



# Kahramanmaraş Sütçü İmam University

## Journal of Engineering Sciences



Geliş Tarihi : 27.09.2023  
Kabul Tarihi : 25.12.2023

Received Date : 27.09.2023  
Accepted Date : 25.12.2023

### İHA TABANLI RAY KONTROLÜ İÇİN EVRİŞİMLİ SİNİR AĞLARI İLE RAY GÖRÜNTÜLERİNİN GERÇEK ZAMANLI SEGMENTASYONU

### REAL-TIME SEGMENTATION OF RAIL IMAGES WITH CONVOLUTIONAL NEURAL NETWORKS FOR UAV BASED RAIL INSPECTION

Selçuk Sinan KIRAT<sup>1</sup> (ORCID: 0000-0003-0106-6995)  
İlhan AYDIN<sup>2\*</sup> (ORCID: 0000-0001-6880-4935)

<sup>1</sup>Milli Eğitim Bakanlığı, İl Millî Eğitim Müdürlüğü, Elazığ, Türkiye  
<sup>2</sup>Fırat Üniversitesi, Bilgisayar Mühendisliği Bölümü, Elazığ, Türkiye

\*Sorumlu Yazar / Corresponding Author: İlhan AYDIN, iaydin@firat.edu.tr

#### ÖZET

Demiryolları insan ve yükünü taşır. Güvenli bir demiryolu seyahati için rayların kontrol edilmesi önemlidir. Raylar genelde insanlar tarafından manuel olarak kontrol edilmektedir. Gelişen teknolojiyle artık İHA'lar birçok görevde insanın yerini almaktadır. Rayların manuel olarak kontrol edilmesi zaman alıcı ve maliyetli bir iştir. Bu nedenle raylar İHA'lar tarafından kontrol edilebilir. İHA'ların rayları kontrol edebilmesi için rayların üzerinde otonom olarak uçuşması gerekir. Bunu yapabilmesi için ray görüntüleri üzerinde segmentasyon yapılmalıdır. Görüntü segmentasyonu bilgisayarlı görü alanında yapılan çalışmalardandır. Bu çalışmalarda derin öğrenmeden faydalanılmaktadır. Derin öğrenme tabanlı evrışimsel sinir ağlarından olan UNet, ICNet ve BiSeNet V2, bilgisayarlı görü uygulamalarında kullanılmaktadırlar. Literatürde gerçek zamanlı görüntü segmentasyonu görevlerinde kullanılan bu ağlar halka açık olarak paylaşılan Railsem19 veri seti özelleştirilerek eğitilmiştir. 1024×512 piksel çözünürlüğündeki görüntüler üzerinde %98 segmentasyon doğruluğuna ulaşan ağlar İHA ile demiryolundan alınan gerçek zamanlı görüntülerde yaklaşık 15 fps hıza ulaşmışlardır. Ağların gerçek zamanlı segmentasyon videosu <https://youtu.be/piVTdsDPzfg> bağlantısından izlenebilir. Çalışmada ayrıca otonom İHA uçuşu bir PID uçuş kontrol sistemi önerilmiştir.

**Anahtar Kelimeler:** Bilgisayarlı görü, derin öğrenme, evrışimsel sinir ağı, gerçek zamanlı segmentasyon

#### ABSTRACT

Railways carry people and their cargo. Checking the rails is important for safe rail travel. Rails are usually controlled manually by humans. With the developing technology, UAVs are now replacing humans in many tasks. Manually checking the rails is time consuming and costly. Therefore, the rails can be controlled by UAVs. In order for UAVs to control the rails, they must fly autonomously on the rails. In order to do this, segmentation must be done on the ray images. Image segmentation is one of the studies in the field of computer vision. Deep learning is used in these studies. UNet, ICNet and BiSeNet V2, which are deep learning-based convolutional neural networks, are used in computer vision applications. These networks, which are used in real-time image segmentation tasks in the literature, were trained by customizing the publicly shared Railsem19 dataset. Networks, which reached 98% segmentation accuracy on 1024×512 pixel resolution images, reached approximately 15 fps on real-time images taken from the railway by UAV. Real-time segmentation video of the networks can be viewed at <https://youtu.be/piVTdsDPzfg>. A PID flight control system for autonomous UAV flight is also proposed in the study.

**Keywords:** Computer vision, deep learning, convolutional neural network, real-time segmentation

## GİRİŞ

Trenler insanlığın ulaşım ve yük taşıma ihtiyacını sağlamak için yüzlerce yıldır demiryollarında ilerlemeye devam etmekte. 21. yy'daki teknolojik gelişmeler, nüfusun artması, şehirlerin büyümesi, insanların daha fazla seyahat ediyor olması, elektrikli hatların ekonomik ve çevreci olması gibi nedenlerden ötürü demiryollarına talep günden güne artmaktadır (Kupriyanovsky vd., 2020). Artık insanlar demiryollarını bazen sadece turistik bir gezi, unutamayacakları farklı bir yolculuk yapma amaçlı olarak da tercih etmekte (Çakmak ve Altaş, 2018). Demiryolları, savaş durumunda mültecilerin ve yaralıların taşınması, mahrumiyet bölgelerine yardım götürülebilmesi gibi görevlerde de insanoğluna büyük katkı sunuyor. Rusya – Ukrayna savaşında mültecilerin %9'u seyahat için demiryollarını kullanmıştır. Savaşın şartları en zorladığı zamanlarda bile demiryolları ile günde 200.000 insan taşınmıştır (International Union of Railways, 2022).

Nüfusu hızla artan ve gelişen ülkemizde özellikle büyük şehirlerde karayolu trafik problemleri ortaya çıkmaktadır. Demiryolları büyük şehirlerin bu problemlerinin çözümüne katkı sunabilmektedir. Ankara-Eskişehir yüksek hızlı tren hattının devreye girmesi ile bu güzergâhta demiryolu yolcu taşıma payı %8'den %72'ye çıkmıştır. 2053 ulaştırma ve lojistik ana planındaki çevreci (sürdürülebilir) senaryoya göre 2053 yılına kadar demiryollarının yük taşımadaki payının %5,08'den %21,93'e, yolcu taşımadaki payının %0,96'dan %6,20'ye çıkarılması hedeflenmektedir. Söz konusu planda 2035 yılına kadar demiryollarında taşınan yük miktarının (ton) %457, yolculuk sayısının ise %645 artması ön görülmektedir (T.C. Ulaştırma ve Altyapı Bakanlığı, 2023). Bu rakamlar diğer tüm ulaşım araçlarında hedeflenen artış miktarından oransal olarak daha yüksektir.

Birbirine traversler ile bağlanmış demir ray çifti üzerinde ilerleyen bir aracın güvenle seyredebilmesi için demiryollarının planlı olarak denetlenmesi gerekir. Demiryolu üzerindeki yapısal bozukluklar, ray kusurları, travers ve bağlantı elemanı kusurlarının tespit edilmesi önemlidir. Ayrıca izinsiz giriş olarak kabul edilen ray üzerine devrilmiş bir ağaç, yuvarlanmış bir kaya, ray üzerinde otomobil gibi bir kara aracı ya da bir yayanın olması demiryolu ulaşımı açısından büyük risk oluşturmaktadır. Tüm bu kusur ve izinsiz girişlerin tren raylardan geçmeden önce tespit edilebilmesi açısından demiryolu denetimlerinin ideal bir sıklıkta ve hassasiyetle yapılması gerekir. Günümüz teknolojisi ile saatte 250km hızı rahatlıkla aşabilen yüksek hızlı trenlerin raylar üzerinde ilerlediği düşünüldüğünde en küçük bir hatanın ciddi can ve mal kayıplarına sebep olabileceği ortadadır. Şekil 1'de ülkemizde 6 Şubat 2023 tarihinde meydana gelen Kahramanmaraş merkezli deprem sonucunda raylarda meydana gelen bükülme görülmektedir.



Şekil 1. Kahramanmaraş Depremi Sonrası Raylardaki Bükülme (Anadolu Ajansı, 2023)

Şekil 1'de Kahramanmaraş depremi sonrasında raylarda meydana gelen bükülme iki farklı açıdan görülmektedir. Görülen konvansiyonel demiryolu hattında meydana gelen bükülme neticesinde ilgili güzergâhtaki seferler raylar düzeltilene kadar iptal edilmiştir.

Demiryollarının denetimi genellikle demiryolu işletmesi personeli tarafından manuel olarak yapılmaktadır. Bu durum gereğinden fazla insan gücü, aşırı zaman ve maliyet sorunlarını ortaya çıkarmaktadır (Ni vd., 2022). Ülkemizde de TCDD personeli tarafından raylar denetlenmektedir. Son yıllarda donanım alanındaki gelişmelerle beraber görüntü işleme alanında derin öğrenme teknikleri rahatlıkla kullanılabilir olmuştur. Bunun neticesinde araştırmacılar otonom ray denetimi için çeşitli bilgisayarlı görü uygulamaları geliştirip önermişlerdir.

İnsansız Hava Aracı (İHA), kendisine verilen uçuş programına göre rotasında otonom ilerleyebilen ya da pilot tarafından uzaktan kontrol edilen hava aracı olarak ifade edilmektedir. Askeri ya da sivil amaçlı olarak kullanılabilen

bu araçlar, günlük hayatta birçok alanda kullanılmaktadır. Gözetleme, keşif gibi askeri görevleri yerine getirmenin yanı sıra sivil alanda bazı e-ticaret firmalarının kargo dağıtım işlerinde dahi kullanılabilir (Murat, 2021). İHA'ların yolların, petrol boru hatlarının ve elektrik hatlarının denetiminde de yaygın olarak kullanıldığı görülmektedir (Yang vd., 2022a). Çalışan personelin can güvenliği açısından erişimi zor olan bölgelerde görev alabilen İHA'lar sayesinde maliyetler azalmaktadır. İHA, üzerinde taşıdığı donanım sayesinde demiryollarını denetlemek için derin öğrenme tabanlı görüntü işleme uygulamalarında da kullanılabilir (Güçlü vd., 2022).

İHA'lar demiryollarını denetlemek için raylar üzerinde ilerlediğinde, kamera tarafından alınan anlık görüntülerin gerçek zamanlı olarak işlenmesi gerekmektedir. İHA'nın rayları takip ederek ilerleyebilmesi, ray görüntülerinin İHA tarafından anlamlandırılması ile mümkün olabilmektedir. Bu bağlamda İHA, içerisinde ray olan bir görüntüdeki ray çiftini, rayların arasında kalan bölgeyi, rayların dışında kalan alanları ayırt edebilmelidir. Bu sayede ray takibi yaparak rotasında ilerleyebilir.

