

ROS-Based Path Planning for Autonomous Underwater Vehicles Using Deep Q-Networks Algorithm

Emre GÖZÜTOK¹ , Fecir DURAN^{2*} 

¹Gazi Üniversitesi, Bilişim Enstitüsü, Bilişim Sistemleri Bölümü, Çankaya/ANKARA

²Gazi Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği Bölümü, Yenimahalle/ANKARA

Makale Bilgisi

Araştırma makalesi
Başvuru: 04/04/2024
Düzeltilme: 08/05/2024
Kabul: 10/05/2024

Anahtar Kelimeler

Pekiştirmeli Öğrenme
Robot İşletim Sistemi
Simülasyon
Su Altı Otonom Araçları
Yapay Zeka

Article Info

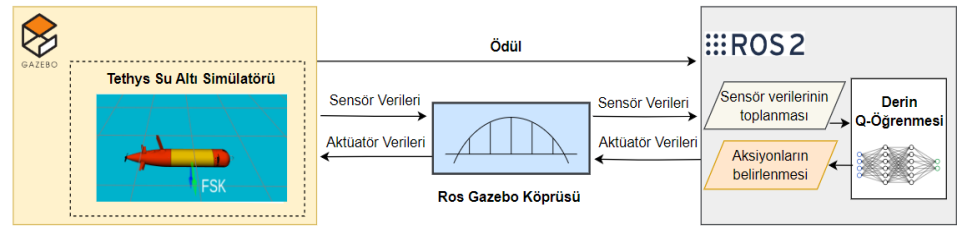
Research article
Received: 04/04/2024
Revision: 08/05/2024
Accepted: 10/05/2024

Keywords

Reinforcement Learning
Robot Operating System
Simulation
Artificial Intelligence
Autonomous Underwater
Vehicles

Grafik Özet (Graphical/Tabular Abstract)

Bu çalışmada otonom su altı araçlarında yol planlamanın otomatize olarak gerçekleştirilebilmesi için ROS tabanlı çevrim içi Derin Q-Öğrenmesi metodu geliştirilmiştir. / A ROS-based online Deep Q-Learning method has been developed to automate path planning in autonomous underwater vehicles.



Şekil A: ROS tabanlı çevrim içi Derin Q-Öğrenmesi ile otonom su altı yol planlama mimarisi / **Figure A:** Autonomous underwater path planning architecture with ROS-based online Deep Q-Learning

Önemli noktalar (Highlights)

- Çevrim içi Derin Q-Öğrenmesi metodu kullanılarak, Uzun Menzilli Otonom Su Altı Araçlarında otonom yol planlaması gerçekleştirilmiştir. / Autonomous path planning in Long-Range Autonomous Underwater Vehicles has been accomplished using the online Deep Q-Learning method.
- Veri akışının sınırlı olduğu durumlar simüle edilerek Derin Q-Öğrenmesi yöntemi ile bu koşullar altında otonom yol planlamanın mümkün olduğu gözlemlenmiştir. / Simulation-based scenarios with limited data flow were simulated, and it was observed that autonomous path planning is possible using the Deep Q-Learning method.
- ROS ve Gazebo entegrasyonu gerçekleştirilerek simülasyon tabanlı bir çevrim içi öğrenme platformu elde edilmiştir. / ROS and Gazebo integration has been achieved, resulting in a simulation-based online learning platform.

Amaç (Aim): Bu çalışmanın amacı, su altı otonom araçları için geleneksel kontrol yöntemleri yerine pekiştirmeli öğrenme tabanlı otonom yol planlama algoritmasının etkinliğini ve potansiyel avantajlarını araştırmaktır. / The aim of this study is to investigate the effectiveness and potential advantages of a reinforcement learning-based autonomous path planning algorithm for underwater autonomous vehicles, as opposed to traditional control methods.

Özgünlük (Originality): Aktif olarak görev yapan bir su altı aracına ait simülasyon üzerinde veri akışının sınırlı olduğu bir senaryo simüle edilerek, bu şartlarda uzun menzilli otonom su altı araçlarında yol planlama için Derin Q-Öğrenmesi metodunun etkin eğitilmesinin mümkün olduğu gözlemlenmiş ve pekiştirmeli öğrenme tabanlı yol planlama algoritmasının performansı ve uygulanabilirliği değerlendirilmiştir. / By simulating a scenario with limited data flow on a simulation of an actively operating vehicle, it was observed that effective training of the Deep Q-Learning method for path planning in long-range autonomous underwater vehicles is possible under these conditions. The performance and applicability of the reinforcement learning-based path planning algorithm were evaluated.

Bulgular (Results): Çalışma Derin Q-Öğrenmesi metodunun otonom su altı araçlarının yol planlamasındaki etkinliğini gösterirken, hiperparametrelerin algoritma performansı üzerindeki kritik etkisini vurgulamaktadır. / The study highlights the effectiveness of Deep Q-Learning in the path planning of underwater autonomous vehicles, emphasizing the critical impact of hyperparameters on algorithm performance.

Sonuç (Conclusion): Sonuç olarak Derin Q-Öğrenmesi yöntemi ile otonom su altı araçlarında yol planlamanın mümkün olduğu görülmüştür. / As a result, it has been observed that path planning in autonomous underwater vehicles is possible with the Deep Q-Learning algorithm.



Su Altı Otonom Araçlarda Derin Q-Ağları Algoritması Kullanılarak ROS Tabanlı Yol Planlama

Emre GÖZÜTOK¹ , Fecir DURAN^{2*}

¹Gazi Üniversitesi, Bilişim Enstitüsü, Bilişim Sistemleri Bölümü, Çankaya/ANKARA

²Gazi Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği Bölümü, Yenimahalle/ANKARA

