

Uzaktan Algılama Görüntülerinden Bina Çıkarımında FPN-ResNeXt50 ve VGG16-UNet Modellerinin Karşılaştırılması

Comparison of FPN-ResNeXt50 and VGG16-UNet Models for Building Extraction from Remote Sensing Images

Serhat Şenses^{1*}, Emine Tanır Kayıkcı²

¹Amasya Üniversitesi, Amasya Teknik Bilimler Meslek Yüksekokulu, 05000, Amasya/Türkiye.

²Karadeniz Teknik Üniversitesi, Mühendislik Fakültesi, Harita Mühendisliği Bölümü, 61080, Trabzon/Türkiye.

ARAŞTIRMA MAKALESİ

*Sorumlu yazar:

Serhat Şenses
serhat.senses@amasya.edu.tr

doi: 10.48123/rsgis.1622130

Yayın süreci

Geliş tarihi: 17.01.2025
Kabul tarihi: 11.09.2025
Basım tarihi: 27.09.2025

Özet

Uzaktan algılama görüntülerinden binaların çıkarılması, arazi kullanımı, iklim ve çevre araştırmaları, afet izleme ve önlemesi imar uygulamaları, gibi çok çeşitli alanlar için gerekli olan bir bilgilerin elde edilmesinde önemli bir veri olarak karşımıza çıkmaktadır. Bu çalışmada, kentsel yerleşim alanları ve kırsal yerleşim alanlarını içeren farklı şehir planlama modellerine sahip Şikago, Austin, Tirol, Viyana ve Kitsap şehirlerine ait 5 farklı kent dokusu için görüntü bölütleme ve sınıflandırma uygulaması yapılmıştır. Görüntülerden binaların çıkarılması için farklı yapay sinir ağları ile bölütleme algoritmaları kullanılarak elde edilen sonuçlar karşılaştırılmıştır. Inria veri seti üzerinde eğitilmiş iki farklı modelin binaların çıkarımında elde ettikleri başarı oranları incelenmiştir. Omurga olarak ResNeXt50 ve FPN mimarisi birlikte kullanılarak bina çıkarımı yapılmış ve sonuçları U net omurgası ile VGG16 mimarisi kullanılarak elde edilen çıkarım sonuçları ile karşılaştırılmıştır. ResNeXt50 ve FPN modeli test verileri üzerinde %96,74 doğrulukla en iyi çıkarım sonucunu vermiştir.

Anahtar kelimeler: Derin öğrenme, Yapay sinir ağları, Bölütleme, Sınıflandırma

Abstract

Extraction of buildings from remote sensing images is an important data in obtaining information required for a wide range of fields such as land use, climate and environmental research, disaster monitoring and prevention, and zoning applications. In this study, image segmentation and classification applications were performed for 5 different urban textures belonging to the cities of Chicago, Austin, Tirol, Vienna and Kitsap, which have different urban planning models including urban and rural settlements. The results obtained by using different artificial neural networks and segmentation algorithms to extract buildings from images were compared. The success rates of two different models trained on the Inria dataset in extracting buildings were examined. Building extraction was performed by using ResNeXt50 and FPN architecture together as the backbone and the results were compared with the extraction results obtained by using the U net backbone and VGG16 architecture. The ResNeXt50 and FPN models gave the best extraction result with 96.74% accuracy on the test data.

Keywords: Deep learning, Artificial neural networks, Segmentation, Classification

1. Giriş

Uzaktan algılama görüntülerinden özellik ve/veya nesne çıkarımı; çevre ve iklim modelleme, topografik haritalama, afet izleme ve değerlendirme, kentsel tasarım, coğrafi bilgi sistemleri, tarım uygulamaları gibi çok çeşitli alanlarda uygulamaları olan bir işlemdir. Bu işlem; görüntünün içeriği, arka planı, dokusu, verilerin kalitesi gibi faktörlere bağlıdır. Binalar da yerleşim planı, arazi kullanımı, kentsel doku özelliklerini etkileyen temel görüntü elemanlarındandır. Literatürde görüntülerden binaların belirlenmesi için, geçmişten günümüze kadar manuel, yarı otomatik ve otomatik

olmak üzere birçok yöntem önerilmiştir. Görüntülerden manuel bina çıkarımı, zaman alan ve yorumlama hatalarına açık olan bir yöntemdir. Otomatik bina çıkarımı yöntemleri son yıllarda çok ilgi çekmiş ve yoğun olarak kullanılmaya başlanmıştır. Ancak uzaktan algılama görüntülerindeki nesne ve arka plan kontrastı, bitişik bina sınırlarındaki ayrım zorlukları ve görüntüdeki gürültü karmaşıklığı gibi sorunlar nedeniyle hala çözülmesi zor olan bir problemdir.

Uzaktan algılama görüntülerinden binaların çıkarımı için birçok yöntem önerilmiştir. Fradkin vd. (2001), farklı bakış açılarından çekilmiş hava fotoğrafı kullanarak bina cephelerinin yeniden oluşturulmasını sağlayan bir yöntem önermişlerdir. Shackelford ve Davis (2003), yüksek çözünürlüklü multispektral görüntülerden kentsel arazilerin sınıflandırılması için, bulanık piksel tabanlı yöntemlerle nesne tabanlı yöntemlerin kullanıldığı hibrit bir yöntem önermiştir. Jin ve Davis (2005), yüksek çözünürlüklü uydu görüntülerinden otomatik bina çıkarımı için spektral, yapısal ve bağlamsal bilgiyi kullanan bütünleştirilmiş bir çözüm önermişlerdir.

Lari ve Ebadi (2007), uzaktan algılama görüntülerinden otomatik bina çıkarımı için sinir ağı tabanlı bir yöntem önermişlerdir. Çalışmada önce görüntü bölgelere ayrılmış, daha sonra her bölge için öncül özellikler çıkarılmış ve binaların belirlenmesi için üç katmanlı bir perception (algılama) ağı eğitilmiştir. Türker ve San(2010), tarafından yüksek çözünürlüklü uydu görüntülerinden dikdörtgen ve dairesel şekilli binaları çıkarılması için Hough dönüşümü kullanan bir çalışma yapılmıştır. Benali vd. (2014), yüksek çözünürlüklü uydu görüntülerinden kentsel alanların çıkarımı için, yarı otomatik bir yöntem sunmuşlardır. Zheng ve Wang (2014), çalışmalarında uzaktan algılama görüntülerinden anlamsal bölütleme yapmak amacıyla, nesne tabanlı bir Markov rastgele alan modeli (OMRF) kullanmışlardır.

Yüksek çözünürlüklü uydu görüntülerinin elde edilmesi ve nesne tabanlı yaklaşımların piksel tabanlı yaklaşımlara göre üstünlüğü nedeniyle, nesne tabanlı yöntemler özellik çıkarımı için sıklıkla kullanılmaya başlamışlardır (Whiteside vd., 2011; Attarzadeh & Momeni, 2012), yüksek çözünürlüklü görüntülerden binaları tanıma ve belirleme işlemini otomatik olarak yapan nesne tabanlı bir yöntem önermişlerdir. Yöntemlerini Quickbird uydu görüntüleri üzerinde test etmişlerdir. Khatraker ve Kumar (2018), yüksek çözünürlüklü WorldView-2 görüntülerinden çoklu çözünürlüklü bölütleme ve kural tabanlı özellik çıkarımını kullanan yarı otomatik bir yöntem kullanmışlardır. Prathiba vd. (2020), çalışmalarında binaların belirlenmesi için nesne tabanlı bir yöntem kullanmışlar ve görüntüyü, piksellerin mekânsal ve spektral benzerliğine göre birçok bileşene ayırmışlardır. Daha sonra nesneleri en yakın komşuluk sınıflandırıcısı kullanılarak sınıflandırılmıştır.

Bina çıkarımı için, piksel tabanlı yöntemler, nesne tabanlı yöntemler, makine öğrenmesi yöntemleri ve derin öğrenme yöntemleri önerilmiştir. Piksel tabanlı yöntemler, spektral değerleri kullanarak görüntüdeki nesneleri ayırt eder. Nesne tabanlı yöntemler, tüm görüntünün farklı boyutlardaki nesnelere ve yamalara (patch) bölünmesini sağlar. Bu nesneler daha sonra spektral, dokusal ve bağlamsal özellikleri göz önüne alınarak sınıflandırılırlar. Makine öğrenmesi yöntemleri, özellik çıkarıcılar ve sınıflandırma algoritmalarını kullanarak görüntüden bilgilerin çıkarılmasını sağlar. Derin öğrenme yöntemlerinden olan evrimsel sinir ağıları (CNN) ise, birden fazla özelliğin çıkarımında ve sınıflandırılmasında evrim katmanlarının kombinasyonunu kullanır.

Derin öğrenme ağıları, çok katmanlı bir ağ yapısı oluşturmak için bazı eğitim işlemleri ve örneklem verileri kullanan bir makine öğrenim sürecidir (Bengio, 2009). Bir tür derin öğrenme tekniği olan evrimsel sinir ağıları (CNN), sınıflandırma için gerekli özellikleri otomatik olarak öğrenebilir ve hassas özellik eşleştirmesi yapabilirler. CNN'lerin uzaktan algılamanın görüntü işleme uygulamalarında kullanımıyla; sınıflandırma, özellik çıkarma ve nesne belirleme alanlarında başarılı sonuçlar elde edilmiştir (Guo vd., 2017; Yao vd., 2017; Tejeswari vd., 2022).

CNN'lerin yanı sıra, üretken çekişmeli ağlar (GANs), görüntü transformatörleri (ViT), Grafik evrim ağıları (GCNs), YOLO ağıları vb. de bu amaçla kullanılmışlardır. Zhong vd. (2016), yüksek çözünürlüklü uzaktan algılama verilerinden binaları çıkarmak için, tamamen evrimsel ağıları (FCN) kullanarak, model performansının; girdi verilerinin boyutlarından, filtre adım sayısından, öğrenme oranından, eğitim süresinden ve ince ayardan nasıl etkilendiğini araştırmışlardır. Li vd. (2019), çok yüksek çözünürlüklü görüntülerden binaların çıkarımı için derin öğrenme temelli semantik bölütleme yöntemi önermişlerdir. Çalışmalarında Paris, Şangay, Las Vegas ve Hartum şehirlerindeki binaların çıkarımı için U-Net modeli kullanılmışlardır. Ps ve Aithal (2023) tarafından, uzaktan algılama görüntülerinden binaları çıkarmak için, evrimsel sinir ağılarını kullanan bir derin öğrenme modeli önerilmiştir. Modeli, metropol alanında bulunan farklı bölgelerden görüntüler üzerinde eğiterek, az yapılaşmanın olduğu alanlarla, yoğun yapılaşma alanlarındaki farklı binaların belirlenmesini amaçlamışlardır. Nurkarim vd. (2023), derin öğrenme yöntemleri kullanarak binaların çıkarımı ve çıkarılan binaların sayımı ile ilgili bir yöntem önermişlerdir. Yöntemleri ile bölütlemeyi, evrimsel sinir ağıları ile birleştirerek WorldView-3 uydu görüntülerinden nesne algılama modeli geliştirmişlerdir. Srivastava vd. (2024), hem geleneksel yöntemleri hem de derin öğrenme yöntemlerini kullanarak bina çıkarımı için kullanılan farklı modelleri tartışmışlardır. Bu çıkarım yöntemlerinin karşılaştığı bazı zorlukları da incelemişlerdir. Li vd. (2024), erken dönem bina çıkarımı uygulamaları ile günümüzde kullanılan optik görüntülerden ve SAR görüntülerinden elde edilen bina çıkarımlarını inceleyerek ayrıntılı bir analiz sunmuşlardır. Çalışmalarında binaların çıkarımı için bina özelliklerini altı ana gruba ayırmışlardır. Binaların arazideki izlerinin belirlenmesi, bina cephelerinin bölütlere (segment) ayrılması, çatı segmentlerinin tanımlanması, bina yüksekliklerinin tahmini, bina türlerinin sınıflandırılması ve binaların zamana göre değişiminin belirlenmesi, üzerinde durdukları özelliklerdir.

FPN-ResNeXt50 modeli, çok ölçekli öznetelik haritalarının birleşimini destekleyen mimarisiyle nesne sınırlarının hassas çıkarımı açısından avantaj sağlamaktadır. VGG16-UNet ise, tıbbi ve coğrafi bölütleme problemlerinde yaygın kullanılan temel bir yapı olup literatürde sıklıkla referans alınan klasik modellerdendir (İsmail vd., 2024; Kaur & Gandhi, 2019; Bansal & Singh, 2023; Sewada & Goyal, 2025). Bu çalışmada, biri daha güncel ve derin (FPN-ResNeXt50), diğeri daha sade ve geleneksel (VGG16-UNet) iki mimari seçilerek karşılaştırma yapılmıştır. FPN-ResNeXt50 ve VGG16-UNet modelleri, Inria veri seti (https://github.com/qubvel/segmentation_models) kullanılarak eğitilmiş ve sonuçlar karşılaştırmalı olarak değerlendirilmiştir.

