmanın bilgisayarlı görü sistemleri alanına hizmet edeceği düşünülmektedir.

2.2. Bilgisayarlı Görü

Tarih sahnesine çıkışından bu yana insanın kurduğu en büyük hayallerden birisi, kendisi gibi davranabilen ve kendi yeteneklerini taklit edebilen bir makine icat etmek olmuştur. Bilgisayarın keşfinden sonra bilgisayarlara kazandırılan hesap yapabilme gibi yetilerle insanoğlunun bu alandaki heyecanı yeniden canlanmıştır. Bu tarz işlemleri yapabilen bir makinenin insanlara has olan görme, işitme, duyma ve bunları anlayarak mantık süzgecinden geçirme işlemlerini de yapabileceği düşünülmeye başlanmıştır. Bu sebeple bilgisayarların tıpkı bir insan gibi dış dünyayı kavrayabilmesi ve elde ettiği bu bilgileri işleyerek doğru çıkarımlar yapması amaçlanmıştır. İnsanlık, bu hayaline ulaşmak için geliştirdiği bilgisayarlı görü, bilgisayarları görsel dünyayı anlayabilmesi için eğitmeye çalışan disiplinlerarası bir kavramdır ve bilgisayarlara normal bir insanın görü mekanizmasının sahip olduğu kabiliyetleri kazandırmayı hedeflemektedir. Sayısal görüntüler ve videolar üzerinde işlemler yapılarak istenen bilgiyi üretmek amaçlı görüntü oluşturma, işleme, analiz etme ve anlamlı hale getirme işlemleri uygulanmaktadır.

Bilgisayarlı görü sistemleri tek bir görüntüden veya bir görüntü dizisinden işimize yaracak bilginin bulunduğu alanı bularak bu bilginin kazanılmasını, analizini ve anlaşılmasını hedeflemektedir. Bu sebeple birden fazla disiplini bir araya getirerek kendiliğinden bu sonuca ulaşmaya çalışılmaktadır. Bilgisayarlı görünün ilgi alanları şunlardır:

Katı Hal Fiziği

Bilgisayarlı görü sistemleri tipik olarak görünür ya da kızılötesi ışık dalga boyundaki elektromanyetik dalgaları algılayacak biçimde çalışan sensörler ile tasarlanmıştır. Bu

kapsamda görüntünün oluşturulmasını algılamak amacıyla optik, kuantum fiziği, yarı iletkenler vb. alanlarda bilgi sahibi olmak gerekir.

Yapay Zeka

Bilgisayarlı görü sistemleri bir insanın zekasını taklit etme ve topladıkları bilgileri anlamlandırma aşamasında yapay zeka alanından faydalanmaktadır. Yapay zekanın alt dallarından olan yapay sinir ağları, bulanık mantık ve genetik algoritmalar gibi sistemler kullanılarak alınan görüntülerden faydalı bilgiler elde edilir.

Sinyal İşleme

Sayısal işaretlerin anlamlandırılmasını hedefleyen bilgisayarlı görünün bir başka ilgi alanı da sinyal işlemedir. Çok boyutlu sinyaller olarak elde edilen sayısal görüntüler bu alan dahilinde pek çok farklı yöntem kullanarak işlenmektedirler.

Robotik

İnsanın yerine geçebilecek ve onun eylemlerini taklit edebilecek makineler yapmayı amaçlayan robotik bilim dalı bilgisayarlı görü ile iç içe geçmiş bir biçimde çalışmaktadır. Çevre ile etkileşimlerini sensörler ve bilgisayarlı görü ile elde eden robotlar, bu verileri işe yarar şekilde anlamlandırmaktadır.

Diğer alanlar

Bilgisayarlı görü sistemleri matematik, geometri, istatistik, optimizasyon, veri bilimi vb. pek çok alanla ilgili olarak çalışmaktadır. Bilgisayarlı görü aynı zamanda pek çok farklı mühendislik alanıyla da uygulama sahasına bağlı olarak çalışmaktadır [25].

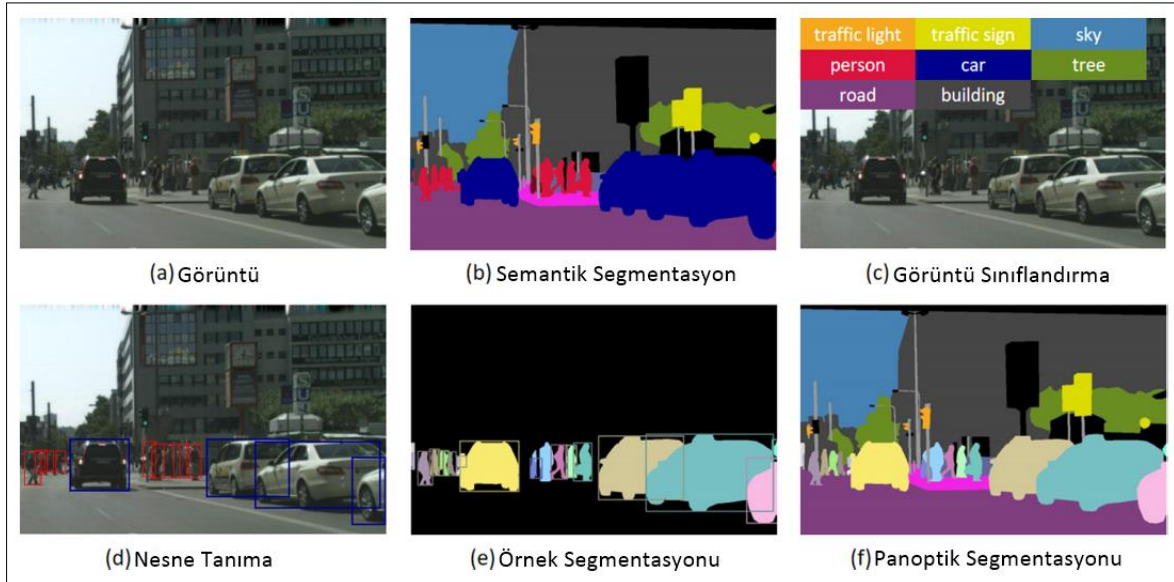
Teknolojinin son dönemlerde hızla gelişmesiyle doğru orantılı olarak maliyetlerin de düşmesiyle birlikte bilgisayarlı görü sistemleri dünya genelinde efektif olarak birçok farklı alanda kullanılmaya başlanmıştır. Bu alanlara örnek olarak aşağıdakiler sayılabilir;

- Nesne ya da canlı tanıma sistemleri,
- Karakter ya da örüntü tanıma sistemleri,

- Robotlarda navigasyon sistemleri,
- Robotların dış dünya ile etkileşimlerinin kurulması,
- Endüstriyel kontrol sistemlerinde süreçlerin izlenmesi ya da işlenmesi,
- İnsanların bilgisayarlarla aracısız etkileşimi,
- Sağlık sektörü,
- Makine görüşü sistemleri,
- Askeri uygulamalar,
- Otonom araçlar,
- Sahnelerin 3 boyutlu modellenmesi,
- NASA'nın dünya dışı görevleri vb.

2.3. Segmentasyon

Görüntü segmentasyonu ya da diğer adıyla görüntü bölütleme, benzer sınıflara ya da etiketlere ait olan piksellerin bir araya getirilerek gruplandırılması işlemidir. Bu işlem sayesinde akıllı sistemler ve robotik alanında bu piksellerden faydalanılarak karar işlemleri gerçekleştirilmektedir [26]. Görüntü bölütleme 3 ana başlık altında gruplandırılır bunlar; semantik segmentasyon, örnek segmentasyonu ve panoptik segmentasyondur.



Şekil 2.3. a) Şehir görüntüsüne ait b) semantik, e) örnek, f) panoptik segmentasyon, c) görüntü sınıflandırma örneği ve d) nesne tanıma örneği gösterimi [27]

Şekil 2.3.b’de görülebileceği gibi Şekil 2.3.a’daki görüntüye anlamsal (semantik) bölütleme uygulanarak kişi, araba, yol, bina, ağaç, trafik ışıkları, trafik levhaları, ağaçlar ve gökyüzü sınıflarına ait olan pikseller gruplandırılmış ve farklı renklerle renklendirilerek görüntülenmeleri sağlanmıştır. Örnek segmentasyonunda (instance segmentation) bölütleme işlemi her bir nesneyi birbirinden ayırt etmeyi amaçlamaktadır (Şekil 2.3.e). Panoptik segmentasyon ise bu iki segmentasyonun özelliklerinin birleştirilmesi ile oluşturulmuştur (Şekil 2.3.f). Semantik segmentasyon aynı sınıfa dahil olan nesneleri benzer şekilde bölütlemeye tabi tutar, örnek segmentasyonunda ise bunlar farklı gruplandırılırlar. Yine Şekil 2.3.b ve Şekil 2.3.e görüntüleri incelendiğinde de görülebileceği gibi araba sınıfına ait farklı arabalar örnek segmentasyonunda farklı renk kodları ile sınıflandırılmışlardır. Anlamsal bölütlemelerde ise nesnelerin birbirinden farklı olmaları bir değişikliğe yol açmamaktadır.

Semantik segmentasyon piksel bazlı kategorileştirme özelliği sayesinde pek çok uygulamada kullanılabilmektedir. Buna örnek vermek gerekirse otonom araçlarda çevreye ait pikselleri sınıflandırmak [28], yayaları ayırt edebilmek [29] ve hatta yüzeydeki eksikliklerin tanımlanması (defect detection) gibi özel görevlerde semantik segmentasyon kullanılmaktadır [30]. Önceleri SVM (Support Vector Machines) ve rastgele ağaçlar (random forest) gibi geleneksel öğrenme metodları kullanılarak anlamsal bölütleme gerçekleştirilmekteydi. Son dönemde ise semantik segmentasyon alanındaki çalışmalar yapay zeka uygulamaları üzerine yoğunlaşmıştır. FCN (Fully Convolutional Networks) öncülüğünde başlayan bu yapay zeka çalışmaları geleneksel yöntemlere oranla dikkate değer bir doğruluk iyileştirmesi sağlamıştır [31].

2.4. Yapay Zeka

Yapay zekanın literatürde birden fazla tanımı vardır. Yani yapay zekanın belirlenmiş ve üzerinde fikir birliğine varılmış tek bir tanımı bulunmamaktadır. Bu bilim dalını tam anlamıyla kavrayabilmek için birkaç terime ait felsefi cevabı bilmek ve anlamak gerekmektedir. Bunlardan en önemli ikisi akıl ve zeka terimleridir.

Akıl (intellect); düşünme, anlama, idrak etme ve karar verme gibi yetenekleri bir arada bulundurur. Akıl sayesinde muhakeme ve bilgi elde etme sağlanır. Akıl doğuştan gelmez, sonradan kazanılır. Sabit ve durağan bir kavram değildir, zamanla değişim gösterir.

Zeka (intelligence); gerçekleri algılama, yargılama ve sonuç çıkarma yeteneklerinin tamamını niteler. Zeka doğuştan gelir. Ancak zeka da akıl gibi durağanlık göstermez ve zamanla değişim gösterir. Çalışılarak, eğitilerek veya deneyimlere dayalı olarak geliştirilebilir. İlk kez karşılaşılan bir olaya uyum yeteneği, öğrenme, anlama, analiz yeteneği ve düşünme becerisi zekaya bağlı olarak çalışır.

Bu kavramların ışığında yapay zeka elektronik donanımlardan, yazılımlardan ve bilgisayarlardan faydalanarak insana özgü niteliklerden olan karar verme, anlam çıkarma, öğrenme ve deneyimleme işlevlerini yerine getirme olarak tanımlanabilir. Gerçekleştirilen bu yapı bilinç, akıl ve duygu gibi insana ait özelliklere sahip olamayacak ancak bunları taklit edebilecektir.

Yapay zeka felsefesini ilk ortaya atan kişi Alan Turing'dir. 1936 yılında yayınladığı "Saptama Problemi Hakkında Bir Uygulamayla Birlikte Hesaplanabilir Sayılar" ("On computable numbers, with an application to the Entscheidungsproblem") isimli bir makalesinde ileride Turing makinesi ismiyle ünlenecek olan teorik ve matematiksel temellere dayalı sanal bir makineyi kuramsal olarak icat etmiştir [32].

Alan Turing'e göre her türlü matematiksel hesap, oluşturulan bu sanal makineyle birlikte çözülebilirdir. 1950 yılında "Mind" adlı felsefe dergisinin Ağustos sayısında yayınlanan Alan Turing'e ait "Hesaplama Mekanizması ve Zeka" ("Computing Machinery and Intelligence") isimli ikinci makalesi ile birlikte yapay zeka hakkında pek çok tartışmalı felsefi konuya açıklık getirilmiştir [33].

Alan Turing, Turing testi olarak adlandırılan bir bilgisayarın insanlarla aynı zihinsel yetiye sahip olup olmadığını ölçen bir test geliştirmiştir. Bu teste göre gönüllü olarak teste katılan bir kişi, makine sorgulayıcının görüş alanının dışında bir alana saklanır. Sorgulayıcı olarak seçilen kişi ise yalnızca soru sormak suretiyle karşısındaki kişinin insan mı yoksa makine mi olduğunu anlamaya çalışır. Sorgulayıcının sorduğu sorular ve aldığı yanıtlar oluşturulan bir klavye sistemi aracılığıyla bir ekranda verilir; dolayısıyla sorgulayıcı karşısındaki kişinin sesini duymaz. Sorgulayıcıya soru-cevap oturumunda elde ettiği bilgiler dışında da bir bilgilendirme yapılmaz. Diziler halinde tekrar eden bu soru-cevaplar sonucunda sorgulayıcı, karşısındakinin insan mı yoksa makine mi olduğunu anlayamadığı zaman makine Turing testini geçmiş kabul edilir. Bu çalışmalar bu alanda yapılan ilk belirleyici çalışmalar

olmuştur. Şu an zeki bir makinenin sadece karşısındaki insan tarafından makine olarak saptanamadığı yani Turing testini başarı ile geçebildiği bilinmektedir. Bu alandaki çalışmalar 1950 sonrasında Shannon, Turing, Mc Carthy, Rochester, Hebb, McCulloch ve Pitts gibi araştırmacıların çalışmalarıyla hız kazanmıştır. 1956 yılında Dartmouth’da gerçekleştirilen çalıştayda ilk kez “Yapay Zeka” terimi kullanılmıştır. Bu kapsamda pek çok yapay zeka tekniği geliştirilmiş ve uygulanmıştır. Bunlara örnek olarak,

- Tavlama Benzetimi (Simulated Annealing),
- Uzman sistemler (Expert Systems),
- Konuşma Tanıma (Speech Recognition),
- Bulanık Mantık (Fuzzy Logic),
- Genetik algoritmalar (Genetic Algorithms),
- Kaotik Modelleme,
- Yapay Sinir Ağları (ANN: Artificial Neural Networks),
- Farklı tekniklerin bir arada kullanıldığı Hibrid sistemler verilebilir.

Görüldüğü üzere farklı birçok yapay zeka tekniği olmasına rağmen, bu çalışmada yalnızca insan sinir sistemini temel alıp onu taklit etmeyi amaçlayan yapay sinir ağları yöntemi ile öğrenme kullanılmıştır.

2.5. Yapay Sinir Ağları

Sibernetik kelimesinin kökü sevk, idare, dümenci gibi anlamlara gelen yunanca “kybernétes” ve latince “gubernare” kelimelerine dayanmaktadır. Sibernetik, dış çevrenin ihtiyaçlarına ve değişimlerine göre kendi kendini kontrol edebilen ve herhangi bir insani müdahaleye ihtiyaç duymadan çalışabilen mekanizmaları inceleyen, özetle kontrol ve yönetme olarak tanımlanan disiplinlerarası bir bilim dalıdır [34]. Sibernetik, canlı ve cansız sistemleri inceler ve bu sistemlerin arasındaki benzerliklerden faydalanmaya çalışır. İnsan vücudu karmaşık bir sistem olarak sibernetiğin ana çalışma alanları arasına girmiştir. Sibernetiğin limitleri çok geniştir, ele aldığı sistemler mekanik, fiziksel, biyolojik, düşünsel ve sosyal sistemler olabilir.

Yapay sinir ağlarının modellenmesi de aslında böyle bir çalışmanın ürünüdür. İnsan beyninin çalışma mekanizması modellenerek sinir hücrelerinin çalışma yapısı elektronik

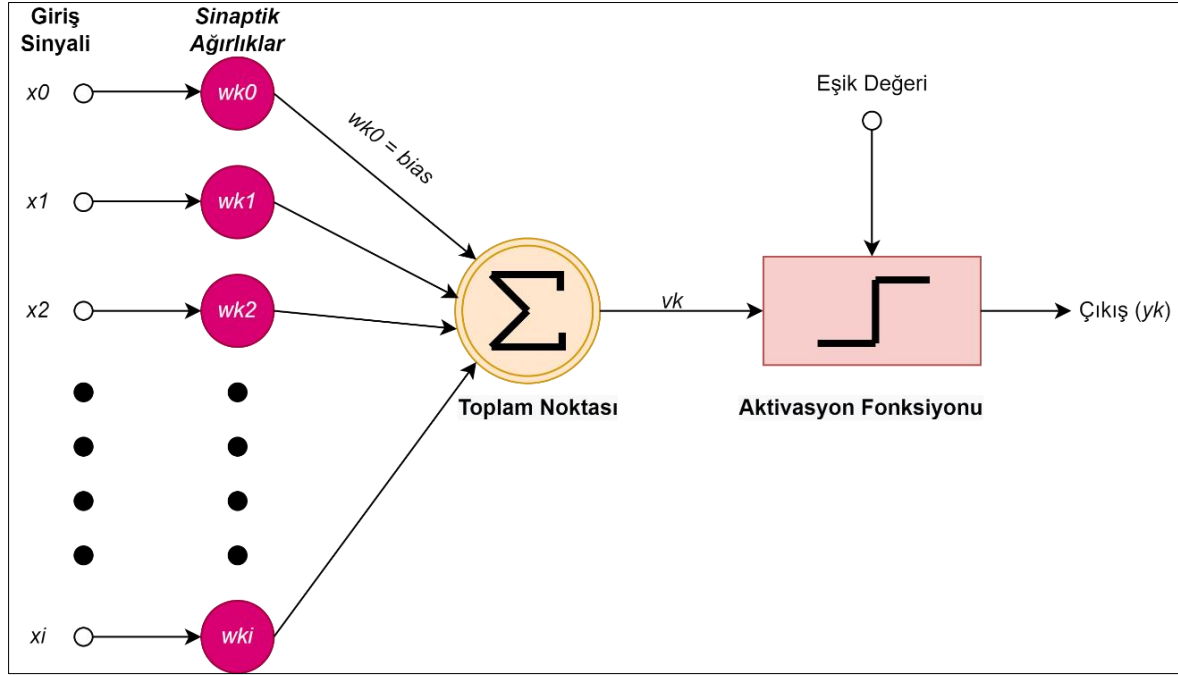
olarak taklit edilmeye çalışılmaktadır. Bu oluşturulan sistemler matematiksel tasarımları sayesinde insanda olduğu gibi kendi kendine öğrenme, değerlendirme yapabilme ve karar verebilme gibi ana görevleri gerçekleştirmeye çalışır.

Yapay sinir ağları insan beyninin temel özelliklerinden olan yeni bilgileri keşfetme, üretme ve oluşturma gibi işlemleri herhangi bir yardıma ihtiyaç duymadan kendiliğinden gerçekleştirebilecek şekilde geliştirilen bilgisayar sistemleridir. Yapay sinir ağları tasarlanırken insan beyni örnek olarak alınmış ve öğrenme süreci matematiksel olarak modellenerek benzetilmeye çalışılmıştır. Yapay sinir ağları diğer bir deyişle biyolojik sinir ağlarının yapısını, belleme, anımsama ve genelleme yeteneklerini taklit eder. Bu alanda 1940 sonrası dönemde başlayan araştırmalar sonucunda pek çok ilerleme kaydedilmesine rağmen ilk temel algoritma üreten çalışma 1943 senesinde gerçekleştirilmiştir. 1943 senesinde McCulloch ve Pitts tarafından yayınlanan makale ile yapay sinir ağlarının temeli atılmıştır. 1949 yılında ise Donald Hebb'in "Organization of Behavior" adlı kitabında "Hebbian öğrenme" adlı yöntemden bahsedilmiştir. Hebb kuralına göre yapay sinir ağları ile öğrenme ve uyum sağlama arasında doğrusal bir uyum olduğu açıkça ortaya konulmuştur. 1951 yılında SNARC isimli yapay sinir ağı temelli çalışan bilgisayar MIT'de Minsky ve Edmonds tarafından üretilebilmiştir.

Farley ve Clark "Calculators" ismini verdikleri hesaplayıcı makineleri Hebbian Ağını simüle etmek için ilk defa kullanmayı başarmışlardır [35]. Bu ilk adımla başlayan hesaplayıcı modeller Rochester, Holland, Habit ve Duda gibi alanında öncü kişiler tarafından geliştirilmeye devam etmiştir. 1957 yılında Perceptron'un Frank Rosenblatt tarafından keşfinden sonra yapay sinir ağı alanındaki çalışmalar hız kazanmıştır. 1958 yılında Rosenblatt ve 1961 yılında Caianiello tarafından adaptif tepki üretme kavramı üzerine çalışmalar yapılmıştır. Rosenblatt'ın geliştirdiği algılayıcı model (perceptron), yapay sinir ağlarının gelişim sürecinde bir dönüm noktası olmuş ve çok katmanlı algılayıcıların geliştirilmesine imkân sağlamıştır.

1960 yılında Widrow ve Hoff tarafından Adaptive Linear Neuron (ADALINE) tek katmalı yapay sinir ağı modeli geliştirilmiştir [36]. Adaptif öğrenmenin temelini oluşturan bu sistem daha sonrasında geliştirilerek MADALINE (Many ADALINE) sisteminin temelini oluşturmuştur.

Marvin Minsky ve Seymour Papert tarafından 1969 yılında yayınlanan “Perceptrons” isimli kitapta, doğrusal olmayan problemlerde yapay sinir ağlarının fayda sağlamadığı eleştirisinde bulunmuşlardır. Doğrusal olmayan lojik birim olan XOR probleminin perceptron tarafından çözülemeyeceğini örnek göstermişlerdir. Bu eleştiriler karşısında yapay sinir ağlarının teorik olarak gelişimi bir süreliğine duraklama devrine girmiştir. 1972 yılında Kohonen ve Anderson yapay sinir ağlarında öğretmensiz öğrenme kurallarını oluşturan çalışmalar gerçekleştirmişlerdir. Grosberg ve Carpenter Adaptif Rezonans Teorisi (ART)’ni geliştirerek denetimsiz öğrenme konusunda döneminin kompleks bir sistemini ortaya koymuşlardır. 1982 ve 1984 yıllarında Hopfield’in ortaya koyduğu çalışmalar ile popüler problemlerden olan gezgin satıcı probleminin çözülebileceği gösterilmiştir. Bu ağ sistemi temel alınarak ilerleyen dönemde Boltzman makinesi üretilmiştir. 1982 yılında Rumelhart ve arkadaşları çok katmanlı algılayıcılar (multilayer perceptron) hakkındaki çalışmaları sayesinde daha önce standart perceptronların çözemediği XOR problemi gibi doğrusal olmayan problemler çözüme kavuşturulmuştur ve bu sayede de Hopfield ağlarının kısıtlamalarından kurtulunması sağlanılmıştır. Bundan dolayı çok katmanlı algılayıcıların üretilmesi yapay sinir ağları için devrim niteliğinde bir gelişmedir. 1988 yılında Broomhead ve Lowe tarafından radyal tabanlı fonksiyonlar modeli ve 1990 yılında Specht tarafından geliştirilen probalistik ağlar ve genel regrasyon ağları ile bu alanda kapsamlı gelişmeler gerçekleştirilmiştir. Tabii ki yapay sinir ağları üzerinde çalışmalar ve geliştirmeler halen devam etmektedir. Bu çalışmaların çoğunluğunun geçmişte teorik olarak sunulmasına rağmen günümüzde gelişen paralel işlem yetenekli donanımlar sayesinde, pek çok çalışma pratik uygulamalara da dönüştürülmüştür.

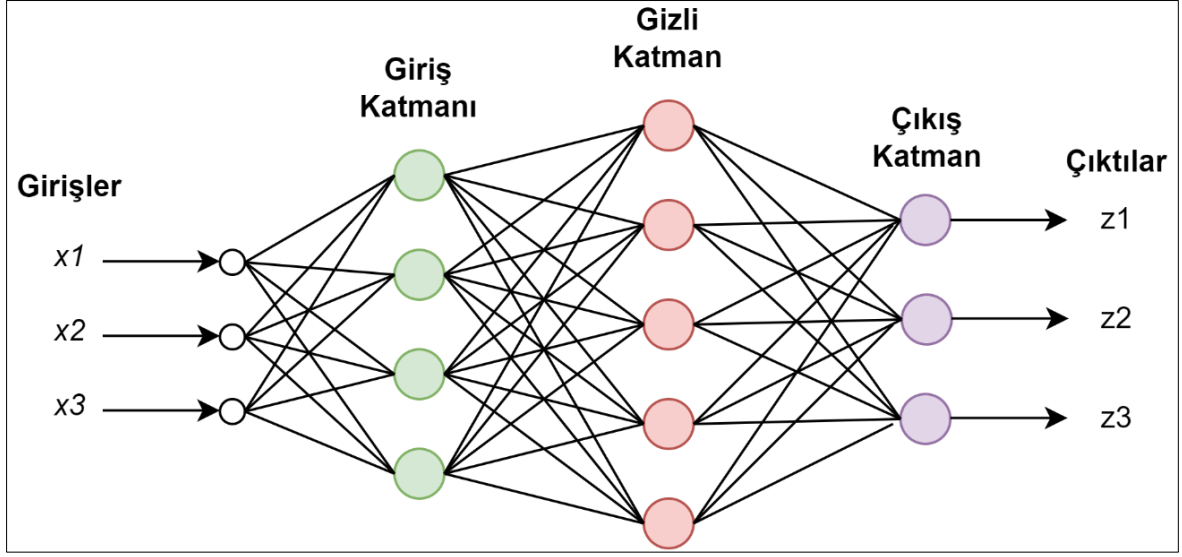


Şekil 2.4. Yapay sinir hücresinin şematik gösterimi

Şekil 2.4'te bir yapay sinir hücresinin şeması verilmiştir. Tek katmanlı bir yapay sinir ağının en temel üyesine perceptron (algılayıcı) ismi verilir. Şekil 2.4'te görüldüğü gibi giriş değerleri, ağırlıklar ve sapmalar, toplama noktası ve aktivasyon fonksiyonundan oluşan bir yapısı bulunmaktadır. Matematiksel olarak ifadesi Eş. 2.1'deki gibidir.

$$f(x) = wx + b \quad (2.1)$$

Eş. 2.1'de $f(x)$ çıkış değeri, b bias, w ağırlık ve x ise girdileri ifade etmektedir. Şekil 2.4'te gösterilen eşik değeri yardımıyla formül işletilir. Eşik değerlerine bağlı olarak sapma ve ağırlıklar güncellenerek girdilere ait benzer çıktıları yapan bir sinir ağı eğitilmiş olur.



Şekil 2.5. Yapay sinir ağı gösterimi

Birden çok girdiye karşılık tek sinir hücresi kullanan algılayıcı (perceptron) yapısı problemler büyüdükçe yetersiz kalmıştır. İlk olarak XOR probleminin çözümü için ortaya atılan çok katmanlı perceptron (Multi-layer Perceptron) yapısı oluşturularak birden çok girdi birden çok nöron için girdi olarak beslenmiş ve çıktılar gizli katmanlarda yine başka sinir hücrelerinden geçirilerek çıkış katmanına ulaştırılmıştır. Her katman birbirinin girdisini üretir ve bilgi çıkış katmanına ulaşır. Çıkış katmanındaki nöron sayısı çıkış sayısını belirlemektedir. Bu yapıda en az 1 adet gizli katman bulunması gerekir. Katmanların sayısı ve katmanlardaki nöron sayıları da çözülmek istenen probleme göre belirlenmektedir. Bu sistemde her nöron kendi aktivasyon fonksiyonuna sahiptir. Bu fonksiyonlar Çizelge 2.1 ile gösterildiği gibi doğrusal, basamak, sigmoid, hiperbolik, ReLu, Sızıntılı ReLu (Leaky ReLu) gibi fonksiyonlar olabilir [37].