Makale Bilgisi

Araştırma makalesi
Başvuru: 04/04/2024
Düzeltilme: 08/05/2024
Kabul: 10/05/2024

Anahtar Kelimeler

Elektro Deşarj İşleme
Ti-6Al-4V Alaşım
Kama Yuvası
Ölçüm Doğruluğu

Öz

Su altı araçları genellikle sınırlı hareket kabiliyetine sahiptir. Suyun yoğunluğu, viskozitesi, kaldırma kuvveti ve akıntı gibi çevresel faktörler su altı araçlarının manevra kabiliyetini kısıtlamaktadır. Bu çalışma, bu problemin çözümüne odaklanmaktadır. Geleneksel denetleyicilerin yerine gerçek zamanlı olarak Pekiştirmeli Öğrenme algoritmalarının kullanılması incelenmiştir. Uzun Menzilli Su Altı Otonom Araçları'nın (UMOSA) yörüngesini belirlemek için Derin Q-Ağları algoritması kullanılmıştır. Gazebo simülasyon ortamındaki su altı aracının kontrolü Robot İşletim Sistemi kullanılarak sağlanmıştır. Sonuçlar geleneksel denetleyicilere kıyasla Pekiştirmeli Öğrenme tabanlı algoritmaların potansiyel avantajlarını göstermektedir. Çalışma sonucunda UMOSA modellerinde Derin Q-Ağları algoritmasının gerçek zamanlı kontrol için verimli olarak kullanılabileceği ve simülasyon ortamında Derin Q-Ağları için gereken eğitim ortamının gerçekleştirilebileceği gözlenmiştir.

ROS-Based Path Planning for Autonomous Underwater Vehicles Using Deep Q-Networks Algorithm

Article Info

Research article
Received: 04/04/2024
Revision: 08/05/2024
Accepted: 10/05/2024

Keywords

Electro Discharge Machining
Ti-6Al-4V Alloy
Keyseat
Measurement Accuracy

Abstract

This study focuses on addressing the limited maneuverability typically associated with underwater vehicles. The maneuverability of underwater vehicles is restricted by environmental factors such as the density, viscosity, buoyancy, and currents of water. RL algorithms were examined in real-time, replacing traditional controllers. The Deep Q-Networks algorithm was employed to determine the trajectory of LRAUV. Control of the Gazebo underwater vehicle was facilitated using the Robot Operating System. Results indicate the potential advantages of RL-based algorithms compared to traditional controllers. The study concludes that the Deep Q-Networks algorithm can be efficiently utilized for real-time control of LRAUV models, and the training environment required for Deep Q-Networks can be achieved in the simulation environment.

1. GİRİŞ (INTRODUCTION)

Uzun Menzilli Su Altı Otonom Araçları (UMOSA), denizaltı keşif ve gözlem görevlerinde, deniz çevresinin izlenmesinde, okyanus ekosistemlerinin araştırılmasında ve sismik araştırmalarda kullanılmaktadır. Bu araçlar genellikle insan müdahalesi olmadan uzun süreler boyunca görev yapabilme yeteneğine sahiptirler. Ancak, bu tür araçların kontrolü, çevresel faktörlerin yanı sıra yerine getirilmesi gereken görev gereksinimlerini de dikkate alarak oldukça karmaşık hale gelebilmektedir [1].

Phillips ve diğerleri (2023), "Autosub Long Range 1500: A continuous 2000 km field trial" başlıklı makalelerinde, Tethys'in Monterey Bay Aquarium Araştırma Enstitüsü (MBARI) tarafından geliştirilen bir Uzun Menzilli Otonom Su Altı Aracı (LRAUV) platformu olduğunu belirtmişlerdir [2]. Bu platform, 120 kg ağırlığında, 0,3 metre çapında ve 2 metre uzunluğundadır. Derinlik derecesi 300 metreye kadar olan bu araç, 0,5 ila 1 m/s hız aralığına sahiptir [2]. Tethys, fitoplankton çiçeklerini izlemek [3], yükselme akıntılarını incelemek [4] ve Arktik ortamlardaki petrol

sızıntılarını değerlendirmek [5] gibi kullanım alanlarına sahiptir.

Su altı otonom araçları ve deniz araçları için yol planlama ve navigasyon üzerine yapılan çeşitli çalışmalar literatürde bulunmaktadır. Qu ve diğerleri tarafından yapılan bir çalışmada, otonom deniz araçları için Pekiştirmeli Öğrenme (Reinforcement Learning - RL) tabanlı bir yol takibi kontrol şeması sunulmaktadır. Surge-heading birleşik rehberlik yöntemi, aracın istenen yolu doğru bir şekilde takip etmesini sağlamaktadır. TD3 algoritması kullanılarak, model belirsizlikleri ve ortam bozulmaları dikkate alınarak yol takip eylemleri üretilir. LSTM-TD3 (Long Short Term Memory – Twin Delayed Deep Deterministic Policy Gradient) tabanlı kontrol şemasının etkinliği ve üstünlüğü, simülasyon sonuçları ve kapsamlı karşılaştırmalarla gösterilmiştir [6].

Ma ve diğerleri tarafından yürütülen araştırma kapsamında, otonom su altı araçları için hassas navigasyonu sağlamak amacıyla derin öğrenme tabanlı bir adaptif navigasyon algoritması önerilmiştir. Algoritma, navigasyon sistemlerinin hata birikimini düzeltmek için derin öğrenme kullanmış ve Doppler Hız Logu (Doppler Velocity Log - DVL) ölçümünün başarısız olup olmadığını belirlemek için χ^2 kuralını benimsemiştir. Ayrıca, navigasyon doğruluğunu daha da artırmak için değişken Bayesian (VB) yöntemine dayalı bir adaptif filtre kullanılmıştır. OSA (Otonom Su Altı Aracı) alan verilerine dayanan deneyler, önerilen algoritmanın sağlam navigasyon performansı ve pozisyon doğruluğunu önemli ölçüde artırdığını göstermiştir [7].

Liu ve diğerleri tarafından gerçekleştirilen çalışmada otonom su altı araçlarında navigasyon için Derin Pekiştirmeli Öğrenme için Derin Deterministik Politika Gradyanı (DDPG) yöntemi kullanılmıştır. Çalışmada sadece iç sensörlerden elde edilen bilgilerle çalışan bir algoritma geliştirilerek çevresel etkileşimler yoluyla davranışların öğrenilmesi amaçlanmıştır. Algoritmanın etkinliğini değerlendirmek için simülasyon ortamında bir dizi deney gerçekleştirilmiş ve elde edilen sonuçlarla DDPG algoritmasının su altı araçlarının kontrol sistemlerinde uygulanabilirliğini ve karmaşık uygulama ortamlarında pontansiyeli doğrulanmıştır [8].