Görüntü anlamlandırma için derin öğrenme tabanlı segmentasyon ağları kullanılmaktadır. Bu ağlar görüntüyü oluşturan her bir pikselin bir sınıfa atanması prensibine göre çalışmaktadır. Bir görüntü milyonlarca pikselden meydana geldiğinden derin ağlarının saniyede onlarca resim içerisindeki pikselleri sınıflandırma görevinde hız sorunu ortaya çıkarmaktadır. Araştırmacılar bu sorunu çözümü için farklı mimaride ağlar tasarlamakta, ağı parametre sayısını azaltarak, boyutunu küçülterek ağı hızlandırmaktadırlar.

Literatürde ray görüntülerinin anlamlandırılması ile ilgili çeşitli çalışmalar yapılmıştır. Rahman ve Mammeri (2021) demiryolu rayları üzeri ve etrafındaki bitki örtüsünün otonom denetimi için İHA ile şehir içinden toplanan hava görüntülerini derin evrimsel sinir ağlarını kullanarak anlamlandırmıştır. Yazarlar şehir görüntülerindeki bitki örtüsünü anlamlandırabildikleri çalışmalarını gelecekte demiryolu görüntüleri üzerinde uygulayacaklarını ifade etmektedirler. Mammeri vd. (2021) İHA ile demiryollarından toplanmış görüntü verilerini dört farklı derin evrimsel sinir ağı ile anlamlandırmış ve modellerin başarılarını karşılaştırmıştır. Yazar farklı demiryollarının kesişim noktalarının anlamlandırılmasında sıkıntı yaşandığını, bunun nedeninin kesişim noktası içeren görüntü verilerinin az olmasından kaynaklandığını belirtmiştir. Guclu vd. (2021) demiryollarında otonom denetim için görüntü verisine ihtiyaç olduğunu, veri toplamak için İHA kullanılabileceğini ifade etmiştir. Çalışmalarında İHA'nın rayları sabit bir yükseklikten sabit bir hızda takip edip görüntü toplayabilmesi için görüntü işleme tekniklerini kullanmışlardır. Gabor filtresi ile ray hatlarını belirginleştirip Hough dönüşümü yapan yazarlar, rayların ufukta kesiştiği noktayı bularak İHA'nın kendisini bu noktaya göre konumlandırarak ilerlemesini sağlamışlardır. Bojarczak ve Lesiak (2021) İHA ile demiryolu üzerinden alınan görüntülerde ezilme, pullanma vb. ray kusurlarını tespit etmişlerdir. Çalışmalarında İHA ile tepeden aşağı doğru dik açı ile çekilmiş ray görüntülerinde rayların konumu belirlenerek derin evrimsel sinir ağı ile anlamlandırılıp kusurun konumu işaretlenmiştir. Aydın vd. (2021) İHA ile ray yüzey kusurlarının tespiti için önerdiği çalışmada, ön ve alt kameraya sahip İHA kullanmıştır. Ön kameradan alınan görüntüleri içerisinde Hough dönüşümü bulunan derin evrimsel sinir ağını kullanarak ray hatlarını anlamlandırmış, hatların kesişim noktasını bulup İHA'nın bu noktaya göre hareket etmesini sağlamıştır. Aynı anda alt kameradan toplanan ray görüntülerini derin evrimsel sinir ağlarını kullanarak anlamlandırmış ve kusur tespiti yapmıştır. Sevi ve Aydın (2022) Gazebo adı verilen simülasyon ortamında İHA'nın rayları takip etmesi için derin evrimsel sinir ağını kullanmıştır. Simülasyon ortamındaki ray görüntülerinin anlamlandırıldığı çalışmada %99 segmentasyon doğruluk oranına ulaşılmıştır. Grandio vd. (2022) mobil haritalama sistemi tarafından demiryolu altyapısından elde edilen nokta bulutu biçimindeki 3B görüntüleri derin evrimsel sinir ağlarında kullanarak anlamlandırmaya çalışmıştır. Yang vd. (2022b) trenler ve İHA'larda bulunan yerleşik kameralardan toplayarak görüntülerle segmentasyon görevi için veri seti (iRailway) hazırlamışlardır. Bu veri seti üzerinde derin evrimsel sinir ağı modellerinden biri olan GAN türündeki RT-GAN ağını kullanarak ray görüntülerini anlamlandırmaya çalışmışlardır. Katar ve Duman (2022) ray görüntülerini anlamlandırmak için derin evrimsel sinir ağı modeli olan U-Net'i farklı versiyonlarla farklı anlamlandırma senaryoları için kullanmıştır. Chen vd. (2022) yüksek hızlı demiryollarından İHA ile alınan görüntüleri anlamlandırmak ve izinsiz giriş tespiti yapmak için derin evrimsel sinir ağı modeli olan All-in-One YOLO'yu kullanmışlardır. Tong vd. (2022) İHA ile toplanan demiryolu görüntülerini gerçek zamanlı anlamlandırma için yeni bir derin evrimsel sinir ağı modeli önermişlerdir. Önerdikleri modelde parametre sayısı ve hesaplama maliyetinin 2 kata kadar azaldığını belirtmişlerdir. Çalışmada düşük donanıma sahip İHA'larda gerçek zamanlı ray anlamlandırma yapılabileceğini göstermişlerdir. Yang vd. (2022a) İHA ile alınan demiryolu görüntüleri anlamlandırma ve takip için ayrıklaştırma-filtreleme-yeniden yapılandırma (Discretization Filtering Reconstruction - DFR) adımlarından oluşan bir yöntem önermişlerdir. DRF yöntemi ile kullanılan hafif bir derin evrimsel sinir ağı sayesinde görüntüler bütünüyle segmente edilmekte, demiryolu olma ihtimali olan bölümleri yamuklarla işaretlenmekte, ray hattının olduğu en uzun yamuk serisi birleştirilerek hat çıkarılmakta ve bu alan ikili olarak tekrar

segmente edilmektedir. Önerdikleri model 640×480'lik giriş görüntülerini gerçek zamanlı anlamlandırma görevinde 12 fps değerini yakalamıştır. Weng vd. (2023) İHA'lerden elde edilen ray görüntülerini anlamlandırmak için derin evrimsel sinir ağı modeli olan DeepLabV3+'nın ana hat çıkarma kısmında sırasıyla MobileNetV2 ve MobileNet3'ü kullanarak segmentasyon başarılarını karşılaştırmıştır. Tong vd. (2023) İHA ile alınan ray görüntülerini gerçek zamanlı anlamlandırmak için yeni bir derin evrimsel sinir ağı modeli geliştirmişlerdir. ARTNet adını verdikleri model 512×512'lik giriş görüntülerini anlamlandırma görevinde 50 fps değerini yakalamıştır. Wu vd. (2023) yüksek hızlı demiryollarındaki potansiyel tehlikelerin tespit edilmesi için İHA ile yüksek mesafelerden alınan görüntüleri kendi geliştirdikleri derin sinir ağı ile anlamlandırmaya çalışmışlardır. YOLARC adını verdikleri ağı gerçek zamanlı segmentasyon görevi için 520×520'lik giriş görüntülerinde 35,9 fps değerini yakalamıştır.

Literatürdeki çalışmalar incelendiğinde araştırmacıların demiryolu denetimlerini otonom olarak sağlamak için İHA'lara odaklandıklarını, özellikle İHA'ların rayı takip edebilmesi için görüntünün anlamlandırılması üzerine gayret ettikleri anlaşılmaktadır. Ayrıca araştırmacıların son zamanlarda gerçek zamanlı görüntü bölütleme görevlerine odaklandıkları görülmektedir. Çünkü İHA'lar bataryasının imkân verdiği ölçüde havada kalabilen ve sınırlı donanım kaynaklarına sahip araçlardır. Bu araçlar üzerinde çok fazla parametre ve hesaplama maliyetine sahip derin evrim ağılarının çalıştırılması iyi sonuçlar vermemektedir. Araştırmacılar bu problemi çözmek ve yüksek çözünürlükte görüntü alabilen İHA üzerinde derin evrim ağılarını çalıştırmak için daha az parametre içeren hafif ağı tasarlama yoluna gitmektedirler (Wu vd., 2023; Chakravarthy vd., 2022).

Bu çalışmada İHA'nın rayları takip edebilmesi için 3 farklı derin evrimsel sinir ağı performansları ray görüntülerini gerçek zamanlı anlamlandırma görevinde karşılaştırılmıştır. Çalışmada hafif bir UNet ile ICNet ve BiseNetV2 ağı kullanılmış olup, halka açık olarak paylaşılan Railsem19 veri seti özelleştirilerek söz konusu ağı eğitilmiştir. Çalışma İHA ile ülkemizde otonom ray denetimi sistemlerinin geliştirilmesine katkı sunabilir. Gerçek zamanlı anlamlandırma görevlerinde çalışacak araştırmacılara fikir verebilir.

## METOT

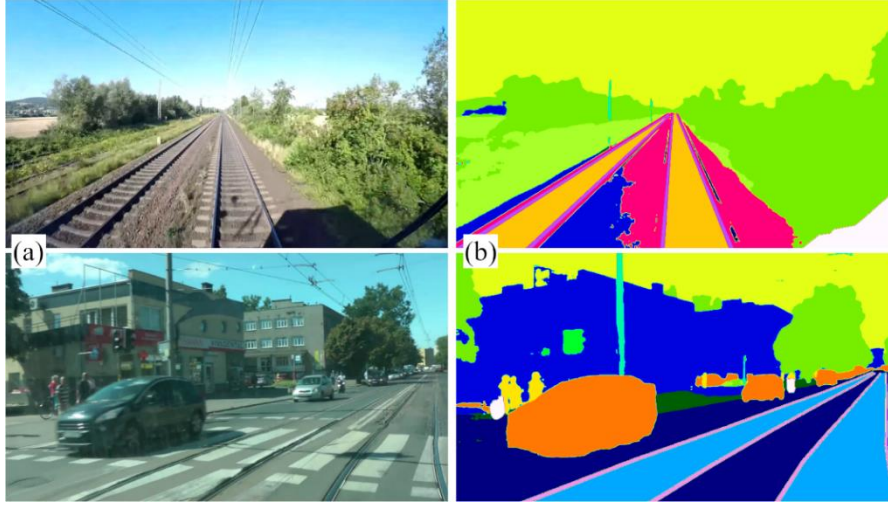
Çalışmada halka açık olarak paylaşılan Railsem19 veri seti kullanılmıştır. Veri seti düzenlenerek belirlenen UNet, ICNet ve BiseNetV2 ağı eğitilmiştir. İHA ile ray takibi için demiryolu üzerinden alınan görüntülerde gerçek zamanlı olarak ağı görüntü anlamlandırma performansları test edilmiştir.