Çizelge 2.1. Aktivasyon fonksiyonları ve matematiksel ifadeleri

Aktivasyon Fonksiyonu İsmi:	Matematiksel İfadesi:	Tanımlı Olduğu Aralık:
Doğrusal Fonksiyon	$f(x) = x$	$(-\infty, \infty)$
Basamak Fonksiyon	$f(x) = \begin{cases} 0 & \text{için } x < 0 \\ 1 & \text{için } x \geq 0 \end{cases}$	$\{0, 1\}$
Sigmoid Fonksiyon	$f(x) = \sigma(x) = \frac{1}{1 + e^{-x}}$	$(0, 1)$
Hiperbolik Tanjant Fonksiyonu	$f(x) = \tanh(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$	$(-1, 1)$
ReLU	$f(x) = \begin{cases} 0 & \text{için } x < 0 \\ x & \text{için } x \geq 0 \end{cases}$	$[0, \infty)$
Leaky (Sızıntılı) ReLU	$f(x) = \begin{cases} 0.01 & \text{için } x < 0 \\ x & \text{için } x \geq 0 \end{cases}$	$(-\infty, \infty)$

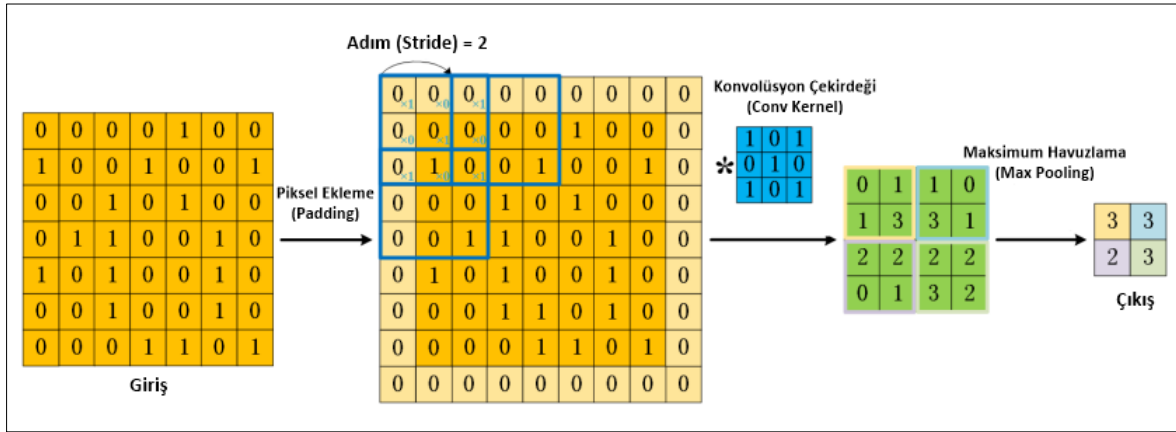
Hangi katmanda hangi aktivasyon fonksiyonunun kullanılacağı da yine problem özelinde avantajı ve dezavantajı göz önünde tutularak belirlenmelidir. Mesela çok sınıflayıcı modellerde hiperbolik tanjant gibi fonksiyonlar kullanılırken ya da daha yavaş ancak kararlı öğrenmeyle sigmoid ya da derin ağlarda işlem yükü sebebiyle ReLU tercih edilebilir [38].

Bu anlamda derin öğrenme üç ana başlık altında gruplandırılabilir. Bunlar;

- 1- Yapay Sinir Ağları (Artificial Neural Networks – ANN)
- 2- Konvolüsyonel Sinir Ağları (Convolution Neural Networks – CNN)
- 3- Tekrarlayan Sinir Ağları (Recurrent Neural Networks – RNN)’dir.

Yapay Sinir Ağları (YSA) Şekil 2.5’te verildiği gibi birden çok perceptron tarafından oluşturulmuş bir sinir ağı modelidir. Bu ağlar genel olarak ileri beslemeli sinir ağları olarak da anılmaktadır. Tekrarlayan Sinir Ağları, özyinelemeli sinir ağları olarak da adlandırılmaktadır. Diğer sinir ağlarında girdilerin birbirleri üzerlerinde etkileri yoktur. Girdinin ağına hangi sırada girdiği önemsizdir. Ancak RNN mimarisinde girdilerin zaman bilgileri bir sonraki ağına girdisinin sonucunu etkilemektedir. Buna örnek olarak bir insanın bir kitap üzerindeki cümleyi okurken sırayla okuması ve kelimeleri bu sıra dahilinde anlamlandırması verilebilir. Her kelimenin bir girdi olduğu düşünülürse her kelimenin sırası

da bir sonraki kelimenin katacağı anlamı etkileyecektir. RNN mimarisi hafızalı bir sinir ağı yapısıdır; böylelikle bir önceki veriye ait bilgileri tutarak bir sonraki girdinin sonucunu değiştirmektedir. Dil üretimi gibi zamana dayalı tekrarlayan problemlerin çözümünde sıkça kullanılan bir yöntemdir. Evrişimsel ya da konvolüsyonel sinir ağları ise farklı filtreleri kullanarak görüntülerdeki özelliklerin (features) çıkarımını yapan ve onları sınıflandıran bir mimaridir. CNN yapısı konvolüsyon, havuzlama (pooling), normalizasyon (normalization) ve tam bağlantı katmanı (fully connected layer) gibi farklı katmanlardan oluşur.



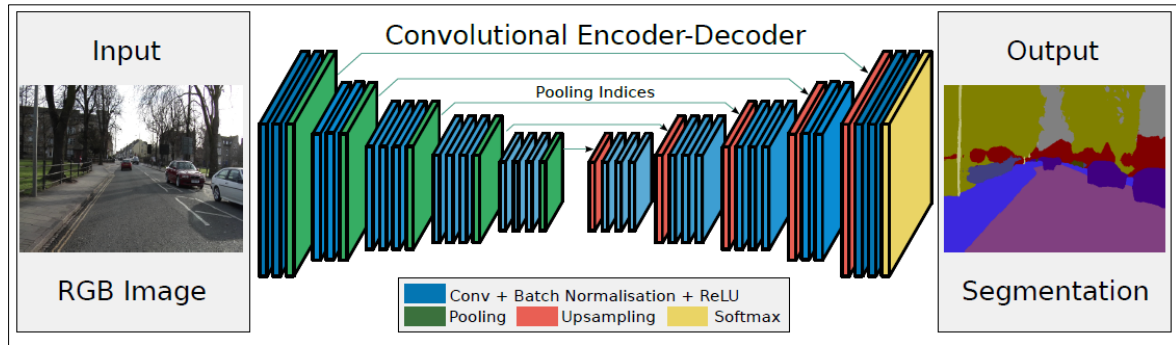
Şekil 2.6. 7 x 7’lik bir giriş matrisine piksel ekleme, konvolüsyon ve havuzlama işlemlerinin uygulanması [39]

Yukarıdaki Şekil 2.6’da da görüleceği gibi 2 boyutlu bir giriş matrisi zero padding (dolgu) yapılarak 3 x 3’lük bir konvolüsyon kernel’i ile filtrelenmiştir. Bu işlem sırasında stride yani adım değeri 2 alınmış ve aktivasyon fonksiyonlarından yararlanılarak 4 x 4’ lük bir sonuç matrisi elde edilmiştir. Bu işlem sırasında kullanılan filtre değeri değiştirilerek özellikler daha iyi tespit edilebilir. Maksimum havuzlama gerçekleştirilerek boyut düşürülerek 2x2 ‘lik bir çıkış matrisi oluşturulmuştur. Havuzlama katmanı çıktıları tam bağlantı katmanlarından (FCL) geçirilerek sınıflandırma çıktılarının verildiği nihai sonuçlar elde edilir. Burada en basit haliyle anlatılan bu mimari yapının günümüzde en iyi sonucu verecek şekle getirilmesi amacıyla pek çok farklı mimari türü oluşturulmuştur: LeNet, AlexNet, VGGNet, GoogLeNet, ResNet vb. [40]

Bu çalışma kapsamında 21 sınıflı ResNet-18 mimarisi kullanılmıştır.

2.6. SegNet Semantik Segmentasyonu

Segmentasyon, görüntülerdeki objeleri ve limitleri tayin etmek ve daha doğru çözümleyebilmek için kullanılır. Semantik segmentasyonda görüntünün içeriğindeki objelerin sınırlarının belirlenmesine ek olarak her bir obje ait oldukları gruplara göre gruplanır. Örnek vermek gerekirse bir görüntünün içeriğinde dört adet kurbağa var ise kurbağaların hepsi tek bir gruba dahil edilir. Semantik segmentasyon, bulunan nesneleri bir sınıfa atar. Robot görüşü, otonom sürüş vazifeleri gibi çalışmalarda sık sık kullanılmaktadır. 2015 senesinde Vijay Badrinarayanan vd. tarafından Cambridge üniversitesinde gerçekleştirilen yapay zeka tabanlı bir çalışmaya ait makalede derin konvolüsyonel kodlayıcı-kod çözücü mimarisi kullanılarak görüntü segmentasyonu gerçekleştirildiği gösterilmiş ve sistemin çalışma performansı detaylıca anlatılmıştır [41].



Şekil 2.7. SegNet semantik segmentasyonu kodlayıcı ve kod çözücü yapısı [42]

SegNet kodlayıcı (encoder) ağı ve onu takip eden bir kod çözücü (decoder) ağına sahiptir (Şekil 2.7). Bu ağlardan geçen görüntü pikselleri sınıflandırma katmanı ile sınıflandırılmaktadır.

Kodlayıcı ağında konvolüsyon, normalizasyon (batch normalisation), ReLU, maksimum havuzlama (pooling) adımları uygulanır. VGG-16 mimarisinden 13 konvolüsyonel katman uygulanmaktadır. Bu adımlar uygulanırken havuzlama işleminde indisler kaydedilir.

Kod çözücü ağıda ise konvolüsyon, normalizasyon (batch normalisation), ReLU, örnek artırımı (upsampling) adımları uygulanmaktadır. Daha önce kodlayıcı ağıda kaydedilen havuzlama indisleri örnek artırımında kullanılır.

Kod çözücü çıktıları son olarak k-sınıfı softmax sınıflandırıcı katmanından geçirilerek her pikselin sınıfı tahmin edilir. Tahmin edilen bu piksel sınıfları RGB olarak renklendirilerek çıkış görüntüsünde gösterilir. Böylelikle anlamsal bir bölütleme görseli elde edilir.

Çizelge 2.2. SegNet çalışması sonuçları [42]

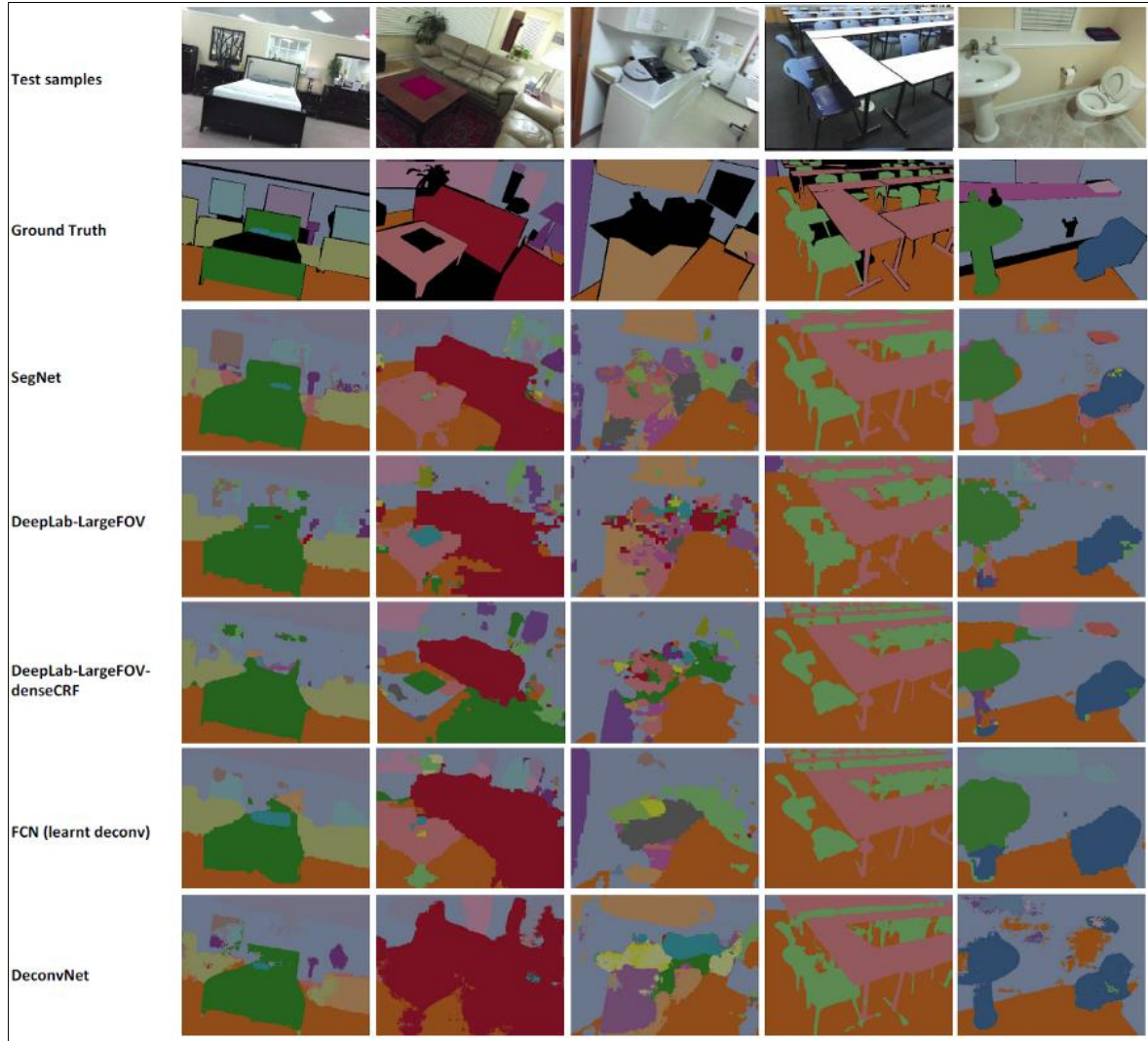
Network/Iterations	80K				140K				>140K				Max iter
	G	C	mIoU	BF	G	C	mIoU	BF	G	C	mIoU	BF	
SegNet	70.73	30.82	22.52	9.16	71.66	37.60	27.46	11.33	72.63	44.76	31.84	12.66	240K
DeepLab-LargeFOV [3]	70.70	41.75	30.67	7.28	71.16	42.71	31.29	7.57	71.90	42.21	32.08	8.26	240K
DeepLab-LargeFOV-denseCRF [3]	not computed								66.96	33.06	24.13	9.41	240K
FCN (learn deconv) [2]	67.31	34.32	24.05	7.88	68.04	37.2	26.33	9.0	68.18	38.41	27.39	9.68	200K
DeconvNet [4]	59.62	12.93	8.35	6.50	63.28	22.53	15.14	7.86	66.13	32.28	22.57	10.47	380K

Bu makalede iki adet veri seti kullanılarak deneysel çalışma gerçekleştirilmiştir. Bu veri setleri SUN RGB-D [43] ve CamVid [44]'dir. Bu tez çalışması kapsamında SUN RGB-D veri seti kullanılmıştır. İlgili makaledeki SUN RGB-D veri setine ait 5250 görüntü üzerinde gerçekleştirilen çalışma çıktıları Çizelge 2.2'deki gibidir. SegNet performansının diğer ağlarla kıyaslaması yapılabilmesi amacıyla 4 adet parametre hesaplanmıştır. Çizelge 2.2'de verilen G küresel doğruluk (global accuracy) değeri doğru sınıflandırılmış piksellerin oranını, C sınıfsal ortalama doğruluk (class average accuracy) değerinin tüm sınıflar üzerinden hesaplanan ortalama tahmini doğruluk değerini, mIoU ise tüm sınıflara ait birleşime oranla kesişimin ortalama değerini (mean intersection over union) ve BF bölgelere ait sınırsal F1 skorunu (boundary F1 score) göstermektedir. Çizelge 2.2'de görülebileceği gibi SegNet sonuçları, FCN, DeconvNet ve DeepLabv1 gibi daha önceki ağların sonuçlarından çok daha iyidir.

Çizelge 2.3. SUN RGB-D veri seti üzerinde gerçekleştirilen çalışma sonucu oluşturulan sınıflar ve doğruluk oranları [42]

Wall	Floor	Cabinet	Bed	Chair	Sofa	Table	Door	Window	Bookshelf	Picture	Counter	Blinds
83.42	93.43	63.37	73.18	75.92	59.57	64.18	52.50	57.51	42.05	56.17	37.66	40.29
Desk	Shelves	Curtain	Dresser	Pillow	Mirror	Floor mat	Clothes	Ceiling	Books	Fridge	TV	Paper
11.92	11.45	66.56	52.73	43.80	26.30	0.00	34.31	74.11	53.77	29.85	33.76	22.73
Towel	Shower curtain	Box	Whiteboard	Person	Night stand	Toilet	Sink	Lamp	Bathtub	Bag		
19.83	0.03	23.14	60.25	27.27	29.88	76.00	58.10	35.27	48.86	16.76		

Çizelge 2.3'te SUN RGB-D veri setine ait segmentasyon sınıfları ve bunların doğruluk değerleri verilmiştir. Fazla sayıda örneğe sahip sınıflar çok daha yüksek doğruluk değerlerine ulaşmaktadır.



Şekil 2.8. SegNet ve diğer ağ çıktılarının karşılaştırılması [42]

Veri setinden seçilen 5 farklı test görüntüsüne (test samples) ait doğruluk değerleri (ground truth), SegNet ve diğer ağların çıktıları Şekil 2.8’de gösterilmiştir. SegNet’in diğer ağlarla karşılaştırıldığında çok daha keskin ve doğru bölütleme gerçekleştirdiği görülmüştür.

Çizelge 2.4. SegNet hafıza ve zaman performansı sonuçları [42]

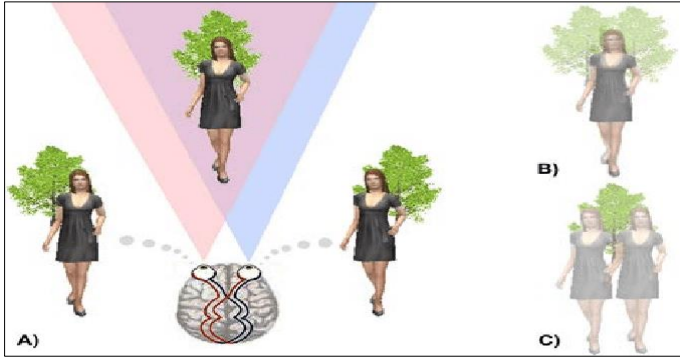
Network	Forward pass(ms)	Backward pass(ms)	GPU training memory (MB)	GPU inference memory (MB)	Model size (MB)
SegNet	422.50	488.71	6803	1052	117
DeepLab-LargeFOV [3]	110.06	160.73	5618	1993	83
FCN (learnt deconv) [2]	317.09	484.11	9735	1806	539
DeconvNet [4]	474.65	602.15	9731	1872	877

Çizelge 2.4’te verilen çalışma sonuçları, zaman ve kullanılan hafıza temelinde değerlendirilmiş ve Çizelge 2.4’te görülebileceği gibi SegNet daha yüksek doğruluk değerlerine ulaşmış fakat eğitim sürelerinde FCN ve DeepLabv1’den daha yavaş sonuç

vermiştir. Buna karşılık SegNet çok daha az hafıza kullanması sebebiyle düşük grafik işlemci üniteleri (GPU) üzerinde rahatlıkla çalışabilmektedir.

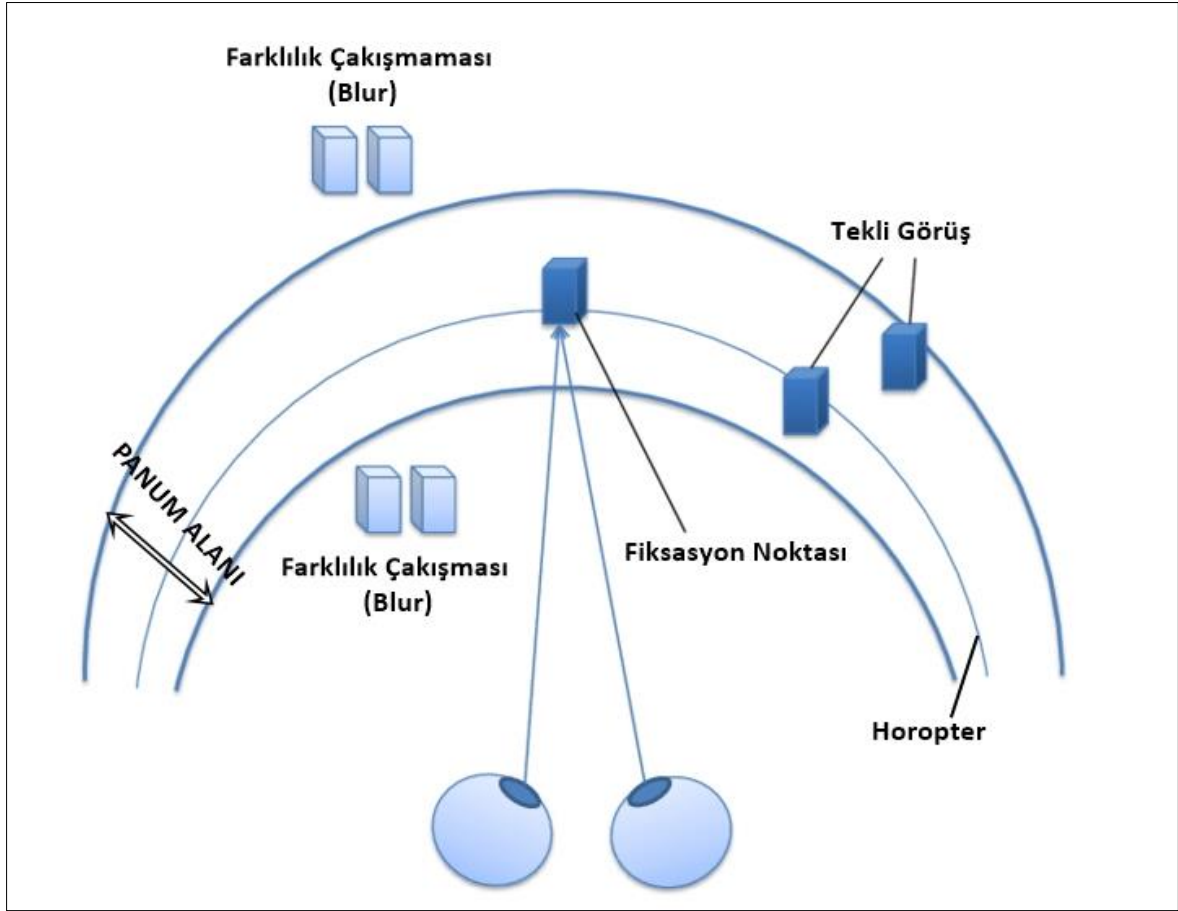
2.7. Derinlik Algılama

Derinlik kelime anlamı olarak Türk Dil Kurumu (TDK) tarafından, bir şeyin dip tarafının yüzeye, ağza olan uzaklığı ya da bir cismin en ve boy dışındaki üçüncü boyutu olarak nitelendirilmektedir. Çevreden kameralar aracılığıyla elde ettiğimiz görüntü ve videolar iki boyutlu bilgi taşırlar. Pek çok uygulama için iki boyutlu bilgi yeterli olmamakta ve çevreye ait olan bilgiyi artırmak adına üçüncü boyut bilgisine ihtiyaç duyulmaktadır. Üçüncü boyut bilgisi olarak ihtiyaç duyulan bu derinlik bilgisinin tahmini için pek çok yöntem geliştirilmiştir. Bu yöntemleri anlamak için öncelikli olarak insanın görme sistemiyle bu bilgiyi nasıl ayırt ettiğini incelemek gerekir.



Şekil 2.9. İnsan görme sistemi örnek görüntüleri [45]

Gözler aracılığıyla alınan sağ ve sol görüntü çiftleri arasındaki farklı bakış açısı nesnelere ait derinlik değeri çıkarımı yapılmasını sağlar. Göz özelinde konuşmak gerekirse teoride alınan görüntülerin horopter eksenini üzerinde olmaması durumunda görüntüde çift görmeye sebebiyet vermesi gerekmektedir. Ancak pratikte Şekil 2.9’da görüldüğü gibi belirli bir tolerans değeriyle tek görme gerçekleşmektedir.



Şekil 2.10. Panum alanı, horoptyer, fiksasyon noktası gösterimi

Şekil 2.10 ile görme sistemine ait panum alanı, horoptyer, tekli görüş ve blur görüntü alanları temsili olarak gösterilmektedir. Panum alanı dışında kalan bölgelerde blurlu görme ya da çoklu görme gerçekleşmektedir [46]. Fiksasyonun sağlanması için gözün bir noktada minimum 100 ms civarında kalması gereklidir. Göz görüntülerinin çakıştırılması (fuse) için fiksasyon mesafesinde bir bölgede derinlik ölçülmelidir. Bunun için gözler vergence yani gözlerin zıt yönlü hareketi ile daha uzak bölgelerde derinlik mesafesi bulabilmektedir. Örneğin 40 cm gibi bir uzaklık için panum alanı +5 ve -5 mm gibi bir aralığa sahiptir. Gözlerin açısının kalması durumunda panum alanı çok kısıtlı bir bilgiye sahip olacaktır. Göz açılarının vergence hareketi sayesinde çok daha uzak mesafelere ait derinlik çıkarımı yapılabilir. Uzak hedefler için iki göz eksenindeki açıyı 28° civarında, yakın hedefler için ise 72° civarında açılara getirilmektedir. Tüm bunlar ek olarak gözler dinamik birer odak mesafesine sahiptir. Bu sayede sol ve sağ görüntü alınmaktadır; daha önceki tecrübeler ve bilgi birikimi ile birlikte nesnelere ait derinlik yaklaşımı yapılmaktadır.

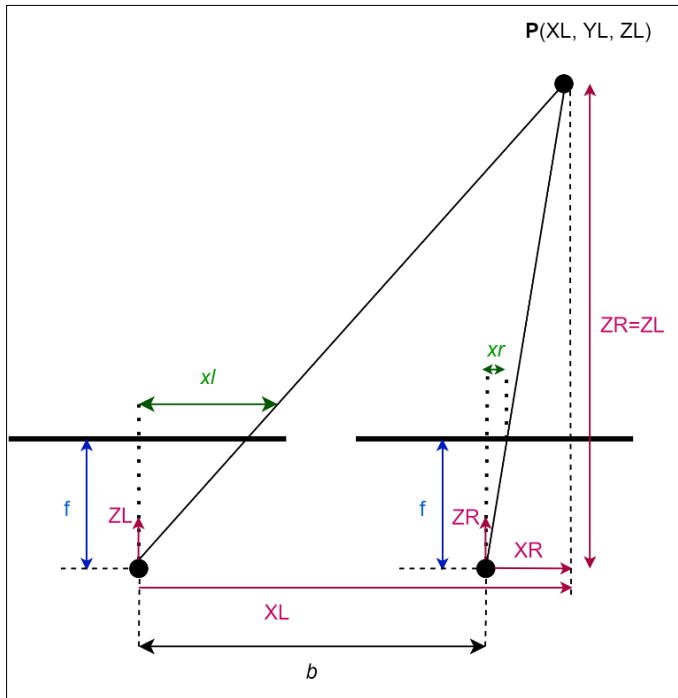
İnsan gözünden hareketle donanımsal ve yazılımsal olarak derinlik tespiti sistemleri geliştirilmeye çalışılmıştır. Donanımsal olarak ikili görüntü elde etmek amacıyla ikili (dual) kamera teknolojileri ve ikili piksel teknolojileri (DPAF-Dual Pixel Autofocus) geliştirilmiştir. Günümüzde cep telefonlarında da bu teknolojiler kullanılmaya başlanmıştır. Bunlara ek olarak sensörler yardımıyla çevreye ait derinlik tespitleri de yaygın olarak kullanılmaktadır. En yaygın kullanılabilecek örnek vermek gerekirse; Kinect, kızıl ötesi projektörü sayesinde çevreye kızıl ötesi noktalardan oluşan bir yayın yapar ve bunlar arasındaki farkı hesaplayarak derinlik tespitinde bulunur. Lidar ve ultrasonik sensörler de derinlik ölçümü için kullanılan sensörlerdendir. Yazılımsal olarak da dijital görüntüler üzerine gerçekleştirilen görüntü işleme yöntemleri ve akıllı algoritmalar ile derinlik değeri elde edilmeye çalışılır. Tek görüntü ve çoklu görüntü üzerinde bu işlemler gerçekleştirilebilir. Yapay zeka alanında gerçekleşen son gelişmeler yardımıyla gözetimli (supervised) ve gözetimsiz (unsupervised) öğrenme yöntemleri de bu alanda sıkça kullanılmaya başlanmıştır.