Zhang ve diğerleri bilinmeyen aktüatör doyumluğu ve çevresel bozulmalar yaşayan az eylemsiz denizaltı araçları için yaklaşım açısı tabanlı üç boyutlu yol-takip kontrol şeması önermişlerdir. Az

eylemsiz denizaltı terimi, aktüatörlerin veya itici güçlerin sayısının aracı hareketlerini tam olarak kontrol etmek için yetersiz olduğu otonom deniz araçlarını tanımlamaktadır [9].

Tian ve ekibinin çalışmasında, OSA'lardaki rota izleme problemini çözmek için kinematik ve dinamik modellere dayanan bir bulanık denetleyici geliştirilmiştir. Bulanık denetleyicinin etkin performansı için, rota uzunluğu, düzgünlük ve çapraz iz pozisyon hatası gibi çoklu optimizasyon performans göstergeleri belirlenmiştir. Üyelik fonksiyonlarının parametrelerini belirlemek için Parçacık Sürü Optimizasyonu (Particle Swarm Optimisation - PSO) algoritması kullanılmıştır. Önerilen algoritmanın performansını test etmek için düz çizgi, sinüs eğrisi, yarım ay şekli, Arşimed spirali ve pratik rotaları içeren çeşitli senaryolar sunulmuştur [10].

Jian ve ekibinin araştırmasında, OSA'lar için geliştirilen bir planlayıcı, hem büyük ölçekli bilinen statik ortamlarda hem de dinamik engellerin bulunduğu ortamlarda çarpışma olmadan bir yol belirleme yeteneği hedeflenmiştir. Bu çalışmada, dinamik engellerin varlığı durumunda OSA'lar için yerel iz ve global yol planlamak için değiştirilmiş ve birleştirilmiş Dinamik Pencere Yaklaşımı (Dynamic Window Approach - DWA) ve Hızlı Keşfedilen Rastgele Ağaç (Rapidly Exploring Random Tree - RRT) yöntemleri kullanılmıştır. Önerilen hibrit yaklaşımda, planlayıcı otomatik olarak çarpışma riskini değerlendirip ve gerektiğinde Dinamik Pencere Yaklaşımı ile arasında geçiş yaparak, dinamik engellemekte ortadan kalktıktan sonra ise tekrar Rastgele Orman yöntemine geçmektedir [11].

Fang ve ekibinin yürüttüğü çalışmada, otonom denizaltı araçları için engel kaçınma amacıyla kendini ayarlayan bir bulanık denetleyici kullanarak basit bir arama tekniği önerilmiştir. Denizaltı aracının hidrodinamik katsayılarını belirlemek için Planar Hareket Mekanizması (Planar Movement Mechanism - PMM) testi kullanılmış ve bu veriler denetim sistemi için önemli veri girişleri olarak kullanılmıştır. Kendini ayarlayan bulanık denetleyici, denizaltı araçlarının itişini yönlendirmek için benimsenmiştir. Engelden kaçınma fonksiyonu, sualtı görüntü tespit yöntemiyle birlikte BK üçgen alt ürünü kullanılarak belirlenmiştir. Yapılan simülasyonlar, kendini ayarlayan kontrol, görünürlük, güvenlik ve akıntı etkisi olmak üzere dört tür hareket kontrol faktörünün manevra performansını incelemek için kullanılmıştır [12].

Bu çalışma, Robot İşletim Sistemi (ROS) çevresinde Gazebo simülasyon ortamında gerçekleştirilen araştırmaya dayanmaktadır. Çalışmada Monterey Körfezi Akvaryumu Araştırma Enstitüsü tarafından geliştirilen Tethys UMOSA [13] Pekiştirmeli Öğrenme algoritmasının kullanılması incelenmiştir. Deneyler Gazebo simülasyon ortamında [14] gerçekleştirilmiştir. Yapılan deneylerde, Paper ve arkadaşları tarafından geliştirilen Tethys UMOSA'nın modellendiği Gazebo su altı simülasyon ortamı [15] kullanılmıştır.

Su altı otonom araçları farklı kategorilerde ele alınmakta olup, UMOSA'lar yapısal farklılıklarından kaynaklanan farklı kinematik yapıya sahiptirler [16]. Bu çalışmada, bu farklılık göz önünde bulundurularak bir UMOSA modeli kullanılmış ve çevrim içi bir kontrol mekanizması kullanılarak aracın su altı gibi dinamik bir ortamda yol planlama yetisini kazanması amaçlanmıştır. Çalışmada, ortamdaki elde edilen sensör verilerinin kısıtlı olduğu senaryolarda dahi (örneğin kamera verisinin suyun bulanıklığı nedeniyle kullanılmadığı) kullanılan yol planlama mekanizmasının geçerli olması hedeflenmiştir. Ayrıca, su altında aktif olarak faaliyet gösteren bir UMOSA modeli üzerinde simülasyon ortamında deneyler yapılarak, geliştirilecek çevrim içi algoritmanın maliyet etkin bir şekilde geliştirilmesi ve test edilmesi amaçlanmıştır.

Monterey Körfezi tarafından oluşturulan Tethys UMOSA simülasyon modeli üzerinde yapılan deneyler, Pekiştirmeli Öğrenme tabanlı bir kontrol yaklaşımının potansiyel avantajlarını değerlendirmektedir. Bu çalışmanın amacı, OSA'nın yörüngesini belirlemek için Derin Q-Ağları algoritmasını kullanarak, Pekiştirmeli Öğrenme tabanlı bir kontrol yaklaşımının etkinliğini göstermektir.

Yapay zeka ve özellikle de Pekiştirmeli Öğrenme algoritmaları, OSA'ların kontrolünde yeni bir yaklaşım sunmaktadır. Pekiştirmeli Öğrenme, denizaltı araçlarının çeşitli denizaltı görevleri için otomatik kontrolünü sağlamak için güçlü bir potansiyele sahiptir. Bu bağlamda, bu çalışmada, geleneksel denetleyicilerin yerine gerçek zamanlı olarak Pekiştirmeli Öğrenme algoritmalarının kullanılmasıyla OSA'ların kontrolüne dair yapay zeka temelli bir yaklaşım sunulmaktadır.