2. Materyal ve Metot

2.1 Veri Seti

Çalışmada, Maggiori vd. (2017) tarafından etiketlenmiş Inria Hava Görüntü veri seti kullanılmıştır. Veri seti 405 km² eğitim, 405 km² test verisi olmak üzere toplam 810 km² kapsama altına sahiptir. Eğitim veri seti (1) bina sınıfı ve (2) bina olmayan sınıf olmak üzere belirlenmiş iki semantik sınıf için gerçek yer görüntü verileriyle, piksel başına 30 cm'lik bir uzamsal çözünürlüğe sahip orto-foto renkli görüntülerden oluşur. Görüntüler, (1) yoğun yerleşim planına sahip bölgelerle, (2) daha az yoğun yerleşime sahip kırsal bölgelerden oluşmuştur. Veri seti, Austin, Şikago, Kitsap, Batı Tirol ve Viyana bölgelerinin her biri için 5000x5000 piksel boyutlu her biri 25000000 piksel içeren 36 renkli görüntü içermektedir. Bu şekilde görüntüler 405 km² lik bir alanı kapsamakta ve toplam 180 görüntüden oluşmaktadır. Gerçek yeryüzü görüntüleri, 225 (bina sınıfı) ile 0 (bina olmayan sınıf) değerlerine karşılık gelen 180 adet, tek bant görüntüsüdür. Eğitim veri seti yazar tarafından geliştirilen Python programı ile modellerin eğitilmesinde kullanılmıştır. Çalışma dönemi itibarıyla GPU erişimi mümkün olmadığından eğitim yalnızca çok çekirdekli CPU altyapısı (paralel 4 CPU 96 çekirdekli, 128 GB RAM) üzerinde gerçekleştirilmiştir. Bu durum eğitim süresini artırmış olsa da modelin doğruluk açısından başarılı sonuçlar verdiği görülmüştür. Modellerin eğitimi sunucu bilgisayarda 27 gün sürmüştür.

Tablo 1. Çalışmada kullanılan eğitim ve test veri setleri

Eğitim Seti	Yama Sayısı	Alan	Test Seti	Yama Sayısı	Alan
Austin	36	81 km ²	Bellingham	36	81 km ²
Şikago	36	81 km ²	San Francisco	36	81 km ²
Kitsap County	36	81 km ²	Bloomington	36	81 km ²
Viyana	36	81 km ²	Innsbruck	36	81 km ²
Batı Tirol	36	81 km ²	Doğu Tirol	36	81 km ²
Toplam	180	405 km²	Toplam	180	405 km²

Test seti olarak da, aynı sayıda olmak üzere, Bellingham, Bloomington, Innsbruck, San Francisco ve Doğu Tirol bölgelerini içeren görüntü bölgeleri kullanılmıştır. Model performansı, birleşim üzerinde kesişim (Intersection over Union, IoU) ve Doğruluk (Accuracy) ölçütleriyle değerlendirilmiştir. Eğitim bölgeleri ve bölgelerin özelliklerine ait bilgiler Ek1'de ayrıntılı olarak verilmektedir.

IoU (Intersection over Union), modelin tahmin ettiği bölgenin gerçek bölgeyle ne kadar örtüştüğünü ölçen bir metriktir ve bölütleme (segmentation) modellerinin başarımını değerlendirmek için yaygın olarak kullanılır. IoU kavramı genellikle nesne tespiti (object detection) alanında çift tespit (duplicate detection) riskini azaltmak için kullanılır. Ancak bölütleme (segmentation) bağlamında, IoU farklı bir şekilde yorumlanır ve modelin tahmin ettiği maskenin gerçek maske ile örtüşme oranını verir. IoU, şu şekilde hesaplanmaktadır:

$$IoU = \frac{GT \cap PM}{GT \cup PM} = \frac{TP}{TP + FP + FN} \quad (1)$$

Burada, GT, gerçek arazi değeri; PM, tahmin edilen değer; TP, gerçek pozitif (Bina olarak sınıflanan gerçekte de bina olan piksel sayısı); FP, yanlış pozitif (Bina olarak sınıflanan gerçekte bina sınıfına ait olmayan piksel sayısı); FN, yanlış negatif (Bina olarak sınıflanmayan ancak gerçekte bina sınıfına ait olan piksel sayısı) olarak alınmıştır. Doğruluk ise doğru sınıflandırılmış piksellerin yüzdesini temsil eder ve aşağıdaki eşitlikle hesaplanır.

$$Doğruluk = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

Daha sonra, eğitim sırasında en iyi performans gösteren modeli seçebilmek için bir 'birleşik' ölçüt kullanılmıştır. Bu ölçüt IoU ve Doğruluk değerlerinin ortalaması alınarak (3) eşitliği ile hesaplanmaktadır.

$$\text{Birleşik Ölçüt} = \frac{(\text{IoU} + \text{Doğruluk})}{2} \quad (3)$$

Birleşik ölçütte doğruluk değeri ile IoU değeri eşit ağırlıktadır. Ancak IoU değeri sonuç metriklerinde daha belirgin bir etkiye sahiptir. Bunun nedeni, (2) nolu eşitlikte yer alan TN (True Negatif) teriminin ölçütlerin hesaplanmasında, sonuçlara etki ağırlığının, diğer terimlere göre daha fazla olmasıdır. Bu durum, IoU değerinin görüntü bölütleme işlemlerinde tercih edilmesi nedeniyle, birleşik ölçütün değerlendirme kriteri olarak kullanılabilmesini sağlamaktadır (Milosavljević, 2020).

İnria veri setinin işlenmesi için uygulanan yöntem, görüntü hazırlama, anlamsal bölütleme için derin Evrişimli Sinir Ağları (CNN) topluluğunun uygulanması ve eğitimi ve son olarak da test veri seti görüntü bölgeleri için bina alanlarını tahmin etmektir.

2.2. Görüntü Hazırlama

Derin anlamsal (semantik) bölütleme yöntemi kullanılarak uzaktan algılama görüntülerinin işlenmesi için ilk adımda, orijinal veriler ve orto-foto gerçek yer görüntülerinden (Şekil 1-Şekil 5), belirlenen boyutlarda bir dizi görüntü yaması (patch) çıkarılmıştır. Burada "yama" terimi, belirli bir boyutta kesilmiş, analiz edilmek üzere oluşturulmuş küçük görüntü parçalarını ifade etmektedir.

Görüntülerin farklı bölümlerinin (örneğin, Tirol 1-6) yamalar içinde çeşitli konumlarda yer almasını sağlamak amacıyla, yamalar belirli bir bindirme oranında üst üste gelecek şekilde ayarlanmıştır.

Bu kapsamda, 5000x5000 piksel boyutundaki veri seti görüntülerinden 512x512 piksel boyutunda yamalar (patches) oluşturulmuş ve minimum %30 örtüşme oranı sağlanmıştır. Bu yöntemle, 19x19 piksel boyutunda toplam 361 adet görüntü yaması ve aynı sayıda orto-foto görüntü yaması elde edilmiştir. Yığın (Batch)/ bölme uyumunu sağlamak ve düzgün kümelenme oluşturmak amacıyla, yamalardan 1 tanesi çıkarılmış ve değerlendirmeye alınmamıştır. Bu çıkarma işleminde, 19x19 boyutlu ızgaranın görüntü sınırında ürettiği dolgu yamalardan (padded patch) en fazla dolguya sahip olan yama (4488, 0) seçilerek değerlendirme dışı bırakılmıştır. Böylece uzamsal dağılım da korunmuştur. Bu işlem sonucunda 360 adet yama ile işleme devam edilmiştir.

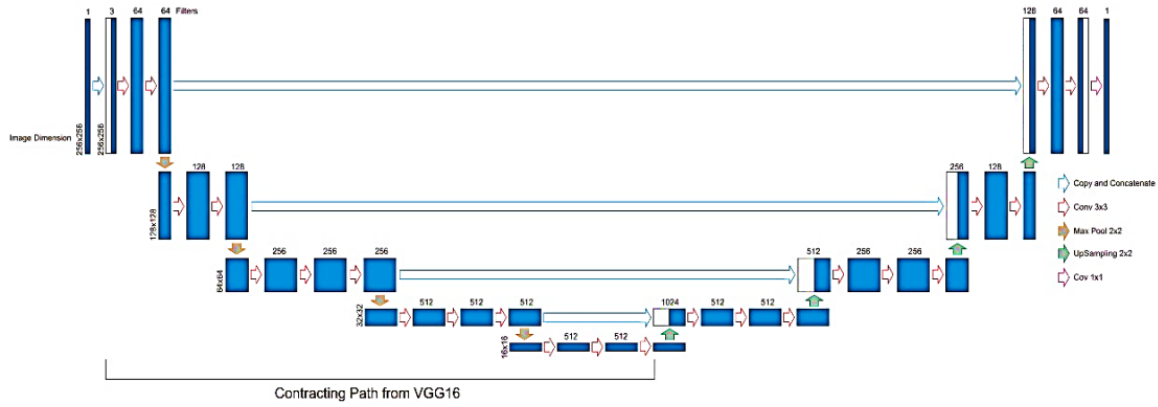
Örtüşme işleminin yanı sıra, eğitim veri kümesinin çeşitliliğini artırmak amacıyla veri büyütme (data augmentation) tekniği uygulanmıştır. Bu amaçla, her bir yama görüntüsü yatay ve dikey olarak çevrilmiş, ayrıca sırasıyla 90, 180 ve 270 derece döndürülerek her orijinal görüntünün beş farklı varyasyonu oluşturulmuştur. Ancak, model ölçek ve dönme değişmezliğine (scale-rotation invariance) sahipse, bu tür dönüşümlerin modelin performansına etkisi sınırlı olabileceği bilindiğinden, model analiz edilmiştir. Analiz sonucunda modelin ölçek ve dönme değişimine duyarlı olduğu görülmüştür.

2.3 Modelin Uygulanması ve Eğitimi

Çalışmada Python programlama dili ve Keras Kütüphanesi kullanılmıştır. Derin anlamsal bölütleme modellerinin uygulanması için Pavel Yakubovskiy'nin Github kitaplığı (https://github.com/qubvel/segmentation_models), bölütleme mimarisi için Keras tabanlı bir Python kitaplığı ve ImageNet omurgası kullanılmıştır. Çalışmada kullanılan başlıca hiperparametreler ve bunlara ait değerler Tablo 2'de verilmiştir.

Tablo 2. Çalışmada kullanılan başlıca hiperparametreler ve bunlara ait değerler

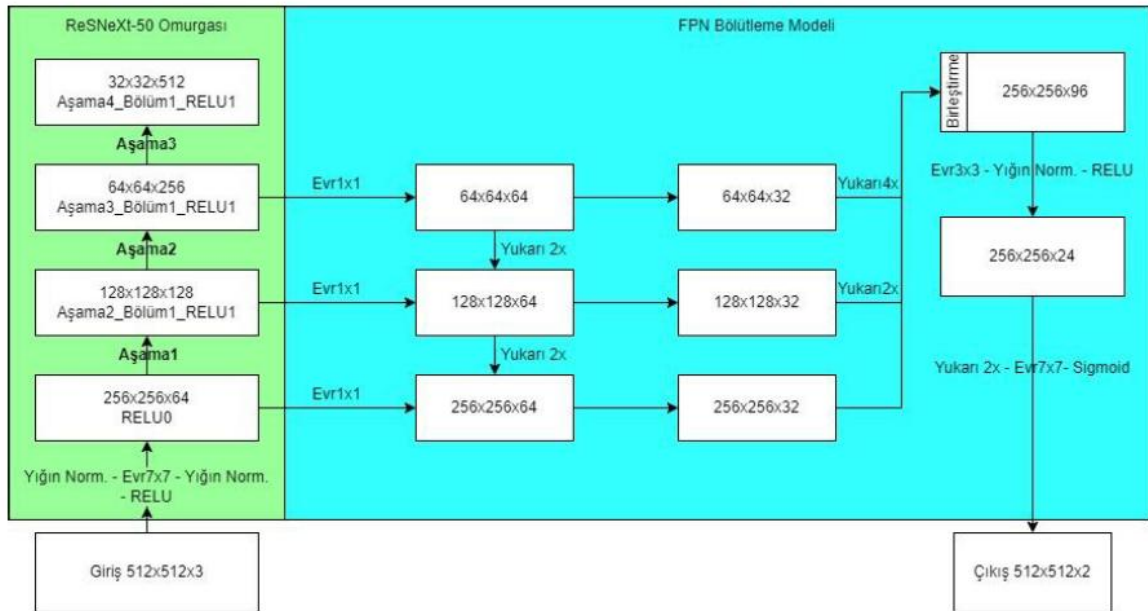
Hiperparametre	Değer
Yığın Boyutu	16
Başlangıç Öğrenme Oranı	0.0001
Öğrenme Oranı Ölçeği	0.01
Öğrenme Oranı Sönümlenme Değeri	0.6
Epok (öğrenme tur sayısı)	5
Önyargı(bias)	0.45
Örtüşme(overlap)	%30
Aktivasyon Fonksiyonu	ReLU
Kayıp Fonksiyonu	Crossentropi+Dice loss
Sınıflandırıcı	Softmax



Şekil 1. U-Net-VGG16 mimarisi (Pravitasari vd., 2020)

Bu mimari, VGG16 tabanlı bir U-Net modelinin görüntü bölütleme (segmentation) görevine uyarlanmış yapısını temsil etmektedir. Model, simetrik bir encoder-decoder (daralan ve genişleyen yol) yapısına sahiptir. Giriş görüntüsü 512×512 boyutunda olup, modelin daralan yolu (contracting path) klasik VGG16 mimarisini ile uyumlu olacak şekilde yapılandırılmıştır. Bu bölümde, 3×3 evrişimsel katmanlar ve 2×2 maksimum havuzlama (max pooling) işlemleri aracılığıyla öznitelikler çıkarılırken boyutsal küçültme gerçekleştirilir. Her evrişim bloğunda filtre sayısı artmakta ve 512×512 boyutunda öznitelik haritalarına ulaşılmaktadır. Bu işlem, modelin daha soyut ve anlamlı öznitelikler öğrenmesini sağlar.

