



Elektrohidrolik bir sistemin pekiştirmeli öğrenme tabanlı denetleyici ile konum denetiminin gerçekleştirilmesi

Reinforcement learning based position control of an electro-hydraulic system

Mustafa Yavuz Coşkun^{1,*} , Mehmet İtik² 

¹ Karadeniz Teknik Üniversitesi, Makine Mühendisliği Bölümü, 61080, Trabzon, Türkiye

² İzmir Demokrasi Üniversitesi, Makine Mühendisliği Bölümü, 35140, İzmir, Türkiye

Öz

Elektrohidrolik sistemler sağladıkları avantajlar sebebiyle endüstrinin vazgeçilmezi olmuştur. Buna karşın hidrolik sistemlerin doğrusal olmayan karakteristik özellikleri ve çok sayıda parametre belirsizliği barındırması bu sistemlerin denetimini zorlaştıran etmenler olarak öne çıkmaktadır. Bu çalışmada ise oransal valf ile sürülen asimetrik bir hidrolik pistonun konumu pekiştirmeli öğrenme ile denetlenmiştir. Pek çok pekiştirmeli öğrenme algoritması olmasına rağmen sürekli uzayda etkinliği ile öne çıkan derin deterministik politika gradyanı yöntemi tercih edilmiştir. İlgili hiper parametreler öncül çalışmalarla belirlenerek çoklu konum referans sinyali için denetleyicinin eğitimi benzetim ortamında gerçekleştirilmiştir. Elde edilen sonuçları kıyaslamak için aynı çalışma PID denetleyici ile de gerçekleştirilmiştir. Çalışmada kullanılan pekiştirmeli öğrenme yöntemi farklı karakteristiklere sahip konum referans sinyalinin takibinde PID denetleyiciden daha %25.51 oranında daha başarılı sonuçlar ortaya koymuştur.

Anahtar kelimeler: Elektrohidrolik sistem, Modelleme, Konum denetimi, Pekiştirmeli öğrenme, Derin deterministik politika gradyanı

1 Giriş

Orta ölçekli ve ağır sanayi kuruluşlarında elektrohidrolik sistemlerden (EHS) sıklıkla yararlanılmaktadır. Ayrıca bu sistemler tarım aletleri [1-2], robotik kollar [3-4] ve kazıcılar [5-6] gibi mobil uygulamalarda da kendisine yer bulmaktadır. EHS'lerin bu derece yaygın kullanıma sahip olmasının nedeni olarak ağırlıklarına karşın sağladıkları yüksek güç çıktısı gösterilebilir [7]. Bu kadar yaygın kullanılmasına rağmen doğrusal olmayan karakteristik özellikleri ve belirsizlikleri nedeniyle EHS'lerin hassas konum ve kuvvet denetimi zorlu bir süreç olarak kalmaya devam etmiştir [8-9].

Literatürde EHS'lerin konum denetim performansını geliştirmek amacıyla farklı denetleyicilerin kullanıldığı çalışmalar yer almaktadır. Kayan kipli denetim sistemleri bunlardan birisidir. Kolsi-Gdoura vd. [10] kayan kipli denetleyici ve PI denetleyiciyi beraber kullanarak 4/3 valfle sürülen simetrik pistonu sahip bir EHS'nin basamak konum referansı takibini gerçekleştirmiştir. PI denetleyici ile elde

Abstract

Electrohydraulic systems have become an inseparable part of the industry due to the advantages they provide. On the other hand, the nonlinear characteristics of the hydraulic systems and the parametric uncertainties make their control troublesome. In this study, the position of an asymmetrical hydraulic piston driven by a proportional valve was controlled by reinforcement learning. Although there are many reinforcement learning algorithms, the deep deterministic policy gradient, which stands out with its effectiveness in continuous space, has been preferred. The hyperparameters were found by preliminary studies and the training of the controller for the designed position reference signal was carried out numerically. The obtained results are compared with the PID controller. The reinforcement learning method reached 25.51% more successful results than the PID controller in terms of tracking the position reference signal with different characteristics.

Keywords: Electrohydraulic system, Modeling, Position control, Reinforcement learning, Deep deterministic policy gradient

edilen sonuçlar kararlı haldeki takip hatası açısından yeterli olmasına rağmen oturma süresinin uzun olması çalışmada işaret edilen bir durum olmuştur. Kayan kipli denetim ise bu etkinin ortadan kaldırılması için kullanılmış ve oturma zamanında %80'e yakın iyileştirme elde edilmiştir. Buna karşın, kayan kipli denetiminin doğası sebebiyle denetim sinyalinin yüksek frekanslı bir şekilde işaret değişimi (çatırtı, *İng.* chattering) göstermesi yöntemin olumsuz yönü olarak ortaya çıkmaktadır. Kolsi-Gdoura vd. [11] devam çalışmasında EHS denetimi için uyarlamalı kayan kipli denetimden yararlanarak çatırtı sorununu en alt düzeye indirmeye başarmıştır. Fakat bu süreçte daha karmaşık bir denetleyici yapısı ortaya çıkmıştır. Shaer vd. [12] kayan kipli denetim yaklaşımını bulanık mantıkla birleştirerek EHS'nin konum ve kuvvet denetimi için yeni bir yaklaşım önermiştir. Çatırtının önüne geçmek için ise bu alanda etkinliği bilinen üstel erişim yaklaşımını (*İng.* exponential reaching law) tercih etmişlerdir. Sundukları sonuçlar önerdikleri yöntemin çatırtıyı olabildiğince azaltarak tatmin edici takip performansına ulaşabildiğini göstermiştir. Buna rağmen

* Sorumlu yazar / Corresponding author, e-posta / e-mail: mycoskun@ktu.edu.tr (M. Y. Coşkun)

Geliş / Received: 17.08.2022 Kabul / Accepted: 09.10.2022 Yayınlanma / Published: 15.01.2023

doi: 10.28948/ngumuh.1163241

çatırtı tamamen ortadan kaldırılamamış, sadece frekansı azaltılmıştır. Bu durum ise takip performansını kötü yönde etkilemiştir.

Bulanık mantık [13] ve ANFIS [14] gibi başarılı denetim yöntemlerinin EHS'ler üzerine uygulaması da literatürde yer almaktadır. Fakat ilgili yöntemler denetim performansını artırmanın yanı sıra karmaşık bir denetim yapısı sunmaktadır. Bununla beraber EHS'den denetleyiciye geri beslenmesi gereken değişken sayısı da artmaktadır. EHS'lerin çalıştığı ortamlar hassas sensörler için uygun olmamakla beraber çoğu kez ilave sensörler için uygun bulmak kolay olmamaktadır. Bu durum ise bahsi geçen denetleyicilerin endüstride karşılık bulmasını güçleştirmektedir.