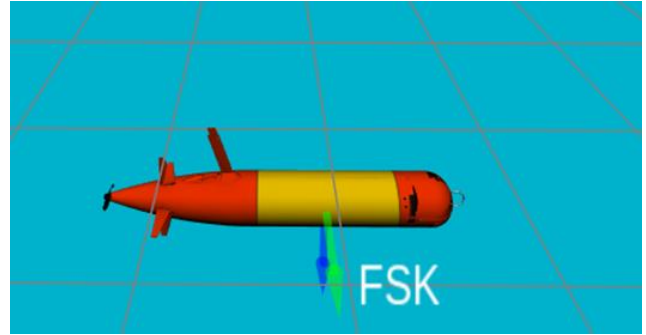
2. MATERYAL VE METOD (MATERIALS AND METHODS)

2.1. Materyal (Material)

Çalışmada, su altı ortamının simüle edilmesi için Gazebo kullanılmıştır. Su altı aracının kontrolü ve veri akışı ise ROS aracılığıyla sağlanmıştır. Tethys UMOSA modeline ait dört adet sensör (manyetometre sensörü, IMU sensörü, navigasyon uydusu sensörü ve batarya sensörü) ile dört adet aktüatör (itici, yatay kanatlar, dikey kanatlar ve kaldırma kuvveti motoru) kullanılmıştır.

2.1.1. Gazebo (Gazebo)

Çalışmada, su altı ortamının simüle edilmesi için Gazebo kullanılmıştır. Su altı aracının kontrolü ve veri akışı ise ROS aracılığıyla sağlanmıştır. Tethys UMOSA modeline ait dört adet sensör (manyetometre sensörü, IMU sensörü, navigasyon uydusu sensörü ve batarya sensörü) ile dört adet aktüatör (itici, yatay kanatlar, dikey kanatlar ve kaldırma kuvveti motoru) kullanılmıştır.



Şekil 1. Gazebo simülasyon ortamında tethys modeli [3] (The tethys model in gazebo simulation environment)

2.1.2. Robot işletim sistemi (ROS) (Robot operating system (ROS))

Çalışmanın uygulanmasında kullanılan su altı aracı kontrol algoritmasının geliştirmek için açık kaynaklı ROS2 Humble sürümü kullanılmıştır. ROS, çeşitli robotik uygulamaları kolayca entegre etmek için geliştirilmiş bir dizi araç ve kütüphaneden oluşmaktadır. ROS barındırdığı iletişim modeli ile düğümler (nodes) ve konular (topics) aracılığı ile dağıtık sistemlerde veri paylaşımını sağlamaktadır.

2.1.3. Sensörler (Sensors)

Manyetometre Sensörü: Manyetik alanı ölçmek için kullanılan bir sensördür. Tethys UMOSA modelinde manyetometre, çevresel manyetik alanı ölçmek ve konum belirleme için kullanılır.

IMU (Inertial Measurement Unit) Sensörü: Hız, ivme ve açısal hız gibi hareket bilgilerini ölçmek için kullanılan bir sensördür. Tethys UMOSA modelinde IMU, denge ve yönlendirme için kullanılır.

Navigasyon Uydu Sensörü: Uydu sinyallerini alarak konum belirleme ve navigasyon için kullanılan bir sensördür. Tethys UMOSA modelinde navigasyon uydu sensörü, hassas konum belirleme ve rota takibi için kullanılır.

Batarya Sensörü: Batarya gerilimi, akımı ve sıcaklığını ölçmek için kullanılan bir sensördür. Tethys UMOSA modelinde batarya sensörü, enerji tüketimini izlemek ve pil güvenilirliğini sağlamak için kullanılır.

2.1.4. Aktüatörler (Actuators)

İtici: Su altında hareket etmek için kullanılan bir aktüatördür. Tethys UMOSA modelinde itici, ilerleme ve manevra yeteneklerini sağlamak için kullanılır.

Yatay Kanatlar: Yüzeyde dengeyi sağlamak ve yönlendirmek için kullanılan aktüatörlerdir. Tethys UMOSA modelinde yatay kanatlar, dengeyi sağlamak ve dalış/açılma işlemlerini kontrol etmek için kullanılır.

Dikey Kanatlar: Dikey hareketi sağlamak ve dalış/açılma işlemlerini kontrol etmek için kullanılan aktüatörlerdir. Tethys UMOSA modelinde dikey kanatlar, derinlik kontrolü ve manevra yeteneklerini artırmak için kullanılır.

Kaldırma Kuvveti Motoru: Sualtı yüzeyinde kalma ve yüzmeyi sağlamak için kullanılan bir aktüatördür. Tethys UMOSA modelinde kaldırma kuvveti motoru, dalış/açılma işlemlerini kontrol etmek ve su üstünde kalma yeteneğini sağlamak için kullanılır.

2.2. Metod (Method)

Simülasyon ortamındaki su altı aracında yol planlamasının gerçekleştirilmesinde bir Pekiştirmeli Öğrenme algoritması olan Derin Q-Ağları algoritması kullanılmıştır.

Pekiştirmeli Öğrenme, yapay zeka alanında önemli bir konudur ve bir ajanın çevresiyle etkileşime girerek öğrenmesini sağlayan bir öğrenme paradigmasıdır [17]. Pekiştirmeli Öğrenme, geniş

bir uygulama yelpazesine sahiptir ve otomatik kontrol, oyunlar ve robotik gibi alanlarda kullanılmaktadır.

Pekiştirmeli Öğrenme, bir ajanın bir çevre içinde bir dizi durumda hareket ederek toplam ödülü maksimize etmeye çalıştığı bir çerçeve sağlar. Pekiştirmeli Öğrenme, Markov Karar Süreci (Markov Decision Process - MDP) modeliyle temsil edilir, burada ajanın hareketlerine dayanarak bir durumdan diğerine geçme olasılıkları ve bu geçişlerin sonuçları belirlenir.

Bellman denklemi, Pekiştirmeli Öğrenme alanında merkezi bir kavramdır ve bir durumun değerini diğer durumların değerleri ile ilişkilendirir. Bellman denklemi, ajanın kararlarını optimize etmek için kullanılan değer fonksiyonunu iteratif bir şekilde günceller.

Bellman denklemi genellikle Eşitlik (1)'deki şekilde ifade edilir:

$$V(s) = \max_a \sum_{s'} P(s'|s, a) (R(s, a, s') + \gamma V(s')) \quad (1)$$

Bu denklemde, $V(s)$ bir durumun değerini, s geçerli durumu, a bir eylemi, s' sonraki durumu, $P(s'|s, a)$ durum geçiş olasılığını, $R(s, a, s')$ ise durumdan duruma bir ödül fonksiyonunu, ve γ ise indirim faktörünü (discount factor) temsil eder.

