T.C.
KONYA TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

DERİN ÖĞRENME TABANLI DİNAMİK İHA KONUMLANDIRMA SİSTEMİ TASARIMI

Muzamil MOHAMMEDELKHATİM

YÜKSEK LİSANS TEZİ

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Haziran-2025
KONYA

TEZ KABUL VE ONAYI

Muzamil MOHAMMEDELKHATİM tarafından hazırlanan "DERİN ÖĞRENME TABANLI DİNAMİK İHA KONUMLANDIRMA SİSTEMİ TASARIMI" adlı tez çalışması …/…/… tarihinde aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri	**İmza**
Başkan Doç. Dr. Akif Durdu (Konya Teknik Üniversitesi)	……..
Danışman Doç. Dr. Ayşe Elif CANBİLEN (Konya Teknik Üniversitesi)	……..
Üye Doç. Dr. HAKKI SOY (Necmettin Erbakan Üniversitesi)	……..

Yukarıdaki sonucu onaylarım.

Prof. Dr. Mevlüt UYAN
Enstitü Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Muzamil MOHAMMEDELKHATİM
Tarih:

ÖZET

YÜKSEK LİSANS TEZİ

DERİN ÖĞRENME TABANLI DİNAMİK İHA KONUMLANDIRMA SİSTEMİ TASARIMI

Muzamil MOHAMMEDELKHATİM

Konya Teknik Üniversitesi
Lisansüstü Eğitim Enstitüsü
Elektrik-Elektronik Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Ayşe Elif CANBİLEN

2025, 54 Sayfa

Jüri
Doç. Dr. Ayşe Elif CANBİLEN
Doç. Dr. Akif DURDU
Doç. Dr. Hakkı SOY

Bu tez çalışması, İnsansız Hava Araçları (İHA) için Takviyeli Öğrenme (Reinforcement Learning, RL) tabanlı bir dinamik konumlandırma sistemi sunarak kablosuz kapsama alanı, enerji verimliliği, adalet ve acil durum müdahale yeteneklerini artırmayı amaçlamaktadır. Derin Q-Öğrenme (Deep Q-learning, DQL) algoritmalarına, özellikle Derin Q-Ağı (Deep Q Network, DQN), Çift DQN (Double Deep Q Network, DDQN) ve Düello DQN, bağlı olarak önerilen sistem, İHA'ların dinamik olarak dağıtılmış yer kullanıcılarına kapsama sağlamak için konumlarını otonom bir şekilde optimize etmesini sağlar. Spesifik olarak, ele alınan sistemde kapsama, enerji tüketimi, adalet ve acil durumlara hızlı yanıtı dengeleyen çok bileşenli bir ödül fonksiyonunu tanımlanmıştır. Simülasyon sonuçları, Düello DQN'nin DQN ve DDQN'den üstün performans gösterdiğini, %94 kapsama, %40 enerji azaltımı ve neredeyse optimum adalet sağlarken sağlam bir acil durum müdahalesi sunduğunu göstermektedir. Önerilen yaklaşım, dinamik ortamlara uyum sağlayarak zorlu koşullarda bağlantı sürekliliği ve afet yönetimi gibi senaryolar için ölçeklenebilir ve verimli bir çözüm sunar.

Anahtar Kelimeler: Derin Öğrenme, İnsansız Hava Aracı (İHA), Takviyeli Öğrenme (RL).

ABSTRACT

MS THESIS

DEEP LEARNİNG BASED DYNAMİC UAV POSİTİONİNG SYSTEM DESİGN

Muzamil MOHAMMEDELKHATİM

Konya Technical University
Institute of Graduate Studies
Department of Electrical-Electronics Engineering

Advisor: Assoc. Prof. Dr. Ayşe Elif CANBİLEN

2025, 54 Pages

Jury
Advisor Assoc. Prof. Dr. Ayşe Elif CANBİLEN
Assoc. Prof. Dr. Akif DURDU
Assoc. Prof. Dr. Hakkı SOY

This thesis presents a Reinforcement Learning (RL)-based dynamic positioning system for Unmanned Aerial Vehicles (UAVs), aiming to improve wireless coverage, energy efficiency, fairness, and emergency response capabilities. Based on Deep Q-Learning (DQL) algorithms, specifically Deep Q-Network (DQN), Double DQN (DDQN) and Dueling DQN, the proposed system enables UAVs to autonomously optimize their positions to provide coverage to dynamically distributed ground users. Specifically, the considered system defines a multi-component reward function that balances coverage, energy consumption, fairness and fast response to emergencies. Simulation results show that Dueling DQN outperforms DQN and DDQN, offering 94% coverage, 40% energy reduction and a robust emergency response while achieving near-optimal fairness. The proposed approach adapts to dynamic environments, providing a scalable and efficient solution for scenarios such as connectivity continuity and disaster management in harsh conditions.

Keywords: Deep Learning, Unmanned Aerial Vehicle (UAV), Reinforcement Learning (RL).

ÖNSÖZ

Tüm çalışma sürecim boyunca bilgi ve tecrübesiyle bana rehberlik eden, her soruma sabırla yanıt vererek yolumu aydınlatan ve desteğini hiçbir zaman esirgemeyen kıymetli danışman hocam Doç. Dr. Ayşe Elif CANBİLEN'e en içten teşekkürlerimi sunarım. Ayrıca bilgi ve emekleriyle akademik gelişimime katkı sağlayan bölümümüzün değerli hocalarına ve her zaman yanımda olan, sevgileriyle güç veren, varlıklarıyla bana cesaret aşılayan canım aileme sonsuz minnettarım.

Muzamil MOHAMMEDELKHATİM
KONYA-2025

İÇİNDEKİLER

SİMGELER VE KISALTMALAR

Simgeler

S	**:** Durumlar kümesi
R	**:** Ödül fonksiyonu
A	**:** Eylemler kümesi
P	**:** Durum geçiş olasılık fonksiyonunu
γ	**:** Azaltma faktörü
$Q(s,a)$	**:** Q-değer fonksiyonu
$V(s)$	**:** Durum değer fonksiyonu
ϵ_{decay}	**:** Epsilon Bozulması
a	**:** Yapılan eylem
r	**:** Alınan anlık ödül
θ	**:** Ana Q-ağı
θ^{-}	**:** Q-değeri hedef ağ
$L(\theta)$	**:** Kayıp fonksiyonu
$\lvert\mathcal{A}\rvert$	**:** Eylem uzayı
$A(s,a;\theta)$	**:** Avantaj fonksiyonu
$x_{İHA}$, $y_{İHA}$, $z_{İHA}$	**:** İHA Koordinasyonları
$\mathcal{W}$	**:** Kullanıcı pozisyonları
$\mathcal{P}$	**:** Kapsama durumu
C	**:** İHA tarafından hizmet verilen kullanıcıların yüzdesi
P	**:** Çalışması için kullanılan enerji
F	**:** Jain Adalet Endeksi
E	**:** Acil durum müdahalelere tepki
$L_{uk}(dB)$	**:** İHA ile kullanıcı arasındaki yol kaybı
ξ_{dB}	**:** Serbest uzaydaki kanal yol kaybı modeli
$P_{uk}^{LoS}(t)$	**:** LoS olasılığı
$P_{uk}^{NLoS}(t)$	**:** NLoS olasılığı
$\theta(t)$	**:** Yükseklik açısı
h_u	**:** İHA yüksekliği

h_k : Kullanıcı yüksekliği

$d_{uk}(t)$: İHA ile kullanıcısı arasındaki üç boyutlu mesafe

$\bar{L}_{uk}(t)$: Ortalama yol kaybı

$C_k(t)$: Kapsama değeri

$P_{cov(i,j)}$: Kapsama olasılığı

K : Kullanıcı sayısı

P_u : Transmisyon gücü

ΔT : Zaman aralığı

T : Süre/tekrarlamalar

v_u : İHA hızı

n_0 : Yol kaybı üssü

f_c : Sistem taşıyıcı frekansı

c : Işık hızı

ρ : Hava yoğunluğu

A : Disk rotoru alanı

χ_{LoS} : LoS zayıflaması

χ_{NLoS} : NLoS zayıflaması

δ , B : Çevresel sabitler

U_{tip} : Uç hızı

v_0 : Ortalama rotor kaynaklı hız

d_0 : Gövde sürükleme oranı

s_0 : Rotor sağlamlığı

P_0 : Kanat profili gücü

P_1 : Türetilmiş güç

$E_{\text{cons}}(t)$: Enerji tüketimi

$T_u(t)$: Uçuş süresi

$E_{\text{res}}(t)$: Artık enerji

E_{max} : Batarya boyutu

$ER_k(t)$: Acil durum müdahale puanı

$Q^*(s,a)$: Optimum eylem-değer fonksiyonu

Kısaltmalar

AoA	**:** Varış Açısı (Angle of Arrival)
3B	**:** Üç boyutlu (three-dimensional, 3D)
3GPP	**:** 3. Nesil Ortaklık Projesi (3rd Generation Partnership Project)
BS	**:** Baz istasyonu (Base Station)
CNN	**:** Konvolüsyonel Sinir Ağları (Convolutional Neural Network)
CONEC	**:** İşbirlikli Ağ Kapsama Alanı Geliştirme Şeması (Collaborative Network Coverage Enhancement Scheme)
DBN	**:** Derin İnanç Ağları (Deep Belief Networks)
DDPG	**:** Derin Deterministik Politika Gradyanı (Deep Deterministic Policy Gradient)
DDQN	**:** Çift Derin Q-Ağı (Double Deep Q-Network)
DL	**:** Derin Öğrenme (Deep Learning)
DNN	**:** Derin sinir ağı (Deep neural network)
DQL	**:** Derin Q-Öğrenme (Deep Q-Learning)
DQN	**:** Derin Q-Ağı (Deep Q-Network)
DRL	**:** Derin Takviyeli Öğrenme (Deep Reinforcement Learning)
DRL	**:** Derin Takviyeli Öğrenme (Deep Reinforcement Learning)
DRL-EC	**:** Kapsama Alanı ve Bağlantı Için DRL Tabanlı Enerji Tasarruflu Kontrol (DRL-Based Energy efficient Control for Coverage and Connectivity)
Düello DQN	**:** Düello Derin Q-Ağı (Dueling Deep Q-Network)
EEFC-TDBA	**:** Yörünge Tasarımı ve Bant Tahsisi Yoluyla Enerji Verimli Adil İletişim (Energy-Efficient Fair Communication Through Trajectory Design and Band Allocation)
GAN	**:** Üretici Karşıt Ağlar (Generative Adversarial Network)
GEO	**:** Jeostatik Yörünge (Geostatıonary Earth Orbit)
GPS	**:** Genellikle küresel konumlama sistemi (Global Positioning System)
HAPS	**:** Yüksek İrtifa Platform İstasyonları (High-Altitude Platform Stations)
İHA	**:** İnsansız Hava Araçları
IoT	**:** Nesnelerin İnterneti (Internet of Things)
LEO	**:** Alçak Dünya Yörüngesi (Low-Earth Orbit)

LoS/NLoS	: Direkt Görüş Hattı/Direkt Görüş Hattı Olmayan (Line-Of-Sight/Non-Line-Of-Sight)
MADDPG	: Çok Ajanlı DDPG (Multi Agent Deep Deterministic Policy Gradient)
MADRL	: Çok Ajanlı DRL (Multi Agent Deep Reinforcement Learning)
MAUC	: Çok Etmenli Derin Takviye Öğrenme Tabanlı Dağıtık İHA-BS Kontrol Yaklaşımı (Multi-Agent Deep Reinforcement Learning-Based Distributed UAV-Bss Control Approach)
MDP	: Markov Karar Süreci (Markov Dicision Proccess)
MEO	: Orta Dünya Yörüngesi (Medium-Earth Orbit)
MILP	: Karma Tamsayılı Doğrusal Programlama (Mixed Integer Linear Programs)
ML	: Makine Öğrenmesi (Machine Learning)
NTN	: Karasal Olmayan Ağlar (Non-Terrestrial Networks)
POMDP	: Kısmen Gözlemlenebilir Markov Karar Süreci (Partially Observable Markov Decision Process)
PPO	: Proksimal Politika Optimizasyonu (Proximal Policy Optimization)
ProxSGD	: Proksimal Stokastik Gradyan İnişi (Proximal Stochastic Gradient Descent)
PSO	: Parçacık Sürü Optimizasyonu (Particle Swarm Optimization)
QoS	: Hizmet Kalitesi (Quality of Service)
RL	: Takviyeli Öğrenme (Reinforcement Learning)
RNN	: Yinelemeli Sinir Ağları (Recurrent Neural Network)
RSSI	: Alınan Sinyal Gücü Göstergesi (Received Signal Strength Indication)
RSU	: Yol Kenarı Üniteleri (Roadside Units)
SA	: Benzetimli Tavlama (Simulated Annealing)
SBG-AC	: Aktör-Eleştirmen ile Durum Tabanlı Oyun (State-Based Game with Actor–Critic)
SUMO	: Kentsel Mobilite Simülasyonu (Simulation Of Urban Mobility)
TD3	: İkiz Gecikmeli DDPG (Twin-Delayed DDPG)
TDoA	: Varış Zaman Farkı (Time Difference of Arrival)
ToA	: Varış Zamanı (Time of Arrival)

U2G : İHA-yer (User to ground)

VANETs : Araçsal Ad Hoc Ağ (Vehicular Ad Hoc Networks)

1.GİRİŞ

Karasal Olmayan Ağlar (Non-Terrestrial Networks, NTN), geleneksel karasal altyapının kısıtlamalarının ötesinde bağlantı ve operasyonel yetenekler sağlayarak iletişim ve algılama teknolojileri alanında devrim yaratmıştır. NTN, her biri farklı yüksekliklerde çalışan ve küresel internet erişimi, afet müdahalesi, çevresel izleme ve askeri gözetleme gibi belirli uygulamalara yönelik olarak tasarlanmış çeşitli platformları kapsamaktadır. NTN'nin başlıca türleri arasında uydu sistemleri, Yüksek İrtifa Platform İstasyonları (High-Altitude Platform Stations, HAPS) ve İnsansız Hava Araçları (İHA) bulunur; her biri kendine özgü özelliklere ve kullanım alanlarına sahiptir.

Uydu tabanlı NTN, yörüngesel yüksekliklerine göre kategorize edilir. Örneğin; Alçak Dünya Yörüngesi (Low-Earth Orbit, LEO) uyduları (160–2,000 km irtifada) düşük gecikme süresi ve yüksek hızlı bağlantı sunarak Starlink gibi geniş bant hizmetleri için idealdir. Orta Dünya Yörüngesi (Medium-Earth Orbit, MEO) uyduları (2,000–35,786 km irtifada) kapsama alanı ve gecikmeyi dengeler, genellikle küresel konumlama sistemi (Global Positioning System, GPS) gibi navigasyon sistemleri için kullanılır. Jeostatik Yörünge (Geostationary Earth Orbit, GEO) uyduları ise (35,786 km irtifada) televizyon yayıncılığı gibi uygulamalar için sabit konumda kapsama alanı sağlar, ancak daha yüksek gecikme süresine sahiptirler. Diğer yandan HAPS, stratosferde yaklaşık 17–22 km yükseklikte çalışan balonları, hava gemilerini ve sabit kanatlı İHA gibi platformları içerir ve afet yönetimi veya kırsal internet erişimi için bölgesel bağlantı sağlar. Düşük irtifalarda (genellikle 1 km'nin altında) çalışan İHA, yüksek derecede hareketli ve çok yönlüdür, yerel görevler için dinamik ağ düğümleri olarak hizmet ederler. Bu platformlar NTN'nin küresel ölçekli iletişimden hassas odaklı operasyonlara kadar geniş bir ihtiyaç yelpazesini karşılamasını sağlar, ancak her tür maliyet, gecikme, kapsama alanı ve dağıtım karmaşıklığı açısından benzersiz dengelemeler sunar.

NTN platformları arasında, gelişmiş sensörler, iletişim modülleri ve yerleşik bilgisayar sistemleri ile donatılmış olan İHA çeviklikleri ve uyum yetenekleri ile öne çıkar. Yere yakın bir şekilde çalışan İHA, hassas tarım için, lojistikte paket teslimatı için, çevre biliminde hava kalitesi izleme için ve savunmada taktik keşif için olmak üzere çeşitli sektörlerde kullanılmaktadır. Esnek röleler olarak hareket eder; ağ kapsama alanını, kara altyapısının bulunmadığı veya zarar gördüğü uzak köyler veya afet bölgeleri gibi hizmet verilmeyen alanlara genişletir. Havada durabilme, karmaşık ortamlarda gezinme ve gerçek

zamanlı olarak pozisyonları ayarlama yetenekleri, onları arama-kurtarma görevlerini koordine etme veya trafik yönetimi için canlı yayın sağlama gibi anında tepki gerektiren uygulamalar için vazgeçilmez kılar. Uydu ve HAPS'a entegre olarak, NTN dayanıklılığını ve erişimini artırır, küresel kapsama alanını yerelleştirilmiş hassasiyetle birleştiren hibrit ağlar oluşturur.

İHA'nın hızlı konuşlandırma yetenekleri; doğal afetler sırasında geçici iletişim bağlantıları kurmak veya altyapı hasarını gerçek zamanlı olarak izlemek gibi acil ihtiyaçları karşılamak için dakikalar içinde fırlatılmalarına olanak tanır. Statik kara baz istasyonlarının aksine, İHA dinamik hareket kabiliyeti sunarak, mobil kullanıcılar için sinyal gücünü optimize etmek veya deniz gözetiminde gemiler gibi hareketli nesneleri takip etmek vb. ağ taleplerine göre yeniden konumlandırılmalarını sağlar. Maliyet etkinliği, İHA'nın uydulardan veya kapsamlı yer altyapısından daha düşük yatırım gerektirmesi nedeniyle bir başka önemli avantajdır, bu da onları küçük ölçekli operatörler için erişilebilir kılar. Alçak irtifalarda çalışan İHA, direkt görüş hattı (Line of sight, LoS) iletişimi sağlar, bu da sinyal güvenilirliğini artırır, gecikmeyi azaltır ve video akışı veya otonom araç koordinasyonu gibi yüksek bant genişliğine sahip uygulamaları destekler. Ayrıca, İHA koordineli sürüler halinde çalışabilir, etkilerini daha büyük alanları kapsayacak şekilde ölçeklendirebilir veya yedeklilik sağlayabilir, bu da büyük ölçekli çevresel haritalama veya kamu etkinliklerinde kalabalık izleme gibi uygulamalarda İHA'yı kullanılabilir kılar.

İHA tabanlı NTN uygulamalarında hassas ve uyarlanabilir konumlandırma ihtiyacının kritik önemi, GPS tabanlı navigasyon veya önceden tanımlanmış uçuş rotaları gibi geleneksel yöntemlerin çevresel değişkenlik veya karmaşık görev gereksinimleriyle başa çıkmada zorlandığını ortaya koymuştur. Bu tür zorlu ortamlar İHA'nın dinamik olarak hareket ettirilmesini zorlaştırmakta ve ayrıca kapsama süreci veya afet müdahale süreci sırasında tüketilen enerji miktarını artırmaktadır. İHA'nın karar verme konusunda tamamen özgür olmayacağı, bunun yerine başlangıçta net olan kararlara tabi olacağı gerçeğe daha yakındır. Bu durum kapsama alanını azaltır ve ayrıca tüm kullanıcılar için adaletin sağlanmasını garanti etmez. Etkin bir İHA konumlandırması, ağ kapsamasını maksimize ederken enerji tüketimini azaltmak, adil kullanıcı hizmeti sağlamak ve acil durumlara hızlı yanıt verebilmek açısından büyük önem taşımaktadır.

Bu tez çalışması, yukarıda bahsedilen geleneksel yaklaşımların eksikliklerini aşmayı ve İHA'nın NTN içindeki güvenilirliğini, verimliliğini ve ölçeklenebilirliğini artırmayı amaçlamaktadır. İHA'nın çevresel etkileşimler yoluyla otonom olarak optimal konumlandırma stratejilerini öğrenmelerini sağlayan Takviyeli Öğrenme (Reinforcement Learning, RL) tekniğini kullanarak, İHA'nın dinamik olarak konumlarını ayarlaması, sinyal kapsama alanını maksimize etmesi, enerji tasarrufu yapması ve öngörülemeyen koşullarda bile istikrarı koruyabilmesine yönelik optimizasyon çalışmaları yürütülmüştür. Elde edilen sonuçlar, kapsama kapasitesinde artış ve İHA çalışması sırasında tüketilen enerji miktarında azalma olduğunu göstermektedir. Önerilen RL tabanlı çözüm; ağdaki tüm kullanıcılar için adaletin sağlanması ve ağ çalışması sırasında acil durumlara ve felaketlere müdahale hızının iyileştirilmesini sağlamıştır.

Tezin ikinci bölümünde, mevcut literatürün detaylı bir incelemesini sunan, mevcut zorlukları tanımlayan ve problemin önemini vurgulayan kaynak araştırması sunulmaktadır. Üçüncü bölüm materyal ve yöntemlere odaklanmakta; sistem tasarımını, kullanılan algoritmaları ve teknikleri açıklamaktadır. Dördüncü bölümde, bulguların analiz edildiği araştırma sonuçları ve tartışma yer almaktadır. Son olarak, beşinci bölümde, çalışmanın ana sonuçları özetlenmekte ve gelecekteki çalışmalar için potansiyel çalışma doğrultuları sunulmaktadır.

2.KAYNAK ARAŞTIRMASI

Bu bölümde; özellikle İHA ağlarında konumlandırma, kapsama alanı optimizasyonu, enerji verimliliği ve adaletin artırılması konularında literatürdeki mevcut çalışmalar ele alınmıştır.

