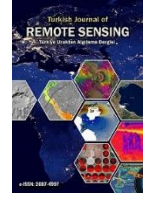




Türkiye Uzaktan Algılama Dergisi

<https://dergipark.org.tr/tr/pub/tuzal>

e-ISSN 2687-4997



Göktürk-1 Uydu Görüntülerinden U-Net Modeli Kullanılarak Binaların Segmentasyonu

Duygu ARIKAN*¹, Ferruh YILDIZ¹

¹Konya Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Harita Mühendisliği Bölümü, Konya, Türkiye

Anahtar Kelimeler:

Derin Öğrenme
U-Net mimarisi
Göktürk-1 Uydu Görüntüsü
Bina çıkarımı

ÖZ

Nüfus artışı, kentsel bölgelerde plansız yapılaşmanın ortaya çıkmasına yol açmaktadır. Bu durum dünya genelinde bir sorun haline gelmiştir. Bu alanların belirlenmesi ve tespit edilmesi, kentsel yönetim ve yeniden yapılanma planlaması için büyük öneme sahiptir. Ancak bu işlemler, arazide maliyetli ve zaman alıcı olabilmektedir. Uzaktan algılama görüntüleri kullanarak kentsel ve kırsal bölgelerde plansız yapılan binaları otomatik olarak tespit etmek ve karakterize etmek oldukça zordur. Son zamanlarda, derin öğrenme yöntemleri sayesinde karmaşık binaların tespiti mümkün hale gelmiştir. Bu çalışmada, Ankara'nın Etimesgut ilçesinden bir bölgenin bina çıkarımı işlemi, U-Net derin öğrenme mimarisi kullanılarak gerçekleştirilmiştir. İşlem için Inria Aerial Image Labeling adlı hazır bir veri seti kullanılmıştır. Eğitim işlemi için farklı sayıda görüntü (500, 1000, 2500, 5000) seçilmiştir. En iyi öğrenme sonucu, 0.5 m uzamsal çözünürlüğe sahip Göktürk-1 uydu görüntüleriyle test edilmiştir. Sonuçlara göre, U-Net modelinin bina segmentasyonunda Jaccard katsayısı 0.862, Dice benzerlik oranı 0.813 olarak bulunmuştur. Hazır veri seti kullanılarak U-Net modelinin derin öğrenme yöntemleri için kullanılabilir olduğu kanıtlanmıştır. Bu çalışma, kentsel alanlardaki binaların tespiti ve haritalanmasında derin öğrenme yöntemlerinin etkinliğini ve potansiyelini göstermiştir.

Segmentation of Buildings Using U-Net Model from Göktürk-1 Satellite Images

Keywords:

Deep Learning
U-Net Model
Gokturk-1 Satellite Image
Building Detection

ABSTRACT

The increase in population has led to unplanned urbanization in urban areas, becoming a global issue. The identification and detection of these areas are of great importance for urban management and redevelopment planning. However, these processes can be costly and time-consuming when conducted on-site. Automatic detection and characterization of unplanned buildings in urban and rural areas using remote sensing imagery is a challenging task. Recently, with the advancements in deep learning methods, the detection of complex buildings has become possible. In this study, the building extraction process of a region from the Etimesgut district of Ankara was performed using the U-Net deep learning architecture. The Inria Aerial Image Labeling dataset, a publicly available dataset, was used for the process. Different numbers of images (500, 1000, 2500, 5000) were selected for the training process. The best learning outcome was tested with Göktürk-1 satellite imagery with a spatial resolution of 0.5 m. According to the results, the U-Net model achieved a Jaccard coefficient of 0.862 and a Dice similarity coefficient of 0.813 for building segmentation. The effectiveness and potential of deep learning methods were demonstrated using the U-Net model with the available dataset. This study showcased the efficiency and potential of deep learning methods in the detection and mapping of buildings in urban areas.

Makale Bilgileri/Article Info

Geliş /Received: 23/05/2023
Kabul/Accepted: 23/06/2023
Yayınlanma/Published: 30/06/2023

Alıntı/Citation:

Arıkan, D. & Yıldız, F. (2023). Göktürk-1 Uydu Görüntülerinden U-Net Modeli Kullanılarak Binaların Segmentasyonu. Türkiye Uzaktan Algılama Dergisi , 5 (1) , 50-58.

1. GİRİŞ

Son yirmi yılda insan nüfusu giderek artmakta ve şehirleşme söz konusu olmaktadır. Şehir planlamalarının yapılmasında, afet alanlarının tespit edilmesinde ya da zamansal olarak kentsel değişikliklerin izlenmesi gibi çalışmalarda binaların, yolların, arazi sınıflarının belirlenmesi önem arz etmektedir. Bina konumlarının, geometrik yapılarının belirlenmesi ve haritalanması birçok disiplin için altlık olarak kullanılmaktadır. Bozulan kentsel alanların yeniden yapılandırılması, plansız yerleşim alanlarının tespit edilmesi şehir planlamacıları ve belediye çalışmalarında gerekli olmaktadır (Kuffer vd., 2016; Mahabir vd., 2018; Wurm vd., 2019).

Örneğin; çalışma alanı olarak seçilen Ankara'nın bazı bölgelerinde yüksek katlı binalar, bazı kısımlarında tek katlı bahçeli yapılar bulunurken, bazı kısımlarında da gece kondu mahalleleri bulunmaktadır. Bu durum şehrin fiziksel ve görsel algısını bozmaktadır. Bu tarz yapıların belirlenmesi, kentsel ya da kırsal alanlardaki binaların tespiti, konumlandırılması, öznelik bilgilerinin belirlenmesi ve sayısallaştırılması işlemi arazide yapılan saha çalışmalarına dayanmaktadır (Wurm vd., 2019). Kentsel bölgelerdeki bina yoğunluğunun fazla olması nedeniyle saha çalışmalarının süre bakımından yavaş olmasına buna bağlı olarak fazla bir maliyet ve emek gerektirmektedir. Bu sebepten, uydu görüntülerinin mekânsal çözünürlüğünün artması ile kentsel alanlarda bina gibi yapıların, ulaşım hatlarının otomatik olarak algılanmasına yönelik çalışmalar uzaktan algılama ile yapılabilmektedir. Bu teknoloji sayesinde düşük maliyet ile büyük alanlara ait bilgiler kısa süre içerisinde elde edilebilir. Ayrıca yapılan işin sürekli güncel olması da sağlanmış olur (Pan vd., 2020; Wurm vd., 2019). Fakat, görüntü üzerinde bulunan meydanlar, yollar gibi detaylar renk (spektral) ve şekil itibarıyla bina özelliklerine benzemektedir. Bu karmaşık arka plan nesnelere binayı tespit etmede zorluk haline getirmektedir.