Derin Q-Ağları, Q-öğrenme algoritmasını derin öğrenme teknikleriyle birleştirerek karmaşık durum-uzaylarında etkili bir şekilde çalışabilen bir Pekiştirmeli Öğrenme algoritmasıdır. Derin Q-Ağları, deneyim tekrarlarını kullanarak bir ağı parametrelerini günceller ve hedef Q değerlerini tahmin etmek için yapay sinir ağlarını kullanır [18].

Pekiştirmeli Öğrenme alanında, Q-öğrenme algoritması belirgin bir şekilde önem taşır; özellikle bir ajanın karar alma sürecindeki eylem-durum değerlerini güncellemek için zaman farkı (TD) hatasını kullanır. Bu yöntem, eylem-değer fonksiyonu $Q(s, a)$ üzerinde iteratif olarak iyileştirmeler yapar ve bunu Eşitlik (2) ile gerçekleştirir. Burada, α öğrenme oranını, r ise anlık ödülü simgeler. Eşitlik (3), zaman farkı hatasını özlü bir şekilde ifade eder, anlık ödül ile tahmin edilen gelecek ödül arasındaki farkı vurgular. Bu denklemler, ajanın ortamındaki optimal karar alma stratejisini belirlemek için temel oluşturur.

$$Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma \max_{a'} Q(s', a') - Q(s, a)] \quad (2) \quad [18]$$

$$TD = r + \gamma \max Q(s', a') - Q(s, a) \quad (3) \quad [19]$$

3. DERİN Q-AĞLARI ALGORİTMASI İLE SU ALTI ARACININ YÖNLENDİRİLMESİ (NAVIGATION OF AN UNDERWATER VEHICLE USING DEEP Q-NETWORK ALGORITHM)

Oluşturulan algoritma yapısında, her bir zaman birimi içerisinde alınan aksiyonlarla birlikte, bir önceki pozisyona göre hedefe yaklaşma durumuna göre algoritmanın beslenmesi sağlanmıştır. Bu beslenme süreci, Derin Q-Ağları algoritmasıyla bir sonraki aksiyonun aracın sensörlerinden gelen verilere göre seçilmesini sağlamıştır.

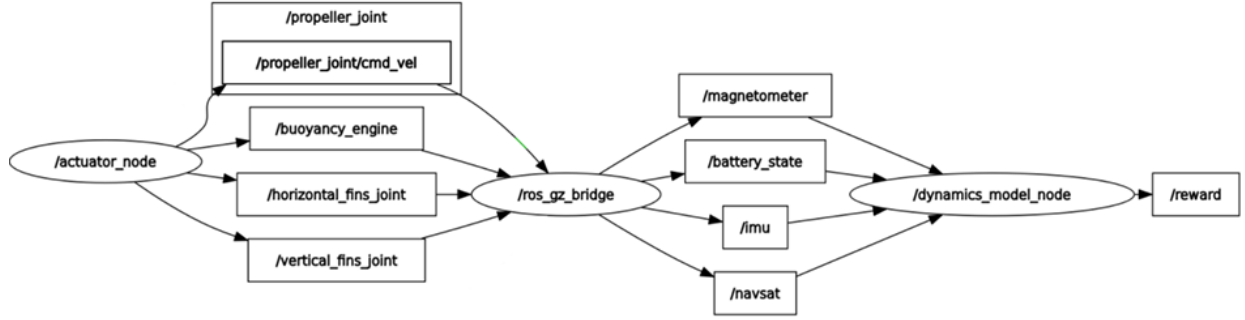
Algoritma: ROS-Gazebo Derin Q-Ağları Tabanlı Kontrol'de ifade edildiği gibi, algoritma simülasyon başlatma, veri akışı ve kontrol adımlarını içermektedir. Başlangıçta, Tethys UMOSA modeliyle birlikte Gazebo simülasyonu başlatılır ve ROS-Gazebo köprüsüyle veri alışverişi sağlanır. Ardından, rastgele bir hedef nokta seçilir ve pekiştirmeli öğrenme algoritması başlatılır. Bu aşamada, durum ve eylem boyutları belirlenir, hiperparametreler ayarlanır ve ortamdan gelen veriler gözlemlenir. Derin Q-Ağları algoritması, bu verileri kullanarak durumlar arasındaki ilişkileri günceller ve optimal kontrol stratejisini belirler. Son olarak, belirlenen sonlanma koşulları kontrol edilir ve gerekirse simülasyon yeniden başlatılır.

Algoritma: ROS-Gazebo Derin Q-Ağları Tabanlı Kontrol

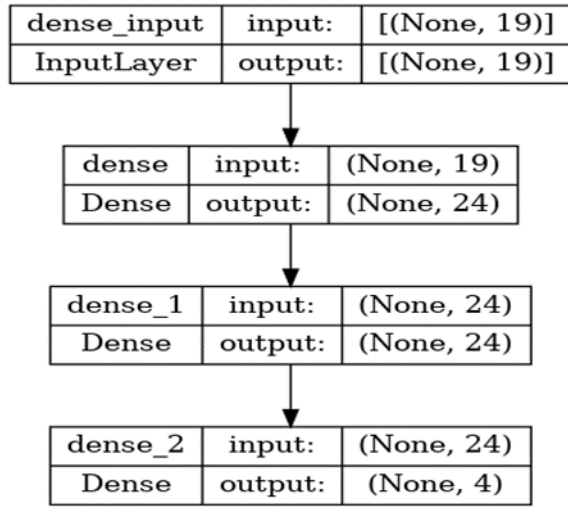
- 1 Gazebo simülasyonunu Tethys UMOSA modeli ile başlat
- 2 ROS-Gazebo köprüsü ile Gazebo ve ROS arasındaki veri akışını başlat
- 3 Simülasyon içerisinde rastgele bir noktayı hedef olarak belirle.
- 4 Pekiştirmeli Öğrenme algoritmasını başlat
- 5 | Durum (state) ve eylem (action) boyutları belirlenir
- 6 | Hiperparametereler (öğrenme oranı, indirim faktörü) ayarlanır
- 7 Bölüm sonlanana kadar ve sonlanma koşulları karşılaşılanaya kadar:
- 8 | Ortamdan mevcut durumu gözlemler
- 9 | | Sensör verilerini vektörel olarak kaydet ve sonraki durumu oluştur
- 10 | Hedef noktaya olan mesafedeki değişime bağlı olarak ortamdan ödül al
- 11 | Derin Q-Ağları algoritmasını (durum, eylem, ödül, sonraki durum) demeti ile güncelle.
- 12 | Sonlanma koşullarını kontrol et:
- 13 | **Eğer** sonlanma koşulları sağlandıysa veya bölüm uzunluğu aşıldıysa:
- 14 | | Yeniden Başla