Liu ve ark. (2018), adil kapsama alanı ve enerji verimliliği için İHA'yı kontrol etmek üzere Derin Deterministik Politika Gradyanı (Deep Deterministic Policy Gradient, DDPG) kullanan derin RL (DRL) tabanlı bir yöntem olan Kapsama Alanı ve Bağlantı İçin DRL Tabanlı Enerji Tasarruflu Kontrol (DRL-Based Energy Efficient Control for Coverage and Connectivity, DRL-EC) yöntemini tanıtmıştır. Yaklaşım, kapsama alanı kazanımlarını, Jain adalet endeksini ve enerji maliyetlerini birleştiren yeni bir ödül fonksiyonunu maksimize ederek, temel yöntemlere göre %28 daha yüksek kapsama alanı ve %30 daha düşük enerji tüketimi elde etmektedir. DRL'nin dinamik İHA kontrolündeki etkinliğini ortaya koymakta, merkezi olmayan ölçeklenebilirlik konusundaki zorlukları vurgulamaktadır.

Liu ve ark. (2019) daha sonra, enerji verimliliği ve coğrafi adaleti sağlarken uzun vadeli iletişim kapsamını en üst düzeye çıkarmayı amaçlayan mobil baz istasyonları olarak hizmet veren birden fazla İHA'nın navigasyonunu optimize etmek için merkezi olmayan bir DRL çerçevesi önermiştir. Merkezi yaklaşımların aksine, geliştirdikleri yöntem her bir İHA'nın yerel gözlemleri kullanarak kontrol politikalarını bağımsız olarak öğrenmesini ve kapsama alanı, adalet ve enerji tüketimini dengeleyen bir ödül fonksiyonuna sahip olmasını sağlamaktadır. Yazarlar, dağıtık yürütme için aktör-kritik ağları içeren bir Kısmen Gözlemlenebilir Markov Karar Süreci (Partially Observable Markov Decision Process, POMDP) modeli tasarlamaktadır. Kapsamlı simülasyonlar aracılığıyla, optimum hiperparametreleri (örneğin, katman başına 160 nöron ve 0.83 azaltma faktörü) belirlermiş; DRL-EC ve açgözlü algoritmalara göre üstün performans göstererek daha yüksek enerji verimliliği ve kapsam adaleti elde eden bir çözüm ortaya koymuşlardır. Temel yenilikler arasında; tamamen dağıtık bir mimari ve gerçek zamanlı İHA bağlantısına ve enerji kısıtlamalarına uyum sağlayarak dinamik ortamlarda otonom çoklu İHA sistemlerini geliştiren dinamik bir ödül mekanizması bulunmaktadır.

Hoseini ve ark. (2021), Nesnelerin interneti (Internet of things, IoT) veri toplama senaryolarında enerji verimliliği ve hizmet önceliği farkındalığı için İHA baz istasyonu yörüngelerini optimize etmek üzere bir Çift Q-Öğrenme yaklaşımı önermektedir. Enerji

tasarrufu ve önceliğe dayalı hizmeti ayrı ayrı ele alan önceki çalışmalardan farklı olarak yazarlar, her iki hedefi de İHA baz istasyonunun, düğümlerin kalan enerji seviyelerine (önceliklerine) ve uçuş enerji maliyetlerine göre yolunu dinamik olarak ayarladığı bir RL çerçevesine entegre etmektedir. Sistem, düğümlerin hizmet önceliklerini modeller ve ağırlıklandırılmış ödüllerle yüksek öncelikli düğümler için gecikmeleri cezalandırır. Simülasyonlar, değiş tokuş optimizasyonu için dengeli bir gelir fonksiyonu ve Q-Öğrenme'nin dinamik öncelik değişikliklerine uyarlanabilirliğinin ampirik olarak doğrulanmasını ortaya koymakta; ölçeklenebilirlik sınırlamaları, daha büyük ağlar için DRL üzerine gelecekteki çalışmaları motive etmektedir.

Bouhamed ve ark. (2020), sürekli eylem alanlarına sahip üç boyutlu (3B) kentsel ortamlarda otonom İHA navigasyonu için Derin Deterministik Politika Gradyanı (deep deterministic policy gradient, DDPG) tabanlı bir DRL çerçevesi önermiştir. Çalışma, hesaplama açısından karmaşık olan veya gerçek zamanlı uyarlanabilirlikten yoksun olan Karma Tamsayılı Doğrusal Programlama (Mixed İnteger Linear Programs, MILP) ve evrimsel algoritmalar gibi geleneksel yol planlama yöntemlerinin sınırlamalarını ele almıştır. Yazarlar, engellerle çarpışmaları cezalandırırken İHA'yı statik veya dinamik hedeflere yönlendirmek için özelleştirilmiş bir ödül fonksiyonu tasarlamıştır. İHA'yı önce engelsiz ortamlarda eğitmek, ardından da engel açısından zengin senaryolara adapte ederek öğrenme verimliliğini artırmak için bir transfer öğrenme yaklaşımı kullanılmıştır. Simülasyon sonuçları, test edilen iki ortamda %84 ve %82'lik görev tamamlama oranları elde ederek İHA'nın otonom olarak gezinme ve engellerden kaçınma becerisini göstermiştir. Çalışma, akıllı şehirlerdeki İHA uygulamaları için ölçeklenebilir bir çözüm sunan DDPG'nin sürekli kontrol ve gerçek zamanlı karar verme için avantajlarını vurgulamıştır.

Ding ve ark. (2020), İHA destekli ağlarda enerji verimli ve adil iletişim için 3B İHA yörüngesini ve frekans bandı tahsisini optimize etmek için Yörünge Tasarımı ve Bant Tahsisi Yoluyla Enerji Verimli Adil İletişim (Energy-Efficient Fair Communication Through Trajectory Design and Band Allocation, EEFC-TDBA) adlı DRL tabanlı bir yaklaşım önermektedir. Çalışma, problemi bir Markov Karar Süreci (Markov Dicision Proccess, MDP) olarak formüle ederek ve bir DDPG algoritması kullanarak, genellikle 3B yörünge optimizasyonunu göz ardı eden veya konveks yaklaşımlara dayanan mevcut yöntemlerin sınırlamalarını ele almaktadır. Simülasyonlar, EEFC-TDBA'nın toplam verim, adalet ve

enerji verimliliğinde temel yöntemlerden daha iyi performans gösterdiğini, kapsama ve kaynak tahsisinde önemli iyileştirmeler sağladığını göstermektedir. Bu çalışma, özellikle kullanıcı hareketliliği ve karmaşık enerji kısıtlamaları olan senaryolarda dinamik İHA ağı optimizasyonu için DRL'nin potansiyelini vurgulamaktadır.

Lee ve ark. (2020), mobil ilk müdahale ekipleri arasında bağlantıyı sürdürmenin kritik olduğu afet yardımı senaryolarında İHA konumlandırmasını optimize etmek için DRL tabanlı bir çerçeve olan DroneDR'yi önermektedir. Çalışma, problemi bir MDP olarak formüle ederek ve İHA ağ bağlantısını sağlarken düğümden düğüme bağlantıları en üst düzeye çıkarmak için Proksimal Politika Optimizasyonu (Proximal Policy Optimization, PPO) kullanarak dinamik İHA röle ağlarının zorluklarını ele almaktadır. Küçük ve büyük ölçekli afet ortamlarında yapılan değerlendirmeler, DroneDR'nin açgözlü sezgisel yöntemlerden (örneğin Steiner ağacı) ve rastgele yol noktası modellerinden daha iyi performans gösterdiğini, 2 kata kadar daha fazla bağlantı, %30 daha yüksek kapsama alanı ve optimuma yakın adalet (Jain endeksi> 0.9) sağladığını göstermektedir. Çalışma, DRL'nin dinamik, gerçek dünya senaryolarında uyarlanabilir İHA ağları için potansiyelini vurgulamakta ve enerji kısıtlamaları, transfer öğrenimi ve donanım doğrulaması dahil olmak üzere gelecekteki yönleri göstermektedir. Bu yaklaşım, acil durum iletişim sistemlerinde teorik optimizasyon ve pratik dağıtım arasındaki boşluğu doldurmaktadır.

Oliveira ve ark. (2021), aşırı yüklü hücresel ağlarda kablosuz bağlantıyı geliştirmek amacıyla İHA'nın dinamik konumlandırmasını optimize etmek için makine öğrenimi (Machine Learning, ML) algoritmalarının kullanımını araştırmıştır. Çalışma, trafiği az kullanılan alanlara yeniden dağıtan bir İHA konumlandırma algoritması önererek yer baz istasyonları arasında eşit olmayan trafik dağılımı sorununu ele almıştır. Yazarlar, kullanıcı hareketliliğini tahmin etmek için çeşitli makine öğrenimi modellerini test etmiş ve Rastgele Orman algoritmasıyla daha uzun eğitim sürelerine rağmen en yüksek doğruluğu elde etmiştir. İHA konumlandırma algoritması aşırı yüklü alanları tespit etmiş, en uygun İHA yerleşimlerini hesaplamış ve trafiği yeniden yönlendirmek için İHA köprüleri inşa etmiştir. Gerçek dünya verileri kullanılarak yapılan performans değerlendirmeleri, algoritmanın özellikle daha az kısıtlayıcı ağ koşullarında bağlantısı kesilen kullanıcı sayısını önemli ölçüde azalttığını göstermiştir. Çalışma, algoritmanın etkinliğini belirlemede tahmin doğruluğunun

ve baz istasyonu bant genişliği ve İHA menzili gibi ağ parametrelerinin önemini vurgulamıştır.

Qin ve ark. (2021), sınırlı İHA baz istasyonlarının bir hedef bölgedeki yer cihazlarına dinamik olarak hizmet vermesi gereken İHA destekli iletişim sistemlerinde kullanıcı düzeyinde adaleti sağlama zorluğunu ele almaktadır. Verim veya enerji verimliliğine odaklanan önceki çalışmalardan farklı olarak yazarlar, verim ve adaleti dengelerken İHA yörüngelerini optimize etmek için Çok Ajanlı DRL (Multi agent DRL, MADRL) yaklaşımı önermektedir.

Lim ve ark. (2021), çarpışmadan kaçınma ve Hizmet Kalitesi (Quality of Service, QoS) gereksinimlerini sağlarken iletişim kapsamını en üst düzeye çıkarmak için birden fazla İHA baz istasyonunun 3B yerleşimini optimize etmek için Benzetimli Tavlama (Simulated Annealing, SA) tabanlı bir algoritma önermektedir. Yazarlar, İHA arası minimum mesafeleri uygulamak için ceza fonksiyonu içeren bir optimizasyon problemi formüle ederek çoklu İHA senaryolarını ele almaktadır. SA algoritması, olasılıksal komşu çözümlerinden ve Boltzmann kabul kriterinden yararlanarak kapsama alanını ve verimi en üst düzeye çıkarmak için İHA konumlarını dinamik olarak ayarlamaktadır. 500×500 m²'lik bir alanda yapılan simülasyonlar, önerilen yöntemin statik İHA baz istasyonu yerleşiminden daha iyi performans gösterdiğini, irtifa ve yatay koordinatları uyarlamalı olarak optimize ederek daha yüksek QoS uyumluluğu ve verim elde ettiğini göstermektedir.

Nemer ve ark. (2022), adil kapsama alanı ve enerji verimliliği için İHA ağlarını optimize etmek üzere hibrit bir oyun teorisi ve DRL yaklaşımı olan Aktör-Eleştirmen ile Durum Tabanlı Oyun (State-Based Game with Actor–Critic, SBG-AC) önermektedir. SBG-AC, İHA etkileşimlerini durum tabanlı potansiyel bir oyun olarak modelleyerek ve bir aktör-eleştiri algoritmasını entegre ederek, enerji kullanımını en aza indirirken kapsamı (8 İHA için %84,6) ve adaleti (%93,4) en üst düzeye çıkarmak için 3B İHA hareketlerini dinamik olarak ayarlar. Simülasyonlar, DRL-EC'den daha iyi performans gösterdiğini ve dinamik ortamlar için ölçeklenebilir bir çözüm sunduğunu göstermektedir, ancak DNN eğitiminden kaynaklanan hesaplama karmaşıklığı bir sınırlama olmaya devam etmektedir.

Islam ve ark. (2022), sınırlı Yol Kenarı Üniteleri (Road Side Units, RSU) konuşlandırması olan alanlarda ağ kapsamını geliştirmek için İHA'yı dinamik olarak konumlandıran bir Araçsal Ad Hoc Ağ (Vehicular Ad Hoc Network, VANET) çerçevesi olan

İşbirlikli Ağ Kapsama Alanı Geliştirme Şemasını (Collaborative Network Coverage Enhancement Scheme, CONEC) önermektedir. Parçacık Sürü Optimizasyonundan (Particle Swarm Optimization, PSO) yararlanan CONEC, bağlanabilirliği en üst düzeye çıkarmak için araç yoğunluğunu, yönelim yönünü ve geçmiş kapsama verilerini dikkate alarak İHA yerleşimini optimize eder. Çerçeve, İHA planlaması ve konum güncellemeleri için algoritmalar içerir ve minimum İHA ile verimli kapsama sağlar. Gerçek dünya, Daegu şehir trafiği, verilerini kullanarak Kentsel Mobilite Simülasyonu (Simulation of Urban Mobility, SUMO) ve Ağ Simülatörü-3'te (Network Simulator 3, ns-3) yapılan simülasyonlar, CONEC'in paket teslim oranını %39,21 artırarak, uçtan uca gecikmeyi %64,42 azaltarak ve verimi optimize ederek temel yöntemlere (sabit / gezinen İHA ve İHA'sız) göre üstünlüğünü göstermektedir. Temel sınırlamalar arasında sabit İHA irtifası ve merkezi kontrole bağımlılık yer almaktadır. Çalışma, İHA'nın özellikle seyrek veya dinamik kentsel ortamlarda araç ağları için uygun maliyetli, çevik röleler olarak potansiyelini vurgulamaktadır.

Liu ve ark. (2022), adil kapsama ve enerji verimliliğine odaklanarak hücresel ağlarda İHA yerleştirme problemini ele almaktadır. Yazarlar, enerji tüketimini en aza indirirken adil kapsama alanını en üst düzeye çıkarmak için ana taşıyıcı kısıtlarını da içeren kısıtlı bir optimizasyon çerçevesi önermektedir. İHA konumlarını yinelemeli olarak optimize etmek için yeni bir Proksimal Stokastik Gradyan İnişi (Proximal Stochastic Gradient Descent, ProxSGD) tabanlı alternatif algoritma tanıtılmış, QoS tahsisinde hesaplama verimliliği ve adalet sağlanmıştır. Algoritma geleneksel yöntemlerden daha iyi performans göstererek daha hızlı yakınsama ve gelişmiş kapsama oranları ortaya koymaktadır.

Bu tez çalışmasında; kapsama alanını, enerji verimliliğini, adaleti ve acil durum müdahale yeteneklerini aynı anda optimize eden, RL tabanlı bir İHA dinamik konumlandırma sistemi geliştirilmektedir. Statik optimizasyon veya tek amaçlı RL kullanan mevcut yaklaşımların aksine, önerilen çerçeve gelişmiş bir Düello DQN (Dueling DQN) mimarisi içinde çok bileşenli bir ödül fonksiyonunu entegre ederek yakınsama hızı (%35 daha hızlı) ve operasyonel kararlılık açısından standart DQN ve Çift DQN (Double DQN) yöntemlerine göre üstün performans göstermektedir. Sistemin gerçek zamanlı acil durum müdahale adaptasyonunu ve adil hizmet dağıtımı için Jain adalet endeksini yenilikçi bir şekilde dahil etmesi ve bu esnada enerji verimliliğini koruması, bu tez çalışmasını benzer RL tabanlı İHA navigasyon çalışmalarından ayırmaktadır. Çizelge 2.1'de bu tez çalışması ile

literatürdeki benzer çalışmalar yöntem, amaç, katkı ve kısıtlamalar açısından karşılaştırılmaktadır.

Çizelge 2.1. Yapılan Çalışmaların Genel Kapsamı

Çalışma	Yöntem/Algoritma	Temel Amaçlar	Ana Katkılar	Sınırlamalar
Nemer ve ark. (2022)	SBG-AC (Oyun Teorisi + DRL, Aktör-Kritik)	Adil kapsama ve enerji verimliliği	3B dinamik İHA hareketi, %84,6 kapsama, %93,4 adalet	Yüksek hesaplama karmaşıklığı
Liu ve ark. (2018)	DRL-EC (DDPG)	Adil kapsama, enerji verimliliği	Kapsama, adalet ve enerji içeren ödül fonksiyonu	Merkezi olmayan yapılarda ölçeklenebilirlik sorunu
Islam ve ark. (2022)	CONEC (PSO tabanlı)	Minimum İHA ile VANET kapsama	Gerçek trafik verileri, gecikmede %64 azalma, teslimde %39 artış	Sabit irtifa, merkezi kontrol bağımlılığı
Qin ve ark. (2021)	MAUC (MADRL)	Dinamik ortamlarda adalet sağlama	Orantılı adalet, dağıtılmış kontrol	—
Liu ve ark. (2019)	Merkezi olmayan DRL (POMDP)	Kapsama, adalet ve enerji dengesi	Yerel gözleme dayalı kontrol, aktör-kritik ağlar	—
Lim ve ark. (2021)	SA tabanlı optimizasyon	3B İHA yerleşimi, çarpışmadan kaçınma	%95 QoS uyumu, uzamsal kısıtlar dikkate alınmış	—
Hoseini ve ark. (2021)	Çift Q-Öğrenme	Öncelik bilincine sahip enerji verimliliği	Yüksek öncelikli düğümler için gecikme azaltma	Büyük ağlara ölçeklenebilirlik sınırlı
Liu ve ark. (2022)	ProxSGD	Adil ve enerji verimli İHA yerleşimi	Kısıtlı optimizasyon, hızlı yakınsama	—

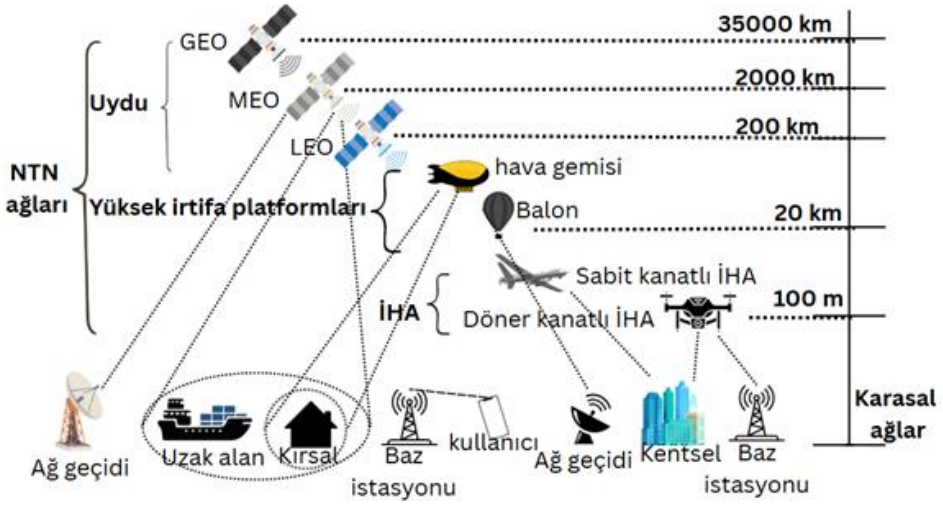
Ding ve ark. (2020)	EEFC-TDBA (DDPG)	3B İHA yörüngesi ve bant tahsisi	Yeni enerji modeli, adalet yardımcı fonksiyonu	—
Lee ve ark. (2020)	DroneDR (PPO)	Afetlerde İHA konumlandırması	2 kat bağlantı artışı, kapsamada %30 artış, Jain> 0.9	Gelecekte aktarım öğrenimi planlanıyor
Oliveira ve ark. (2021)	ML (Random Forest, GB)	Aşırı yüklenmiş trafik için İHA konumlandırma	Trafik tahmini ile İHA yönlendirme	Doğruluk ve ağ parametrelerine bağlılık
Bouhamed ve ark. (2020)	DDPG + Aktarım Öğrenimi	Kentsel alanlarda otonom İHA navigasyonu	%84 görev başarımı	Gerçek zamanlı adaptasyon var, enerji optimizasyonu yok
Bu Çalışmada	DQN, DDQN, Düello DQN	Kapsam, adalet, enerji, acil durum müdahalesi	Gerçek zamanlı adaptasyon, Jain adaleti, %98 kapsama, %48 enerji tüketimi	—

3.MATERYAL VE YÖNTEM

Bu bölümde, çalışmada kullanılan sistem, algoritmalar, simülasyon ortamı ve sistem bileşenlerinin nasıl çalıştığı detaylı olarak açıklanmıştır.

3.1. Karasal Olmayan Ağ (NTN)

NTN, bağlantı sağlamak için hava veya uzay kaynaklı platformları kullanan iletişim ağlarını ifade eder. Bu platformlar arasında uydular, HAPS ve İHA yer almaktadır. NTN, özellikle Beşinci Nesil (Fifth Generation, 5G) ve ötesi haberleşmede, karasal altyapının sınırlı olduğu veya pratik olmadığı uzak, kırsal ve yetersiz hizmet alan bölgelere ağ kapsamını genişletmek için 3. Nesil Ortaklık Projesi (3rd Generation Partnership Project, 3GPP) standartlarına entegre edilmiştir. NTN küresel bağlantı sağlayarak karasal ağları tamamlar; geniş bant internet, afet müdahalesi ve Nesnelerin İnterneti (Internt of Things, IoT) iletişimi gibi uygulamaları destekler. Şekil 3.1'de örnek bir NTN bir mimarisi ve taksonomisi sunulmuştur.



Şekil 3.1. Örnek bir heterojen NTN mimarisi.

NTN, üç ana bileşen etrafında yapılandırılmıştır. Birinci bileşen, uzay/hava platformlarıdır. Bunlar arasında uydular, HAPS (örneğin balonlar ve hava gemileri) ve İHA yer alır. Bu platformlar, sinyal iletimi için baz istasyonu veya röle görevi görür. İkinci bileşen,

yer segmentidir ve NTN platformlarını karasal ağlara veya internet omurgasına bağlayan ağ geçidi istasyonlarından oluşur. Üçüncü bileşen ise kullanıcı segmentidir; radyo arayüzleri aracılığıyla NTN platformlarıyla iletişim kuran son kullanıcı cihazlarını içerir (örneğin akıllı telefonlar, IoT sensörleri veya özel terminaller).