Bu nedenle kentsel alanlardaki binaların tespit edilmesi ve uzaktan algılama tekniğinden yararlanılarak harita üretim işlemi önem arz etmekle birlikte literatürde uzun bir geçmişe sahiptir. Yüksek yapı binalar ya da bölgedeki bina miktarının yoğunlaştığı kısımlarda, geleneksel görüntü segmentasyon ya da sınıflandırma yöntemlerinin süre yönünden yeterli olmadığı tespit edilmiştir. Bina segmentasyonda, araştırmacılar piksel tabanlı özellik tanımlayıcıları tasarlamışlardır. Örneğin; Tuermer vd. (2013), çalışmada kentsel bölgelerdeki nesnelere tespit etmek amacıyla eğitim histogramını (HOG-Histogram of oriented Gradients) kullanmışlardır. Zhang, vd. (2011) çalışmada ise, bina çıkarma işlemi için renk, doku ve şekil özellikleri birleştirilerek daha iyi sınıflandırma performansına ulaşabilmeyi hedeflemişlerdir. Klasik yöntemlere ek olarak, son zamanlarda, bu alanda yapılan çalışmalarda makine öğrenmesi ve derin

öğrenme yöntemleri popülerlik kazanmıştır. Mitra vd. (2004) uydu görüntüleri kullanılarak yapılacak denetimli piksel sınıflandırması çalışmalarında gerekli olan etiketli verinin azlığı problemini çözmek amacıyla makine öğrenmesi algoritması olan destek vektör makinesini kullanmışlardır. Pal (2005), arazi sınıflarını tespit etmek amacıyla rastgele orman algoritmasını kullanmışlardır. Yapılan çalışmalarda geleneksel makine öğrenme algoritmalarının, çözümlenemediği ve eksik kaldığı kısımlar olduğu anlaşılmıştır. Etiketli veriyi oluşturmak için manuel bir işlem adımı olması dezavantaj olarak görülmektedir. Ayrıca bu çalışmalarda kullanılan donanımın da iyi olması gerekmektedir. Çünkü verilerin ya da görüntülerin işlenmesi süre bakımından önem arz etmektedir. Bu nedenle, GPU hesaplama hızındaki artış, Google Colab GPU kullanımının belirli sınırlamalar dahilinde ücretsiz kullanılması araştırmacıları derin öğrenme algoritma yöntemlerine yönlendirmiştir. Derin öğrenme algoritmalarından biri olan evrişimli sinir ağları (ESA) özellikle bina çıkarımı, yol çıkarımı gibi nesne çıkarımında başarılı bir şekilde kullanılmaktadır (Kattenborn vd., 2021; X. Li vd., 2023; Pan vd., 2020; Wurm vd., 2019). İlk olarak sağlık alanındaki görüntülerde, daha sonra yapılan çalışmalarda otonom araçlarda, artırılmış gerçeklik çalışmalarında, arazi sınıflarının belirlenmesi gibi çalışmalarda ESA tercih edilmektedir (Sariturk & Seker, 2022). Yöntemde ağ eğitimi için girdi verilerine ait etiketlenmiş veriler arasında ilişkiler kurulmaktadır. Ağ eğitimi kullanılan bilgisayarın performansına göre değişkenlik göstermekte ve dezavantaj olarak görülmektedir (Zhu vd., 2017). Bu sebepten ötürü farklı ağ mimarileri oluşturulmuştur ve bunlardan bir tanesi de U-Net mimarisidir. Görüntü segmentasyonu için tasarlanmış ve tam evrişimli olan bu mimarinin yapısı kodlayıcı (sol taraf)-kod çözücü (sağ taraf) iki ana kısımdan oluşmaktadır. Simetrik yapısından dolayı ismini buradan almıştır (Hou, vd., 2021; Ronneberger vd., 2015).

Kentsel alandaki çalışmalarda U-Net algoritması kullanılarak birden çok çalışma bulunmaktadır. Temeldeki amaç kentsel alanlardaki bina çıkarımı olsa da kullanılan materyal-metot farklı olabilmektedir. Çeşitli uydu görüntüleri, havadan alınan görüntüler ya da hava lidar teknolojilerinden yararlanılabilmektedir. Örneğin; L. Li vd. (2019), yaptıkları çalışmada Sentinel-1 SAR uydu görüntüsünü kullanırken, Abdollahi ve Pradhan (2021) ve Wu vd. (2018) hava görüntülerini kullanarak sınıflandırma işlemini gerçekleştirmişlerdir. Ayrıca U-Net ve diğer derin öğrenme algoritmalarını geliştirmek amacıyla kullanıcılara ücretsiz sunulan veriler de bulunmaktadır. Massachusetts (Mnih, 2013), Inria (Maggiore, vd., 2017) ve WHU bina veri kümeleri (Ji, Wei, & Lu, 2018) mevcut ve kullanıcılara açıktır.

Sinir ağlarının ve uydu verilerinin çeşitlilik kazanması, erişim imkanının kolay olması nedeniyle bu çalışmada, Inria Aerial Image Labeling Dataset

kullanılmıştır. Binaların segmentasyon işlemi U-Net mimarisi ile gerçekleştirilmiştir. Bu model kullanılırken eğitim verisinin ve epok sayısının sonuçlara olan etkisi irdelenmiştir. Ek olarak, en iyi optimizasyon sonucuyla, Göktürk-1 uydu görüntüsünden yararlanılarak bina segmentasyonu işlemi gerçekleştirilmiştir. Çalışmada mimarinin oluşturulması ve eğitilmesi, test işlemleri MSI dizüstü bilgisayarında gerçekleştirilmiştir. Bu çalışmanın literatüre katkısı U-Net mimarisi kullanılarak ilk kez Göktürk-1 uydu görüntüsünden bina tespiti çalışması gerçekleştirilmiştir. Yapılan bu çalışma 3 ana kısımdan oluşmaktadır. Birinci bölümünde, kullanılan veri seti, U-Net mimari ve çeşitli optimizasyon yöntemlerinin açıklanması, İkinci bölümünde farklı sayıda kullanılan veriler ile epok sayısının sonuçları ve Göktürk uydu görüntüsünde test edilmesi, son bölümde ise sonuçlar ve tartışma kısmıdır.