Bu çalışmada pasif derinlik algılama sistemlerinden olan stereo kameralar ile derinlik algılama sistemleri kullanılmıştır. İki kameradan aynı zaman kesitine ait alınan görüntüler üzerinden derinlik çıkarımı yapılmaktadır. Bu alanda geliştirilen yapay zeka tabanlı algoritmaların performanslarının gerçek zamanlı mobil sistemler için yeterli seviyede olmaması sebebiyle blok eşleme yöntemleri temel alınarak derinlik görüntüsü elde edilmiştir.

2.8. Stereo Görüntüleme

Robotik sistemlerde çevreyle olan etkileşimin artırılması adına derinlik tespitine ihtiyaç duyulmaktadır. Algılanan bu derinlik değeri üçüncü boyut bilgisi de elde edilerek otonom sürüş, üç boyutlu sahne modellenmesi, engelden kaçınma, haritalama, artırılmış gerçeklik vb. pek çok uygulamada kullanılmaktadır. Bilgisayarlı görü sistemlerinde bir nesnenin kameraya olan uzaklığının yani derinliğinin tespit edilmesi gerekmektedir. Derinlik ölçme sistemleri genel olarak ikiye ayrılır. Bunlardan ilki aktif sistemlerdir. Aktif sistemler nesneye mikrodalga, kızılötesi, ses dalgası ve lazer gibi ışınlar göndermektedirler ve yansımaların kendisine ulaşması ile geçen zamanı orantılayarak derinlik değerlerini tespit etmektedirler. İkinci bir yöntem ise pasif sistemlerdir. Pasif sistemler derinlik değerini, ortama yayın yapmadan aldığı bilgileri işleyerek elde eden sistemlerdir. Monoküler ve stereo görüntüleme

sistemleri bunlara örnek verilebilir. Monoküler sistemler tek bir kameradan gelen görüntü değerlerini işleyerek ortama ait derinlik tahmininde bulunmaktadır. Orantılama, kenarlar, köşeler ve hareket gibi anahtar özellikler kullanılarak derinlik değerleri elde edilmeye çalışılmaktadır. Stereo görüntüleme sistemleri ise görüş alanları çakışan iki kameradan meydana gelmektedirler. Bir nesneye ait görüntü bu iki farklı konumda olan kameralardan aynı anda alınmaktadır. İki görüntüde de nesnenin konumu tespit edilir ve aradaki piksel farkından kamera ile nesne arasındaki mesafe elde edilir. Aynı nesneye ait görüntüler arasındaki piksel eşitsizliği (disparity) değeri ile nesnenin kameraya olan uzaklığı (derinliği) ters orantılıdır.



Şekil 2.11. Stereo görüntüleme sistemi koordinat eksenleri ve değerleri

Şekil 2.11’de P nesnesinin stereo görüntüleme sistemi üzerinde oluşturacağı görüntüler koordinat ekseninde gösterilmiştir. Piksel eşitsizliği (d) Eş. 2.2’de $(x_r - x_l)$ ile gösterilmiştir. P noktasının kameralarda oluşturduğu görüntünün merkeze uzaklığı x_r ve x_l ile verilmiştir. Kamera sensörlerinin P noktası ile olan gerçek uzaklık değerleri X_L , X_R , Z_R ve Z_L değerleridir (Şekil 2.11). Birbirine eş odak uzaklığına sahip bu iki kamerada P noktasının derinlik değerine (z) Eş. 2.10 ile geçilebilmektedir.

$$d = x_r - x_l \quad (2.2)$$

$$xl = f \frac{XL}{ZL} \quad (2.3)$$

$$xr = f \frac{XR}{ZR} \quad \rightarrow \quad xr = f \frac{(XL+b)}{z} \quad (2.4)$$

$$ZL = ZR = Z \quad (2.5)$$

İki kamera arasındaki taban uzaklığı b 'dir (Şekil 2.11).

$$XR = XL + b \quad (2.6)$$

Eş. 2.2'deki denklem yeni bilgiler ışığında P noktasına olan gerçek mesafeler XL ile yeniden yazılır (Eş. 2.7 – Eş. 2.8) ve ilgili işlemler yapılırsa farklılık değeri (piksel eşitsizliği) bulunur (Eş. 2.9).

$$d = \left(f \frac{(XL+b)}{z} \right) - \left(f \frac{(XL)}{z} \right) \quad (2.7)$$

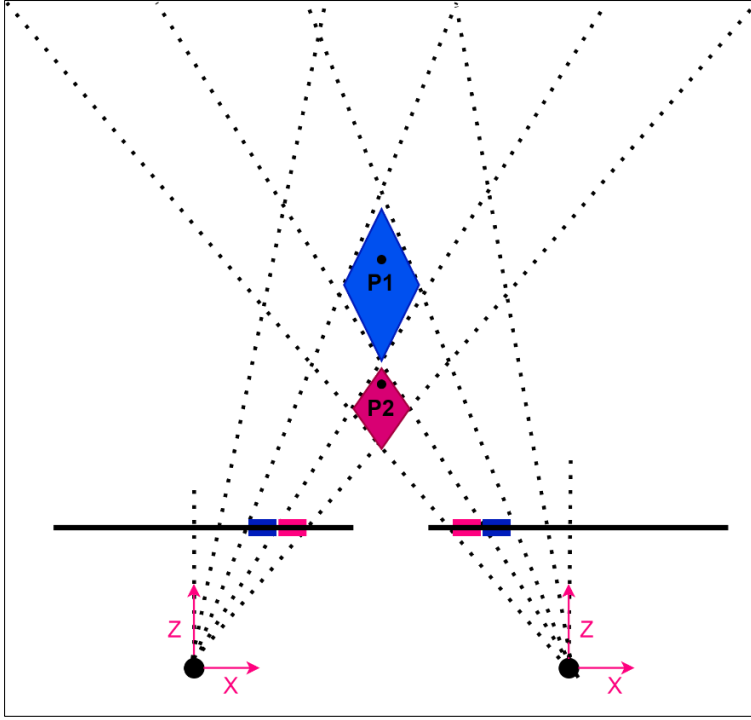
$$d = \left(f \frac{(XL+b-XL)}{z} \right) \quad (2.8)$$

$$d = f \frac{b}{z} \quad (2.9)$$

Bulunan farklılık değeri Eş. 2.9'dan derinlik değeri bilgisi elde edilebilir (Eş. 2.10).

$$z = f \frac{b}{d} \quad (2.10)$$

Eş. 2.9'dan de görülebileceği gibi disparity (farklılık) değeri arttıkça derinlik değeri azalmaktadır. Bir nesnenin iki kamera arasındaki eşitsizlik değeri, kameraların odak uzaklığı ve taban uzaklığı (baseline width) bilindiği takdirde nesnenin derinlik değeri hesaplanabilmektedir.



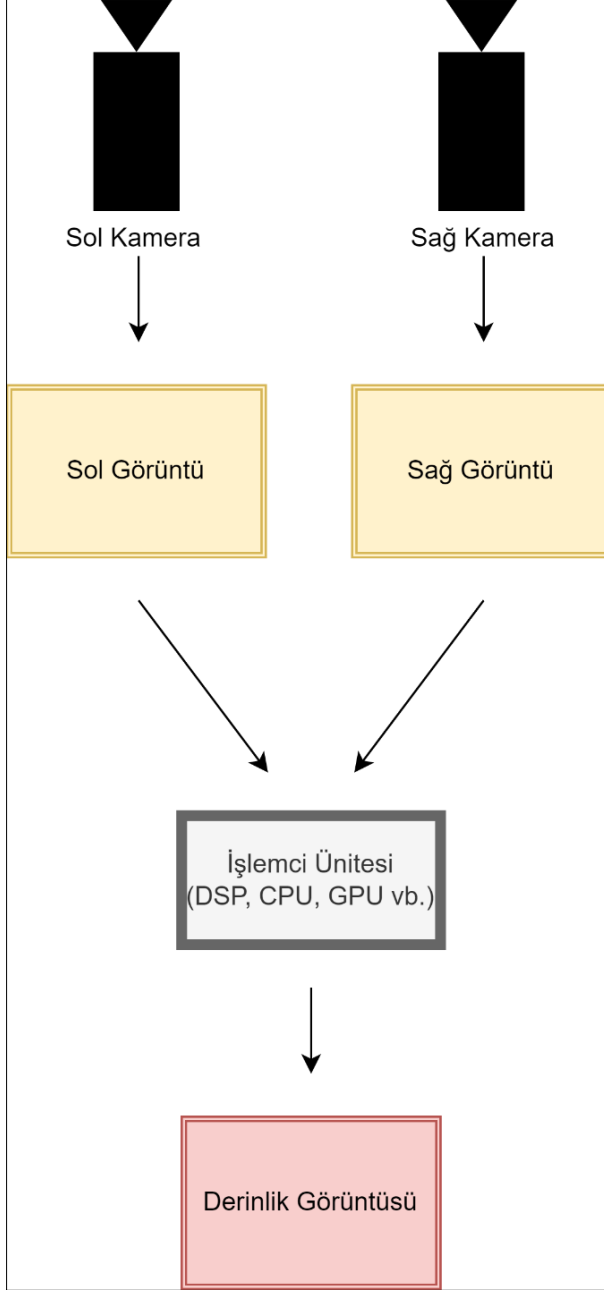
Şekil 2.12. Stereo görüntülemeye belirsizlik

Şekil 2.12’de de görüldüğü gibi nesneye ait piksel değerleri nesnenin mutlak konumunu belirleyememektedir. Şekil 2.12’deki P2 noktasına ait piksel verisi belirsiz bir bölgeyi niteler. P2 noktası yerine P1 noktası alındığında ise noktaya ait net konumun belirsizliğinde artış meydana gelir. Bu belirsizlik değeri derinlik, odak uzaklığı ve taban uzaklığına bağlı değişir. Disparity (farklılık) değeri azaldıkça belirsizlik artmaktadır.

$$\frac{d}{b} = \frac{f}{z} \quad (2.11)$$

Kameralar arası mesafe ve odak uzaklığı yukarıdaki Şekil 2.11’de gösterilmiştir. Burada uygulanacak bir benzerlik teoreminden de Eş. 2.11 elde edilir. Stereo görüntülemeye daha düşük çözünürlük değerleri daha düşük detay anlamına gelmektedir. Çözünürlükteki düşüş de belirsizlikte (ambiguity) artışa sebep olmaktadır. Belirsizlikteki artışla birlikte alınan derinlik değeri bilgisi de daha düşük olmaktadır. Eğer çözünürlük değerleri, 1080p gibi değerlerden VGA ya da QVGA gibi ölçülere getirilirse derinlik bilgisinde artış sağlamak amacıyla ya taban uzaklığı (b, baseline width) değeri artırılmalı ya da odak uzaklığı (f, focal length) değeri artırılmalıdır. Odak uzaklığı kameraların üretim aşamasından gelmekte ve başlangıç durumunda tespiti yapılmaktadır. Değiştirilmesi daha kolay olan taban mesafesindeki değişim ise görüş alanında çakışmayı düşürmektedir. Benzer şekilde odak

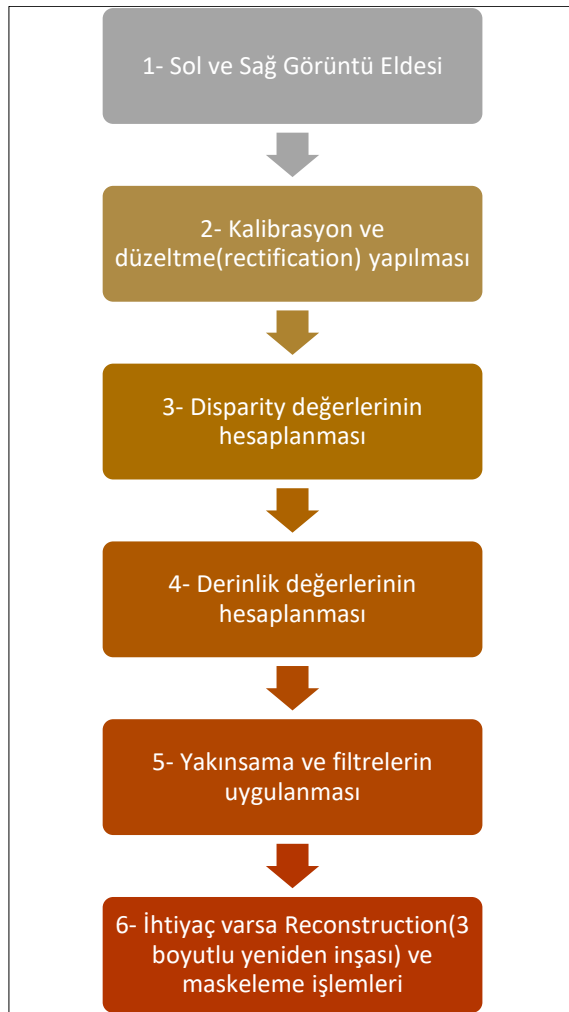
uzaklığındaki artış da benzer görüş alanı daralmasına sebebiyet verecektir. Bu amaçla bütünlük cihazlarda belirli kısıtlar çerçevesinde kullanım alanına yönelik olarak bu değerler belirlenmektedir. Bütünlük stereo görüntüleme sistemlerinde bu değerler değiştirilemeyen parametrelerdir ve bu değerlere bilgi formlarından (datasheet) ulaşılabilir.



Şekil 2.13. Kameralardan derinlik görüntüsü eldesi şematik gösterimi

Şekil 2.13'te görüldüğü gibi stereo görüntüler aynı anda tetiklenerek alınmaktadır. Belirsizliği azaltmak için yüksek çözünürlüklü alınan bu görüntülerin üzerinde işlem yapabilmek için DSP, FPGA ve GPU'lar kullanılmaktadır. Derinlik değerlerinin elde

edildiği görsellerde titremeler (flickering) gerçekleşebilmektedir. Bunun sebebi ortamın iyi aydınlatılmamış olmasıdır. Stereo görüntüleme sistemleri için ortam aydınlatması çok büyük önem arz etmektedir. Kameraları saturasyona uğratmayacak şekilde harici bir yüksek ışıklı aydınlatma ile bu titreme etkisi de önlenebilmektedir. Şekil 2.13'te sol ve sağ görüntülerin bir donanım birimi üzerinde işlenmesi sonrası derinlik değerlerinin elde edilişi gösterilmiştir. Derinlik değerini elde etmek için gerçekleştirilen ana işlemler sırasıyla Şekil 2.14 ile gösterildiği gibidir:



Şekil 2.14. Derinlik değeri elde etme adımları

Bu tez kapsamında kullanılmış olan bu ana adımlara ait genel bilgilendirmeler bu bölümün alt başlıklarında verilmiştir.

2.8.1. Sol ve sağ görüntünün eldesi

Stereo kameralar aynı zaman diliminde tetiklenerek görüntüler elde edilmektedir. Sensörler aracılığıyla elde edilen görüntülere ait elektriksel sinyaller, sayısal görüntü işaretlerine çevrilerek bellekte kaydedilirler.

2.8.2. Kalibrasyon ve düzeltme (rectification) yapılması

İkili kameradan alınan bu görüntüler idealde hiçbir zaman kusursuz çıktı vermemektedirler çünkü lens, kayıt cihazı, kamera ve ortam koşullarına bağlı olarak bozulmalara uğramaktadırlar. Bazı bozulmalar, cihaz üzerinde değişikliklere müdahale ederek düzeltilebilse de fabrikasyon kaynaklı bozulma sebeplerini düzeltme imkânı bulunmamaktadır. Bu tip bozulmalar yazılımsal olarak görüntü işleme yöntemleri kullanılarak düzeltilebilirler.

Radyal görüntü bozuklukları

Radyal görüntü bozukluklarından en sık karşılaşılanlar kamera optiği ve lensi kaynaklı bozulmalardır. En sık görülen çeşitleri fıçı bozulması, yastık bozulması ve bıyık bozulmasıdır.

Fıçı bozulması (barrel distortion)

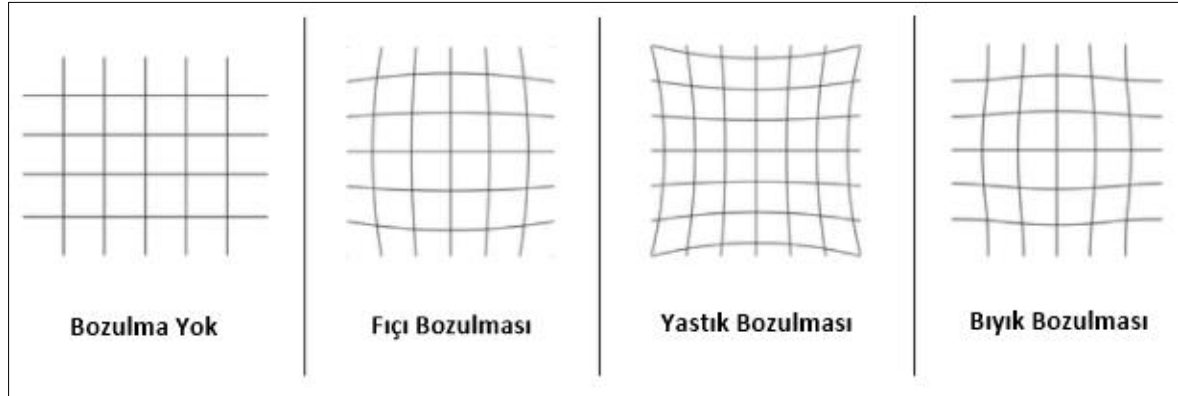
En çok balıkgözü lenslerde görülen bu bozulmada görüntüler merkeze doğru büyürken kenarlarda küçülmektedir. Dış bükey bozulma ya da varil bozulması olarak da adlandırılmaktadırlar.

Yastık bozulması (pincshion distortion)

Görüntü merkezinden kenarlara doğru ilerledikçe iç bükülmeler gösteren bozulmalardır. İğne yastığı ve iç bükey bozulma olarak da adlandırılmaktadırlar.

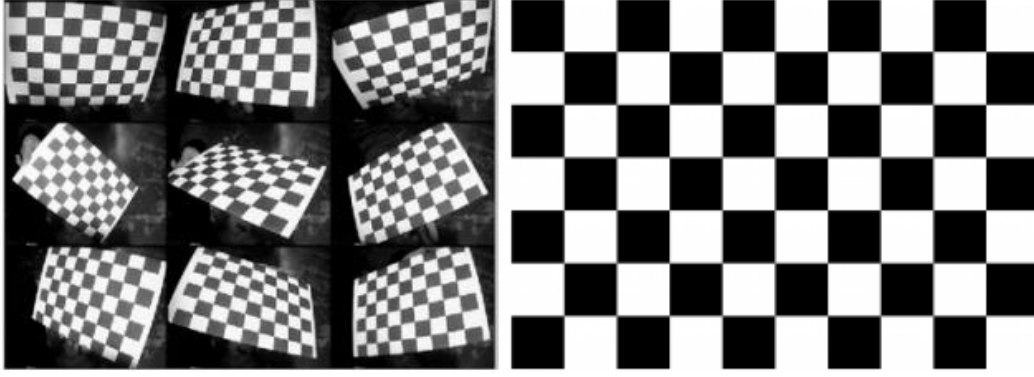
Bıyık bozulması (mustache distortion)

Görüntü üzerinde merkezden kenarlara ilerledikçe inişli çıkışlı bir eğrilme gösteren bozulmalardır. Eğrilme şablonu, şekli itibariyle bıyığa benzetildiği için böyle isimlendirilmiştir. Tespiti diğerlerinden daha zordur.



Şekil 2.15. Radyal görüntü bozulmaları örnek gösterimleri

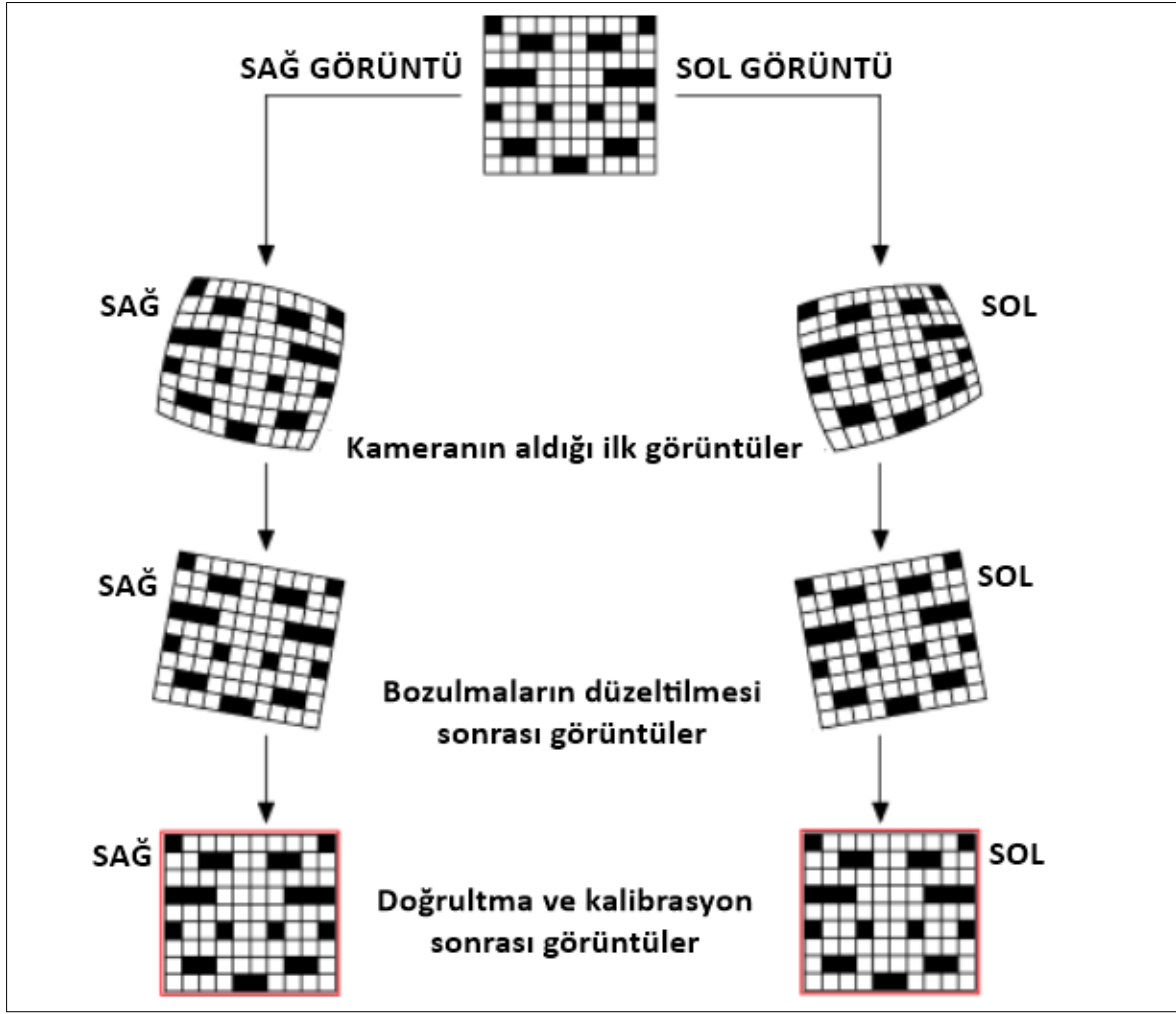
Şekil 2.15'te en yaygın görülen örnekleri verilen radyal bozulmaları düzeltmek amacıyla kameralara kalibrasyon işlemi uygulanmaktadır. Kameralara uygulanan kalibrasyon sonrasında düzeltme işleminde kullanılacak bilgiler elde edilmektedir. Kalibrasyon kameraya özgü gerçekleştirilmektedir ve başka bir kamerada uygulanması mümkün değildir. Kamera matrisi ve bozulma katsayılarının matrisleri elde edilmesi için işlem uygulanmaktadır. Literatürde farklı yöntemler bulunmasına rağmen çoğunluğu referans noktaları tespit ederek bu noktaların gerçek değerlerle olan bağıntısından yararlanarak düzeltme işlemi yapılmasını öngörmektedir. Kalibrasyon yöntemlerinden en iyi bilineni boyutları bilinen çift ve tek kare sayılarına sahip satranç tahtası kullanarak gerçekleştirilen kalibrasyondur.



Şekil 2.16. Kalibrasyonda kullanılan satranç tahtası görüntüleri

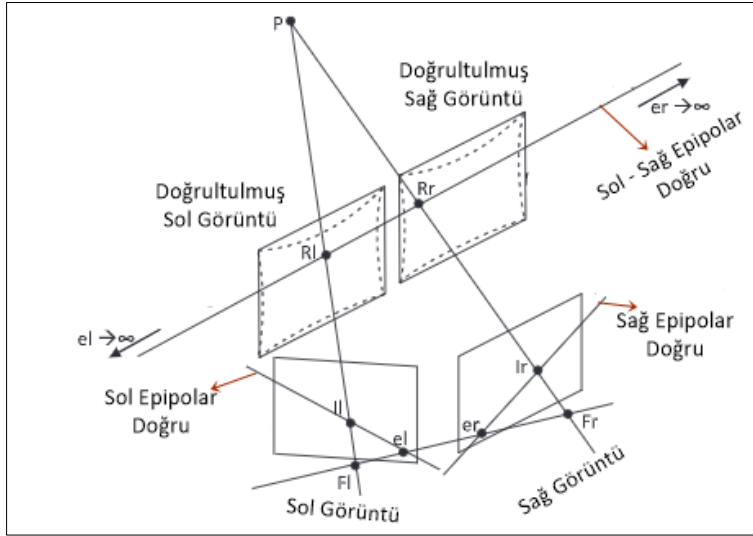
Şekil 2.16'daki gibi bir satranç tahtasının karelerinin köşeleri referans noktaları olarak tespit edilmek için kameraya farklı açı, uzaklık ve eğimde tutularak tahtanın tamamı görülecek şekilde görüntüleri alınır. Alınan bu görüntülerden referans köşe noktaları çıkarımı yapılarak düzeltme matrisleri elde edilir.

Elde edilen bu düzeltme verilerinden faydalanarak görüntü düzeltme (undistorted) işlemi gerçekleştirilir (Şekil 2.17). Yapılan araştırmada beşten fazla görüntüyle gerçekleştirilen kalibrasyon işlemlerinde çok daha sağlıklı sonuçlar alındığı görülmüştür [47].



Şekil 2.17. Kalibrasyon ve doğrultma adımları örnek gösterimi

Kalibrasyonu gerçekleştirilen görseller elde edilen kamera parametreleri rektifikasyonu (düzeltmesi) yapılmaktadır. Rektifikasyon, epipolar doğruların görüntü satırlarında eşleşmesi amacıyla yapılan stereo görüntü çiftinin yeniden örneklenmesi işlemidir. Rektifikasyon matrisi oluşturabilmek için sistemin dışsal parametrelerine ihtiyaç duyulmaktadır [48]. Bu parametreler yardımıyla düzeltilen görüntüler aynı epipolar düzleme oturtulmuş olmaktadır (Şekil 2.18).