### *Veri Setinin Hazırlanması*

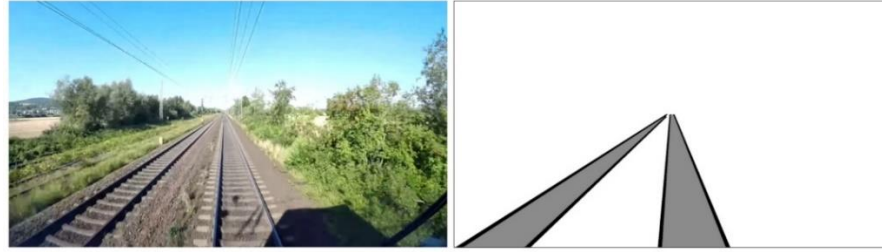
Railsem19 veri seti trenler ve tramvayların ön perspektifinden alınan 1920×1080 piksel en/boy oranındaki 8500 farklı görüntü ve maskelerinden oluşmaktadır. Bu görüntüler, içerisinde karmaşık demiryolu görüntüleri, yetersiz aydınlatma koşulları, şehir içi sahneler gibi zorlukları barındırmaktadır. Railsem19 veri seti görüntüleri maskelerinde yol, kaldırım, yapı, tramvay hattı, çit, direk, trafik ışığı, trafik işareti, bitki örtüsü, arazi, insan, gökyüzü, tren yolu hattı, araba, kamyon, tren yolu ray zemini, ray üzerindeki araç, tren yolu rayları, tramvay rayları yüzeyi olmak üzere 19 sınıf piksel düzeyinde etiketlenmiştir. Piksel düzeyinde etiketleme, bu veri setindeki örnekler için görüntüyü oluşturan tüm piksellerin 1 ile 19 arasındaki değerler ile ifade edilmesi anlamına gelmektedir. Her görüntü bu sınıflardan bazılarını içerisinde barındırmaktadır (Zendel vd., 2019). Şekil 2'de Railsem19 veri setinden 2 örnek görüntü ve maskeleri görülmektedir.

Şekil 2-a'da tren yolu görülmektedir. Şekil 2-b'de ise şehir içi trafikteki tramvay yolu görülmektedir. En sağdaki renkli kısımlar ise gerçek görüntülerdeki sınıfları ifade eden maskelerdir. Maskeler uzmanlar tarafından gerçek görüntülerdeki çeşitli piksel alanlarına belirlenen sınıf etiket numarasının atanmasıyla oluşturulmuştur. Söz konusu etiket numarasına göre ilgili pikseller farklı renklerde görülmektedir. Şekil 2-b'deki görüntü maskesinde otomobil sınıfının turuncu renkle ifade edildiği görülmektedir. Bu çalışma İHA'nın tren raylarında ilerlemesine katkı sunmayı amaçladığından, Railsem19 içerisinde tramvay yolu görüntüsü içerenler gereksizdir. Ayrıca İHA'nın ilerlemesi için ray yüzeyi ile raylar arasında kalan bölgenin (tren yolu hattı) anlamlandırılması yeterlidir. Bu nedenle Railsem19 veri seti içerisindeki şehir içi tramvay görüntüleri çıkarılmış, geriye kalan tren yolu görüntü maskeleri sadece ray yüzeyi ve tren yolu hattı kalacak şekilde yeniden yapılandırılmıştır. Son durumda çalışmada kullanılan veri setinde 7773 adet tren yolu içeren görüntü ile bu görüntülerin 3 sınıflı maskeleri kalmıştır. Şekil 3'te çalışmada kullanılan bir örneğe ait görüntü ile maskesi görülmektedir.

Şekil 3'te görüldüğü gibi maskeler 3 sınıf kalacak şekilde yeniden etiketlenmiştir. İHA'nın ray takibi için ray yüzeyi ve ray hattının olması yeterlidir. Geriye kalan tüm pikseller "diğer" sınıfı olarak kabul edilebilir.



Şekil 2. a. Railsem19 Veri Seti Görüntüleri b. Görüntülerin Maskeleri



Şekil 3. Çalıřmada Kullanılan Bir Görüntü ve Maskesi

Şekil 3'te görüntü maskesinin çizgi izleyen robotların ilerlediği parkura kısmen benzediği söylenebilir. Çizgi izleyen robotun parkurunda sadece beyaz bir alan ve üzerinde siyah çizgi ya da tam tersinin olduğu 2 sınıflı bir yüzey vardır. Robot sadece çizgiye odaklanmaktadır, geriye kalan alanlar aslında robotun ilgilenmediği kısımlardır. Bu çalışmada da benzer biçimde İHA kamerasından alınan görüntü Şekil 3'teki maskeye benzer bir şekilde segmente edildiğinde, İHA çizgi izleyen robot mantığı ile rayı takip edecektir. Dolayısıyla “diğer” sınıfı esasında İHA'yı bu görev bağlamında ilgilendirmeyen tüm alanlardır. Çizgi izleyen robot parkurunda sadece tek bir çizgi varken bu görevde ray yüzeyi ve ray hatları da görülmektedir. Bunun nedeni İHA'nın kendi konumunu ray yüzeyi sınıfına göre ray hattının üzerinde ortalaması gerekliliğidir. Ayrıca demiryollarında çim, bitki örtüsü, çakıl fazlalığı gibi nedenlerle ray hattı bazen görünmeyebilir. Bu gibi durumlarda İHA'nın rotasını kaybetmemesi için ray yüzeyine göre kendini konumlandırması daha etkili bir çözüm olacaktır. Maskelerde kullanılan renkler istenildiği gibi değiştirilebilir. Burada önemli olan görüntü içerisindeki tüm piksellerin belirli kategorik renklerle ifade edilebilmesidir.

### Derin Evrimsel Sinir Ağlarının Seçimi ve Parametre Ayarları

Bölütleme amaçlı kullanılan ağlarda çoğunlukla klasik kodlayıcı – kod çözücü mimarisine sahip UNet ağı omurga olarak kullanılmaktadır. Bu nedenle bu çalışmada da UNet'in kullanımı uygun görülmüştür. Ek olarak literatürde gerçek zamanlı görüntü bölütleme görevlerinde kullanılmış BiseNetV2 ile ICNet ağları da çalışma kapsamına alınmıştır. Seçilen bu ağlara Tablo 1'deki parametreler verilerek derlenmiş, hazırlanan verilerle eğitilmiştir.

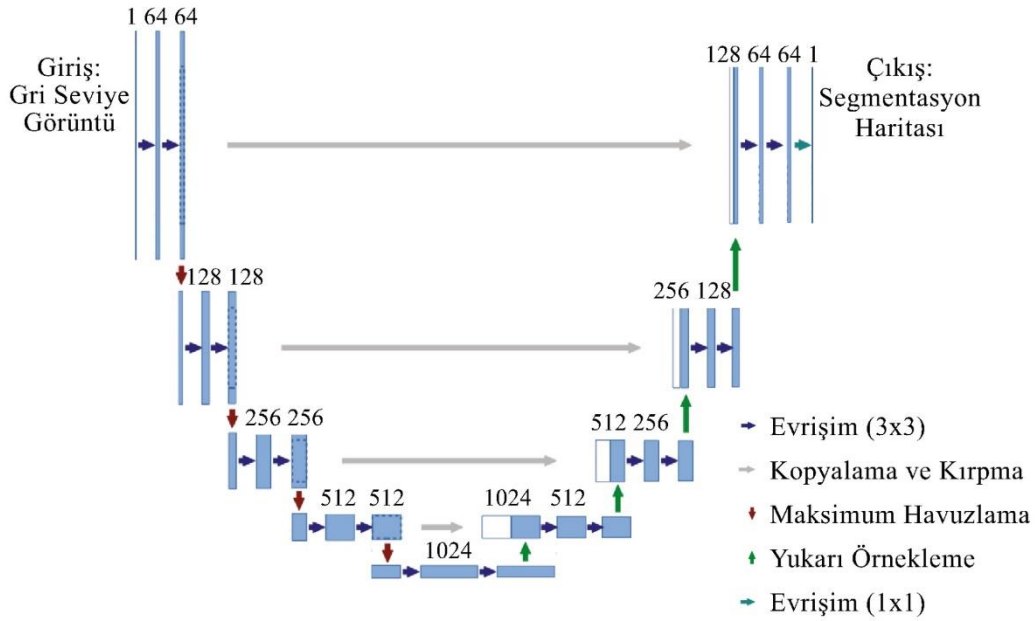
Tablo 1. Ağlar İçin Tercih Edilen Parametreler

Parametre Adı	Parametre Değeri
Giriş Görüntüsü Genişliği	1024
Giriş Görüntüsü Yüksekliği	512
Sınıf Sayısı	3
Yığın Boyutu	16
Dönem	100
Erken Durdurma için Bekleme Dönemi	16
Öğrenme Oranı	0,001
Öğrenme Oranı Bekleme Dönemi	4
Öğrenme Oranı Çarpanı	0,85
Optimize Edici	Adam

Tablo 1’den anlaşılacağı gibi veri setindeki  $1920 \times 1080$  ölçülerine sahip görüntüler,  $1024 \times 512$  ölçülerinde tekrar boyutlandırılarak ağlara verilmiştir. Hatayı en aza indirmek için başta 0,001 olarak belirlenen öğrenme oranı, eğitimdeki 4 dönem boyunca hatada azalma olmaz ise 0,85 ile çarpılarak azaltılmaktadır. Bu sayede eğitimin daha iyiye gitmediği her 4 dönemde bir, daha küçük öğrenme oranı ile eğitim yapılmaktadır. Daha küçük öğrenme oranı hatayı azaltmak için ağın daha hassas öğrenmesini sağlayacaktır. Ayrıca eğitim sürecinde eğer 16 dönem boyunca hata azalmaz ise eğitim otomatik olarak durdurulacaktır. Bu durumda gereksiz yere 100 dönem eğitim yapılmadan da eğitim sonlandırılabilir. Tablo 1’de görülen parametreler deneysel olarak iyi sonuç alındığı için tercih edilmişlerdir. Tüm ağlar için giriş ve çıkış görüntüsü  $1024 \times 512 \times 3$  olarak ayarlanmıştır.

