

YAPAY ZEKÂ DESTEKLİ YARIM TAŞIT SİSTEMİ: HİDROLİK EYLEYİCİLER VE PEKİŞTİRMELİ ÖĞRENME İLE İVME KONTROLÜ

Bartu IŞIK

Talha OĞUZ

ÖZET

Bu çalışmada, hidrolik pistonlar ile entegre edilmiş yarım taşıt modelinin doğrusal olmayan dinamikleri incelenmiş ve yapay zekâ (YZ) tabanlı kontrol stratejileri uygulanmıştır. Model, servo valflerden gelen basınç değişimlerinin pistonlara aktarıldığı ve bunun sonucunda yarım taşıt modelinin titreşim ve hareket karakteristiklerinin etkilendiği karmaşık bir yapıyı temsil etmektedir. Klasik doğrusal kontrol yöntemlerinin yetersiz kaldığı bu sistemde, özellikle pekiştirmeli öğrenme (Deep Deterministic Policy Gradient - DDPG) yaklaşımı ele alınarak hem basınç farklarının hem de hedef ivme değerlerinin kontrolü amaçlanmıştır. Çalışmada farklı ödül ve gözlem tasarımları, davranış klonlama ve evrimsel strateji gibi yöntemlerle karşılaştırılarak en uygun kontrol politikasının elde edilmesi hedeflenmiştir. Deneysel sonuçlar, doğru tanımlanmış bir ödül fonksiyonu ve gözlem uzayı yardımıyla, hidrolik eyleyicinin etkin kullanımının mümkün olduğunu ve doğrusal olmayan ivme kontrolünde YZ destekli yöntemlerin başarılı performans sergilediğini ortaya koymuştur.

Anahtar Kelimeler: Yapay Zekâ, Pekiştirmeli Öğrenme, İvme Kontrolü.

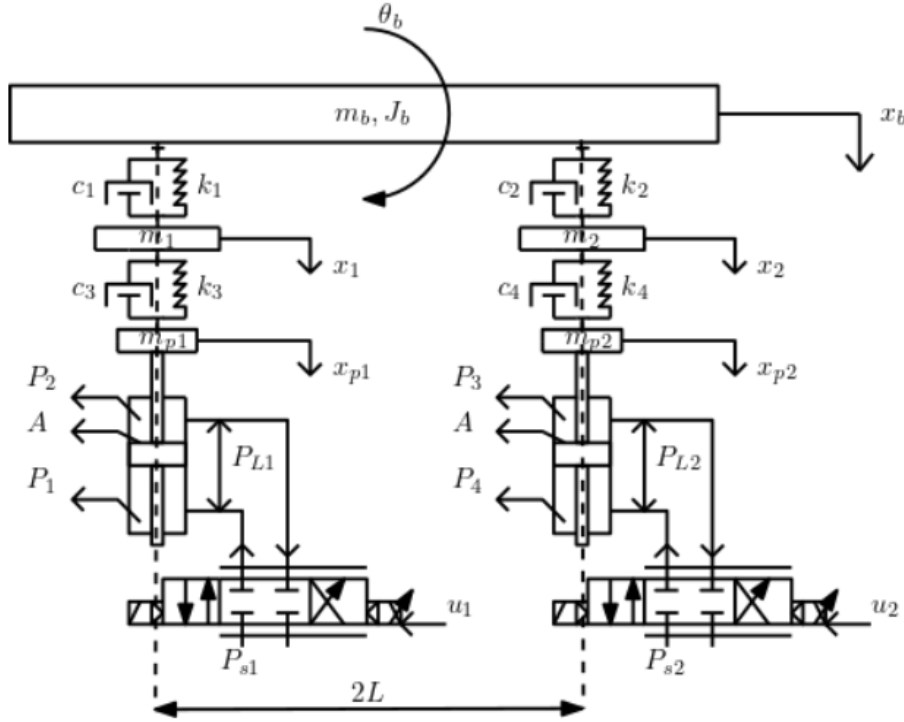
ABSTRACT

In this study, the nonlinear dynamics of a semi-vehicle model integrated with hydraulic pistons were investigated and artificial intelligence (AI) based control strategies were implemented. The model represents a complex structure where pressure changes from servo valves are transferred to the pistons and as a result, the vibration and motion characteristics of the semi-vehicle model's are affected. In this system where classical linear control methods are inadequate, the reinforcement learning (Deep Deterministic Policy Gradient - DDPG) approach was taken into consideration and the control of both pressure differences and target acceleration values was aimed. In the study, it was aimed to obtain the most appropriate control policy by comparing different reward and observation designs with methods such as behavior cloning and evolutionary strategy. Model results revealed that with the help of a correctly defined reward function and observation space, effective use of the hydraulic actuator is possible and AI-supported methods exhibit successful performance in nonlinear acceleration control.