Bu noktaya kadar etkili fakat karmaşık yapıya ve/veya istenmeyen etkiler oluşturabilecek denetleyicilerden bahsedilmiştir. Bu dezavantajları barındırmayan ve uygulama kolaylığı sayesinde endüstride karşılık bulan PID denetleyici EHS denetiminde yüksek performans sergilemektedir [15-16]. PID denetleyiciler doğrusal denetim yöntemleri altında yer almaktadır, EHS ise doğrusal olmayan bir sistemdir. Bu durum PID ile EHS denetiminde dikkat edilmesi gereken bir husustur. PID denetleyiciyi basamak veya rampa veya sinüs gibi tek bir referans sinyali takip edecek şekilde optimize etmek bu durumda tercih edilen bir yaklaşımdır. Böylelikle, doğrusal denetleyici ile doğrusal olmayan bir sistemi denetlerken ortaya çıkabilecek denetim performansı kaybı en aza indirgenebilir. Fakat bu çalışmadaki gibi birden çok sayıda (rampa, basamak ve sinüs) referans sinyali takibi gerçekleştirilmek istendiğinde PID denetleyicinin denetim performansı gerileyecektir. Bu noktada ise veriye dayalı bir yöntem olan pekiştirmeli öğrenme tabanlı denetleyiciler potansiyel gösteren bir yöntem olarak öne çıkmaktadır. Pekiştirmeli öğrenmenin doğrusal olmayan farklı sistemlerdeki başarılı uygulamaları bu önermeyi destekler niteliktedir [17-19].

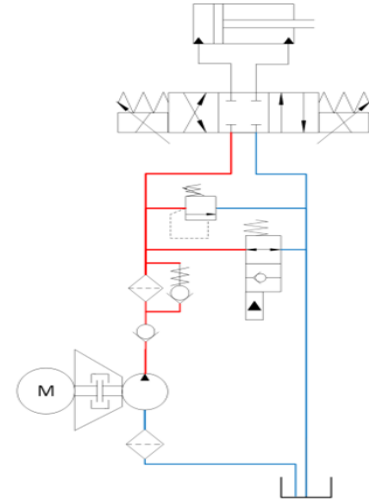
Yapılan bu çalışmada geleneksel bir EHS'nin konum denetimi benzetim ortamında pekiştirmeli öğrenme kullanılarak gerçekleştirilmiştir. Sürekli çözüm uzayına uygunluğu sebebiyle de pekiştirmeli öğrenme içerisinde yer alan derin deterministik politika gradyanı (DDPG) yaklaşımı kullanılmıştır. Bu amaçla, öncelikle EHS'nin doğrusal olmayan matematiksel modeli elde edilmiştir. Sonrasında DDPG yönteminin hiper parametreleri model üzerinde yapılan öncül çalışmalarla belirlenmiştir. Belirlenen hiper parametreler kullanılarak DDPG denetleyicinin eğitimi çoklu konum referans sinyali kullanılarak gerçekleştirilmiştir. DDPG denetleyicinin ortaya koyduğu performansı kıyaslayabilmek için aynı çalışma PID denetleyiciyle tekrarlanmıştır.

Çalışma giriş kısmıyla beraber beş kısım olarak planlanmıştır. Bu plana bağlı olarak, EHS'nin doğrusal olmayan modelinin eldesine ikinci kısımda değinilmiştir. Üçüncü kısımdaysa DDPG yöntemi anlatılarak çalışma için belirlenen hiper parametreler verilmiştir. DDPG ajanının eğitimi için yapılan çalışmalar ve gerçekleştirilen konum referans takibinin sonuçları ise dördüncü kısımda paylaşılmıştır. Son kısımda ise elde edilen sonuçlar

değerlendirilmiş ve gelecek çalışmalar için önerilere yer verilmiştir.

2 Elektrohidrolik sistem modeli

Bu çalışmada matematiksel modeli elde edilen EHS temel olarak sabit hızlı-sabit deplasmanlı dişli pompa, 4/3 oransal valf, diferansiyel hidrolik silindir ve diğer yardımcı elemanlardan oluşmaktadır (Şekil 1). Pompadan gelen hidrolik akışkanın basıncı basınç emniyet valfi ile ayarlanmaktadır. Sonrasında akışkan oransal valf yardımıyla hidrolik silindirin 1 veya 2 nolu odasına yönlendirilerek pistonun uzama veya çekilme yönünde hareket etmesi sağlanmaktadır. Oransal valfin içerisinde yer alan makaranın (Ing. spool) konumuna bağlı olarak valf içerisindeki açıklıkların alanı değişmektedir. Böylelikle, hidrolik pistonun yönü ve hızı denetlenebilmektedir.



Şekil 1. 4/3 oransal valf ile sürülen geleneksel bir EHS'nin şematik gösterimi (Kırmızı = Yüksek basınç hattı, Mavi = Düşük basınç hattı)

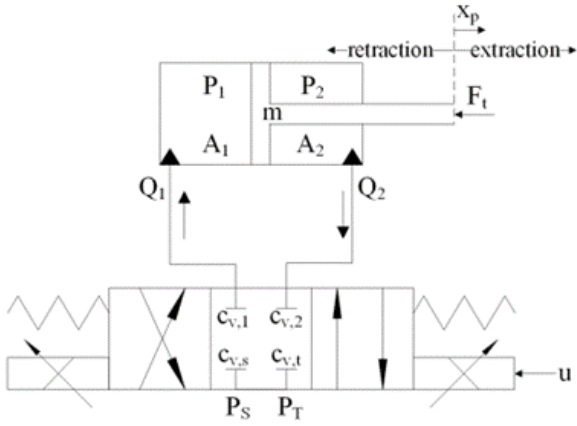
Çok karmaşık bir model ortaya çıkaracağı için EHS'nin eksiksiz bir matematiksel modelini elde etmek tercih edilen bir yaklaşım değildir. Buradaki amaç denetime etki eden baskın dinamiklerin modelde yer almasıdır. Bu nedenle modelleme öncesinde bazı ön kabullerde bulunmak yerinde bir yaklaşım olacaktır [20]. Bu bağlamda aşağıdaki kabuller yapılmıştır.

-Hidrolik sıvının sıkıştırılabilirliği basınç ve sıcaklık değişimlerinden etkilenmeksizin sabittir,

-Besleme basıncı bütün çalışmalar süresince sabittir ve herhangi bir dalgalanma göstermemektedir,

-Hidrolik akışkanın sıcaklığı ve viskozitesi çalışma süresince sabittir.

Şekil 1'de belirtilen geleneksel EHS bütün temel ekipmanları (aktarım hatları, filtre, yardımcı valfler) içermektedir. Bu ekipmanların matematiksel modele dahil edilmesi modelin doğruluğunu arttırsa dahi modellenen aynı oranda karmaşıklaştırmaktadır. Bu yüzden literatürde benimsenen yaklaşımdan yararlanarak EHS'nin matematiksel modeli Şekil 2'de verilen şematik temel alınarak elde edilmiştir.