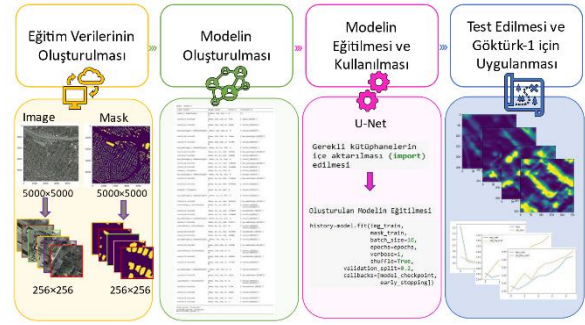
2. YÖNTEM

2.1. Veri Seti

Çalışmada eğitim verisi olarak "Inria Aerial Image Labelling Dataset (Etiketli Hava Görüntüsünün Veri Kümesi)" kullanılmıştır. Bu veri seti içerisinde bölgeler yoğun nüfuslu bölgeler (Austin, Chicago gibi) ve kırsal kasaba (Avusturya Tyrol) olacak şekilde farklı alanları içermektedir (Maggiori vd., 2017). Bu sayede farklı lokasyonlarda bulunan, farklı bina türlerinin bulunduğu görüntülerinin kullanılması modellemenin gücünü arttırmaktadır. Kullanılan görüntüler ve maskeler 5000×5000 piksel boyutunda olup, her biri $1500 \text{ m} \times 1500 \text{ m}$ 'lik bir alanı kaplamaktadırlar. Ayrıca birbirleriyle örtüşmeyecek şekilde ayarlanmış ve görüntülerin mekansal çözünürlüğü $0,3 \text{ m}$ 'dir. Çalışma kapsamında eğitim ve test amaçlı kullanılan görüntüler 3 bantlı (kırmızı, yeşil ve mavi) ve tek bantlı maskelenmiş görüntülerdir.

Çalışmanın son kısmında Inria veri setinde eğitilmiş modellerden en yüksek doğruluklu model alınarak, Göktürk-1 uydusunda test edilmiş ve sonuçları incelenmiştir. Bu uydu 2018 yılında fırlatılmış olup, RGB'de $0,5 \text{ m}$, pankromatikte de 2 m konumsal çözünürlüğe sahiptir. Uydu görüntüsü Ankara iline aittir. Şehir, Türkiye'nin başkenti olup, kentsel bölgelerde yoğun yapıların yer aldığı, bazı kısımlarında ise gece kondu mahallerinin bulunduğu bölgeler incelenmiştir. Çoğu makalede kullanılan görseller temel olarak üç kanaldan (Kırmızı, Yeşil ve Mavi) oluşur. Bu nedenle çalışmamızda hem örnek veri setinde hem de Göktürk-1 uydu görüntüsünden RGB görüntüler kullanılmıştır. Çalışmanın iş akış diyagramı Şekil 1'de sunulmuştur. Derin öğrenme ağı eğitim ve modellemesi aşamaları, dört çekirdekli

Intel Core i7 işlemci, NVIDIA GTX 850M GPU ve 32GB hafızaya sahip MSI dizüstü bilgisayarda yapılmıştır.