### U-Net Mimarisi

Ronneberger vd. (2015) biyomedikal görüntü segmentasyonu görevindeki başarılı performansı ile U-Net’i tanıtmışlardır. U-Net, mikroskop görüntülerindeki hücre izleme yarışmasını büyük farkla kazanmıştır. U-Net, U harfine benzeyen şekli nedeni ile bu ismi almış olup ağın mimarisi Şekil 4’te görülmektedir.



Şekil 4. U-Net Mimarisi (Ronneberger vd., 2015)

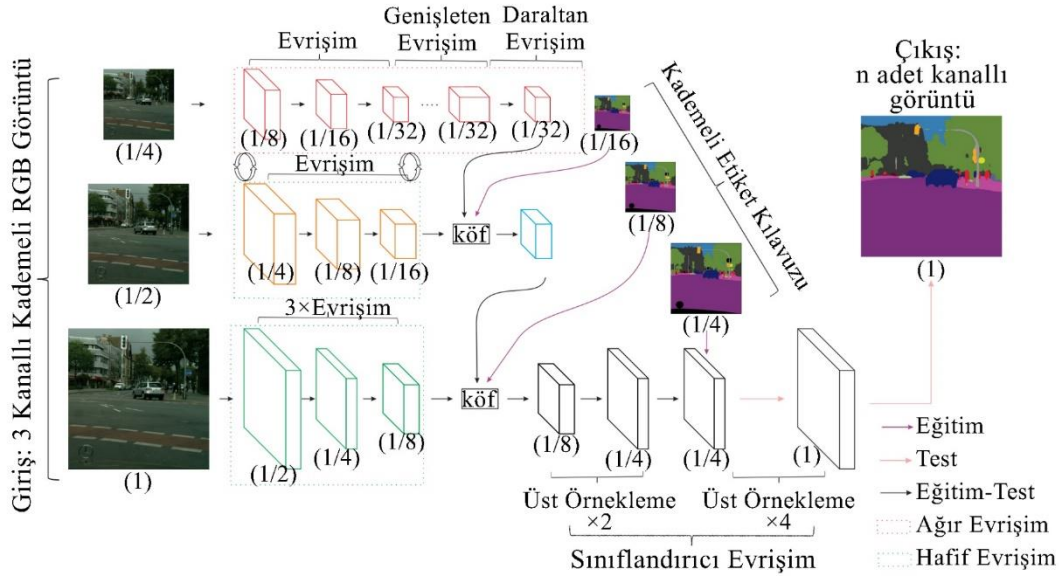
Şekil 4’te görülen U-Net mimarisine göre mavi dikdörtgenler üzerinde yazan sayılar görüntünün kanal sayısını ifade etmektedir. Giriş görüntüsünün evrişim operatörleri (mavi ok) ile kanal sayısı artırılmakta, maksimum havuzlama (kırmızı ok) ile boyutu yarıya düşmektedir. Bu kısım daralma yolu olarak ifade edilmektedir. Mimarinin en alt kısımda iyice boyutu küçülen görüntüler yukarı örneklemeye (yeşil ok) yapılarak boyutu tekrar iki katına çıkarılmaktadır. Bu kısım ise genişleme yolu olarak ifade edilmektedir.

Küçültülen bir görüntüde veri kayıpları meydana gelmektedir. U-Net veri kaybının dezavantajını kopyalama ve kırpma operatörleri yardımı ile azaltmaktadır. Kopyalama ve kırpma adımı, daralma yolundaki her aşamada evrişim işlemlerinden hemen en sonra kopyası alınmakta, gereksiz kısımları kırılarak genişleme yolunda tam karşıya eklenmektedir. Bu sayede ağı çıkışından önce genişleme yolundaki her seviyede daralma yolundaki yüksek çözünürlüklü görüntü bilgileri de kullanılabilir.

### ICNet Mimarisi

Zhao vd. (2018) şehir manzaraları, CamVid, COCO-Stuff gibi veri setlerinde gerçek zamanlı segmentasyon görevini yerine getirmek için ICNet (Image Cascade Network) adını verdikleri ağı tanıtmışlardır. Şehir manzaraları veri setinde %69,5 segmentasyon doğruluğu ve 30,3 fps hızı, CamVid veri setinde %67,1 segmentasyon doğruluğu ve 27,8 fps hızı, COCO-Stuff veri setinde ise %29,1 doğruluğa ve 35,7 fps hızı ulaşmıştır. Ağı mimarisi Şekil 5’te görülmektedir.

Şekil 5’te görülen ICNet mimarisinde adından da anlaşılacağı üzere ağı girişindeki görüntü kademeli bir şekilde 1, 1/2 ve 1/4 oranlarına ölçeklenerek birbirine paralel evrişim operatörlerinden oluşan dallarda işlenmektedir. Parantez



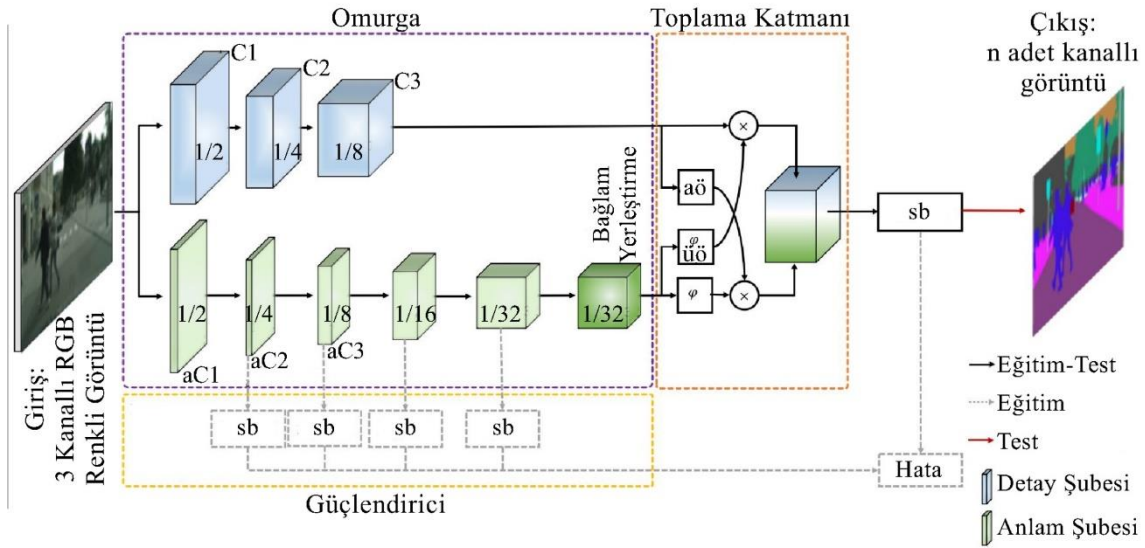
Şekil 5. ICNet Mimarisi (Zhao vd., 2018)

içerisindeki sayılar her katmandaki görüntünün giriş görüntüsüne oranını ifade etmektedir. Ağda yüksek çözünürlüklü görüntülerin işlenmesi hesaplama karmaşıklığını artırmaktadır. Düşük çözünürlüklü görüntülerde ise hesaplama karmaşıklığı azalırken çok kaba segmentasyon sonuçları ortaya çıkmaktadır. Hız ve doğruluk arasında ideal bir denge kurmayı amaçlayan ICNet geliştiricileri, tasarladıkları ağır mimarilerinde görüntünün yüksek, orta ve düşük çözünürlüklü ölçeklerini birbirine paralel dallarda birlikte kullanmakta, kademeli özellik füzyonu (köf) adımı verdikleri yapılarda ise farklı çözünürlük ölçeklerindeki görüntülerden elde edilen özellikleri birleştirmektedir. Hesaplama karmaşıklığını azaltmak adına ICNet'te en üst daldaki derin katmanlı evrişimlerde görüntünün düşük çözünürlüklü hali kullanılarak anlamsal çıkarımlar yapılmaktadır. Orta ve yüksek çözünürlüklü görüntülerin kullanıldığı dallarda ise daha ayrıntılı bilgiler toplanmaktadır. Tüm dalların çıkışındaki özellik haritaları ve bu haritalarla aynı boyutlara sahip kademeli etiket kılavuzları sistematik olarak köf birimlerinde bir araya getirilmektedir. Üst örnekleme yapılarak modelin çıktısında görüntü istenilen boyutlara tekrar ölçeklenmektedir.

### BiSeNet V2 Mimarisi

Yu vd. (2021) halka açık olarak paylaşılan şehir manzaraları veri seti ile segmentasyon görevinde yüksek doğruluk ve hıza ulaşan BiSeNet V2 (Bilateral Segmentation Network Version 2) ağını tanıtmışlardır. Söz konusu veri setinde  $2048 \times 1024$  çözünürlüğündeki görüntülerde %72.6 segmentasyon başarısına ve 156 fps hıza ulaşabilmiştir. Şekil 6'da BiSeNet V2 mimarisi görülmektedir.

Şekil 6'da görüldüğü gibi BiSeNet V2, omurga, toplama katmanı ve güçlendirici olmak üzere üç ana bölümden oluşmaktadır. Şubeler içerisinde yazan sayılar o şubedeki görüntünün giriş görüntüsüne oranını ifade etmektedir. C1, C2 ve C3 kanal sayısını ifade ederken, aC1, aC2 ve aC3 ise detay şubelerindeki kanal sayılarının bir a katsayısı ile çarpılarak elde edilen yeni kanal sayısını ifade etmektedir. Geliştiriciler mevcut yaklaşımların görüntülerdeki düşük seviyeli detayları ihmal ettiğini, bu durumun doğruluk düşüşüne neden olduğunu ifade etmişlerdir. Gerçek zamanlı anlamsal bölütleme görevinde yüksek doğruluk elde edebilmek için görüntülerdeki uzamsal detayların ve kategorik anlamların ayrı ele alındığı BiSeNet V2'de omurga bölümü detay şubeleri ve anlam şubelerini içerisinde barındırmaktadır. Omurga bölümünün amacı geniş kanallara ve sık katmanlara sahip detay şubeleri ile görüntüdeki düşük seviyeli detayları, dar kanallara ve derin katmanlara sahip anlam şubeleri ile görüntüdeki üst düzey anlamsal bağlamı yakalamaktır. Toplama katmanında detay ve bağlam şubelerinden gelen görüntü temsilleri bir araya getirilir. Anlam şubeleri daha derin olduğu için görüntü son şubede detay şubesine göre daha küçüktür. Bu nedenle toplama katmanında alt örnekleme (aö) ve üst örnekleme (üö) ile şubelerden gelen görüntü boyutları eşitlenerek element bazında çarpma işlemi yapılarak bir araya getirilir. Segmentasyon başlıklarından (sb) alınan bilgiler ile hata ölçülür. Güçlendirici olarak adlandırılan bölümde anlam şubelerine bağlanmış segmentasyon başlıkları hatayı daha hassas ölçmek ve azaltmaya katkı sağlamak için kullanılmaktadır. Bu bölüm sadece eğitim aşamasında kullanıldığından test aşamasında ağırlık performansını olumsuz etkilememektedir.