Genişleyen yol (expanding path) ise yukarı örnekleme (upsampling) ve evrişimsel işlemlerle karakterizedir. Bu aşamada, encoder katmanlarında elde edilen öznitelik haritaları, decoder tarafında uygun seviyelerde tekrar kullanılarak (skip connections – kopyala ve birleştir) hem konumsal hem de anlamsal bilgi korunur. Bu strateji, modelin bölütleme sınırlarını daha hassas bir şekilde öğrenmesini sağlar. Her bir yukarı örnekleme adımında 2×2 ölçekli büyütme yapılır ve ardından 3×3 evrişim uygulanarak detayların yeniden yapılandırılması sağlanır. Son katmanda 1×1 evrişim uygulanarak, çıktı istenilen sınıf sayısına (örnekte 1 kanal) indirgenir. Bu yapı, özellikle görüntü analizi gibi detaylı bölütleme gerektiren uygulamalarda yüksek doğrulukla sonuç üretmek üzere tasarlanmıştır.



Şekil 2. ResNeXt50-FPN mimarisi (Milosavljević, 2020)

Sunulan mimari şeması, görüntü bölütleme amacıyla tasarlanmış ve ResNeXt-50 omurgası ile yapılandırılmış bir Feature Pyramid Network (FPN) modelinin detaylı yapısını göstermektedir. Modelin girdi boyutu 512×512×3 olup, bu giriş görüntüsü ilk olarak ResNeXt-50 mimarisini temel alınarak oluşturulan bir omurga yapısından geçirilir. Bu omurga, derin evrişimsel katmanlardan oluşmakta ve çok ölçekli özelliklerin çıkarımına olanak sağlamaktadır. Aşamalı olarak sırasıyla 256×256×64, 128×128×128, 64×64×256 ve 32×32×512 boyutlarında öznitelik haritaları üretilmektedir.

Bu çok ölçekli öznetelikler, FPN bölümlene modülüne aktarılmakta ve her seviyede 1×1 evrişim uygulanarak kanal sayıları azaltılmaktadır. Ardından, bu öznetelikler belirli oranlarda yukarı örneklenerek ($2 \times$ ve $4 \times$), boyutsal olarak hizalanmakta ve birleşim aşamasında 256×256 boyutunda birleştirilmiş bir öznetelik haritası elde edilmektedir. Bu birleşik özellik haritası, sırasıyla 3×3 evrişim, yığın normlaştırma ve ReLU aktivasyonu ile işlenerek $256 \times 256 \times 24$ boyutuna indirgenir. Son aşamada ise, yukarı örnekleme ve 7×7 evrişimsel katmanlarla birlikte sigmoid aktivasyon fonksiyonu uygulanarak, $512 \times 512 \times 2$ boyutunda nihai bölütleme çıktısı üretilir. Bu yapı hem düşük hem de yüksek seviyeli öznetelikleri etkin bir biçimde birleştirerek detaylı ve doğruluk oranı yüksek bölütleme çıktıları elde etmeyi amaçlamaktadır.

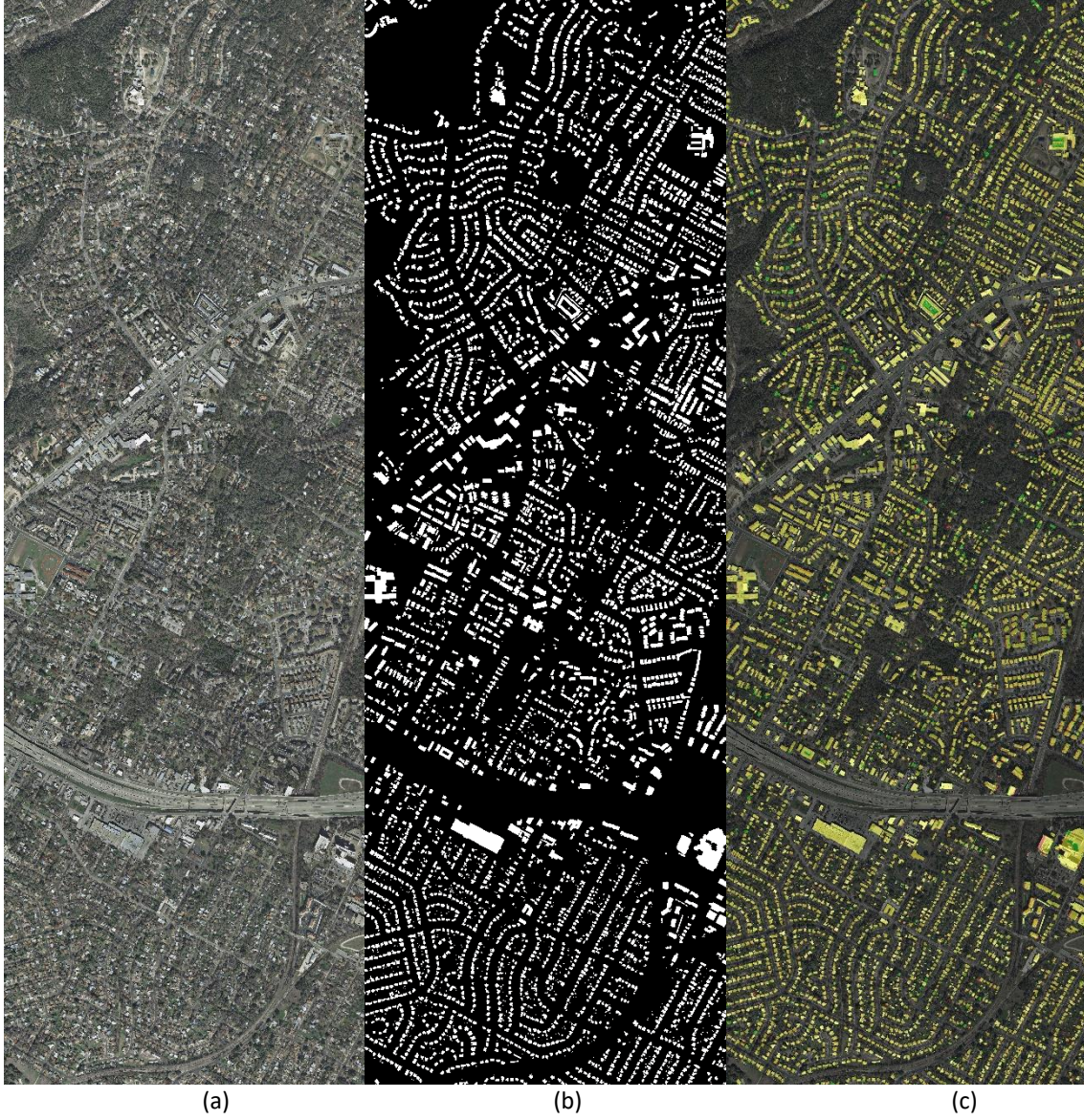
3. Bulgular

Bu çalışmada FPN-ResNeXt50 ve VGG16-UNet modelleri, derin öğrenme tabanlı bölütleme (segmentation) mimarileri olarak kullanılmıştır. ResNeXt50 ve VGG16 backbone ağları görüntüden anlamlı özellikler çıkarmaktadır. Elde edilen özellik haritaları, FPN ya da UNet gibi üst yapılar aracılığıyla işlenerek bölütleme gerçekleştirilmiştir. Söz konusu modellerde, ResNeXt50 ve VGG16 omurga ağları görüntülerden yüksek seviyeli ve anlamlı özellikler çıkarmaktadır. Elde edilen özellik haritaları, ilgili üst yapılar (FPN veya UNet) aracılığıyla çok ölçekli biçimde işlenmekte ve böylece bölütleme işlemi gerçekleştirilmektedir. Bu yaklaşım, omurga ağlarının güçlü temsil yetenekleri ile üst yapıların ayrıntı koruma ve çok ölçekli bilgi işleme kabiliyetlerini bir araya getirerek, segmentasyon performansını artırmayı hedeflemektedir.

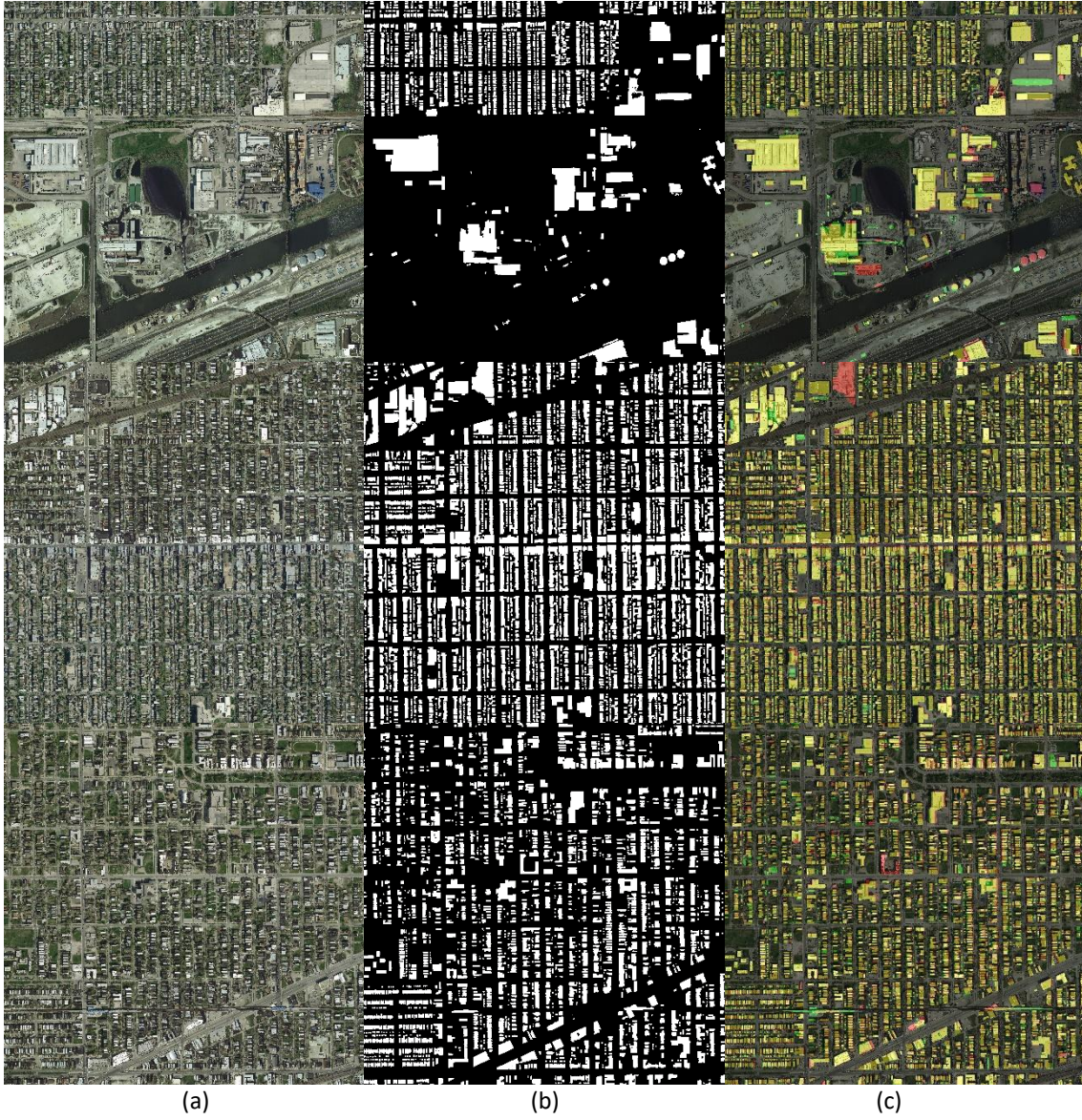
Her bir bölgeye ait eğitim verileri için doğruluk, IoU ve birleşik ölçüt değerleri hesaplanmış ve Tablo 4-8'de sunulmuştur. Bu tablolarda, ölçütlerin hesaplanmasında kullanılan her bölgeye ait TP, TN, FP ve FN değerleri de gösterilmiştir. Ayrıca, tüm bölgeler için ilgili ölçüt değerleri Tablo 5'te verilmiştir.

Bu çalışma, sınıflandırma (classification) veya nesne tespiti (object detection) değil, bölütleme (segmentation) üzerine odaklanmaktadır. Bölütleme işlemi, giriş görüntüsündeki piksellerin belirli sınıflara atanması ile gerçekleştirilmiştir. Mimari, omurga ağları tarafından çıkarılan derin özellikleri kullanarak, her pikselin ait olduğu bölgeyi belirlemektedir.