Key Words: Artificial Intelligence, Reinforcement Learning, Acceleration Control.

1. GİRİŞ - HİDROLİK EYLEYİCİ ENTEGRELİ YARIM TAŞIT MODELİ

Bu çalışmada, Şekil 1'de gösterildiği üzere, hidrolik-piston entegreli yarım taşıt sisteminin tasarımı ve matematiksel modellenmesi ele alınmaktadır. Model, dikey hidrolik pistonlar üzerine yerleştirilen taşıtın karmaşık dinamiklerini, gerçek yol koşullarını simüle edecek şekilde tanımlamaktadır. Bu yapıda, sistemin performansı, hidrolik piston bölmeleri arasındaki basınç farkları (P_{L1} ve P_{L2}) ile belirlenir. Basınç farkları, pistonların düşey doğrultuda hareket etmesine ve taşıtın gövdesinde titreşimlerin oluşmasına neden olur.



Şekil 1. Hidrolik-Piston Entegreli Yarım Taşıt Test Sistemi Modeli

2. MODEL DİNAMİKLERİ VE DENKLEMLERİ

Sistemin matematiksel temsili, Denklem 1-8 arasında açıklanmaktadır. Temel durum değişkenleri, taşıt gövdesinin düşey (x_b) ve dönme (θ_b) yer değiştirmesi ile pistonların (x_{p1}, x_{p2}) ve tekerleklerin (x_1, x_2) hareketlerini içermektedir. Süspansiyonlar ve lastikler için yay-sönüm elemanları kullanılarak sertlik (k_1, k_2, k_3, k_4) ve sönüm (c_1, c_2, c_3, c_4) etkileri modellenmiştir. Hidrolik kuvvetler ise, hacimsel akış (Q) ve yığın modülü (β) üzerinden türetilen f_1, f_2, f_3 terimleri ile temsil edilmektedir ve pistonların taşıt dinamikleriyle etkileşimini açıklamaktadır.

Gövde dinamikleri şu şekilde ifade edilmektedir:

$$m_b \ddot{x}_b = -k_1(x_b - x_1) - c_1(\dot{x}_b - \dot{x}_1) - k_2(x_b - x_2) - c_2(\dot{x}_b - \dot{x}_2) \quad (1)$$

$$J_b \ddot{\theta}_b = -k_1 L(-x_b + x_1 + L\theta_b) - c_1 L(-\dot{x}_b + \dot{x}_1 + L\dot{\theta}_b) - k_2 L(+x_b - x_2 + L\theta_b) - c_2 L(+\dot{x}_b - \dot{x}_2 + L\dot{\theta}_b) \quad (2)$$

$$m_1 \ddot{x}_1 = -k_1(x_1 - x_b) - c_1(\dot{x}_1 - \dot{x}_b) - k_3(x_1 - x_{p1}) - c_3(\dot{x}_1 - \dot{x}_{p1}) \quad (3)$$

$$m_2 \ddot{x}_2 = -k_2(x_2 - x_b) - c_2(\dot{x}_2 - \dot{x}_b) - k_4(x_2 - x_{p2}) - c_4(\dot{x}_2 - \dot{x}_{p2}) \quad (4)$$

$$m_{p1} \ddot{x}_{p1} = -k_3(x_{p1} - x_1) - c_3(\dot{x}_{p1} - \dot{x}_1) + Ax_3 \quad (5)$$

$$m_{p2} \ddot{x}_{p2} = -k_4(x_{p2} - x_2) - c_4(\dot{x}_{p2} - \dot{x}_2) + Ax_4 \quad (6)$$

$$\dot{x}_3 = f_1 \dot{x}_{p1} + f_2 x_3 + f_3 \sqrt{P_s - \text{sign}(u_1) x_3} u_1$$

$$f_1 = -\frac{4\beta A}{V}, f_2 = -\frac{4\beta C_{tl}}{V}, f_3 = \frac{4\beta C_d w k_i}{V\sqrt{\rho}} \quad (7)$$

$$\dot{x}_4 = f_1 \dot{x}_{p2} + f_2 x_4 + f_3 \sqrt{P_s - \text{sign}(u_2) x_4} u_2$$

$$f_1 = -\frac{4\beta A}{V}, f_2 = -\frac{4\beta C_{tl}}{V}, f_3 = \frac{4\beta C_d w k_i}{V\sqrt{\rho}} \quad (8)$$