NTN, kullanılan platforma göre sınıflandırılabilir:

Uydu Tabanlı NTN: kendi içinde üç sınıfa ayrılabilir:

- Alçak Dünya Yörüngesi (Low Earth Orbit, LEO) Uyduları: 500-2.000 km yükseklikte çalışır, düşük gecikme süresi (~20-50 ms) ve yüksek veri hızları sunar. Örnekler arasında Starlink ve OneWeb takımyıldızları bulunmaktadır.
- Orta Dünya Yörüngesi (Medium-Earth Orbit, MEO) Uyduları: 2,000-35,786 km'de konumlandırılmış, navigasyon (örn. GPS) ve bazı iletişim hizmetleri için kullanılır.
- Jeostatik Yörünge (Geostationary Earth Orbit, GEO) Uyduları: 35.786 km'de bulunur, geniş kapsama alanı sağlar ancak gecikme süresi daha yüksektir (~600 ms). Yayın hizmetleri için uygundur.

HAPS Tabanlı NTN: Stratosferde 20-50 km yükseklikte çalışır ve uydulardan daha düşük gecikme süresiyle bölgesel kapsama alanı sağlar. Örnekler arasında Google'ın Loon projesi bulunmaktadır.

İHA Tabanlı NTN: Afet senaryoları gibi geçici veya yerel kapsama için düşük irtifalarda (<10 km) konuşlandırılan dronlar veya İHA haberleşmelerini içerir.

NTN küresel iletişim yeteneklerini geliştiren önemli avantajlar sunar. Karasal altyapının pratik veya ekonomik olmadığı uzak, kırsal ve deniz bölgelerinde bağlantı sağlayarak geniş kapsama alanı sunarlar. Bu da NTN'yi dijital uçurumun kapatılması ve yetersiz hizmet alan bölgelerde geniş bant internet ve IoT gibi uygulamaların desteklenmesi için hayati hale getirmektedir. Ek olarak, NTN doğal afetler sırasında dayanıklılık gösterir, çünkü havadan veya uzaydan taşınan platformlar, karasal ağlar hasar gördüğünde iletişimi hızla yeniden kurabilir. Ölçeklenebilirlikleri, artan talepleri karşılamak için büyük uydu takımyıldızlarının veya HAPS'lerin konuşlandırılmasına olanak tanıyarak; IoT ve yüksek hızlı veri hizmetleri için devasa cihaz bağlantısını mümkün kılar. Ayrıca, NTN kesintisiz hareketliliği destekleyerek havacılık ve deniz lojistiği gibi sektörler için kritik öneme sahip olan gemi, uçak ve araç gibi hareketli platformlar için güvenilir iletişim sağlar.

Avantajlarına rağmen NTN, etkinliklerini sınırlayan çeşitli zorluklarla karşı karşıyadır. Uyduların fırlatılması veya HAPS ve İHA'nın işletilmesi gibi süreçleri barındıran NTN altyapısının konuşlandırılma ve bakım işlemlerinin yüksek maliyeti, özellikle LEO takımyıldızları gibi büyük ölçekli sistemler için önemli bir finansal engel oluşturmaktadır. Gecikme, özellikle yaklaşık 600 ms'lik gecikmeler sergileyen ve bu nedenle gerçek zamanlı oyun veya otonom araç kontrolü gibi gecikmeye duyarlı uygulamalar için uygun olmayan GEO uyduları için, bir endişe kaynağı olmaya devam etmektedir. Spektrum tahsisi ve uluslararası koordinasyon da dahil olmak üzere düzenleyici karmaşıklıklar, çatışmaları önlemek için küresel standartlar ve anlaşmalar gerektiğinden NTN dağıtımını daha da zorlaştırmaktadır. Ayrıca NTN, özellikle LEO uydu sistemlerinde parazit yönetiminde ve hızlı hareket eden platformlar arasında sorunsuz geçişler sağlamada teknik zorluklarla karşılaşmakta, bu da hizmet sürekliliğini kesintiye uğratabilmekte ve kullanıcı deneyimini bozabilmektedir.

3.2. İnsansız Hava Araçları (İHA)

İHA, geçici veya yerelleştirilmiş kablosuz bağlantı sağlamak için kullanılan havadan platformlardır. NTN bağlamında İHA, karasal altyapının mevcut olmadığı, hasar gördüğü veya yetersiz olduğu senaryolarda iletişimi kolaylaştıran mobil baz istasyonları veya röleler olarak hizmet vermektedir. Düşük irtifalarda (genellikle 10 km'nin altında) çalışan İHA, afet müdahalesi, olay tabanlı bağlantı ve hassas tarım gibi uygulamaları destekleyerek 5G ve ötesi ağların ayrılmaz bir parçasıdır. Esneklikleri ve hızlı konuşlandırma yetenekleri, onları özellikle kısa vadeli veya dinamik iletişim ihtiyaçları için NTN mimarilerinin kritik bir bileşeni haline getirmektedir.

İHA, belirli iletişim senaryolarına uygun şekilde tasarlanır ve operasyonel yeteneklerine göre kategorize edilir. Küçük uçaklara benzeyen sabit kanatlı İHA uzun uçuş dayanıklılığı için tasarlanmıştır ve geniş alanları kapsayabilir, bu da onları büyük ölçekli izleme veya kırsal bölgelerde bağlantı sağlamak için ideal hale getirir. Buna karşılık, döner kanatlı İHA, havada asılı kalma ve dikey kalkış/inişte mükemmeldir ve kentsel ortamlarda veya afet müdahale operasyonları sırasında hassas, yerelleştirilmiş bağlantı sağlar. Hibrit İHA, sabit kanatlı ve döner kanatlı tasarımların özelliklerini bir araya getirerek dayanıklılık

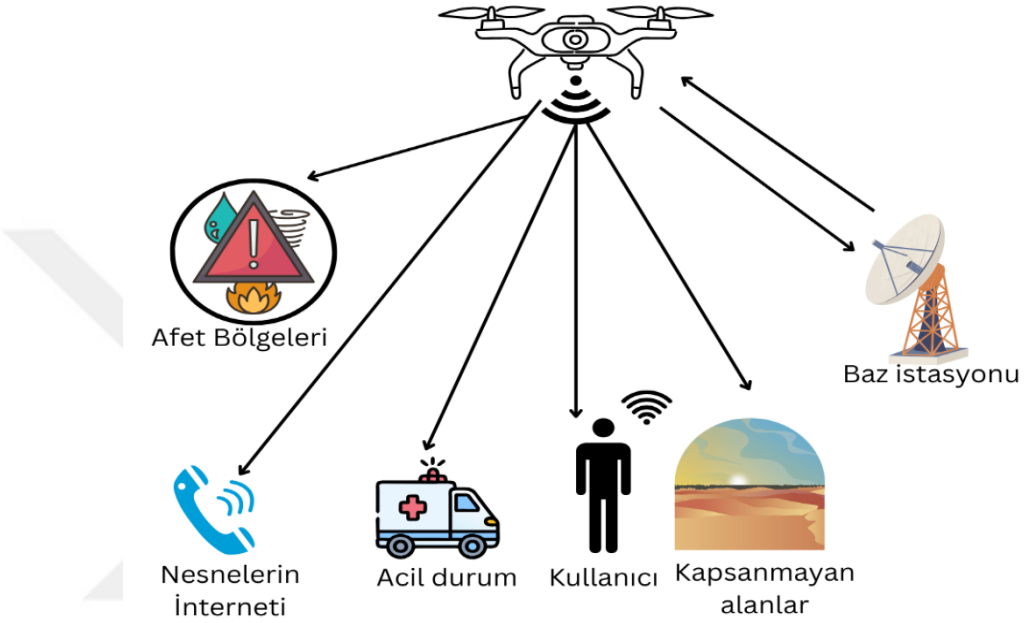
ve manevra kabiliyetini dengeleyerek çeşitli uygulamalar için çok yönlülük sunar. Ek olarak, sürü İHA koordineli gruplar halinde çalışarak, birden fazla İHA'nın sağlam ağ performansı sağlamak için birlikte çalıştığı büyük halk etkinlikleri veya arama-kurtarma görevleri gibi dinamik ortamlarda kapsama alanını, kapasiteyi ve esnekliği işbirliği içinde artırır.

İHA, esneklikleri ve hızlı konuşlandırma yetenekleri sayesinde NTN bağlamında belirgin avantajlar sunar. İHA, deprem veya sel gibi doğal afetlerden etkilenen ve karasal ağların kesintiye uğradığı bölgelerde, acil durum müdahale ekipleri ve etkilenen topluluklar için kritik iletişimi sağlamak amacıyla geçici bağlantı sağlamak üzere hızlı bir şekilde konuşlandırılabilirler. Alçak irtifada çalışmaları, uydu tabanlı NTN'ye kıyasla daha düşük gecikme süresi sağlar, bu da onları video akışı veya uzaktan algılama gibi gerçek zamanlı veri alışverişi gerektiren uygulamalar için uygun hale getirir. İHA aynı zamanda son derece hareketlidir, bu da kapsama alanını optimize etmek veya büyük kamu etkinlikleri gibi değişen ağ taleplerine yanıt vermek için dinamik yeniden konumlandırmaya imkan sağlar. Ayrıca, İHA'nın uydu veya yüksek irtifa platform sistemlerine kıyasla göreceli olarak düşük maliyeti, onları kısa vadeli veya yerel bağlantı ihtiyaçları için ekonomik bir çözüm haline getirir.

İHA, NTN'deki etkinliklerini etkileyen birkaç sınırlama ile karşı karşıyadır. Ana zorluk, özellikle serbest uçan İHA için sınırlı uçuş dayanıklılığıdır; bu İHA genellikle yalnızca 20-60 dakika çalışma süresi destekleyen bataryalara dayanır, bu da uzun vadeli bağlantı için kullanımını kısıtlar. İHA'nın ayrıca sınırlı yük kapasitesi vardır, bu da taşıyabilecekleri iletişim ekipmanlarının boyutunu ve karmaşıklığını kısıtlar, bu da diğer NTN platformları olan uydulara kıyasla ağ kapasitesini veya sinyal menzilini azaltabilir. Hava sahası düzenlemeleri ve spektrum tahsisi gibi düzenleyici kısıtlamalar, İHA operasyonlarının ulusal ve uluslararası havacılık kurallarına uyması gerektiğinden, önemli engeller teşkil eder ve bu da konuşlandırmayı geciktirebilir. Ayrıca, İHA güçlü rüzgarlar veya aşırı hava koşulları gibi çevresel faktörlere duyarlıdır, bu da uçuş stabilitesini ve ağ güvenilirliğini özellikle zorlu arazilerde veya olumsuz koşullarda bozabilir.

Şekil 3.2, acil durumlar ve uzak bölgelerdeki çeşitli İHA uygulamalarını göstermektedir. İHA, afet bölgelerine bağlanarak sel, yangın veya fırtına gibi kriz alanlarında hızlı değerlendirme ve iletişim sağlamaktadır. Acil servisleri destekleyerek ambulanslar ve diğer müdahale birimlerine kritik bilgileri ileterek koordinasyonu kolaylaştırmakta ve tepki süresini artırmaktadır. Ayrıca, IoT yoluyla etkilenen bireylere internet erişimi sunarak

geleneksel ağların başarısız olduğu yerlerde bağlantıyı sağlamaktadır. İHA, aynı zamanda farklı konumlardaki kullanıcılar için iletişim yeteneklerini genişletmektedir. Keşfedilmemiş veya erişilemez alanlarda, İHA uydu aracılığıyla baz istasyonlarıyla bağlantı kurarak veri iletimi ve izlemeyi geleneksel erişimin ötesindeki bölgelerde kolaylaştırmaktadır.



Şekil 3.2. İHA'nın bazı kullanımları

3.3. Derin Öğrenme (DL) Yöntemleri

DL, büyük veri hacimlerinden otomatik olarak özellikleri öğrenmek ve çıkarmak için çok katmanlı yapay sinir ağlarını kullanan bir ML dalıdır. Geleneksel ML modellerinin aksine, DL, hiyerarşik veri temsillerini doğrudan girdiden öğrenerek manuel özellik mühendisliği ihtiyacını ortadan kaldırır. Bu, görüntü tanıma, doğal dil işleme, otonom kontrol ve ardışık karar verme gibi karmaşık görevleri yönetmede son derece etkili olmasını sağlar.

DL teknikleri genellikle öğrenme paradigmalarına göre aşağıdaki kategorilere ayrılır:

- *Denetimli Öğrenme*: Bu türde, model etiketlenmiş veriler kullanılarak eğitilir, burada her giriş bilinen bir çıktı ile ilişkilendirilmiştir. Denetimli öğrenmede kullanılan DL modelleri örnekleri arasında görüntü sınıflandırması için Konvolüsyonel Sinir Ağları

(Convolutional Neural Network, CNN) ve dizilim tahmini için Tekrarlayan Sinir Ağları (Recurrent Neural Network, RNN) bulunur.

- *Denetimsiz Öğrenme*: Bu yöntemler, etiketlenmiş çıktılar olmadan verilerden desenler öğrenir. Kümeleme ve boyut indirgeme gibi görevlerde kullanılırlar. Örnekler arasında Otomatik Kodlayıcılar, Derin İnanç Ağları (Deep Belief Networks, DBN) ve Üretici Karşıt Ağlar (Generative Adversarial Network**,** GAN) bulunur.
- *Takviyeli Öğrenme (RL)*: Yukarıdaki iki yöntemden farklı olarak RL, bir ajanı bir ortamda eylemler gerçekleştirmesi için eğiterek toplam ödülü maksimize etmeyi amaçlar. Ajan, deneme yanılma yoluyla öğrenir ve ödüller veya cezalar şeklinde geri bildirim alır.

DL yöntemleri, akıllı ve uyarlanabilir karar vermeyi optimizasyonda önemli avantajlar sağlar. Bu yöntemler, yörünge planlama ve kaynak tahsisi için hassas stratejiler oluşturmak üzere kullanıcı konumları, çevresel koşullar ve İHA durumları gibi büyük miktarda heterojen veriyi işleyebilir, böylece kapsama alanını en üst düzeye çıkarır ve adil bağlantı sağlar. DL modelleri, kullanıcı hareketliliği veya hava durumu değişiklikleri gibi gerçek zamanlı değişikliklere uyum sağlayarak dinamik ortamlarda mükemmeldir ve bu da İHA esnekliğini ve verimliliğini artırır. Ayrıca, RL tabanlı DL yaklaşımları, açık programlama gerektirmeden enerji tasarruflu politikaları öğrenerek güç tüketimini en aza indirmek gibi uzun vadeli hedefleri optimize eder. Geleneksel optimizasyon teknikleriyle karşılaştırıldığında DL yöntemleri, NTN sistemlerindeki karmaşık, doğrusal olmayan ilişkileri ele alarak üstün ölçeklenebilirlik ve esneklik sunar; bu da akıllı şehirler ve hassas tarım gibi acil durum müdahalesinin ötesindeki uygulamalar için kritik öneme sahiptir.

DL yöntemleri, İHA optimizasyonuna uygulandığında çeşitli zorluklar ortaya çıkarmaktadır. DL modellerini eğitmenin ve konuşlandırmanın yüksek hesaplama karmaşıklığı, sınırlı yerleşik donanıma sahip kaynak kısıtlı İHA için pratik olmayabilecek önemli bir işlem gücü gerektirir. DL yöntemleri ayrıca eğitim için büyük, yüksek kaliteli veri kümeleri gerektirir ve koşulların büyük ölçüde değiştiği dinamik NTN ortamlarında bu tür verileri elde etmek zor ve zaman alıcı olabilir. Ayrıca, DRL'nin kara kutu doğası, yorumlanabilirliği azaltarak karar verme süreçlerini anlamayı veya doğrulamayı zorlaştırır, bu da İHA navigasyonu gibi güvenlik açısından kritik uygulamalar için bir endişe kaynağıdır. Son olarak, DL modelleri hiperparametrelere duyarlıdır ve kapsamlı ayarlama gerektirir ve

performansları eksik veya gürültülü verilerin olduğu senaryolarda düşebilir, bu da potansiyel olarak kaynak tahsisinde adaleti ve güvenilirliği etkiler.

3.3.1. Derin Takviyeli Öğrenme (DRL) Yöntemleri

DRL, yüksek boyutlu ve sürekli durum veya eylem uzaylarını işlemek için derin sinir ağlarını entegre eden geleneksel RL'nin güçlü bir uzantısıdır. DRL'de bir ajan, bir ortamla etkileşime girerek, mevcut durumu gözlemleyerek, bir eylemde bulunarak, bir ödül alarak ve yeni bir duruma geçerek bir dizi karar vermeyi öğrenir. Amaç, zaman içinde kümülatif ödülü maksimize eden optimal bir politika öğrenmektir.

Tablo yöntemlerine veya doğrusal yaklaşımlara dayanan geleneksel RL algoritmalarının aksine DRL; politika fonksiyonunu, değer fonksiyonunu veya her ikisini de yaklaşık olarak hesaplamak için derin sinir ağları kullanır. Bu, geleneksel yöntemlerin başarısız olduğu karmaşık ortamlarda ajanların etkili bir şekilde ölçeklenmesini sağlar.

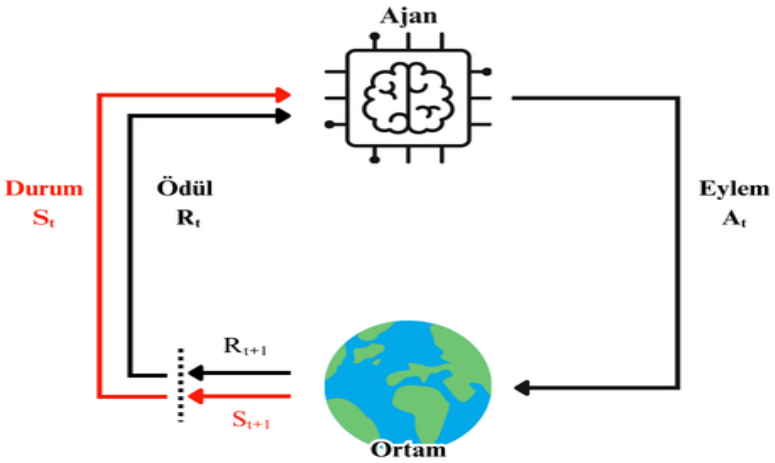
Bir RL problemi tipik olarak bir demet ile tanımlanan bir MDP vektörü olarak (S, A, P, R, γ) parametreleriyle modellenir. Burada S olası durumlar kümesi ve A olası eylemler kümesidir. P durum geçiş olasılık fonksiyonunu, R ödül fonksiyonunu temsil ederken; $\gamma \in [0,1]$ ise anlık ve gelecekteki ödülleri dengeleyen azaltma faktörüdür.

RL'deki ana hedef, her durumda alınacak en iyi eylemi tanımlayan ve beklenen toplam ödülü maksimize eden optimal bir politika π* bulmaktır, bu genellikle getiri olarak adlandırılır.

Şekil 3.3'te DRL sisteminin temel bileşenleri görülmektedir. DRL bileşenleri şu şekilde detaylandırılabilir:

- ***Ajan ve Ortam (Environment)***: Ajan, çevreden aldığı duruma (state) göre bir eylem (action) seçer ve bu eyleme karşılık ortamdan bir ödül (reward) ve yeni bir durum alır. Ajanın amacı, uzun vadede toplam ödülü maksimize etmektir.
- ***Durum Temsili (State Representation)***: Ortamın mevcut durumunu temsil eden vektördür. Karmaşık uygulamalarda (örneğin görüntüler, sensör verileri) bu durumlar sinir ağları aracılığıyla işlenerek anlamlı özellikler çıkarılır.
- ***Politika Ağı ($\pi_\theta(s)$)***: Girdisi durum olan ve çıktısı eylem olan sinir ağıdır. Bu ağ, ajan için bir politika öğrenir ve sürekli ya da kesikli eylem alanlarında kullanılabilir. Aktör-eleştirmen (actor-critic) ve politika tabanlı yöntemlerde temel bileşendir.

- ***Değer Fonksiyonu Ağı*** ***($V(s)$, $Q(s,a)$)***: Belirli bir durumun ya da durum-eylem çiftinin beklenen ödülünü tahmin eder. Değer tabanlı yöntemlerde (örneğin DQN), bu ağ kullanılarak en yüksek değere sahip eylem seçilir.
- ***Tekrarlama Belleği (Replay Buffer)***: Ajanın geçmiş depolayan belleğidir. Bu bellekteki örnekler, eğitim sırasında rastgele seçilerek kullanılır. Bu yöntem, veriler arasındaki korelasyonu azaltır ve eğitimi daha istikrarlı hale getirir.
- ***Hedef Ağ (Target Network)***: DQN ve DDPG gibi algoritmalarda kullanılan bu yapı, ana ağdan yavaşça güncellenen ayrı bir sinir ağıdır. Eğitim sırasında kararlılığı artırmak için kullanılır.
- ***Kayıp Fonksiyonu ve Optimizasyon***: Öğrenme sürecinde kullanılan hata fonksiyonudur (örneğin zamansal fark hatası). Bu hata, geri yayılım (backpropagation) ve optimizasyon algoritmaları (örneğin Adam) ile minimize edilir.
- ***Ödül Fonksiyonu (Reward Function)***: Ortamın ajana verdiği geri bildirimdir. Ajanın istenilen davranışları öğrenebilmesi için bu fonksiyon dikkatli bir şekilde tasarlanmalıdır.
- ***Keşif Stratejisi (Exploration Strategy)***: Ajanın öğrenme sırasında yeni eylemleri keşfetmesi için kullanılan yöntemdir. Keşif ve sömürü (exploration vs. exploitation) arasında denge kurmak esastır. Epsilon-Greedy, Ornstein-Uhlenbeck süreci veya Gaussian gürültü gibi stratejiler kullanılır.



Şekil 3.3. DRL sisteminin temel bileşenleri

DRL algoritmaları genel olarak aşağıdaki Şekil 3.4'teki gibi sınıflandırılabilir. Burada yer alan modelden bağımsız DRL türlerinden aşağıda kısaca bahsedilmektedir.