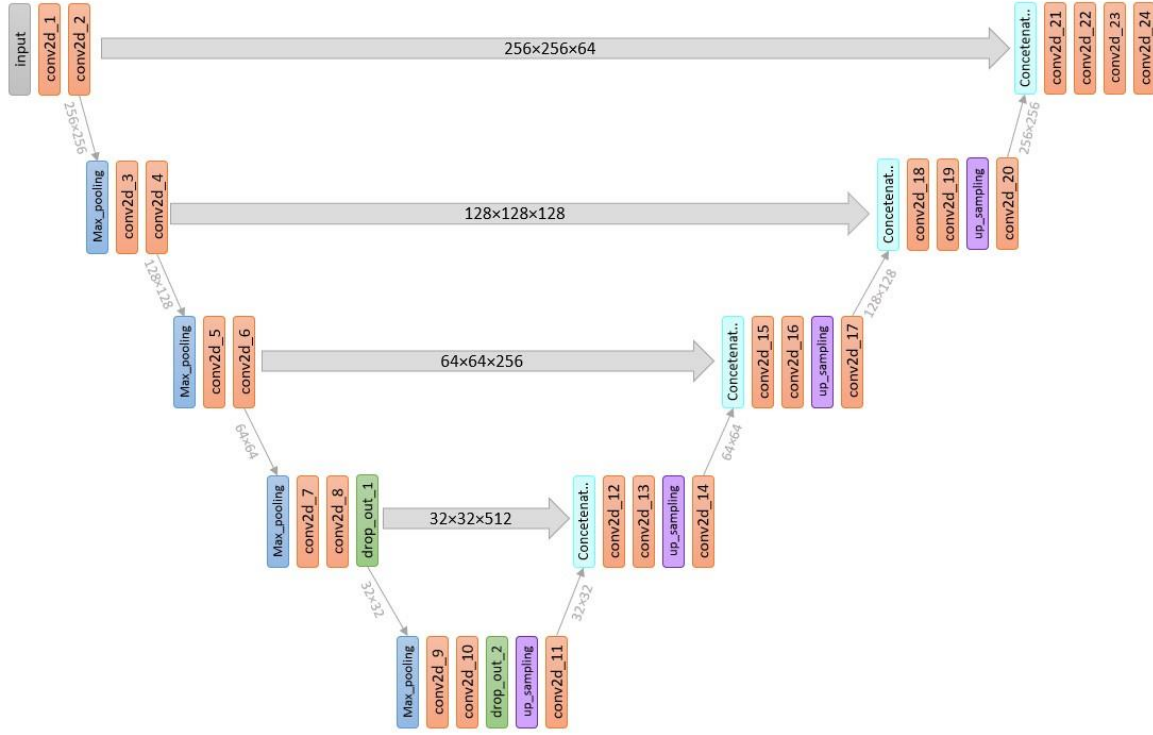


Şekil 1. İş - akış diyagramı

2.2. U-Net Mimarisi

Çalışmada U-Net mimarisi kullanılmıştır. Bu mimari yapı, Ronneberger ve diğerleri tarafından geliştirilmiş olan tekrarlayan evrişimli bloklardan oluşan bir sinir ağı mimarisidir (Ronneberger vd., 2015b; Sariturk & Seker, 2022). U-Net mimarisi daha çok segmentasyon, sınıflandırma ve görüntüler üzerinde çalışmalar yapmak amaçlı geliştirilmiş bir görüntü segmentasyon algoritmasıdır (Yan vd., 2022). Şekil 2'de yapılan çalışmadaki katman adımları verilmiştir ve modelin 'U' harfi şeklinde olduğu açıkça görülmektedir. Kod oluşturucu kısımda eğitim verilerinden özelliklerin çıkartılması sağlanır ve bir kodlayıcı blok dizi aracılığıyla giriş görüntülerinin soyut bir temsilini öğrenir. Her bir kod oluşturucu (enkoder) kısımda ReLU aktivasyon fonksiyonu tarafından takip edilen 3×3 evrişim ağından oluşmaktadır. RELU fonksiyonu eğitim verilerinin genelleştirilmesine yardımcı olurken, aynı zamanda ağında doğrusal olmamasına imkân vermektedir. Bu işlemlerden sonra özellik haritalarının ya da görsellerin boyutlarının yarıya indirildiği 2×2 maksimum havuzlama kısmı gelir. Kod çözümlenici blok kısmı 2×2 evrişim ağı ile başlar ve kodlayıcı bloğun atlamalı bağlantı özellik haritasıyla birleştirilir. Bundan sonraki her bir adımda evrişimin ReLU aktivasyon kodu tarafından takip edildiği 3×3 evrişim ağı kullanılır (Atlan, vd., 2020).

Kod çözücü kısımda ise eğitilen verilerin maskeleri oluşturulur. Bu kısım 2×2 devrik evrişim ağı ile başlar. Sonrasında kod oluşturucu bloktaki özellik haritasıyla birleştirilir. Bundan sonra, her bir evrişimin bir ReLU aktivasyon fonksiyonu tarafından takip edildiği iki 3×3 evrişim kullanılır. Son kod çözücünün çıkışı, sigmoid aktivasyonu ile 1×1 evrişimden geçer. Sigmoid aktivasyon işlevi, piksel bazında sınıflandırmayı temsil eden segmentasyon maskesini verir. Oluşturulan model Şekil. 2'de sunulmuştur.



Şekil 2. Çalışma da kullanılan U-Net Modelinin Katmanları

2.3. Optimize Edilen Parametreler

Yapay zekâ uygulamalarında, veri setinin aynı anda işlenmesi ve öğrenme durumu için model eğitiminde optimizasyon seçenekleri vardır. Bunlar; epok sayısı, batch size ve aktivasyon fonksiyonları olarak ifade edilen optimizasyon yöntemleri bulunmaktadır. İfade edilen seçeneklerden herhangi bir parametrenin değiştirilmesi modelin başarı performansını etkilemektedir (Öz, 2021).

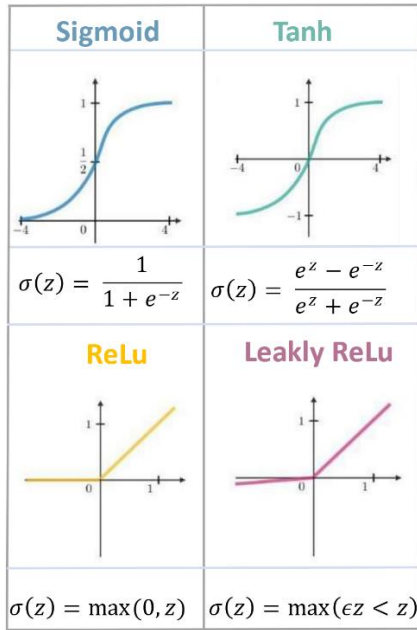
Batch size genellikle epok sayısı ile karıştırılan iki farklı kavramdır. Batch size öğrenme işlemi sırasında her iterasyon için hesap işleminin yapılmasını ifade ederken, epok sayısı da bir öğrenme için gerçekleştirilecek işlemidir. Batch size birden fazla girdi verisinin parçalar halinde işlenmesini ifade etmektedir. Böylelikle öğrenme işlemi gerçekleştirilirken her bir iterasyon işleminde ağırlık değerleri güncellenmekte ve hata oranları minimum seviyeye indirgenmektedir. Kısacası batch size verilerin her bir adımda işleme miktarını belirtirken, epok sayısı tam bir eğitim veri setinin kaç kez işleneceğini belirtir. Örneğin; 2500'lük bir eğitim veri setinin batch size değeri 5 seçildiğinde, bu veri seti 500'lük gruplara bölünür. 1 epoku tamamlamak için de 5 adet batch size işleminin yapılması gerekmektedir.

Bu parametrelerin sabit bir değer seçilmesi gibi bir durum söz konusu değildir. Modelin performansı aynı zamanda eğitim verisinin boyutlarına ve öğrenme hızına göre değişkenlik göstermektedir. Batch size değerinin 1 seçilmemesi tavsiye edilir. Çünkü, modelin genelleştirme yeteneklerini sınırlandırmaktadır (Öz, 2021).

Aktivasyon fonksiyonları, sinir ağlarında bir nöronun tetiklenip tetiklenmeyeceğine karar vermek amacıyla girdi verilerinin ve ağırlıklarının toplamını hesaplamak için kullanılan bir işlev görevindedir (Nwankpa, vd., 2018). Bu nedenle aktivasyon fonksiyonlarının seçimi ve kullanılması derin öğrenme algoritmalarında önemli bir yere sahiptir (Ramachandran, vd., 2017). Çünkü, algoritma içerisinde çok fazla katman (girdi, gizli ve çıkış) ve ağırlıklar tanımlanmaktadır. Ağ içerisindeki her bir katmana aktivasyon işlemi uygulanır ki, daha sonrasında bu katman bir başka katman için giriş katmanı olarak kullanılabilir.

