

Derin Q Öğrenmeye Dayalı Akıllı Trafik Sinyal Kontrol Stratejileri

Intelligent Traffic Signal Control Strategies Based on Deep Q Learning

Ali Aras Fırat¹, Ahmet Fatih Olgunöz¹, İlhan Tunç², Mehmet Turan Söylemez¹

¹Kontrol ve Otomasyon Mühendisliği Bölümü
İstanbul Teknik Üniversitesi, İstanbul
firata18@itu.edu.tr
olgunoz18@itu.edu.tr
soylemezm@itu.edu.tr

²Mekatronik Mühendisliği Bölümü
Bursa Teknik Üniversitesi, Bursa
ilhan.tunc@btu.edu.tr

Özetçe

Trafik sıkışıklığı sorununun önümüzdeki yıllarda daha karmaşık hale gelmesi beklenmektedir. Trafik sıkışıklığı problemlerinin en yoğun şekilde yaşandığı alanlar trafik kavşaklardır. Bu sorunu hafifletmek için bir kavşak kontrol yöntemi gereklidir. Bu çalışmada, kavşaklardaki trafik sıkışıklığını ortadan kaldırmak için üretilen akıllı kontrol sistemlerinde kullanılmak üzere gündemde olan Derin Q Öğrenme tekniğinin çeşitli varyantları denenmiştir. Bu teknik ile tek bir kavşak için elde edilebilecek optimal ajan belirlenmeye çalışılmıştır. Bu amaçtan yola çıkılarak ajanın durum parametreleri, durum sayısı, ödül fonksiyonu seçimi üzerinde çeşitli kıyaslamalar yapılmıştır ve her bir kombinasyon ayrı bir strateji olarak belirlenmiştir. Ele alınan bu varyantlar birbirleri ile toplam bekleme süresi, ortalama kuyruk uzunluğu, CO_2 emisyonu ve kavşaktaki araçların ortalama uzunluğu cinsinden kıyaslanmıştır ve en iyi sonuçları veren stratejiler belirlenmiştir.

Abstract

The traffic congestion problem is expected to become even more complex in the coming years. Traffic congestion problems usually occur in traffic intersections. An intelligent traffic intersection control method can mitigate these problems. In this study, various intelligent controllers that rely on Deep Q Learning technique have been developed and simulated, with the aim of reducing or eliminating the traffic congestion at a single intersection. To determine an optimal controller for a single intersection, different properties for the agent were tested, such as state parameters, number of states, and the selection of the reward function. Obtained variants are then compared with each other in terms of total waiting time, average queue length, CO_2 emissions and average queue length of vehicles at the intersection, to determine the best strategy for the control method properties.

1. Giriş

Günümüzün en büyük sorunlarından biri, nüfusu yüksek kentlerdeki trafik yoğunluğu problemidir. Trafik yoğunluğu ciddi boyutlarda zaman ve para kaybına neden olmaktadır ve en büyük sebepleri arasında kazalar ve bireysel araç kullanımının fazlalığı gibi insan kaynaklı faktörlerin yanında, trafik kavşaklarının teknik limitleri ve bu kavşakların kontrol yöntemlerinin yetersizliği gösterilebilir [1].

Bu çalışma, trafik kavşaklarında kullanılabilecek potansiyel akıllı sistemleri incelemektedir. Bahsedilen akıllı trafik sinyal sistemleri, sensörler ve kameralar yardımıyla kavşağa gelmekte veya kavşakta beklemekte olan araçları tespit eden, daha sonra ise çeşitli kontrol yöntemleriyle trafik ışığını kontrol ederek kavşakta oluşabilecek trafik yoğunluğunu azaltmaya çalışan sistemlerdir. Uzun yıllardır kullanılan klasik, sabit zamanlı trafik ışıkları yerine, akıllı yöntemlerle kontrol edilen, bazı yenilikçi kavşak kontrolü yöntemleri incelenmiş ve bu yöntemler kendi aralarında kıyaslanmıştır.