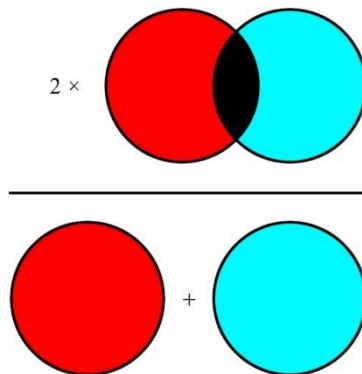


Şekil 6. BiSeNet V2 Mimarisi (Yu vd., 2021)

### Değerlendirme Metriği

Segmentasyon görevi klasik görüntü sınıflandırma görevinden farklıdır. Görüntü sınıflandırma görevinde araştırmacı, eğittiği ağın örnek bir çiçek görüntüsü için doğru etiket tahmini yapmasını amaçlar. Bu durumda her bir çiçek görüntüsü için tek bir etiket tahmini yapılır. Segmentasyon görevinde ise bir görüntünün etiketini tahmin etmek yerine içerisindeki tüm pikselleri belirlenen sınıflar için seçilmiş sayılar ile etiketlemek gerekir. Dolayısıyla bu görev daha karmaşık bir süreçtir ve bir görüntü içerisindeki bitki örtüsü, hayvan, insan, araç, bina vb. nesnelere oluşan piksellerin o nesne için seçilmiş sayı ile etiketlenmesini amaçlar. Süreç tamamlandığında ortaya çıkan görüntü belirlenen sınıf sayısı kadar renk çeşitliliğine sahiptir. Ortaya çıkan bu renklerle makine için dış dünyadan alınan görüntü anlam kazanır. Bu sayede otonom sürüş yapacak araçlar için karar verme süreci kolaylaşır.

Klasik görüntü etiketleme görevinde ağ, her bir görüntü için doğru ve yanlış etiket tahminlerini yapar. Yapılan tahminlerdeki doğru ve yanlış sayıları arasındaki çeşitli oransal hesaplamalarla ağın doğruluk, hassasiyet gibi metrikleri ölçülür. Ölçülen metriklere göre ağın başarısı değerlendirilir. Segmentasyon görevinde ise görüntü içerisindeki pikseller etiketlendiği için, ağın çıkışında segmente edilmiş görüntü içerisindeki etiketlenmiş piksel gruplarının gerçekte olması gerekenle (maske görüntüleri) alansal olarak ne kadar örtüştüğü ölçülmeye çalışılır. Bu metriklerden birisi de zar katsayısı (dice coefficient) olarak adlandırılmaktadır. Zar katsayısı Şekil 7’de gösterilmeye çalışılmıştır.



Şekil 7. Zar Katsayısı Gösterimi (Tiu, 2019)

Şekil 7’de gösterilmeye çalışına zar katsayısı, her bir sınıf için maske görüntüsü ile ağ tarafından tahmin edilen görüntüde örtüşen piksel sayısının iki katının, maske ve tahmin görüntülerindeki toplam piksel sayısına oranı bulunarak hesaplanır. Her bir sınıf için hesaplama yapıldıktan sonra tüm sınıfların ortalaması alınarak tek bir görüntü için zar katsayısı bulunmuş olur. Şekil 7’ye göre mavi alan ağın tahmin ettiği sınıf piksellerini, kırmızı alan maskede gösterilen sınıfın gerçek piksellerini ifade ediyor olsun. Buna göre zar katsayısı Denklem 1 ile hesaplanabilir:



$$DC = 2 \times |X \cap Y| / (|X| + |Y|) \quad (1)$$

Denklem 1'e göre tahmin edilen sınıf bölgesindeki pikseller ile gerçekteki sınıf pikselleri örtüştükçe (küme kesişimi büyüdükçe) zar katsayısı da artmaktadır. Görüntüdeki nesnelere için tahmin ile gerçeğin örtüşmesi tam olduğunda zar katsayısı 1, örtüşme olmadığında 0 değerini almaktadır. Dolayısıyla zar kat sayısı 0 ile 1 arasında değer alabilmektedir

### Demiryolundan Görüntü Toplama

Parrot Anafi4K İHA ile Elazığ merkezden geçen demiryolu üzerinden görüntü alınmıştır. Parrot Anafi4K maksimum 15 m/s yatay hızda ilerleyebilmekte, toplam 25 dk havada kalabilmektedir. Görüntü alınırken İHA ve kamerası evrimsel sinir ağlarının eğitildiği özelleştirilmiş Railsem19 veri setindeki görüntülere uygun şekilde, yaklaşık 3m yükseklikten karşı perspektife bakacak bir açıyla konumlandırılmıştır. 5 m/s hızla uçurulmuştur. Şekil 8'de bu süreç görülmektedir.

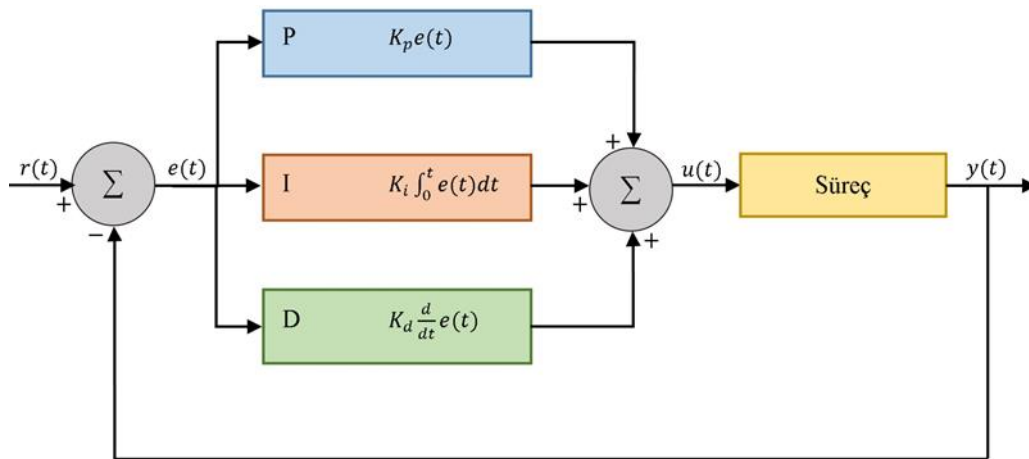


Şekil 8. İHA ile Elazığ Demiryolundan Görüntü Toplama

Demiryolundan alınan video, eğitilen ağların görüntü segmentasyon görevi için kullanılmıştır. Ayrıca bu çalışma kapsamında tasarlanan PID uçuş kontrol sisteminin test edilmesinde kullanılmıştır. Parrot Anafi4K 180 derece dönüş açısına sahip 4K görüntü alabilen kamera, bilgisayara anlık görüntü aktarabilme, SDK'sı aracılığı ile bilgisayardan uçuş komutu alabilme gibi avantajlara sahiptir

### PID Uçuş Kontrolü Tasarımı

Oransal integral türev kontrolcüsü (Proportional Integral Derivative - PID) İHA'ların otonom olarak hareket etmesi için motorların kontrol edilmesinde kullanılmaktadır. Adından anlaşılacağı üzere oran, integral ve türev hesabı yapılarak motorlar için çıkış üretilir. Üretilen çıkış ile istenilen değer arasındaki fark hata olarak PID kontrolcüsüne tekrar verilerek tekrar çıkış hesaplanır. Söz konusu döngü hata sıfır olana kadar devam eder. Bir PID kontrolcüsünün yapısı Şekil 9'da görülmektedir.

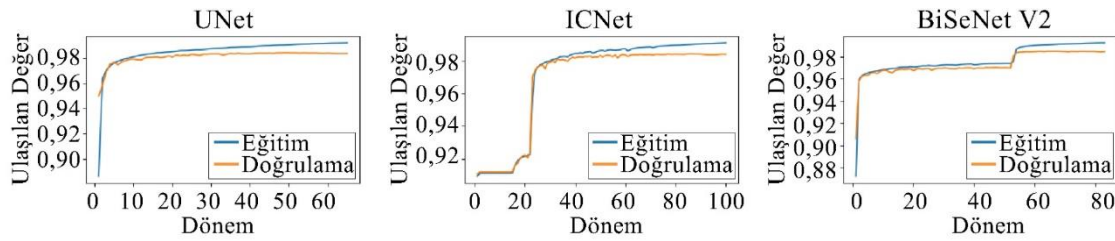


Şekil 9. PID Kontrol Sistemi

Şekil 9'da görüldüğü gibi PID sisteminde P, I ve D olmak üzere 3 ana bileşen mevcuttur. Sisteme  $t$  anında  $r(t)$  referans değeri girer. Sistemin çıkış sinyali  $y(t)$  bu referans değerden çıkarılarak hata  $e(t)$  bulunur. Sistemin P bileşeninde  $e(t)$ , bir  $K_p$  katsayısı ile çarpılır. I bileşeninde bir  $K_i$  katsayısı,  $e(t)$ 'nin integrali ile çarpılır. D bileşeninde ise  $K_d$  katsayısı  $e(t)$ 'nin  $t$ 'ye göre alınan türevi ile çarpılmaktadır. Bu çarpımlardan elde edilen sonuçlar toplandığında  $u(t)$  kontrol çıkışı ve buna bağlı sistem çıkış sinyali  $y(t)$  bulunur. P bileşeni çıkışa sadece belirtilen katsayı ile mevcut hatanın çarpımı kadar etki eder. I bileşeni ise en baştan o ana kadarki tüm hataları integral işleminin doğası gereği hesaba katar ve belirlenen katsayı ölçüsünde çıkışa etki eder. D bileşeni gelecekteki hatayı yordamaya çalışır. Bunu yapabilmek için bir önceki hata ile o anki hatanın zamana göre türevini bulup belirlenen katsayı ile çarpılarak çıkışa etki eder (Bayati, 2019).