Değer Tabanlı DRL Algoritmaları:

Bu algoritmalar, belirli bir durumda belirli bir eylemi gerçekleştirmenin beklenen getirisini tahmin eden değer fonksiyonunu tanımlar. Politika, en yüksek değere sahip eylemlerin seçilmesiyle dolaylı olarak elde edilir. Temel örnekler şunlardır:

- DQN: DeepMind tarafından tanıtılan DQN, ayrık eylem uzaylarında eylem değerlerini tahmin etmek için Q-Öğrenmeyi derin sinir ağlarıyla birleştirir.
- DDQN: Eylem seçimi ve değerlendirmeyi birbirinden ayırarak Q-değerlerindeki aşırı tahmini ele alır.
- Düello DQN: Bir durumun değerinin tahminini ve her bir eylemin avantajını birbirinden ayırarak öğrenme verimliliğini artırır.

Politika Tabanlı DRL Algoritmaları:

Bu algoritmalar, bir değer fonksiyonu kullanmadan durumları eylemlere eşleyen politikayı doğrudan öğrenir. Özellikle sürekli eylem uzayları için kullanışlıdırlar. Örnekler şunları içerir:

- REINFORCE: Tam bölüm getirilerine dayalı olarak politikayı güncelleyen bir Monte Carlo politika gradyan yöntemidir.
- Proksimal Politika Optimizasyonu: Büyük politika değişimlerini önlemek için güncellemeleri sınırlayan ileri düzey ve kararlı bir politika gradyan yöntemidir.

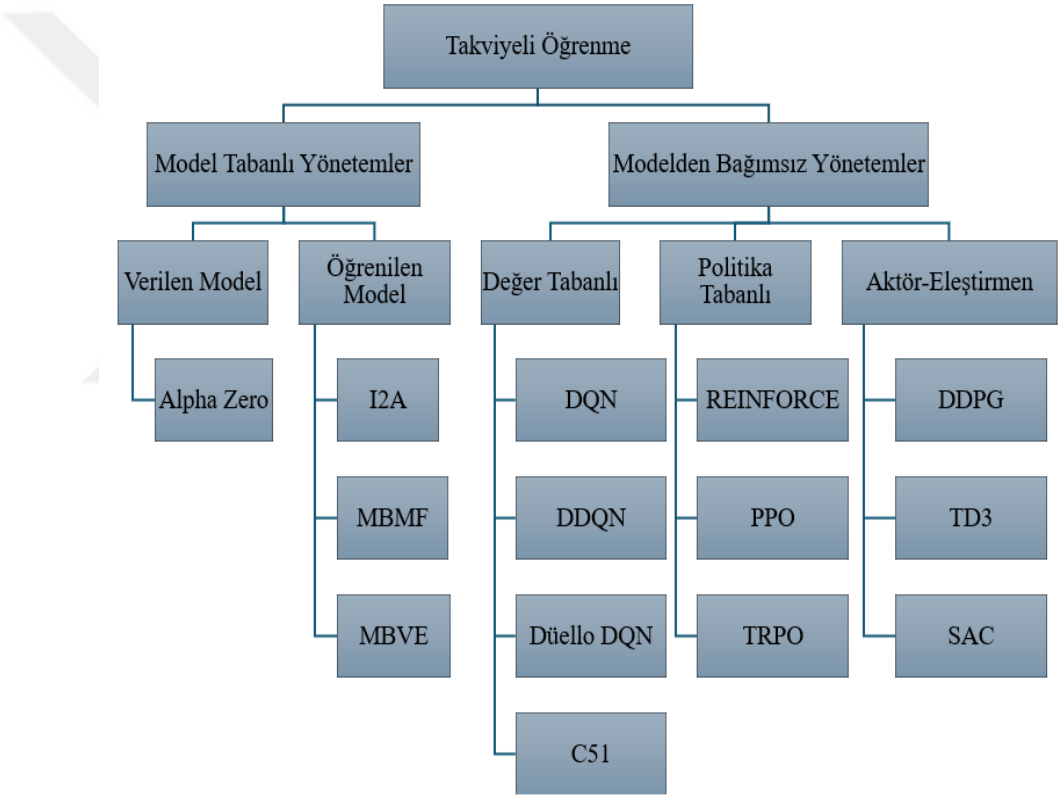
Actor-Critic DRL Algoritmaları:

Bunlar hem değer tabanlı hem de politika tabanlı yöntemleri birleştirir. Aktör ağı politikayı güncellerken, "eleştirmen" aktör tarafından alınan eylemleri değerlendirmek için değer fonksiyonunu tahmin eder. Örnekler arasında şunlar yer alır:

- DDPG: Aktörün sürekli eylemler ürettiği ve eleştirmenin bunları değerlendirdiği sürekli eylem uzayları için uygundur.
- İkiz Gecikmeli DDPG (Twin-Delayed DDPG): Aşırı tahmin ve eğitim istikrarsızlığını ele alarak DDPG'yi geliştirir.
- Yumuşak Aktör-Eleştirmen (Soft Actor Critic, SAC): Entropi maksimizasyonu yoluyla keşfi teşvik eder, ödül ve politika rastgeleliğini dengeler.

- MADDPG: DDPG'yi çoklu ajan ortamlarına genişleterek ajanların ortak ortamlarda işbirlikçi veya rekabetçi davranışlar öğrenmesine olanak tanır.

RL; robotik, otonom araçlar ve İHA kontrolü gibi gerçek zamanlı etkileşim ve dinamik ortamlara uyum sağlamanın kritik olduğu otonom karar verme görevleri için özellikle uygundur. Bu tez çalışmasında değer tabanlı DRL algoritmaları kullanılmıştır ve bu bölümde sırayla alt başlıklar halinde açıklanmıştır.



Şekil 3.4. DRL algoritmaları

3.3.1.1 Derin Q-Öğrenme (DQL)

Derin Q-öğrenme (Deep Q-Learning, DQL), klasik Q-öğrenme algoritmasının DL ile güçlendirilmiş bir versiyonudur. Geleneksel Q-öğrenme, her bir durum-eylem çiftine karşılık gelen ödül beklentisini (Q-değerini) saklamak için bir Q-tablosu kullanır. Ancak, büyük ve sürekli durum alanlarında bu tabloyu tutmak hem bellek açısından hem de hesaplama

açısından mümkün değildir. Bu problemi çözmek için DQL, Q-fonksiyonunu yaklaşık olarak öğrenen bir derin sinir ağı kullanır.

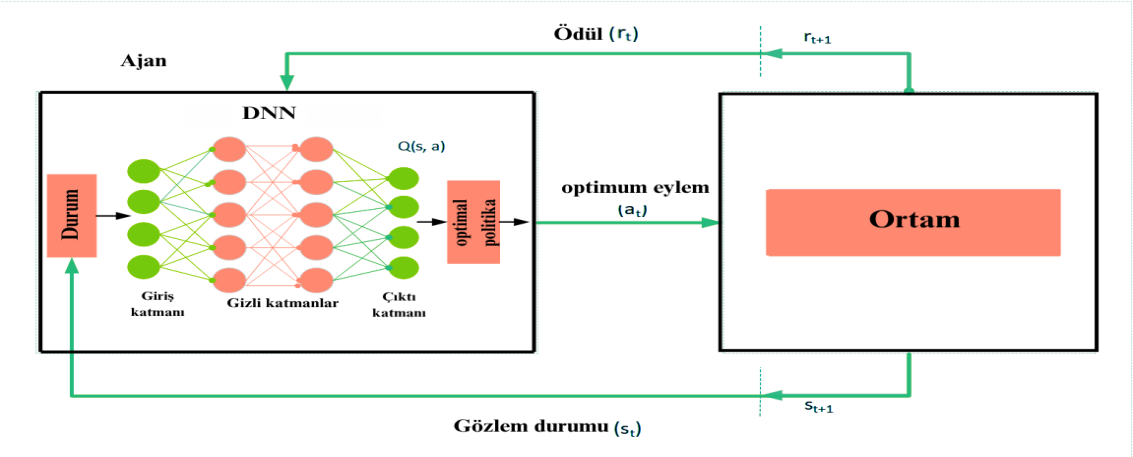
DQL'nin amacı, ajan için ideal bir politika öğrenmektir. Bunun için ajan her durumda hangi eylemi seçeceğini belirlemek üzere bir Q-ağı $Q(s,a;\theta)$ kullanır. Bu ağ, Bellman denklemine göre eğitilir:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left(r + \gamma \max_{a'} Q(s',a';\theta^{-}) - Q(s,a;\theta)\right) \quad (3.1)$$

Bu denklemde s mevcut durum, a yapılan eylem, r alınan anlık ödül, s' yeni durum, γ azaltma faktörü (gelecek ödüllere verilen önem), θ ana ağın parametreleri, θ^{-} sabit hedef ağın parametreleri (belirli aralıklarla güncellenir) olarak adlandırılır.

DQL, temel bileşenleriyle etkili bir öğrenme süreci sunar. Q-ağı, durum ve eylemleri giriş olarak alarak Q-değerlerini çıkış olarak üretir ve ajanın karar mekanizmasını oluşturur. Hedef Ağı, eğitimin kararlılığını artırmak için Q-ağının sabit bir kopyası olarak kullanılır ve belirli aralıklarla güncellenir. Deneyim tekrarı, ajanın geçmiş deneyimlerini bir tampon bellekte saklayarak, eğitim sırasında bu deneyimlerden rastgele örnekler seçip öğrenmeyi sağlar. Epsilon-Greedy Stratejisi ise ajanın çevreyi keşfetmesi için rastgele eylemler (ϵ olasılığıyla) seçmesini veya en iyi eylemi ($1-\epsilon$ olasılığıyla) tercih etmesini sağlayarak keşif ve sömürü arasında denge kurar. Bu bileşenler, DQL'nin güçlü ve istikrarlı bir öğrenme algoritması olmasını sağlar.

Şekil 3.5, belirli bir s durumunda bir a eylemini gerçekleştirmenin gelecekteki beklenen ödüllerini tahmin eden Q-değer fonksiyonu $Q(s,a)$'ya yaklaşmak için bir Derin sinir ağı (Deep neural network, DNN) kullanan bir RL yaklaşımı olan DQN çerçevesini temsil eder. DNN, en uygun eylemi belirlemek için çevreden gelen mevcut s_t durumunu; girdi katmanları, gizli katmanlar ve çıktı katmanları aracılığıyla işler. Ajan bu eylemi seçer ve uygular, r_{t+1} ödülünü alır ve bir sonraki duruma (s_{t+1}) geçer. Süreç, ajanın zaman içinde karar verme sürecini iyileştirmek için ödül sinyaline dayalı olarak Q-değeri tahminlerini iyileştirmesiyle tekrar eder.



Şekil 3.5. DQL Temel Öğrenme Mekanizması

3.3.1.1.1. Derin Q-Ağı (DQN)

DQN, Q-öğrenme algoritmasının DL teknikleri ile güçlendirilmiş versiyonudur ve derin sinir ağları kullanarak karmaşık durum-eylem ortamlarında öğrenme yapılmasını sağlar. DQN, büyük veya sürekli durum uzaylarını işlemek için Q-Öğrenmeyi derin sinir ağlarıyla birleştiren bir RL algoritmasıdır. Bir sinir ağı kullanarak Q-değeri fonksiyonuna yaklaşır ve öğrenmeyi stabilize etmek için deneyim tekrarı ve hedef ağları kullanır. DQN'nin en önemli avantajı, geleneksel Q-Öğrenmesini Atari oyunları veya İHA konumlandırma görevleri gibi karmaşık ortamlara ölçeklendirme yeteneğidir. Bununla birlikte, DQN, Q-değerlerini tahmin ederken aşırı tahmin önyargısından muzdarip olabilir, bu da belirli ortamlarda optimal olmayan politika kararlarına ve dengesiz eğitime yol açabilir.

DQN, bir durumu girdi olarak alıp, her mümkün eylem için tahmin edilen Q-değerlerini çıktı olarak veren bir derin sinir ağı kullanır. Bu ağın amacı, ajan için en yüksek uzun vadeli ödülü verecek eylemleri seçmesini sağlamaktır. Eğitilen bu ağ $Q(s,a;\theta)$ fonksiyonunu temsil eder.

DQN, klasik Q- öğrenme algoritmasında olduğu gibi Bellman denklemini temel alır. Ancak burada Q-değeri bir tablo ile değil, sinir ağı ile tahmin edilir. Öğrenme süreci, (3.1)'deki hedef Q-değeri ile yapılır. Bu hedef ile mevcut tahmin arasındaki fark, Ortalama Karesel Hata (Mean Squared Error, MSE) kaybı kullanılarak minimize edilir:

$$L(\theta) = \mathbb{E}_{(s,a,r,s')}\left[\left(y - Q(s,a;\theta)\right)^2\right] \quad (3.2)$$

DQN, etkinliğini sağlayan temel bileşenlerden oluşur. Ana Q-ağı, öğrenilen politikayı temsil eder ve her eğitim adımında güncellenir. Hedef Ağ, ana Q-ağının bir kopyası olup belirli aralıklarla (τ) güncellenerek eğitimin kararlılığını artırır ve aşırı güncellemeyi önleyerek öğrenme sürecinin dengesini sağlar. Deneyim tekrarı, ajanın geçmiş deneyimlerini (s, a, r, s') dörtlüleri şeklinde bir hafızada (replay buffer) saklar ve eğitim sırasında rastgele örnekler seçer; bu yöntem, veriler arasındaki korelasyonu azaltır, ajanın önceki durumları tekrar görmesini sağlar ve eğitimi daha kararlı hale getirir. Epsilon-Greedy Eylem Seçimi ise ajanın çevreyi keşfetmesi için ϵ olasılığıyla rastgele eylem seçmesini, kalan $(1 - \epsilon)$ olasılıkta ise Q-ağının önerdiği en yüksek ödüllü eylemi tercih etmesini sağlar; ϵ değeri zamanla azaltılarak keşiften sömürüye geçiş yapılır.

DQN, DRL alanında önemli başarılar elde etmiştir. Q-Öğrenme algoritmasının genelleştirme eksikliğini, DL tekniklerini entegre ederek aşmış ve bu sayede karmaşık problemlerde etkili çözümler üretmiştir. Ayrıca, deneyim tekrarı ve hedef ağ gibi yenilikçi yapılar, derin pekiştirmeli öğrenme sistemlerinde bir standart haline gelerek, daha stabil ve verimli öğrenme süreçlerinin önünü açmıştır. Özellikle robotik navigasyon sistemlerinde otonom hareket kabiliyeti sağlamak, otomatik sürüş sistemlerinde araçların çevresel karar alma süreçlerini optimize etmek, dinamik kaynak yönetimi ile verimli kaynak dağıtımı gerçekleştirmek ve İHA'nın konumlandırması ile enerji verimliliğini artırmak gibi alanlarda da etkin bir şekilde uygulanmaktadır.

DQN, önemli başarılarına rağmen bazı kısıtlarla karşı karşıyadır. İlk olarak, Q-değerlerinin aşırı tahmini (overestimation bias) sorunu ortaya çıkabilir; DQN, Q-değerlerini bazen olması gerekenden daha yüksek tahmin ederek politikaların sapmasına yol açabilir. Bu sorun, DDQN gibi sonraki versiyonlarla giderilmeye çalışılmıştır. Ayrıca, ağ kararsızlığı bir diğer önemli kısıttır; eğitim sürecinde Q-değerlerinin sürekli değişmesi, sinir ağı eğitimini dengesiz hale getirebilir. Bu nedenle, hedef ağ ve deneyim tekrarı gibi yapılar, eğitimin stabilitesi için kritik öneme sahiptir.

3.3.1.1.2 Çift Derin Q-ağı (DDQN)

DDQN, orijinal algoritmanın doğasında bulunan aşırı tahmin önyargısını azaltmak için tasarlanmış gelişmiş bir DQN sürümüdür. Bunu, Q-değeri güncellemeleri sırasında eylem seçimi ve eylem değerlendirmesini birbirinden ayırarak eylem-hedef ağı (action-target

network) kullanarak değerlendirirken çevrimiçi ağı kullanarak en iyi eylemi seçer. Bu da daha doğru değer tahminleri ve daha istikrarlı bir eğitim sağlar. DDQN genellikle birçok görevde DQN'den daha iyi ve daha tutarlı performans gösterse de yine de dikkatli ayarlama gerektirir ve oldukça karmaşık veya durağan olmayan ortamlarda zorlanabilir.

Bu yöntemde, eylemi seçmek ve Q-değerini tahmin etmek için iki farklı ağ kullanılır:

$$Q_{\text{Double}}(s,a) = r + \gamma Q\left(s', \underset{a'}{\text{argmax}} Q(s',a';\theta);\theta^{-}\right) \quad (3.3)$$

Burada farklı aksiyon seçimini ana ağ (θ) yapar, Q-değeri hedef ağ (θ^{-}) tarafından tahmin edilir ve daha istikrarlı ve doğru Q-değerleri tahmini sağlar.

3.3.1.1.3 Düello DQN

Düello DQN, durum değer fonksiyonunu $V(s;\theta)$ ve avantaj fonksiyonunu $A(s,a;\theta)$ ayrı ayrı tahmin eden ve daha sonra Q-değerleri üretmek için birleştirilen bir sinir ağı mimarisi sunar. Bu, özellikle bazı eylemler sonucu önemli ölçüde etkilemediğinde, ajanın eylemden bağımsız olarak bir durumun değerini daha iyi değerlendirmesine yardımcı olur. Düello DQN'nin ana avantajı, özellikle benzer değerli birçok eylemin bulunduğu ortamlarda öğrenme verimliliğinin ve politika performansının iyileştirilmesidir. Bununla birlikte, mimari karmaşıklık ekler ve daha basit görevlerde veya durum-eylem değerleri zaten iyi ayrılmış olduğunda önemli kazanımlar sunmayabilir. Özellikle hangi eylemin daha iyi olduğunu seçmenin zor olduğu durumlarda daha iyi performans sağlar.

Q-değeri bu iki bileşenin ($V(s;\theta)$ ve $A(s,a;\theta)$) birleşimiyle elde edilir:

$$Q(s,a;\theta) = V(s;\theta) + \left(A(s,a;\theta) - \frac{1}{|\mathcal{A}|}\sum_{a'\in|\mathcal{A}|} A(s,a';\theta)\right) \quad (3.4)$$

Burda $V(s;\theta)$ durum değer fonksiyonu, s durumunda olmak için beklenen ödülü tahmin edip $A(s,a;\theta)$ avantaj fonksiyonu, a eylemini gerçekleştirmenin s durumundaki ortalama eyleme kıyasla ne kadar daha iyi olduğunu ölçer. Ayrıca $|\mathcal{A}|$, eylem uzayında olası eylemlerin sayısı ve $\frac{1}{|\mathcal{A}|}\sum_{a'\in|\mathcal{A}|} A\,(s,a';\theta)$ terimi, avantaj fonksiyonunu normalleştirmek için kullanılan tüm eylemler üzerindeki ortalama avantajı temsil eder.

Q-değerinin hesaplanma yöntemi, özellikle bazı durumlarda hangi eylemin seçileceğinin önemli olmadığı, sadece durumun değerli olup olmadığının önemli olduğu durumlarda öğrenmeyi daha verimli hale getirir. Ayrıca, hangi eylemin daha iyi olduğunu seçmenin zor olduğu durumlarda daha iyi performans sağlar.

3.4 Konumlandırma Yöntemleri ve Sistemleri

Konumlandırma, bir nesnenin, kullanıcının veya cihazın fiziksel ortam üzerindeki yerinin belirlenmesi işlemidir. Günümüzde bu ihtiyaç, başta mobil iletişim, acil durum yönetimi, akıllı ulaşım sistemleri, askeri uygulamalar ve IoT gibi birçok alanda kritik öneme sahiptir. Konumlandırma sistemleri, kullanılan teknolojiye, hedeflenen doğruluk düzeyine ve uygulama senaryosuna bağlı olarak farklı sınıflandırmalara tabi tutulmaktadır.