Şekil 2. Oransal valfle sürülen asimetrik hidrolik silindirin modelleme parametreleri

Newton'un 2.hareket denklemi kullanıldığında hidrolik silindir içerisindeki kuvvet dengesi Denklem (1)'deki gibi bulunur.

$$m\ddot{x}_p = P_1A_1 - P_2A_2 \quad (1)$$

Burada m [kg] piston kütlesi, \ddot{x}_p [m/s²] piston konumunun ikinci türevi, P_x [N/m²] hidrolik silindir x nolu bölgesindeki basınç ve A_x [m²] hidrolik pistonun x nolu bölgedeki efektif çalışma alanı olarak belirtilmiştir. Hidrolik silindire etkileyen basınç alanı herhangi bir iç ve dış kaçak olmadığı durum göz önünde bulundurulduğunda aşağıdaki gibi ifade edilmektedir.

$$\frac{V_{01} + x_p A_1}{E} \dot{P}_1 = Q_1 - A_1 \dot{x}_p \quad (2)$$

$$\frac{V_{02} - x_p A_2}{E} \dot{P}_2 = -Q_2 + A_2 \dot{x}_p \quad (3)$$

Denklem (2) ve (3)'te V_{0x} [m³] hidrolik silindirin x nolu tarafındaki başlangıç hacmini, E [N/m²] hidrolik sıvının sıkıştırılabilirliğini (Bulk modülü) ve Q_x [m³/s] ise hidrolik silindirin x nolu tarafındaki debiyi göstermektedir.

Elektrohidrolik sistemin çalışma şartları göz önünde bulundurularak $P_s > P_1 > P_T$, $P_s > P_2 > P_T$ ve $P_T = 0$ kabulleri yapılabilir. Buna ek olarak Şekil 2'de belirtilen $c_{v,\{s,t,1,2\}}$ terimlerinin birbiriyle aynı ve c_v 'ye eşit olduğu ifade edildiğinde hidrolik sıvı debisi aşağıdaki bağıntılar kullanılarak gösterilebilir.

$$Q_1 = c_v x_v [s_g(x_v) \sqrt{P_s - P_1} + s_g(-x_v) \sqrt{P_1}] \quad (4)$$

$$Q_2 = c_v x_v [s_g(x_v) \sqrt{P_2} + s_g(-x_v) \sqrt{P_s - P_2}] \quad (5)$$

Burada c_v [m³s⁻¹N^{-0.5}] valf portlarının akış katsayısını, x_v [m] valf sürgüsünün konumunu, P_s ve P_T ise sırasıyla besleme ve tank basınçlarını ifade etmektedir. Son olarak, s_g ise Denklem (4) ve (5)'te verilen debi yönünü düzenlemek

için kullanılan parçalı bir fonksiyon olup aşağıdaki biçimde ifade edilmektedir.

$$s_g = \begin{cases} 0, & x \geq 0 \\ 1, & x < 0 \end{cases} \quad (6)$$

Valf tepki süresinin elektrohidrolik sistemin genelinden çok daha düşük olması sebebiyle sistem modeline dinamik valf modeli dahil edilmemiştir [21]. Bunu yerine bir fonksiyon kullanılarak valf modellenmiştir.

$$x_v = k_v(u) \quad (7)$$

Denklem (7)'de k_v fonksiyonu u [V] denetim girdisini kullanarak valf sürgüsünün konumunu belirlemek için kullanılmaktadır. Çalışmada modellenen oransal valf kapalı merkezli ve üst üste binen (İng. overlap) sürgülüdür. Sürgü türünden dolayı meydana gelen ölü bölge denetim performansı üzerinde kötüleştirici bir etkiye sahiptir [22] ve önerilen denetim yönteminin bu süreksizliğin üstesinden gelmesi gerekmektedir.

Denklem (1), (2), (3), (4) ve (5) bir arada değerlendirildiğinde $x_1 \equiv x_p$, $x_2 \equiv \dot{x}_p$, $x_3 \equiv P_1$ ve $x_4 \equiv P_2$ olmak üzere elektrohidrolik sistemin uzay durum denklemi aşağıdaki gibi elde edilmektedir.

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \frac{1}{m}(x_3 A_1 - x_4 A_2) \\ \dot{x}_3 &= \frac{1}{V_{01} + x_1 A_1} (Q_1(x_3) - A_1 x_2) \\ \dot{x}_4 &= \frac{1}{V_{02} - x_1 A_2} (-Q_2(x_4) + A_2 x_2) \end{aligned} \quad (8)$$

Benzetim çalışması için elektrohidrolik sistemin durum uzay denklemleriyle beraber Tablo 1'de belirtilen parametreler kullanılmıştır.

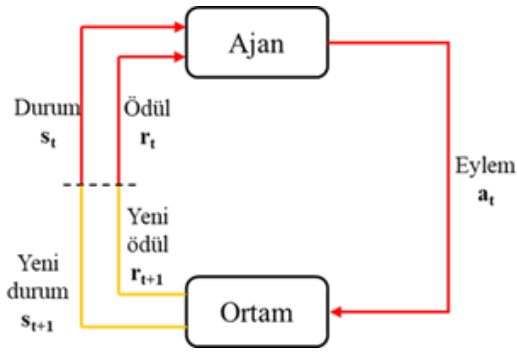
Tablo 1. Elektrohidrolik sistemin model parametreleri

| | |
|--|---------------------------------------|
| Besleme basıncı, P_s | 30 bar |
| Tank basıncı, P_T | 0 bar |
| Sıkıştırılabilirlik modülü, E | 1.5e ⁹ N/m ² |
| 1 nolu bölgenin başlangıç basıncı, P_{01} | 13.5 bar |
| 2 nolu bölgenin başlangıç basıncı, P_{02} | 18 bar |
| 1 nolu bölgedeki efektif piston alanı, A_1 | 1256.64e ⁻⁶ m ² |
| 2 nolu bölgedeki efektif piston alanı, A_2 | 942.48e ⁻⁶ m ² |
| 1 nolu bölgenin başlangıç hacmi, V_{01} | 6.594e ⁻⁵ m ³ |
| 2 nolu bölgenin başlangıç hacmi, V_{02} | 6.594e ⁻⁵ m ³ |

3 Pekiştirmeli öğrenme

Makine öğrenmesi veya yapay öğrenme veriye dayalı olarak öğrenmeyi mümkün kılan algoritmalara verilen genel bir isimdir. Kendi içerisinde gözetimli (İng. supervised), gözetimsiz (İng. unsupervised) ve pekiştirmeli (İng. reinforcement) öğrenme gibi alt dallara ayrılmaktadır. Bu çalışmada kullanılan pekiştirmeli öğrenme yöntemi ise ödül tabanlı öğrenme yapısı sayesinde denetim problemlerinde yaygın olarak kullanılmaktadır.

Pekiştirmeli öğrenme temel olarak ortamdan alınan geri bildirimlere bağlı olarak çıktı oluşturan bir fonksiyonun veri kullanılarak eğitilmesini amaçlar. Yapısal olarak incelendiğinde ise 3 temel bileşenden meydana geldiği görülür: Ajan (*İng.* agent), durum (s_t , *İng.* state) ve ortam (*İng.* environment) [23]. Öğrenilen fonksiyon politika (μ , *İng.* policy) olarak adlandırılır. Politika gözlenen durum (s_t) karşısında uzun vadede en yüksek ödül (r_t , *İng.* reward) getirecek olan aksiyonun (a_t) seçilmesini sağlar [24]. Belirlenen politika tercih edilen pekiştirmeli öğrenme yöntemine göre stokastik veya deterministik olabilir. Pekiştirmeli öğrenmenin eğitilmesi için kullanılan genel yapının gösterimi Şekil 3'te verilmiştir.