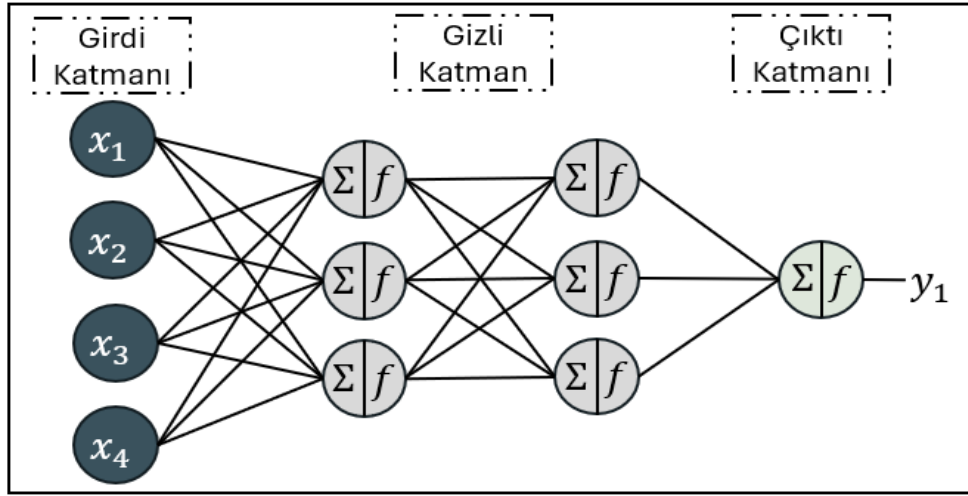
Hidrolik sistem, servo valfler ve pistonlardan oluşmakta olup, elektriksel giriş sinyallerini (u_1, u_2) basınç değişimlerine dönüştürmektedir. Akış hızları ve basınçlar, pistonlara uygulanan net kuvvetleri belirler ve bu kuvvetler, taşıtın genel hareketine katkıda bulunur. Hidrolik aktüasyonun taşıtın yapısal dinamikleriyle entegrasyonu, sistemin doğrusal olmayan (nonlinear) davranış sergilemesine neden olmaktadır. Piston hareketine bağlı hidrolik basınç farklarının zamanla değişmesi, sistemin karmaşıklığını artırarak titreşim analizi ve kontrol tasarımını daha hassas bir şekilde ele almayı gerektirir. İlerleyen aşamalarda bu matematiksel denklemler kullanılarak MATLAB / Simulink modeli kurulmuştur.

3. YAPAY ZEKÂ DESTEKLİ DOĞRUSAL OLMAYAN KONTROL

Robotik, havacılık, otonom sistemler ve benzeri birçok alanda karmaşık dinamik sistemler oluşturulmaktadır. Bu sistemlerin sahip olduğu doğrusal olmayan dinamikler yüzünden geleneksel doğrusal kontrol teknikleri yetersiz kalabiliyor. Örneğin bir pinomatik sistemde kullanılan hava basıncını kontrol eden valfin havanın sıkışabilirliğinin değişkenliğinden dolayı kontrolü zor hale getirecek ve doğrusal olmayan sonuçlar doğuracaktır. Buna benzer bir şekilde hidrolik silindirin iç keçesinde meydana gelen bir aşınma doğrusal olmayan davranışlar oluşmasını sağlar. Bunlar hidrolik ve pinomatik sistemlerde doğrusal olmayan kontrolün gerekliliğini ortaya koyar.

Doğrusal olmayan kontrol tekniklerinin kullanılması büyük bir teorik altyapıya ihtiyaç duyduğu için bu çalışmada kontrol tekniklerini kullanmadan yapay zekâ (YZ) ile nasıl bir kontrol sağlanabileceği ortaya koyulacaktır. Bunun için ilk olarak yapay sinir ağlarının matematik modelini ortaya koymak gerekir. Denklem 9'da x_i nörona bir önceki nörondan gelen giriş değeri ve w_i o nöronun katkısını ortaya koyan katsayıdır, b önyargı katsayısıdır [2]. Bir nörona kendinden önce bağlanan her bir nöronun değerlerinin katsayıları ile toplanması ve bias önyargı katsayısının eklenmesi ile nöronun değeri hesaplanır. Nöron içerisinde yer alan bir aktivasyon fonksiyonu aktif hale gelir ve bu değer nörondan çıkacağı çıktı değerini oluşturur. Aktivasyon fonksiyonları nöronların değerlerin çok büyümesini önler, gidişattaki doğrusallığı ortadan kaldırır ve böylece çözülmesi daha zor modellerin çözülmesine imkân verir. En yaygın kullanılan türleri sigmoid fonksiyonu, hiperbolik tanjant, rectified linear unit formülleridir. Sonuç olarak Denklem 1'de yer alan basit bir matematik işlemi her nöron için yapılır ve aktivasyon fonksiyonu ile elde edilen sonucu bir sonraki nörona iletilir. Bu süreç Şekil 2'de gösterilmektedir. İlk olarak bilgiler girdi katmanından girer, gizli katmandan geçer ve çıktı katmanından çıkarak sonuç ortaya çıkar.