Derin öğrenme algoritmalarında görüntü, video, ses gibi karmaşık yapıdaki bilgileri aktivasyon fonksiyonları sayesinde anlamlı özellikleri çıkartabilmektedir. Eğer bu fonksiyonlar kullanılmazsa, sinir ağı sınırlı öğrenme gücüne sahip tek dereceli bir polinom fonksiyonu işlemi yürütecektir. Fakat, girdi verileri sadece doğrusal yapıda olmadığı için, bu durum da ağı iyi eğitilememesini ifade eder. Bu nedenle doğrusal olmayan durumlarında öğrenilebilmesi için aktivasyon fonksiyonlarından yararlanılmaktadır. Yapay sinir ağlarında sıkça kullanılan aktivasyonlar, sigmoid, tanh, relu ve Leaky ReLU olarak ifade edilebilir. Aktivasyon fonksiyonları ve matematiksel modelleri Şekil 3'te sunulmuştur. Literatürde lojistik fonksiyon ya da ezici fonksiyon olarak da bilinen (logistic function or squashing function) sigmoid fonksiyonu karar vermeye yönelik olasılıksal bir yaklaşım olup, [0-1] aralığında değer almaktadır (Nwankpa vd., 2018). Eğer sürekli türevi alınabiliyorsa, bu fonksiyon girdi verisini öğrenmeye başladığı anlamına gelmektedir. Derin öğrenme

mimarilerinde daha çok çıktı katmanında yer almaktadır (Nwankpa vd., 2018). Tanh fonksiyonu sigmoid fonksiyonuna yapı olarak çok benzemektedir. Fakat tanh fonksiyonunda değer aralığı $[-1,1]$ ve orijin etrafında simetrik (Nwankpa vd., 2018). Bu fonksiyonun en önemli problemi uç noktalarda türevi sifıra yaklaşmaktadır. Bu yüzden modeldeki katman sayısına bakılarak bu fonksiyon seçimi gerçekleştirilir. Eğer model çok katmanlı bir ağdan oluşuyorsa sigmoid ve tanh fonksiyonlarının kullanılmaması tercih edilir. Daha çok evrişimli sinir ağlarında tercih edilen ReLu fonksiyonu tanh ve sigmoid fonksiyonuna göre daha verimli ve hızlı çalışmaktadır (Ramachandran vd., 2017; Zeiler vd., 2013). Bu neden ara katmanlarda bu fonksiyon tercih edilmektedir. Tanh fonksiyonunun tersine ReLu fonksiyonu çok katmanlı sinir ağlarında tercih edilmektedir. Fakat, ReLu fonksiyonu da geri yayılım sırasında pozitif olduğu durumda aktif işlem gerçekleşirken, negatif olduğu durumda sadece sıfır olması dezavantajı olarak karşımıza çıkmaktadır (Madhu, vd., 2023). Leaky (sızıntı) ReLu ise, ReLu fonksiyonundaki sınırlamayı kaldırmak amaçlı geliştirmiş olup, sıfır olmayan bir gradyana sahiptir (Maas, vd., 2013).



Şekil 3. Aktivasyon fonksiyonları ve matematiksel modelleri (Nwankpa vd., 2018)

2.4. Metrikler

Uydu görüntüleri kullanılarak segmentasyon işlemlerinde metriklerin hesaplanması gerekmektedir. Çünkü yöntemin performansının nicel olarak değerlendirilmesi gerekmektedir. Bunun için doğruluk, kesinlik, F-ölçüsü, Jaccard indeksi (IoU), Dice benzerlik katsayısı metriklerinden yararlanılmaktadır. Bu çalışma kapsamında görüntü segmentasyonunda yaygın olarak tercih edilen doğruluk ölçütü Dice benzerlik ve Jaccard katsayıları kullanılmıştır.

Dice Benzerlik katsayısı, maskelenmiş görüntü ile tahmin edilen görüntü arasındaki çakışma derecesini ifade eder. 0 ve 1 arasında değer almaktadır (Ataş, 2023). 1'e yakın olması durumunda model sınıflandırmasının iyi olduğunu ifade ederken, 0'a yakın olması durumunda tam tersi durum geçerlidir (Zhang et al., 2019). Matematiksel formülü Eşitlik 1'de verilmiştir.

$$DBK = \frac{2|X \cap Y|}{(|X| + |Y|)} \quad (1)$$

Jaccard katsayısı ise, benzerlik ölçmek için kullanılan bir istatistiksel yöntemdir (Zhang et al., 2019). İki kümenin kesişimlerinin birleşimlerine oranı olarak ifade edilmektedir (Ataş, 2023). Doğruluk değeri 0 ve 1 arasındadır. Matematiksel formülü Eşitlik 2'de verilmiştir.

$$J(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \quad (2)$$

Eşitliklerde kullanılan X maskelenmiş görüntüleri ve Y ise model sonucunda tahmin edilen görüntüyü ifade etmektedir.

U-Net mimarisi ve yapılan tüm segmentasyon işlemleri için Python programlama dili kullanılmıştır. Python, açık erişim tabanlı olup, kullanımı kolay, nesne yönelimli ve etkileşimi yüksek bir programlama dilidir. İçerisinde bulunan kütüphaneler vasıtasıyla, birçok işlemi doğrudan gerçekleştirilebilir (Bayraktar, 2018). Bu çalışmada kodlama işlemi yapılırken farklı kütüphanelerden yararlanılmıştır. Bunlar; numpy, os, cv2, h5py, matplotlib, keras ve tensorflowdur.

İşlemlerdeki hesaplama maliyetini azaltmak ve veri kümesinin boyutunu artırmak amacıyla 5000 × 5000 piksellik görüntüler ve etiketlenmiş görüntülerin boyutu 256 × 256 piksel olacak şekilde yeniden boyutlandırıldı. Toplamda 20.090 görüntü bulunmaktadır. Fakat çalışmada kullanılan görüntülerin sayısı Tablo 1'de sunulurken, dört kısımda ele alınmıştır. 1. Kısım da görüntüler eğitilirken yaklaşık %50'si eğitim, %25'i test ve %25'i de doğrulama verisi olarak kullanılmıştır. Sonraki kısımlarda test ve doğrulama sayıları aynı tutulmuş, fakat eğitim verisindeki görüntü sayıları artırılmıştır. Sırasıyla 1000, 2500 ve 5000 olarak seçilmiştir. Çalışmada farklı eğitim ve test verilerinin sonuca katkısı incelenmiş olup, çalışmada kullanılan görüntülerin sayıları tabloda verilmiştir.