## BULGULAR

Çalışmada NVIDIA RTX A5000 GPU donanımı kullanılmıştır. Ağlar Python TensorFlow ile hazırlanmıştır. Tablo 1'deki parametreler verilerek derlenip eğitilen ağların zar katsayılarının eğitim dönemine bağlı grafikleri Şekil 10'da görülmektedir.



Şekil 10. Seçilen Ağların Eğitim Grafikleri

Şekil 10'a göre UNet ile BiSeNet V2'nin daha ilk eğitim dönemlerinde yüksek segmentasyon doğruluğuna (zar katsayısı) ulaşabildikleri dikkat çekmektedir. Tablo 1'de ağların gereksiz dönem eğitimi sürdürmemesi için erken durdurma kullanıldığı görülmektedir. Ağların eğitim sonucunda Tablo 2'deki verilere ulaşılmıştır.

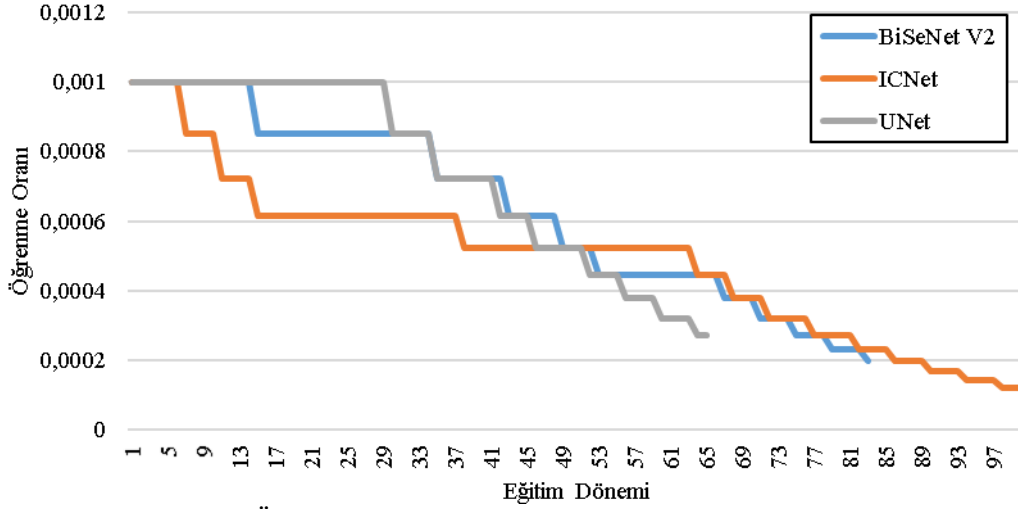
Tablo 2. Seçilen Ağların Eğitim Sürecine Dair Veriler

Elde edilen veriler	UNet	ICNet	BiSeNet V2
Segmentasyon doğruluk oranı (Zar katsayısı)	0,9837	0,9835	0,9843
Eğitim dönemi sayısı	65	100	83
Eğitilebilir parametre sayısı	1.944.083	6.713.353	3.087.675
Bir eğitim döneminin ortalama gerçekleşme süresi	280s	180,9s	378,3s

Tablo 2'ye göre tüm ağların segmentasyon başarısını ifade eden zar katsayısı sonuçları birbirine çok yakındır. En yüksek başarıyı %98,43 oranı ile BiSeNet V2 elde etmiştir. ICNet'in %98,35'lik segmentasyon başarısına ulaşmak için 100 dönem eğitim yaptığı görülmektedir. UNet ile BiSeNet V2'nin ise çok daha erken dönemlerde maksimum segmentasyon başarılarına ulaştığı dikkat çekmektedir. Ağların eğitilebilir parametre sayısına göre bir eğitim döneminin ortalama gerçekleşme süresi incelendiğinde ICNet'in 6 milyondan fazla eğitilebilir parametreyi sadece 180,9 saniyede tamamlayabildiği görülmektedir.

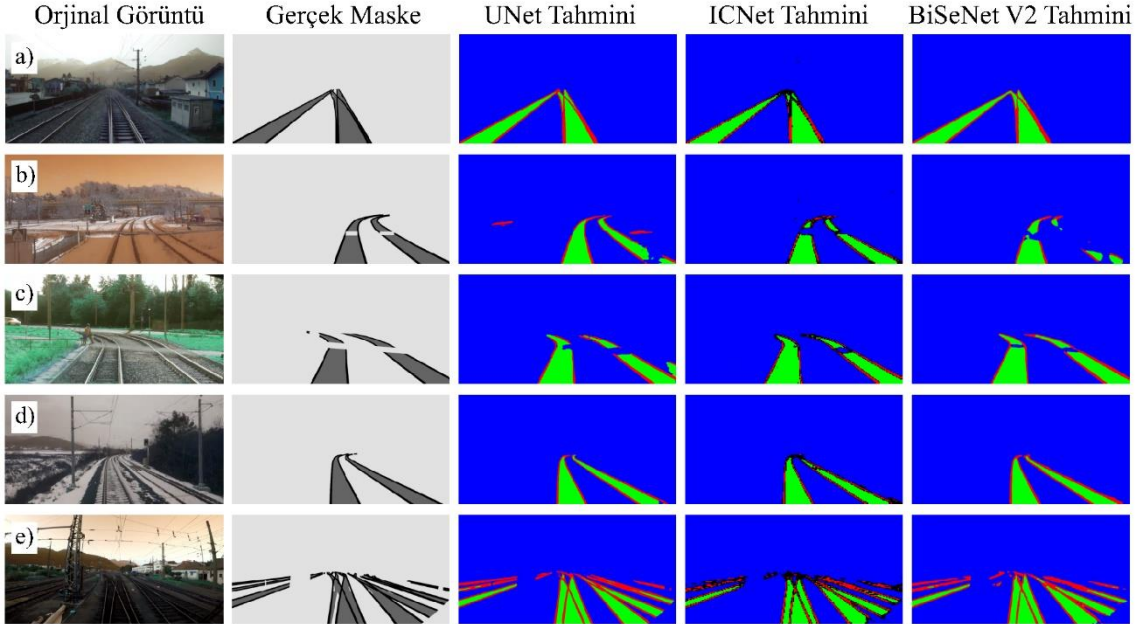
UNet'in parametre sayısının az olmasında her katman için belirlenen çekirdek sayılarının küçük seçilmiş olması etkili olmuştur. Şekil 4'teki orijinal mimaride her katmandaki kanal sayıları sırasıyla 64, 128, 256, 512, 1024 olarak görülmektedir. Orijinal mimariye göre model kullanıldığında eğitilebilir parametre sayısı 31.043.651 olduğundan dolayı donanım kaynakları yetersiz kalmıştır. Bu denklemlerle deneysel olarak en ideal sonucun alınabildiği çekirdek sayısı belirlenerek UNet'in katmanlarındaki görüntü kanalı sayıları 16, 32, 64, 128, 256'ya düşürülmüştür. Ağlardaki eğitilebilir parametre sayısı katmanlardaki çekirdek ve buna bağlı kanal sayıları dışında giriş görüntüsünün boyutuyla da ilgilidir. Bu çalışmada kullanılan giriş görüntüleri 3 kanallı 1024×512 piksel ölçülerindedir. Bunun anlamı 1024×512×3 adet (1.572.864) pikselin sayı değeri ağların girişinde işleme alınacaktır. Eğitilebilir parametre sayısının azaltılması için görüntü boyutlarının küçültülmesi denenebilir. Fakat bu durumda da görüntülerde veri kaybı olacak, ağlar ray konumunu öğrenmekte zorlanacaktır. Bu da deneysel olarak araştırmacının kullandığı donanım kaynaklarına göre belirlemesi gereken bir parametredir. Donanımsal gelişmelere paralel olarak araştırmacılar daha büyük görüntüleri ağına girişine verebilmektedirler.

Tablo 1’de öğrenme oranının eğitimde 4 dönem boyunca daha iyi bir sonuç alınmadığında 0,85 katsayısı ile çarpılarak küçültüldüğü belirtilmiştir. Buna göre seçilen ağların eğitim dönemleri boyunca öğrenme oranlarındaki değişimi gösteren grafik Şekil 11’de görülmektedir.



Şekil 11. Öğrenme Oranlarının Eğitim Dönemlerinde Değişimi

Şekil 11’e göre ICNet’in öğrenme oranı 13 kez, BiSeNet V2’nin öğrenme oranı 10 kez, UNet’in öğrenme oranı 8 kez azalmıştır. Grafik dikkatli incelendiğinde aslında ICNet ile BiSeNet V2’nin 62-68. dönemlerden sonra kayda değer bir öğrenme gerçekleştiremedikleri için öğrenme oranının sürekli düşürülerek daha hassas öğrenme denemeleri yaptığı göze çarpmaktadır. UNet ise eğitimin daha ilk dönemlerinde çok süratli öğrenmiş, yaklaşık 50. dönemden sonra daha küçük öğrenme oranı ile daha hassas öğrenme denemeleri yapıp segmentasyon doğruluğunu artırmayı denemiştir. Şekil 12’de hazırlanan veri setindeki bazı test görüntüleri için ağların tahminleri görülmektedir.



Şekil 12. Seçilen Ağların Segmentasyon Tahminleri

Şekil 12’de görüldüğü gibi BiSeNet V2 ile UNet genel olarak daha iyi segmentasyon tahminleri ortaya çıkarmışlardır. Şekil 12-b’deki görüntüde ise ICNet’in daha başarılı olduğu görülmektedir. Dünyanın farklı yerlerinde, farklı mevsimlerde, günün farklı zamanlarında alınan görüntüler çeşitli zorluklar barındırmaktadır. Derin evrimsel sinir ağları ise mimarilerine göre görüntülerde istenileni öğrenme kabiliyetine sahiptirler. Şekil 12-b da görülen aydınlatma ve hava koşulları durumunda ICNet rayların görüntü üzerindeki konumunu daha iyi öğrenmiştir.