Tethys UMOSA üzerindeki sensörler ve aktüatörlerin Robot İşletim Sistemi'ne yansıtılması için Gazebo Simülasyon sistemi ve ROS_Gazebo_Köprüsü kullanılarak bir deney düzeneği oluşturulmuştur. Bu şekilde Robot İşletim

Sistemi ile oluşturulan algoritma tarafından belirlenen aksiyonların simülasyon sistemine, sensör verilerinin ise simülasyon sisteminden Robot İşletim Sistemi'ne doğru aktığı Şekil-2'deki yapı elde edilmiştir.



Şekil 2. ROS-Gazebo köprüsü ile elde edilen iletişim grafiği (Communication graph obtained with ROS-Gazebo bridge)



Şekil 3. Yapay sinir ağları mimarisi (Artificial neural network architecture)

Durum uzayı ile beslenen bu model yapısında çıktı olarak 4 boyutlu bir aksiyon uzayı elde edilmiş ve bu aksiyon uzayında her bir aktüatör için elde edilen değer Robot İşletim Sistemi üzerinden Gazebo simülasyon sistemine iletilerek batch eğitiminde geri dönüşüm (feedback) kontrolü sağlanmıştır.

Derin Q-Ağları algoritması için kullanılan hiperparametreler ise aşağıdaki gibidir:

- durum boyutu (state size): 19
- eylem boyutu (action size): 4
- öğrenme oranı (learning rate): 0,1
- indirim faktörü: 0,99
- keşif oranı (exploration rate): 1,0
- keşif azalma oranı (exploration decay rate): 0,99

4. BULGULAR VE TARTIŞMA (RESULTS AND CONCLUSION)

Elde edilen sonuçlar, bu belirlenen hiperparametrelerin algoritmanın performansı üzerinde belirgin bir etkiye sahip olduğunu göstermektedir. Özellikle, düşük öğrenme oranlarının daha istikrarlı ancak yavaş bir öğrenme sürecine yol açtığı, yüksek indirim faktörlerinin uzun vadeli ödülleri daha fazla vurguladığı ve belirlenen keşif oranı ve keşif azalma oranının dengeli bir performans sağladığı gözlenmiştir.

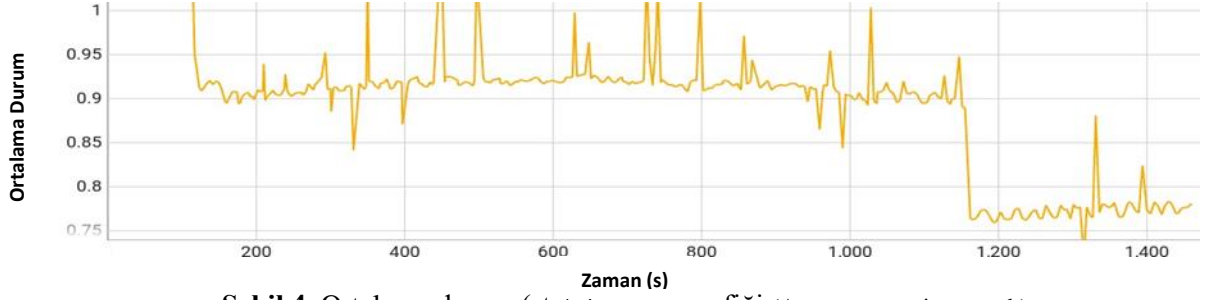
Öncelikle, öğrenme oranı parametresinin, algoritmanın konverjans hızı ve istikrarı üzerinde önemli bir etkiye sahip olduğu gözlenmiştir. Düşük öğrenme oranları daha istikrarlı ancak yavaş bir öğrenme sürecine yol açarken, yüksek öğrenme oranları daha hızlı ancak daha değişken bir öğrenme sürecine neden olmuştur.

Ayrıca, indirim faktörü parametresinin, algoritmanın uzun vadeli ödülleri dikkate alma yeteneği üzerinde kritik bir rol oynadığı gözlenmiştir. Yüksek indirim faktörleri uzun vadeli ödülleri daha fazla vurgularken, düşük indirim faktörleri daha fazla anlık ödüllere odaklanmaya neden olmuştur.

Keşif ve kıyaslama dengesini sağlamak için belirlenen keşif oranı ve keşif azalma oranı parametrelerinin, algoritmanın başlangıçta keşif yapmasını ve daha sonra kıyaslama yapmasını sağlayarak dengeli bir performans sağladığı gözlenmiştir.

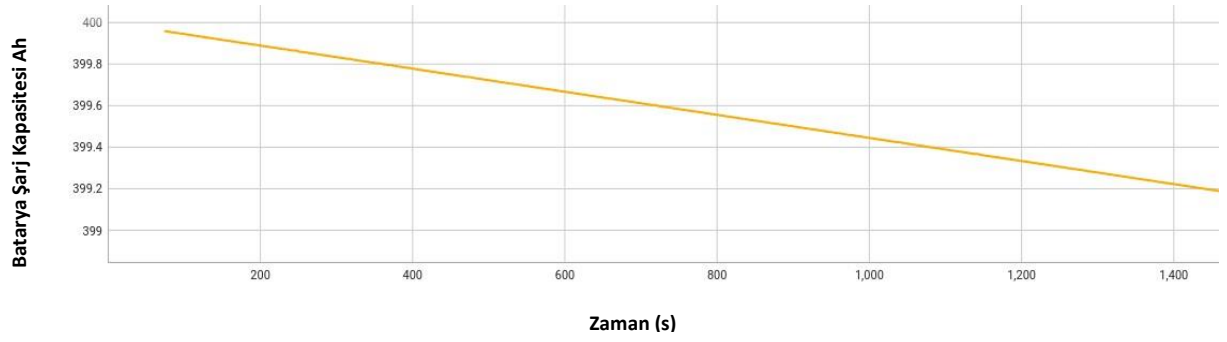
Durum uzayında alınabilecek değerler sensörlerin en küçük ve en büyük değerleri arasındadır. Durum uzayı 0-1 aralığında normalize edildiğinde durum demetindeki değerlerin ortalamasının zamana bağlı değişimi Şekil-4'te verilmiştir. Epsilon-greedy

yaklaşımı ile simülasyon başlangıcında iken simülasyon hedefe yaklaşılması ile gözlemlenen ortalama durum 0,9 ve 1,0 aralığında gözlemlenen ortalama durum 0,75 bandındadır.