3.4.1 İHA Tabanlı Konumlandırma Sistemleri

İHA, esnek hareket kabiliyetleri ve geniş görüş alanları sayesinde konumlandırma teknolojilerinde önemli bir rol üstlenmektedir. Özellikle kara altyapısının yetersiz olduğu bölgelerde ya da Küresel Navigasyon Uydu Sistemi (Global Navigation Satellite System, GNSS) sistemlerinin çalışmadığı senaryolarda, İHA mobil konumlandırma platformları olarak öne çıkar. İHA tabanlı konumlandırma sistemleri, yer kullanıcılarından alınan çeşitli sinyal ölçümlerine (alınan sinyal gücü göstergesi, varış zamanı, varış zaman farkı, varış açısı vb.) dayanarak hedef konumunu belirlemeye çalışır. Bu sistemlerin temel amacı; yüksek doğruluk, düşük gecikme ve geniş kapsama alanı ile gerçek zamanlı konum bilgisini elde etmektir. Günümüzde İHA tabanlı konumlandırma sistemleri çeşitli yöntemlerle sınıflandırılabilir. Bu yöntemler, kullanılan algoritma yapısına ve karar mekanizmalarına göre farklılık gösterir. Aşağıda bu yöntemler detaylı olarak ele alınmıştır.

a) Kural Tabanlı İHA Konumlandırma Sistemleri:

Kural tabanlı konumlandırma sistemlerinde İHA'nın hareketi, önceden tanımlanmış kurallar ve senaryolar üzerinden yürütülür. Bu kurallar genellikle sabit matematiksel ifadeler, koşullu ifadelere (if-else yapıları) dayalı hareket stratejileri veya belirli eşiğe göre yapılan konum tahminleri içerir. Örneğin, sinyal gücü belli bir eşiğin altına düştüğünde İHA belirli

bir açıyla pozisyonunu değiştirir ya da kullanıcı yoğunluğu bir bölgeye kaydığında İHA otomatik olarak o bölgeye yönelir. Bu sistemler, gerçek zamanlı uygulamalarda düşük hesaplama maliyeti ve uygulama kolaylığı sunar; ancak çevresel değişkenliklere veya belirsizliklere karşı uyarlanabilirlikleri sınırlıdır. Doğruluk oranları genellikle sistem tasarımına ve kuralların çevreye uygunluğuna bağlıdır.

b) Kuralsız (İstatistiksel veya Veri Tabanlı) Konumlandırma Sistemleri:

Kuralsız konumlandırma sistemleri, önceden tanımlı sabit kurallar yerine, geçmişten toplanan veriler ve istatistiksel çıkarımlar üzerinden çalışır. Bu tür sistemlerde İHA'nın davranışları, çevreden elde edilen verilere göre şekillenir. Örneğin, belirli bir kullanıcı dağılımı ve sinyal yoğunluğuna göre İHA, daha önceki benzer durumlarda verilen tepkilere benzer hareketler sergiler. Bu yöntemler genellikle regresyon modelleri, olasılık dağılımları veya k-ortalamaları gibi sınıflandırma algoritmalarıyla desteklenir. Esnek olmalarına rağmen, bu sistemler büyük veri kümeleri gerektirir ve anlık değişimlere yanıt verme süresi kural tabanlı sistemlere göre daha uzundur.

c) Yapay Zekâ ve RL Tabanlı Konumlandırma Sistemleri:

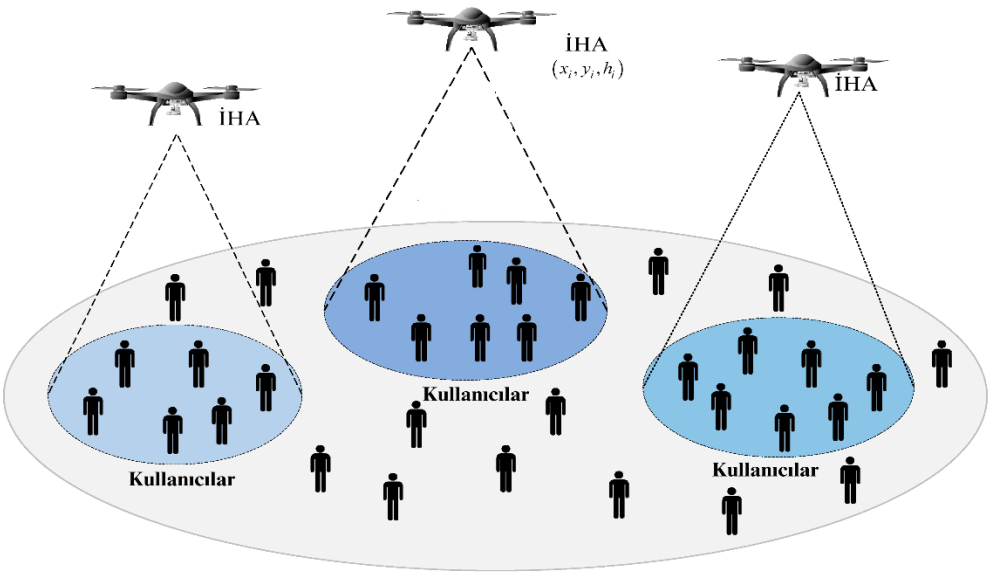
Yapay zekâ destekli sistemler, özellikle DRL algoritmaları ile birlikte, İHA'nın otonom şekilde çevreyi öğrenmesini ve en uygun konumlandırma kararlarını almasını sağlar. Bu sistemlerde ajan (İHA), çevreden aldığı durum bilgisine (örneğin kullanıcı konumu, enerji seviyeleri, sinyal gücü) göre aksiyon alır ve bu aksiyonun sonucuna göre ödül veya ceza alarak gelecekteki kararlarını günceller. Bu yaklaşımlar arasında Q- Öğrenme, DQN, DDPG, PPO ve MADDPG gibi çeşitli algoritmalar öne çıkmaktadır. Yapay zekâ tabanlı sistemlerin en büyük avantajı, dinamik ortamlara adaptasyon yeteneklerinin yüksek olmasıdır. Özellikle kullanıcıların sürekli yer değiştirdiği, enerji seviyelerinin dalgalandığı veya acil durumların meydana geldiği senaryolarda yüksek performans gösterirler. Bununla birlikte, bu sistemlerin eğitilmesi için yüksek işlem gücü ve yeterli sayıda simülasyon verisi gereklidir.

d) Hibrit ve Ortam Farkındalıklı Konumlandırma Sistemleri:

Bazı İHA tabanlı konumlandırma sistemleri, yukarıdaki yöntemlerin kombinasyonlarını kullanarak hem doğruluğu hem de uyarlanabilirliği artırmayı hedefler.

Örneğin bu hibrit sistemlerde, temel kural tabanlı karar yapısı bir güvenlik ağı olarak kullanılırken, yapay zekâ modülü dinamik çevre koşullarına göre daha esnek kararlar alabilir. Aynı zamanda, ortam farkındalıklı sistemler çevredeki coğrafi engelleri, hava durumu koşullarını ve kullanıcı yoğunluğunu analiz ederek konumlandırma performansını artırır. Bu tür sistemler; özellikle afet yönetimi, arama-kurtarma operasyonları ve geniş alan kapsama uygulamalarında oldukça etkilidir.

Bu tez çalışmasında geliştirilen İHA tabanlı konumlandırma sistemi, yukarıda açıklanan yöntemler arasından RL tabanlı konumlandırma kategorisine girmekte olup, özel olarak DQL algoritmaları temelinde tasarlanmıştır. DQL algoritması, çevresel dinamikleri (kullanıcı konumu, yoğunluk vb.) göz önünde bulundurarak İHA'nın optimum konumunu öğrenmesini sağlar. Bu sayede sistem, kullanıcıların dağılımına ve anlık ihtiyaçlarına göre kendi stratejisini güncelleyerek, dinamik ve enerjiye duyarlı bir kapsama ağı oluşturur. DQL tabanlı yaklaşım, özellikle sabit kurallarla veya istatistiksel modellerle sınırlandırılmış sistemlerin aksine, belirsizlik altındaki ortamlarda yüksek performans ve esneklik sunar. Böylece, önerilen yöntem; enerji verimliliği, kapsama adaleti ve acil durumlara hızlı yanıt gibi hedefleri bir arada gerçekleştirme potansiyeline sahiptir. Dronların alanları ve kullanıcıları nasıl kapsadığına dair bir örnek Şekil 3.5'te gösterilmektedir.



Şekil 3.5. İHA Tabanlı Konumlandırma

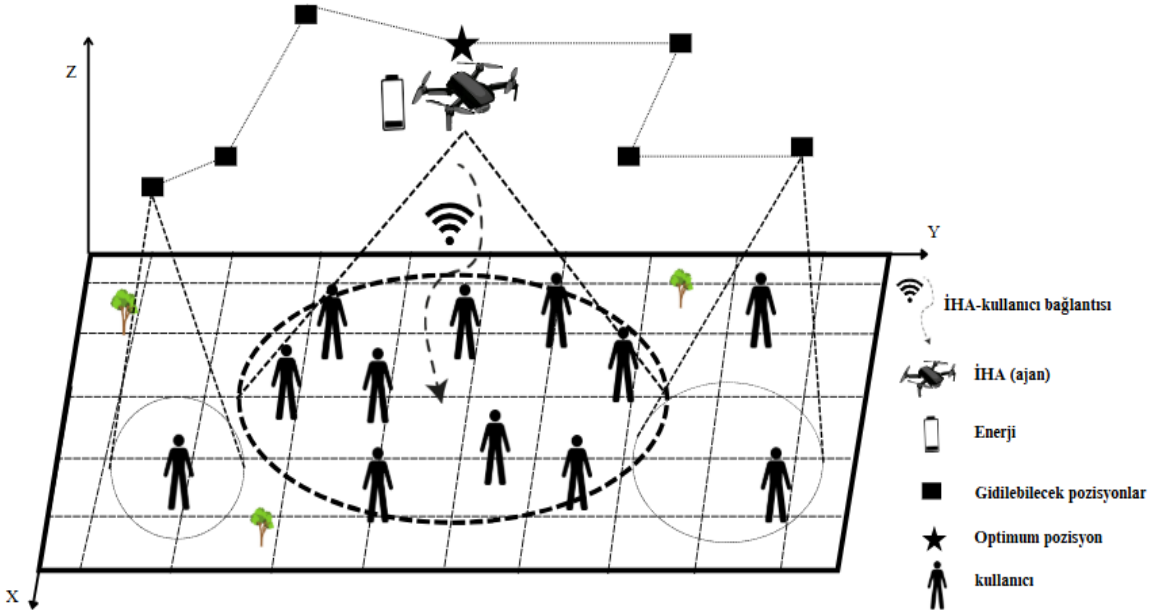
4.ARAŞTIRMA SONUÇLARI VE TARTIŞMA

Bu tez çalışmasında, yer kullanıcılarının kapsama alanını en üst düzeye çıkarmak, güç tüketimini en aza indirmek, tüm kullanıcılar arasında adaleti sağlamak ve afet olaylarında bir acil müdahale mekanizması sağlamak için İHA-destekli bir haberleşme ağı modeli önerilmektedir.

Bu bölümde sistem modelinin bileşenleri detaylandırılmakta ve ağ performansını iyileştirmek için önerilen çözüm yaklaşımı tartışılmaktadır. Ayrıca, çözüm için kullanılan DQN, DDQN ve Düello DQN RL algoritmalarının uygulanması ile ulaşılan deneysel sonuçlar tartışılmaktadır.

4.1 Sistem Modeli

Önerilen sistem, kablosuz hizmet kapsamını geliştirmek, güç tüketimini azaltmak ve dinamik bir ortamda kullanıcılar arasında adaleti sağlamak için tasarlanmış bir RL tabanlı İHA dinamik konumlandırma modeli sunmaktadır. Sistem, tek bir İHA'nın yerde rastgele dağıtılmış 100'e kadar kullanıcıya hizmet vermek üzere konuşlandırıldığı 4 kilometrekarelik bir 3B alanda çalışmaktadır. Bu kullanıcılar zaman içinde alana girip çıkabilirler, bu da İHA'nın optimum hizmet sunumu için konumunu sürekli olarak uyarlamasını gerektirir. İHA, kapsama alanını en üst düzeye çıkarmak, toplam enerji kullanımını en aza indirmek ve kullanıcı erişiminde adaleti sağlamak amacıyla RL algoritmaları, özellikle DQN, DDQN ve Düello DQN, kullanarak gerçek zamanlı olarak kendini akıllıca yeniden konumlandırmak üzere eğitilmiştir. İHA, kullanıcı konumları ve yoğunluk dağılımları gibi çevresel girdileri alır ve zaman içinde performansı artırmak için en uygun hareket kararlarını verir. Ek olarak, sistem, İHA'nın konumunu ani afet bölgelerine veya acil müdahale gerektiren yüksek talep alanlarına doğru hızla kaydırmasını sağlayan bir acil durum müdahale mekanizması içerir. Uygulama Python'da gerçekleştirilmiş ve performans; kapsama oranı, güç tüketimi, adalet endeksi ve acil durum müdahale süresi gibi ölçütler kullanılarak değerlendirilmiştir. Şekil 4.1'te, tez çalışması kapsamında ele alınan senaryo görselleştirilmiştir.



Şekil 4.1. Tez kapsamında gerçekleştirilen çalışmalar için oluşturulan dinamik İHA konumlandırma senaryosu.

4.1.1 Kanal ve Kapsama Modeli

Herhangi bir kullanıcı ile İHA arasındaki bağlantının LoS ve LoS olmayan yayılım modellerini izlediği varsayılır. Mozaffari ve ark. (2016)'nın çalışmasında ayrıntılı olarak açıklandığı üzere, serbest uzaydaki bir kanal yol kaybı modeli dB cinsinden aşağıdaki gibi ifade edilebilir:

$$\xi_{uk} = 10 \times n_0 \times \log\left(\frac{4\pi f_c d_{uk}}{c}\right) \quad (4.1)$$

Burada f_c sistem taşıyıcı frekansını, d_{uk} ise İHA ile kullanıcı arasındaki mesafeyi temsil eder. Ayrıca, c ışık hızıdır ve n_0 ortama özgü (kırsal, kentsel, yoğun kentsel) yol kaybı üssüdür. İHA-Yer (User to Ground, U2G) bağlantılarını modellemek için kullanılan popüler bir yaklaşım olasılıksal LoS ve NLoS'tur (Mozaffari ve ark., 2016). Burada NLoS genellikle ortamdaki gölgelenme ve kırınım sorunlarından kaynaklanır ve NLoS'ta ortaya çıkan zayıflamalar, LoS'a kıyasla İHA'lar üzerinde daha büyük bir etkiye sahiptir. İHA ile kullanıcı k arasındaki yol kaybı aşağıdaki gibi ifade edilebilir (Challita ve ark., 2017):

$$L_{uk}(dB) = \begin{cases} \xi_{uk} + \chi_{LoS}\,, & LoS\ hattı \\ \xi_{uk} + \chi_{NLoS}\,, & NLoS\ \text{hattı} \end{cases} \quad (4.2)$$

Burada ξ_{uk}, Eşitlik 4.1'ye göre hesaplanırken; χ_{LoS} ve χ_{NLoS} gölgelenme sorununun neden olduğu ek zayıflamayı temsil eder. Bu sistemde, LoS bağlantısının olasılığı, kullanıcının ve İHA'nın konumları ve kullanıcı ile İHA arasındaki yükseklik açısı gibi ortama dayalı bir dizi değişkene bağlıdır. Bu nedenle, LoS ve NLoS olasılıkları aşağıdaki gibi ifade edilebilir (Nemer ve ark., 2022):

$$P_{uk}^{LoS} = \delta \times \left(\frac{180}{\pi} \times \theta_k\right)^B \quad (4.3)$$

$$P_{uk}^{NLoS} = 1 - P_{uk}^{LoS} \quad (4.4)$$

Burada δ ve B çevreye bağlı sabit parametreler olup yükseklik açısı θ_k aşağıdaki gibi elde edilebilir:

$$\theta_k = \sin^{-1}\left(\frac{h_u - h_k}{d_{uk}}\right) \quad (4.5)$$

Burda h_u ve h_k sırasıyla İHA yüksekliğini ve kullanıcı yüksekliğini gösterir. d_{uk}, İHA ile k kullanıcısı arasındaki 3B mesafeyi temsil eder ve şu şekilde hesaplanabilir:

$$d_{uk} = \sqrt{(x_{\text{İ}HA} - x_k)^2 + (y_{\text{İ}HA} - y_k)^2 + h_u^2} \quad (4.6)$$

Burada, $(x_{\text{İ}HA}\,, y_{iHA}\,)$ ve $(x_k\,, y_k)$ sırasıyla İHA'nın ve kullanıcının yatay koordinatlarını temsil eder.

Tüm bunlar dikkate alındığında, Mozaffari ve ark. (2016) ile Zhang ve ark. (2020)'nın çalışmalarında olduğu gibi, ortalama yol kaybı aşağıdaki gibi ifade edilebilir:

$$\bar{L}_{uk} = P_{uk}^{LoS} \times L_{uk}^{LoS} + P_{uk}^{NLoS} \times L_{uk}^{NLoS} \quad (4.7)$$

Kullanıcının kapsama olasılığı, kullanıcı ile İHA arasındaki ortalama yol kaybı kullanılarak değerlendirilebilir. Kullanıcı İHA'nın iletişim menziline girdiğinde, bu kullanıcının kapsandığını kabul edilmiştir. Herhangi bir kullanıcının kapsama değeri; P_{jk},

İHA *j*'nin kullanıcı *k*'yi kapsama olasılığı ve *N* tane İHA bulunan bir ağ için bir kontrol stratejisi olarak kabul edilerek ve aşağıdaki gibi ifade edilebilir (Nemer ve ark., 2022):

$$C_k = \left(1 - \Pi_{j \in N}\left(1 - P_{jk}\right)\right) \tag{4.8}$$

4.1.2 Adalet

Çalışmanın temel amacı, minimum enerji gereksinimi ile genel kapsama puanını maksimize etmektir. Ancak, bu motivasyon, bölgedeki bazı kullanıcılar için adil olmayan kapsama alanına neden olabilir. Başka bir deyişle, bazı kullanıcılar uzun süre kapsanırken, diğerleri görev süresi boyunca nadiren kapsanabilir. Bu nedenle, bölgedeki tüm kullanıcılar için adil bir kapsamayı garanti etmek gerekir. Bu, Jain adalet endeksi adı verilen bir ölçüm metriği ile gerçekleştirilebilir (Jain ve ark., 1984):

$$\mathrm{F} = \frac{\left(\sum_{k=1}^{K} C_k\right)^2}{K\left(\sum_{k=1}^{K} C_k^2\right)} \tag{4.9}$$

Bu eşitlikte $\mathrm{F} \in [0,\ 1]$ adalet indeksi ne kadar büyükse kapsamın o kadar adil olduğu kolayca görülebilir. K, ağdaki toplam kullanıcı sayısıdır.

4.1.3 Güç Tüketimi

Genel olarak, İHA'ların enerji tüketim modeli iki ana bölümden oluşmaktadır: hareket/itki bölümü ve iletişim bölümü. Hareket/itki gücü ise üç kısımdan oluşmaktadır: indüklenen güç, profil gücü ve parazit güç. İndüklenen güç, itilen havanın aşağıya doğru itilmesiyle sonuçlanır. Profil gücü, pervanenin dönen kanatlarının maruz kaldığı dönme sürtünmesinin üstesinden gelir. Parazit güç de, İHA ile rüzgar arasında göreceli bir öteleme olduğunda gövde sürüklenmesini önler. Önceki açıklamaya dayanarak, İHA'nın hareket / itme enerjisi, hareketler sırasında yerçekimi ve sürüklenmenin üstesinden gelmek için itme sağlamak için kullanılır. Bir İHA'nın uçuş hareketleri yalnızca yatay hareket, havada asılı kalma ve dikey hareket olabilir. Güç tüketim modelini; faydalı yükler, İHA'nın ağırlığı ve uçuş süresi gibi bazı faktörler etkilemektedir. Öte yandan, iletişim için harcanan enerji; sinyal yayılımı/alımı, sinyal işleme ve iletişim devresinden kaynaklanmaktadır. Bu tüketilen enerji genellikle uçuş enerjisine kıyasla daha küçüktür (Lyu ve ark., 2016). Bu nedenle, bu tezde

iletişim enerjisi ihmal edilmiştir. Dolayısıyla, yaklaşık hareket/itiş gücü matematiksel olarak aşağıdaki gibi ifade edilebilir (Alzenad ve ark., 2017; Qin ve ark., 2021):

$$P_u(v_u) = P_0\left(1 + \frac{3v_u^2}{U_{tip}}\right) + P_1\sqrt{\left(\sqrt{1 + \frac{v_u^4}{4v_0^4}} - \frac{v_u^2}{2v_0^2}\right)} + \frac{d_0 \rho s_0 A v_u^3}{2} \quad (4.10)$$

Burada v_u İHA'nın uçuş hızı, U_{tip} İHA üzerindeki rotor kanadının uç hızı, d_0 her bir rotor için gövde sürükleme oranı, ρ hava yoğunluğu, s_0 rotor katılığı, A her bir rotor için disk alanı, v_0 havada asılı kalma modunda rotor kaynaklı ortalama hız ve P_0 ve P_1 sırasıyla kanat profili gücü ve türetilmiş güçtür. Dolayısıyla, anlık enerji tüketimi aşağıdaki gibi yazılabilir (Qin ve ark., 2021):

$$E_{\text{cons}} = P_u(v_u)T_u + P_u(0)(\Delta T - T_u) \quad (4.11)$$

Burada T_u ve İHA için uçuş süresini temsil eder. Ayrıca, ilgili zaman dilimindeki artık enerji (E_{res}), tüketim enerjisi (E_{cons}) ve batarya kapasitesi ($E_{\max}$) açısından aşağıdaki gibi tanımlanabilir:

$$E_{\text{res}} = E_{max} - E_{\text{cons}} \quad (4.12)$$

Bu tez çalışmasında gerçekleştirilen simülasyonlarda, tüm algoritmalar için aynı süre boyunca tüketilen enerji hesaplanmıştır.

4.1. 4 Acil Durum Müdahalesi

Dinamik ve görev açısından kritik ortamlarda, İHA'nın doğal afetler, acil kullanıcı ihtiyaçları veya beklenmedik ağ kesintileri gibi ani acil durum senaryolarına otonom olarak yanıt verme yeteneği çok önemlidir. Bunu ele almak için, önerilen RL tabanlı İHA kontrol sistemi, bir acil durum müdahale mekanizmasını doğrudan ödül yapısına entegre ederek İHA'nın yüksek öneme sahip bölgelere dinamik olarak öncelik vermesini sağlar.

Önerilen İHA sistemi, dinamik olarak tetiklenen acil durum bayraklarına tepki veren bir acil durum müdahale mekanizmasını entegre etmektedir. Bu bayraklar, çevredeki belirli bir alan afetten etkilenmiş olarak işaretlendiğinde üretilir ve acil İHA müdahalesi gerektirir.

Herhangi bir zaman adımında, bir afet alanı k konumuyla ilişkili ikili bir bayrak sinyali $ER_k \in \{0,1\}$ ile ifade edilir. Bu sinyalde, ER_k=1 değeri, k alanında aktif bir acil durumun varlığını gösterirken; ER_k=0 değeri, o alanda acil bir durum olmadığını belirtir.

İHA bu bayrağı tespit etmeli ve işaretli alandaki kullanıcılara hizmet vermek için hareketini yeniden önceliklendirmelidir. Acil durum bölgeleri geçicidir ve öngörülemeyen gerçek dünya olaylarını simüle ederek rastgele yerlerde görünebilir.