Akıllı kavşak kontrol yöntemleriyle ilgili literatür taraması gerçekleştirildiğinde, konunun incelenmesinde çeşitli konseptlerin ve stratejilerin kullanıldığı görülmektedir. Kullanılan yöntemlerdeki farklılıklara rağmen genel bakış açısı, kavşak sistemini bir kontrol sistemi ile kapalı çevrime alarak, araç kuyruk uzunluğu gibi parametrelerin hata sinyali olarak, yeşil ve kırmızı ışık sürelerinin ise kontrol sinyali olarak belirlenmesi yönündedir. 2015 yılında yapılan bir çalışmada, geri beslemeli kontrol sistemleri bakımından basit sayılabilecek, PID kontrolcüye dayanan kavşak kontrol yönteminin dahi, sabit zamanlı klasik kavşaklara göre belirgin bir şekilde avantajlı sonuçlar verdiği görülmüştür. [2] 2014' te yayımlanan bir başka makalede, çeşitli bulanık mantık kontrolcülere denenmiştir, bulanık mantığa dayanan kontrolcülere kavramak bakımından önemli detaylara yer veren makale, yine bu akıllı yöntemlerin en uygun sonuçlar verdiğini göstermektedir. [3] Benzer bir yaklaşımla, 2019' da bir çalışmada, tek kavşaklı bir ortamda, bulanık mantık ile kontrol edilen kavşak, PI kontrolöre sahip sistem ile kıyaslanmıştır ve

bulanık mantığa dayanan sistemin daha iyi sonuç verdiği görülmüştür [4].

Bulanık mantık temelli kontrolcülerin, trafik sinyali için kullanıldığı bir başka çalışmada, iki farklı strateji ile çalışan kontrolcü incelenmiş ve bu stratejilerden birinin düşük trafik yoğunluğunda, diğerinin ise daha yüksek trafik yoğunluğunda avantajlı sonuçlar verdiği görülmüştür. [5] Bu bakımdan tasarlanan kontrolcülerin parametreleri belirlenirken, uygulanacak kavşağın ortalama yoğunluk seviyesinin de dikkate alınmasının önemi görülmektedir. Bir başka çalışma, durum geri beslemeli kontrolörlerin trafik kavşaklarında uygulanmasını incelemiş ve belirli koşullarda bulanık mantık temelli kontrolcülerden daha iyi sonuçlar elde edilebildiğini, aynı zamanda hem durum geri beslemeli hem de bulanık mantık temelli kontrolcülerin, sabit zamanlı, klasik trafik sinyali yöntemine göre belirgin şekilde avantajlı olduğunu göstermiştir [6].

Gerçek hayat uygulamalarında, yayaların da trafik ışıklarıyla kontrol edilen kavşaklarda önemli bir rolü vardır. 2022 yılında yapılan bir çalışmada, bahsedilen akıllı trafik kavşakları ile, yayaların etkileşimi incelenmiştir. Derin Q Öğrenme ağı mimarisi ile tasarlanan kontrolcü ile, yayaların da dahil edildiği bir ortamda, diğer yöntemlere kıyasla olumlu sonuçlar elde edilmiştir. [7] 2023 yılında Tunç ve Söylemez tarafından yapılan çalışma, Derin Q Öğrenme ve bulanık mantık temelli kontrolcülerini birleştirerek karmaşık bir sistem ortaya koymuştur. Bu şekilde bir arada kullanılan yöntemlerin doğru planlandığı takdirde, tek yöneme dayanan sistemlere göre öne çıktığı gösterilmiştir [8].