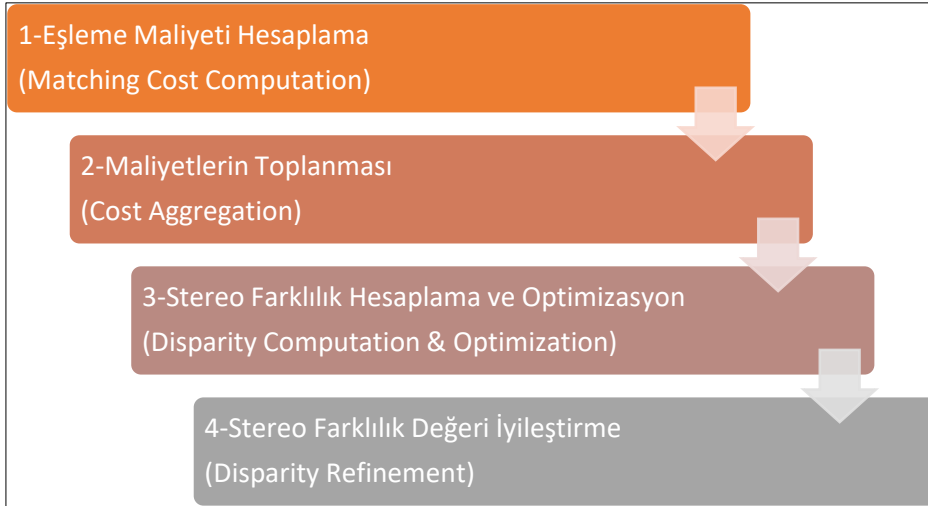
Şekil 2.18. Sol ve sağ görüntünün epipolar düzlemde gösterimi [49]

Aynı epipolar düzleme oturtulmuş olan kalibrasyonu gerçekleştirilmiş sol ve sağ görüntüler yardımıyla farklılık değerlerinin hesaplanması adımına geçilir.

2.8.3. Disparity (farklılık) değerinin hesaplanması

Disparity (farklılık) değerlerinin hesaplanması için kalibrasyonu ve rektifikasyonu yapılmış sol ve sağ görüntüdeki eş piksellerin bulunması gerekmektedir. Bu eşleme (correspondence) işlemi ne kadar doğru yapılırsa derinlik görüntüsü de o derece doğru olmaktadır.

2002 yılında Daniel Scharstein ve Richard Szeliski tarafından stereo eşleşme algoritmaları üzerine yayınlanan “A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms” isimli makalede dört adımdan oluşan bir taksonomiden bahsedilmektedir [50]. Şekil 2.19’da iyi bir eşleşme için gerçekleştirilecek bu dört adım verilmiştir.



Şekil 2.19. Farklılık değeri hesaplanması adımları

Her piksel üzerinde bu dört adım gerçekleştirilerek farklılık haritası (disparity map) elde edilmektedir. İlk adım olan eşleme maliyeti hesaplanmasında her iki resimde aynı cisme ait pikseller aranır ve bu piksellerin ne kadar iyi eşleştikleri hesaplanır. İkinci adımda ise eşleme sırasında bulunan değerler için toplam maliyet hesaplanır. Üçüncü adımda en düşük maliyetli sıra için farklılık değeri bulunur. Son adım olan dördüncü adımda ise bu değerlerin iyileştirilmesi yapılarak gerçek derinlik değerlerine en yakın sonuç elde edilmeye çalışılır. Global algoritmalarda ikinci adım uygulanmadan farklılık hesaplama adımına geçilmektedir.

Stereo görüntüler arasında eşleme değeri hesaplanması amacıyla pek çok farklı matematiksel hesaplama yöntemi ve algoritması tasarlanmıştır. Tasarlanan bu algoritmalar bilgisayarlı görü alanına taşınmıştır.

Bu alanda en sık kullanılan algoritmalar arasında özellik eşleme algoritmaları ve alan eşleme algoritmaları bulunmaktadır [51].

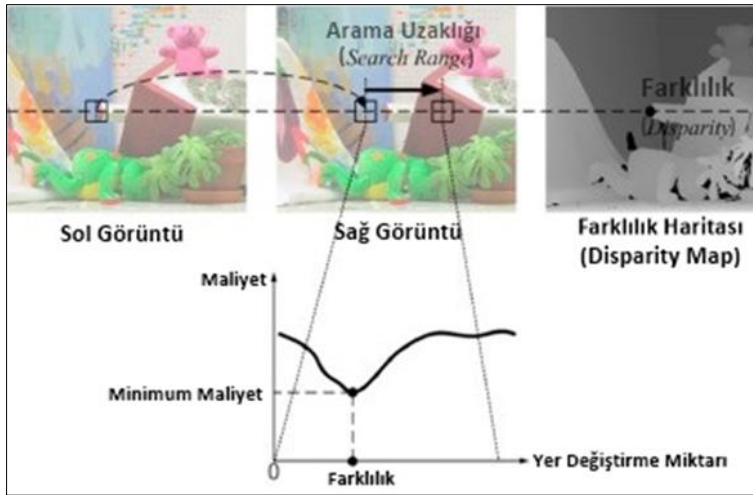
Özellik Eşleme Algoritmaları

Kenar, köşe gibi özellikler taşıyan pikseller için eşleme yapılmaktadır. Bu özel noktalar (key feature points) için derinlik haritası oluşturulmaktadır. Çok hızlı çalışan algoritmalar oldukları için gerçek zamanlı derinlik tespitinde sıkça kullanılmaktadır. Dış etkenler (ışıklandırma vb.) bu özellik noktalarının çıkarılmasını daha az etkilediği için algoritma farklı koşullarda da verimli çalışabilmektedir.

Alan Eşleme Algoritmaları

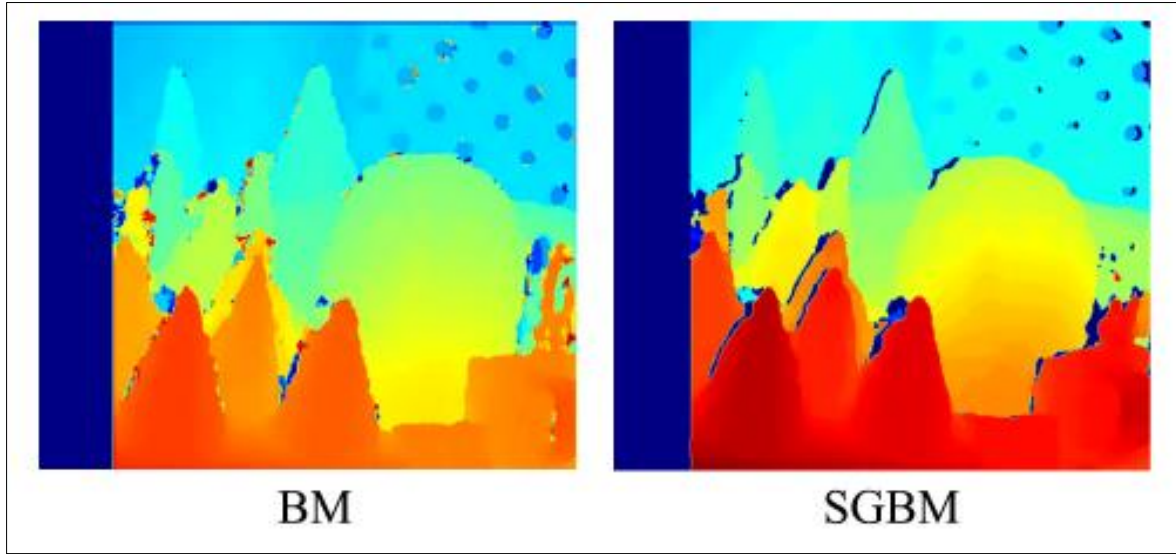
Özellik eşlemenin aksine iki görüntü arasındaki her bir piksel için eşleme gerçekleştirilir. Bu algoritma ile üretilen derinlik haritası yardımıyla üç boyutlu sahnenin yeniden oluşturulması sağlanabilir fakat dış etkenlerden daha fazla etkilenmektedir. Görüntülenen sahnenin detayları da derinlik haritasını etkilemektedir.

Bu hesaplamalar yapılırken bölgeye özel (lokal) yaklaşımlar ile piksel bazlı çözümler getirilebilmektedir. Bu tip yaklaşımlarda resmin tamamı ile ilgilenilmediği için hızlı sonuçlar üretilmektedir. Global yaklaşımlarda ise global optimizasyon fonksiyonu ile tüm resim için bir çözüm üretilmektedir. Bu yönleriyle global yaklaşımlar çok fazla işlem gücüne ihtiyaç duyarlar ve yavaş çalışmaktadırlar.



Şekil 2.20. Sol ve sağ görüntü farklılık değeri maliyet hesaplaması [52]

Birbirinden farklı bu yöntemleri temel alarak geliştirilen ve optimize edilerek yüksek hızlarda çalışması sağlanan OpenCV blok eşleme algoritması bir kamerada alınan görüntüye ait her bir pikselin diğer kameradaki eşini bulmaya çalışmaz. Bunun yerine komşu piksellerle bir blok oluşturur ve bu bloğun benzerini diğer kamera üzerindeki bloklarla karşılaştırır (Şekil 2.20). Burada OpenCV pek çok kısıtın parametresel olarak değiştirilmesine imkan vermektedir.



Şekil 2.21. Blok eşleme (BM) ve Yarı Global Blok Eşleme (SGBM) örnekleri

OpenCV üzerinde tanımlı olan bir başka optimize edilmiş algoritma da Yarı-Global Blok Eşlemedir (Semi-Global Block Matching). 2005 yılında H. Hirschmüller tarafından yayınlanan makale temel alınarak geliştirilmiştir [53]. Blok eşleme algoritmalarından daha yavaş çalışan bu algoritma çok daha yüksek doğruluklu sonuçlar vermektedir (Şekil 2.21).

2.8.4. Derinlik değerinin hesaplanması

Bir önceki adımda hesaplanan farklılık (disparity) değerleri her piksel için Eş. 2.11’de yerine konularak derinlik değerlerine geçiş yapılır. Böylelikle derinlik haritası elde edilmiş olunur. Belirlenen bu derinlik değerleri arttıkça belirsizliğin artacağı unutulmamalıdır.

2.8.5. Yakınsama ve filterelerin uygulanması

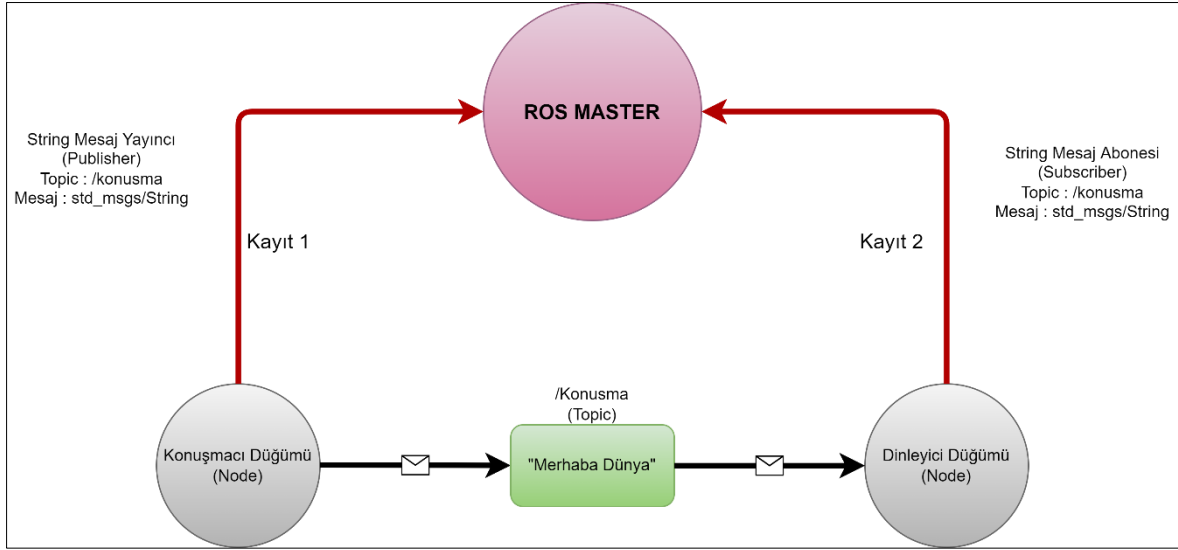
Elde edilen bu derinlik değerleri öncesinde uygulanan yöntemlere göre gürültülü ve doğruluktan uzak durumdadır. Derinlik hesaplamadaki kısıtlar sebebiyle alınan derinlik değerlerinde kopukluklar ve hatalar meydana gelmektedir. Bu değerlere yakınsama işlemleri ve doğruluğu artırıcı filtreler uygulanarak gerçek derinlik değerleriyle örtüşen bir derinlik haritası elde edilmiş olur.

2.8.6. Reconstruction (üç boyutlu yeniden inşa) ve maskeleme işlemleri

İhtiyaç dahilinde elde edilen bu nihai derinlik haritasından faydalanarak piksel piksel üç boyutlu bir görüntü inşası gerçekleştirilebilmektedir. Bu sayede çevreye ait derinlik değerleri ile kameradan alınan iki boyutlu görünür bant renkli görseller, eşleştirilerek üç boyutlu bir haritalama gerçekleştirilmektedir. Bu aşamada tüm piksellerin yerine ihtiyaç duyulan nesne, insan ya da bölgeye ait pikseller de ayrıştırılarak üç boyutlu hale getirilebilmektedir.

2.9. Robot İşletim Sistemi

Robot İşletim Sistemi, ingilizce isminin baş harfleri olan ROS (Robot Operation System) ile anılmaktadır. ROS'dan önce geliştirilen her yazılım robotik platforma özgü durumdaydı. Bu sebeple yapılan çalışmalar hedefinden sapmakta ve geliştiriciler platformlara özgü problemlerin üstesinden gelmekle uğraşmaktaydı. Bir platforma özgü geliştirilen algoritmanın başka bir donanım ortamında çalışmaması gibi durumlar çokça yaşanmaktaydı. ROS üzerine ilk çalışmalar Stanford Yapay Zeka Laboratuvarı'nda 'STAIR' isimli projeyle 2007 yılında başlamıştır [54]. Bu çalışmalar 2008 yılında Willow Garage ve ekibinin üstlendiği çalışmalarla sürdürülmüştür. 2010 yılında ilk versiyonunu yayınladıkları Box Turtle ile ROS sistemi robotik alanında hızla yayılmıştır. ROS içerdiği gecikmelerden dolayı gerçek zamanlı bir işletim sistemi olarak adlandırılmamaktadır. ROS, robotik alanında yazılımcıların ihtiyaç duyduğu araçları (paketler, kütüphaneler, datalar, veri paylaşım yöntemleri vb.) içeren platformdan bağımsız olarak çalışabilen açık kaynak kodlu bir yazılım iskeleti şeklinde tanımlanabilir. Bu çatı sayesinde platformların kontrolleri, sensör okumaları, haberleşme işlemleri, alınan verilerin anlamlandırılması, haritalama, navigasyon vb. işlemler gerçekleştirilebilmektedir. Bu yazılım çatısı altında bulunan Laser Scan Matcher, Robot Pose EKF, AMCL, gMapping ve Navigation Stack vb. paketler herhangi bir robotik platform üzerinde de çalıştırılabilmektedir [55]. Tüm bunlara ek olarak TurtleBot, ClearPath HUSKY, PR2 vb. robotik araçlara özel yazılımlar da ROS içerisinde barındırılmaktadır. ROS sisteminin sağladığı kolaylıklar sayesinde araştırmacılar çalışır sistemin oluşturulmasından daha çok kendi araştırma konularına odaklanabilmektedir. ROS bir işletim sistemi değildir ve bir işletim sistemi altında çalışması gerekmektedir. ROS kütüphaneleri UNIX temelli işletim sistemleri üzerinde çalışmaktadır. Ubuntu Linux sistemi ile yaygın olarak kullanılmaktadır [56].



Şekil 2.22. ROS mimarisi basit gösterimi

ROS ağ yapısında döğümler, master, parametre sunucuları, mesajlar, hizmetler, konular ve depolar bulunmaktadır. ROS mimarisi yayınlama ve abonelik sistemi üzerine kuruludur [57]. ROS mimarisinin bileşenleri aşağıdaki gibidir:

Roscore ya da Rosmaster

- Döğümlerin ve programların bir koleksiyonudur.
- Tüm iletişimi ayarlar ve yönetir. IP adreslerini tutar ve aktif olan döğümler arası bağlantı sağlar.
- Kimin yayın yaptığını, kimin abone olduğunu bilir. Çalışma anında hangi servislerin ve konuların aktif ve ulaşılabilir olduğunu takip eder.
- Her ROS projesinde roscore çalışmalıdır. Bir aktif ağ yapısı içinde yalnızca bir roscore çalışır.

Döğümler (Nodes)

- Tek başlarına derlenebilir dosyalardır.
- Çalıştırıldıklarında üzerlerine tanımlı görevleri gerçekleştirirler. Bu görevlere örnek vermek gerekirse; sensör verisinin okunup yayınlanması, abone olunarak alınan bir verinin kullanılarak harita oluşturulması ya da elindeki verileri işleyerek en kısa yolu bulan bir yol bulma bunlar arasında sayılabilir.

- Bir düğüm veriyi başka bir düğümden alabilir ya da gönderebilir. Bunu servisler ve konular yardımıyla yaparlar.

Konular (Topics)

- Düğümler arası iletişimler için kullanılırlar.
- Yayıncıları ya da aboneleri olan kavramlardır.
- Konuların yalnızca bir yayıncısı olabilir; buna zıt olarak bir konuya birden çok abone olabilir.
- Konular tek yönlü olarak çalışırlar. Bir yayıncı konu üzerinden aboneye mesajını iletebilir. Ancak abone yayıncıya bir mesaj göndermek isterse başka bir konuya ihtiyaç vardır.

Mesajlar (Messages)

- Bir yayıncı neyi konu üzerine yayınlıyor ya da bir abone hangi veriyi alıyor onu tanımlamak için kullanılır.
- Gönderilen mesajın türüne göre ROS üzerinde farklı veri tipleri kütüphanelerle belirlenmiştir. Örnek vermek gerekirse, bir düğüm Kartezyen koordinat sistemindeki bir aracın x, y ve z değerlerini float türünde pozisyon bilgisi olarak konuya bir mesaj içerisinde yollar. Böylelikle aboneler bu mesajlara belirlenen ölçütlerde ulaşabilir.

Servisler (Services)

- ROS servisleri en basit tabiriyle senkronize çalışan istemci ve sunucu sistemidir.
- Yayıncı ve abonenin mesajlara ulaşabileceği konular (topicler) ile benzer yapıda bir bloktur.
- Konulardan farklı olarak istemci tarafında gelen isteği (request) bekleyerek mesajlarını iletirler.
- Genelde işlem tabanlı ve hızlı tepki gerektiren işlemlerde kullanılırlar.

Düğüm en küçük işlem birimleri olarak düşünülebilir. Bir ROS içerisinde birden fazla düğüm birbiriyle haberleşerek çalışabilmektedir. Örnek olarak, bir düğümün kameradan görüntü alırken başka bir düğümün yol kestirimi yapmasını verebiliriz. Şekil 2.22'deki gibi farklı düğüm yapılarının haberleşebilmesi ve senkronize olarak çalışabilmeleri için master

yapısı bir sunucu görevi görmektedir. En basit anlamda ROS'un çalışma yapısını anlamak için yayınlama ve abonelik üzerine kurulu çalışma sisteminin anlaşılması gerekir. Şekil 2.22'de iki konuşmacı düğümü bir string tipinde bir mesaj üretir ve ilgili konu (topic) üzerinden yayınlanır. Dinleyici düğümü ise bu konuya abone olur ve mesajı (message) okur. Bu düğümler aynı Rosmaster üzerinde kayıtlıdır. Bu yapı sayesinde her düğüm aynı konuya bağlı olarak birbirinden bağımsız bir şekilde çalışmaktadır.

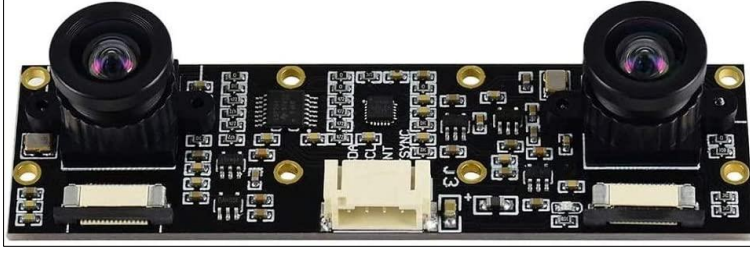
3. GELİŞTİRİLEN UYGULAMALAR

Tez kapsamında yapay zeka yöntemleri ve derinlik algılama sistemlerinden faydalanarak hareket edebilen ve engelden kaçınabilen otonom bir kara aracı geliştirilmiştir. Sistem pasif algılayıcılardan olan stereo kameralardan faydalanmaktadır. Stereo kameralar, Bölüm 2.8’de verildiği gibi birbirine paralel görüş alanı eksenleri olan benzer odak uzaklıklarına sahip iki adet kameradan oluşmaktadır. Bu kameraların birbirlerine olan mesafeleri sabittir ve gözlemlenen bölgede görüş alanlarını çakıştırarak bölgedeki nesnelerin derinlik değerlerinde çıkarım yapmaktadırlar. Otonom araçlar bu derinlik değerlerini hareketlerini sürdürmede, yol planlaması yapmada ya da konumlarının kestirimini gerçekleştirmede kullanabilirler. İnsansız araçlar için çevreyle olan etkileşimde anlamsal verilere de ihtiyaç vardır. Örneğin LIDAR sistemi kullanan bir otonom araç için zemine ait hiçbir anlamsal kestirimde bulunulamaz. Bu araçlar aktif yayılımlarının geri dönüşlerini toplayarak zamanın çok küçük değişimlerinde engellerin tespitini gerçekleştirmektedirler. İşte çevreyle olan bu etkileşimin sağlanması hususunda stereo ile derinlik tespitinin avantajlı olduğu durumlar bulunmaktadır. Benzer şekilde görünür bantta çalışan kamera verileri yapay zeka tabanlı çalışan semantik segmentasyon işlemlerine tabi tutularak kişi, yol, ağaç vb. sınıflandırma işlemleri gerçekleştirilebilmektedir.

Bu çalışma üç ana bölüme sahiptir. Bunlar sırasıyla; gerçek zamanlı stereo derinlik haritalandırılması, yapay zeka tabanlı semantik segmentasyon ile gerçek zamanlı yol tespiti ve ROS temelli çalışan donanımın gerçekleştirilmesidir.

3.1. Gerçek Zamanlı Stereo Derinlik Tespiti

Stereo görüntüleme ile derinlik tespitinin gerçekleştirilmesi için gerçek dünyaya ait sahnelerin stereo kamera yardımıyla elektriksel işaretler olarak alınması ve bu işaretlerin sayısal verilere çevrilerek matrisler içerisinde saklanması gerekmektedir. Ardından alınan bu matrisler aracılığıyla işlem yapabilen donanımsal birimler üzerinde istenen yöntemlerle işlenerek hedefe yönelik gerekli çıkarımlar yapılmaktadır.

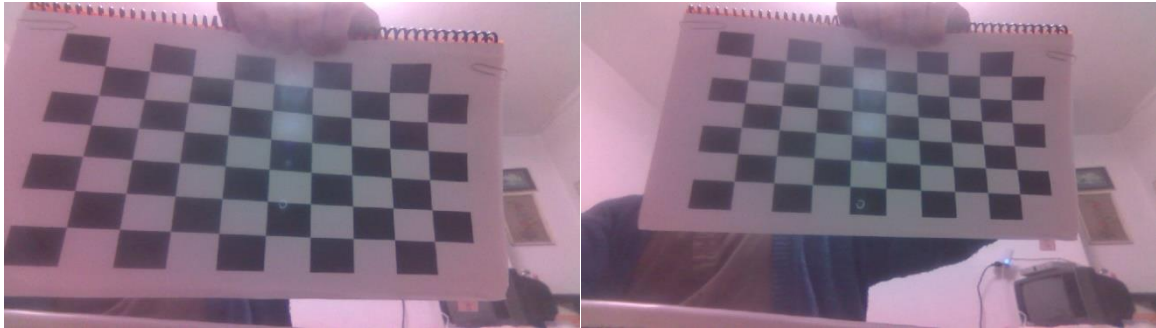


Şekil 3.1. Stereo kamera modülü IMX29-83

Bu çalışmada stereo kamera olarak IMX219-83 modülü kullanılmıştır (Şekil 3.1). Bu modül; 8 megapiksel, 3280 x 2464 çözünürlüklü, 83/73/50 derece (diyagonal/yatay/dikey) görüş açısına sahip, 60 mm taban mesafeli (baseline length), 2,6 mm odak uzaklığı bulunan 24 x 85 mm boyutlarında bir kameradır. Kamera üzerine entegre olarak bulunan CSI portu ve I2C arayüzünün SDA ve SCL pinleri kullanılarak görüntü alınmıştır.

Stereo kameradan alınan görüntülerin işlenmesi amacıyla NVIDIA Jetson Nano 4 GB yapay zeka geliştirme kartı kullanılmıştır. Bu kart üzerinde entegre iki adet CSI girişi sayesinde her iki kamera verisi de senkronize bir şekilde alınabilmektedir. Yine kart üzerine kablosuz bağlantı sağlanabilmesi için iki adet patch anten içeren wi-fi modülü takılmıştır.

16 GB'lık bir microSD kart üzerine Ubuntu 18.05 işletim sistemi kurlumu yapılmış ve gerekli güncellemeler gerçekleştirilmiştir. Öncelikli olarak kameradan CSI portu üzerinden görüntüler alınmış; böylelikle elde edilen görüntülerle fokus ayarları gerçekleştirilmiştir. Gerçekleştirilen fokus ayarı sonrası görüntülerin kırmızılıklara sahip olduğu görülmüştür (Şekil 3.2).



Şekil 3.2. Kırmızı renkli görüntü bozuklukları

Üretici firmanın düzeltme için sunduğu kamera düzeltme dosyaları (camera overrides) yüklenmiş ve bozulmalar (kırmızılık) düzeltilmiştir. İşletim sistemi üzerinde Python

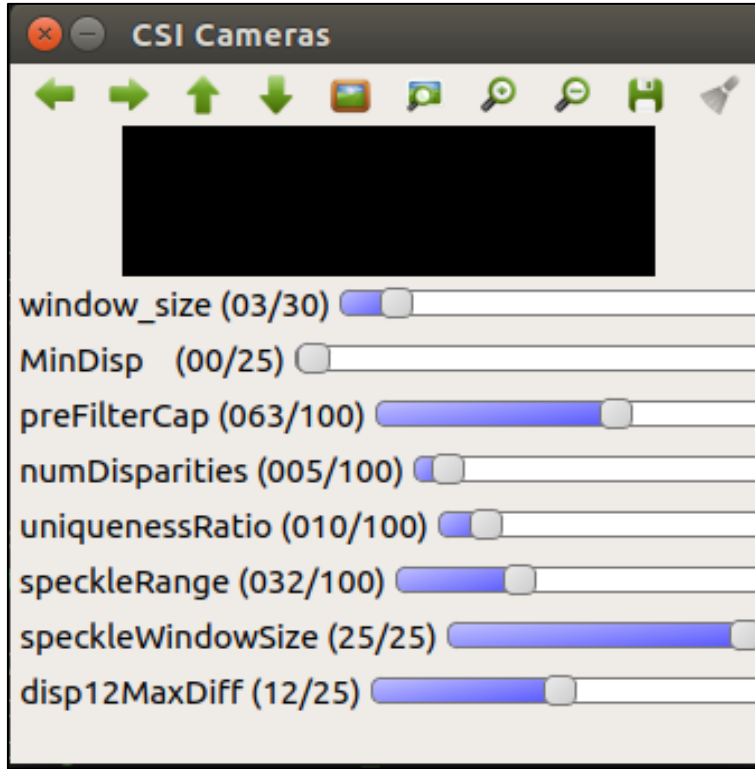
programlama dili kullanılmış ve VSCode geliştirme ortamı kullanılarak çalışma gerçekleştirilmiştir. Python ortamında OpenCV, PIL, numpy, threading, rospy, roslib, time, glob vb. kütüphaneler kullanılmıştır. Bu kütüphanelerin yükleme işlemleri için Jetson Nano üzerinde hafıza takas (memory swap) işlemleri gerçekleştirilmiştir ve kütüphane kurulumları sağlanmıştır.