$$Z = \sum_{i=1}^n w_i x_i + b \quad (9)$$



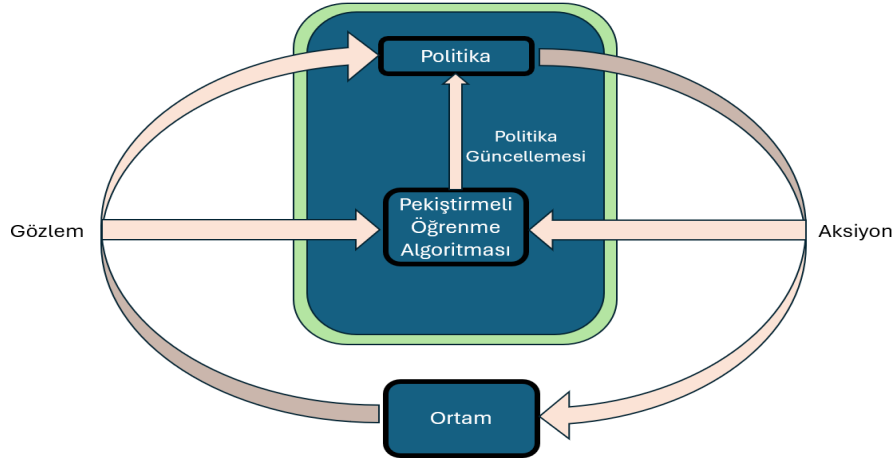
Şekil 2. Yapay Sinir Ağları

YZ destekli kontrol sistemleri, geleneksel modellerden daha az bilgi ve tecrübe ile kurula bilinmesine olanak vermektedir. YZ karmaşık dinamiklerin, kontrol edilmesinde ve optimize edilmesinde kullanılmaktadır. Geri besleme sayesinde, veriye dayalı bir öğrenme sağlar ve kendi cevaplarını iyileştirmeye odaklanır. AI destekli kontrolcüler, -yaşam boyu öğrenme- metotları ile sistemdeki parametre değişikliklerine ve çevresel etkenlere kolaylıkla uyum sağlayabilir. Sistemlerde insan müdahalesi olmadan karar vermesini sağlar. Elimizde bir veri seti olmadığında, sistemin matematik modeli mevcut olsa bile kontrol modeli kurulamaz veya istenilen başarı seviyelerine ulaşamaz. Bu durumda YZ çözümlerinden nasıl faydalanabiliriz?

Pekiştirmeli Öğrenme (PÖ) bir YZ algoritmasının veri setine ihtiyaç duymadan, algoritmanın sistemle iletişim kurarak çıktıları optimize etmek için öğrendiği metot çeşididir. PÖ sistemi şu bileşenler üzerine kurulur;

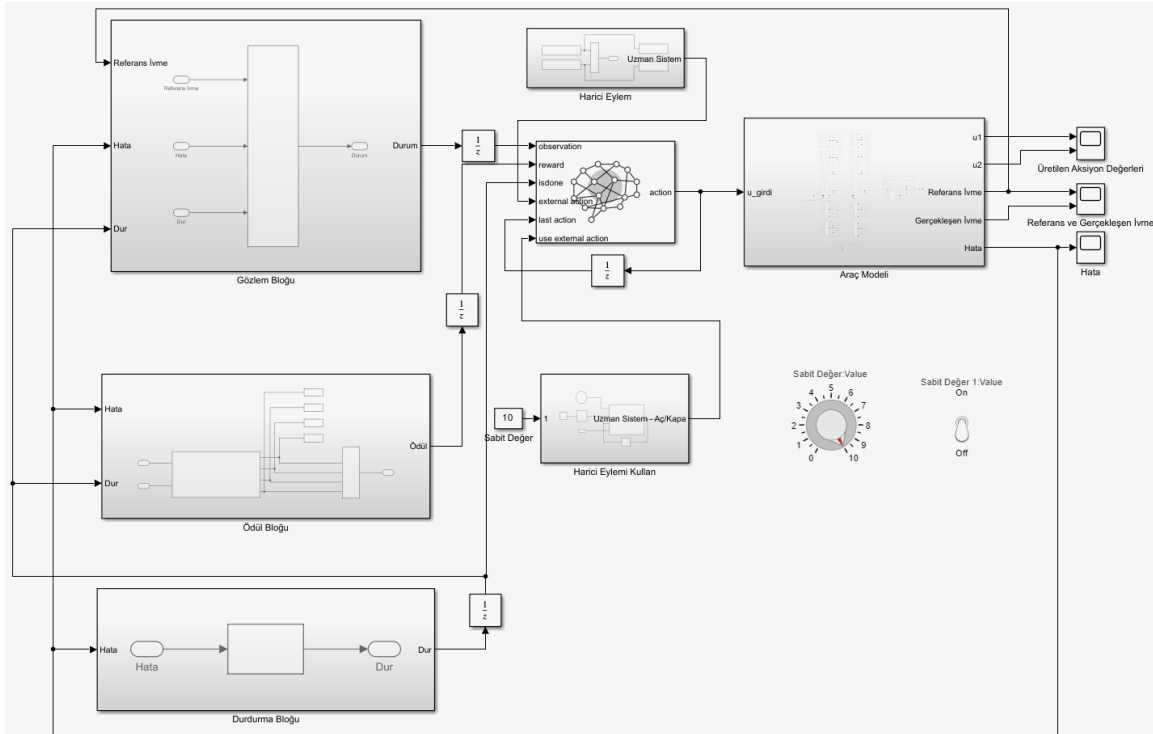
1. Ajan: Ortamdaki bilgileri kullanarak karar veren birim.
2. Ortam: Ajanın etkileşimde bulunduğu sistem.
3. Durum: Ortamın anlık hali.
4. Eylem: Ajanın ortamdaki duruma göre üretmiş olduğu aksiyon.
5. Ödül: Ajanın verdiği tepkinin sonucunda ortamdaki aldığı geri bildirim.