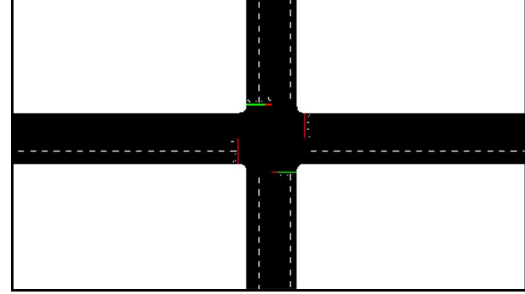
Akıllı kavşak kontrolü sistemleri, tek kavşaktan oluşan ulaşım yapılarıyla sınırlı değildir. Çok kavşaklı sistemler için, hem bahsedilen yöntemlerin, hem de yeni yöntemlerin tasarlanıp denendiği çalışmalar mevcuttur. Bu çalışmalardan biri, bahsedilen Derin Q Öğrenme modelini, çoklu kavşaklı sistemde, kavşaklar arasında Q değeri transferi yöntemini de katarak tasarlanmış ve kavşaklar arasında olacak bir haberleşme yönteminin faydalarını göstermiştir. [9] Derin Q Öğrenme yönteminin yanında, daha alışılmış olan derin pekiştirmeli öğrenme yöntemi de kavşak kontrol sistemlerinde kullanılabilir. Derin Q Öğrenme yönteminin avantajı, kavşaklar arası haberleşme için Q değeri transferi yöntemine olanak sağlaması ve bu yöntemin, haberleşme yöntemleri arasında nispeten daha az işlem yükü gerektirmesidir [10].

Bu çalışmada Derin Q öğrenme temelli trafik ışık kontrolörü önerilmiştir. Derin Q öğrenme trafik ışık kontrol metodunda farklı durum değerleri ve farklı ödül değerleri test edilmiş ve sonuçlar karşılaştırılmıştır. SUMO (Simulation of Urban MObility) programı vasıtası ile benzetim ortamı oluşturulmuş ve sonuçlar test karşılaştırılarak test edilmiştir. Bu bildirinin kalan kısmının organizasyonu şu şekildedir. İkinci bölümde Trafik ışık kontrol sisteminden bahsedilmiştir. Üçüncü ve dördüncü bölümlerde ise sırası ile Derin Q öğrenme yönteminden ve benzetim sonuçları açıklanmıştır. Son olarak dördüncü bölümde sonuçlar tartışılmış ve yorumlar yapılmıştır.

2. Trafik Işık Kontrol Sistemi

Trafik ışık benzetimi için SUMO ortamında tasarlanan kavşak Şekil 1’de verilmiştir. SUMO, gerçek bir haritadan içe aktarılan ağlar üzerinde çalışabilen, küçük ölçekli çeşitli çalışma alanları sağlayan, trafikteki çeşitli araçların yanı sıra aynı zamanda yaya simülasyonuna da olanak tanıyan, erişilebilir açık kaynaklı bir simülasyon paketidir. Her kavşakta 4 şerit vardır, sola dönüş

için bu şeritlerden en soldaki kullanılırken düz gidiş için sağdaki 3 şerit, sağa dönüş için ise en sağdaki şerit kullanılmalıdır. Trafik ışık kontrol sistemi, ışıkların kırmızı ve yeşil yanma sürelerini Derin Q Öğrenme ile kontrol edecektir. Modellenen trafik kavşağının sınırlamaları ve varsayımları Şekil 1’deki gibidir.



Şekil 1: Benzetim ortamı

Bu sistemde, karşıya direkt geçişler ile sağa dönüşlerin sekans içinde aynı anda olduğu ve sola dönüşlerin sekans içinde farklı fazlarda gerçekleştiği senaryo ele alınmıştır. Sarı ışık süresi 3 saniyedir. Kıyaslama açısından klasik trafik ışık sisteminin fazları Tablo 1’de verilmiştir:

Tablo 1: Trafik ışık fazları

FAZ	D	DS	K	KS	B	BS	G	GS
1	Y	K	K	K	Y	K	K	K
2	S	K	K	K	S	K	K	K
3	K	Y	K	K	K	Y	K	K
4	K	S	K	K	K	S	K	K
5	K	K	Y	K	K	K	Y	K
6	K	K	S	K	K	K	S	K
7	K	K	K	Y	K	K	K	Y
8	K	K	K	S	K	K	K	S