4.2 Optimizasyon Problemi

Değer tabanlı RL yaklaşımının optimizasyon amacı, *s* durumunda *a* eylemini gerçekleştirerek ve sonrasında en uygun politikayı izleyerek elde edilebilecek maksimum beklenen kümülatif ödülü temsil eden $Q^*(s,a)$, optimum eylem-değer fonksiyonunu öğrenmek için İHA'yı eğitmektir. $Q^*(s,a)$ şu şekilde ifade edilebilir:

$$Q^*(s,a) = \max_{\pi} \mathbb{E}[\sum_{t=0}^{T} \gamma^t r_t \mid s_0 = s, a_0 = a, \pi] \quad (4.13)$$

Burada $\gamma \in (0,1)$ azaltma faktörüdür ve r_t zaman adımındaki ödüldür ve temel performans göstergelerinin (kapsama alanı, güç tüketimi, adalet ve acil durum yanıtı) ağırlıklı bir toplamı olarak hesaplanır. İHA, gerçek dünyadaki operasyonel kısıtlamaları yerine getirirken bu beklenen getiriyi en üst düzeye çıkaran eylemleri seçmeyi amaçlamaktadır. Burada mevcut kısıtlar şu şekilde verilebilir:

$$\sum_{t=1}^{T} E_{\text{cons}}(t) \leq E_{max} \quad (4.14)$$

$$C_k \geq C_{\min} \quad (4.15)$$

$$\mathrm{F}_k \geq \mathrm{F}_{\min} \quad (4.16)$$

$$\mathrm{ER} \geq D_{\min} \quad (4.17)$$

ER, İHA tarafından hizmet verilen kritik kullanıcıların oranı olarak tanımlanan acil durum müdahale puanıdır ve $D_{\min}$ kabul edilebilir minimum eşiktir. İHA, bu eşiğin operasyon boyunca tutarlı bir şekilde karşılanmasını sağlamalıdır.

Bu değer tabanlı optimizasyon, sinir ağları kullanarak Q-fonksiyonuna yaklaşan DRL algoritmaları (DQN, DDQN, Düello DQN) kullanılarak çözülür. Öğrenilen politika, İHA'nın

dinamik 3B ortamlarda verimlilik, adalet ve yanıt verebilirliği dengeleyen akıllı hareket kararları almasını sağlar.

4.3 Problemin Çözümünde Kullanılan Algoritmalar

Bu bölümde, geliştirilen sistemin çözüm sürecinde kullanılan RL algoritmaları ve çalışma adımları detaylı olarak açıklanmıştır.

4.3.1 Durum Uzayı, Eylem Uzayı ve Ödül Fonksiyonunun Tanımı

Durum uzayı, sistemin herhangi bir andaki durumunu tanımlamak için kullanılır. Bu çalışma için durum uzayı, İHA'nın üç boyutlu uzaydaki konumunu (İHA'nın üç boyutlu koordinatları), kullanıcıların konumlarını (burda $\mathcal{W}$ kullanıcı pozisyonları ve $\mathcal{P}$ ise kapsama durumunu temsil edilir) ve E enerji durumunu içermektedir. Bu nedenle, durum uzayı şu şekilde tanımlanabilir:

$$s = [x_{\text{İ}HA}, y_{\text{İ}HA}, z_{\text{İ}HA}, \mathcal{W}, \mathcal{P}, \text{ER}, E] \tag{4.18}$$

Eylem uzayı ise, bir İHA'nın gerçekleştirebileceği bir dizi eylemi tanımlar. Bu çalışmada, İHA'nın yedi olası eylemi vardır: yukarı (a_0), aşağı (a_1), ileri (a_2), geri (a_3), sola (a_4), sağa (a_5) ve çapraz (a_6). Her eylem, İHA'nın üç boyutlu uzaydaki konumunda bir değişikliğe karşılık gelir. Belirli bu eylemler şu şekilde tanımlanır:

$$A = \{a_0, a_1, a_2, a_3, a_4, a_5, a_6\} \tag{4.19}$$

Ödül fonksiyonları, İHA'nın belirli konumlardaki performansını değerlendirmek için kullanılır ve birden fazla dengeleme hedefi olabilir. Bu doğrultuda tez çalışmasının temel hedefleri; kapsanan kullanıcı sayısını maksimize etmek, enerji tüketimini minimize etmek, enerji seviyeleri arasında adil kapsama sağlamak ve acil durum kullanıcılarına hızlı yanıt vermek olarak belirlenmiştir. Bu dört faktörü birleştiren ödül fonksiyonu şu şekilde tanımlanmıştır:

$$r = \alpha_1 \times C_k - \alpha_2 \times E_{cons} + \alpha_3 \times \text{F}_k + \alpha_4 \times ER_k \tag{4.20}$$

DQN'de Q-değerlerini tahmin etmek için tek bir sinir ağı kullanılır. İstikrar, bir deneyim tekrar tamponu ve sabit bir hedef ağ kullanılarak sağlanır. DDQN'de ise Q-değeri

güncellemelerinde maksimum Q-değerinin aşırı tahminini azaltmak için iki farklı ağ (çevrimiçi ve hedef) kullanılır. Bu, daha doğru politika öğrenimi ile sonuçlanır. Son olarak Düello DQN, Q-değerini avantaj ve değer fonksiyonlarına ayrıştırarak durumların ne kadar değerli olduğunu daha doğru bir şekilde öğrenmeyi amaçlamıştır. Bu yaklaşım, özellikle durağan ortamlarda daha hızlı yakınsama sağlamıştır.

Her algoritma, sabit hiperparametrelerle 600 döngü (episode) boyunca eğitilmiştir. Her döngü 100 zaman adımı içermekte ve her adımda İHA, güncel duruma göre en uygun hareketi gerçekleştirmektedir. Eğitim sırasında elde edilen performans metrikleri, kapsama oranı, ortalama enerji tüketimi, adalet skoru ve acil müdahale süresi üzerinden değerlendirilmiştir.

DQN, DDQN ve Düello DQN algoritmaları için güncelleme formülleri daha önce sırasıyla Eşitlik (3.1), (3.3) ve (3.4)'te belirtilmiştir ve sözde kodları aşağıda gösterilmiştir:

Algoritma 1: DQN Sözde Kod

Q-ağı başlat ve ağırlıkları θ olarak ayarla
Hedef Q-ağı başlat ve ağırlıkları $\theta^{-} = \theta$ olarak ata
Deneyim belleği D'yi başlat
Her bir döngü için (1'den M'ye):
  Ortam başlat ve ilk durumu s al
  Her bir zaman adımı için (1'den T'ye):
    ε olasılığıyla rastgele bir eylem a seç
    Aksi halde, $a = \text{argmax}_a Q(s, a; \theta)$
    Eylem a ortamda uygula, ödül r ve yeni durum s′ al
    Geçiş (s,a,r,s') belleğe kaydet
    Bellekten rastgele bir minibatch (s,a,r,s') örnekle
    Hedef hesapla:
      $y = r + \gamma * \max_{a'} Q\_target(s', a'; \theta^{-})$
    Belirli adımlarda hedef ağı güncelle:
    $\theta^{-} \leftarrow \theta$
    $s \leftarrow s'$
  Zaman adımı bittiğinde döngüye devam et
Döngüler tamamlandığında işlemi bitir

Algoritma 2: DDQN Sözde Kod

Q-ağı başlat ve ağırlıkları θ olarak ayarla

Hedef Q-ağı başlat ve ağırlıkları $\theta^- = \theta$ olarak atanır

Deneyim belleği D'yi başlat

Ayarla: parametreler: γ ,ε ,ε _ min ,ε _ *decay* ,α
Her bir döngü için (1'den M'ye):
Ortam başlat ve ilk durum s al
Her bir zaman adımı için (1'den T'ye):
ε olasılığıyla rastgele bir eylem a seç
Aksi halde, $a = \text{argmax}_a Q(s, a; \theta)$
Eylem a ortamda uygula, ödül r ve yeni durum s′ al
Geçiş (*s*,*a*,*r*, *s*′) belleğe kaydet
Bellekten rastgele bir minibatch (*s*,*a*,*r*, *s*′) örnekle
Hedef hesapla:
$a^* = \text{argmax}_a Q(s', a; \theta)$
$y = r + \gamma * Q_{target}(s', a^*; \theta^-)$

Hedef ağ Q′ ağırlıklarını Q ile eşleşecek şekilde periyodik olarak güncelle

Belirli adımlarda hedef ağı güncelle: $\theta^- \leftarrow \theta$

$s \leftarrow s'$

Zaman adımı bittiğinde döngüye devam et

Döngüler tamamlandığında işlemi bitir

Algoritma 3: Düello DQN Sözde Kod

İki çıkış koluna sahip bir ağ tanımla:

Değer fonksiyonu çıkışı: V(s)

Avantaj fonksiyonu çıkışı: A(s, a)

Birleştirilmiş çıkış: $Q(s,a) = V(s) + [A(s,a) - \text{ort}_a A(s,a)]$

Düello ağı θ ile başlat

Hedef Düello ağı θ^- ile başlat

Deneyim belleği D'yi başlat

Her bir döngü için (1'den M'ye):

Ortam başlat ve ilk durumu s al

Her bir zaman adımı için (1'den T'ye):

ε olasılığıyla rastgele bir eylem a seç

Aksi halde, $a = \text{argmax}_a Q(s, a; \theta)$

Eylem a ortamda uygula, ödül r ve yeni durum s′ al

Geçiş (s,a,r, s') belleğe kaydet

Bellekten rastgele bir minibatch (s,a,r, s') örnekle

Hedef hesaplanır (DQN ya da DDQN tarzında):

$a^* = \text{argmax}_a Q(s_t, a; \theta)$

Q-değeri (3.4) numaralı denklemden hesaplanır

Kayıp fonksiyonu ile öğrenme yap (isteğe bağlı):

$L(\theta) = (y_j - Q(s_j, a_j; \theta))^2$

Belirli adımlarda hedef ağı güncelle: $\theta^- \leftarrow \theta$

$s \leftarrow s'$

Zaman adımı bittiğinde döngüye devam et

Döngüler tamamlandığında işlemi bitir

4.4 Simülasyon Sonuçları

Bu bölümde her algoritmanın (DQN, DDQN ve Düello DQN) performansı sunulmuş ve aralarında bir karşılaştırma yapılmıştır. Çizelge 4.1'de, sistemin ve eğitim sürecinin kapsamını ve işlevselliğini tanımlayan temel parametreler tanıtılmaktadır. Çizelge 4.2'de, simülasyonun gerçekleştirilmesi sırasında kullanılan temel parametreler ve bu parametrelerin değerleri verilmiştir. Bu ayarlar; kullanıcı sayısı, alan boyutu, zaman aralığı ve yükseklik açısı gibi önemli bileşenleri içermektedir.

Çizelge 4.1. Temel Parametreler

Parametre	Değer
Öğrenme Oranı (α)	0.001
İndirim Faktörü (γ)	0.95
Parti (Batch) Boyutu	64
Bellek Boyutu	10,000
Epsilon Bozulması (ϵ_{decay})	0.995
Minimum Epsilon (ϵ_{min})	0.01
Güncelleme Frekansı	Her 10 adımda

Çizelge 4.2. İHA ağı için simülasyon ayarları

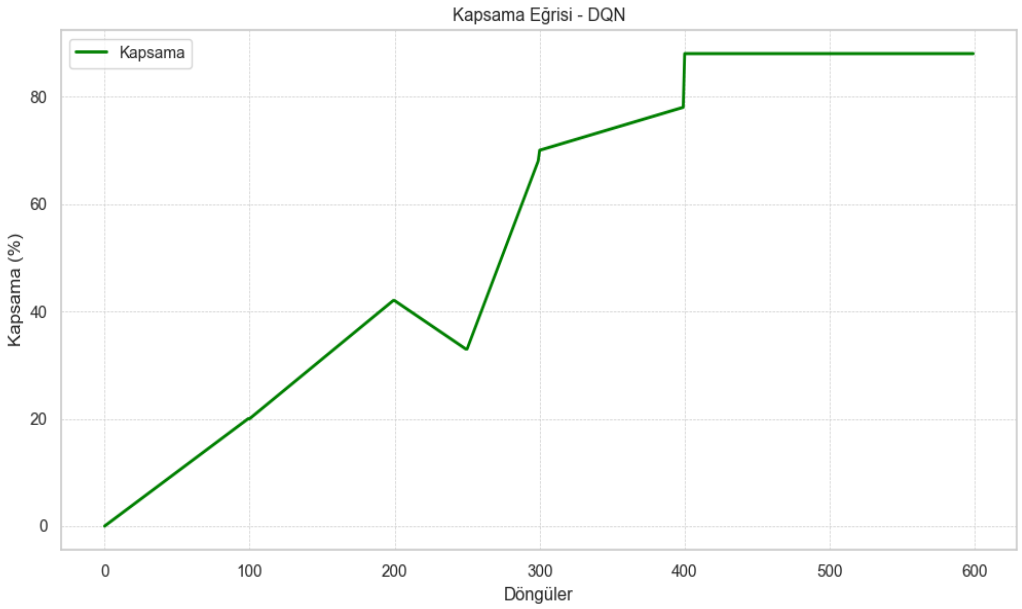
Parametre	Simge	Değer
Kullanıcı sayısı	K	100
Transmisyon gücü	P_u	32 dBm
Alan Boyutu	-	2km x 2km
Zaman aralığı	ΔT	1 s
Süre/tekrarlamalar	T	600
İHA hızı	v_u	10 m/s
Yol kaybı üssü	n_0	2.5
Sistem taşıyıcı frekansını	f_c	2 GHz
Işık hızıdır	c	0.3 Gm/s
Hava yoğunluğu	ρ	1.225 kg/m³
Disc rotoru alanı	A	0.5 m²
LoS zayıflaması	χ_{LoS}	1 dB
NLoS zayıflaması	χ_{NLoS}	20 dB
Çevresel sabitler	δ , B	0.11, 0.6
Yükseklik açısı	θ	45
Uç hızı	U_{tip}	120 m/s
Ortalama rotor kaynaklı hız	v_0	0.002 m/s
Gövde sürükleme oranı	d_0	0.48
Rotor sağlamlığı	s_0	0.0001
Kanat profili gücü	P_0	99.66 W
Türetilmiş güç	P_1	120.16 W

4.4.1 Derin Q-Ağı (DQN) Sonuçları

DQN, her bir durum-eylem çifti için Q değerlerini yaklaşık olarak hesaplamak için bir sinir ağı kullanır. Maksimum operatörü aşırı tahmin önyargısına yol açabilir, bu nedenle kararsız ve gürültülü eğitim yapar.

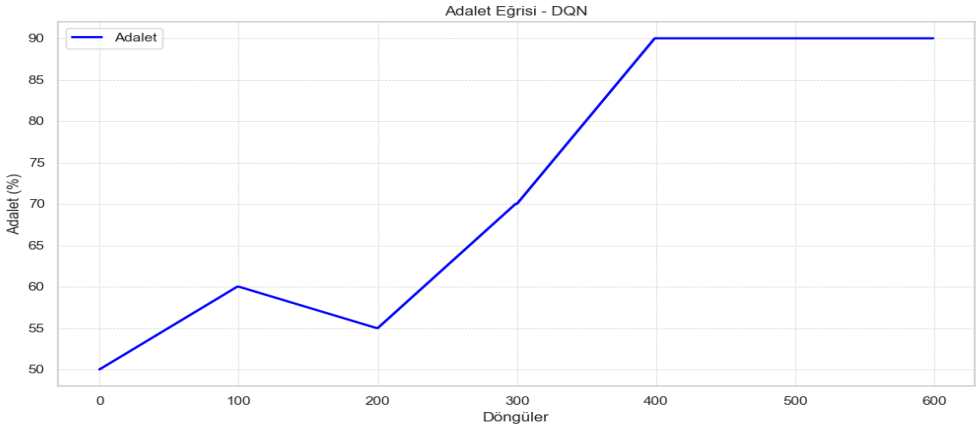
Şekil 4.1, DQN modeli tarafından 600 eğitim döngüsü boyunca elde edilen kapsama yüzdesinin gelişimini göstermektedir; x ekseni döngü sayısını ve y ekseni %0'dan %100'e kadar kapsamayı göstermektedir. Başlangıçta %0 civarında başlayan kapsama oranı, önemli sıçramalarla belirginleşen istikrarlı bir artış sergilemekte ve nihayetinde 400 döngüde yaklaşık %88'de sabitlenmektedir. Şekil 4.1, DQN'nin görev ortamını etkili bir şekilde

keşfettiğini ve kapsamlı bir şekilde kapsamayı öğrendiğini ve bundan sonra minimum düzeyde daha fazla gelişme gösterdiğini göstermektedir. Gözlemlenen eğilim, modelin sağlam keşif stratejisini vurgulamakta ve eğitimin son aşamalarında optimuma yakın bir politikaya yakınsadığını göstermekte ve bu da onu kapsamlı çevresel kapsama gerektiren görevler için umut verici bir yaklaşım haline getirmektedir.



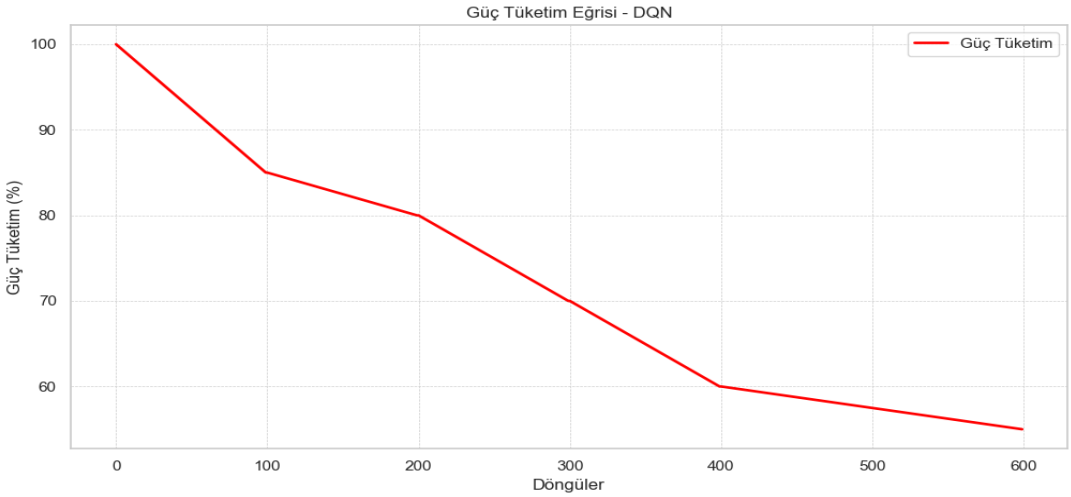
Şekil 4.1. DQN kapsama Eğrisi

Şekil 4.2, 600 döngü boyunca DQN modelinin adalet yüzdesini göstermektedir. Adillik yaklaşık %50 ile başlamakta ve 300 döngüde yaklaşık %70'e çıkmakta ve 400 döngü sonrasında daha hızlı bir artışla %90'a yakın bir seviyede sabitlenmektedir. Bu örüntü, modelin karar verme sürecindeki ilk istikrarsızlık aşamasını ve ardından adil sonuçlara öncelik vermek için başarılı bir adaptasyonu yansıtmaktadır. Eğitimin sonunda ulaşılan yüksek ve istikrarlı adalet seviyesi, DQN'nin diğer hedeflerle birlikte adaleti dengeleme kabiliyetinin altını çizmekte ve onu adil kaynak tahsisi veya karar vermenin kritik olduğu uygulamalar için uygun hale getirmektedir.



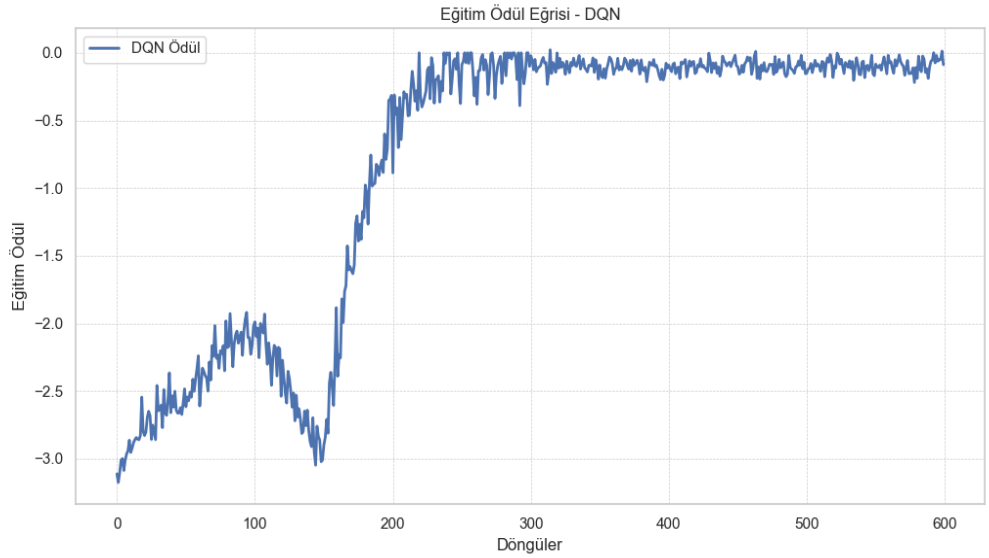
Şekil 4.2 DQN Adalet Eğrisi.

Şekil 4.3'te görülen güç tüketim eğrisi, DQN modelinin güç tüketim yüzdesini gösterir. Bu şekle göre, DQN ile güç tüketimi kademeli olarak azalmakta, 200 döngü civarında %80'e kadar düşmekte ve 400 döngüde yaklaşık %60'a düşerek istikrarlı bir azalma eğilimi göstermektedir. Güç kullanımındaki bu tutarlı azalma, modelin eğitim ilerledikçe enerji verimliliğini optimize etme yeteneğini gösterir ve sonunda %54'lık bir düşüşe ulaşır. Düşüş eğilimi, DQN'nin zaman içinde daha kaynak verimli bir politika öğrendiğini ve bu sayede sürdürülebilirliğin öncelikli olduğu enerji kısıtlı ortamlar veya uygulamalar için cazip bir seçenek haline geldiğini göstermektedir.



Şekil 4.3 DQN Güç Eğrisi.