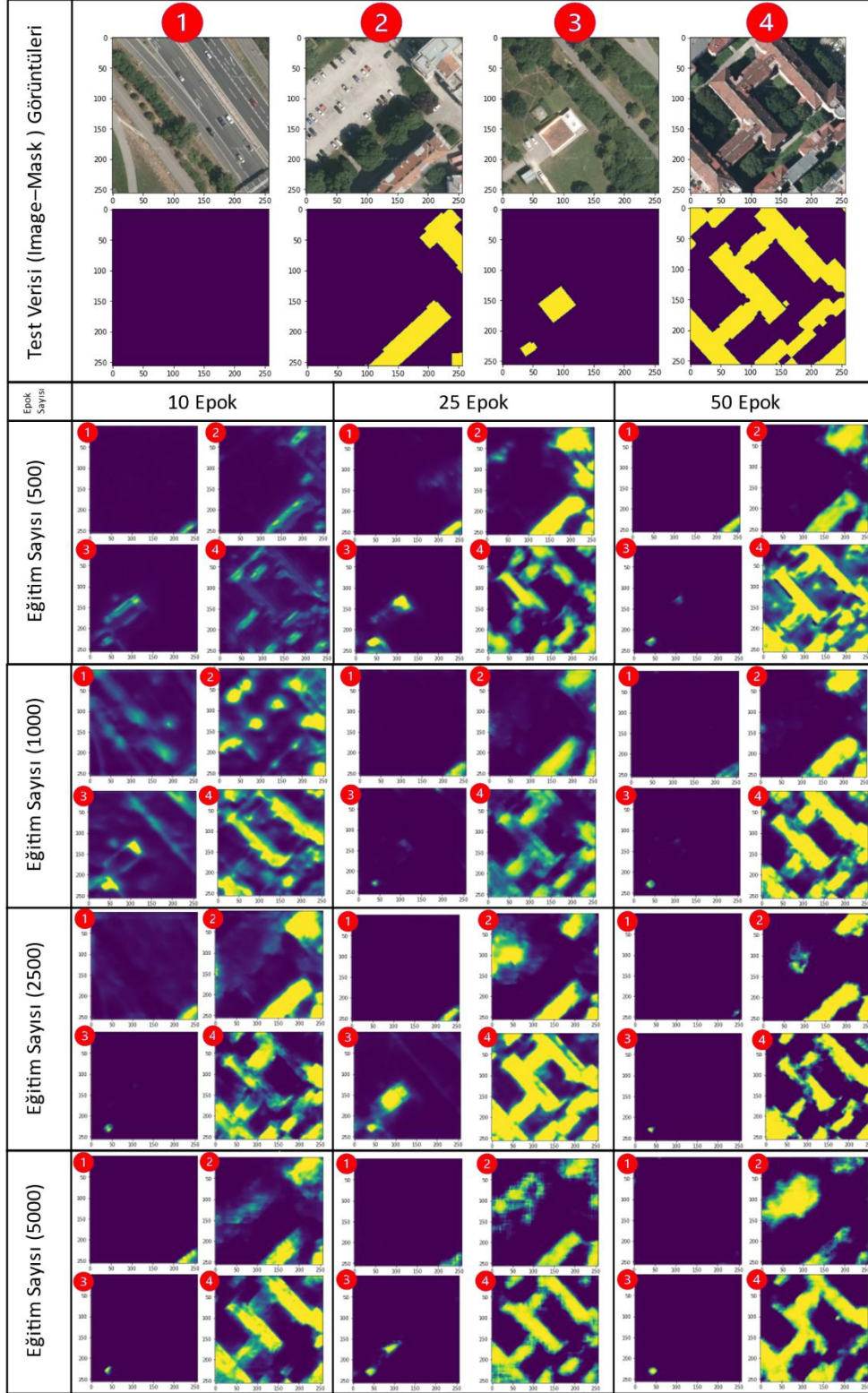
Tablo 1. Modelde kullanılan eğitim, test ve doğrulama verilerinin sayısı

	Eğitim Verisi	Test Verisi	Doğrulama Verisi
1. Kısım	500	250	250
2. Kısım	1000	250	250
3. Kısım	2500	250	250
4. Kısım	5000	250	250

3. BULGULAR

Farklı sayıdaki her bir eğitim verisi için model çalıştırılmış ve Şekil 4'te sunulmuştur. Çok fazla görüntü olduğu ve sonuçları görsel olarak ifade edebilmek için, test verisi içerisinde bina yoğunluğun fazla olduğu alanlardan, hiç olmayacak şekilde 4 adet görüntü seçilmiştir ve model sonunda

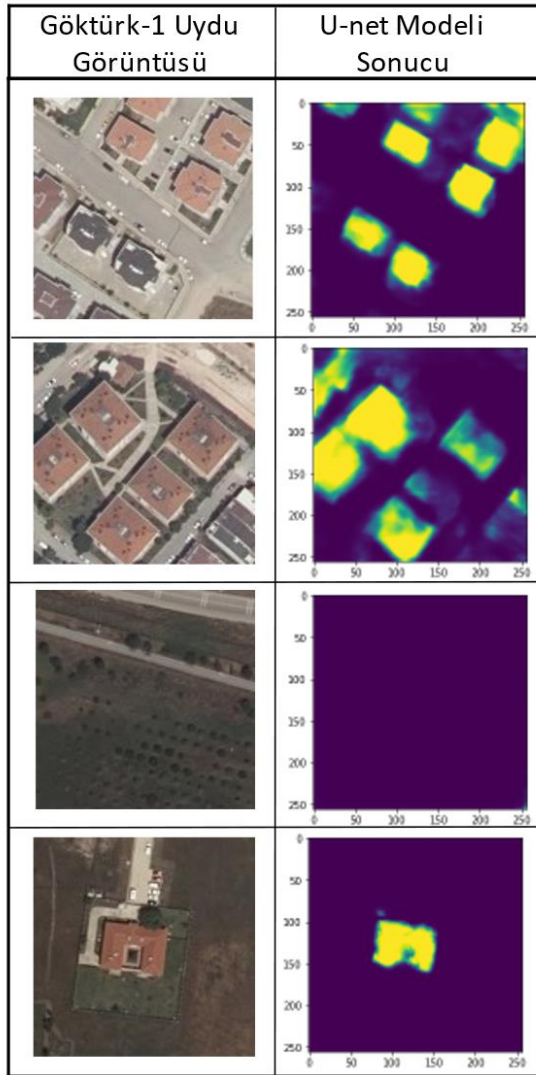
öğrenme durumları verilmiştir. Şekil 4'te 1 numaralı görüntü de bina bulunmamaktadır, 2 ve 3 numaralı görüntüde küçük iki bina vardır. 4 numaralı görüntüde ise karmaşık bir bina görüntüsü yer almaktadır. 2 numaralı görüntü de ayrıca otoparkın bulunması spektral ve dokusal olarak binaya benzemektedir. Bundan dolayı model sonucunda doğruluğu test etmek amacıyla özellikle seçilmiştir.