Araştırma kapsamında stereo kameranın sol ve sağ kamerasından 960 x 540 boyutlarında üç kanal RGB görüntüler alınmış ve satranç tahtası yöntemi kullanılarak kalibrasyon matrisleri elde edilmiştir [58]. Kalibrasyon matrislerinin hesaplanması için sol ve sağ kameradan olmak üzere 15 tane farklı kalibrasyon standartlarına uygun satranç tahtası görüntüsü alınmıştır (Şekil 3.3).



Şekil 3.3. Kalibrasyonda kullanılan satranç tahtası görüntüsü örneği

Bu görüntülerden elde edilen matrisler hafıza üzerinde XML formatında kaydedilmiştir. Bu matrisler yardımıyla kalibrasyon (calibration) ve doğrultma (rectify) işlemleri gerçekleştirilerek doğrultulmuş kalibre görüntüler elde edilebilmektedir. Sağ ve sol kameraya ait doğrultulmuş bu görüntüler bu süreçte aynı epipolar düzlem üzerinde yer almaktadır. Bu görüntüler üzerinde blok eşleme yöntemleri kullanılarak farklılık değerlerinin hesabı gerçekleştirilebilmektedir. Bu araştırma kapsamında OpenCV kütüphanesi üzerinde BM (Block Matching) ve SGBM (Semi Global Block Matching) algoritmalarının çalışma performansı incelenmiştir. BM algoritması yüksek çerçeve hızı sunsa bile doğruluk olarak SGBM'den çok daha düşük olan bir algoritmadır [59]. Bu sebeple SGBM metodunda karar kılınmış ve farklılık (disparity) değerlerinin hesabı gerçekleştirilmiştir. Bu değerlerin hesaplanmasında kullanılan farklı parametrelerin sonuçlarının gözlemlenebilmesi amacıyla bir ara yüz hazırlanmıştır (Şekil 3.4).



Şekil 3.4. OpenCV SGBM parametreleri değişikliği için hazırlanan arayüz

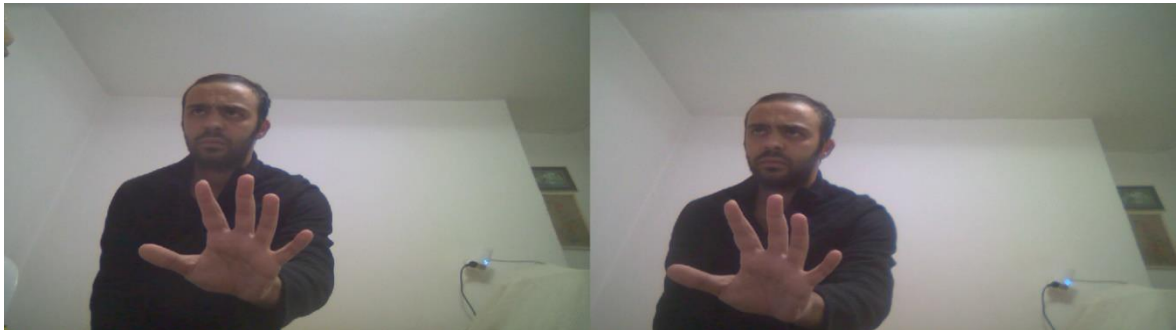
Stereo eşleme algoritmaları pek çok kısıttan etkilenmektedir. Bu kısıtların etkilerinin azaltılarak ölçülen derinlik haritasının doğruluğunu artırmak için farklı parametrelerle bu eşleşmeler hesaplanmaktadır. Blok eşleme gibi yöntemlerde çerçeve boyutu, farklılık sayısı ve eşsizlik parametreleri gibi parametreler elde edilen farklılık değerlerini etkilemektedir. OpenCV üzerinde oluşturulan StereoSGBM nesnesine ait değiştirilebilir parametreler tanımlanmıştır [60]. Tanımlanan bu parametreleri gerçek zamanlı görüntünün alımı sırasında değiştirilmiş; bu değişimin etkilerinin incelenmesi için Şekil 3.4'te gösterilen arayüz oluşturulmuş ve belirlenen aralık değerlerinde uygun parametre değerlerinin değiştirilerek denenmesi sağlanmıştır. Kullanılan parametreler ise aşağıdaki gibi açıklanabilir:

- minDisparity parametresi normalde sıfır olarak alınmaktadır. Ancak doğrultma algoritmaları gibi görüntü üzerinde değişikliğe sebep olan daha önce uygulanan görüntü işleme metodları sebebiyle görüntülerde kayma olmakta ve bu kayma miktarına göre bu değer ayarlanması gerekmektedir.
- numDisparities parametresi maksimum farklılıktan minimum farklılık çıkarılması ile bulunmaktadır. Bu değer sıfırdan büyük olmalı ve 16 ile bölünebilen bir değer olmalıdır.

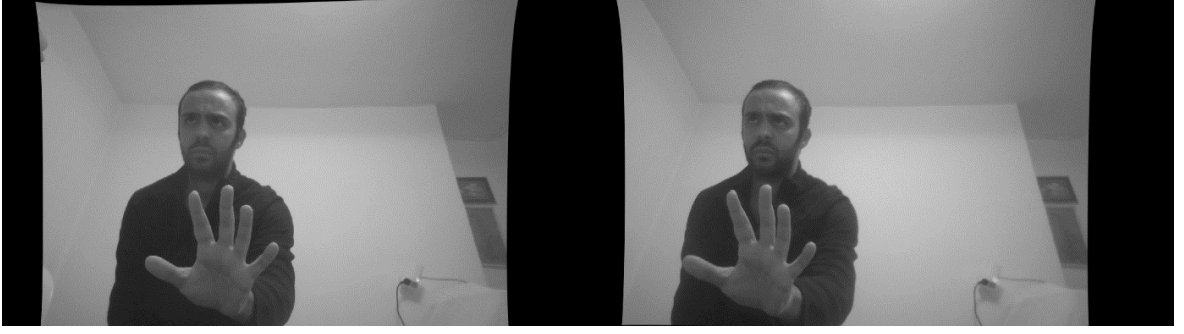
- blockSize parametresi windowSize olarak da bilinen eşleşen blokların boyutlarını belirler; 1'e eşit ya da büyük bir tek sayı ile ifade edilmektedir. Genelde iyi performans vermesi sebebiyle bu değer 3 ile 11 arasında belirlenmektedir.
- P1 ve P2 değerleri farklılıkların yumuşak geçişli olmasını sağlamaktadır. P2 değerinin P1 değerinden büyük olması gerekmektedir.
- Disp12MaxDiff değeri sol ve sağ görüntülerine ait farklılık hesabında maksimum izin verilen farktır.
- preFiltCap parametresi ile görüntü piksellerinin filtrelenmesi için gerekli eşik değeri belirlenmektedir.
- uniquenessRatio parametresi ise hesaplanan maliyet fonksiyonu değerinin, ikinci değerler karşılaştırılmasında yenebilmesi için gereken yüzdesel orandır. Bu değer 5 - 15 arasında bir değer alır ve görüntü kümesine ait benzersizlikleri ayırt etmede etkindir.
- SpeckleWindowSize değeri yumuşak geçişli bölgelerde yüksek değerleri elemek için belirlenecek değerdir. Geçişlerdeki yumuşaklığın büyüklüğünü belirlemektedir. Genelde 50 ile 100 arasında bir değer almaktadır. Bu değerın sıfıra çekilmesi durumunda filtreleme iptal edilmiş olur.

Tüm bu parametreler üzerinde deneysel olarak çalışılmış ve hem BM hem de SGBM algoritmalarına ait en iyi sonuç veren parametre değerleri belirlenmiştir. Bu tez çalışmasında SGBM parametreleri Şekil 3.4 ile verilen değerlerle kullanılmıştır.

SGBM metodunun kullanılması için sağ ve sol kameralardan RGB tonlamalı görüntüler alınmıştır (Şekil 3.5).

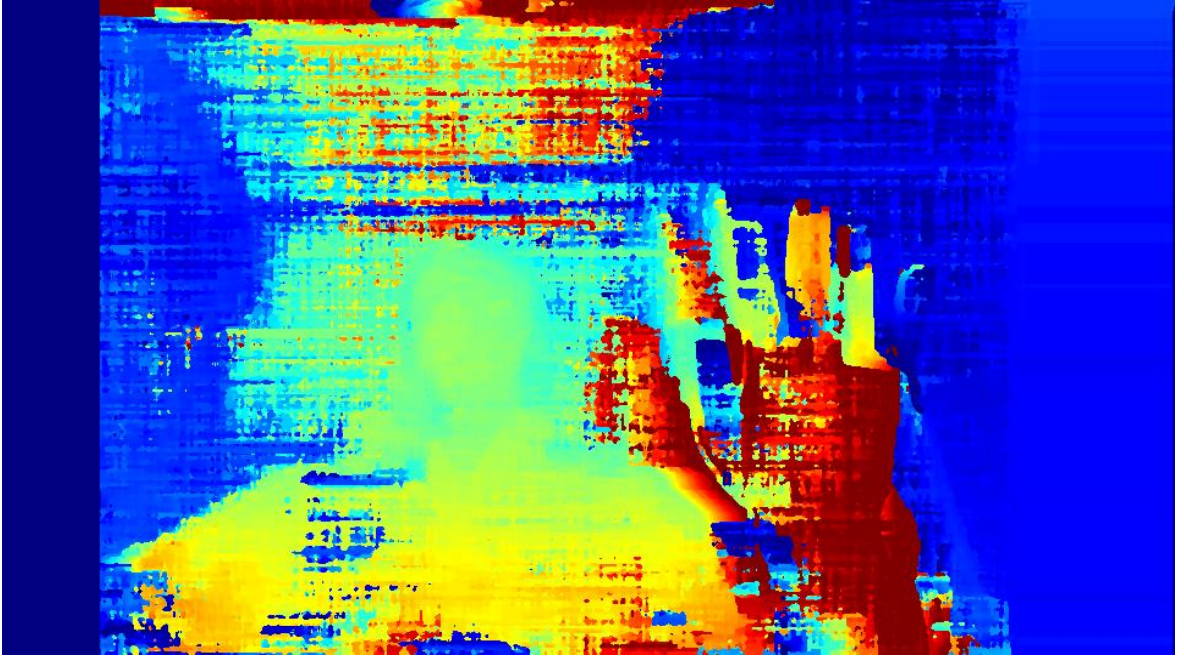


Şekil 3.5. Sol ve sağ görüntü örneği



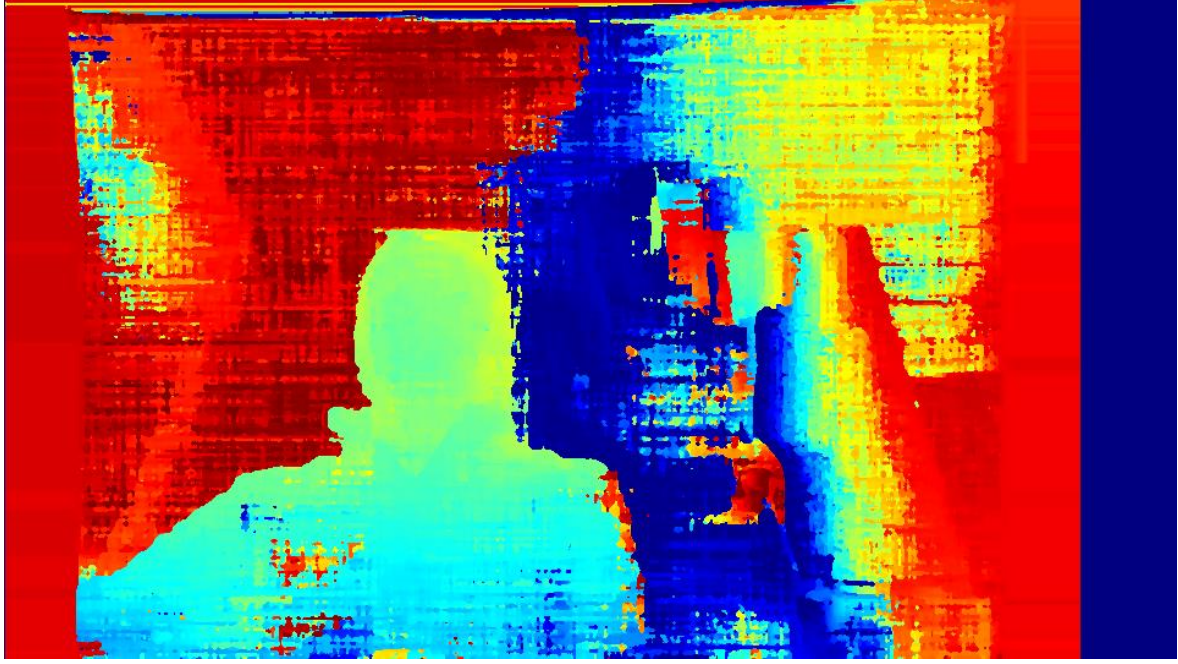
Şekil 3.6. Doğrultulmuş ve kalibre edilmiş gri tonlamalı sol ve sağ görüntü

Sol ve sağ kameradan alınan RGB tonlamalı görüntüler gri tonlamalı tek kanal görüntüler haline getirilmiştir. Bu görüntüler SGBM algoritmasına girdi olarak verilmeden önce kalibre edilmiş ve doğrultulmuş gri tonlamalı girdi görüntülerine dönüştürülmüştür (Şekil 3.6).



Şekil 3.7. Sol eşleme kullanılarak elde edilen görüntü

Sol kamera görüntüsü OpenCV üzerinde SGBM uygulanarak Şekil 3.7'deki farklılık değerleri elde edilmiştir. Farklılık değerleri üzerindeki değişimlerin daha iyi anlaşılmasını sağlamak amacıyla görüntülere renklendirme yapılmıştır.



Şekil 3.8. Sağ eşleme kullanılarak elde edilen görüntü

Şekil 3.8’de sağ görüntüler baz alınarak farklılık değeri çıkarımı gerçekleştirilmiştir. 2008 yılında Min D. ve arkadaşları tarafından WLS filtre kullanarak farklılık değerlerinin hesaplanması üzerine ve üst üste çakılı (occlusion) bölgelere ait farklılık değerlerinde iyileştirme öngören bir çalışma gerçekleştirilmiştir [61]. Bu çalışmada elde edilen farklılık değerleri ile sol görüntüye ait gri görüntü, WLS filtresinden geçirilerek iyileştirilmiş farklılık görüntüleri elde edilmiştir. Bu makaleden yola çıkarak derinlik değerleri üzerinde iyileştirme gerçekleştiren yöntem gerçek zamanlı çalışabilir hale getirilmiştir.



Şekil 3.9. WLS filtre çıktısı derinlik haritalanması

Şekil 3.9’da ağırlıklandırılmış en küçük kareler (WLS - Weighted Least Square) filtresi uygulanarak farklılık değerlerinde iyileştirme gerçekleştirilmiş ve farklılık değerleri renklendirilerek gösterilmiştir. Şekil 3.10’da kameraya yakın mesafede olan nesnelere ait piksellerdeki yüksek farklılık değerleri beyaz renk, uzaktaki nesnelere ait piksellerin düşük farklılık değerleri ise siyah renk ile gösterilmektedir. Derinlik değerlerinin kolay görüntülenebilmesi için gri tonlamalı derinlik görüntüleri üzerinde renklendirilme gerçekleştirilmiştir (Şekil 3.11).



Şekil 3.10. Nihai farklılık değerleri gri tonlamalı gösterimi

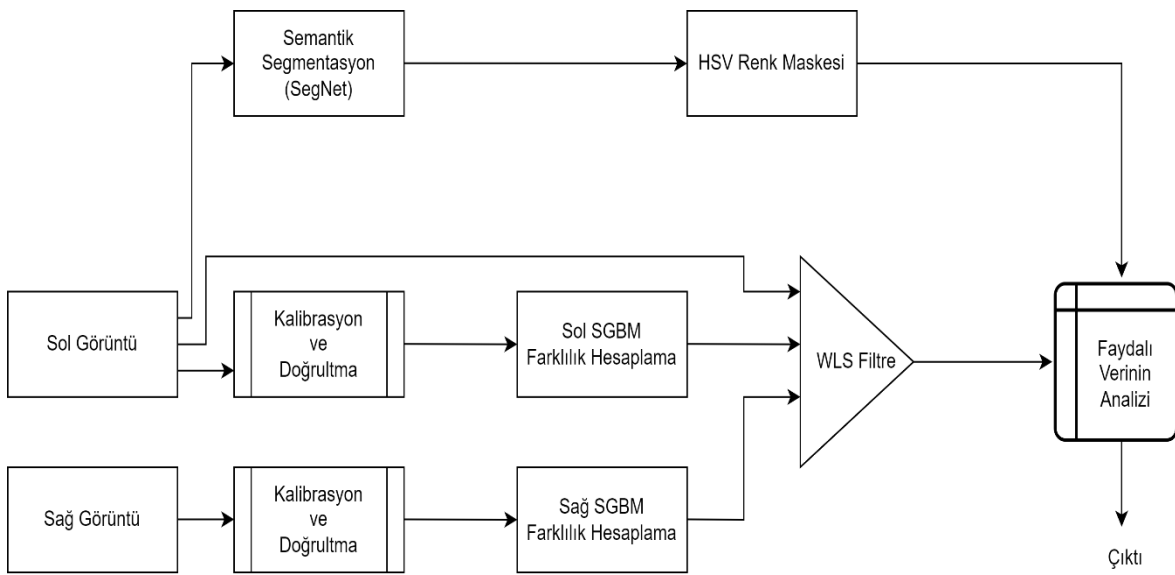


Şekil 3.11. Farklılık değerleri renklendirilmiş gösterimi

Elde edilen farklılık değerleri üzerinde Bölüm 2.8’de verilen Eş. 2.11’deki gibi odak uzaklığı ve taban mesafesi değerlerinden faydalanarak derinlik değerlerine geçiş sağlanmıştır. Görüntülerin eldesinden WLS filtre çıkışında derinlik görüntüsü eldesine kadar olan çalışma Şekil 3.12’de blok şema ile gösterilmiştir.

3.2. YZ Tabanlı Semantik Segmentasyon ile Gerçek Zamanlı Yol Tespiti

Elde edilen derinlik haritalaması otonom araç için en önemli belirleyicilerden biridir. Bu işlem sırasında aynı sahneye ait semantik segmentasyonun da gerçekleştirilmesi gerekmektedir. Derinlik hesabında kullanılan sol kamera görüntüsü semantik segmentasyon modeli için de girdi olarak beslenmektedir. Gerçekleşen işlemlere ait blok şema Şekil 3.12 ile gösterilmiştir. Ancak işlemleri gerçekleştiren donanımın performansının bu işlem için sınırlayıcı bir durum olduğu unutulmamalıdır.



Şekil 3.12. Sistem algoritması şematik gösterimi

Semantik segmentasyonun gerçekleştirilmesi amacıyla Jetson inference yapısı içerisinde optimize edilmiş segmentasyon algoritmasından faydalanılmıştır. Yapay zeka modeli SUN RGB-D veri seti kullanılarak eğitilmiştir [62]. Kodlayıcı (encoder) ve kod çözücü (decoder) yapısına sahip ResNet-18 yapay zeka mimarisinin ağırlıkları belirlenmiş ve kaydedilmiştir. Daha önce ağırlıklandırılmış bu çıktı, sistem çalıştırıldığında yüklenerek model girdi görüntülerini bölütlemeye hazır hale getirilmiştir. Oluşturulan bu model sol kamera tarafından alınan her çerçeve görüntüsü tarafından beslenerek çıktı üretmektedir. Elde edilen segmentasyon görüntüsü piksel bazlı olarak zemin (floor) sınıfına ait pikselleri bölütleyebilmektedir.



Şekil 3.13. Test alanında semantik segmentasyon sonuçları

Şekil 3.13'ten de görülebileceği gibi zemine ait pikseller yeşil renk ile renklendirilmiştir. Renklendirilmiş bu görüntüde ayrımın gerçekleştirilmesi için RGB'den HSV'ye çevrim yapılmıştır ve OpenCV maskeleye fonksiyonlarından faydalananak HSV renk maskeleye gerçekleştirilmiştir.

Bu maske görüntüsü her bir piksel için intensity (yoğunluk) değeri olarak 1 ve 0'lardan oluşan bir maske oluşturmaktadır. Aracın ilerleyebileceği zemine ait olan pikseller 1 değerine, diğerleri ise 0 değerine sahiptirler.

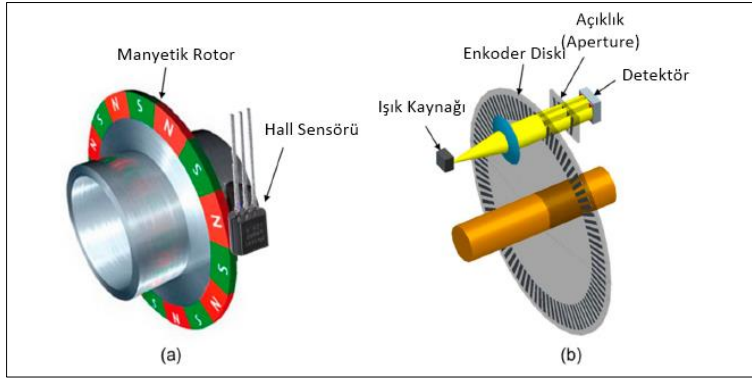
Pre-Process	CPU	0.06651ms	CUDA	2.52630ms
Network	CPU	44.88515ms	CUDA	41.30526ms
Post-Process	CPU	1.24682ms	CUDA	1.57562ms
Visualize	CPU	0.08687ms	CUDA	10.76068ms
Total	CPU	46.28536ms	CUDA	56.16787ms

Şekil 3.14. Semantik segmentasyon işlem süreleri

Bu aşamaya kadar Jetson Nano üzerinde derinlik haritalaması çıktısı ve yapay zeka temelli semantik segmentasyon çıktısı üretilmiştir. Çıktı 24 FPS değerlerinde çalışmaktadır ve işlemlerde geçen zamanlar Şekil 3.14 ile verilmiştir. Elde edilen derinlik ve semantik segmentasyon görüntüleri faydalı verinin analizi bloğunda incelenerek ROS için gerekli çıktı üretilmiştir. Bu veri ROS tarafından kullanılabilmesi için anlık olarak kaydedilir. Veri üzerindeki inceleme ve ROS sistemi Bölüm 2.9'da anlatılmıştır.

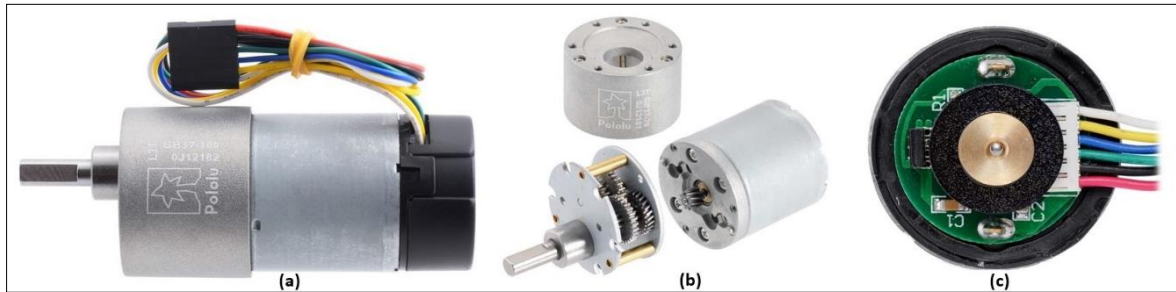
3.3. ROS Tabanlı Donanımın Oluşturulması

Bölüm 3.2’de verildiği gibi Jetson Nano yapay zeka geliştirme kartı üzerinde elde edilen görüntüler ve bunlar üzerinde gerçekleştirilen işlemler sonrası elde edilen nihai sonuçlar değerlendirilerek otonom sürüşün gerçekleştirilmesi amaçlanmıştır. Bu iş kapsamında basit bir 2 tekerlekli araç (2 wheels car) tasarımı planlanmıştır. Aracın gövdesi, işlenmesinin kolay olması ve maliyetinin düşüklüğü sebebiyle pleksiglas malzemedan oluşturulmuştur. Bu taban levhası üzerindeki yere dik bölümlerde araç tekerleklerinin montajı için yuva açımı ve motor sabitlenmesi için 3 mm çapında vida delikleri açılmıştır. Proje kapsamında hareket anında motora ait hareketin miktarının ve yönünün tespit edilmesi gerekmektedir. Araç tekerleklerinin hareketinin sağlanması için motor araştırması gerçekleştirilmiştir. Öncelikli olarak Nema-17 step motor ile adım kontrollü olarak tekerlek hareketi sağlanmış ancak hareketin fiziki olarak geri dönüşü okunamamıştır. Step motor kullanılması durumunda sistem kendi verdiği emrin karşılığını okuyarak değil; varsayımsal olarak konum tespiti gerçekleştirebilecektir. Bu gibi durumlarda ya fiziki hareket kontrolü tespit sisteminin geliştirilmesi gerekir ya da işlemci tarafından verilen komutun fiziki olarak değişimini geri besleme şeklinde ileten bir motor seçimi yapılması gerekmektedir. Seçenekler arasından en uygun motorun enkoderli doğru akım (DC) motorları olacağına karar kılınmıştır. Enkoderler döner veya doğrusal hareket gerçekleştiren sistemlerde hareketi tespit eden ve bunu sayısal işaretlere dönüştüren sistemlerdir. Hız, mesafe, yön vb. hareket parametreleri enkoder sistemlerinin ürettiği sinyaller yardımıyla izlenmektedir. Enkoder sinyallerinin oluşturulmasında mekanik, optik, manyetik vb. sensörler kullanılmaktadır. Bu sensörlerin birbirlerine karşı avantaj ve dezavantajları bulunmaktadır. Bu sebeple tez çalışmasına uygun sensör araştırılmıştır. Piyasada gerçekleştirilen araştırmada yaygın olarak optik ve manyetik enkoderli iki tip DC motor kullanıldığı görülmüştür.



Şekil 3.15. a) Manyetik ve b) optik enkoder iç yapısı [63]

Şekil 3.15 ile optik ve manyetik enkoderlerin çalışma sistemleri verilmiştir. Şekil 3.15.b’de görüldüğü gibi optik enkoderlerde bir ışık kaynağından yayılan ışık motora sabitlenmiş bir dönüş hareketi yapan enkoder diskinden geçirilerek dedektör tarafından alınır. Dedektörün aldığı bu ışık, elektrik sinyaline dönüştürülür. Manyetik enkoderlerde ise dönel disk sıralı S-N kutuplu manyetik bir rotordur. Buradaki dönel hareketle değişen manyetik değerler Hall etkisine dayanarak çalışan bir sensör tarafından ölçülmekte ve sinyale dönüştürülmektedir (Şekil 3.15.a).



Şekil 3.16. a) Enkoderli DC motor, b) redüktör ve c) enkoder yapısı

Bu çalışma kapsamında 2 adet 12 V, 37 mm, 100 RPM değerlerine sahip 100:1 redüktör oranlı DC motor kullanılmıştır (Şekil 3.16.b ve 3.16.c). Şekil 3.16.c’de gösterilen enkoder devresi sayesinde motorun dönme hızı, dönme yönü ve tur sayısı değerlerine ulaşılabilir. Motor boşa 200 mA akım çekmektedir ve 34 kg/cm zorlanma torkuna sahiptir. Bu motorun üzerindeki enkoder 64 CPR’lık bir çözünürlüğe sahiptir. CPR devir başına döngü sayısıdır. Motor dişli oranı 102,083:1 olarak verilmiştir. Bu iki değer çarpılarak motorun bir tam turunu tamamladığından enkoderden alınacak değer 6533 darbe olarak bulunacaktır. Şekil 3.16.c’de gösterilen bağlantı kablolarından siyah (-) ve kırmızı (+) motor gücü, mavi (+) ve yeşil (-) enkoderin güç kablosu, sarı ve beyaz renk ise sensörün A

ve B kanalındaki çıkışlarıdır. Motorun çalışma gerilimi 12 V'tur ve zorlama akımı olarak 5,5 A ihtiyaç duymaktadır. Enkoder sistemi 3,5 – 20 V arası çalışabilmekte ve 10 mA civarında akım çekmektedir.