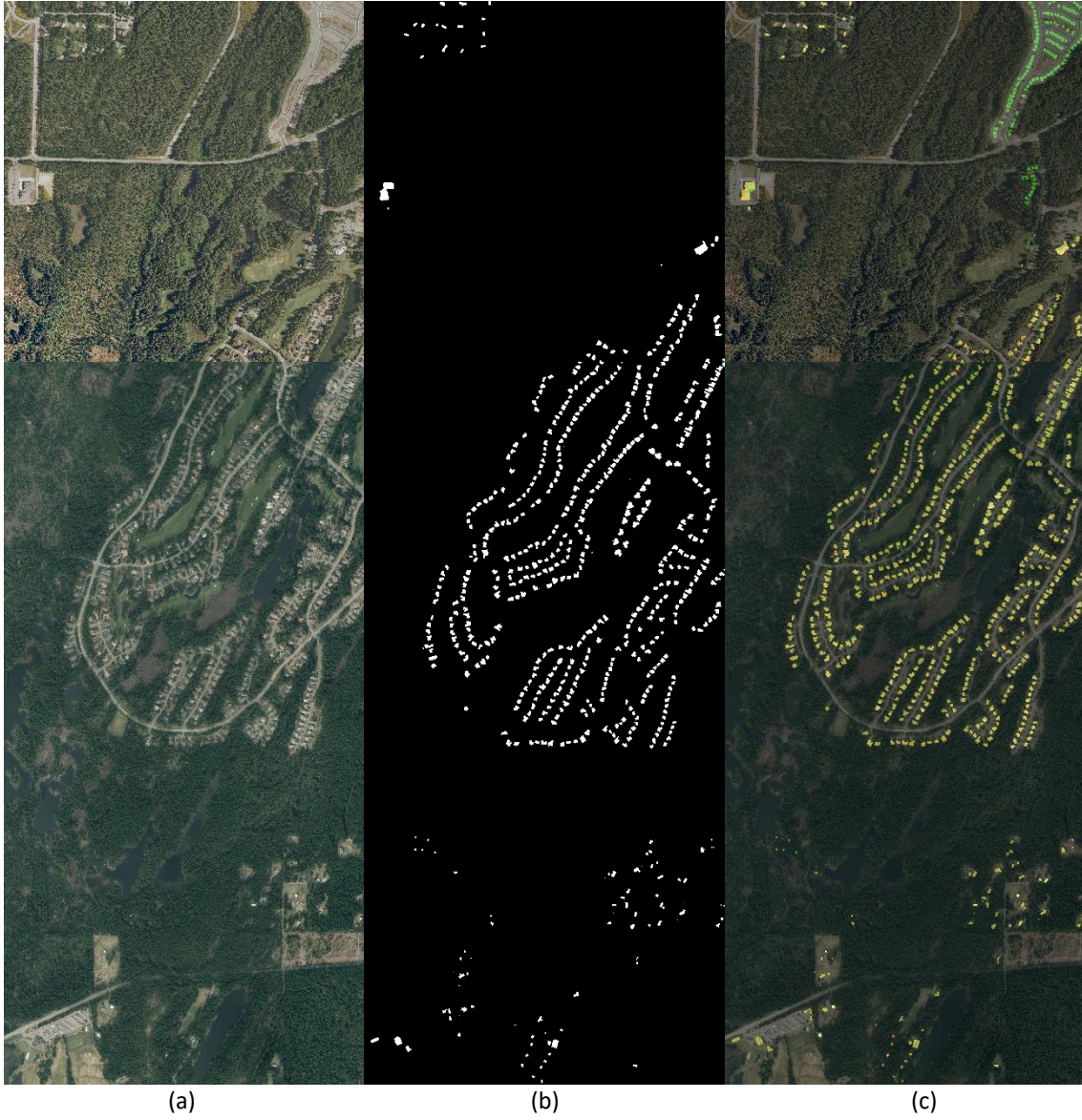
Şekil (3-7)'de girdi görüntüsü, bölütlenmiş görüntü ve bölütlenmiş görüntü ile orto-foto gerçek yeryüzü görüntüsünün karşılaştırılması ile elde edilen görüntüler verilmiştir. Şekillerdeki (c) sütununda verilen karşılaştırılmış görüntüde yeşil olarak gösterilen alanlar gerçekte bina olan ancak model tarafından belirlenemeyen binaları, kırmızı olan alanlar ise, bina olmadığı halde model tarafından bina olarak etiketlenen alanları göstermektedir.



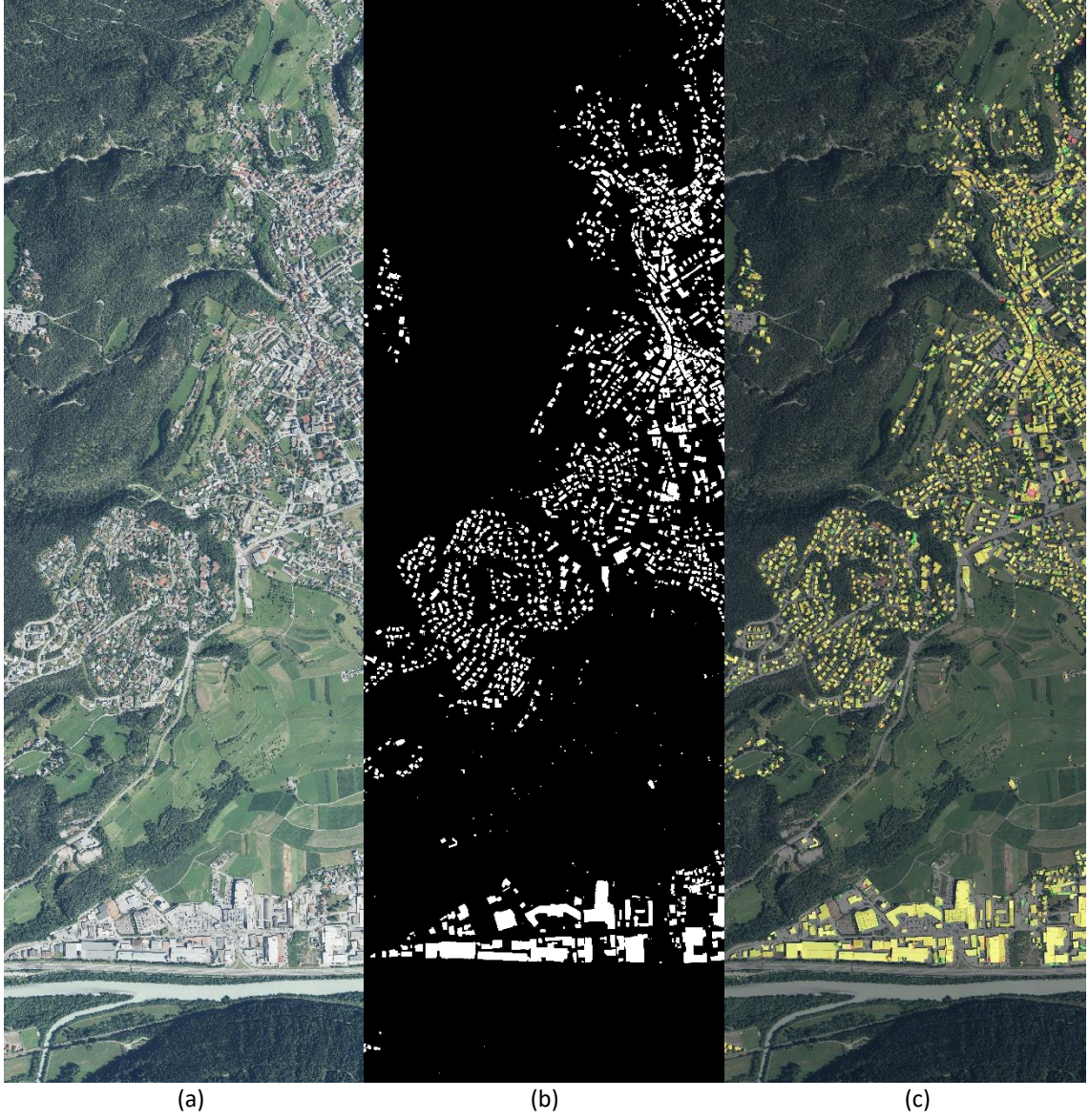
Şekil 3. Austin için binaların belirlenmesi uygulamasında: a) girdi görüntüsü, b) bölütlenmiş görüntü ve c) bölütlenmiş görüntü ile orto-foto gerçek yeryüzü görüntüsünün çakıştırılmasıyla elde edilmiş görüntü



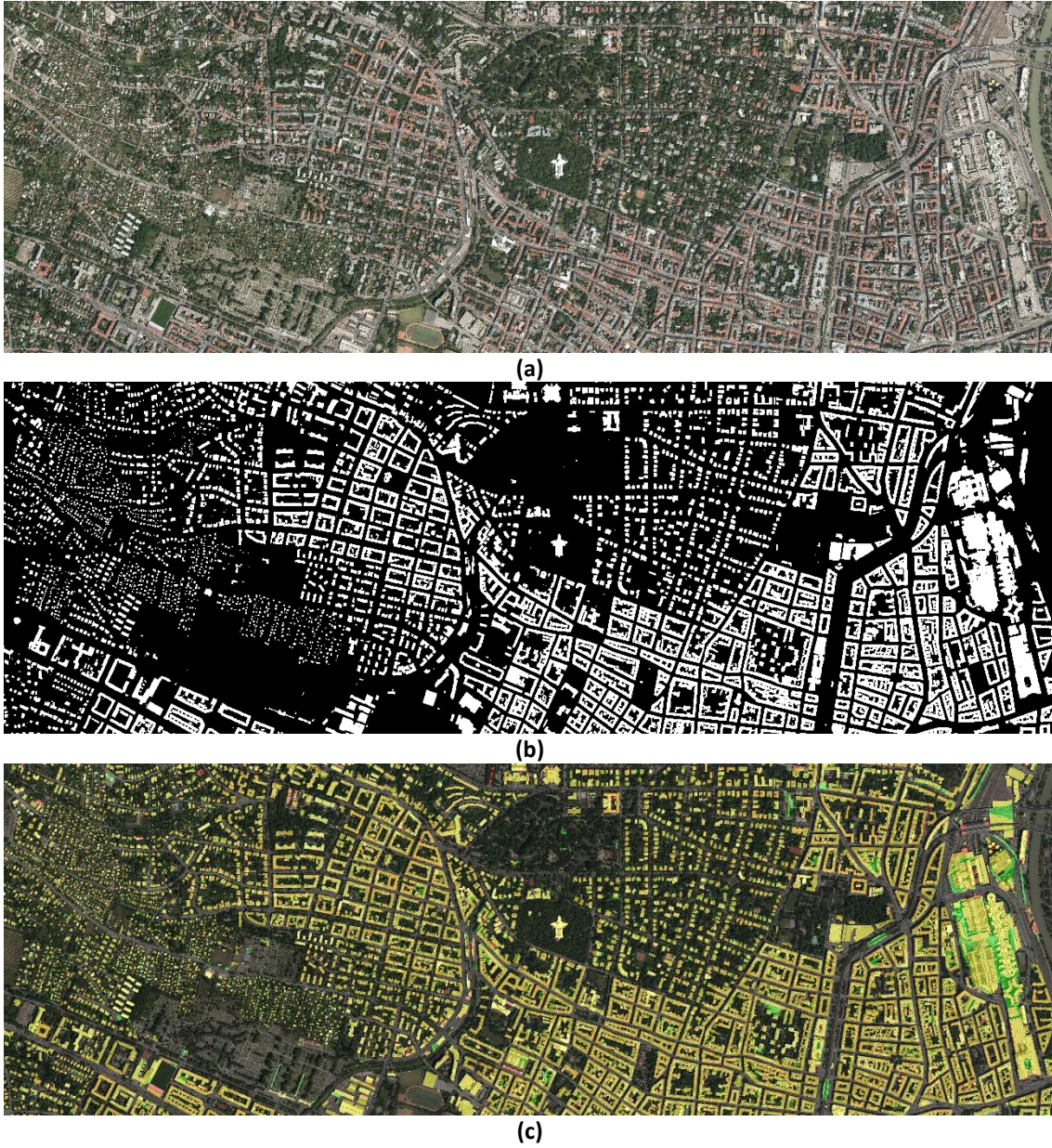
Şekil 4. Şikago için binaların belirlenmesi uygulamasında: a) girdi görüntüsü, b) bölütlenmiş görüntü ve c) bölütlenmiş görüntü ile orto-foto gerçek yeryüzü görüntüsün karşılaştırılmasıyla elde edilmiş görüntü



Şekil 5. Kitsap için binaların belirlenmesi uygulamasında: a) girdi görüntüsü, b) bölütlenmiş görüntü ve c) bölütlenmiş görüntü ile orto-foto gerçek yeryüzü görüntüsün karşılaştırılmasıyla elde edilmiş görüntü



Şekil 6. Tirol için binaların belirlenmesi uygulamasında: a) girdi görüntüsü, b) bölütlenmiş görüntü ve c) bölütlenmiş görüntü ile orto-foto gerçek yeryüzü görüntüsün çakıştırılmasıyla elde edilmiş görüntü



Şekil 7. Viyana için binaların belirlenmesi uygulamasında: a) girdi görüntüsü, b) bölütlenmiş görüntü ve c) bölütlenmiş görüntü ile orto-foto gerçek yeryüzü görüntüsün karşılaştırılmasıyla elde edilmiş görüntü

Austin, Şikago, Kitsap, Batı Tirol ve Viyana şehirlerinin eğitim verilerine ait bilgiler Tablo 3'te verilmiştir. Her şehrin alt bölgelerine ait eğitim veri sonuçlarının detaylı gösterimleri Ekler bölümünde verilmiştir.