Şekil 4. Öğrenme sonucunda elde edilen görüntüler

Çalışma da epok sayısındaki miktarının sonuca olan etkisi incelenmiştir. Epok sayısı arttıkça eğitim ve doğrulamanın doğruluğu artarken, kayıp oranında azalmalar olduğu tespit edilmiştir. Çalışmada eğitim sayısı 500 ve epok sayısı 10 iken, öğrenme durumu neredeyse olmamıştır. Fakat epok sayısı arttıkça binaların bulunduğu kısımlar belirginleşmeye başlamıştır. Sadece öğrenme işlemi yeterli olmamıştır. Bu nedenle çalışmada eğitim verisinin miktarı artırılmıştır. Eğitim sayısı 1000 ve epok sayısı 10 olduğunda 1 numaralı görüntü de bina olmamasına rağmen, yol ağlarının rengini bina olarak algılamış, aynı şekilde 2 numaralı görüntüde otopark kısmını ve bazı araçlar, 3 numaralı görüntü de ise yol ağı bina olarak algılanmıştır. 10 epok için tüm eğitim verileri incelendiğinde eğitim sayısı 5000 olan modelin sonucu diğerlerine göre daha iyi olduğu anlaşılmaktadır.

Doğruluk artışı ve kayıp düşüşü başlangıçta hızlı bir şekilde değişime uğrarken, ardından olarak sabit hale geldi. 10 epokluk eğitimde 4. epoktan sonra, 25 epokluk eğitimde 11. epoktan sonra ve 50 epokluk eğitimde 21. epoktan sonra doğruluk artışı ve kayıp miktarındaki değişimler azalmaya başlamıştır.



Şekil 5. U-Net modeli sonucu Göktürk-1 görüntülerinden bina çıkarımı

Eğitim verisi 2500 ve epok sayısı 25 için Göktürk-1 verilerinden U-Net modeli kullanılarak bina çıkarımı gerçekleştirilmiştir. U-Net modelinin bina segmentasyonu için Jaccard katsayısı için 0.862 iken, Dice benzerlik oranı 0.813 çıkmıştır. Her iki doğruluk ölçütü de benzer sonuçlar verdiği görülmüştür. Maskelenmiş görüntüler ve tahmin edilen görüntülerin %81 oranında doğru eşlendiği tespit edilmiştir. Bundan dolayı aynı model Göktürk için de uyarlanmış ve test edilmiştir. Şekil 5' te sonuçları verilmiştir.

4. SONUÇLAR

Kentsel bölgelerde otomatik olarak bina çıkarım işlemi uzun ve zahmetli olmaktadır. Bu nedenle son zamanlar uydu görüntüsünden nesne çıkarımı kullanılan birçok yapay zekâ teknolojisi kullanılmıştır. Derin öğrenme yöntemleri diğer tekniklere göre nispeten yeni olmasına rağmen kullanılan mimariler hızlı bir şekilde geliştirilmektedir.

Yapılan bu çalışmada hazır veri seti kullanılarak, eğitim modellemesi gerçekleştirilmiş olup Göktürk-1 uydu görüntüleri test amacıyla kullanılmıştır. Sonuç olarak, ağ modeli, binaları tespit etmede iyi performans göstermiştir. Göktürk-1'den tahmin edilen görüntüler, kentsel ortamları sınıflandırmada oldukça iyi performans gösterdiği görülmüştür. Şekil 5'te görüldüğü üzere, çoğu binanın doğru bir şekilde ayrıldığı görülmektedir. Buda, hazır veri setinden yararlanılarak Göktürk-1 uydu görüntüsünün ve U-Net ağının yüksek yoğunluklu kentsel alanlarda bu tür binaları ayırt etme yeteneğini göstermektedir. Ayrıca, modelde eğitim ve epok sayısının, sonuçlara olan etkisi de ele alınmış olup değerlerin artması sonuçlara olumlu katkı sağlamıştır. Fakat bir modeli eğitmek için şu kadar veri veya epokta yapılması gerektiğini söylemek zordur. Çünkü, kullanılan ağ mimarisinde eğitim verisinin miktarına, kullanılan görüntülerin çözünürlüğüne bağlı olarak duyarlılık analizin yapılması gerekmektedir. Yeterli miktarda veri yoksa eğer, veri arttırma yöntemleri uygulanabilir. Tüm bunlar sonuçların farklı olmasında etkindir. Süreç bakımından ele alındığında geleneksel yöntemlere göre daha az manuel müdahale gerektirmektedir. Bu sebepten ötürü, derin öğrenme yöntemleri kullanılarak uydu görüntülerinden, nesne çıkarımı gerçekleştirmek daha kolay olmaktadır. Çalışma doğrultusunda sonuçların iyi ve doğru olması da umut vericidir. Binaların otomatik olarak belirlenmesi, özellikle kentsel bölgelerde yapılan çalışmalarda altlık olarak kullanılmasını sağlar.

Bilgilendirme/Teşekkür

Bu araştırma Duygu ARIKAN'ın "Yüksek Çözünürlüklü Uydu Görüntülerinden Arazi Sınıflarının Derin Öğrenme Yöntemiyle Belirlenmesi" başlıklı tez çalışmasından

gerçekleştirilmiştir. Ayrıca, yazarlar Göktürk-1 uydu görüntülerinin teminini sağlayan Harita Genel Müdürlüğü'ne ve makalenin değerlendirme sürecine zaman ayırarak katkı sunan hakemlere teşekkür eder.

Araştırmacıların katkı oranı beyan özeti

D. Arıkan: Veri temini, Veri analizi, Araştırma, Modelleme, Makale yazımı;

F. Yıldız: Makalenin dil bilgisi kontrolü.

Çıkar çatışması beyanı

Makale ile ilgili olarak, herhangi bir kurum, kuruluş, kişi ile mali çıkar çatışması yoktur ve yazarlar arasında çıkar çatışması bulunmamaktadır.