Şekil 3.17. 12 V DC (doğru akım) motora ait 64 CPR enkoderin A ve B kanalı osiloskop görüntüsü [64]

Şekil 3.17’de görülebileceği gibi çıkış olarak A ve B kanallarından 0 V ile besleme gerilimi aralığında 90° faz farklı kare dalgalar üretilmektedir. Bu iki kanal arasındaki öncelik motorun yönünü, kare dalgaların sayısı ise motorun hızını vermektedir.

DC motorlar besleme gerilimlerinin ayarlanabilmesi, motor hız kontrolü, motor yön kontrolü ve motor sürücülerine ihtiyaç duymaktadırlar. Bu amaçla elektronik projelerde sıkça kullanılan L298n sürücü kartında karar kılınmıştır. L298n, doğru akım ve step motor kontrolünde sıkça kullanılan birbirinden bağımsız olarak iki motoru kontrol edebilen bir sürücü kartıdır. 24 V’a kadar olan motorlar bu kart yardımıyla sürülebilir ve kanal başına 2 A’lık maksimum akım verebilmektedir. L298n dahili regülatörü, soğutucusu, kısa devre koruması, motor dönüş yönü kontrolü gibi özellikleri sebebiyle tercih edilmiştir. Yapılan

deneylerde motorun anlık akımlarının yüksek olması ve sürücü kartın maliyetinin ucuz olması sebebiyle her motorun ayrı bir sürücü ile beslenmesinde karar kılınmıştır [65].

Motor kontrolleri için sürücülerle iletişimde olacak, gerekli olan veri okuma ve yazma işlemlerini yapacak mikroişlemci olarak öncelikli olarak Arduino Nano kullanılmış fakat bu işlemcinin yeterli kesme pini olmaması sebebiyle Arduino Mega 2560'a geçiş yapılmıştır. Arduino Mega üzerindeki 2, 3, 18, 19, 20, 21 numaralı pinler enkoder verilerini hızlı bir şekilde okuyabilmek amacıyla kesme pinleri olarak kullanılabilir. Arduino Mega mikrodenetleyici olarak Atmega2560 kullanılmaktadır. Bu mikrodenetleyici, 15 adet PWM (Pulse Width Modulation) çıkış üretebilen 54 sayısal pine, 16 analog giriş pinine, 16 MHz'lik saat hızına sahip 10,6 x 53,4 mm'lik 36 gram ağırlığında bir devre kartıdır. Bu kart üzerinden seri haberleşme yapılabilmesi, bu port üzerinden programlanabilmesi ve açık kaynak kodlu arayüzü bulunması sebebiyle tercih edilmiştir. Arduino Mega'nın bu sistemde 3 ana görevi bulunmaktadır. Bu görevlerden ilki kullanıcının ya da otonom sistemin hareketinin bilgisini tespit etmektir. Görevlerinden ikincisi, bu hareket isteri karşısında işlemleri gerçekleştirerek motor sürücülerine gerekli sinyaller göndermek ve motorların dönüş hızlarını ve yönleri belirlemektir. Son görevi ise dönüş gerçekleştiren motorların enkoder kanallarından gelen bilgiyi okumaktır. Tüm bu görevlerinin yanı sıra Arduino Mega 2560 işlemci kartının 5 V ve ground çıkışları devre içerisinde düşük akımlı güç kaynağı görevi de görmektedir. Arduino, enerjisini seri port konektörü üzerinden bağlı olacağı asıl yönetici birim olan Raspberry Pi 4 işlemcisinden almaktadır.

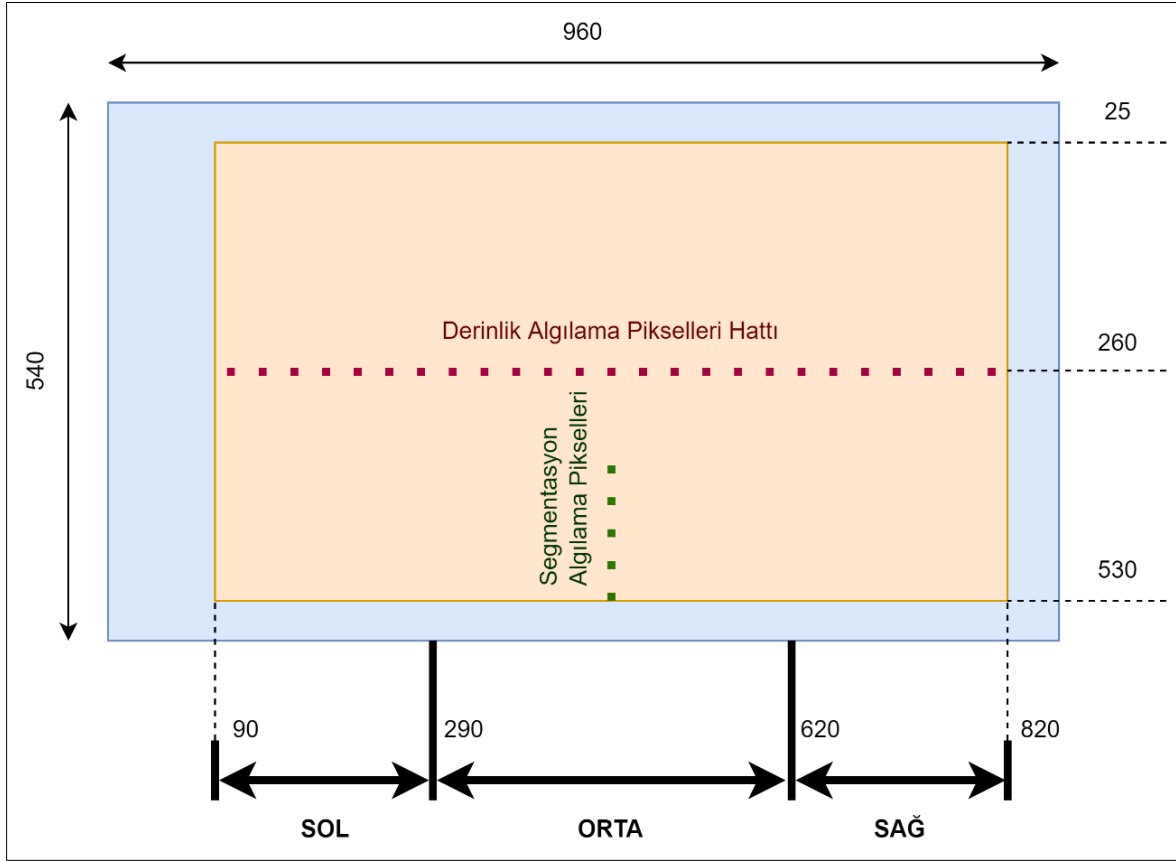
Sistemin ilk çalışmaları doğru akım üreteçleri üzerinde masaüstünde gerçekleştirilmiştir. Araç üzerinde 12 V'luk güç kaynağı ihtiyacının karşılanması amacıyla 12 V 2.8 Ah TNL Kuru Tip Akü tedarik edilmiştir. Bu aküler çok sayıda doldur-boşalt yapılabilen harici güç kaynaklarıdır. 70 x 48 x 98 mm boyutlarındadır.

Arduino Mega ve sürücülerle kurulu bu hareket sisteminin kontrol komutları tek kart bilgisayar olarak tanımlanan Raspberry Pi üzerinden gelmektedir. Bu projede Raspberry Pi 4 kartının 2 GB RAM özellikli olanı kullanılmıştır. Raspberry Pi 4 kartı üzerinde dört adet USB portu barındırmakta ve bunlardan iki tanesi yüksek hızlı aktarıma (USB 3) izin vermektedir. Yine üzerinde iki adet 4K yayına izin veren micro HDMI portu, ethernet portu, GPIO pinleri ve güç için USB-C tipi güç girişi içermektedir. Üzerinde bulunan micro SD kart girişi yardımıyla işletim sistemi kurulumu yapılarak masaüstü bir bilgisayar gibi

kullanılabilmektedir. 4 çekirdekli ve 64 bit işlemcili kart 1,5 Ghz frekansında çalışmaktadır. Raspberry Pi 4 kartı dahili olarak kablosuz ağ ve bluetooth destekli olarak gelmektedir. Bu sistemde Raspberry Pi 4 kartının 2 temel görevi bulunmaktadır. Bunlardan ilki, Arduino için gerekli enerjiyi üreterek Arduino ile USB portu üzerinden seri haberleşme kurmasıdır. Bu seri haberleşme sayesinde, Raspberry Pi 4 Arduino'nun motor kontrolü için gerekli olan komutları gönderecek ve Arduino'nun okuduğu sol ve sağ motora ait enkoder verilerini alacaktır. Bu veriler sayesinde aracın odyometrik haritalanması sağlanacaktır. İkinci görevi ise, Bölüm 3.2'de anlatılan Jetson Nano'nun görüntü işleme sonrası yayınlayacağı verileri kablosuz ağ üzerinden okuyarak otonom sürüş için gerekli verileri üretmek olacaktır.

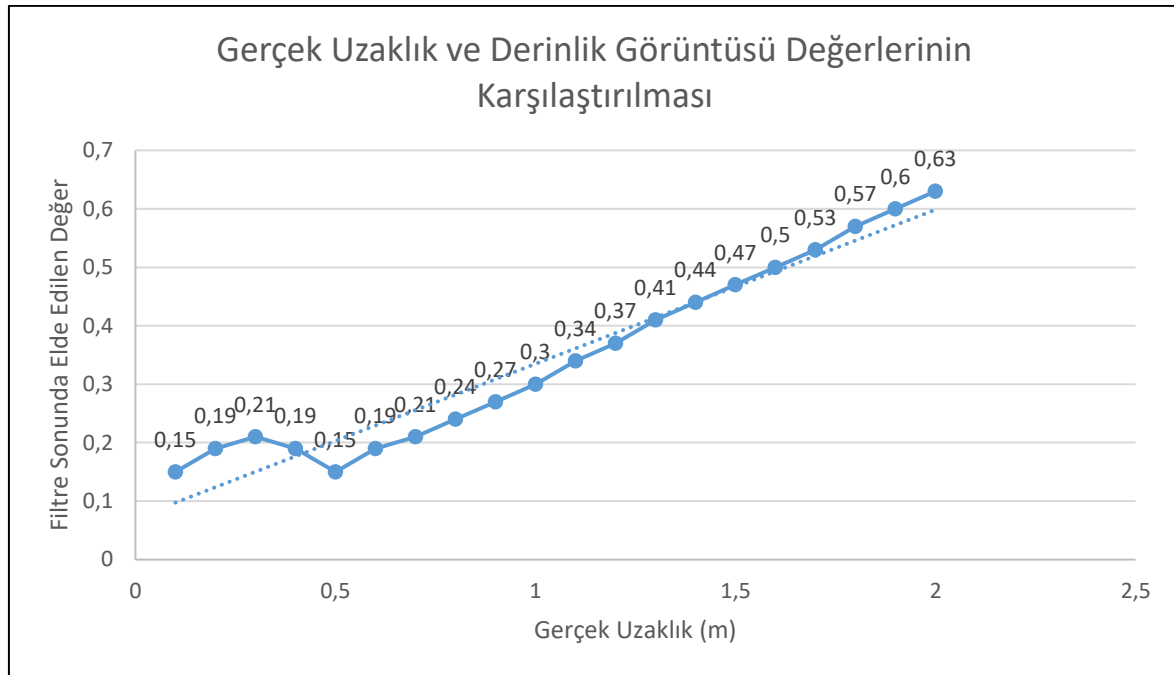
Sistemin ilk testleri motor sürücü testlerinde olduğu gibi masaüstünde gerçekleştirilmiş ve bu aşamada Jetson Nano ve Raspberry Pi 4 kendi güç adaptörleri ile şebekeden beslenerek çalıştırılmıştır. Ancak araç üzerinde mobil bir sistem kurulabilmesi amacıyla her iki işlemciye de USB çıkışları üzerinden 5 V'luk maksimum 3 A'lik çıkış verebilen bir batarya temin edilmiştir. Bu güç kaynağı her iki portundan da işlemcilerin ihtiyacı olan maksimum 3 A'lik çıkışı desteklemektedir. 10000 mAh'lık 15 W'lık Lityum Ion 20 x 16 x 8 cm boyutlarında taşınabilir bir şarj cihazı kullanılmıştır.

Derinlik görüntüsü ve segmentasyon çıktısı Bölüm 3.2'de belirtilen şekilde elde edilmiştir. Bu görüntüler üzerinde faydalı veri çıkarımı gerçekleştirilmiştir.



Şekil 3.18. Görüntü piksel ayrıştırması şeması

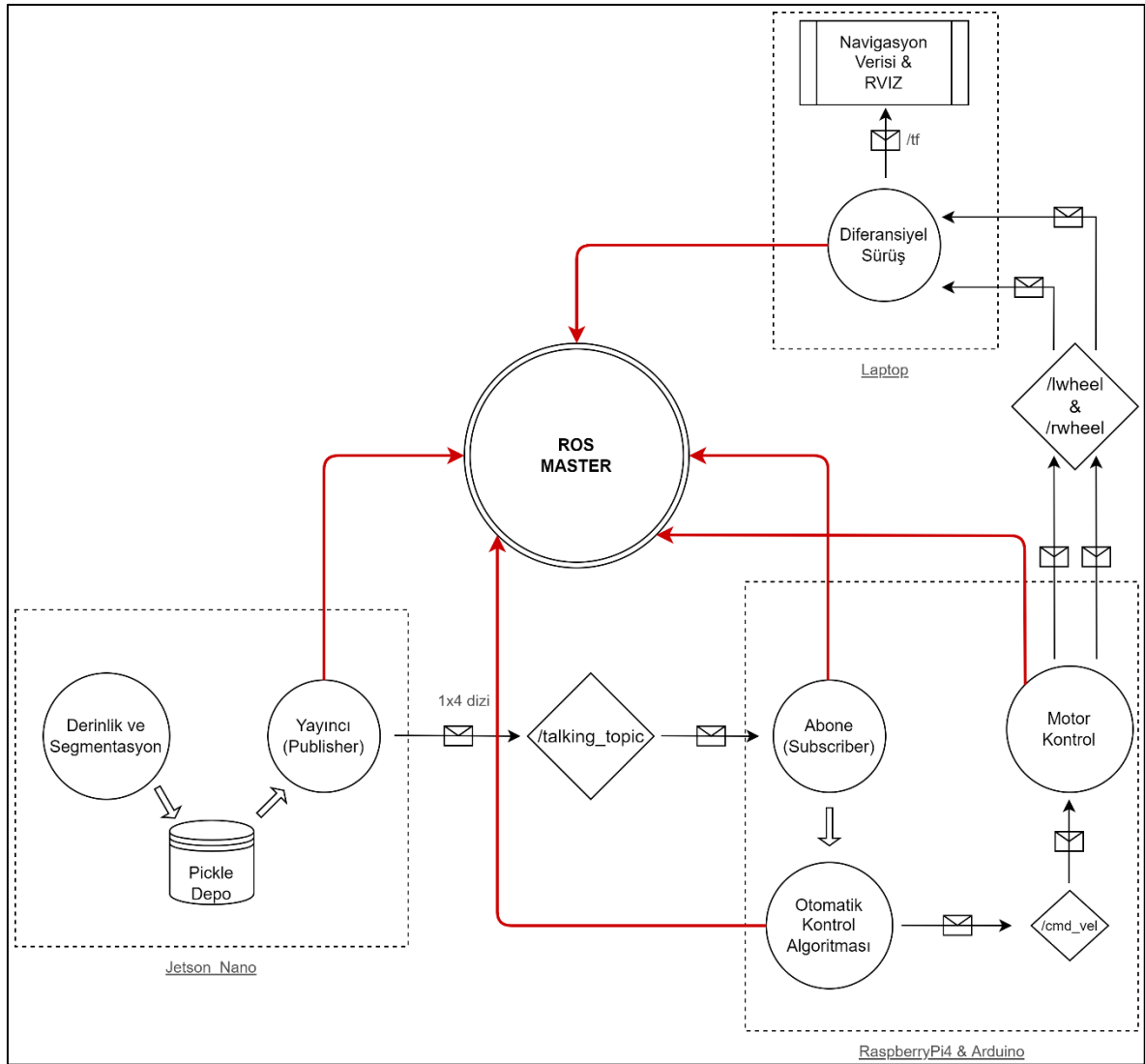
Araç üzerindeki kameranın hizası ve görüş alanı dikkate alınarak tüm bir satıra ait pikseller işlem hızını artırmak adına her 10 pikselde bir örneklenmiştir (Şekil 3.18). Derinlik görüntüsündeki bu piksellerin disparity değerlerinden derinlik değerleri elde edilmiştir. Fakat gerçekleşen filtreleme işlemlerinin derinlik değerlerinde kaymalara sebep olduğu tespit edilmiştir.



Şekil 3.19. Gerçek uzaklık ve filtre çıktısı derinlik sonucu karşılaştırması

Bu sebeple mesafeye ait derinlik çıktılarının değerleri deneysel olarak belirlenmiştir. Şekil 3.19’da görülen derinlik algılama pikselleri hattında yer alan bu örneklenmiş piksellerin bir alt fonksiyonda Şekil 3.19’da gösterilen 3 farklı bölge için (sol, orta ve sağ bölgeler) ayrımı yapılmaktadır. Bu piksellerin bu aşamada belirlenen eşik mesafesini ihlal edip etmedikleri kontrol edilmekte ve kaç adet pikselin bu eşik mesafesini geçtiği sayılmaktadır. Şekil 3.19’da verilen grafikte de görülebileceği gibi kararlı doğrusal durumun başladığı eşik değeri olarak 0,2 seçilmiştir. Böylelikle araç, önündeki engeli gecikme sebebiyle tespit edememesi durumunda 0,2 değerini yakınlamayla ikinci kez test ederek engel tespiti gerçekleştirebilmektedir. Bu işlem sonucunda ilgili görüş alanına ait belirlenen d mesafesinden yakında olan engelin tespiti gerçekleştirilmektedir. Bu çalışmada eşik değeri olarak belirlenen 0,2 değerine karşılık engele yaklaşma ile 2 farklı d mesafesi değeri test edilebilmektedir. Bunlardan ilki yaklaşık olarak 0,25 m, ikincisi ise yaklaşık olarak 0,65 m’dir. Sonuç olarak 1 x 3’lük bir integer dizisi elde edilir. Bu dizide ilk terim sol bölgedeki d mesafesinden yakın engel piksel sayısını, ikinci terim orta bölgedeki d mesafesinden yakın piksel sayısını ve son terim ise sağ bölgedeki d mesafesinden yakın piksel sayısını ifade eder. Buna ek olarak segmentasyon çıktısı da sol kameradan alınan araç hareket bölgesindeki piksellerin zemin düzlemine ait olup olmadığı çıktısını üretir. Alınan piksellerin ortalaması hesaplanarak yapay zeka modelinin kararsızlıkları ortadan kaldırılmaya çalışılmıştır. Hesaplanan ortalama değerinin 1’den büyük olması durumunda boolean 1 (True) değeri

döndürülerek ilerlenecek güzergahta bir yol bulunduğu verisi oluşturulmaktadır. Daha önce oluşturulan matrise bu aşamada bulunan 1 - 0 değeri de eklenerek her frame için 1 x 4 boyutlarında bir dizi elde edilmektedir. Elde edilen bu matris değeri başka bir alt fonksiyon yardımıyla pickle dosya formatına çevrilerek saklanır. ROS tarafında oluşturulacak yayıncı (publisher) bu pickle binary dosyasını okuyarak işlemlerini gerçekleştirmektedir.

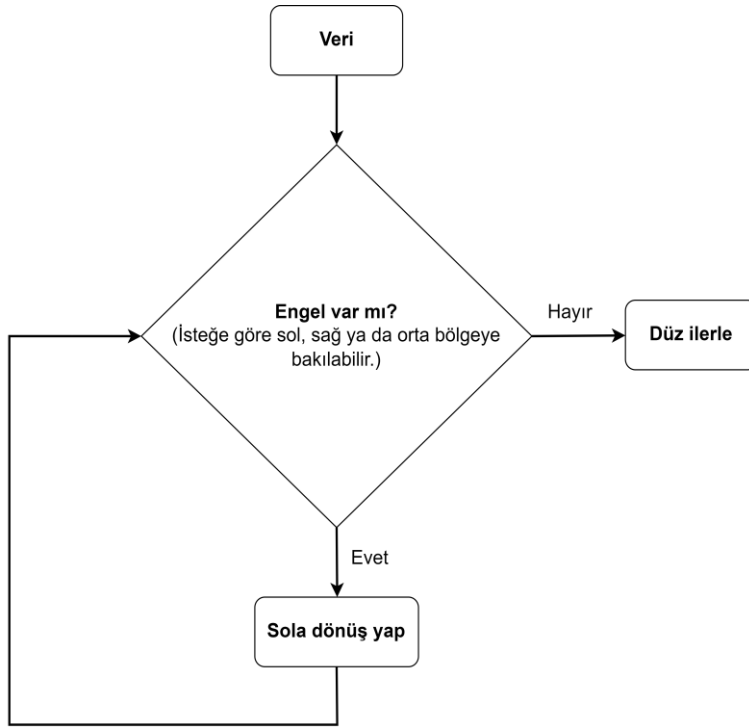


Şekil 3.20. Çalışmada tasarlanan ROS sistemi şematik gösterimi

Şekil 3.20’de ROS sisteminin çalışma diyagramı ana hatlarıyla gösterilmiştir. Yayıncı, abone, otomatik kontrol algoritması ve diferansiyel sürüş isimli düğümler (node) ROS Master’a kayıtlı (registered) durumdadırlar. Sistem üzerinde dört farklı işlemci bulunmaktadır. Bunlar kesikli çizgilerle gösterilmiş Jetson Nano, Raspberry Pi 4, Arduino Mega 2560 ve dizüstü bilgisayar (laptop) ’dır. Pickle dosyasında depolanan veri yayıncı

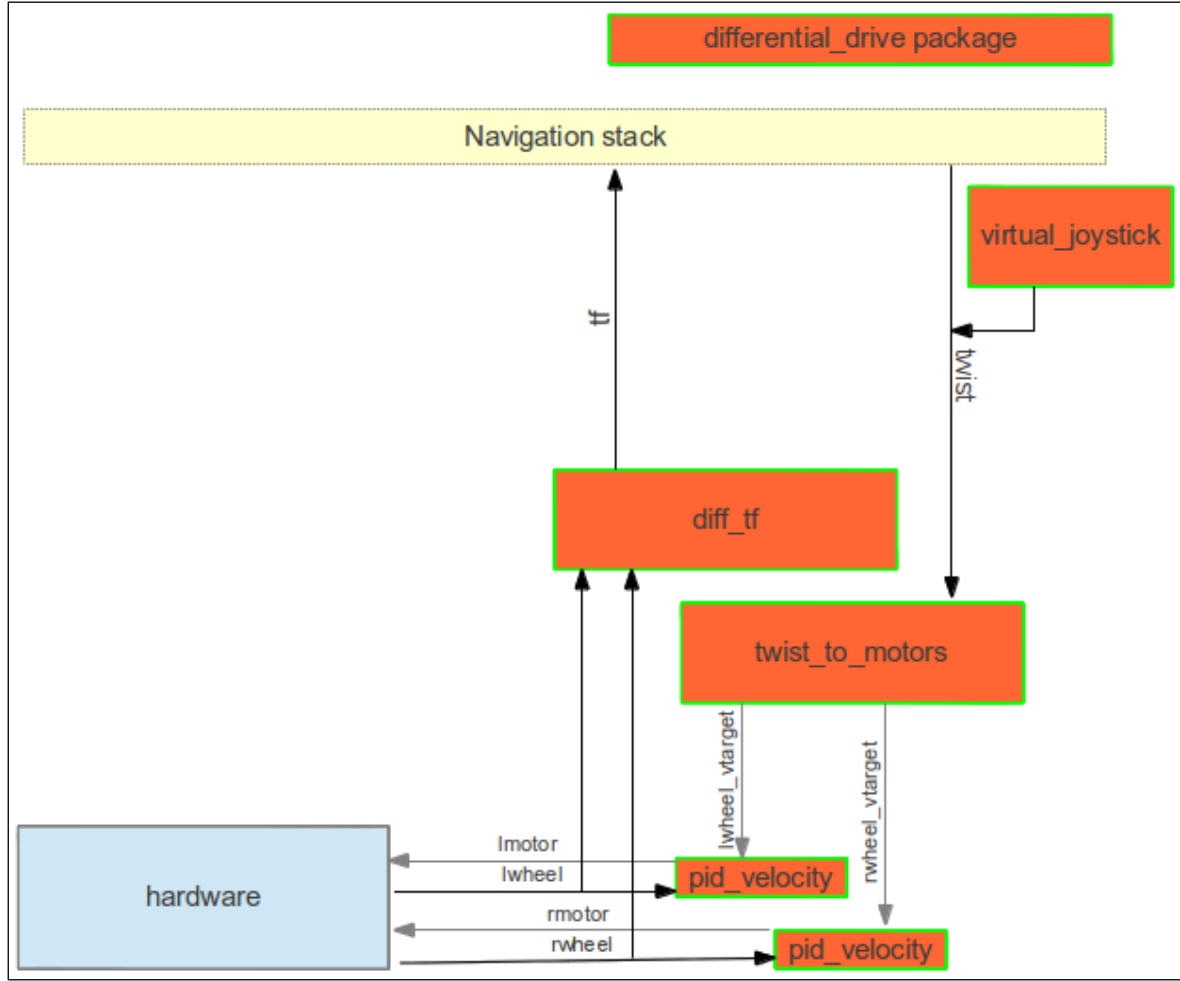
tarafından okunur ve `talking_topic` isimli konuda (topic) yayınlanır. Burada gönderilen mesaj, integer list türündedir. Bu topic'in abonesi olan Abone düğümü Raspberry Pi üzerinde yer alır ve aldığı bu integer list verisini otomatik kontrol algoritmasına ileterek kontrol için gerekli olan verinin üretilmesini sağlar.

Yol planlama ve ortam haritalama, güncel ve üzerinde yoğun olarak çalışılan çok geniş kapsamlı konulardır. Bu tez kapsamında engelden kaçınma işleminin gerçekleştirilmesi için Şekil 3.21 ile verilen algoritma oluşturulmuş ve otomatik karar mekanizması bu algoritma dahilinde gerçekleştirilmiştir.



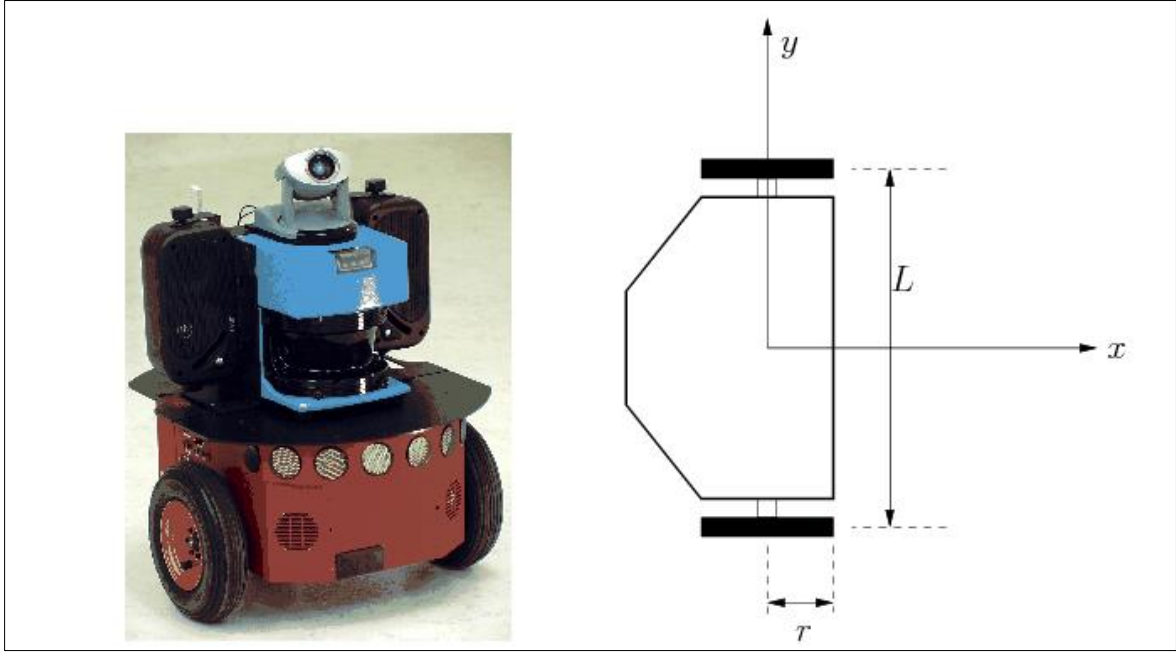
Şekil 3.21. Kaçınma algoritması

Raspberry Pi tarafından kaçınma algoritması yardımıyla üretilen bu komutlar Arduino ile seri porttan okunarak motor sürücülerine gerekli olan talimatlar verilir ve kontrol sağlanır. Motor kontrol algoritması “*lwheel*” ve “*rwheel*” isminde iki adet yayın yapmaktadır. Bu yayınlar motorların enkoder verilerini ve yönlerini göstermektedir. Şekil 3.22’de enkoder verilerinden RVIZ için odyometrik veri üreten ROS paketinin gereksinimlerini göstermektedir.