Bu bileşenler bir araya gelir ve ajanın öğrenmesini maksimuma çıkaracak bir politika geliştirilir [3]. Deep Deterministic Policy Gradient (DDPG) sürekli ortamlarda çalışan bir PÖ modelidir. DDPG politikayı ve ödülü optimize etmek için aktör ve kritik bağlantısını kullanır [1]. Aktör ağı politikayı optimize ederek elde edilecek ödülü maksimize etmeye çalışır. Kritik ağ yapısı o anki durum ve eylemin ne kadar ödül getireceğini tahmin eder ve bu aktörün cevaplarını değerlendirir. Geçmiş deneyimler depolanır ve bu sayede daha kararluya giden bir öğrenme süreci oluşturulmaya çalışılır. Şekil 3'te gösterilen politika güncellenme süreci tatmin edici sonuçlar elde edinceye kadar devam eder. Politika, bir ajandanın ortamdaki duruma bağlı olarak nasıl bir eylem seçmesi gerektiğini belirleyen bir fonksiyondur.



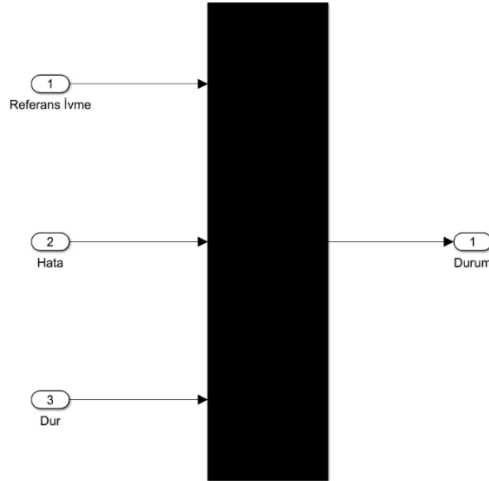
Şekil 3. Politika Güncellemesi

4. YAPAY ZEKÂ MODELİ VE DETAYLARI



Şekil 4. Pekiştirmeli Öğrenme Ortamı

Şekil 4'te PÖ ortamı gösterilmektedir. "Araç Modeli" bloğunda yarım taşıt matematik modeli yer almaktadır ve herhangi bir kontrolcü yer almamaktadır. PÖ ile "u_girdi" aksiyonu üretilmekte, bu iki adet valfin giriş sinyalinin ifade eder. Bu sayede ivme çalacak kontrolcü oluşturulması amaçlanır. Şekil 5'te gözlem bloğu yer almaktadır.