D: Doğu, B: Batı, K: Kuzey, G: Güney, S: Sola Dönüş

K-S-Y: Kırmızı, Sarı, Yeşil

3. Derin Q Öğrenme ile Trafik Işık Kontrolü

Derin Q Öğrenme pekiştirmeli öğrenme yöntemlerinden biridir. Bu algoritmanın temel amacı aksiyonları inceleyerek bu aksiyonlara göre kazanılacak ödülü belirlemek ve bu ödül en üst düzeye çıkarmak için karar almaktır. Q öğrenmenin temel kuralı denklem 1’de verilmiştir:

$$Q_{t+1}(s_t, a_t) = Q(s_t, a_t) + \alpha (r_{t+1}) + \gamma \max_A Q(s_{t+1}, a_t) - Q(s_t, a_t) \quad (1)$$

Denklem 1’de $Q(s_t, a_t)$ değeri her döngüde azalan öğrenme oranı α ’ya bağlı olarak güncellenir. Denklem 1’deki \max_A

ifadesi olası aksiyonlar arasında en değerli aksiyon anlamına gelir. γ önem faktörüdür, öğrenme esnasında mevcut ödülün sonraki döngülerde elde edilen ödüllere kıyasla ne kadar önemli olduğunu belirtir ve 0-1 arası değerler alır.

Bu tip problemlerde ana amaç toplam ödülün maksimize edilmesidir. Mevcut ödülün gelecekteki ödül ile ilgisi denklem 2'de verilmiştir:

$$R_t = \sum_{k=t_0}^{t_{\infty}} \gamma^{k-t} r_k = r_t + \gamma R_{t+1} \quad (2)$$

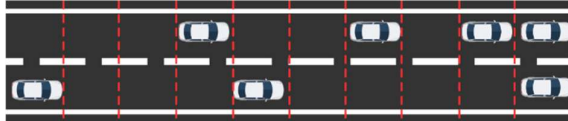
Gelecekteki elde edilecek tüm ödülleri hesaplamak imkansızdır ve bu nedenle bir tahmin algoritmasına ihtiyaç duyulmaktadır ve bu algoritma da Q öğrenmedir.

Şimdiye kadar Derin Q Öğrenme ile tek kavşakta yapılan çalışmalarda üzerinde durulan ana tema olarak araçların konumları kullanılmış ve bu durumların nasıl optimize edileceği ve ödül fonksiyonlarının türleri incelenmiştir. Bu çalışmada tekli kavşakta izole edilmiş optimal bir kontrolör tasarlamak için denenebilecek stratejiler incelenip kıyaslanmıştır.

Bu çalışmada, ajanın optimal sonuç vermesi için değiştirilebilecek 3 etmen incelenmiştir. Ajanın durumlarının seçimi, ödül fonksiyonunun seçimi ve şartlardaki araçların konum bilgisinin seçim şekli kıyaslanarak bu parametrelere bağlı en uygun ajanın tespiti incelenmiştir.

Ödül durumunda kuyruk uzunluğu ve bekleme süresi göz önüne alınmıştır. Kuyruk uzunluğu, ışıklarda bekleyen araçların ortalama uzunluğu, bekleme süresi ise ışıklarda bekleyen araçların toplam bekleme süresi olarak ifade edilebilir.

Bölme uzunluğu parametresi, kavşaktaki araçların konumlarının durumlara ayrılmasında seçilecek uzunluk parametresidir. Sabit bölme uzunluğu kavşağın Şekil 2'deki gibi 10 eşit bölmeye ayrıldığı senaryo iken değişken bölme uzunluğu kavşağa yaklaştıkça bölmelerin küçüldüğü senaryodur.