Şekil 3. Pekiştirmeli öğrenme için kullanılan eğitim yapısı [24]

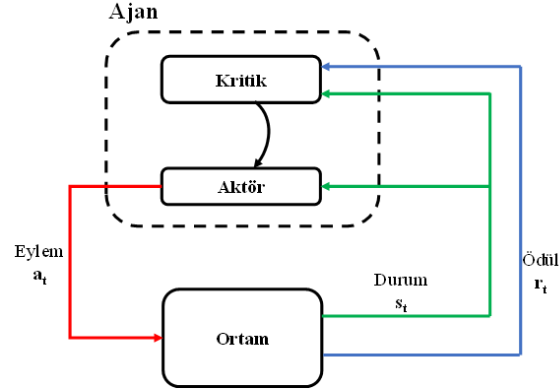
Pekiştirmeli öğrenme kendi içerisinde de sürekli (*İng.* continuous) ve süreksiz (*İng.* discrete) çözüm uzayı içeren sistemler için farklı algoritmalara sahiptir. Bu çalışmada modellenen EHS ± 10 aralığında süreklilik gösteren bir denetim sinyali ile denetlenmektedir. Bu nedenden dolayı sürekli çözüm uzayında etkinliği bilinen derin deterministik politika gradyanı algoritması EHS'nin denetimi için tercih edilmiştir.

3.1 Derin deterministik politika gradyanı

Derin deterministik politika gradyanı (Deep Deterministic Policy Gradient, DDPG) temel olarak deterministik politika gradyanı (Deterministic Policy Gradient, DPG) ile derin Q-ağı (Deep Q-Network, DQN) yöntemlerinin birleşimidir ve aktör-kritik (*İng.* actor-critic) tabanlı yöntemler içerisinde yer alır [25]. Süreksiz aksiyon uzayında yüksek performans gösteren DQN'in sürekli aksiyon uzayını kapsayacak biçimde genişletilirken DPG kaynaklı kararsızlık sorunları giderilerek bu yöntem ortaya koyulmuştur. Yöntemin kararlı ve gürbüz olması için kendisini oluşturan iki yöntemin özelliklerinden (i) bellekten tekrarlamaya (*İng.* replay buffer) ve (ii) çıktı üreten ağıncı güncellenmesinin ikincil bir Q ağıncının kullanılması yaklaşımları korunmuştur [25]. Bellekten tekrarlamaya ile aktör ağıncının eğitim esnasında birbirini takip eden örneklemeleri kullanarak ezberleyici bir tutum sergilemesinin önüne geçilmiş ve ikincil bir ağıncı yardımıyla eğitimin hedeflenen doğrultuda yüksek değişimlerden etkilenmeden sürmesi sağlanmıştır.

DDPG'nin de içinde yer aldığı aktör – kritik yaklaşımı aktör yapay sinir ağıncının (YSA) durum karşısındaki

eyleminin (a_t) kritik YSA tarafından değerlendirilmesi üzerine kuruludur. Kritik YSA değerlendirme amacıyla ortamdan gelen durum (s_t) bilgisini ve gerçekleştirilen eylem sonucunda elde edilen ödül (r_t) sürekli olarak izler. Elde ettiği bilgileri kullanarak aktör YSA'nın başarımlarını hesaplar. Hesaplanan gradyan ve geri dağılım (*İng.* back-propagation) algoritmasıyla aktör YSA'nın ağırlık ve bias değerlerinin değişimi gerçekleştirilir. Aktör – kritik tabanlı pekiştirme öğrenme yöntemlerinin genel işleyişi Şekil 4'te verilmiştir.



Şekil 4. Aktör-kritik tabanlı pekiştirmeli öğrenme yöntemlerinin genel işleyişi [24]

DDPG yönteminde aktör – kritik yaklaşımı 2 yerine 4 farklı YSA ile sağlanmaktadır. Bunun nedeni ise aktör YSA'nın eğitiminde ana girdi olan gradyanın tutarlı bir şekilde hesaplanabilmesinin istenmesidir. Bu amaçla aktör ve kritik YSA'larla aynı ağırlık ve bias değerlerine sahip 2 adet YSA oluşturulur: Aktör hedef (*İng.* actor target) YSA ve kritik hedef (*İng.* critic target) YSA. Böylelikle, optimize edilen aktör ve kritik YSA'lar kendi başarımlarını ölçmek için kullanılmaz ve YSA'ların ani olarak değişmesinin önüne geçilmiş olur [25]. Belirlenen sıklıkla hedef YSA'ların ağırlıkları seçilen öğrenme oranı kullanılarak güncellenir. DDPG algoritmasının işleyişi Şekil 5'te verilmiştir.

DDPG için gerekli parametrelerinin belirlenmesi amacıyla benzer problemler incelenmiş ve elde edilen bilgiler başlangıç noktası olarak kullanılmıştır. Sonuç olarak elde edilen DDPG hiper parametreleri Tablo 2'de ve görev başarımlarını ölçmek için geliştirilen ödül fonksiyonu Denklem (9)'da verilmiştir.

Tablo 2. DDPG hiper parametreleri.

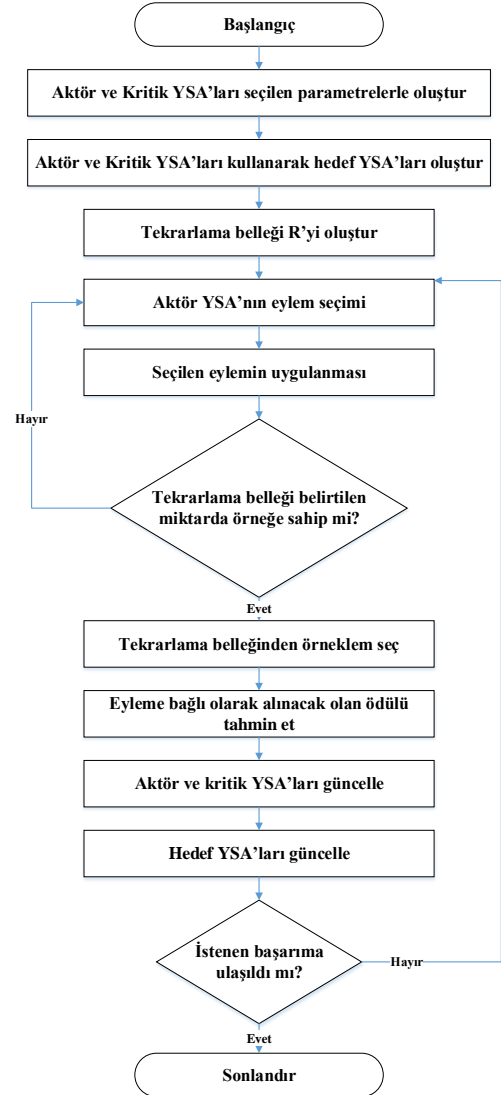
| Hiper Parametre | Değer |
|--|-------|
| Aktör YSA öğrenme oranı | 1e-4 |
| Kritik YSA öğrenme oranı | 1e-3 |
| Aktör YSA'nın örneklem zamanı | 1e-2 |
| Tekrarlamaya belleği boyutu | 1e6 |
| Mini örneklem yığıncı boyutu | 128 |
| İndirim oranı | 0,99 |
| Orstein-Uhlenbeck (O-U) gürültü varyansı | 1e-2 |
| O-U gürültü varyansı azalma oranı | 1e-5 |