Şekil 3.22. Diferansiyel sürüş ROS şematik gösterimi [66]

Şekil 3.23'te iki tekerlikli bir gezing robotu ait görüntü ve üstten görünüşünün teknik çizimi verilmiştir. Burada L değeri iki teker arasındaki mesafe, r değeri ile tekerleklere ait yarıçaptır. Şekil 3.23'te x ve y koordinatları gösterilmektedir.



Şekil 3.23. İki tekerlekli araç görüntüsü ve üst görünüşüne ait koordinat sisteminde değerlerinin gösterilmesi [67]

Harekete ait vektör her iki tekerleğe bağlı olarak $u = (ur, ul)$ 'dir. Çizelge 3.1 ile gösterildiği gibi bu değerler arasındaki ilişkiye göre aracın hareketi şekillenmektedir.

Çizelge 3.1. Hareket yönleri ve vektör bağlantısı

$ul = ur > 0$	İleri Yönde Düz Hareket
$ul = ur < 0$	Geri Yönde Düz Hareket
$ul = -ur \neq 0$	Saat Yönünde Dönüş Hareketi
$-ul = ur \neq 0$	Saat Yönünün Tersine Dönüş Hareketi

Bu dönüşlerdeki çevirme hızı tekerleklerin açısal hızlarının ortalamasına bağlıdır. Θ değeri tekerleklerin açısal hızlarını temsil etmektedir.

Jon Stephan tarafından geliştirilmiş bir Arduino Kütüphanesi sayesinde bu yayınlardan tf verisi üretilir. Bu veri üretilirken Eş. 3.1, Eş. 3.2 ve Eş. 3.3 ile verilen geçiş denklemlerinden faydalanarak aracın odyometrik verileri elde edilmektedir [68].

$$\dot{x} = \frac{r}{2}(ul + ur)\cos\theta \quad (3.1)$$

$$\dot{y} = \frac{r}{2}(ul + ur)\sin\theta \quad (3.2)$$

$$\dot{\theta} = \frac{r}{L}(ur - ul) \quad (3.3)$$

Elde edilen bu odyometrik sahne verileri RVIZ üzerinde gerçek zamanlı olarak gözlemlenebilir.

4. BULGULAR VE İRDELEME

Bu bölümde iki alt başlık altında iç mekan ortamlarında otonom hareket gerçekleştiren robotun sahada aldığı görüntüler, detaylı araç yerleşim planı ve gerçek zamanlı harekete ait sonuçlar gösterilmiştir.

4.1. Robotun Sahada Aldığı Görüntüler

Gerçekleştirilen çalışmada bilgisayarlı görü donanımlarının otonom araç üzerine entegrasyonu gerçekleştirilerek hareketli iç mekân testlerindeki sonuçları kaydedilmiştir. Araç üzerindeki stereo kameranın sağ ve sol kameralarından alınan RGB görüntüler Şekil 4.1’de verilmiştir.



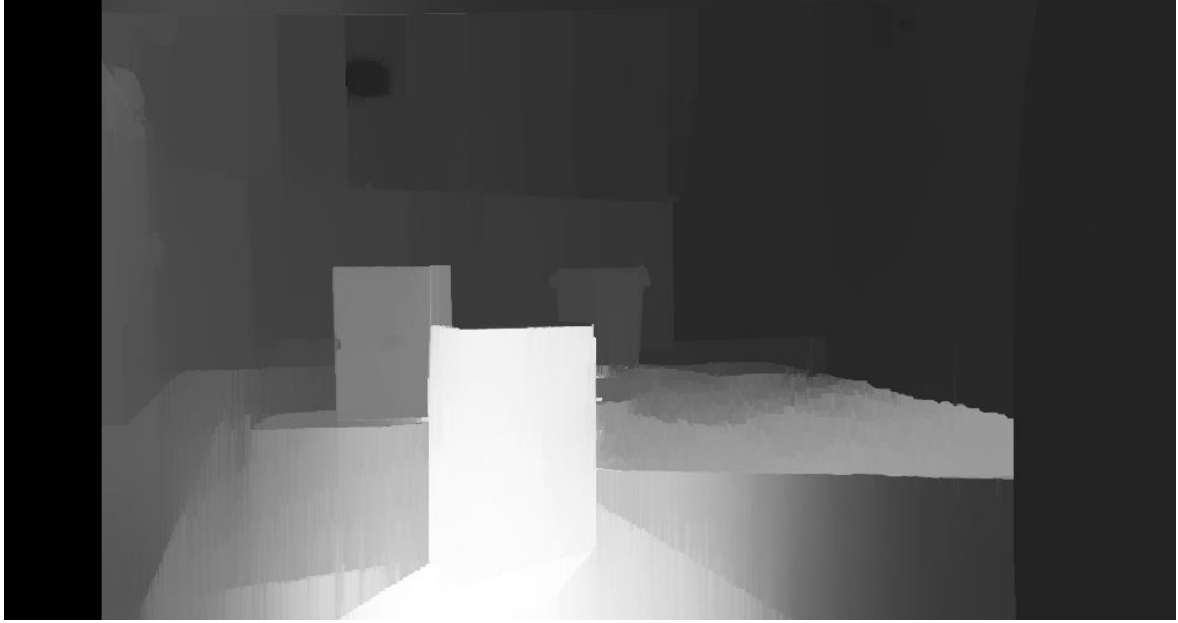
Şekil 4.1. Test alanında engellerle alınan sol ve sağ görüntü örneği

Alınan RGB kamera görüntüleri kalibre edilmiş ve doğrultulmuştur. Elde edilen bu görüntüler Şekil 4.2’de gösterilen gri tonlamalı görüntülere çevrilmiştir.



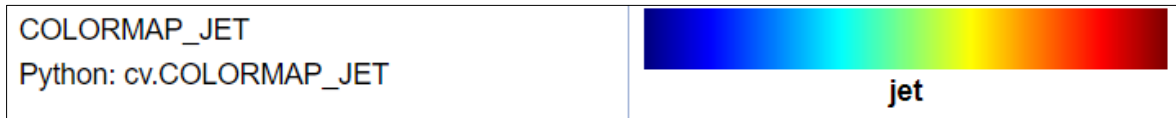
Şekil 4.2. Test alanında engellerle alınan sol ve sağ görüntünün kalibre edilip doğrultulmuş örneği

Sağ ve sol kameraya ait bu doğrultulmuş, kalibre edilmiş gri tonlamalı görüntülerin nihai farklılık değerleri Şekil 4.3 ile gösterildiği gibidir. Yakın olan cisimlerin farklılık değerleri yüksektir. Alınan bu değerler 0 - 255 aralığında normalize edilerek Şekil 4.3'te gösterimi gerçekleştirilmiştir. Otonom araca yakın olan cisimlere ait pikseller beyaz renkte, uzaktaki bölgeler ise siyah renktedir.



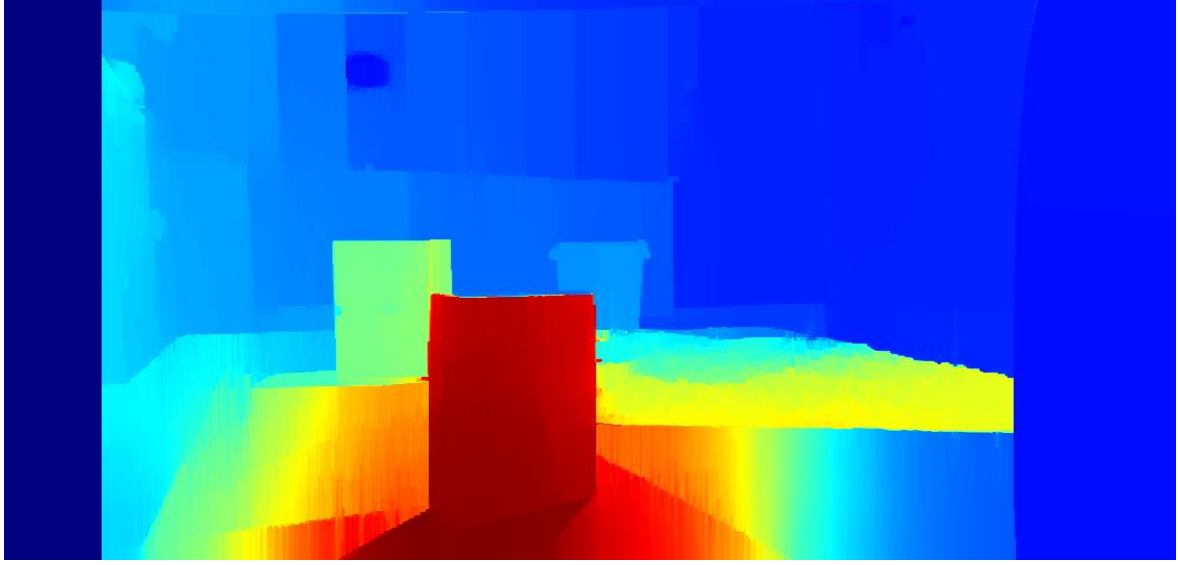
Şekil 4.3. Test alanına ait derinlik haritası gri tonlamalı gösterimi

Elde edilen bu farklılık değerlerinin incelenmesini kolaylaştırmak amacıyla piksellere ait değerlere renk haritalaması uygulanmıştır (Şekil 4.4).



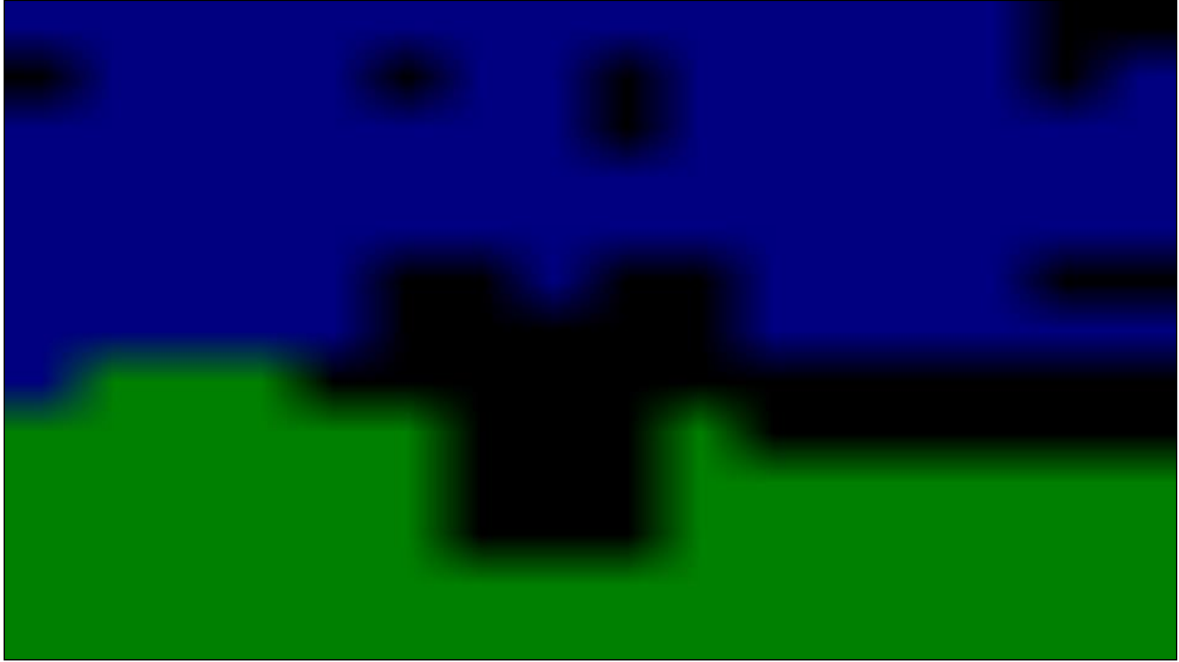
Şekil 4.4. Colormap Jet renk haritası gösterimi

COLORMAP_JET haritalaması uygulanarak yakındaki nesnelere ait piksel değerleri kırmızı, uzaktaki nesnelere ait pikseller ise mavi ile gösterilmektedir (Şekil 4.5).



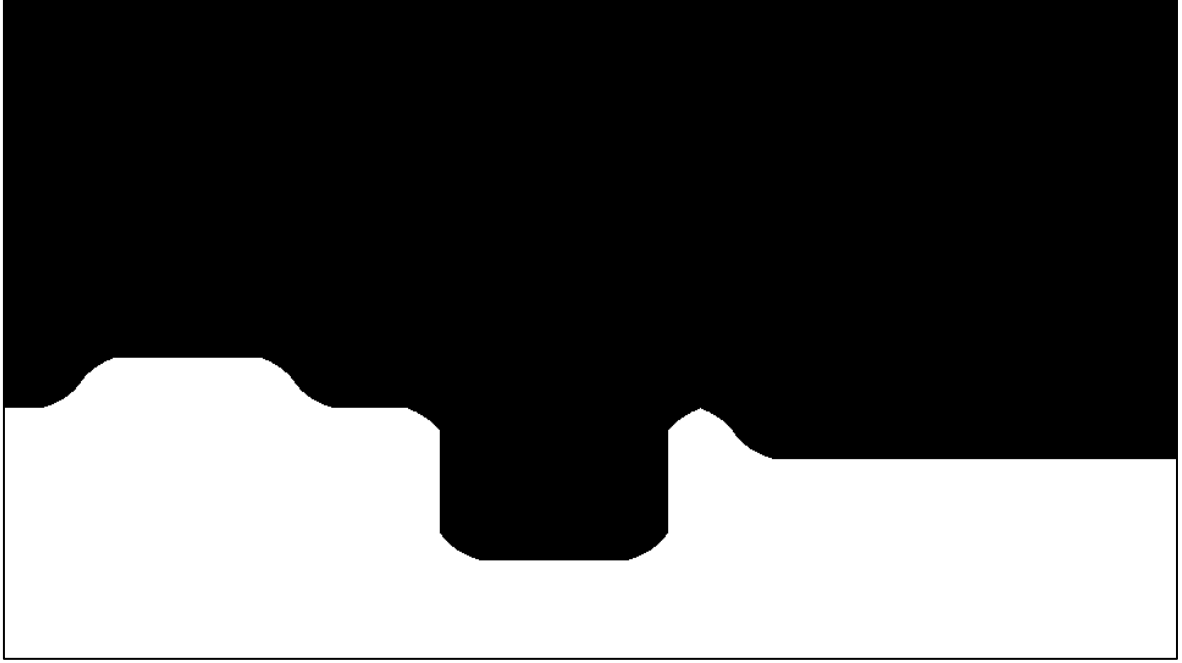
Şekil 4.5. Test alanına ait derinlik haritası Colormap Jet tonlamalı gösterimi

Araç üzerindeki sol kamera aracılığıyla alınan aynı ana ait görüntü SegNet semantik segmentasyonuna tabi tutularak otonom aracın hareket edebileceği zemine ait pikseller tespit edilmektedir. Şekil 4.6’da gösterilen görüntüde yeşil renk ile hareket edilebilir zemin pikselleri, mavi renk ile çevreye ait engeller (duvar, masa, koltuk vb.) ve siyah renk ile tespit edilebilir sınıfların dışında bulunan arka plana ait pikseller gösterilmektedir.



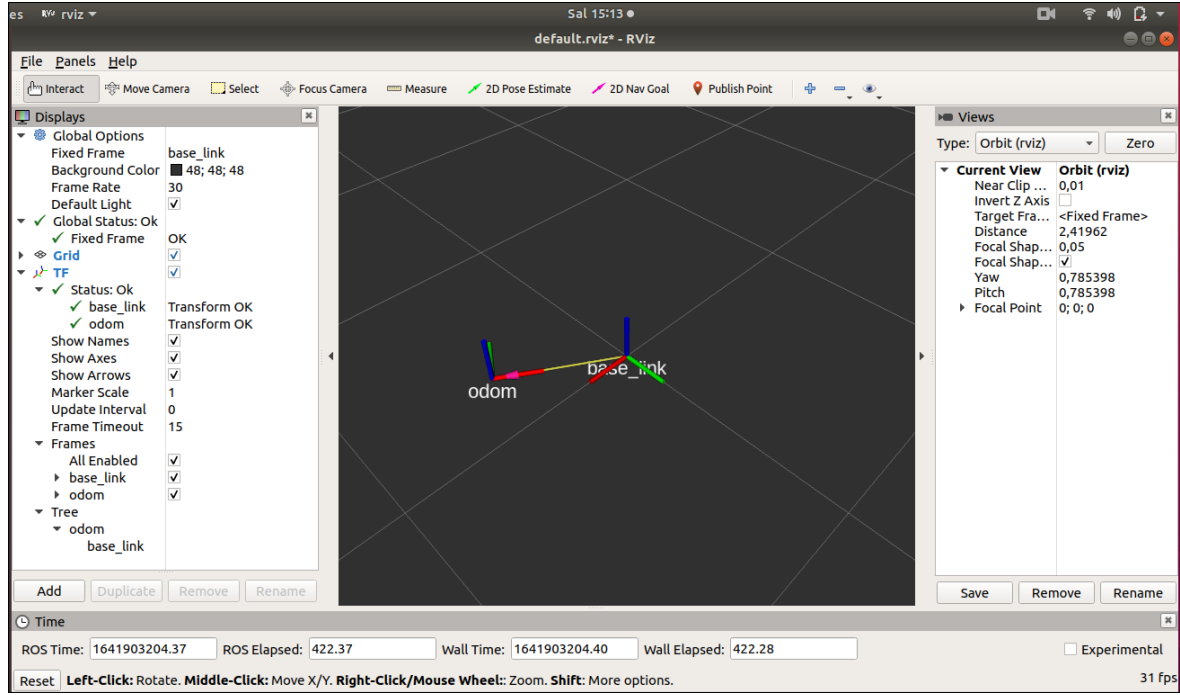
Şekil 4.6. Test alanında semantik segmentasyon çıktısı üretimi

Araç için kritik öneme sahip bu hareket edilebilir zemine ait pikseller, engel ve arkaplan piksellerinden ayırt edilmektedir.



Şekil 4.7. Test alanında semantik segmentasyon çıktısının maskelenmesi sonucu

Bu amaçla HSV maskelemesi uygulanarak 1 – 0’lardan oluşan maske çıktısı görüntüsü Şekil 4.7’deki gibidir. Bu tez kapsamında otonom araç hareketlerinin uzaktan bağlantı yardımıyla bilgisayar ortamında görüntülenmesi planlanmıştır. Araca ait odyometrik veriler kablosuz bağlantı yardımıyla alınarak ROS çatısı altında RVIZ ortamında görüntülenmiştir. Şekil 4.8’de base link olarak başlangıç durumu sabitlenmiş ve bu konuma bağlı olarak gerçekleştirilen yer değiştirme odom isimli dönüşüm ağacı sembolü (TF – Transform Tree) ile gösterilmiştir.

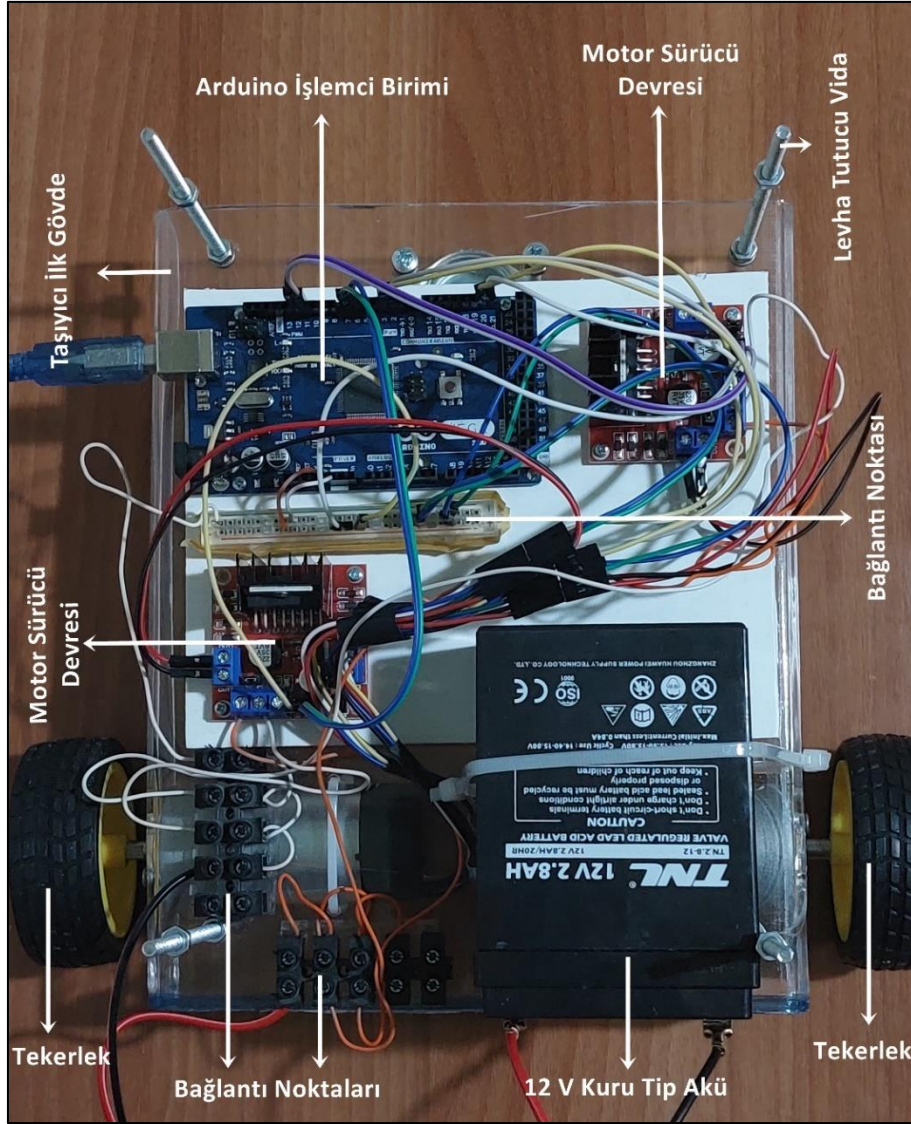


Şekil 4.8. Otonom robota ait odyometrik verilerin RVIZ programı aracılığıyla görselleştirilmesi

RVIZ programı dizüstü bilgisayar üzerinden çalıştırılmaktadır. Birbirinden farklı Raspberry Pi, Jetson Nano ve Laptop işlemci üniteleri bağlı bulundukları kablosuz bağlantı üzerinde IP değerleri kullanılarak ROS master ve ROS hostname atamaları yapılarak çalıştırılmıştır. Bu yapı sayesinde aracın uzaktan kontrolü mümkün olmakta ve RVIZ üzerinde hareket takip edilebilmektedir. Aynı kablosuz ağ sayesinde gerekli görülen çalışmalar SSH bağlantısı kurularak uzaktan gerçekleştirilmektedir.

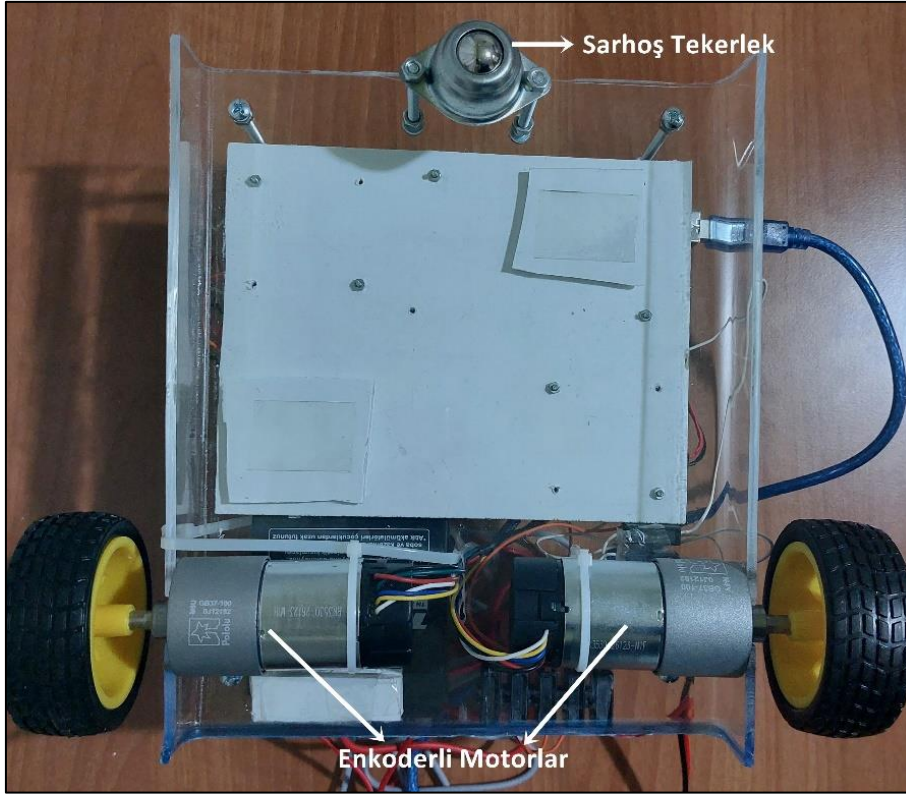
4.2. Araç Yerleşim Planı ve Gerçek Zamanlı Hareket Sonuçları

Araç iki katlı bir yapıda tasarlanmıştır. Otonom araca ait ilk kat için pleksiglas malzeme sıcak büküm ile iki uzun kenarından 3 cm pay bırakılarak -90° bükülmüştür. Bu iki bölümde motor tutucu hazneler açılarak 3 mm çapında motor tutucu vida gözleri delinmiştir. Motor gövdeleri bu iki bölüme vida ile tutunmaktadır. Yine bu kat üzerinde Arduino işlemci birimi, motor sürücü devreleri, güç bağlantı noktaları (5 V ve 12 V) ve 12 V Akü bulunmaktadır (Şekil 4.9).