0	0	0	1	0	0	1	0	1	1
1	0	0	0	1	0	0	0	0	1

+

			0.9			0.6		0.1	0
0.5				0.8				0	0

Şekil 2: Konum ve hıza göre durumların elde edilmesi

Durumlar parametresi ise ajanın durum olarak araçların konumu mu, ortalama hızı mı yoksa hem konum hem ortalama hızının mı hesaba katılacağını inceleyen parametredir.

Çalışmada 12 adet strateji denenmiştir. Her strateji için parametre seçimleri Tablo 2'de verilmiştir.

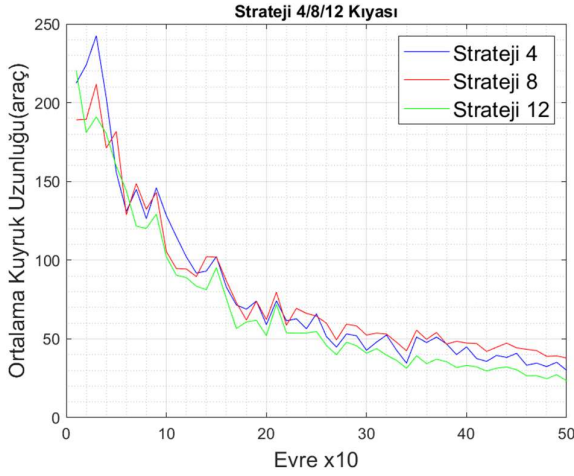
Tablo 2: Trafik ışık kontrol stratejileri

Strateji Numarası	Durumlar	Bölme Uzunluğu	Ödül
1	Konum	Sabit	Kuyruk Uzunluğu
2	Konum	Sabit	Bekleme Süresi
3	Konum	Değişken	Kuyruk Uzunluğu
4	Konum	Değişken	Bekleme Süresi
5	Hız	Sabit	Kuyruk Uzunluğu
6	Hız	Sabit	Bekleme Süresi
7	Hız	Değişken	Kuyruk Uzunluğu
8	Hız	Değişken	Bekleme Süresi
9	Konum+Hız	Sabit	Kuyruk Uzunluğu
10	Konum+Hız	Sabit	Bekleme Süresi
11	Konum+Hız	Değişken	Kuyruk Uzunluğu
12	Konum+Hız	Değişken	Bekleme Süresi

4. Benzetim Sonuçları

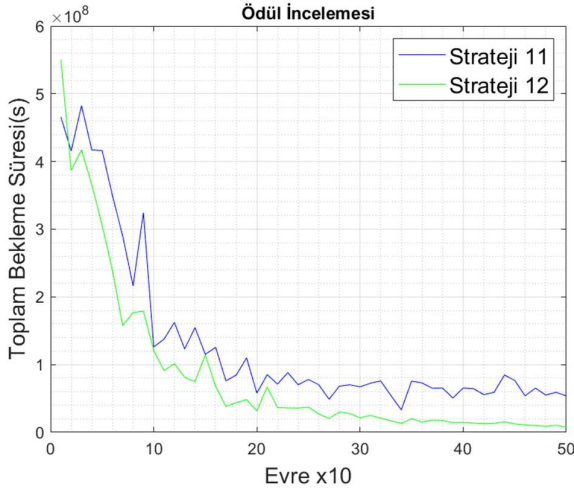
Önerilen tüm stratejiler benzetim ortamında test edilmiştir. Toplam bekleme süresi, ortalama kuyruk uzunluğu, CO_2 emisyonu, araçların ortalama hızı, karşılaştırma kriterleri olarak kullanılmıştır.

Durum parametresine bağlı kıyaslamalar Şekil 3'teki grafikte verilmiştir. Kıyaslamalar yapılırken ödül olarak bekleme süresi, bölme uzunluğu için değişken bölme uzunluğu seçilmiştir. Karşılaştırma için araç cinsinden ortalama kuyruk uzunluğu seçilmiştir. Karşılaştırma sonucunda elde edilen bulgulara göre konum ve ortalama hızın hesaba katıldığı durumlar en iyi sonuç verdiği gözlemlenmiştir. En kötü sonucun ise ortalama hızın durum olarak seçildiği ajanlarda olduğu gözlemlenmiştir.