$$\begin{aligned}
 P_1 &= \min\left(0, -\frac{d}{dt}|e(t)|\right) \\
 P_2 &= -20|e(t)| \\
 P_3 &= -\frac{3}{20}\left|\frac{d}{dt}u(t-1)\right| \\
 R_1 &= \frac{1}{(\max(0.2, |e(t)|))^2} \\
 R &= \sum (P_1 + P_2 + P_3 + R_1)
 \end{aligned} \quad (9)$$

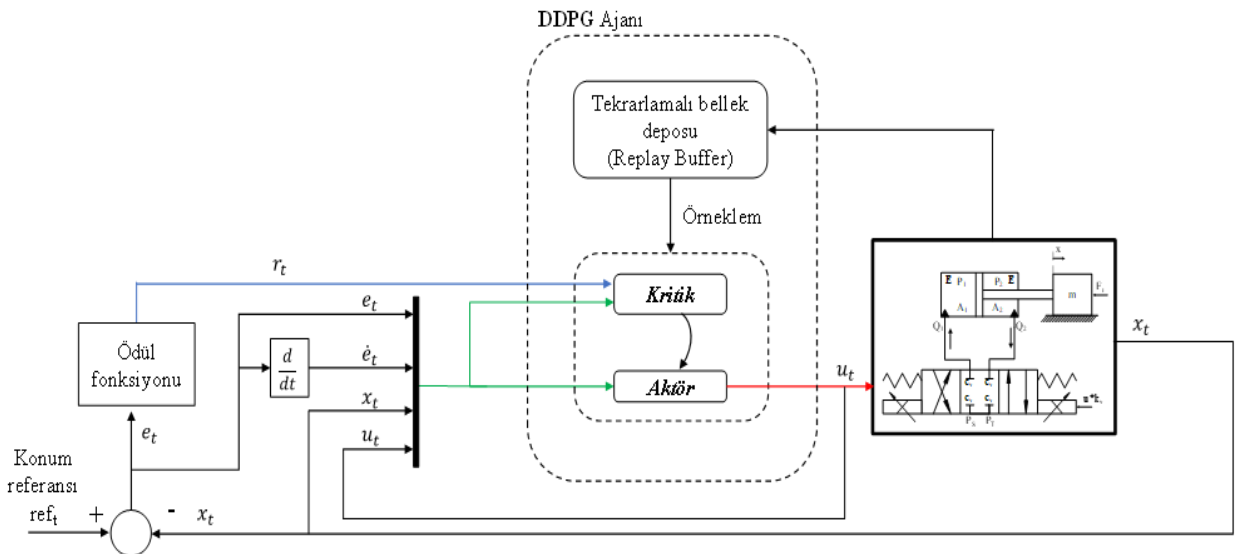
Gözlemlenerek eğitimde kullanmak üzere DDPG'ye beslenecek durum değerleri olarak ise konum takip hatası (e_t), konum takip hatasının türevi (\dot{e}_t), denetim sinyali (u_t) ve piston konumu (x_t) seçilmiştir (Şekil 6). Verilen bilgiler doğrultusunda oluşturulan aktör ve kritik yapay sinir ağları (YSA) Şekil 7 ve Şekil 8'de verilmiştir.

4 Benzetim çalışmaları

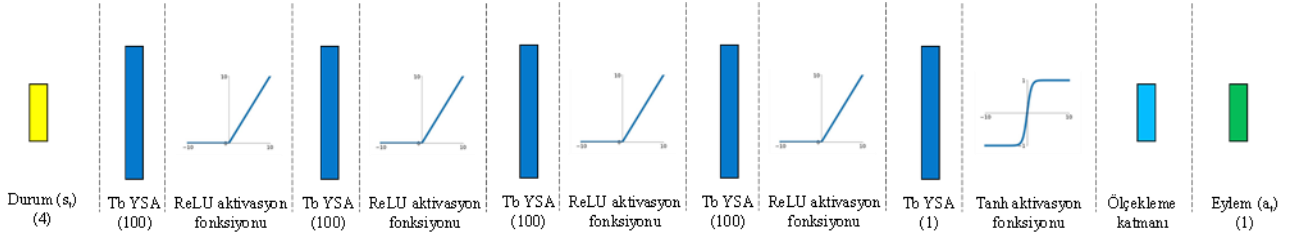
EHS'lere yönelik denetim çalışmaları incelendiğinde temel olarak tek bir referans sinyalinin kullanıldığı görülmektedir. Bunun nedeni ise EHS'lerin doğrusal olmayan yapısı itibarıyla denetleyicilerin her referans için ayrı ayrı tasarlanması gerekliliğidir. Bu çalışmada önerilen yöntem ise aynı anda farklı karakteristiklerdeki konum referans sinyallerini takip edebilecek potansiyele sahiptir. Bu sebeple hiper parametreleri ve YSA yapısı tanımlanan DDPG denetleyici rampa, sinüs ve basamak olmak üzere 3 farklı konum referans sinyalini takip etmesi için eğitilmiştir. Benzetim ortamında gerçekleştirilen eğitim için zaman adımı 0.01 s olarak seçilmiştir (Şekil 9). Şekil 9'da 180. bölüme kadar DDPG denetleyicinin aldığı ödülün aynı seviyede kaldığı görülmektedir. Sonrasında ise ödül istenen düzeye kadar yükselmektedir. 300. bölümden sonra ise DDPG denetleyici genel olarak birbirine yakın ödüller almayı başarmıştır. DDPG denetleyicinin eğitimi 527 bölüm ve 900000 adım sürmüştür.