Tablo 3. Şehirlerin modellere göre eğitim verileri ölçütleri

Şehirler	Model-Bölütleme	Doğruluk	IoU	Birleşik Ölçüt
Austin	FPN-ResNeXt50	0,9741	0,8040	0,8890
	VGG16-UNet	0,9459	0,5964	0,7695
Şikago	FPN-ResNeXt50	0,9300	0,7297	0,8298
	VGG16-UNet	0,9222	0,7011	0,8116
Kitsap	FPN-ResNeXt50	0,9938	0,6609	0,8274
	VGG16-UNet	0,9928	0,6243	0,8086
Batı Tirol	FPN-ResNeXt50	0,9867	0,8257	0,9062
	VGG16-UNet	0,9831	0,7861	0,8846
Viyana	FPN-ResNeXt50	0,9512	0,7992	0,8752
	VGG16-UNet	0,9421	0,7724	0,8572

Veriler incelendiğinde, tüm ölçütlerde FPN-ResNeXt50 modelinin daha iyi sonuçlar verdiği görülmüştür.

Tablo 4. Tüm bölgelere ait toplam eğitim verileri

Model-Bölütleme	Doğruluk	IoU	Birleşik Ölçüt	Kesişim	Birleşim	Toplam Piksel	Gerçek Pozitif	Gerçek Negatif	Yanlış Pozitif	Yanlış Negatif
FPN-ResNeXt50	0.9672	0.7814	0.8743	88040457	112677229	750000000	88040457	637322771	13105024	11531748
VGG16- UNET	0.9572	0.7204	0.8388	82681763	114766562	750000000	82681763	635233438	15194357	16890442

Sunulan karşılaştırmalı değerlendirme, FPN-ResNeXt50 ve VGG16-UNET modellerinin görüntü bölütleme (segmentation) görevindeki performanslarını çeşitli metrikler üzerinden incelemektedir. Analiz sonuçları, FPN-ResNeXt50 modelinin tüm temel ölçütlerde VGG16-UNET modeline kıyasla üstün bir performans sergilediğini ortaya koymaktadır. Özellikle doğruluk (accuracy) oranı %96,72 olan FPN-ResNeXt50, %95,72 doğruluk elde eden VGG16-UNET modelini geride bırakmıştır. Benzer şekilde, FPN-ResNeXt50 modelinin Intersection over Union (IoU) değeri 0.7814 olup, VGG16-UNET'in 0.7204'lük değerinden anlamlı ölçüde yüksektir. Bu durum, FPN-ResNeXt50 modelinin nesne sınırlama başarımının daha yüksek olduğunu göstermektedir. Ayrıca, FPN-ResNeXt50 modelinin 0.8743'lük değeri, VGG16-UNET'in 0.8388'lik değerine kıyasla daha iyi bir genel performansa işaret etmektedir. Gerçek pozitif (TP) ve gerçek negatif (TN) piksel sayılarında FPN-ResNeXt50 daha yüksek değerler üretmiş, buna karşın yanlış pozitif (FP) ve yanlış negatif (FN) sayılarında daha düşük değerler kaydetmiştir. Bu bulgular, FPN-ResNeXt50 modelinin hem hedef sınıf tespiti hem de arka plan ayrımı açısından daha isabetli ve dengeli bir bölütleme gerçekleştirdiğini göstermektedir. Sonuç olarak, FPN-ResNeXt50 modeli, görüntü bölütleme uygulamalarında daha yüksek doğruluk ve güvenilirlik sunarak tercih edilebilirliğini artırmaktadır.

Her kent ve kentlerin alt bölgeleri için ayrı ayrı elde edilen değerlerin ortalaması alınarak bulunan sonuçlar Tablo 4'te verilmiştir. Eğitim verileri üzerindeki genel başarımların gösterildiği bu tablo irdelendiğinde en iyi sonuçların ResNeXt50-FPN modeli tarafından üretildiği belirlenmiştir. Bu model tüm eğitim verileri kullanılarak uygulanmış ve görüntüdeki binaların çıkarımı için bölütleme ve sınıflandırma yapılmıştır. Modelde kullanılan hiperparametreler tablo 5'te verilmiştir.

Tablo 5. Modelde kullanılan hiperparametreler

Hiperparametre	Değer
Giriş ve çıkış veri boyutu	512x512
Yığın Boyutu	16
Başlangıç Öğrenme Oranı (eğitim aşaması)	0.0001
Başlangıç Öğrenme Oranı (ince ayar aşaması)	0.00001
Öğrenme Oranı Ölçeği	0.01
Öğrenme Oran Sönümlenme Değeri	0.6
Epok (öğrenme tur sayısı)	5
Önyargı (bias)	0.45
Örtüşme (overlap)	%30

Tablo 5'te sunulan hiperparametreler, modelin hem eğitim hem de ince ayar (fine-tuning) aşamalarındaki davranışını doğrudan etkileyen temel ayarlardır. Giriş ve çıkış boyutu olarak 512x512 piksel seçilmiş; bu, yüksek çözünürlüklü detayların korunmasını ve modelin daha isabetli sınırlama yapabilmesini sağlamıştır. Yığın boyutu (batch size) 16 olarak belirlenmiş, bu değer bellek kullanımını ve eğitim stabilitesini dengelemiştir. Başlangıç öğrenme oranı eğitim aşamasında 0.0001, ince ayar aşamasında ise daha hassas ağırlık güncellemeleri için 0.00001 olarak seçilmiştir. Öğrenme oranı ölçeği 0.01 ve sönümlenme katsayısı 0,6 olarak belirlenerek, modelin öğrenme sürecinde aşırı sıçramadan kaçınması hedeflenmiştir. Epok sayısı 5 olarak belirlenmiştir. Önyargı (bias) değeri 0,45 olarak belirlenmiş, bu değer farklı sınıflar arasındaki dengesizliği düzenlemek amacıyla kullanılmıştır. Ayrıca, görüntü yama üretim sürecinde minimum %30 örtüşme uygulanarak hem örnek sayısı artırılmış hem de modelin mekânsal genelleme yeteneği güçlendirilmiştir. Bu parametreler, literatürde önerilen değerlere ve yapılan ön test sonuçlarına göre belirlenmiş olup, modelin kararlı ve yüksek doğrulukta öğrenmesini sağlamak üzere optimize edilmiştir. Model 5 öğrenim dönemi (epok) için Tablo 5'te belirtilen parametreler kullanılarak çalıştırılmış ve Tablo 6'da verilen sonuçlar elde edilmiştir.

Tablo 6. En iyi sonuca sahip FPN-ResNeXt50 modelinin 5 epok çalışma sonuçları

Dosya	Doğruluk	IoU	Birleşik Ölçüt	Kesişim	Birleşim	Toplam Piksel	Gerçek Pozitif	Gerçek Negatif	Yanlış Pozitif	Yanlış Negatif
Austin1	0.9727	0.8316	0.9021	3366413	4048249	25000000	3366413	20951751	359491	322345
Austin2	0.9676	0.8244	0.8960	3802265	4612323	25000000	3802265	20387677	388573	421485
Austin3	0.9664	0.8047	0.8856	3457844	4296820	25000000	3457844	20703180	374285	464691
Austin4	0.9785	0.8166	0.8976	2395924	2933853	25000000	2395924	22066147	254248	283681
Austin5	0.9787	0.7795	0.8791	1880705	2412598	25000000	1880705	22587402	251602	280291
Austin6	0.9770	0.7507	0.8638	1728538	2302700	25000000	1728538	22697300	285922	288240
Şikago1	0.9609	0.7652	0.8630	3187468	4165799	25000000	3187468	20834201	574869	403462
Şikago2	0.8958	0.7213	0.8085	6743956	9350044	25000000	6743956	15649956	1586000	1020088
Şikago3	0.9213	0.6921	0.8067	4422871	6390454	25000000	4422871	18609546	1168138	799445
Şikago4	0.9367	0.7145	0.8256	3962950	5546611	25000000	3962950	19453389	913902	669759
Şikago5	0.9535	0.7548	0.8541	3577702	4740123	25000000	3577702	20259877	625359	537062
Şikago6	0.9183	0.7386	0.8284	5768699	7810788	25000000	5768699	17189212	1164456	877633
Kitsap1	0.9993	0.6971	0.8482	39558	56746	25000000	39558	24943254	7222	9966
Kitsap2	0.9981	0.8109	0.9045	203302	250713	25000000	203302	24749287	26871	20540
Kitsap3	0.9906	0.8525	0.9215	1352964	1587145	25000000	1352964	23412855	134094	100087
Kitsap4	0.9866	0.3239	0.6553	160114	494279	25000000	160114	24505721	46265	287900
Kitsap5	0.9916	0.4677	0.7297	183750	392905	25000000	183750	24607095	59386	149769
Kitsap6	0.9952	0.7561	0.8756	374853	495757	25000000	374853	24504243	61413	59491
Tirol-W1	0.9831	0.7815	0.8823	1510284	1932588	25000000	1510284	23067412	188121	234183
Tirol-W2	0.9901	0.7908	0.8904	938565	1186828	25000000	938565	23813172	126195	122068
Tirol-W3	0.9774	0.8041	0.8908	2323814	2889802	25000000	2323814	22110198	292910	273078
Tirol-W4	0.9860	0.8298	0.9079	1704781	2054447	25000000	1704781	22945553	149692	199974
Tirol-W5	0.9925	0.8979	0.9452	1647928	1835241	25000000	1647928	23164759	82798	104515
Tirol-W6	0.9895	0.8356	0.9125	1336668	1599681	25000000	1336668	23400319	129817	133196
Viyana1	0.9600	0.7970	0.8785	3927592	4927813	25000000	3927592	20072187	416975	583246
Viyana2	0.9561	0.8507	0.9034	6259261	7357908	25000000	6259261	17642092	550456	548191
Viyana3	0.9290	0.8209	0.8749	8141615	9917570	25000000	8141615	15082430	695892	1080063
Viyana4	0.9566	0.8614	0.9090	6751836	7837905	25000000	6751836	17162095	679556	406513
Viyana5	0.9623	0.8270	0.8946	4503708	5446160	25000000	4503708	19553840	511714	430738
Viyana6	0.9515	0.6619	0.8067	2373613	3586038	25000000	2373613	21413962	781461	430964
TOPLAM	0.9674	0.7828	0.8751	88029541	112459888	750000000	88029541	637540112	12887683	11542664

Tablo 6, FPN-ResNeXt50 modelinin eğitim süreci sonrasında test veri seti üzerinde elde ettiği performans değerlendirmelerini sunmaktadır. Eğitim veri seti üzerinde elde edilen %96,72 doğruluk değeriyle diğer modelden daha yüksek performans gösteren FPN-ResNeXt50 modeli ile test veri seti üzerinde %96,74 oranında doğruluk değeri elde edilmiştir. Model, her bölge için 5 epokluk eğitimden geçirilmiş ve ardından test bölgelerine uygulanarak doğruluk (Accuracy), kesişim birlik (IoU) ve birleşik ölçüt değerleri hesaplanmıştır. Test bölgeleri, Austin, Şikago, Kitsap, Batı Tirol ve Viyana şehirlerinin daha önce eğitim verisinde yer almayan farklı bölümlerinden seçilmiştir. Tablo, ayrıca her test örneği için gerçek pozitif (TP), yanlış pozitif (FP), yanlış negatif (FN) ve gerçek negatif (TN) piksel sayılarıyla birlikte toplam piksel sayısı, kesişim ve birleşim değerlerini de içermektedir. Bu sayede modelin sınıflandırma doğruluğu, bina tespitindeki hassasiyeti ve genel başarımları detaylı şekilde analiz edilmiştir. Elde edilen %96,74 doğruluk ve %78,28 IoU ortalama değerleri, modelin yüksek genelleme kabiliyetine ve farklı kentsel dokulara başarılı bir şekilde uyum sağladığına işaret etmektedir.