Şekil 3: Diğ er parametreler sabitken durum parametresinin karşılařtırılması

Ödül kıyaslaması içinse daha iyi sonuç veren konum ve ortalama hızın birlikte kullanıldıđı bölme uzunluđu olarak ise deđişken bölme uzunluđu seçilmiřtir. Toplam bekleme süresi cinsinden yapılan kıyaslamada bekleme süresinin kuyruk uzunluđundan daha iyi sonuç verdiđi, Şekil 4'te görülmektedir.



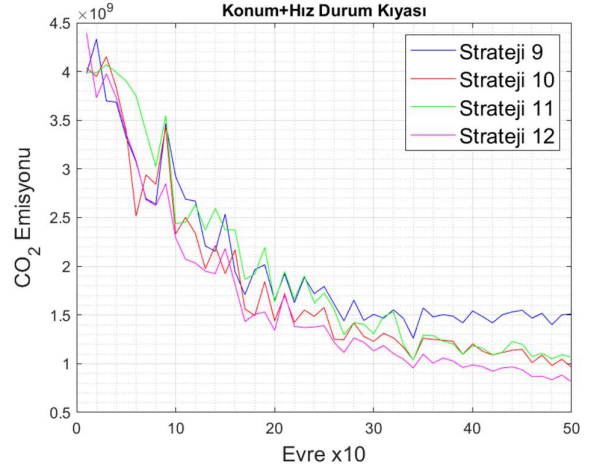
Şekil 4: Diğ er parametreler sabitken ödül parametresinin karşılařtırılması

Bölme uzunluđu kıyaslaması için ise daha iyi sonuç veren konum ve ortalama hızın birlikte kullanıldıđı ödül olarak ise bekleme süresi seçilmiřtir. Ortalama hız cinsinden yapılan Şekil 5'teki kıyaslamada, deđişken kuyruk uzunluđunun sabit kuyruk uzunluđundan daha iyi sonuç verdiđi gözlemlenmiřtir.



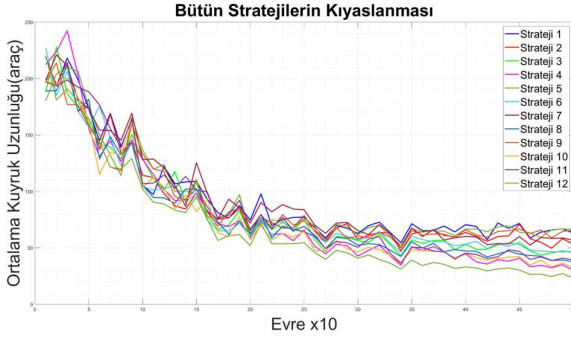
Şekil 5: Diğ er parametreler sabitken bölme uzunluđu parametresinin karşılařtırılması

Bu arařtırmada yenilik olarak sunulan durum olarak konum ile birlikte ortalama hız kullanılmasının denendiđi tüm stratejilerin CO_2 (mg/s) emisyonu cinsinden kıyaslanması Şekil 6'da verilmiřtir. Bu kıyaslamada önceki çıkarımları destekler şekilde ödül olarak bekleme süresinin kuyruk uzunluđundan, bölme uzunluđu parametresinde ise deđişken bölme uzunluđunun sabit bölme uzunluđundan daha iyi sonuç verdiđi gözlemlenmiřtir.