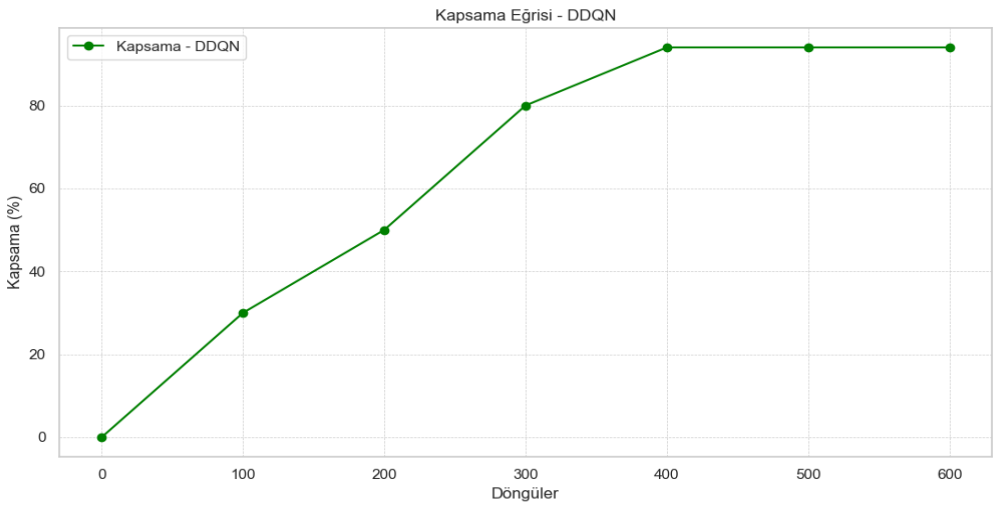
Şekil 4.4'te verilen eğitim ödül eğrisi, DQN modelinin 600 döngü boyunca ödül ilerlemesini göstermektedir. Ödül yaklaşık olarak -3.0 ile başlar ve model farklı eylemleri keşfettikçe ilk 200 döngüde -3 ile -0.5 arasında dalgalanarak yüksek değişkenlik gösterir. 300 döngüden itibaren küçük dalgalanmalarla 0 civarında sabitlenir. Bu şekilde verilen eğri, ilk keşif aşamasını, etkili öğrenmeyi ve keşif ile sömürüyü dengeleyen optimal bir politikaya yakınsamayı yansıtmaktadır. Ancak eğri, keşif-sömürü ödünleşimi nedeniyle gürültülüdür.



Şekil 4.4 DQN Eğitim Ödül Eğrisi.

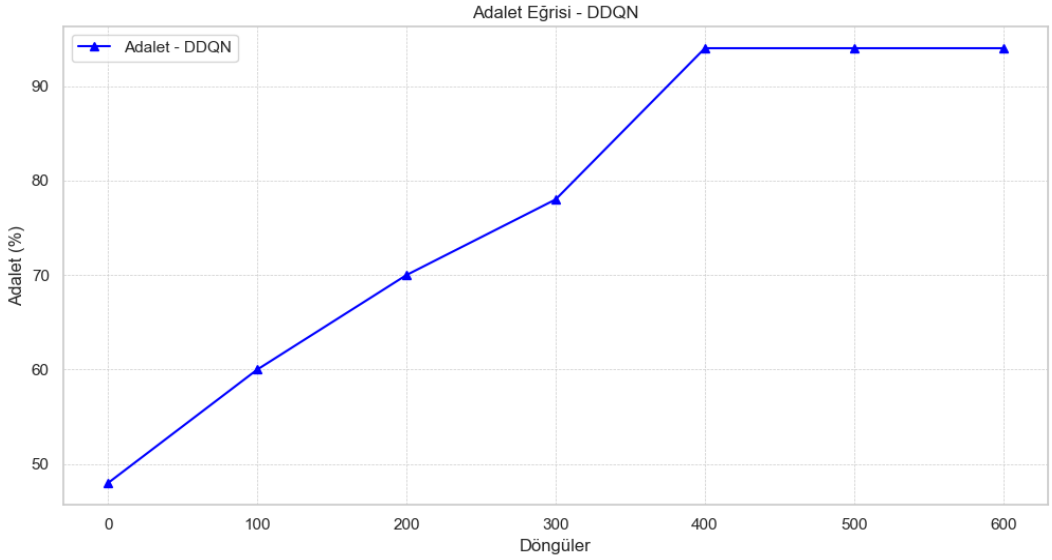
4.4.2 Çift Derin Q-Ağı (DDQN) Sonuçları

DDQN modeli tarafından 600 döngü boyunca elde edilen kapsama yüzdesi Şekil 4.5 ile verilmiştir; kapsama (yüzde olarak) y ekseninde ve döngü sayısı x ekseninde çizilmiştir. Kapsam, modelin başlangıçta durum uzayının yalnızca küçük bir bölümünü keşfettiğini gösterecek şekilde başlar ve istikrarlı bir şekilde artarak 300. döngüde yaklaşık %80'e ve 400. döngüde %93'e ulaşır. Döngü 400'den 600'e kadar %93 civarında sabit kalır. Bu tutarlı artış, modelin zaman içinde daha geniş bir durum veya eylem yelpazesini keşfetmeyi etkin bir şekilde öğrendiğini göstermektedir. Böylece Şekil 4.5, modelin verilen ortamda maksimum kapsama alanına ulaştığını, kapsamlı ve başarılı bir keşif stratejisi ortaya koyduğunu vurgular.



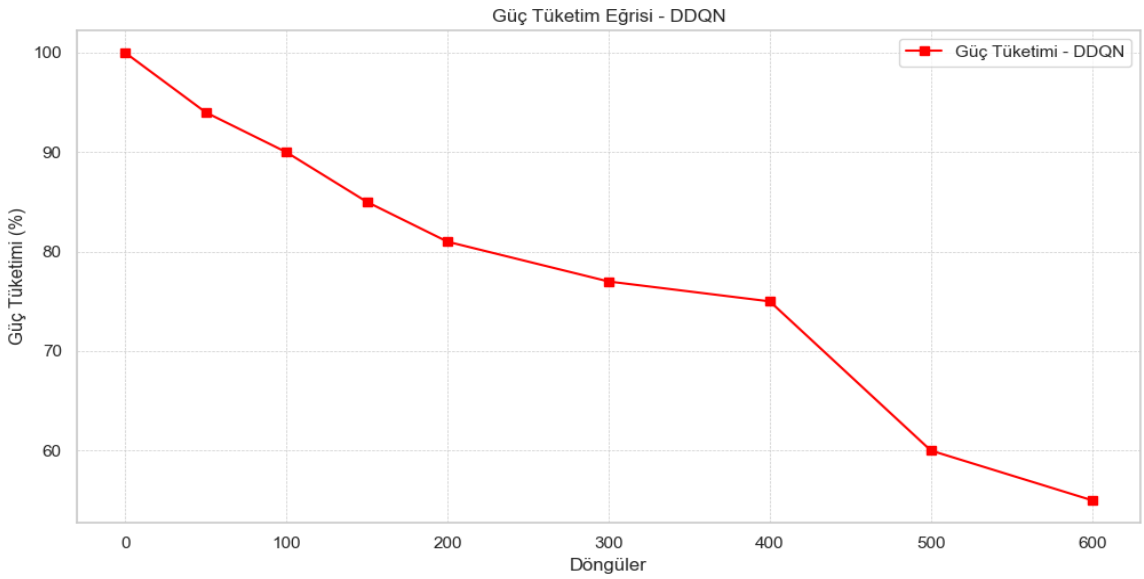
Şekil 4.5 DDQN Kapsama Eğrisi.

Şekil 4.6, DDQN'nin 600 döngü boyunca elde edilen adalet yüzdesini göstermektedir. Başlangıçta adil karar verme eksikliğine işaret eden adalet, istikrarlı bir artış göstererek 400. döngüde %95'in üstüne çıkmaktadır. Bu noktadan sonra, adalet stabilize olur ve tutarlı bir şekilde değerini korur. Başlangıçtaki bu düşük adalet, modelin adil stratejileri öğrenmek için zamana ihtiyaç duyduğunu gösterirken, etkili eğitim ve parametre ayarlaması sayesinde önemli bir iyileşmeyi vurgulamaktadır.



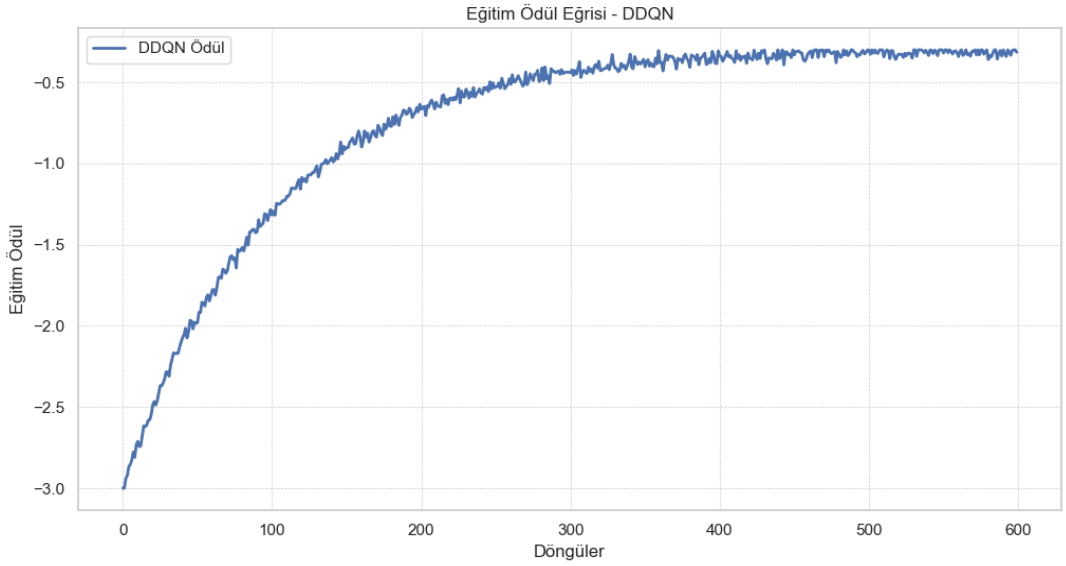
Şekil 4.6 DDQN Adalet Eğrisi.

DDQN'nin Şekil 4.7 ile verilen güç tüketimi eğrisi, DDQN modelinin 600 döngü boyunca güç tüketimi yüzdesini temsil eder. Güç tüketimi, yüksek başlangıç enerji kullanımını yansıtacak şekilde %100'den başlar ve kademeli olarak azalarak 600. döngüde yaklaşık %55'e ulaşır. Bu düşüş eğilimi, modelin eğitim ilerledikçe, enerji verimliliğini optimize ettiğini göstermektedir. İstikrarlı düşüş, performansı korurken güç kullanımını en aza indirmek için başarılı bir uyarlama yapıldığını göstermektedir ki bu, enerji verimliliğinin önemli bir husus olduğu gerçek dünya uygulamaları için çok önemlidir.



Şekil 4.7 DDQN Güç Eğrisi.

Şekil 4.8 ile verilen eğitim ödül eğrisi, DDQN modelinin 600 döngü üzerindeki eğitim ödülünü göstermektedir; y ekseni eğitim ödülünü (negatif değerler) ve x ekseni döngü sayısını göstermektedir. Ödül, erken keşif ve optimal olmayan eylemleri yansıtan -3.0 civarında başlar ve ilk 100 döngü içinde hızla artar. Bu iyileşme, artan ödülün de gösterdiği gibi modelin etkili stratejiler öğrendiğini göstermektedir. 300. Döngüden sonrası, modelin optimale yakın bir politikaya ulaştığını gösterir, ancak bazı değişkenlikler devam eden keşiflere veya eğitim sürecini etkileyen çevresel gürültüye işaret edebilir.



Şekil 4.8 DDQN Eğitim Ödül Eğrisi.

4.4.3 Düello Derin Q-Ağı (Dueling DQN) Sonuçları

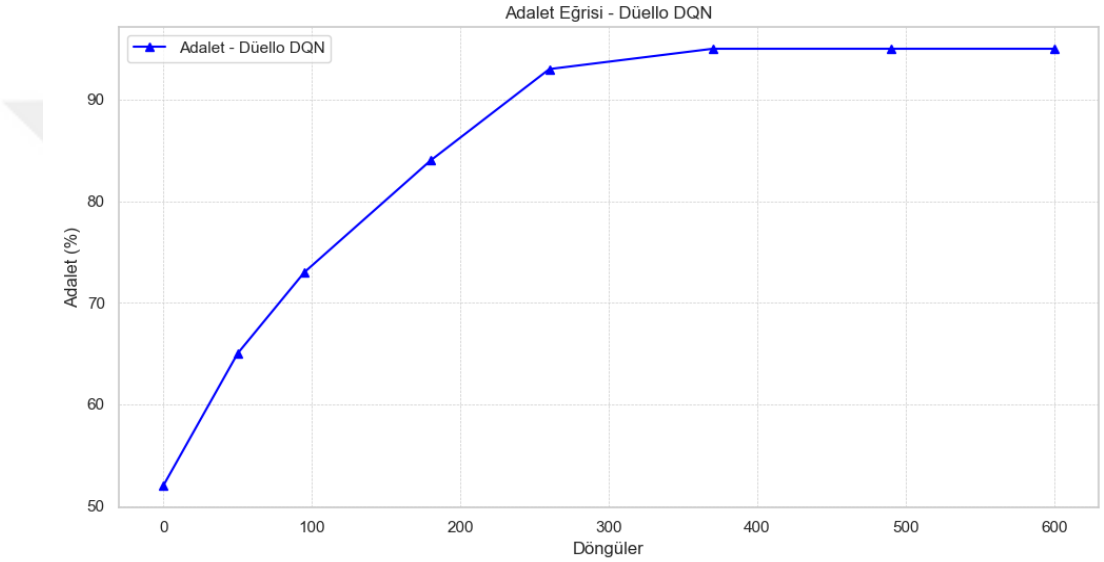
Düello DQN ağa iki akış getirir: değer akışı V(s) ve avantaj akışı A(s,a). Bu, eylemler biraz farklı olduğunda ajanın durum değerini daha iyi tahmin etmesine yardımcı olur.

Düello DQN modelinin performansı 600 döngü boyunca dört grafikte gösterilmiştir. İlk olarak Şekil 4.9'da adalet eğrisi verilmiş; adalet yüzdesinin %50 civarında başlayıp istikrarlı bir şekilde yükselerek 300 döngüde yaklaşık %95'e ulaştığı gözlenmektedir. Bu, başlangıçta tutarlı bir adalet gelişimi ve ardından istikrar olduğunu göstermektedir.

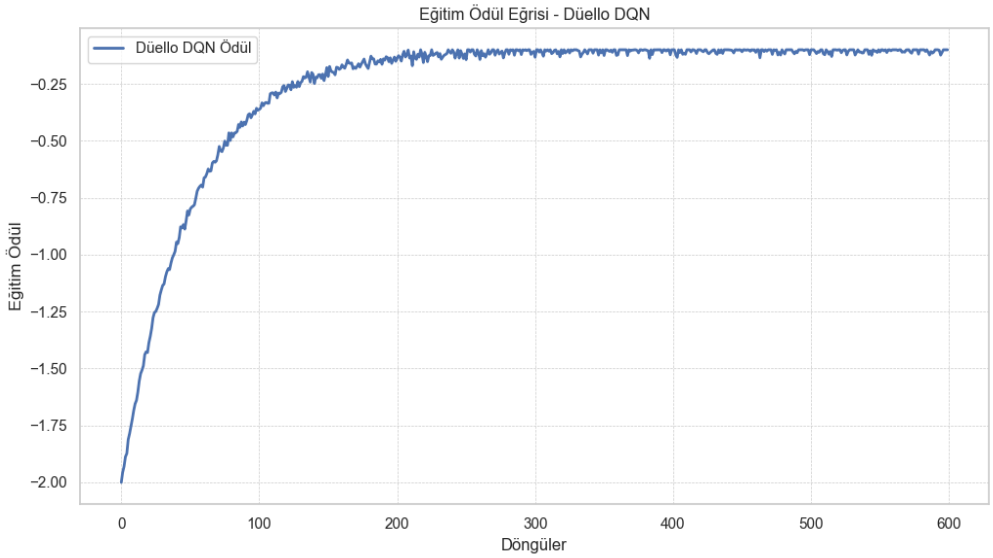
İkinci olarak Şekil 4.10'da eğitim ödülü eğrisi sunulmuştur. Bu şekle göre ilgili ödül -2.0'dan başlamakta ve 200 döngü sonunda -0.1 civarında sabitlenmektedir. Bu da modelin tutarlı bir performans aralığına yerleşmeden önce ödülleri optimize etmeyi hızlı bir şekilde öğrendiğini gösterir.

Şekil 4.11'den; %100 ile başlayan güç tüketiminin 600 döngüde %50'nin altında olan seviyelerine çekilebildiği görülmektedir. Bu eğilim, modelin zaman içinde enerji kullanımını azaltma ve eğitim ilerledikçe önemli ölçüde güç verimliliği elde etme yeteneğini yansıtmaktadır.

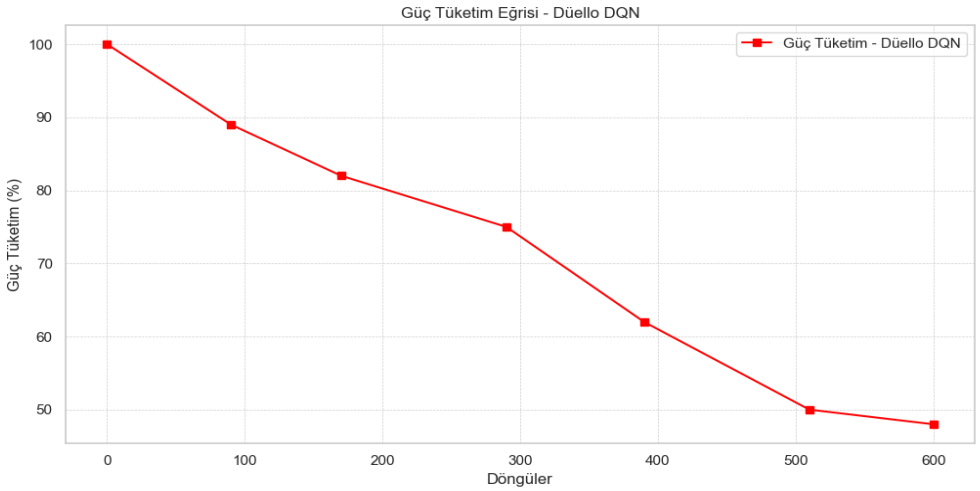
Son olarak Şekil 4.12'de kapsama oranının 200 döngü ile %60'a çıkarılabildiği görülmektedir. Daha sonra kademeli olarak artmaya devam ederek 600 döngüde %98'e çıkmaktadır. Bu da modelin kapsama alanını iyileştirmedeki etkinliğini ve eğitim süresinin sonunda neredeyse tam kapsama alanına ulaştığını göstermektedir. Bu grafikler birlikte Düello DQN'nin adaleti, ödül optimizasyonunu, güç verimliliğini ve kapsamı dengeleme becerisini vurgulamaktadır.



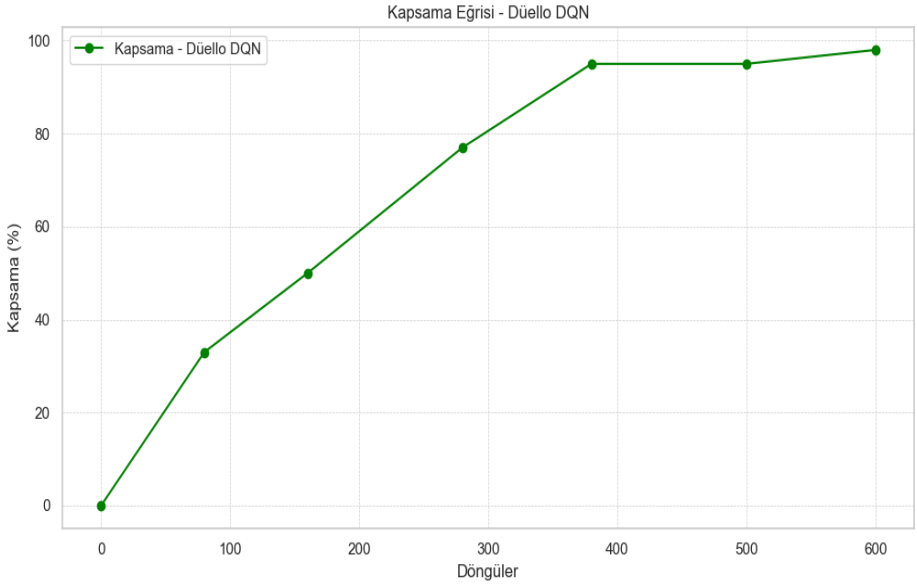
Şekil 4.9 Düello DQN Adalet Eğrisi.



Şekil 4.10. Düello DQN Eğitim ödülü eğrisi.



Şekil 4.11. Düello DQN Güç eğrisi.



Şekil 4.12. Düello DQN Kapsama eğrisi.