4. Bulgular ve Tartışma

Çalışma sonuçlarından elde edilen bulgularla;

- Yerleşim düzenlerindeki farklılıkların sınıflandırma sonuçlarına etki ettiği görülmüştür. Düzenli yapılaşmanın olduğu, binaların şekil ve büyüklüklerinin yaklaşık olarak benzer olduğu alanlarda başarımlar yüksek çıkmıştır.
- Yerleşim alanlarında binaların diğer objelerle yansıtım özelliklerinin farklı olduğu yüksek kontrasta sahip alanlarda bina çıkarımı sonuçlarının daha iyi olduğu görülmüştür.
- Binaların bulunduğu alanlardaki bitki örtüsünün özellikle ağaç kanopisinin tamamen ya da kısmen örttüğü binalar tespit edilememiştir. Bu gibi alanlarda geometriler kısmen ya da tamamen belirsizleştiklerinden çıkarım sonuçlarını etkilemiştir.
- Binalara birleşik olan ve aynı yansıtım değerlerine sahip olan bahçe otopark gibi alanlar bina geometrileri içerisinde gösterilmiştir. Bu gibi yerlerin çıkarım doğruluklarını arttırmak için yeni özniteliklerin tanımlanması ve bölütleme aşamasında bu öznitelikleri de belirli ağırlıklandırma ile dikkate alacak algoritmaların kullanılmasının uygun olacağı düşünülmektedir.
- Farklı bölgeler için elde edilen sonuçların farklı özellikteki alanlar için kullanımı anlamındaki genelleme doğruluklarının artırılması için daha fazla eğitim verisi ile modelin eğitilmesi sonuç doğruluklarını arttıracaktır.
- Kullanılacak eğitim verisinin, bitişik nizamda olan ve belirgin sınırlarla ayrılmamış ve/veya ağaç kanopisi ile örtülen binaların bulunduğu bölgeleri içeren veriler olması, bu nedenlerden dolayı yanlış negatif sınıflandırmanın oranını azaltılacağı düşünülmektedir.
- Modelin doğruluğu yanında, işlem hızı da önemlidir. İstatistiksel olarak anlamlı farkların olduğu sonuçların elde edilmesi daha az zaman maliyeti ile yapılabilir. Bu amaçla derin öğrenme modellerinin tasarlanmasında kullanılacak hiperparametreler de önemlidir. Bunların seçimi zaman/donanım maliyetlerini doğrudan etkilemektedir.
- Karşılaştırılan modellerde örneğin 10 eğitim dönemi için harcanan zaman ile 3 eğitim dönemi için harcanan zaman karşılaştırıldığında yaklaşık 3 kata varan bir maliyet söz konusudur. Hiperparametre seçimi model seçimi kadar önemlidir.
- Kullanılan donanım da zaman maliyeti açısından önemlidir.

Elde edilen sonuçlar literatürde aynı veri seti üzerinde ve aynı değerlendirme ölçütlerinin kullanıldığı çalışmalar ile karşılaştırılmış ve aşağıdaki sonuçlar elde edilmiştir.

Tablo 7. Literatürde farklı modellerle elde edilen IoU değerleri

Çalışma	Model	IoU
Xu vd. (2022)	CVNet	%77,61
Wang vd. (2022)	BuildFormer	%75,74-%91,44
Li, Z. vd. (2022)	HCRB-MSAN	%79,92-%91,79
Li, Q. vd. (2022)	Kodlayıcı-Kod Çözücü Mimari	%72,03
de Souza vd. (2022)	FLIM-U	%55,29
Milosavljević (2020)	ResNeXt50-FPN (10 Epok)	%82,23
Wang vd. (2025)	CANet	%81,42-%77,84-%62,55-%57,97
Bu Çalışmada	ResNeXt50-FPN (5 Epok)	%78,28

Tablo 7’deki sonuçlar değerlendirildiğinde;

- Xu vd. (2022), CVNet modeli ile aktif kontur yaklaşımı kullanarak kenar belirlemeye dayalı bina çıkarımı ile IoU değerini %77,61 olarak elde etmiştir. Bu çalışmada elde edilen sonuçlar ise, Xu. vd. (2022)’de elde edilenden 0,67 puan daha iyi olarak bulunmuştur.
- Wang vd. (2022), kayıp fonksiyon olarak entropi kayıp fonksiyonu (cross-entropyloss), aktivasyon fonksiyonu olarak ReLU6, özellik birleştirme fonksiyonu olarak özellik piramid ağı (FPN) kullanarak yapılandırılan bir çeşit görüntü transformatörü olan Build Former modeli ile IoU değerini, Inria veri seti için uygulamıştır. Çalışmada Massachusetts ve WHU veri setleri için IoU sırasıyla %75,74 ve %91,44 olarak bulmuşlardır. Elde edilen sonuçlardan Massachusetts veri seti için bulunan değer bu çalışmada bulunan değerden 2,54 puan daha düşüktür. WHU veri setinde bulunan değer ise 13,16 puan daha yüksektir. Aynı modelin farklı veri setleri için çok farklı başarımlar elde ettiği görülmektedir.
- Li, Z. vd. (2022), Inria veri seti için %79,92 IoU değeri, WHU veri seti için ise %91,79 IoU değeri elde etmiştir. Model kullanılan eğitim setlerinin içerdiği bina yapı düzenleri ve arka plan özelliklerine göre farklı başarımlar elde

etmiştir. Model 120 öğrenme dönemi sonunda bu sonuçları elde etmiştir. Model, Inria veri seti için, 5 öğrenme dönemi ile bu çalışmada elde edilenden 1,64 puan daha iyidir. Çalışmada 5 öğrenme dönemi için elde ettiğimiz sonuçla karşılaştırıldığında zaman maliyeti ve elde edilen başarımlar arasındaki farkın istatistiksel olarak anlamlı olup olmadığı değerlendirilmelidir.

- Bu çalışmada önerilen model %78,28 IoU değeri ile (Li, Q. vd., 2022)'de önerilen modelle elde edilenden 6,25 puan daha iyidir. Li, Q. vd. (2022), 31 adet eğitim seti kullanmıştır. Bu çalışmada kullanılan eğitim veri seti sayısından az olması başarımlarındaki farkın oluşmasında etken olduğu düşünülmektedir. Eğitim veri setinin az olmasının yetersiz öğrenmeye neden olarak başarımlarını etkilediği değerlendirilmektedir. Bunun yanında az eğitim verisi zaman maliyetini azaltmaktadır. Ancak başarımların nispeten anlamlı bir şekilde düşük olarak elde edilmiştir. Bu sonuç zaman maliyetine belirli bir düzeyde katlanmak gerekliliğini ortaya koymaktadır. Katlanması gereken zaman maliyeti için optimum düzey daha sonraki çalışmalarda araştırılması gereken bir problem olacaktır.
- de Souza vd. (2022), FLIM-U mimarisi ile kullanarak eğitilen verilerle HRNet-48 modelinden %55,29 IoU değeri elde etmiştir. Bu değer çalışmada elde edilen %78,28 değerinden oldukça kötü bir değerdir. Bunun nedeninin, bölütleme algoritması olarak geliştirilmiş U-Net mimarisinin (FLIM-U) kullanılması olduğu değerlendirilmektedir. Bu çalışmada modelin seçimi için yapılan değerlendirmede de en az doğruluk U-Net bölütleme mimarisi ile elde edilmiştir. Ayrıca çalışmada eğitim veri seti olarak belirli bölgeler seçilmiştir. Bu seçimin de modelin genelleme becerisini etkilediği düşünülmektedir.
- Milosavljević (2020)'den elde edilen sonuçlarda da IoU değeri %82,23 ile bu çalışmada elde edilenden 3,95 puan daha yüksek olarak bulunmuştur. Milosavljević (2020), çalışmasında 384x384 boyutlu veriler kullanılmış ve 10 eğitim dönemi için öğrenme yapılmıştır. Bu çalışmada ise 512x512 boyutlu veriler 5 öğrenme dönemi için eğitilmişlerdir. Zaman ve donanım maliyeti, çalışmamızdan yaklaşık 2,5 kat daha fazladır.
- Wang vd. (2025), önerdikleri CANet modelini dört farklı veri seti üzerinde uygulamışlar ve farklı IoU değerleri elde etmişlerdir. Bu da veri setlerine göre modelin farklı sonuçlar verdiğini göstermektedir.

Her kent için ayrı ayrı değerlendirme yapıldığında elde edilen değerlerle, topluca değerlendirme yapılması durumunda elde edilen değerler karşılaştırıldığında oluşan farklılıklar görülmektedir. Hatta her kentin farklı bölgeleri için bile farklı başarımlar ortaya çıkmıştır. Bunlar;

- Yerleşim alanlarının farklı coğrafi bölgelerde olması,
- Bölgelerin tarihsel, kültürel ve sosyal farklılıklardan dolayı ortaya çıkan şehircilik politikaları ve imar planı anlayışları ve uygulamaları olması ve bunun da bir sonucu olarak binaların geometrilerindeki ve yerleşim düzenlerindeki çeşitlilik,
- Görüntüdeki farklı yansıma, ışıma koşulları ve topoğrafya dolayısıyla ortaya çıkan gölgelenme problemlerinin etkileri ile spektral farklılıklardan dolayı ortaya çıkan kontrast problemi,
- Özellikle yeşil alanları fazla olan yerleşim bölgelerinde, yeşil bitki örtüsü kanopisinin kısmen ya da tamamen binaların görünürlüğünü etkilemesi gibi nedenler, sonuçlardaki bu farklılıkların sebebi olarak değerlendirilmektedir.

Ayrıca değerlendirme için seçilen ölçütlerin belirlenmesi aşamasında, test ve eğitim veri seti seçiminde her ne kadar örneklem büyütme yöntemleriyle, algoritmanın çok çeşitli mekânsal, uzamsal ve spektral özellikler çıkarması sağlanmaya çalışılsa da öğrenmenin genelleştirilmesi ve küresel ölçekte uygulanabilir olmasının sağlanabilmesi için, bu ölçütlerin daha yerel özellikler üzerinden belirlenip genele yayılması yaklaşımlarının da değerlendirilmesi gerektiği düşünülmektedir. Mevcut çalışmalarda daha genel yaklaşımlar çerçevesinde bu ölçütler belirlenip, kent ölçeğinde hatta daha küçük ölçeklerde alanlar için birer kriter olarak kullanıldıkları bilinmektedir.

Derin öğrenme algoritmalarının ve bir derin sinir ağının eğitilmesi aşamasında kullanılan fonksiyonların da işlevselliğinin daha da iyileştirilmesinin bölütleme sonuçlarının doğruluğunu arttıracığı öngörülmektedir.

5. Sonuç ve Öneriler

Uzaktan algılama görüntülerinden nesnelerin çıkarılması işlemi oldukça karmaşık ve yoğun hesaplama gerektiren işlemlerdir. Elde edilen sonuçların değerlendirilmesi standart bir ölçüte göre yapılamamakta ve dolayısıyla karşılaştırmalar göreceli olmaktadır. En iyi sonucun elde edilip edilemediği yalnızca insan gözü sistemleri temel alınarak söylenebilmektedir. Gelişen sensörlerle birlikte çok fazla veri ve istenmeyen ölçüde detay elde edildiğinden bu durum hesaplama ve değerlendirmelerin yapıldığı bilgisayar sistemlerine ek yük getirmektedir. Bunun önlenmesi için;

- Öncelikle verilerin standartlaştırılmasını sağlayacak ve verileri istenen uygulamalar açısından süzebilecek bir yöntemin geliştirilmesi, sonrasında uygulanan yöntemlerin hantallığını azaltabilecektir.
- Yüksek çözünürlüklü görüntülerden nesnelerin çıkarılmasında, nesne özelliklerinin iyi belirlenmesi, modellenmesi ve değerlendirilmesi büyük önem taşımaktadır.

- Var olan nesne özelliklerine ek, yeni ve modellendirilebilir özelliklerin tanımlanması, uygulamaların başarısını arttıracaktır. Bu nedenle insan yapımı nesnelerin sınıflandırılmasında yeni geometrik ve semantik özellik bilgilerinin ortaya konması gelecekte yapılacak çalışmalar için hedefler arasında düşünülmelidir. Özellikle yerleşim alanlarında planlama ve uygulamaların bir düzen içerisinde ve belirli planlamalarla yapılması verilerin standartlaşmasını sağlayacak ve böylece, farklı bölgeler için farklı özellik ve yöntem kullanılması gereğini ortadan kaldıracaktır.
- Mevcut uygulamaların etkinliklerinin artırılması ve karmaşık uygulamaları basitleştirici yeni yazılımlar geliştirilmesi de üzerinde çalışılabilecek konulardandır.

Sonuç olarak, birçok yöntemin birleştirilmesi ile çok fazla girdiye sahip yöntemlerin uygulaması sonuçları iyiye götürüyor gibi olsa da var olan yöntemler ile uygulanacak bileşik yöntemlerin doğruluğa etkileri bir noktadan sonra anlamlı bir şekilde farklılaşmayacak, sadece çözüm aşamasında ek yükler getirecektir. Derin öğrenme modellerinin eğitilmesi sırasında hem omurganın hem de bölütleme mimarisinin seçimi de önemlidir. Ayrıca modelin farklı katmanlarında kullanılan filtreler, aktivasyon fonksiyonları, kayıp fonksiyonları için çok çeşitli alternatifler vardır. Bu alternatifler arasında en uygun olan kombinasyonun belirlenmesi, çok fazla deneme gerektirmektedir. Buna ek olarak belirli bir eğitim veri seti için en optimum algoritmanın başka veri setleri için de en uygun olacağı ileri sürülememektedir. Bu nedenle derin öğrenme modellerinin genelleme yeteneğinin tüm coğrafyalar ve yerleşim yerleri için tekil bir sonuca ulaşamayacağı değerlendirilmektedir. En uygun modelin, çalışılan bölgeyi en iyi temsil eden eğitim veri seti ile eğitilmiş ve bazı hiperparametre değişiklikleri ile iyileştirilmiş model olacağı sonucuna varılmıştır.