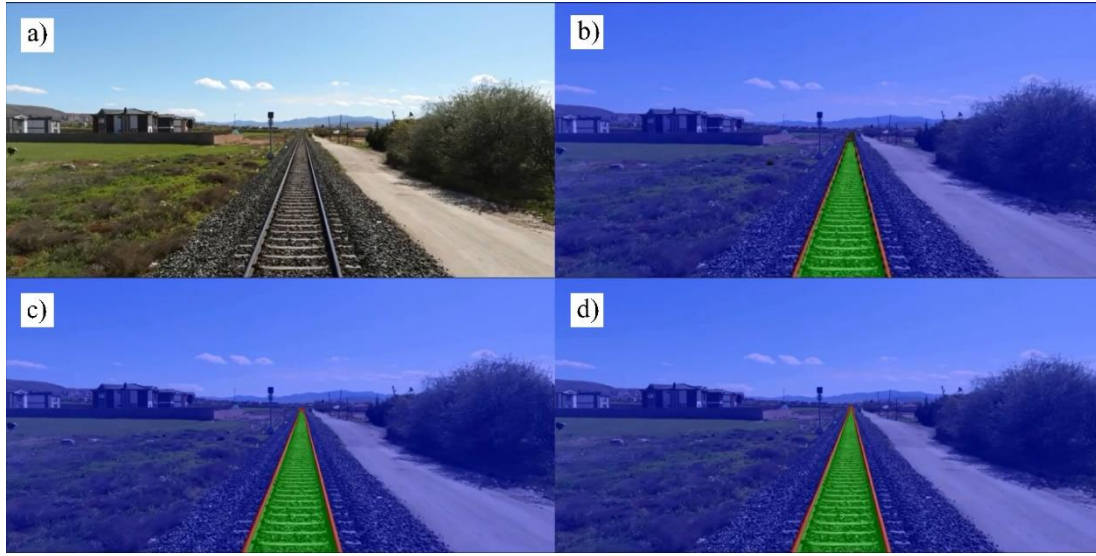
Katar ve Duman (2022) otonom demiryolu araçları için UNet'i Railsem19 veri setini özelleştirerek eğitmişlerdir. Söz konusu çalışmada her görüntü 256×256 piksellik parçalara ayrılarak UNet'e segmente ettirilmiştir. Çalışmada UNet'in gerçek zamanlı segmentasyon performansı belirtilmemiştir. Otonom araçlar ve İHA görevleri için ağların gerçek zamanlı segmentasyon başarıları büyük önem arz etmektedir. Bu nedenle çalışmamızda görüntüler dilimlenmemiş, gerçek zamanlı segmentasyon için UNet'in çekirdek sayısı azaltılmış, BiSeNet V2 ve ICNet ağları da kullanılarak karşılaştırmalar yapılmıştır.

İHA yardımı ile Elazığ merkezden geçen demiryolu üzerinde 3m yükseklikten sabit uçuş ile 5m/s yatay hız ve 30fps görüntü yakalama hızında alınan bir video için ağların segmentasyon hızları Tablo 3'te verilmiştir.

**Tablo 3.** Ağların Video Görüntülerindeki Segmentasyon Performansları

Karşılaştırılan Nitelik	UNet	ICNet	BiSeNet V2
Gerçek zamanlı segmentasyonu hızı	13 FPS	15 FPS	14 FPS

Tablo 3'e göre ICNet'in 1024×512 piksellik görüntülerde 15 FPS'ye kadar ulaştığı görülmektedir. Tablo 2'ye göre ICNet en fazla eğitilebilir parametreye sahipken gerçek zamanlı segmentasyon görevinde diğerlerine göre daha hızlı olduğu görülmüştür. Ağların Elazığ demiryolu videosu için segmentasyon sonuçları Şekil 13'te görülmektedir.



**Şekil 13.** Elazığ Demiryolu Görüntüsünün Anlamlandırılması **a.** Orijinal Görüntü **b.** ICNet Segmentasyonu **c.** BiSeNet V2 Segmentasyonu **d.** UNet Segmentasyonu

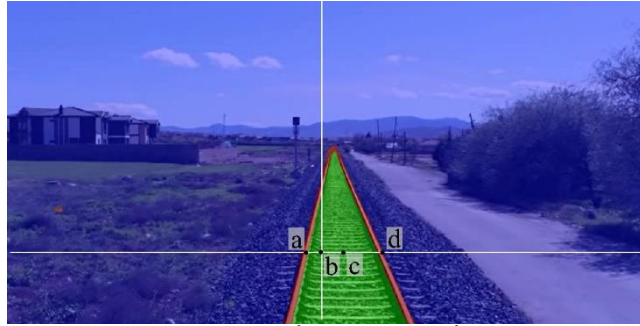
Şekil 13.a Elazığ demiryolu videosundaki anlık bir görüntüdür. Bu görüntü çalışmada seçilen ağlar tarafından segmente edilmiştir. Segmentasyon renkleri Şekil 10'dakine uygun olup, gerçek görüntü üzerine şeffaflaştırılarak bindirilmiştir. Çalışma videosuna YouTube linkinden ulaşılabilir (Kırat, 2023). Genel olarak ICNet daha stabil bir segmentasyon sağlamaktadır. BiSeNet V2 ile UNet ise ara sıra ray dışındaki çok küçük bölgeleri de ray varmış gibi segmente etmektedir. Paylaşılan video bağlantısından söz konusu durum gözlemlenebilir. Tablo 4'te bu çalışmadan elde edilen sonuçlar ile literatürdeki bazı çalışmaların sonuçları görülmektedir.

**Tablo 4.** Literatürdeki Çalışmalarla Karşılaştırma

Çalışma	Model	Giriş boyutu	Donanım	Segmentasyon Başarısı	FPS
(Yang vd., 2022a)	DFR Xception	640×480	-	0,8650	6,9
	DFR CSPDarkNet-S			0,8590	12,5
(Mammeri vd., 2021)	Rail UNet	832×832	RTX 2080 GPU	0,9594	25,85
(Yang vd., 2022b)	RT-GAN	720×560	Tesla V100 GPU	0,8807	34,66
(Tong vd., 2022)	FDRNet	1024×512	RTX 2060 GPU	0,5882	32,3
(Tong vd., 2023)	ARTNet	512×512	RTX 3090 GPU	0,7612	>50
(Wu vd., 2023)	YOLARC	520×520	RTX 3080 Ti GPU	-	35,9
	ICNet			0,9835	14,8
Çalışmamız	BiSeNet V2	1024×512	RTX A5000 GPU	0,9843	14,5
	UNet			0,9837	13,3

Tablo 12'ye göre gerçek zamanlı segmentasyon görevinde kullanılan evrişimli sinir ağı modeli, giriş görüntüsü boyutu, kullanılan donanım, segmentasyon başarısı ve saniyede işlenen kare sayısı (fps) değerleri görülmektedir. Özellikle giriş görüntü boyutunun büyümesi saniyede işlenen kare sayısını düşürmektedir. Aynı giriş görüntülerinde farklı fps değerleri model mimarilerine ve kullanılan donanıma göre de değişmektedir. Aynı giriş görüntüsündeki yüksek fps değerinde segmentasyon başarısının düşük olabildiği dikkat çekmektedir.

PID sistemine veri sağlamak için bir dizi görüntü işleme tekniği kullanmak gerekir. Ağlardan alınan segmente edilmiş görüntülerdeki renkler bu iş için kullanılmaktadır. Şekil 14'te İHA ile alınan video görüntüsündeki ICNet ağı çıktısındaki bir an için görüntü işleme adımlarının yapılacağı noktalar gösterilmektedir.



Şekil 14. PID Kontrolü İçin Görüntü İşleme Noktaları

Şekil 14'te görülen dikey beyaz çizgi 1024×512 ölçüsündeki giriş görüntüsünün yataydaki simetri merkezi olup soldan sağa 512. piksellerin bulunduğu sütundur. İHA'nın rayı ortalaması için bu çizgiye göre kendini ayarlaması gerekir. Yatay beyaz çizgi ise görüntünün yukarıdan aşağıya 412. sıradaki piksellerinden oluşan satırdır. Dikey eksenin simetri merkezi perspektif bakışa göre uzakta kaldığından daha yakın bir nokta olan 412. satır seçilmiştir. Bu referans satırı daha yukarıda seçmek, daha uzak bir noktayı seçmek anlamına geleceğinden İHA'nın çok hızlı ilerlemesine neden olabilir. Deneysel olarak bu satır uygun görülmüştür. a noktası kırmızı renkle segmente edilmiş sol ray ile yatay beyaz çizginin kesişimini, d noktası sağ ray ile yatay beyaz çizginin kesişimini ifade etmektedir. b noktası beyaz renkte belirlenen satır ve sütunun kesişim noktasıdır. c noktası ise a ve d noktasının tam ortasını ifade etmektedir. İHA'nın o anki konumu b noktasına göre olup, istenilen durum c ile b noktasını mümkün olduğunca çakışacak hale getirmektir. b ve c noktaları arasındaki fark ise hatadır. Şekil 14'e göre hesaplanacak hata sonucunda İHA'nın hafif sağa yönelerek c noktasını b noktasına bindirmesi gerekmektedir. c noktası Denklem 2 ile, hata ise Denklem 3 ile kolaylıkla hesaplanabilir.

$$c = a + (d - a)/2 \quad (2)$$

$$e(t) = 512 - c \quad (3)$$

Denklem. 2'de görülen 512 değeri, Denklem 1'de kullanılacak olan  $r(t)$  referans değeridir. Denklem 2'deki c değeri ise Denklem 1'deki  $y(t)$ 'ye karşılık gelmektedir. Buna göre İHA tarafından alınan video görüntüsünün her karesinde PID sistemi çalışarak hata hesaplanır ve İHA'ya gerekli hareketi yapması için komut gönderilir. Bu sayede İHA ray üzerinde otonom olarak ilerleyebilir.

## SONUÇLAR

Demiryolu görüntüsündeki ray hatları, bakış perspektifine göre aşağıdan yukarıya daralan birer çizgi çifti olarak görülmektedir. Derin öğrenme alanında yapılan görüntü segmentasyonu görevlerinde donanım kaynaklarının sınırlı olması nedeni ile genelde giriş görüntüleri çok küçük boyutlara ölçeklenmektedir. Bu durumda görüntüde veri kaybı olmakta, ray hatları küçülen resimde incecik iki çizgi haline gelmektedir. Rayları otonom olarak takip edecek bir İHA'nın aynı anda ray yüzeyindeki kusurları da tespit edebilmesi için daha yüksek çözünürlükteki görüntülerin kullanılması önem taşımaktadır. Gelişen teknoloji ile donanım kaynakları gittikçe artmakta, görüntü alan cihazlar daha yüksek çözünürlükte görüntü alabilmekte, bilgisayarların GPU'ları daha büyük kapasitelere ulaşmaktadır. Evrişimsel sinir ağları ise esnek yapıları ile çok farklı mimarilerde tasarlanabilmektedir. Gerçek zamanlı segmentasyon görevlerinde daha yüksek görüntülerde daha iyi FPS değerlerine ulaşabilmek için araştırmacıların güçlenen donanımlar ile farklı evrişimsel sinir ağlarını denemeye devam etmeleri gerekmektedir.