Şekil 4. Ortalama durum (state) zaman grafiği (Average state time graph)

Şekil-5 ve Şekil-6'da gösterilen batarya doluluk gözlenmiştir. Başlangıçtaki batarya şarj kapasitesi oranında zamana bağlı olarak linear bir düşüş 400 Ah'tır.



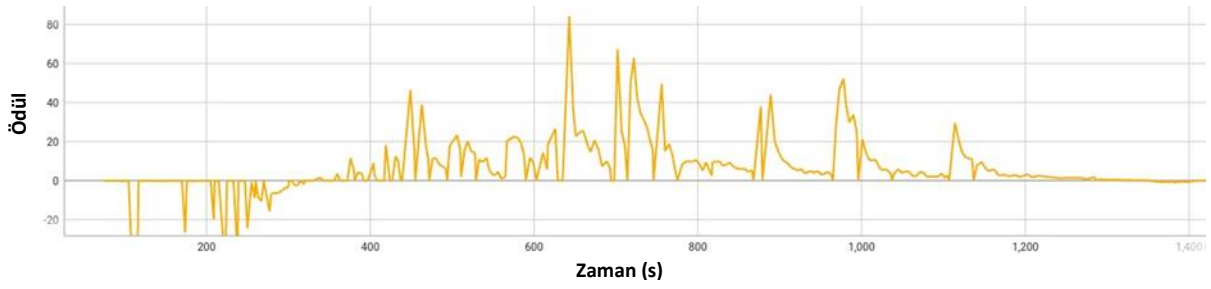
Şekil 5. Batarya şarj kapasitesi zaman grafiği (Battery charge capacity over time graph)



Şekil 6. Batarya doluluk yüzdesi grafiği (Battery capacity percentage graph)

Epsilon-greedy yaklaşımı ile başlangıçta çevreyi keşfetmesi amacı ile rastgele eylemleri seçen su altı aracında alınan ödüllerin başlangıçta negatif olduğu ancak keşfetme oranının düşmesi ve Derin Q-Ağlarında önceki aksiyon-durum değerleri baz

alınarak aksiyon alındığında alınan ödüllerin arttığı Şekil-7'de görülmektedir. Hedefe olan uzaklığa orantılı olarak negatif ödül verildiğinde verilen ödüllerin (-20,50) aralığında olduğu gözlenmiştir.



Şekil 7. Ödül, zaman grafiği (Reward, time graph)

Şekil-8’de görülebileceği gibi başlangıçta 700 metre olan hedefe olan uzaklık, ilk 200 saniye

içinde 1100 metreye kadar artmış, sonrasında ise hedefe 100 metreye kadar yaklaşmıştır.



Şekil 8. Hedefe uzaklık zaman grafiği (Distance to target over time graph)

Elde edilen sonuçlar, Derin Q-Ağları tabanlı yol planlama yönteminin su altı araçlarında başarılı bir şekilde uygulanabileceğini göstermektedir. Derin Q-Ağları, çeşitli su altı ortamlarında görevlerin başarılı bir şekilde tamamlanması için etkili rotalar belirleme yeteneğine sahiptir.

Ayrıca, bu çalışma, Derin Q-Ağları algoritmasının su altı yol planlama problemlerindeki performansını etkileyen anahtar hiperparametrelerin belirlenmesine odaklanmıştır. Özellikle, öğrenme oranı, indirim faktörü ve keşif oranı gibi hiperparametrelerin, Derin Q-Ağları algoritmasının başarımı ve konverjansı üzerinde belirgin bir etkisi olduğu gözlenmiştir. Bu hiperparametrelerin dikkatlice ayarlanması, Derin Q-Ağları tabanlı su altı yol planlama algoritmalarının en iyi performansı sergilemesini sağlamak için kritik öneme sahiptir.

Yapılan deneyler sonucunda, hiperparametrelerin yanı sıra tercih edilen öğrenme mekanizması ve ödül verme algoritmasının yol planlaması üzerinde önemli bir etkiye sahip olduğu gözlenmiştir. Epsilon-greedy öğrenme stratejisinin kullanılmasıyla, Şekil-7’de görüldüğü gibi simülasyon başlangıcında hedefe olan uzaklık artmıştır. Ancak, çevrenin keşfedilmesi ve toplanan ödül verilerinin kullanılmasıyla birlikte hedefe olan uzaklıkta bir düşüş gözlenmiştir.

Ayrıca, ödül mekanizmasının geliştirilmesinde robotun hareketsizlik durumunun da göz önünde

bulundurulmasının önemli olduğu görülmüştür. Uzaklığa bağlı olarak ödül verilmesi durumunda, robotun negatif ödülü tercih etmemesi ve sıfır aksiyon almasına bağlı olarak durması, eğitim süresinin uzamasına neden olabilir. Bu nedenle, hareketsizlik durumu da dikkate alınarak ödül mekanizması geliştirilmelidir.

Bu sonuçlar, belirtilen hiperparametrelerin dikkatlice seçilmesinin, Derin Q-Ağları algoritmasının başarıyla uygulanmasında kritik öneme sahip olduğunu vurgulamaktadır.