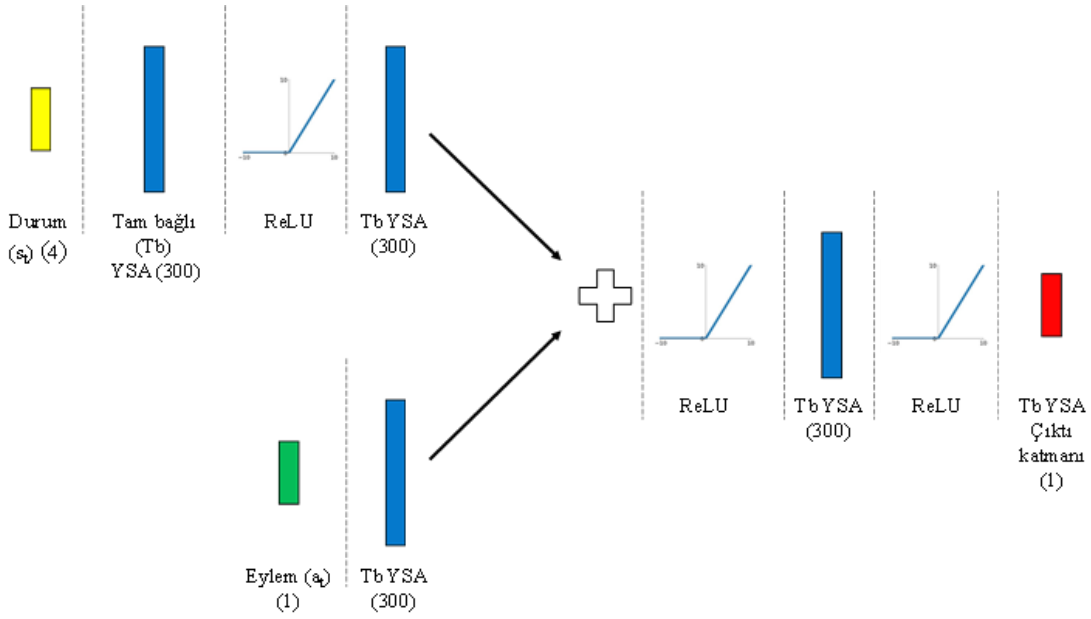
Şekil 5. DDPG yaklaşımının akış diyagramı [26]



Şekil 6. DDPG ile geleneksel hidrolik sistemin eğitiminde kullanılan benzetim ortamı. Kırmızı çizgi ajanın ürettiği eylemi, mavi çizgi eylem sonucunda alınan ödülü ve yeşil çizgi ise DDPG içindeki aktör – kritik YSA'larına iletilen durum verisini göstermektedir



Şekil 7. Aktör YSA olarak kullanılan ağ yapısı

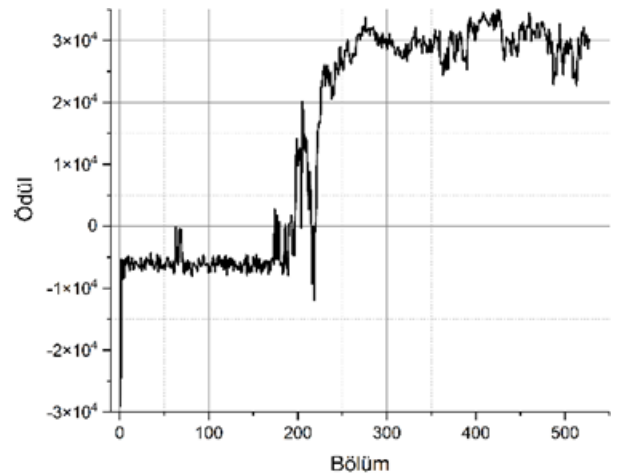


Şekil 8. Kritik YSA olarak kullanılan ağ yapısı

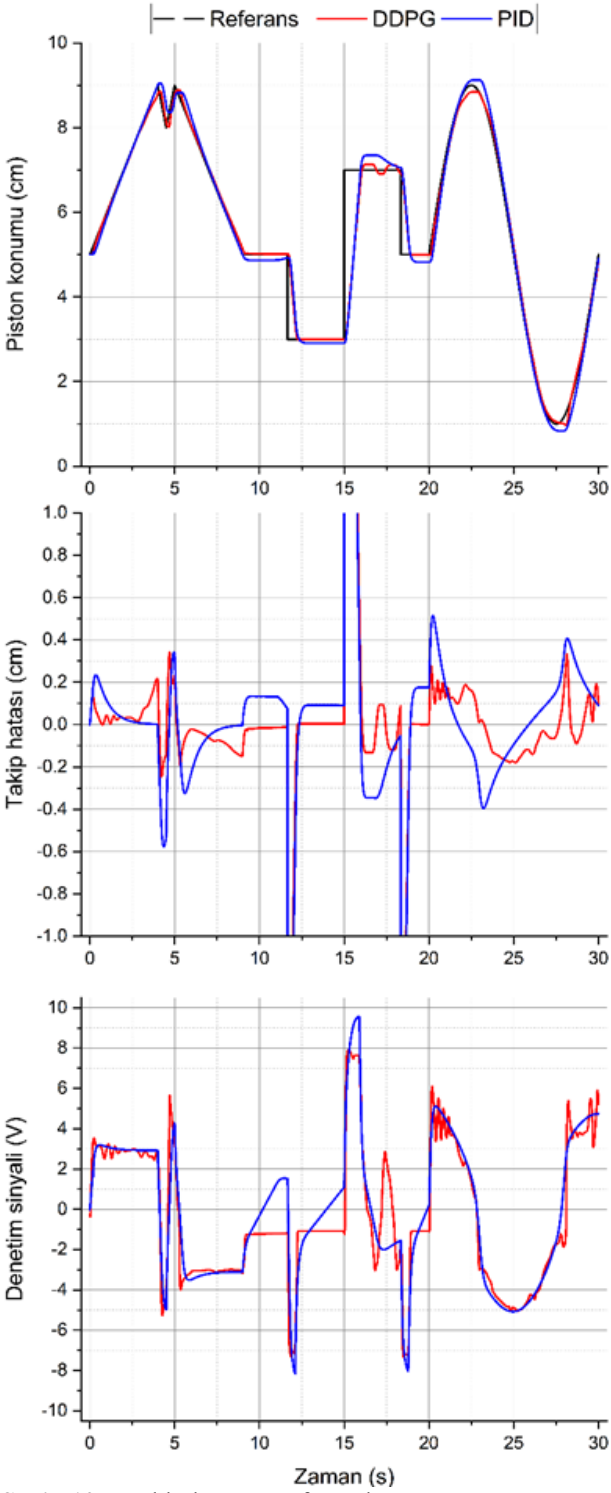
Bu çalışmada denetim performansı sunulan denetleyici eğitim süresince en başarılı takip performansının sergilendiği 400-425 bölümleri arasındaki kısımdan seçilmiştir. Buna ek olarak Ziegler-Nichols yöntemi ile parametreleri belirlenmiş bir PID denetleyici de DDPG denetleyici ile kıyaslamak için kullanılmıştır. Ziegler-Nichols ile belirlenen PID denetleyici parametreleri denetim çıktısının ± 10 V sınırları içerisinde kalmasını sağlayacak biçimde düzenlendiğinde $P = 10.2$, $I = 11.8$ ve $D = 0.8$ denetleyici parametrelerine ulaşılmıştır.