Araştırma ve yayın etiği beyanı

Yapılan çalışmada yazarlar, araştırma ve yayın etiğine aykırı bir durum olmadığını ve çalışmanın etik kurul izni gerektirmediğini beyan etmektedir.

KAYNAKÇA

- Abdollahi, A., & Pradhan, B. (2021). Integrating semantic edges and segmentation information for building extraction from aerial images using UNet. *Machine Learning with Applications*, 6, 100194.
- Atlan, F., Hançer, E., & Pençe, İ. (2020). U-Net ile Çekirdek Segmentasyonunda Hiper Parametre Optimizasyonu Etkisinin Değerlendirilmesi. *Avrupa Bilim ve Teknoloji Dergisi*, 60-69.
- Ataş, İ. (2023). Performance Evaluation of Jaccard-Dice Coefficient on Building Segmentation from High Resolution Satellite Images. *Balkan Journal of Electrical and Computer Engineering*, 11(1), 100-106.
- Bayraktar, U. (2018). Derin Öğrenme Tabanlı Kanserli Hücre Tespiti. no. December, 2019.
- Hou, Y., Liu, Z., Zhang, T., & Li, Y. (2021). C-UNet: Complement UNet for remote sensing road extraction. *Sensors*, 21(6), 2153.
- Ji, S., Wei, S., & Lu, M. (2018). Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set. *IEEE Transactions on Geoscience and Remote Sensing*, 57(1), 574-586.
- Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS Journal of photogrammetry and remote sensing*, 173, 24-49.
- Kuffer, M., Pfeffer, K., & Sliuzas, R. (2016). Slums from space—15 years of slum mapping using remote sensing. *Remote Sensing*, 8(6), 455.
- Li, L., Wang, C., Zhang, H., & Zhang, B. (2019). Residual UNet for urban building change detection with Sentinel-1 SAR data. Paper presented at the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium.
- Li, X., Zhang, G., Cui, H., Hou, S., Chen, Y., Li, Z., . . . Wang, H. (2023). Progressive fusion learning: A multimodal joint segmentation framework for building extraction from optical and SAR images. *ISPRS Journal of photogrammetry and remote sensing*, 195, 178-191.
- Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. Paper presented at the Proc. icml.
- Madhu, G., Kautish, S., Alnowibet, K. A., Zawbaa, H. M., & Mohamed, A. W. (2023). NIPUNA: A Novel Optimizer Activation Function for Deep Neural Networks. *Axioms*, 12(3), 246.
- Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017). Can semantic labeling methods generalize to any city? the inria aerial image labeling benchmark. Paper presented at the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS).
- Mahabir, R., Croitoru, A., Crooks, A. T., Agouris, P., & Stefanidis, A. (2018). A critical review of high and very high-resolution remote sensing approaches for detecting and mapping slums: Trends, challenges and emerging opportunities. *Urban Science*, 2(1), 8.
- Mitra, P., Shankar, B. U., & Pal, S. K. (2004). Segmentation of multispectral remote sensing images using active support vector machines. *Pattern recognition letters*, 25(9), 1067-1074.
- Mnih, V. (2013). Machine learning for aerial image labeling: University of Toronto (Canada).
- Nwankpa, C., Ijomah, W., Gachagan, A., & Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. arXiv preprint arXiv:1811.03378.
- Oliphant, T. E. (2006). A guide to NumPy (Vol. 1): Trelgol Publishing USA.
- Öz, M. (2021). Eye segmentation using deep neural networks.

- Pal, M. (2005). Random forest classifier for remote sensing classification. *International journal of remote sensing*, 26(1), 217-222.
- Pan, Z., Xu, J., Guo, Y., Hu, Y., & Wang, G. (2020). Deep learning segmentation and classification for urban village using a worldview satellite image based on U-Net. *Remote Sensing*, 12(10), 1574.
- Ramachandran, P., Zoph, B., & Le, Q. V. (2017). Searching for activation functions. *arXiv preprint arXiv:1710.05941*.
- Ronneberger, O., Fischer, P., & Brox, T. (2015a). U-net: Convolutional networks for biomedical image segmentation. Paper presented at the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18.
- Ronneberger, O., Fischer, P., & Brox, T. (2015b). U-net: Convolutional networks for biomedical image segmentation. Paper presented at the International Conference on Medical image computing and computer-assisted intervention.
- Sariturk, B., & Seker, D. Z. (2022). Comparison of Residual and Dense Neural Network Approaches for Building Extraction from High-Resolution Aerial Images. *Advances in Space Research*.
- Tuermer, S., Kurz, F., Reinartz, P., & Stilla, U. (2013). Airborne vehicle detection in dense urban areas using HoG features and disparity maps. *IEEE Journal of selected topics in applied earth observations and remote sensing*, 6(6), 2327-2337.
- Wu, G., Shao, X., Guo, Z., Chen, Q., Yuan, W., Shi, X., . . . Shibasaki, R. (2018). Automatic building segmentation of aerial imagery using multi-constraint fully convolutional networks. *Remote Sensing*, 10(3), 407.
- Wurm, M., Stark, T., Zhu, X. X., Weigand, M., & Taubenböck, H. (2019). Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. *ISPRS Journal of photogrammetry and remote sensing*, 150, 59-69.
- Yan, X., Tang, H., Sun, S., Ma, H., Kong, D., & Xie, X. (2022). After-unet: Axial fusion transformer unet for medical image segmentation. Paper presented at the Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision.
- Zeiler, M. D., Ranzato, M., Monga, R., Mao, M., Yang, K., Le, Q. V., . . . Dean, J. (2013). On rectified linear units for speech processing. Paper presented at the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing.
- Zhang, L., Zhang, L., Tao, D., & Huang, X. (2011). On combining multiple features for hyperspectral remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3), 879-893.
- Zhang, J., Du, J., Liu, H., Hou, X., Zhao, Y., & Ding, M. (2019). LU-NET: An improved U-Net for ventricular segmentation. *IEEE Access*, 7, 92539-92546.
- Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE geoscience and remote sensing magazine*, 5(4), 8-36.



© Author(s) 2023.

This work is distributed under <https://creativecommons.org/licenses/by-sa/4.0/>