5. SONUÇ (RESULT)

Sonuç olarak, bu çalışma, UMOSA’ların yol planlaması için Derin Q-Ağları algoritmasının etkinliğini ROS2 çevresinde Gazebo simülasyon ortamında değerlendirmiştir. Yapılan deneyler, Derin Q-Ağları tabanlı kontrol yaklaşımının, geleneksel denetleyicilere kıyasla potansiyel avantajlar sunduğunu göstermektedir. Özellikle, öğrenme oranı, indirim faktörü ve keşif oranı gibi hiperparametrelerin, algoritmanın performansını belirgin bir şekilde etkilediği gözlenmiştir. Ayrıca, ödül mekanizmasının ve öğrenme stratejisinin yol planlaması üzerinde önemli bir etkiye sahip olduğu tespit edilmiştir. Bu bulgular, Derin Q-Ağları ağlarının su altı araçlarının gelecekteki görevlerinde önemli bir rol oynayabileceğini ve daha etkili, verimli ve güvenilir bir şekilde hareket etmelerine katkıda bulunabileceğini öne sürmektedir. Ancak, farklı mimarilerin karşılaştırılması ve

hiperparametrelerin etkisi gibi konular daha detaylı araştırmaya ihtiyaç duyabilir. Ayrıca, su altı ortamındaki belirsizliklerin daha iyi modellenmesi için Derin Q-Ağları algoritmalarının iyileştirilmesi gerekmektedir. Bu yönde yapılabilecek çalışmalar, su altı görevlerinin daha güvenilir ve etkili bir şekilde gerçekleştirilmesine katkı sağlayabilir.

ETİK STANDARTLARIN BEYANI (DECLARATION OF ETHICAL STANDARDS)

Bu makalenin yazarı çalışmalarında kullandıkları materyal ve yöntemlerin etik kurul izni ve/veya yasal-özel bir izin gerektirmediğini beyan ederler.

The author of this article declares that the materials and methods they use in their work do not require ethical committee approval and/or legal-specific permission.

YAZARLARIN KATKILARI (AUTHORS' CONTRIBUTIONS)

Emre GÖZÜTOK: Deneyleri yapmış, sonuçlarını analiz etmiş ve maklenin yazım işlemini gerçekleştirmiştir.

He conducted the experiments, analyzed the results and performed the writing process.

Fecir DURAN: Deneyleri yapmış, sonuçlarını analiz etmiş ve maklenin yazım işlemini gerçekleştirmiştir.

He conducted the experiments, analyzed the results and performed the writing process.

ÇIKAR ÇATIŞMASI (CONFLICT OF INTEREST)

Bu çalışmada herhangi bir çıkar çatışması yoktur.

There is no conflict of interest in this study.

KAYNAKLAR (REFERENCES)

- [1] [Watson, S.; Duecker, D.A.; Groves, K. Localisation of Unmanned Underwater Vehicles (UUVs) in Complex and Confined Environments: A Review. Sensors 2020
- [2] Phillips, A.B., vd. (2023). "Autosub Long Range 1500: A continuous 2000 km field trial." Ocean Engineering, 280, 114626.
- [3] Godin, M.A., vd. (2011). "Real-time sensing of upwelling from a moving autonomous platform." Limnology and Oceanography: Methods, 9(1), 1-13.

- [4] Zhang, Y., vd. (2012). "Using AUVs to study frontal dynamics." Journal of Field Robotics, 29(6), 1035-1048.
- [5] Kukulya, A., vd. (2016). "AUVs in the Arctic: A platform for interdisciplinary science." OCEANS 2016 MTS/IEEE Monterey.
- [6] Qu, Xingru, et al. "A Deep Reinforcement Learning-Based Path-Following Control Scheme for an Uncertain Under-Actuated Autonomous Marine Vehicle." Journal of Marine Science and Engineering 11.9 (2023): 1762.
- [7] Ma, Hui, Xiaokai Mu, and Bo He. "Adaptive navigation algorithm with deep learning for autonomous underwater vehicle." Sensors 21.19 (2021): 6406.
- [8] Liu, Tao, Yuli Hu, and Hui Xu. "Deep reinforcement learning for vectored thruster autonomous underwater vehicle control." Complexity 2021 (2021): 1-25.
- [9] Zhang, Jialei, et al. "Approach-angle-based three-dimensional indirect adaptive fuzzy path following of under-actuated AUV with input saturation." Applied Ocean Research 107 (2021): 102486.
- [10] Ma, H.; Mu, X.; He, B. Adaptive Navigation Algorithm with Deep Learning for Autonomous Underwater Vehicle. Sensors 2021, 21, 6406
- [11] Tian, Q.; Wang, T.; Song, Y.; Wang, Y.; Liu, B. Autonomous Underwater Vehicle Path Tracking Based on the Optimal Fuzzy Controller with Multiple Performance Indexes. J. Mar. Sci. Eng. 2023, 11, 463
- [12] Fang, Ming-Chung, et al. "Applying the self-tuning fuzzy control with the image detection technique on the obstacle-avoidance for autonomous underwater vehicles." Ocean Engineering 93 (2015): 11-24.
- [13] B. W. Hobson, J. G. Bellingham, B. Kieft, R. McEwen, M. Godin, and Y. Zhang, "Tethys-class long range AUVs - extending the endurance of propeller-driven cruising AUVs from days to weeks," in 2012 IEEE/OES Autonomous Underwater Vehicles (AUV), 2012, pp. 1-8

- [14] Open Source Robotics Foundation, "Gazebo." [Çevrimiçi]. Erişim: <https://gazebosim.org>
- [15] Player, T. R., Chakravarty, A., Zhang, M. M., Raanan, B. Y., Kieft, B., Zhang, Y., & Hobson, B. (2023, May). From Concept to Field Tests: Accelerated Development of Multi-AUV Missions Using a High-Fidelity Faster-than-Real-Time Simulator. In 2023 IEEE International Conference on Robotics and Automation (ICRA) (pp. 3102-3108). IEEE.
- [16] Panda JP, Mitra A, Warrior HV. A review on the hydrodynamic characteristics of autonomous underwater vehicles. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*. 2021;235(1):15-29.
- [17] Tonkal, Ö., & Polat, H. (2021). Traffic Classification and Comparative Analysis with Machine Learning Algorithms in Software Defined Networks. *Gazi University Journal of Science Part C: Design and Technology*, 9(1), 71-83.
- [18] Dayan, Peter, and C. J. C. H. Watkins. "Q-learning." *Machine learning* 8.3 (1992): 279-292.
- [19] Tesauro, Gerald. "Td-gammon: A self-teaching backgammon program." *Applications of neural networks*. Boston, MA: Springer US, 1995. 267-285.