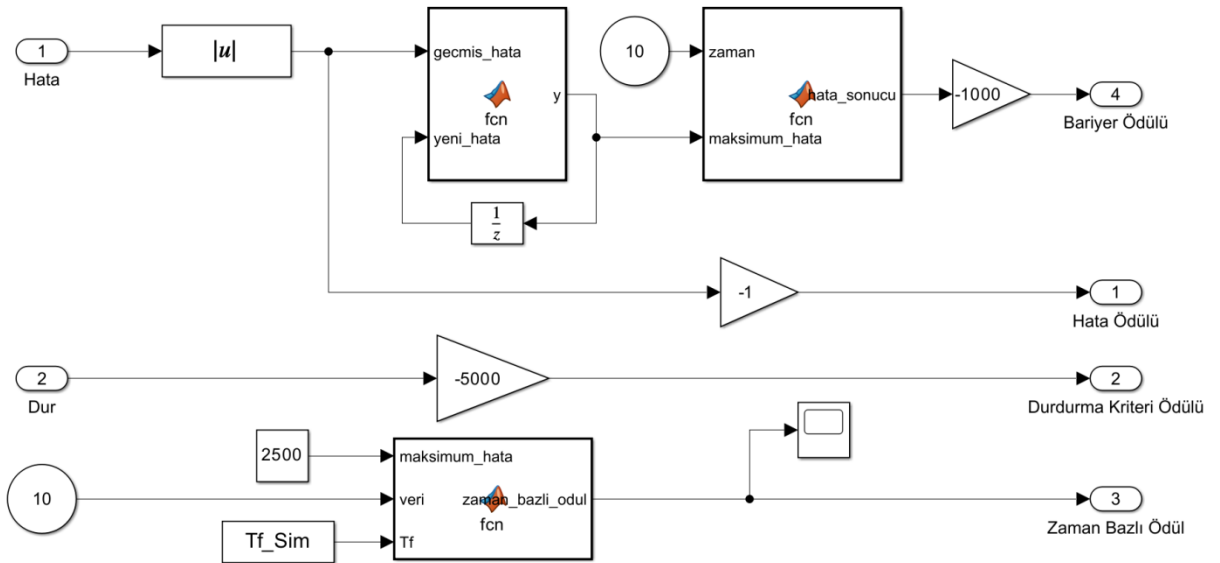


Şekil 5. Gözlem Bloğu

Gözlem bloğunda;

- Referans/istenilen ivme değerleri,
- İstenilen ile elde edilen ivme değerlerinin farkı ve
- Durdurma kriteri sorgusu yer almaktadır.

Simülasyonun örnekleme zamanı 2 saniye olarak belirlenmiş ve örnekleme değeri 1000 Hz. olarak ayarlanmıştır. Gözlem uzayı, alt ve üst sınırları sırasıyla $[-3, -3, 0]$ ve $[3, 3, 1]$ olarak belirlenmiştir. Eylem uzayı, sistemin kontrol sinyallerini temsil eden iki boyutlu bir vektör olarak tanımlanmıştır. Eylemlerin alt ve üst sınırları sırasıyla -1 ve 1 olarak ayarlanmıştır. Kritik ağı, gözlem ve eylem yollarından oluşan derin bir sinir ağıdır. Gözlem yolu, 512, 256, 128, 64 ve 32 nöronlu tam bağlı katmanlar içermektedir. Eylem yolu ise 512, 128 ve 32 nöronlu katmanlardan oluşmaktadır. Bu yollar, birleştirilerek ortak bir yol oluşturulmuştur. Aktör ağı, gözlemleri eylemlere dönüştüren bir sinir ağıdır. Bu ağ, 256, 64, 32 ve 2 nöronlu tam bağlı katmanlar içermektedir. Ağın çıkışı, tanh aktivasyon fonksiyonu ile sınırlandırılmıştır. Kritik optimizasyonu için Adam optimize edici kullanılmış ve öğrenme oranı $1e-3$ olarak ayarlanmıştır. Aktör optimizasyonu için Adam optimize edici kullanılmış ve öğrenme oranı $1e-2$ olarak ayarlanmıştır. Mini-batch boyutu 512, DiscountFactor 0.975 ve hedef yumuşatma faktörü (TargetSmoothFactor) 0,7 olarak belirlenmiştir.



Şekil 6. Ödül Bloğu

YZ politikasının ne kadar iyi olursa o kadar istenilen sonuçlar elde edilebilir. Politikanın başarılı mı başarısız mı olduğunu ise ödül değerinden anlaşılabilir. Yapılan 2 saniyelik eğitimin 1000 Hz. 'de, ödül değerleri elde edilmektedir. Ödül politikası şu şekildedir;

- Durdurma kriteri tetiklenmediği her an: 2,5 puan kazanılır.
- Durdurma kriteri tetiklendiğinde: 5000 puan ceza uygulanır.
- İvme farkının mutlak değeri: Her örnekleme anında, mutlak farkın eksi hali ödül olarak eklenir.
- Bir saniye içindeki maksimum ivme farkı: Mutlak değer -1000 katı ödül olarak eklenir.