Bu çalışmada demiryolu işletmelerinin ray kontrollerini İHA ile otonom yapabilmesi için gereken görüntü anlamlandırma görevi 3 farklı evrimsel sinir ağı ile denenmiştir. Railsem19 veri seti çalışmaya uygun biçimde yeniden düzenlenerek literatürde görüntü segmentasyonu görevinde kullanılan UNet, ICNet, BiSeNet V2 ağları eğitilmiş ve Elazığ demiryolundan İHA ile alınan gerçek görüntüler üzerinde denenmiştir. Çalışmada kullanılan ağlar veri setindeki test görüntülerinde %98 segmentasyon doğruluğuna, 1024×512 çözünürlüğüne sahip video görüntülerinin segmentasyonunda yaklaşık 15 fps değerine ulaşılmıştır. Ayrıca çalışmada otonom sürüş için bir PID sistemi önerilmiştir. Çalışma, görüntü anlamlandırma görevinin demiryollarında uygulanması açısından önem taşımakta, alternatif bir ray kontrolörü olarak İHA'ların işe koşulması ile otonom ray kontrolü açısından demiryolu işletmelerine katkı sağlamakta, bu alanda çalışacak araştırmacılara fikir vermektedir.

## TEŞEKKÜR

Bu çalışma Fırat Üniversitesi Bilimsel Araştırma Projeleri tarafından ADEP.22.02 proje numarası ile desteklenmiştir.

## KAYNAKLAR

- Anadolu Ajansı. (2023). Deprem tren raylarını tel gibi büktü. [https://www.ntv.com.tr/galeri/turkiye/deprem-tren-raylarini-tel-gibi-buktu,j6Y22jcDNk2TPmVE6ZCoA/JUO5LUJ0SkyT\\_r2jLbTUWg](https://www.ntv.com.tr/galeri/turkiye/deprem-tren-raylarini-tel-gibi-buktu,j6Y22jcDNk2TPmVE6ZCoA/JUO5LUJ0SkyT_r2jLbTUWg) Erişim: 17.04.2023.
- Aydın, I., Sevi, M., Sahbaz, K., & Karakose, M. (2021). Detection of Rail Defects with Deep Learning Controlled Autonomous UAV. *2021 International Conference on Data Analytics for Business and Industry, ICDABI 2021*, 500–504. <https://doi.org/10.1109/ICDABI53623.2021.9655796>
- Bayati, A. M. A. (2019). *Evrimsel Sinir Ağları Kullanarak Drone Tarafından Elde Edilen Görüntülerde Nesne Tanıma*. Yüksek Lisans Tezi. Selçuk Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Müh. A.B.D., Konya 72s.
- Bojarczak, P., & Lesiak, P. (2021). UAVs in rail damage image diagnostics supported by deep-learning networks. *Open Engineering*, 11(1), 339–348. <https://doi.org/10.1515/eng-2021-0033>
- Çakmak, V., & Altaş, A. (2018). Sosyal Medya Etkileşiminde Tren Yolculukları: DOĞU EKSPRESİ İle İlgili Youtube Paylaşım Videolarının Analizi. *Journal of Tourism and Gastronomy Studies*, 6(1), 390–408. <https://doi.org/10.21325/jotags.2018.194>
- Chakravarthy, A. S., Sinha, S., Narang, P., Mandal, M., Chamola, V., & Yu, F. R. (2022). DroneSegNet: Robust Aerial Semantic Segmentation for UAV-Based IoT Applications. *IEEE Transactions on Vehicular Technology*, 71(4), 4277–4286. <https://doi.org/10.1109/TVT.2022.3144358>
- Chen, P., Wu, Y., Qin, Y., & Yang, H. (2022). All-in-One YOLO Architecture for safety Hazard Detection of Environment along High-Speed Railway. *2022 Global Reliability and Prognostics and Health Management (PHM-Yantai)*, 1–7. IEEE. <https://doi.org/10.1109/PHM-Yantai55411.2022.9941973>
- Grandio, J., Riveiro, B., Soilán, M., & Arias, P. (2022). Point cloud semantic segmentation of complex railway environments using deep learning. *Automation in Construction*, 141. <https://doi.org/10.1016/J.AUTCON.2022.104425>
- Guclu, E., Aydın, I., & Akin, E. (2021). Development of Vision-Based Autonomous UAV for Railway Tracking. *2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, 3ICT 2021*, 120–125. <https://doi.org/10.1109/3ICT53449.2021.9581919>
- Güçlü, E., Aydın, İ., & Akin, E. (2022). Mask R-CNN Algoritmasını Kullanarak Demiryolu Travers Eksikliklerinin Tespiti İçin Otonom İHA Tasarımı. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 34(1), 409–420. <https://doi.org/10.35234/fumbd.1039995>
- International Union of Railways. (2022). 8th UIC Refugee Task Force meeting held on 6 May 2022. <https://uic.org/com/enews/article/8th-uic-refugee-task-force-meeting-held-on-6-may-2022> Accessed: 10.05.2022.
- Katar, O., & Duman, E. (2022). Automated Semantic Segmentation for Autonomous Railway Vehicles. *TECHNICAL JOURNAL*, 16, 484–490. <https://doi.org/10.31803/tg-20220329114254>
- Kırat, S. S. (2023). Elazığ Demiryolu Segmentasyonu. <https://youtu.be/piVTdsDPzfg> Erişim: 09.04.2023.
- Kupriyanovsky, V., Pokusaev, O., Klimov, A., Dobrynin, A., Lazutkina, V., & Potapov, I. (2020). BIM on the world's railways-development, examples, and standards. *International Journal of Open Information Technologies*, 8(5), 57–80.

- Mammeri, A., Jabbar Siddiqui, A., & Zhao, Y. (2021). UAV-assisted Railway Track Segmentation based on Convolutional Neural Networks. *IEEE Vehicular Technology Conference, 2021-April*, 1–7. IEEE. <https://doi.org/10.1109/VTC2021-Spring51267.2021.9448887>
- Murat, S. (2021). *İnsansız Hava Aracı Görüntülerinden Derin Öğrenme ile Nesne Tanıma*. Yüksek Lisans Tezi. Maltepe Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, İstanbul 108.
- Ni, X., Liu, H., Ma, Z., Wang, C., & Liu, J. (2022). Detection for Rail Surface Defects via Partitioned Edge Feature. *IEEE Transactions on Intelligent Transportation Systems*, 23(6), 5806–5822. <https://doi.org/10.1109/TITS.2021.3058635>
- Rahman, M. A., & Mammeri, A. (2021). Vegetation Detection in UAV Imagery for Railway Monitoring. *Proceedings of the 7th International Conference on Vehicle Technology and Intelligent Transport Systems*, 457–464. SCITEPRESS - Science and Technology Publications. <https://doi.org/10.5220/0010439904570464>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. İçinde N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Ed.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-24574-4>
- Sevi, M., & Aydın, I. (2022). Rail Tracking and Detection with Drone in Gazebo Environment. *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, 1450–1454. IEEE. <https://doi.org/10.1109/DASA54658.2022.9765014>
- T.C. Ulaştırma ve Altyapı Bakanlığı. (2023). 2053 Ulaştırma ve Lojistik Ana Planı <https://www.uab.gov.tr/uploads/pages/bakanlik-yayinlari/20221025-2053-ulasirma-ve-lojistik-ana-planı-tr.pdf>  
Erişim: 05.04.2023.
- Tiu, E. (2019). Metrics to Evaluate your Semantic Segmentation Model. <https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2> Accessed: 17.01.2022
- Tong, L., Jia, L., Geng, Y., Liu, K., Qin, Y., & Wang, Z. (2023). Anchor-adaptive railway track detection from unmanned aerial vehicle images. *Computer-Aided Civil and Infrastructure Engineering*, 1–19. <https://doi.org/10.1111/mice.13004>
- Tong, L., Wang, Z., Jia, L., Qin, Y., Wei, Y., Yang, H., & Geng, Y. (2022). Fully Decoupled Residual ConvNet for Real-Time Railway Scene Parsing of UAV Aerial Images. *IEEE Transactions on Intelligent Transportation Systems*, 23(9), 14806–14819. <https://doi.org/10.1109/TITS.2021.3134318>
- Weng, Y., Li, Z., Huang, X., & Chen, X. (2023). *Improved DeepLabV3+ based Railway Track Extraction to Enhance Railway Transportation Safety*. [https://doi.org/10.1007/978-981-99-0272-9\\_16](https://doi.org/10.1007/978-981-99-0272-9_16)
- Wu, Y., Meng, F., Qin, Y., Qian, Y., Xu, F., & Jia, L. (2023). UAV imagery based potential safety hazard evaluation for high-speed railroad using Real-time instance segmentation. *Advanced Engineering Informatics*, 55(May 2022), 101819. <https://doi.org/10.1016/j.aei.2022.101819>
- Yang, H., Li, X., Guo, Y., & Jia, L. (2022a). Discretization–Filtering–Reconstruction: Railway Detection in Images for Navigation of Inspection UAV. *IEEE Transactions on Instrumentation and Measurement*, 71, 1–13. <https://doi.org/10.1109/TIM.2022.3220295>
- Yang, H., Li, X., Guo, Y., & Jia, L. (2022b). RT-GAN: GAN Based Architecture for Precise Segmentation of Railway Tracks. *Applied Sciences*, 12(23), 12044. <https://doi.org/10.3390/app122312044>
- Yu, C., Gao, C., Wang, J., Yu, G., Shen, C., & Sang, N. (2021). BiSeNet V2: Bilateral Network with Guided Aggregation for Real-Time Semantic Segmentation. *International Journal of Computer Vision*, 129(11), 3051–3068. <https://doi.org/10.1007/s11263-021-01515-2>
- Zendel, O., Murschitz, M., Zeilinger, M., Steininger, D., Abbasi, S., & Beleznaï, C. (2019). RailSem19: A Dataset for Semantic Rail Scene Understanding. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Vienna: Austrian Institute of Technology.
- Zhao, H., Qi, X., Shen, X., Shi, J., & Jia, J. (2018). ICNet for Real-Time Semantic Segmentation on High-Resolution Images. İçinde *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*: C. 11207 LNCS (ss. 418–434). [https://doi.org/10.1007/978-3-030-01219-9\\_25](https://doi.org/10.1007/978-3-030-01219-9_25)