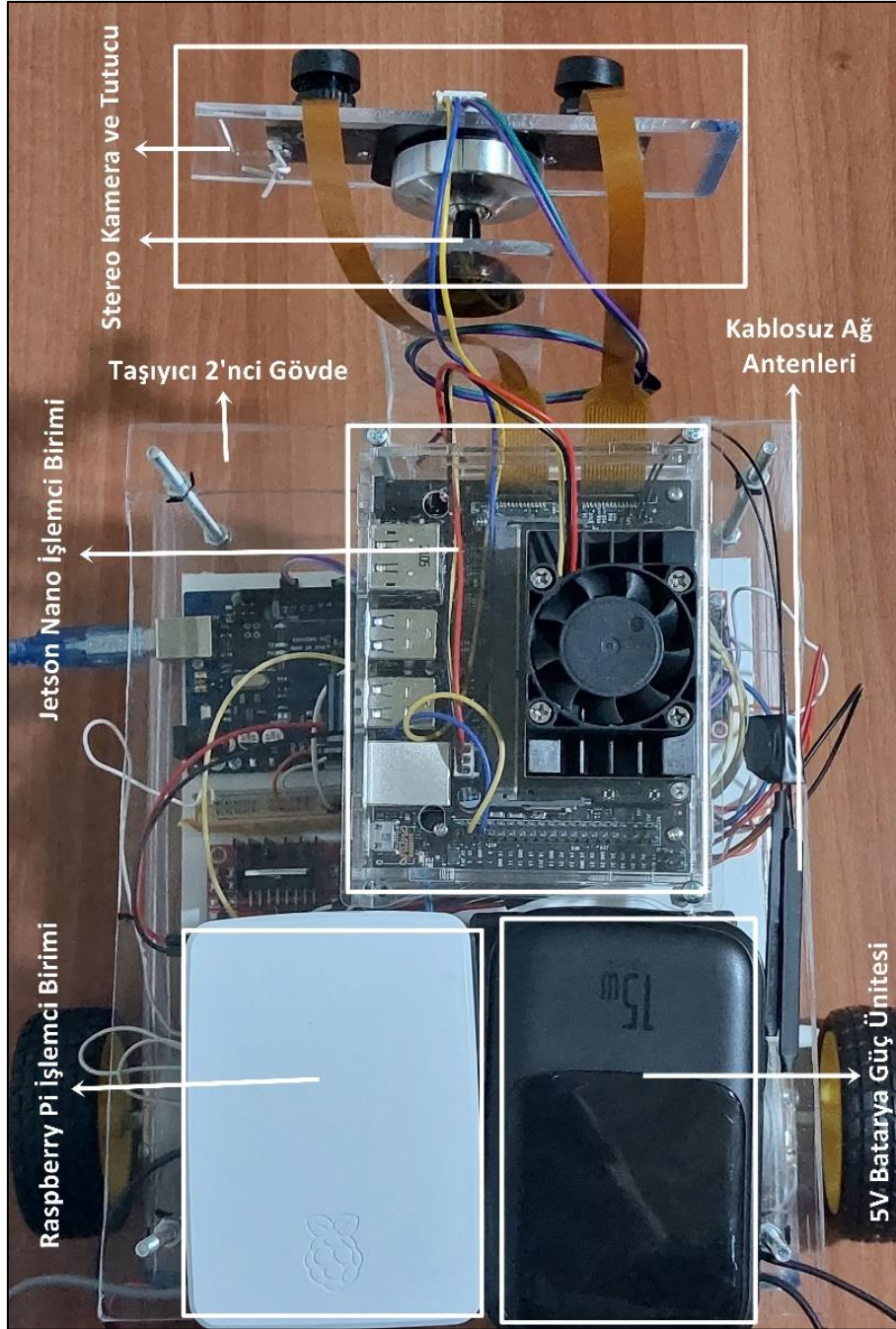
Şekil 4.9. İlk kat araç yerleşim üst görüntüsü

İlk kat levhası alttan görüntüsü ve sarhoş tekerlek montajı Şekil 4.10 ile gösterilmiştir. İkinci katı taşıyacak olan 4 adet vida ise yine bu plaka üzerinde monte edilmiş durumdadır.



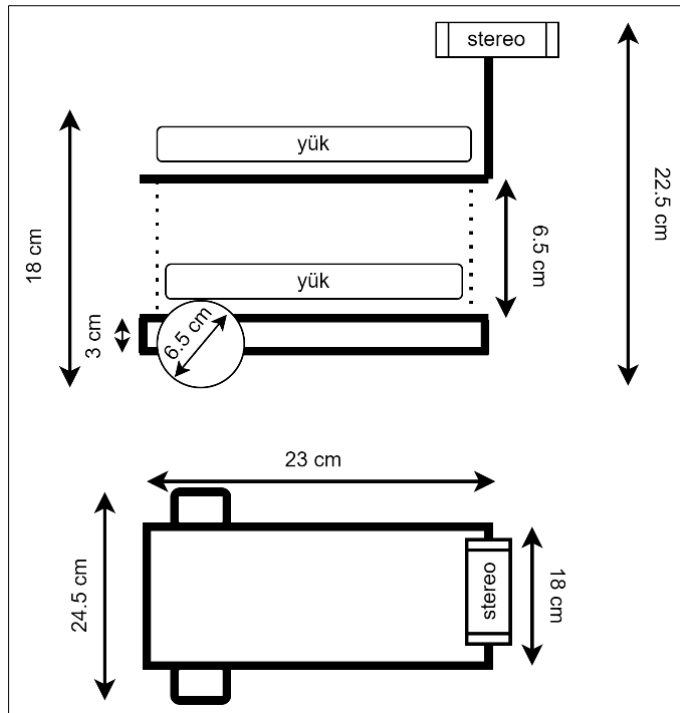
Şekil 4.10. İlk kat araç yerleşim alt görüntüsü

Aracın üstten görüntüsü ve ikinci gövde parçasına ait yerleşim planı Şekil 4.11’deki gibidir. 2’nci kat stereo kamera tutucusunu, Raspberry Pi işlemci birimini, Jetson Nano işlemci birimini ve bunlara enerjisini veren 5 V batarya güç ünitesini ihtiva etmektedir. Araç ön tekerleğinin sarhoş tekerlek olması sebebiyle ağır donanımlar olan güç üniteleri enkoderli motorlar ile aynı eksen üzerine yerleştirilmiştir. Böylelikle araç dengesinde bozulmanın engellenmesi hedeflenmiştir.



Şekil 4.11. İkinci kat araç yerleşim üst görüntüsü

Araç Jetson Nano üzerinden 10 Watt, Raspberry Pi ise maksimum 5.5 Watt güç tüketiminde bulunmaktadır. Hareket halinde araç motorları maksimum 30 Watt güç tüketiminde bulunmaktadır. Toplamda maksimum güç tüketimi 45,5 Watt civarında gerçekleşmektedir. Aracın karar anlarında duruyor olması ve sabit hareket durumlarında motorların çektiği akımların düşük olması gibi sebeplerle ortalama güç tüketimi çok daha düşük olmaktadır. Otonom aracın ölçüleri Şekil 4.12'deki gibidir.



Şekil 4.12. Araç ölçüleri teknik resim görüntüsü

Şekil 11.42’de görülebileceği gibi tüm araç 23 x 18 x 22,5 cm boyutlarında üretilmiştir. Araç, iç mekan ortamında gerçek zamanlı olarak test edilmiştir. Aracın otonom sürüş modunda başarılı bir şekilde engelden kaçındığı ve yapay zeka yardımıyla zemine ait anlamsal bölütleme gerçekleştirerek hareket edilebilir bölge tespiti gerçekleştirdiği görülmüştür. Deneye ait video görüntüsünün çerçeveleri Şekil 4.13’te verilmiştir.



Şekil 4.13. Gerçek zamanlı harekete ait kaydedilen video görüntülerinin çerçeve görüntülerinden örnekleri

Toplamda 65 saniye süren aracın test sürüşü video kaydına alınmıştır. Bu sürüş esnasındaki en önemli 10 zamana ait çerçeve görüntüleri anlık olarak yakalanmıştır. Şekil 4.13'te otonom sürüş modunda hareket eden araca ait bu 10 çerçeve görüntüsü gösterilmiştir. Her çerçeveye ait numara ve süre bilgisi çerçeve görüntülerinin altında verilmiştir. Gerçek zamanlı hareket 0'ıncı saniyede başlamakta ve araç 7'nci saniyede engeli görmesiyle durarak sola dönüş gerçekleştirmektedir. Ardından bir sonraki engele kadar hareket sürdürülmektedir. 27'nci saniyede araç tekrar durmakta ve kaçınma algoritması uyarınca sola dönüş gerçekleştirmektedir. 34'üncü saniyede aracın önünde hala engel bulunmaktadır. Araç tarafından engel başarıyla algılanmış ve tekrar sola dönüş gerçekleştirilmiştir. Ardından 60'nci saniyeye kadar hareket etmiş ve engelin tespitiyle sola dönüş gerçekleştirmiştir. Bu çerçeve görüntülerinden de görülebileceği gibi araç otonom engelden kaçınma görevini başarıyla gerçekleştirmektedir.

5. SONUÇ VE ÖNERİLER

Son dönemde gerçekleşen teknolojik gelişmeler ile işlemci hızlarındaki artış, donanımların boyutlarındaki küçülme ve teknolojik malzemelerin ucuzlaması ile bilgisayarlı görü ve robotik gibi alanlara artık çok daha fazla insan ulaşabilmekte ve dolayısıyla bu alandaki çalışmalar gün geçtikçe artmaktadır. Buna ek olarak son yıllarda yapay zeka alanında kaydedilen ilerlemeler sayesinde de bilgisayarlı görü sistemleri eskisinden çok daha verimli çalışabilmektedir. Bu çalışmada kapalı ortamlarda (iç mekan) çalışabilecek otonom bir kara aracı geliştirilmiştir.

Geliştirilen araç otonom olarak engelden kaçınabilmekte ve hareket edilebilir zemine ait çıkarımda bulunabilmektedir. Gerçek zamanlı olarak çalışabilen bu sistem dinamik ortamlara rahatça uyum sağlayabilmektedir. Geliştirilen gezgin robot çevreye ait engellerin derinlik tespitini sadece stereo görüntüleme sistemi kullanarak yapmaktadır. Araç klavye üzerinden el ile kontrol edilebilmekte ve atanan tuş yardımıyla otonom sürüşe geçebilmektedir. Otonom hareket sırasında stereo kamera aracılığıyla engelin tespitini yapabilmektedir. Ardından daha önce planlanmış kaçınma algoritması çalışmaktadır. Bunun yanı sıra, araç hareket bölgesindeki zeminin hareket edilebilir bir zemin olup olmadığını da SegNet isimli yapay zeka tabanlı semantik segmentasyonu sayesinde belirleyebilmektedir. Böylelikle çevreye ait piksellerde anlamsal bir çıkarım yapabilmektedir.

Semantik segmentasyon sayesinde iç mekanlarda derinlik algılama sistemlerinin ayırımına varamadığı zemin bilgisi yapay zeka sayesinde tespit edilmektedir. Aktif ya da pasif derinlik algılama sistemlerinin ayırt edemediği merdiven boşlukları, çim yüzeyler, ince nesneler ve aracın hareketinin istenmediği farklı kaplamaya sahip bölgeler ayırt edilerek aracın hareketleri kısıtlanabilmektedir.

Geliştirilen sistemde otonom araç çevreye ait bilgileri kablosuz bağlantılarla almaktadır; bu yönüyle de araç üzerinde olmayan farklı kameralar yardımıyla araca hareket sağlanabilir. Nesnelerin hareketli olduğu ortamlarda gerçek zamanlı olarak engelin tespiti gerçekleştirilebilmektedir.

Çalışma Jetson Nano, Raspberry Pi ve Arduino işlemci üniteleri kullanılarak gerçekleştirilmiştir. Algoritmalar gerçekleştirilirken C++, C# ve Python programlama dilleri

kullanılmıştır. Çalışma ROS üzerinde gerçekleştirilmiştir. Robota kablosuz ağ ile uzaktan bağlantı gerçekleştirilebilmesinin yanında ROS üzerinden yayınlanan parametreler ile robotun hareketi ve konumu RVIZ aracılığıyla uzaktan izlenebilmektedir.

Bu çalışmadaki en büyük zorlayıcı durum ana işlemcilerin yetersiz olmalarıdır. Araç görüntü işleme sürecini yaklaşık olarak 1 FPS civarında gerçekleştirebilmektedir. İşlemcinin çalışma performansına göre tüm bu işlem süreleri değişkenlik gösterebilmektedir. Bu sebeple araç hareketi öncesi minimum 2 saniyelik gecikmeler verilerek derinlik çıkarımlarının doğruluğu artırılmıştır. OODA çevrim sürecinin toplamda 3 saniye civarında gerçekleşmesi deneysel açıdan uygun görülmüştür. Özellikle yapay zeka algoritmaları ve görüntü işleme algoritmaları üzerinde koşulan Jetson Nano yerine daha gelişmiş bir işlemci (Jetson TX2, Jetson AGX Xavier vb.) kullanılması önerilmektedir. Böylelikle ROS sisteminin performansı ve elde edilen görüntülerin çerçeve hızları artırılabilir. Buna ek olarak, Jetson Nano güç beslemesi olarak USB-C güç girişi ile beslendiğinde 2 A üzerinde akım çekilmesi durumunda sistemin yüksek akım uyarısı verdiği görülmüştür. Bu sebeple sistemin daha yüksek akım sağlayan bir güç kaynağı ile güç portundan beslenmesi önerilmektedir. Sistemin çalışmasında enkoderli motorların eş olmasına dikkat edilmelidir. Farklı motorlar aynı dönüş tur sayısında farklı enkoder kare dalga sayısı çıktısı verebilmektedir. Bu da RVIZ üzerinde ve gerçek dünyada kaymalara sebebiyet vermektedir.

Sistemin donanımsal dezavantajlarının yanı sıra SGBM algoritmasının ortam şartlarına bağlı olarak değişkenlik göstermesi, çok yakın cisimlerin derinliklerinin tespit edilememesi ve SegNet çıktılarındaki hata oranı sistemin geliştirilmesi gereken diğer özellikleridir.

Sistemi literatürdeki engelden kaçınma sistemlerinden ayıran en önemli farklılık, tüm bu donanımsal kısıtlara rağmen bölütleme ve stereo derinlik algılama sistemlerinin avantajlarını bir araya getirmesidir. Geliştirilen bu sistem sayesinde dinamik ortamlarda sadece görüntüleme sistemleri kullanılarak engelden kaçınma işlemi gerçekleştirilebilmektedir. Sistem üzerinde çalışan yapay zeka modeli farklı veri setleri yardımıyla özelleştirilebilir. Bu sayede sistem dağ yollarını, asfaltı, çimi ya da istenilen desendeki bir zemini ayırt edebilir hale getirilerek istenilen görevlerde kullanılabilir.

Gelecek çalışmalarda otonom aracın engelleri tanınması, ortama ait haritalamanın gerçekleştirilmesi ve yol planlaması (path planning) yapılması planlanmaktadır. Daha

yüksek performanslı işlemciler ile 3 boyutlu olarak görüntü çatma gerçekleştirerek 3 boyutlu sanal haritalama yapılması da hedeflenmektedir. Bunun yanı sıra stereo derinlik tespiti algoritmalarında yapay zeka tabanlı derinlik çıkarımı gerçekleştiren AnyNet [69] ve StereoNet [70] gibi mobil cihazlarda çalışabilecek derinlik tahmin sistemlerinin oluşturulması ve bunların performanslarının artırılması da gelecek çalışmalar arasında planlanmaktadır.

KAYNAKLAR

1. Yıldız, H. (2012). *Endüstri Ürünleri Tasarımı Kapsamında Biyomimetik Tasarımın Yeri ve Metodolojisi*, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 5-9.
2. Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 7068349, 1-2.
3. Massa, J. S., Wallace, A. M., Buller, G. S., Fancey, S. J., & Walker, A. C. (1997). Laser depth measurement based on time-correlated single-photon counting. *Optics Letters*, 22(8), 543-545.
4. Zhao, C., Sun, Q., Zhang, C., Tang, Y., & Qian, F. (2020). Monocular depth estimation based on deep learning: An overview. *Science China Technological Sciences*, 63(9), 1-16.
5. Levinson, J., Askeland, J., Becker, J., Dolson, J., Held, D., Kammel, S., & Thrun, S. (2011). *Towards fully autonomous driving: Systems and algorithms*. In 2011 IEEE Intelligent Vehicles Symposium (IV), Baden, 163-168.
6. Gao, H., Cheng, B., Wang, J., Li, K., Zhao, J., & Li, D. (2018). Object classification using CNN-based fusion of vision and LIDAR in autonomous vehicle environment. *IEEE Transactions on Industrial Informatics*, 14(9), 4224-4231.
7. Milella, A., & Reina, G. (2014). 3D reconstruction and classification of natural environments by an autonomous vehicle using multi-baseline stereo. *Intelligent Service Robotics*, 7(2), 79-92.
8. Kriegman, D. J., Triendl, E., & Binford, T. O. (1989). Stereo vision and navigation in buildings for mobile robots. *IEEE Transactions on Robotics and Automation*, 5(6), 792-803.
9. Onoguchi, K., Takeda, N., & Watanabe, M. (1998). Planar projection stereopsis method for road extraction. *IEICE Transactions on Information and Systems*, 81(9), 1006-1018.
10. Kumano, M., Ohya, A., & Yuta, S. (2000). *Obstacle avoidance of autonomous mobile robot using stereo vision sensor*. In 2nd International Symposium on Robotics and Automation, Kitakyushu, 497-502.
11. Burschkal, D., Lee, S., & Hager, G. (2002). *Stereo-based obstacle avoidance in indoor environments with active sensor re-calibration*. In Proceedings 2002 IEEE International Conference on Robotics and Automation, Washington, Cat. No. 02CH37292, Vol. 2, 2066-2072.
12. Sabe, K., Fukuchi, M., Gutmann, J. S., Ohashi, T., Kawamoto, K., & Yoshigahara, T. (2004). *Obstacle avoidance and path planning for humanoid robots using stereo vision*. In IEEE International Conference on Robotics and Automation (ICRA'04), New Orleans, Vol. 1, 592-597.

13. Chao, C. H., Hsueh, B. Y., Hsiao, M. Y., Tsai, S. H., & Li, T. H. S. (2009). Fuzzy target tracking and obstacle avoidance of mobile robots with a stereo vision system. *International Journal of Fuzzy Systems*, 11(3), 183-191.
14. Uddin Haque, A., & Nejadpak, A. (2017). Obstacle Avoidance Using Stereo Camera. *arXiv Preprint arXiv:1705.04114*.
15. Aytekin, B. (2009). *Bilgisayarlı Görü Temelli Araç Belirleme ve Takibi*, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
16. İyidir, İ. K. (2012). *Visual Obstacle Detection and Avoidance for Indoor Mobile Robots*, Yüksek Lisans Tezi, Işık Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
17. Barry, A. J., Florence, P. R., & Tedrake, R. (2018). High-speed autonomous obstacle avoidance with pushbroom stereo. *Journal of Field Robotics*, 35(1), 52-68.
18. Ruf, B., Monka, S., Kollmann, M., & Grinberg, M. (2018). Real-time on-board obstacle avoidance for UAVs based on embedded stereo vision. *arXiv Preprint arXiv:1807.06271*.
19. Dirik, M. (2020). *Development of Vision-Based Mobile Robot Control and Path Planning Algorithms in Obstructed Environments*, Doktora Tezi, İnönü Üniversitesi Fen Bilimleri Enstitüsü, Malatya.
20. Hiçdurmaz, A. (2020). *Gezgin Robotun Görme Tabanlı ve Gerçek Zamanlı Olarak Engelden Sakınımı*, Yüksek Lisans Tezi, Yalova Üniversitesi Fen Bilimleri Enstitüsü, Yalova.
21. Mohammed, M.T. (2020). *Obstacle Avoidance with Object Detection Using 3-D Stereo Cameras for Indoor Mobile Robots*, Doktora Tezi, Bahçeşehir Üniversitesi Lisansüstü Eğitim Enstitüsü, İstanbul.
22. Gönül, E. (2019). *Vision Based Real Time Obstacle Detection and Human Motion Tracking System for Movement of Indoor Autonomous Mobile Robot*, Yüksek Lisans Tezi, Yeditepe Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.
23. Ulusoy, U., Demirhan, A. (2021). *Stereo Kamera ve Yapay Zeka Tabanlı Derinlik Algılama Sistemi İle İnsan Mesafe Tespiti*. 12'nci Uluslararası Bilimsel Araştırmalar Kongresi, 17-18 Aralık, Ankara, 40-47.
24. İnternet: Piksel. URL: <https://en.wikipedia.org/wiki/Pixel>, Son Erişim Tarihi: 21.01.2022.
25. Szeliski, R. (2010). *Computer vision: algorithms and applications*. Londra: Springer Science & Business Media. 3-19.
26. Jokić, A., Petrović, M., & Miljković, Z. (2022). Semantic segmentation based stereo visual servoing of nonholonomic mobile robot in intelligent manufacturing environment. *Expert Systems with Applications*, 190, 116203.

27. Kirillov, A., He, K., Girshick, R., Rother, C., & Dollár, P. (2019). *Panoptic segmentation*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, 9404-9413.
28. Li, B., Liu, S., Xu, W., & Qiu, W. (2018). *Real-time object detection and semantic segmentation for autonomous driving*. In MIPPR 2017: Automatic Target Recognition and Navigation, International Society for Optics and Photonics, Vol. 10608, 106080.
29. Brazil, G., Yin, X., & Liu, X. (2017). *Illuminating pedestrians via simultaneous detection & segmentation*. In Proceedings of the IEEE International Conference on Computer Vision, Venice, 4950-4959.
30. Tao, X., Zhang, D., Ma, W., Liu, X., & Xu, D. (2018). Automatic metallic surface defect detection and recognition with convolutional neural networks. *Applied Sciences*, 8(9), 1575.
31. Hao, S., Zhou, Y., & Guo, Y. (2020). A brief survey on semantic segmentation with deep learning. *Neurocomputing*, 406, 302-321.
32. Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *J. of Math*, 58(345-363), 5.
33. Turing, A. M. (2009). *Computing machinery and intelligence*. In *Parsing the turing test*. Dordrecht: Springer, 23-65.
34. Tür, Mehmet Rıdâ, Çakmak, F. (2017). *Bilim tarihinde ilk sibernetik ve otomatik kontrol sistemleri*. Genç Akademisyenler Sempozyumu-3, Mardin, 248.
35. Farley, B. W. A. C., & Clark, W. (1954). Simulation of self-organizing systems by digital computer. *Transactions of the IRE Professional Group on Information Theory*, 4(4), 76-84.
36. Widrow, B. (1960). An adaptive adaline neuron using chemical memistors. *Stanford Electronics Laboratories Technical Report*. Stanford. 6-20.
37. İnternet: Aktivasyon fonksiyonları. URL: <https://www.analyticsvidhya.com/blog/2020/01/fundamentals-deep-learning-activation-functions-when-to-use-them/>, Son Erişim Tarihi: 21.01.2022.
38. İnternet: Aktivasyon fonksiyonları kullanım alanları. URL: <https://ayyucekizrak.medium.com/ad%C4%B1m-ad%C4%B1m-google-colab-%C3%BCcretsiz-tpu-kullan%C4%B1m%C4%B1-621dc6e5487d>, Son Erişim Tarihi: 21.12.2021.
39. Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, doi: 10.1109/TNNLS.2021.3084827, 2.
40. Dhillon, A., & Verma, G. K. (2020). Convolutional neural network: a review of models, methodologies and applications to object detection. *Progress in Artificial Intelligence*, 9(2), 85-112.

41. Badrinarayanan, V., Handa, A., & Cipolla, R. (2015). Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. *arXiv Preprint arXiv:1505.07293*.
42. Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12), 2481-2495.
43. Song, S., Lichtenberg, S. P., & Xiao, J. (2015). *Sun rgb-d: A rgb-d scene understanding benchmark suite*. IEEE Conference on Computer Vision and Pattern Recognition, 567-576.
44. Brostow, G. J., Fauqueur, J., & Cipolla, R. (2009). Semantic object classes in video: A high-definition ground truth database. *Pattern Recognition Letters*, 30(2), 88-97.
45. Majernik, J. (2013). 3D Virtual Projection and Utilization of Its Outputs in Education of Human Anatomy. *Multimedia Technology*, 2(1), 1-8.
46. Mikkola, M., Boev, A., & Gotchev, A. (2010). *Relative importance of depth cues on portable autostereoscopic display*. 3rd Workshop on Mobile Video Delivery, 63-68.
47. Zhang Z. (1998). A Flexible New Technique for Camera Calibration, *MSR-TR-98-71 Microsoft Research Technical Report*. Redmond. 10.
48. Kanay, N., & Ayten, U. E. (2016). *Görme Engelliler İçin Engel Mesafe Tespiti Sistemi*, 016 National Conference on Electrical, Electronics and Biomedical Engineering (ELECO), 120-124.
49. Zicari, P. (2013). Efficient and high performance FPGA-based rectification architecture for stereo vision. *Microprocessors and Microsystems*, 37(8), 1144-1154.
50. Scharstein, D., & Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1), 7-42.
51. Özgündüz, E. (2008). *Stereo Görme ile Hareketli Görüntülerde Engellerin Uzaklık ve Boyutlarının Gerçek Zamanlı Bulunması*. Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, 8-25.
52. Seo, Y. H., Yoo, J. S., & Kim, D. W. (2015). A new parallel hardware architecture for high-performance stereo matching calculation. *Integration the VLSI Journal*, 51(1), 81-91.
53. Hirschmuller, H. (2005). *Accurate and efficient stereo processing by semi-global matching and mutual information*. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, 807-814.
54. İnternet: Robot İşletim Sistemi. URL: https://tr.wikipedia.org/wiki/Robot_%C4%B0%C5%9Fletim_Sistemi, Son Erişim Tarihi: 20.01.2022.

55. Çakmak, F., Uslu, E., Balcılar, M., Yavuz, S., & Amasyalı, M. F. (2014). *ROS Uyumlu Robot Platformu Gerçeklenmesi*. Türkiye Otonom Robotlar Konferansı (TORK), Ankara, 1.
56. Koca, Y. B., Gökçe, B., & Aslan, Y. (2020). ROS/Gazebo Ortamında Tank Sürüş Özellikli Mobil Bir Robotun Simülasyonu. *Journal of Materials and Mechatronics: A*, 1(1), 29-41.
57. Fernandez, E., Crespo, L. S., Mahtani, A., & Martinez, A. (2015). *Learning ROS for robotics programming*. Birmingham: Packt Publishing. 32-39.
58. İnternet: Camera Calibration and 3D Reconstruction URL: https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html, Son Erişim Tarihi: 18.12.2021.
59. Zhang, Y. (2014). Stereo Rig Final Report. *Kastner Research Group Report*. San Diego. 5-7.
60. İnternet: Stereo SGBM Algoritması. URL: https://docs.opencv.org/4.x/d2/d85/classcv_1_1StereoSGBM.html, Son Erişim Tarihi: 20.12.2021.
61. Min, D., & Sohn, K. (2008). Cost aggregation and occlusion handling with WLS in stereo matching. *IEEE Transactions on Image Processing*, 17(8), 1431-1442.
62. Silberman, N., Hoiem, D., Kohli, P., & Fergus, R. (2012). *Indoor segmentation and support inference from rgb-d images*. Computer Vision – ECCV 2012, Lecture Notes in Computer Science, vol 7576, 746-760.
63. Seybold, J., Büla, A., Fritz, K. P., Frank, A., Scherjon, C., Burghartz, J., & Zimmermann, A. (2019). Miniaturized optical encoder with micro structured encoder disc. *Applied Sciences*, 9(3), 452.
64. İnternet: Enkoderli DC motor. URL: <https://www.pololu.com/product/2822>, Son Erişim Tarihi: 20.12.2021.
65. İnternet: Motor sürücü. URL: <http://pdf.datasheetcatalog.com/datasheet/stmicroelectronics/1773.pdf>, Son Erişim Tarihi: 20.12.2021.
66. İnternet: Diferansiyel sürüş paketi. URL: http://wiki.ros.org/differential_drive, Son Erişim Tarihi: 20.01.2022.
67. İnternet: İki tekerli araç yol planlama. URL: <http://planning.cs.uiuc.edu/node659.html>, Son Erişim Tarihi: 20.01.2022.
68. Torres, E. O. C., Konduri, S., & Pagilla, P. R. (2014). *Study of wheel slip and traction forces in differential drive robots and slip avoidance control strategy*. In 2014 American Control Conference, Portland, 3231-3236.

69. Wang, Y., Lai, Z., Huang, G., Wang, B. H., Van Der Maaten, L., Campbell, M., & Weinberger, K. Q. (2019). *Anytime stereo image depth estimation on mobile devices*. In 2019 International Conference on Robotics and Automation (ICRA), 5893-5900.
70. Khamis, S., Fanello, S., Rhemann, C., Kowdle, A., Valentin, J., & Izadi, S. (2018). *Stereonet: Guided hierarchical refinement for real-time edge-aware depth prediction*. In Proceedings of the European Conference on Computer Vision (ECCV), 573-590.



GAZİ GELECEKTİR..