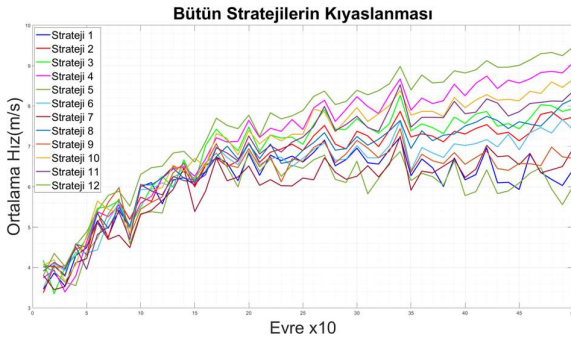
Şekil 6: CO_2 emisyonu karşılařtırılması

Şekil 6'da verilen 4 grafikten çıkan ortak sonuç ajanların çeřitli parametreler açısından kıyaslandıđında benzer şekilde sıralandıđı gözlemlenmiřtir. Bu durum sonucunda ajanların performanslarının deđerlendirme ölçeklerinin birbiriyle ilintili olduđu ve dolayısıyla ajanın tasarımı iyileřtirildikçe kıstaslarda daha iyi performans verecekleri çıkarımları yapılabilir.



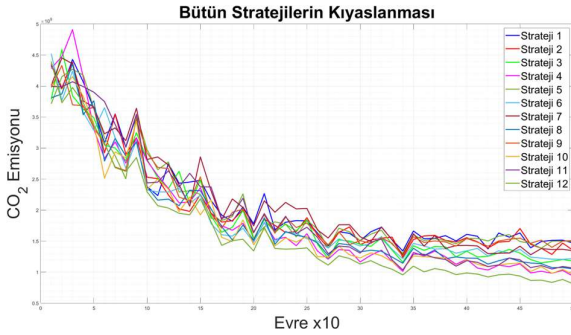
Şekil 7: Tüm stratejilerin ortalama kuyruk uzunluğu karşılaştırılması

Şekil 7’de görüldüğü gibi, tüm stratejiler ortalama kuyruk uzunluğu bakımından incelendiğinde, her birinin öğrenme konusunda belli bir başarı gösterdiği, ve rastgele seçilen ilk değerlerine kıyasla başarılı bir performans gösterdikleri söylenebilir.



Şekil 8: Tüm stratejilerin ortalama hızlarının karşılaştırılması

Aynı şekilde tüm stratejilerde ortalama araç hızları incelendiğinde sonuç Şekil 8’deki gibidir. Burada stratejiler arasındaki performans farklılığı daha belirgin bir biçimde görülmektedir. Tüm stratejilerin karbon dioksit emisyonu cinsinden karşılaştırılması ise Şekil 9’da görülmektedir.



Şekil 9: Tüm stratejilerin CO2 emisyonu karşılaştırılması

Tablo 3’te bu araştırma için tasarlanan tüm stratejilerin son 50 benzetimde kuyruk uzunluğu, karbon dioksit emisyonu ve ortalama hız sonuçlarının ortalaması verilmiştir.

Tablo 3: Benzetim sonuçları

Strateji Numarası	Ort. Kuyruk Uzunluğu (m)	Karbon Emisyonu (kg)	Ortalama Hız (m/s)
1	48.9803	1,224,755,740	7.2725
2	54.1106	1,361,928,154	7.8531
3	41.3613	1,103,652,192	8.2960
4	27.2589	883,095,514	9.2458
5	45.6191	1,190,789,767	7.3415
6	78.7450	1,658,405,392	5.2620
7	49.5519	1,232,274,867	7.2112
8	41.9637	1,115,059,752	7.7039
9	62.8719	1,459,685,460	6.5737
10	31.3054	947,328,261	8.7617
11	38.2572	1,042,225,239	8.3552
12	24.1331	825,133,972	9.4839

5. Sonuçlar

Çalışmada farklı parametre ve kriter kombinasyonlarını barındıran çok sayıda kontrolcü, tek kavşaklı bir trafik sistemi benzetiminde test edilerek karşılaştırılmıştır. Bu karşılaştırmalar sonucunda kontrolcü tasarlanırken daha iyi sonuç veren özellikler belirlenmiştir. Durum ve bölme uzunluğu parametreleri sabit tutulan senaryolarda, ödül olarak bekleme süresinin seçilmesinin, kuyruk uzunluğu ödülünden daha iyi sonuç verdiği görülmüştür. Durum ve ödül parametreleri sabit tutulan senaryolarda, değişken kuyruk uzunluğunun, sabit kuyruk uzunluğundan daha iyi sonuç verdiği görülmüştür. Ödül ve bölme uzunluğu parametreleri sabit tutulan senaryolar için, hem konum hem hız durum seçiminin en iyi durum seçimi olduğu, en kötü durum seçiminin ise sadece hız durum seçimi olduğu görülmüştür.