Ayrıca derin öğrenme modelleri kullanılarak görüntü işleme yüksek maliyetli bilgisayar donanımı da gerektirmektedir. Kullanılan fonksiyonların hesaplama süreleri ve geri besleme algoritmalarının mantığı gereği bellek ihtiyacı da artmaktadır. Görüntülerin işlenmesi ayrıca yüksek oranda GPU belleği de gerektirmektedir. Bu maliyetlerin azaltılması ve kişisel bilgisayarlar kullanılarak uygun sürelerde çözümler üretebilmek için hesaplama süresi ve bellek ihtiyacını nispeten azaltacak yeni matematiksel ve istatistiksel yaklaşımların geliştirilmesi uygun olacaktır. Gelecekte önerilecek yaklaşımların bu konular üzerinde de yoğunlaşmasının derin öğrenme konusundaki gelişmelerden olacağı değerlendirilmektedir.

Teşekkür

Bu makale birinci yazarın doktora tezinden üretilmiştir.

Kaynaklar

- Attarzadeh, R., & Momeni, M. (2012). Object-based building extraction from high resolution satellite imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 39, 57–60.
- Bansal, K., & Singh, A. (2023). Development of VGG-16 transfer learning framework for geographical landmark recognition. *Intelligent Decision Technologies*, 17(3), 799–810.
- Benali, A., Dermeche, H., Belhadj, S., Adnane, A., & Amar, R. H. E. (2014, April 14–16). *Buildings extraction of very high spatial resolution satellite images* [Conference presentation]. 2014 International Conference on Multimedia Computing and Systems (ICMCS), Marrakesh, Morocco.
- Bengio, Y. (2009). *Learning deep architectures for AI*. Now Foundations and Trends.
- de Souza, I. E., Cazarin, C. L., Veronez, M. R., Gonzaga, L., & Falcao, A. X. (2022). User-guided data expansion modeling to train deep neural networks with little supervision. *IEEE Geoscience and Remote Sensing Letters*, 19, Article 6515505. <https://doi.org/10.1109/LGRS.2022.3201437>
- Fradkin, M., Maitre, H., & Roux, M. (2001). Building detection from multiple aerial images in dense urban areas. *Computer Vision and Image Understanding*, 82(3), 181–207.
- Ghiasi, G., Lin, T. Y., & Le, Q. V. (2019, June 15–20). *Nas-fpn: Learning scalable feature pyramid architecture for object detection* [Conference presentation]. 2019 IEEE/CVF conference on computer vision and pattern recognition, Long Beach, CA, USA.
- Guo, J., Pan, Z., Lei, B., & Ding, C. (2017). Automatic color correction for multi-source remote sensing images with Wasserstein CNN. *Remote Sensing*, 9(5), Article 483. <https://doi.org/10.3390/rs9050483>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016, June 26–July 1). *Deep residual learning for image recognition* [Conference presentation]. IEEE conference on computer vision and pattern recognition, Las Vegas, Nevada.
- Ismail, A. R., Nisa, S. Q., Shaharuddin, S. A., Masni, S. I., & Amin, S. A. S. (2024). Utilising VGG-16 of convolutional neural network for medical image classification. *International Journal on Perceptive and Cognitive Computing*, 10(1), 113–118.

- Jin, X., & Davis, C. H. (2005). Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information. *EURASIP Journal on Advances in Signal Processing*, 2005, Article 745309. <https://doi.org/10.1155/ASP.2005.2196>
- Kaur, T., & Gandhi, T. K. (2019, December 19–21). *Automated brain image classification based on VGG-16 and transfer learning* [Conference presentation]. 2019 International Conference on Information Technology (ICIT), Bhubaneswar, India.
- Khatriker, S., & Kumar, M. (2018). Building footprint extraction from high resolution satellite imagery using segmentation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 123–128.
- Lari, Z., & Ebadi, H. (2007, May 29–June 1). *Automated building extraction from high-resolution satellite imagery using spectral and structural information based on artificial neural networks* [Workshop presentation]. ISPRS Hannover Workshop, Hannover, Germany.
- Li, Q., Shi, Y., & Zhu, X. X. (2022, July 17–22). *Feature and output consistency training for semi-supervised building footprint generation* [Symposium presentation]. IGARSS 2022 – IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia.
- Li, Q., Mou, L., Sun, Y., Hua, Y., Shi, Y., & Zhu, X. X. (2024). A review of building extraction from remote sensing imagery: Geometrical structures and semantic attributes. *IEEE Transactions on Geoscience and Remote Sensing*, 62, Article 4702315. <https://doi.org/10.1109/TGRS.2024.3369723>
- Li, W., He, C., Fang, J., Zheng, J., Fu, H., & Yu, L. (2019). Semantic segmentation-based building footprint extraction using very high-resolution satellite images and multi-source GIS data. *Remote Sensing*, 11(4), Article 403. <https://doi.org/10.3390/rs11040403>
- Li, Z., Zhang, Z., Chen, D., Zhang, L., Zhu, L., Wang, Q., ... & Peng, X. (2022). HCRB-MSAN: Horizontally connected residual blocks-based multi-scale attention network for semantic segmentation of buildings in HSR remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 5534–5544.
- Liu, S., Bai, X., Fang, M., Li, L., & Hung, C. C. (2022). Mixed graph convolution and residual transformation network for skeleton-based action recognition. *Applied Intelligence*, 52(2), 1544–1555.
- Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017, July 23–28). *Can semantic labeling methods generalize to any city? The INRIA aerial image labeling benchmark* [Symposium presentation]. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA.
- Milosavljević, A. (2020). Automated processing of remote sensing imagery using deep semantic segmentation: A building footprint extraction case. *ISPRS International Journal of Geo-Information*, 9(8), Article 486. <https://doi.org/10.3390/ijgi9080486>
- Nurkarim, W., & Wijayanto, A. W. (2023). Building footprint extraction and counting on very high-resolution satellite imagery using object detection deep learning framework. *Earth Science Informatics*, 16(1), 515–532.
- Prathiba, A. P., Rastogi, K., Jain, G. V., & Govind Kumar, V. V. (2019). Building footprint extraction from very-high-resolution satellite image using object-based image analysis (OBIA) technique. In J. Ghosh & I. da Silva (Eds.), *Applications of Geomatics in Civil Engineering* (pp. 517–529). Springer Singapore.
- Pravitasari, A. A., Iriawan, N., Almuhyar, M., Azmi, T., Irfamah, I., Fithriasari, K., ... & Ferriastuti, W. (2020). UNet-VGG16 with transfer learning for MRI-based brain tumor segmentation. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 18(3), 1310–1318.
- Ps, P., & Aithal, B. H. (2023). Building footprint extraction from very high-resolution satellite images using deep learning. *Journal of Spatial Science*, 68(3), 487–503.
- Ronneberger, O., Fischer, P., & Brox, T. (2015, October 5–9). *U-net: Convolutional networks for biomedical image segmentation* [Conference presentation]. 18th International Conference on medical image computing and computer-assisted intervention (MICCAI 2015), Munich, Germany.
- Shackelford, A. K., & Davis, C. H. (2003). A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 41(10), 2354–2363.
- Srivastava, V., Avudaiammal, R., & V George, S. (2024). Investigations on extraction of buildings from RS imagery using deep learning models. *International Journal of Remote Sensing*, 45(1), 68–100.
- Sewada, R., & Goyal, H. (2025). A novel VGG-16 adaptation for multi-band satellite image classification: Optimized preprocessing and class-specific augmentation. *Journal of Computational and Cognitive Engineering*. Advance online publication. <https://doi.org/10.47852/bonviewJCCCE5202480>
- Tahir, M. Z., Lyu, X., Nasir, M., & Zhang, S. (2025). Advanced image enhancement and a lightweight feature pyramid network for detecting microaneurysms in diabetic retinopathy screening. *International Journal of Imaging Systems and Technology*, 35(1), Article e70004. <https://doi.org/10.1002/ima.70004>

- Tejeswari, B., Sharma, S. K., Kumar, M., & Gupta, K. (2022). Building footprint extraction from space-borne imagery using deep neural networks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 641–647.
- Turker M., San D. (2010) Building detection from pan-sharpened IKONOS imagery through support vector machines classification. *Int Arch Photogramm Remote Sens Spat Inf Sci*, 38(Part 8), 841–846.
- Wang, L., Fang, S., Meng, X., & Li, R. (2022). Building extraction with Vision Transformer. *IEEE Transactions on Geoscience and Remote Sensing*, 60, Article 5625711. <https://doi.org/10.1109/TGRS.2022.3186634>
- Wang, S., Shao, Z., Hou, D., Wang, Y., Wang, J., & Cai, B. (2025). CANet: A spatial structure constraint and local semantic awareness based network for weakly supervised building extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 63, Article 4501919. <https://doi.org/10.1109/TGRS.2025.3537099>
- Whiteside, T. G., Boggs, G. S., & Maier, S. W. (2011). Comparing object-based and pixel-based classifications for mapping savannas. *International Journal of Applied Earth Observation and Geoinformation*, 13(6), 884–893.
- Yao, Y., Jiang, Z., Zhang, H., Cai, B., Meng, G., & Zuo, D. (2017, July 23–28). *Chimney and condensing tower detection based on Faster R-CNN in high-resolution remote sensing images* [Symposium presentation]. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA.
- Zheng, C., & Wang, L. (2014). Semantic segmentation of remote sensing imagery using object-based Markov random field model with regional penalties. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(5), 1924–1935.
- Zhong, Z., Li, J., Cui, W., & Jiang, H. (2016, July 10–15). *Fully convolutional networks for building and road extraction: Preliminary results* [Symposium presentation]. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China.
- Zhou, X., & Zhang, L. (2022). SA-FPN: An effective feature pyramid network for crowded human detection. *Applied Intelligence*, 52(11), 12556–12568.
- Xu, Z., Xu, C., Cui, Z., Zheng, X., & Yang, J. (2022, June 18–24). *CVNet: Contour vibration network for building extraction* [Conference presentation]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA.

EKLER:**Ek Tablo 1. Eğitim bölgeleri ve özellikleri**

Bölgeler	Şekil	Özellikleri
Austin1	Şekil 1	Yoğun yerleşim alanlarının olduğu, ağaç örtüsü etkisi bulunan, az sayıda büyük ve farklı biçimli binaların bulunduğu bina çatılarının farklı spektral yansıtımı olan (açık-koyu görünümlü) bölge
Austin2		Konut alanları ve farklı spektral yansıtımı olan çatılara sahip büyük boyutlu binaların olduğu, yeşil alan görünümü yoğunluklu bölge
Austin3		Austin2 bölgesi ile benzer özellikli bölge
Austin4		Bir ana ulaşım arterinin her iki yanında topoğrafik olarak farklı düzeyde yerleşimlerin bulunduğu bölge
Austin5		Bir nehir etrafında yerleşmiş spektral yansıtım özelliği açısından koyu görüntü veren farklı büyüklük ve dokuya sahip binaların bulunduğu bölge
Austin6		Nehir ve bir iç su yolu bulunan, nehir üzerinde benzer özellikli yapılaşmış binalar olan, yeşil alan için ayrılmış büyükçe bir bölge kenarına yapılanmış spektral yansıtımı benzer ve nispeten belirsiz olan binaların bulunduğu bölge
Şikago1	Şekil 2	Fabrikalar gibi büyük boyutlu ve farklı biçim özellikli binalarla, konut amaçlı kullanılan binaların birlikte bulunduğu bölge
Şikago 2		Konut amaçlı kullanılan binaların daha yoğun olduğu, binaların doku özelliklerinin belirgin olduğu bölge
Şikago 3		Sadece yerleşim amaçlı yapılan binaların bulunduğu, binaların büyüklük ve biçimlerinin homojen olduğu çok sık yerleşim düzenine sahip bölge
Şikago 4		Aynı biçim ve büyüklükteki binaların nispeten seyrek yapılaşma düzeninde yapılaştığı bölge
Şikago 5		Farklı biçim ve boyutlarda binaların 2. Bölgeye göre nispeten az yoğunluklu yerleşim bölgesi
Şikago 6		Konut ve diğer amaçlı olarak yapılmış büyük binaların bulunduğu bölge
Kitsap1	Şekil 3	Topoğrafik olarak engebeli, yoğun orman alanları içerisinde seyrek olarak binaların bulunduğu, yayla yerleşimi özellikli bölge
Kitsap2		Yoğun orman alanı ile birlikte az sayıda konut dışı alanın ve konut alanlarının bulunduğu bölge
Kitsap3		Orman alanı ile bütünlük olarak yerleşmiş planlı yapılaşma bölgesi
Kitsap4		Orman alanı içerisinde farklı bölgelerde yerleşmiş az sayıda farklı biçim ve büyüklükte binaların olduğu bölge
Kitsap5		Yoğun orman alanı içerisinde çoğu ağaç kanopisi örtüsünden etkilenmiş az sayıda bina bulunan bölge
Kitsap6		Orman alanı içerisinde ağaç kanopisinin etkisi altında kalmış ve tek tek bölgeye yayılmış binaların bulunduğu bölge
Viyana1	Şekil 4	Kent içi peyzaj düzenlemelerinin olduğu, farklı biçim ve büyüklükte ve farklı spektral yansıtım değerine sahip binaları içeren ve imar adalarının boyutlarının benzer olduğu bölge
Viyana2		Genel olarak bakıldığında iki farklı imar düzeni ile yapılaşmış, benzer spektral yansıtım yapan bina çatılarına sahip, farklı doku özellikli binaların ağaçlarla iç içe bulunduğu bölge
Viyana3		Viyana2 Bölgesi ile benzer özellikli nispeten daha az sayıda ağaç bulunduran, dolayısıyla ikinci bölgeye nazaran ağaç örtüsü altında kalabilecek bina sayısının az olduğu, benzer spektral yansıtıma sahip çatılı binaların bulunduğu bölge
Viyana4		Aynı imar planı yerleşim düzeniyle yapılaşmış, iki ve üçüncü bölgedeki farklı iki yapılaşmanın tek başına aynı özellikli olanının bulunduğu bölge
Viyana5		Nehir kenarı yerleşiminin olduğu, farklı biçim, büyüklük ve doku özellikli binaların bulunduğu, yeşil alanların çok fazla olmadığı bölge
Viyana6		Nehir kenarına yerleşmiş konut binaları ile konut dışı büyük boyutlu binaların olduğu bölge
Tirol1	Şekil 5	Orman alanları ile bina alanlarının sınır olduğu, topoğrafik olarak engebeli olan bölge
Tirol2		Kıyı-kenar çizgisi bulunan, bu çizginin kara tarafında tarım arazilerinin olduğu, düzenli yapılaşmış binaların bulunduğu bölge
Tirol3		Dağlık bir orman alanı, orman ağaçları ile çevrilmiş bir yerleşim bölgesi ve tarım arazisine komşu yerleşim alanının birlikte bulunduğu bölge
Tirol4		Dağlık bir ormanlık alanla bütünlük konut dışı büyük biçimli binalarında bulunduğu yerleşim bölgesi
Tirol5		Tarım alanları ile sınırlanmış yoğun olarak konut dışı binaların Fabrika, sosyal tesis... vb.) bulunduğu bölge
Tirol6		Tarım alanları ile bütünlük, konut ve konut dışı binaların orman bölgeleri ile sınırlanmış binalarla birlikte bulunduğu bölge