DDPG ve PID denetleyicilerle gerçekleştirilen konum takip çalışmasından elde edilen sonuçlar Şekil 10'da verilmiştir. DDPG denetleyici rampa konum referansı ($t = 0-9$ s) takibinde PID denetleyiciden daha iyi performans göstermiştir. Her iki denetleyicinin rampa sinyalinin yön değiştirdiği ($t = 4.5 - 5.5$ s) kısımlardaki takip hatası incelendiğinde bu durum daha belirginleşmektedir. PID denetleyicinin takip hatası ± 0.6 cm seviyesindeyken DDPG denetleyici hatayı ± 0.2 cm civarında tutmayı başarmıştır. DDPG denetleyici başarımını basamak konum referansı ($t = 9-20$ s) takibinde de sürdürmektedir. PID denetleyici basamak sinyali takibinde kararlı hal hatası (> 0.1 cm) ve aşma sorunları sergilerken DDPG denetleyici kararlı hal hatasını (< 0.02 cm) ortadan kaldırmaktadır. PID denetleyiciyle %12.5 aşma meydana gelen kısımda ise DDPG denetleyici aşmayı %2.5 seviyesine indirmektedir. Benzer yaklaşımla denetleyicilerin sinüs konum sinyali ($t =$

20-30 s) takip performansı incelendiğinde, önceki sonuçlarla uyumlu olarak, DDPG denetleyici öne çıkmaktadır. Yapısı itibarıyla sinüs sinyalinin eğimi sürekli değişmektedir. DDPG denetleyici bu değişime uyum sağlayarak takip hatasını çoğunlukla ± 0.2 cm aralığında tutmaktadır. PID denetleyicinin takip performansı incelendiğinde ise takip hatası bandının ± 0.4 cm'e genişlediği görülmektedir.



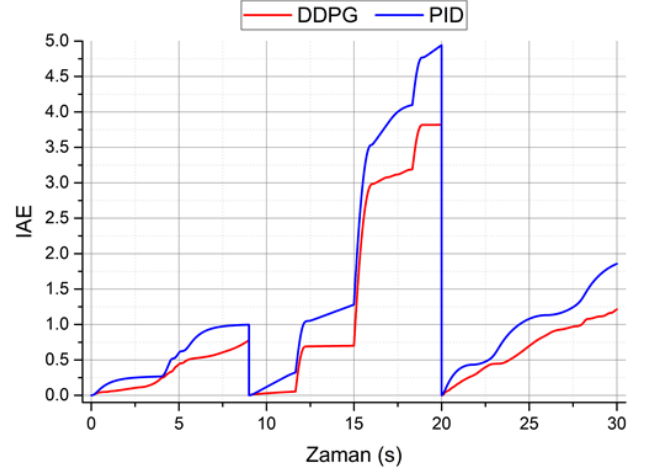
Şekil 9. DDPG denetleyicinin eğitim sürecindeki başarımı aldığı ödülle birlikte artmaktadır



Şekil 10. Farklı konum referanslarının DDPG ve PID denetleyiciler ile takibi

Denetleyicilerin performansının hatanın mutlak değer toplamına (IAE, integral absolute error) göre karşılaştırması Şekil 11’de verilmiştir. IAE değeri her referans sinyali değişiminde sıfırlanmıştır. Bu bağlamda, 0-9 saniye aralığında rampa, 9-20 saniye aralığında basamak ve 20-30 saniye aralığında ise sinüs konum referans sinyallerine ait IAE değerleri zamana bağlı olarak çizilmiştir. İlgili sayısal değerler Tablo 3’te de ifade edilmiştir. Bu bağlamda,

geliştirilen DDPG denetleyicinin IAE performans kriterine göre PID denetleyiciden rampa sinyal takibinde %22.17, basamak sinyal takibinde %22.74 ve son olarak sinüs sinyal takibinde %34.68 daha iyi sonuç verdiği görülmektedir.



Şekil 11. DDPG ve PID denetleyicilerin konum referans sinyali takibinin IAE performans kriterine göre değerlendirilmesi

Tablo 3. IAE performans kriterine göre DDPG ve PID denetleyicilerin konum takibi performansı

| | DDPG | PID | % |
|---------|-------|-------|-------|
| Rampa | 0.776 | 0.997 | 22.17 |
| Basamak | 3.818 | 4.942 | 22.74 |
| Sinüs | 1.213 | 1.857 | 34.68 |
| Toplam | 5.807 | 7.796 | 25.51 |

5 Sonuçlar

EHS’nin konum denetim için gerçekleştirilen bu çalışmada rampa, basamak ve sinüs olmak üzere 3 farklı konum referans sinyalinin takibi bir pekiştirmeli öğrenme yöntemi olan DDPG ile gerçekleştirilmiştir. Bununla beraber, elde edilen sonuçların karşılaştırılabilmesi için Ziegler-Nichols yöntemiyle parametreleri belirlenmiş PID denetleyici de çalışmaya dahil edilmiştir. Elde edilen sonuçlar, kısa sayılabilecek bir eğitime rağmen, DDPG denetleyicinin belirlenen bütün konum referanslarının takibinde PID denetleyiciden genel olarak daha iyi performans sergilediğini göstermiştir. DDPG denetleyici doğrusal olmayan yapısı ve modellenen sistemin dinamiklerini öğrenebilmesi sayesinde bu performansa ulaşabilmiştir.

Bu çalışmada DDPG denetleyicinin sistem parametrelerinin değişimine ve sensor gürültüsüne karşı dayanımı göz ardı edilmiştir. İlerleyen çalışmalarda bu hususlar üzerine yoğunlaşarak kısa sürede sonuç verebilen YSA yapısı ve ödül fonksiyonun geliştirilmesi planlanmaktadır. Diğer bir husus ise DDPG denetleyicinin etkinliğinin deneysel olarak gözlenmesidir. Benzetim çalışması için geliştirilen EHS modeli baskın sistem dinamiklerini yansıtabilmede fakat deneysel sistemde yer alan bütün etkileri aktarmakta eksik kalabilmektedir. Önerilen DDPG denetleyicinin deneysel olarak sınanması yöntemin yetkinliğiyle ilgili daha kapsamlı bilgiler

verecektir. Son olarak pekiştirmeli öğrenme tabanlı denetleyicilerin üstünlüğü sistem modeline ihtiyaç duymadan sadece veri tabanlı olarak çalışabilmesidir. Bu durum ise pekiştirme öğrenme tabanlı denetleyicilerin deneysel olarak eğitimini mümkün kılmaktadır. İlerleyen çalışmalarda DDPG denetleyicinin gerçek zamanlı olarak eğitilmesi bu yöntemin pratik ve hızlı bir biçimde sahadaki uygulamalarda kullanımının önünü açabilecek potansiyele sahiptir.

Çıkar çatışması

Yazarlar çıkar çatışması olmadığını beyan etmektedir.