Yapılan tüm karşılaştırmalar sonucunda, durum parametresinin doğru seçimi, ajanın davranışını etkileyen en önemli faktörken bu parametreyi performans etkisi bakımından sırasıyla ödül seçimi ve bölme uzunluğu parametreleri takip etmektedir. Tasarlanan tüm ajanların öğrenme konusunda belli bir ölçüde başarılı olduğu gözlemlenmiştir. Bu çalışma, ilerleyen aşamalarda birden fazla kavşaktan oluşan daha büyük sistemlerde, kavşaklar arası haberleşmenin de sağlandığı daha gelişmiş yöntemlerle geliştirilecektir.

Kaynakça

- [1] Y. Wang, X. Zhu, L. Li, and B. Wu, "Reasons and Countermeasures of Traffic Congestion under Urban Land Redevelopment," *Procedia - Social and Behavioral Sciences*, vol. 96, pp. 2164–2172, Nov. 2013, doi: 10.1016/j.sbspro.2013.08.244.
- [2] L. Norena, A. Sarrazola, A. Acosta, A. Marquez, J. R. Espinosa, and J. Espinosa, *Urban traffic control in the city of Medellín: A PID control approach*. 2015, doi: 10.1109/wea.2015.7370145.
- [3] H. S. Talab and H. Mohammadkhani, "Design optimization traffic light timing using the fuzzy logic at a Diphasic's Isolated intersection," *Journal of Intelligent and Fuzzy Systems*, vol. 27, no. 4, pp. 1609–1620, Jan. 2014, doi: 10.3233/ifs-141127.
- [4] İ. Tunc and M. T. Söylemez, Simulation of Fuzzy Logic and PI Control Methods in Intersection Management. 21. *Otomatik Kontrol Ulusal Toplantısı TOK 2019*, 2019, pp. 91-96.
- [5] I. Tunc, A. Y. Yesilyurt, and M. T. Soylemez, Intelligent Traffic Light Control System Simulation for Different Strategies with Fuzzy Logic Controller. 2019, doi: 10.23919/eleco47770.2019.8990632.
- [6] I. Tunc and M. T. Soylemez, *State feedback control for intelligent traffic light systems*. 2020, doi: 10.1109/med48518.2020.9183224.
- [7] G. Han, Q. Zheng, L. Liao, P. Tang, Z. Li, and Y. Zhu, "Deep reinforcement learning for intersection signal control considering pedestrian behavior," *Electronics*, vol. 11, no. 21, p. 3519, Oct. 2022, doi: 10.3390/electronics11213519.
- [8] I. Tunc and M. T. Soylemez, "Fuzzy logic and deep Q learning based control for traffic lights," *Alexandria Engineering Journal*, vol. 67, pp. 343–359, Mar. 2023, doi: 10.1016/j.aej.2022.12.028.
- [9] H. Ge, Y.-M. Song, C. Wu, J. Ren, and G. Tan, "Cooperative deep Q-Learning with Q-Value transfer for Multi-Intersection signal control," *IEEE Access*, vol. 7, pp. 40797–40809, Jan. 2019, doi: 10.1109/access.2019.2907618.
- [10] T. Haddad, D. Hedjazi, and S. Aouag, "A deep reinforcement learning-based cooperative approach for multi-intersection traffic signal control," *Engineering Applications of Artificial Intelligence*, vol. 114, p. 105019, Sep. 2022, doi: 10.1016/j.engappai.2022.105019.