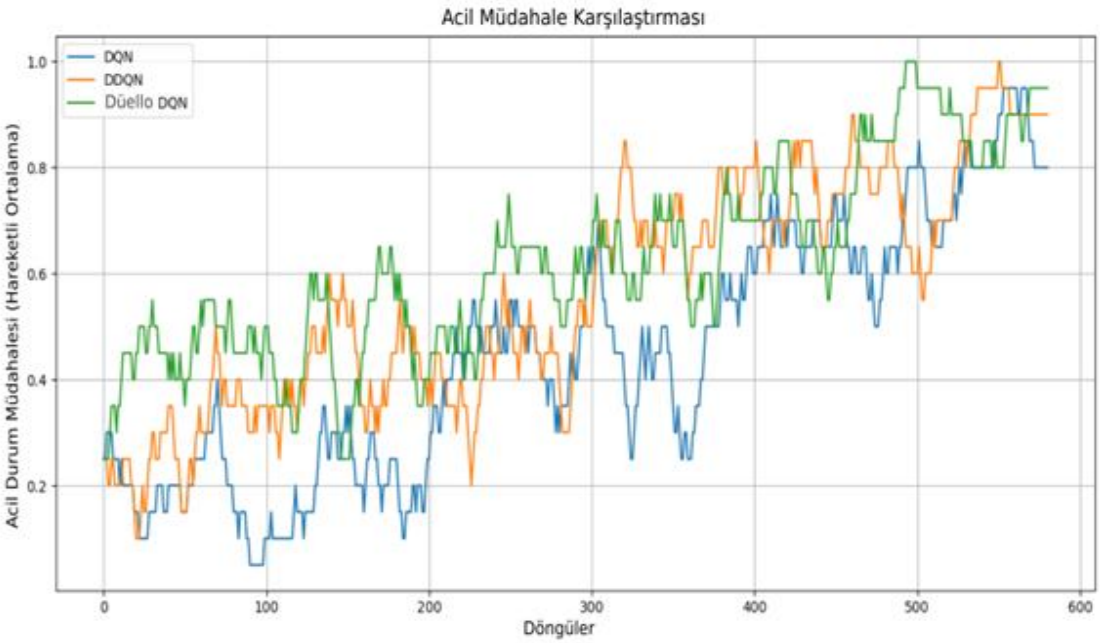
4.4.4 Acil Durum Müdahalesi

Üç DQN modelinin (DQN, DDQN ve Düello DQN) acil durum yanıt performansı 600 döngü boyunca Şekil 4.13 ve 4.14'te karşılaştırılmıştır. İlk olarak şekil 4.13'te her üç model de acil durum yanıtında (hareketli ortalama) yüksek değişkenlik göstermekte ve döngüler boyunca 0.0 ile 1.0 arasında dalgalanmaktadır. DQN, DDQN ve Düello DQN, acil durum yanıt sürelerinde dalgalanmalara sahip olsa da döngü sayısı artıkça müdahalenin

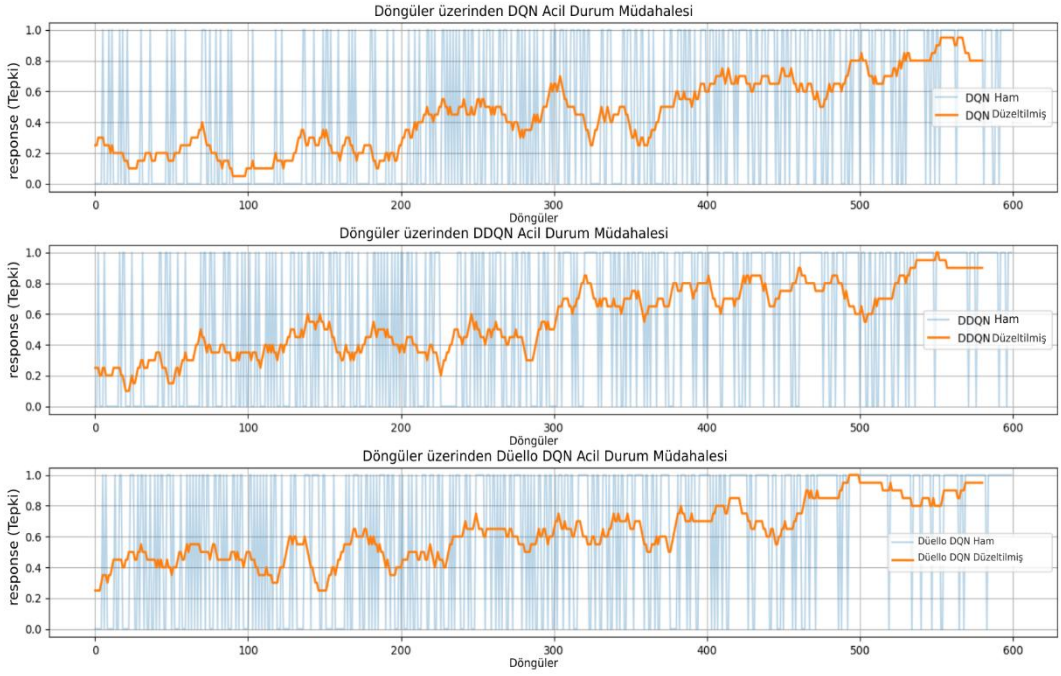
iyileştiği gözlenmiştir. İkinci grafik setinde (Şekil 4.14'te), bu değerlendirmede, sistemin çevrede rastgele ortaya çıkan ani afet alanlarını veya acil durum bölgelerini tespit etmesi ve bunlara yanıt vermesi beklenmiştir. Her bir algoritmanın performansı iki eğri ile gösterilmiştir: ham veriler (raw data) ve düzeltilmiş veriler (smoothed data).

Ham veri eğrisi (mavi), her bir eğitim episodu için elde edilen acil durum yanıt puanlarını doğrudan göstermektedir; bu da İHA'nın o döngüde ne kadar etkili yanıt verdiğini ifade eder. Özellikle eğitim sürecinin erken aşamalarında çevrenin dinamik doğası ve keşif davranışları nedeniyle bu değerlerde önemli dalgalanmalar görülmektedir. Buna karşılık, düzeltilmiş veri eğrisi (turuncu), ham verilerin belirli bir pencere boyutundaki (örneğin, 10 döngü) hareketli ortalamasını temsil eder ve kısa vadeli dalgalanmaları azaltarak genel öğrenme eğiliminin daha net görünmesini sağlar. Grafiklerin dikey ekseni, 0 ile 1 arasında değişen acil durum yanıt puanlarını göstermektedir. Örneğin, 0.5 değeri, İHA'nın o döngüde gerçekleşen acil durum olaylarının %50'sine başarılı bir şekilde yanıt verdiğini ifade eder. Daha yüksek puanlar, sistemin acil durumlara daha duyarlı ve güvenilir şekilde yanıt verdiğini göstermektedir.

Elde edilen sonuçlar, eğitim bölümleri arttıkça tüm algoritmaların acil durum yanıt performanslarında belirgin bir iyileşme gösterdiğini ortaya koymaktadır. DQN algoritması başlangıçta düşük ve düzensiz yanıtlar vermesine rağmen zamanla performansını artırmıştır. DDQN algoritması ise daha istikrarlı bir öğrenme eğrisi sergilemiş, aşırı tahmin hatalarını azaltarak son döngülerde daha yüksek puanlara ulaşmıştır. Düello DQN modeli ise tüm algoritmalar arasında en iyi performansı göstermiş; kararlı bir gelişim sergileyerek eğitim sonunda neredeyse 1.0 değerine yakın puanlar elde etmiştir. Bu, Düello DQN tarafından yönlendirilen İHA sisteminin neredeyse her zaman acil durum bölgelerini doğru bir şekilde tespit ettiğini ve başarılı müdahalede bulunduğunu göstermektedir.



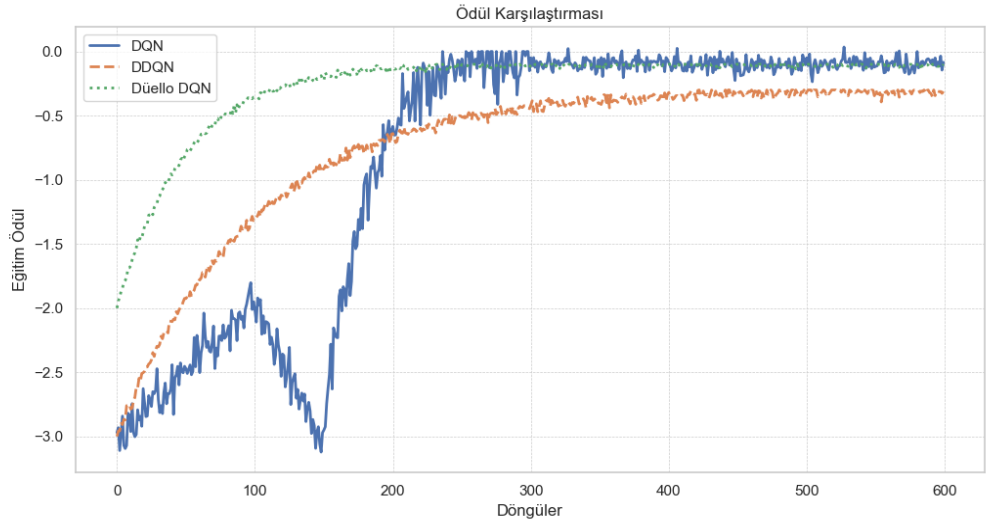
Şekil 4.13. Acil durum müdahalesi eğrisi.



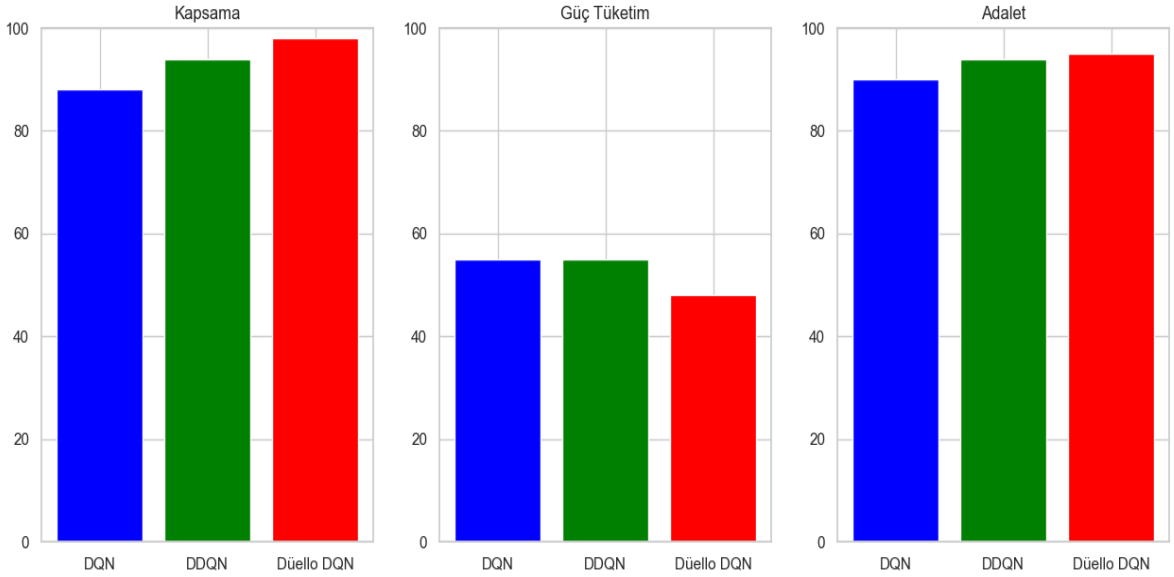
Şekil 4.14. Her modelin acil durum tepkisi.

4.5. Karşılaştırmalı Grafikler

Bu tez çalışmasında yapılan değerlendirmelere dayanarak kullanılan yöntemlerin eğitim ödülü ile kapsama, güç tüketimi ve adalet yüzdeleri Şekil 4.15 ve 4.16'da karşılaştırmalı olarak verilmiştir. Elde edilen sonuçlara göre Düello DQN dört temel faktörün tamamında en dengeli ve üstün performansı göstermiştir. Kapsama açısından, gelişmiş durum-değer tahmini sayesinde sürekli olarak daha geniş ve daha istikrarlı bir alan kapsaması elde etmiş ve hem DQN hem de DDQN'den daha iyi performans göstermiştir. Adalet konusunda, Düello DQN İHA hizmetini kullanıcılar arasında daha adil bir şekilde dağıtarak belirli bölgelere veya kullanıcı gruplarına yönelik hizmet önyargısını azaltırken, DDQN orta düzeyde adalet iyileştirmeleri sunmuş ve DQN en az adaleti göstermiştir. Afet müdahalesi için, Düello DQN, değer ve avantaj işlevleri arasındaki daha iyi politika ayrımının bir sonucu olarak ani acil durumlara yanıt olarak daha hızlı ve daha doğru İHA yeniden konumlandırması sergilerken, DDQN makul derecede iyi yanıt vermiştir. Ancak Düello DQN ile karşılaştırıldığında kritik yanıt süresinde gecikmeler göstermiştir. Son olarak, enerji azaltımında, hem Düello DQN hem de DDQN gereksiz İHA hareketlerinden kaçınarak kayda değer enerji verimliliği elde ederken, Düello DQN daha istikrarlı öğrenme nedeniyle DDQN'den biraz daha iyi performans göstermiş, DQN ise nispeten istikrarsız politika güncellemeleri nedeniyle en fazla enerjiyi tüketmiştir. Genel olarak, Düello DQN, karmaşık ve çok amaçlı senaryolarda dinamik İHA konumlandırması için en etkili yaklaşım olarak ortaya çıkmıştır.



Şekil 4.15. Ödül eğrileri karşılaştırması.



Şekil 4.16. genel karşılaştırma.

5. SONUÇLAR VE ÖNERİLER

5.1 Sonuçlar

Bu çalışma, kablosuz kapsama alanını geliştirmeyi, enerji verimliliğini optimize etmeyi ve dinamik ve öngörülemeyen bir ortamda kullanıcılara hizmet vermede adaleti sağlamayı amaçlayan RL tabanlı bir İHA dinamik konumlandırma sistemi sunmaktadır. Model, 4 kilometrekarelik bir 3B simülasyon alanına tek bir İHA konuşlandırarak, RL tekniklerini kullanarak gerçek zamanlı olarak akıllı hava kapsama alanının fizibilitesini göstermektedir. İHA, rastgele konumlandırılmış ve mobil 100 kadar kullanıcıyı kapsamakla görevlendirilmiştir ve performans ölçümlerini en üst düzeye çıkarmak için konumu hakkında sürekli kararlar vermesini gerektirmektedir.

Akıllı ve uyarlanabilir hareketi mümkün kılmak için İHA üç RL algoritması kullanılarak eğitilir: DQN, DDQN ve Düello DQN. Bu algoritmaların her biri, Q değerlerinin aşırı tahmin edilmesi ve değere dayalı eylem seçiminin iyileştirilmesi gibi zorlukları ele alarak İHA'nın karar verme yeteneğinin geliştirilmesine katkıda bulunur. Bu algoritmaların karşılaştırmalı analizi, yüksek kullanıcı hareketliliği ve yoğunluk değişimleri ile dinamik senaryoların ele alınmasındaki etkinliklerinin değerlendirilmesine olanak tanır.

Sistemin bir diğer özelliği de, İHA'nın ani kullanıcı artışları veya felaketin vurduğu bölgeler gibi kritik durumlara hızla adapte olmasını sağlayan entegre acil müdahale mekanizmasıdır. Bu, İHA'nın yalnızca tutarlı bir kapsama alanı sağlamakla kalmayıp aynı zamanda yüksek öncelikli taleplere anında tepki verebilmesini sağlayarak sistemin görev açısından kritik iletişim ve kamu güvenliği alanlarında gerçek dünyada uygulanabilirliğini artırmaktadır.

Sistem, kapsama yüzdesi, güç tüketimi, adalet endeksi ve acil durum yanıt süresi gibi temel ölçütlere göre değerlendirilmiştir. Sonuçlar, RL'nin, özellikle de DQN'nin gelişmiş versiyonlarının, İHA'nın dinamik bir ortamda otonom olarak gezinme ve kullanıcılara verimli bir şekilde hizmet verme yeteneğini önemli ölçüde geliştirebileceğini göstermektedir. Test edilen üç model arasında Düello DQN; istikrar, kapsama alanı ve adalet açısından üstün sonuçlar sunarak sürekli olarak en iyi performansı elde etmiştir. Bu çalışma, gelecekteki kablosuz ağlarda ve afet müdahale çerçevelerinde daha sofistike İHA dağıtım sistemleri için temel oluşturmaktadır.

5.2 Öneriler

Bu tez çalışmasında, dinamik İHA konumlandırmasında RL algoritmalarının (DQN, DDQN ve Düello DQN) etkinliğini gösterirken, sistem performansını ve gerçek dünyada uygulanabilirliği daha da artırmak için çeşitli uygulamalar araştırılabilir. Umut verici yönlerden biri, çoklu ajan RL (MARL) tekniklerini kullanarak koordinasyon içinde çalışan birden fazla İHA'nın entegrasyonudur. Bu, daha büyük veya daha yoğun nüfuslu alanların kapsanmasında daha iyi ölçeklenebilirlik ve verimlilik sağlamanın yanı sıra acil durumlara yanıt olarak işbirlikçi davranışa olanak tanıyacaktır.

Gelecekteki bir başka potansiyel çalışma da simülasyonu daha gerçekçi hale getirmek için engeller, hava koşulları veya 3B arazi varyasyonları gibi çevresel faktörlerin dahil edilmesidir. Ayrıca, kullanıcı davranışı modellemesi, kalabalık hareketi veya önceliğe dayalı hizmetler gibi modeller eklenerek geliştirilebilir ve İHA'nın bağlama duyarlı stratejiler öğrenmesine olanak sağlanabilir.

Algoritmik bir bakış açısıyla, PPO veya SAC gibi daha gelişmiş DRL teknikleri; kararlılıklarını, yakınsama hızlarını ve dinamik ortamlara uyarlanabilirliklerini karşılaştırmak için araştırılabilir. Ayrıca, batarya limitleri ve şarj politikaları da dahil olmak üzere İHA'nın enerji kısıtlamaları, pratik dağıtım sınırlamalarını yansıtacak şekilde modellenebilir.

Son olarak, gelecekteki araştırmalar, önerilen modelin canlı ağ koşullarındaki performansını ve sağlamlığını doğrulamak için gerçek İHA'lar ve uç cihazları kullanarak döngü içi donanım simülasyonları veya saha deneyleri yoluyla gerçek dünya dağıtım senaryolarını dikkate alabilir.

6. KAYNAKLAR

Alzenad, M., El-Keyi, A., Lagum, F. ve Yanikomeroglu, H., 2017, 3-D placement of an unmanned aerial vehicle base station (UAV-BS) for energy-efficient maximal coverage, IEEE Wireless Communications Letters, 6 (4), 434–437.

Bouhamed, O., Ghazzai, H., Besbes, H. ve Massoud, Y., 2020, Autonomous UAV Navigation: A DDPG-based Deep Reinforcement Learning Approach, IEEE International Symposium on Circuits and Systems (ISCAS), October 2020.

Challita, U. ve Saad, W., 2017, Network formation in the sky: Unmanned aerial vehicles for multi-hop wireless backhauling, GLOBECOM 2017 - IEEE Global Communications Conference, Singapore, 1–6.

Ding, R., Gao, F. ve Shen, X. S., 2020, 3D UAV Trajectory Design and Frequency Band Allocation for Energy-Efficient and Fair Communication: A Deep Reinforcement Learning Approach, IEEE Transactions on Wireless Communications, 19 (12), 7796–7809.

Hoseini, S., Bokani, A., Hassan, J., Salehi, S. ve Kanhere, S., 2021, Energy and Service-priority aware Trajectory Design for UAV-BSs using Double Q-Learning, IEEE, Las Vegas, NV, USA, Mart 2021.

Islam, Md. M., Khan, M. T. R., Saad, M. M., Tariq, M. A. ve Kim, D., 2022, Dynamic positioning of UAVs to improve network coverage in VANETs, Vehicular Communications, 36, 100498.

Jain, R. K., Chiu, D. M. W. ve Hawe, W. R., 1984, A quantitative measure of fairness and discrimination, Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA, 21 (1), 2022–2023.

Johnson, W., 1994, Helicopter theory, Dover Publications, New York.

Lee, I., Babu, V., Caesar, M. ve Nicol, D., 2020, Deep Reinforcement Learning for UAV-Assisted Emergency Response, Association for Computing Machinery, December 2020, 327–336.

Lim, N. H. Z., Lee, Y. L., Tham, M. L., Chang, Y. C., Sim, A. G. H. ve Qin, D., 2021, Coverage Optimization for UAV Base Stations using Simulated Annealing, IEEE 15th Malaysia International Conference on Communication (MICC), Malaysia, Aralık 2021, 43–48.

Liu, C. H., Chen, Z., Tang, J., Xu, J. ve Piao, C., 2018, Energy-Efficient UAV Control for Effective and Fair Communication Coverage: A Deep Reinforcement Learning Approach, IEEE Journal on Selected Areas in Communications, 36 (9), 2059–2070.

Liu, C. H., Ma, X., Gao, X. ve Tang, J., 2019, Distributed Energy-Efficient Multi-UAV Navigation for Long-Term Communication Coverage by Deep Reinforcement Learning, IEEE Transactions on Mobile Computing, 19 (6), 1274–1285.

Liu, Y., Huangfu, W., Zhou, H., Zhang, H., Liu, J. ve Long, K., 2022, Fair and Energy-Efficient Coverage Optimization for UAV Placement Problem in the Cellular Network, IEEE Transactions on Communications, 70 (6), 4222–4235.

Lyu, J., Zeng, Y., Zhang, R. ve Lim, T. J., 2016, Placement optimization of UAV-mounted mobile base stations, IEEE Communications Letters, 21 (3), 604–607.

Mozaffari, M., Saad, W., Bennis, M. ve Debbah, M., 2016, Efficient Deployment of Multiple Unmanned Aerial Vehicles for Optimal Wireless Coverage, IEEE Communications Letters, 20 (8), 1647–1650.

Nemer, I. A., Sheltami, T. R., Belhaiza, S. ve Mahmoud, A. S., 2022, Energy-Efficient UAV Movement Control for Fair Communication Coverage: A Deep Reinforcement Learning Approach, Sensors, 22 (5).

Oliveira, F., Luís, M. ve Sargento, S., 2021, Machine Learning for the Dynamic Positioning of UAVs for Extended Connectivity, Sensors, 21 (13), 4618.

Qin, Z., Liu, Z., Han, G., Lin, C., Guo, L. ve Xie, L., 2021, Distributed UAV-BSs Trajectory Optimization for User-Level Fair Communication Service with Multi-Agent Deep Reinforcement Learning, IEEE Transactions on Vehicular Technology, 70 (12), 12290–12301.

Qin, Z., Liu, Z., Han, G., Lin, C., Guo, L. ve Xie, L., 2021, Distributed UAV-BSs trajectory optimization for user-level fair communication service with multi-agent deep reinforcement learning, IEEE Transactions on Vehicular Technology, 70 (12), 12290–12301.