Benzerlik oranı (iThenticate): %5

Kaynaklar

- [1] J. Činkelj, R. Kamnik, P. Čepon, M. Mihelj and M. Munih, Closed-loop control of hydraulic telescopic handler. *Automation in Construction*, 19, 954–963, 2010. <https://doi.org/10.1016/j.autcon.2010.07.012>
- [2] M. Borghi, B. Zardin, F. Pintore and F. Belluzzi, Energy savings in the hydraulic circuit of agricultural tractors. *Energy Procedia*, Elsevier B.V. 45, 352–361, 2014. <https://doi.org/10.1016/j.egypro.2014.01.038>
- [3] T. Boaventura, J. Buchli, C. Semini and D.G. Caldwell, Model-Based hydraulic impedance control for dynamic robots. *IEEE Transactions on Robotics*, IEEE. 31, 1324–1336, 2015. <https://doi.org/10.1109/TRO.2015.2482061>
- [4] A. Bayrak, E. Tatlicioglu and E. Zergeroglu, Backstepping control of electro-hydraulic arm. 2018 6th International Conference on Control Engineering and Information Technology, CEIT 2018, IEEE. pp. 1-4, 2018. <https://doi.org/10.1109/CEIT.2018.8751833>
- [5] D. Janosevic, R. Mitrev, B. Andjelkovic, and P. Petrov, Quantitative measures for assessment of the hydraulic excavator digging efficiency. *Journal of Zhejiang University: Science A*, 13, 926–942, 2012. <https://doi.org/10.1631/jzus.A1100318>
- [6] M. Heikkilä and M. Linjama, Displacement control of a mobile crane using a digital hydraulic power management system. *Mechatronics*, 23, 452–461, 2013. <https://doi.org/10.1016/j.mechatronics.2013.03.009>
- [7] M. Jelali and A. Kroll, *Hydraulic Servo-systems*. Springer London, London. 2003. <https://doi.org/10.1007/978-1-4471-0099-7>
- [8] M. Fallahi, M. Zareinejad, K. Baghestan, A. Tivay, S.M. Rezaei and A. Abdullah, Precise position control of an electro-hydraulic servo system via robust linear approximation. *ISA Transactions*, Elsevier Ltd. 80, 503–512, 2018. <https://doi.org/10.1016/j.isatra.2018.06.002>
- [9] C. Onat, M. Daskin and A. Turan, Gain scheduling linear model of an electro-hydraulic actuator. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, 7, 301–306, 2017. <https://doi.org/10.11121/ijocta.01.2017.00487>
- [10] E. Kolsi-Gdoura, M. Feki and N. Derbel, Sliding mode-based robust position control of an electrohydraulic system. 10th International Multi-Conferences on Systems, Signals & Devices 2013 (SSD13), IEEE. pp. 1–5 2013. <https://doi.org/10.1109/SSD.2013.6564127>
- [11] E. Kolsi Gdoura, M. Feki and N. Derbel, Sliding mode control of a hydraulic servo system position using adaptive sliding surface and adaptive gain. *International Journal of Modelling, Identification and Control*, 23, 248–259, 2015. <https://doi.org/10.1504/IJMIC.2015.069946>
- [12] B. Shaer, J.-P. Kenné, C. Kaddissi and C. Fallaha, A chattering-free fuzzy hybrid sliding mode control of an electrohydraulic active suspension. *Transactions of the Institute of Measurement and Control*, 40, 222–238, 2018. <https://doi.org/10.1177/0142331216652468>
- [13] M. Kalyoncu and M. Haydim, Mathematical modelling and fuzzy logic based position control of an electrohydraulic servosystem with internal leakage. *Mechatronics*, Elsevier Ltd. 19, 847–858, 2009. <https://doi.org/10.1016/j.mechatronics.2009.04.010>
- [14] L. Yu, L. Ding, F. Yu, J. Zheng and Y. Tian, Force tracking control for electrohydraulic servo system based on adaptive neuro-fuzzy inference system (ANFIS) controller. *International Journal of Intelligent Computing and Cybernetics*, 14, 1–16, 2021. <https://doi.org/10.1108/IJICC-09-2020-0132>
- [15] Y.J. Liu, Y.D. Xie and H. Wang, Fuzzy PID control for valve-controlled cylinder hydraulic system. *Applied Mechanics and Materials*, 212–213, 1244–1248, 2012. <https://doi.org/10.4028/www.scientific.net/AMM.212-213.1244>
- [16] B.K. Sarkar, P. Mandal, R. Saha, S. Mookherjee and D. Sanyal, GA-optimized feedforward-PID tracking control for a rugged electrohydraulic system design. *ISA Transactions*, Elsevier. 52, 853–861, 2013. <https://doi.org/10.1016/j.isatra.2013.07.008>
- [17] A. Rodriguez-Ramos, C. Sampedro, H. Bavle, P. de la Puente and P. Campoy, A Deep Reinforcement Learning Strategy for UAV Autonomous Landing on a Moving Platform. *Journal of Intelligent & Robotic Systems*, 93, 351–366, 2019. <https://doi.org/10.1007/s10846-018-0891-8>
- [18] X. Wu, S. Liu, T. Zhang, L. Yang, Y. Li and T. Wang, Motion Control for Biped Robot via DDPG-based Deep Reinforcement Learning. 2018 WRC Symposium on Advanced Robotics and Automation, WRC SARA 2018 - Proceeding, IEEE. 40–45, 2018. <https://doi.org/10.1109/WRC-SARA.2018.8584227>
- [19] Y. Hou, L. Liu, Q. Wei, X. Xu and C. Chen, A novel DDPG method with prioritized experience replay. 2017 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2017, 2017-Janua, pp. 316–321, 2017. <https://doi.org/10.1109/SMC.2017.8122622>
- [20] P. Wos and R. Dindorf, Modeling and Analysis of the Hydraulic Servo Drive System. *Advances in Intelligent Systems and Computing*, Springer International Publishing, 253-262 2015. <https://doi.org/10.1007/978-3-319-26886-6>

- [21] L. Lu and B. Yao, Energy-saving adaptive robust control of a hydraulic manipulator using five cartridge valves with an accumulator. *IEEE Transactions on Industrial Electronics*, 61, 7046–7054, 2014. <https://doi.org/10.1109/TIE.2014.2314054>
- [22] U. Pinsopon, T. Hwang, S. Cetinkunt, R. Ingram, Q. Zhang, M. Cobo, D. Koehler and R. Ottman, Hydraulic actuator control with open-centre electrohydraulic valve using a cerebellar model articulation controller neural network algorithm. *Proceedings of the Institution of Mechanical Engineers Part I: Journal of Systems and Control Engineering*, 213, 33–48, 1999. <https://doi.org/10.1243/0959651991540368>
- [23] S. Armoogum and X. Li, Big Data Analytics and Deep Learning in Bioinformatics With Hadoop. *Deep Learning and Parallel Computing Environment for Bioengineering Systems*, Elsevier. pp. 17–36 2019. <https://doi.org/10.1016/B978-0-12-816718-2.00009-9>
- [24] R.S. Sutton and A.G. Barto, *Reinforcement Learning: An Introduction*. The MIT Press, London, 2018.
- [25] T.P. Lillicrap, J.J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver and D. Wierstra, Continuous control with deep reinforcement learning. 2015. <https://doi.org/10.48550/arXiv.1509.02971>
- [26] S. Guo, X. Zhang, Y. Zheng and Y. Du, An autonomous path planning model for unmanned ships based on deep reinforcement learning. *Sensors (Switzerland)*, 20, 2020. <https://doi.org/10.3390/s20020426>