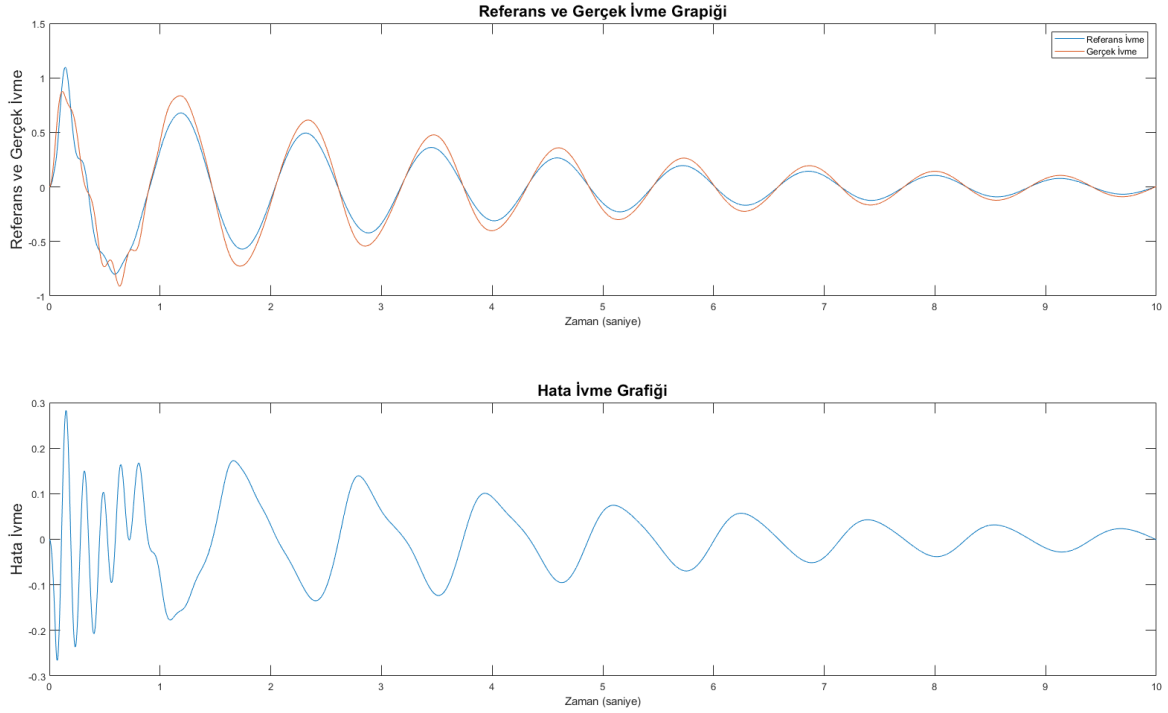
Sonuç olarak elde edilebilecek en yüksek değer 5000 tam puan alabilmektir. Ödül bloğu Şekil 6'te gösterilmektedir. Sonuç olarak 4200 puan civarlarında başarıya ulaşabilen algoritmalar elde edilmiştir. Modelde başarılı olmak için çeşitli ödül ve gözlem blok kombinasyonları denenmiştir. Örnek olarak, gözlem bloklarında oluşturulan ivmedeki hata farkının integrali, istenilen referans ivme değeri, üretilen aksiyon değerleri kullanılmıştır. Ödül bloğunda çeşitli kombinasyonlar yine bir araya getirilmiş ama en başarılı sonuç şu anki kombinasyon ile elde edilmiştir. Sisteme bir adım önce üretmiş olduğu aksiyon değerleri geri beslemeyle giriyor bu sayede kopmaların önüne geçilmesi hedeflenmiştir. Uzun-Kısa Vadeli Bellek (LSTM) ile kurulan ağ yapıları, çalışmalar devamlı bir ortamda yapıldığı için daha başarılı olması beklenmiştir. Ancak, normal yapay sinir ağı yapısı ile örülmüş yapı daha başarılı sonuçlar vermiştir. Bu sistem, daha başarılı kararlar almış ve oldukça stabil bir şekilde çalışmıştır.

5. EĞİTİM VE SONUÇLAR

Çeşitli eğitimler yapılmıştır. Bunlar arasında "Yapay Zekâ Modeli ve Detayları" kısmında bahsedilen ödül ve gözlem blok çeşitleri kullanılmış, bunlara ek olarak davranış klonlama yöntemleri kullanılmış ve bir kontrol sistemi cevaplarını sistemi eğitmek için kullanıp kullanılmayacağı test edilmiştir. Davranış klonlama için yine 2 saniye eğitim yapılmış ve kalan 8 saniye için test yapılmıştır. Eğitim sırasında ajan 1 kere klondan öğrenme, 1 kere 2 saniyelik validasyon, 2 kere klondan öğrenme 1 kere 2 saniyelik validasyon, 3, 4, 5 şeklinde farklı adetlerle gözlemlenmiştir. Ajan bu eğitimde çok hızlı bir şekilde öğrenme gerçekleştirmiştir ve validasyon bölümünde başarılı sonuçlar vermiştir. Sonrasında 10 saniyelik testi tamamen kendisinin karar vermesi istendiğinde ikinci saniyeden sonra sapıtmalar yaşamış ve aşırı sapmadan dolayı durdurma kriteri devreye girmiştir. Davranış klonlama tekniği gerekli başarıyı sağlayamamıştır. Bu yüzden başka bir teknik olan evrimsel strateji tercih edilmiştir. Bu teknikte her eğitim sınıfında belirli sayıda iterasyon gerçekleştirilir ve en başarılı birkaç iterasyonun ortalaması ya da en başarılı iterasyondaki ajan tercih edilerek eğitime devam edilir. Bu ajanın yapmış olduğu yanlış denemelerin eğitim geçmişine etkisini azaltarak hep başarılı sonuçlar ile öğrenmesini arttıracak şekilde yardımcı olur.

Şekil 7'de evrimsel strateji sonucunda elde edilen ajanın vermiş olduğu kararların sonuçları gösterilmektedir. Ajan yine ilk 2 saniye ile eğitime tutulmuş ve 10 saniyelik test süreci Şekil 7'de gösterilmektedir. İlk grafikte referans ivmesini ne kadar iyi takip ettiği gözükmektedir. İkinci grafikte aradaki hata farkı gösterilmektedir. İstenilen ivme değeri 0'da doğru gitmekte ve ajan bu formun takibini sağlayabilmektedir. Maksimum hata farkı ilk saniyedeki patlamada meydana gelmekte ve $0,28 \frac{m}{s^2}$ civarı bir değere denk geliyor.

Sonuç olarak farklı ajan türleri, gözlem ve ödül türleri, eğitim stratejileri denenmiş ve elde edilen en başarılı sonuç paylaşılmıştır. Bu çalışma doğrusal olmayan ivme kontrolünün matematiksel kontrol yöntemlerini kullanmadan nasıl yapılabileceği ile ilgili bir bakış açısı sunmakta ve bunları paylaşmaktadır.



Şekil 7. Referans ve Gerçek İvme Grafiği, Hata Farkı İvme Grafiği

KAYNAKLAR

- [1] Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., Wierstra, D., *Continuous control with deep reinforcement learning*, arXiv preprint, Cilt: 1509, Sayı: 02971, Yayın Yılı: 2015.
- [2] LeCun, Y., Bengio, Y., Hinton, G., *Deep learning*, Nature, Cilt: 521, Sayı: 7553, Say. 436–444, Yayın Yılı: 2015.
- [3] Kiumarsi, B., Modares, H., Lewis, F.L., Karimpour, A., Davoudi, A., *Optimal and autonomous control using reinforcement learning: A survey*, IEEE Transactions on Neural Networks and Learning Systems, Cilt: 29, Sayı: 6, Say. 2042–2062, Yayın Yılı: 2018.

ÖZGEÇMİŞ

Bartu IŞIK

1999 yılı İstanbul doğumludur. 2022 yılında Yeditepe Üniversitesi Mühendislik Fakültesi Makine Mühendisliği Bölümü'nden mezun olmuştur. Lisans eğitiminin ardından aynı üniversitede Makine Mühendisliği alanında bütünleşik doktora programına kabul edilerek lisansüstü çalışmalarına devam etmektedir. 2022 yılından itibaren Mert Teknik Fabrika Malzemeleri Tic. ve San. A.Ş.'de Ar-Ge Departmanında görev yapmaktadır. Akademik ve sektörel çalışmaları, özellikle yapay zekâ tabanlı kontrol sistemleri, pekiştirmeli öğrenme algoritmaları (Reinforcement Learning), robotik sistemlerin optimizasyonu üzerine yoğunlaşmıştır.

Talha OĞUZ

1998 yılı İstanbul doğumludur. 2022 yılında Yeditepe Üniversitesi Mühendislik Fakültesi Makine Mühendisliği Bölümü'nden mezun olmuştur. Lisans eğitiminin ardından Yıldız Teknik Üniversitesi Makine Mühendisliği alanında yüksek lisans programına kabul edilerek lisansüstü çalışmalarına devam etmektedir. 2023 yılından itibaren Mert Teknik Fabrika Malzemeleri Tic. ve San. A.Ş.'de Ar-Ge Departmanında görev yapmaktadır. Çalışma alanları arasında sistem dinamiği, kontrol sistemleri ve hidrolik test teknolojileri yer almaktadır.