Ek Tablo 2. Austin bölgesi eğitim verileri

Model - Bölütleme	Dosya	Doğruluk	IoU	Birleşik Ölçüt	Kesişim	Birleşim	Toplam Piksel	Gerçek Pozitif	Gerçek Negatif	Yanlış Pozitif	Yanlış Negatif
FPN - ResNeXt50	Austin 1	0.9737	0.8354	0.9046	3344201	4002877	25000000	3344201	20997123	314119	344557
VGG16 - UNET	Austin 1	0.9427	0.6385	0.7906	2528040	3959440	25000000	2528040	21040560	270682	1160718
FPN - ResNeXt50	Austin 2	0.9681	0.8267	0.8971	3782139	4578676	25000000	3782139	20421324	354926	441611
VGG16 - UNET	Austin 2	0.9385	0.6648	0.8013	3037299	4574101	25000000	3037299	20425899	350351	1186451
FPN - ResNeXt50	Austin 3	0.9672	0.8078	0.8875	3441328	4260277	25000000	3441328	20739723	337742	481207
VGG16 - UNET	Austin 3	0.9254	0.5524	0.7389	2302433	4167928	25000000	2302433	20832072	245393	1620102
FPN - ResNeXt50	Austin 4	0.9789	0.8197	0.8993	2393403	2919872	25000000	2393403	22080128	240267	286202
VGG16 - UNET	Austin 4	0.9568	0.6278	0.7923	1822746	2903307	25000000	1822746	22096693	223702	856859
FPN - ResNeXt50	Austin 5	0.9791	0.7817	0.8804	1868595	2390512	25000000	1868595	22609488	229516	292401
VGG16 - UNET	Austin 5	0.9613	0.5864	0.7739	1372384	2340227	25000000	1372384	22659773	179231	788612

Ek Tablo 3. Şikago bölgesi eğitim verileri

Model- Bölütleme	Dosya	Doğruluk	IoU	Birleşik Ölçüt	Kesişim	Birleşim	Toplam Piksel	Gerçek Pozitif	Gerçek Negatif	Yanlış Pozitif	Yanlış Negatif
FPN-ResNeXt50	Şikago1	0.9614	0.7670	0.8642	3180334	4146236	25000000	3180334	20853764	555306	410596
VGG16- UNET	Şikago1	0.9512	0.7220	0.8366	3170964	4391687	25000000	3170964	20608313	800757	419966
FPN-ResNeXt50	Şikago2	0.8925	0.7178	0.8051	6835028	9522838	25000000	6835028	15477162	1758794	929016
VGG16- UNET	Şikago2	0.8854	0.7007	0.7931	6709042	9574282	25000000	6709042	15425718	1810238	1055002
FPN-ResNeXt50	Şikago3	0.9202	0.6904	0.8053	4449823	6444834	25000000	4449823	18555166	1222518	772493
VGG16- UNET	Şikago3	0.9121	0.6635	0.7878	4331934	6529056	25000000	4331934	18470944	1306740	890382
FPN-ResNeXt50	Şikago4	0.9366	0.7164	0.8265	4001478	5585380	25000000	4001478	19414620	952671	631231
VGG16- UNET	Şikago4	0.9285	0.6848	0.8066	3884710	5672889	25000000	3884710	19327111	1040180	747999
FPN-ResNeXt50	Şikago5	0.9528	0.7522	0.8525	3584109	4764544	25000000	3584109	20235456	649780	530655
VGG16- UNET	Şikago5	0.9475	0.7268	0.8371	3490737	4803042	25000000	3490737	20196958	688278	624027
FPN-ResNeXt50	Şikago6	0.9163	0.7343	0.8253	5779207	7870670	25000000	5779207	17129330	1224338	867125
VGG16- UNET	Şikago6	0.9082	0.7089	0.8086	5589100	7884037	25000000	5589100	17115963	1237705	1057232

Ek Tablo 4. Kitsap bölgesi eğitim verileri

Model - Bölütleme	Dosya	Doğruluk	IoU	Birleşik Ölçüt	Kesişim	Birleşim	Toplam Piksel	Gerçek Pozitif	Gerçek Negatif	Yanlış Pozitif	Yanlış Negatif
FPN - ResNeXt50	Kitsap1	0.9993	0.7085	0.8539	39928	56359	25000000	39928	24943641	6835	9596
VGG16 - UNET	Kitsap1	0.9993	0.6841	0.8417	39753	58112	25000000	39753	24941888	8588	9771
FPN - ResNeXt50	Kitsap2	0.9982	0.8143	0.9062	202731	248949	25000000	202731	24751051	25107	21111
VGG16 - UNET	Kitsap2	0.9976	0.7722	0.8849	201387	260804	25000000	201387	24739196	36962	22455
FPN - ResNeXt50	Kitsap3	0.9907	0.8525	0.9216	1349056	1582438	25000000	1349056	23417562	129387	103995
VGG16 - UNET	Kitsap3	0.9877	0.8137	0.9007	1340515	1647441	25000000	1340515	23352559	194390	112536
FPN - ResNeXt50	Kitsap4	0.9873	0.3367	0.6625	160645	477176	25000000	160645	24522824	29162	287369
VGG16 - UNET	Kitsap4	0.9866	0.3121	0.6494	151990	486927	25000000	151990	24513073	38913	296024
FPN-ResNeXt50	Kitsap5	0.9918	0,4870	0.7394	193731	397812	25000000	193731	24602188	64293	139788
VGG16- UNET	Kitsap5	0.9913	0,4420	0.7166	173185	391848	25000000	173185	24608152	58329	160334
FPN-ResNeXt50	Kitsap6	0.9954	0,7665	0.8810	374698	488838	25000000	374698	24511162	54494	59646
VGG16- UNET	Kitsap6	0.9944	0,7215	0.8580	363457	503725	25000000	363457	24496275	69381	70887

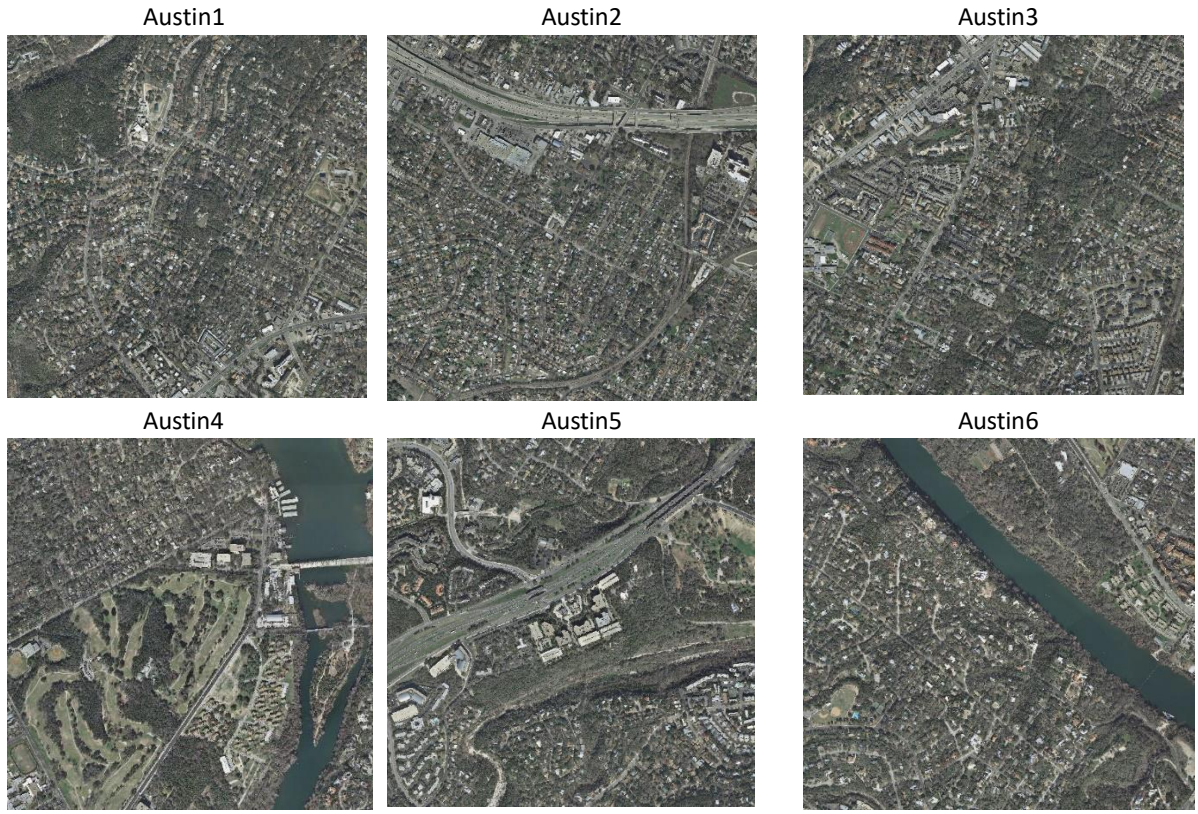
Ek Tablo 5. Batı Tirol bölgesi eğitim verileri

Model-Bölütleme	Dosya	Doğruluk	IoU	Birleşik Ölçüt	Kesişim	Birleşim	Toplam Piksel	Gerçek Pozitif	Gerçek Negatif	Yanlış Pozitif	Yanlış Negatif
FPN-ResNeXt50	BatıTirol1	0.9830	0.7804	0.8817	1511420	1936761	25000000	1511420	23063239	192294	233047
VGG16- UNET	BatıTirol1	0.9806	0.7591	0.8699	1527447	2012183	25000000	1527447	22987817	267716	217020
FPN-ResNeXt50	BatıTirol2	0.9905	0.7968	0.8936	932761	1170648	25000000	932761	23829352	110015	127872
VGG16- UNET	BatıTirol2	0.9874	0.7462	0.8668	928729	1244603	25000000	928729	23755397	183970	131904
FPN-ResNeXt50	BatıTirol3	0.9789	0.8161	0.8975	2339869	2867114	25000000	2339869	22132886	270222	257023
VGG16- UNET	BatıTirol3	0.9734	0.7796	0.8765	2353141	3018553	25000000	2353141	21981447	421661	243751
FPN-ResNeXt50	BatıTirol4	0.9859	0.8274	0.9066	1693554	2046781	25000000	1693554	22953219	142026	211201
VGG16- UNET	BatıTirol4	0.9818	0.7856	0.8837	1669573	2125236	25000000	1669573	22874764	220481	235182
FPN-ResNeXt50	BatıTirol5	0.9925	0.8970	0.9447	1642251	1830834	25000000	1642251	23169166	78391	110192
VGG16- UNET	BatıTirol5	0.9891	0.8539	0.9215	1596295	1869441	25000000	1596295	23130559	116998	156148
FPN-ResNeXt50	BatıTirol6	0.9895	0.8362	0.9128	1340828	1603534	25000000	1340828	23396466	133670	129036
VGG16- UNET	BatıTirol6	0.9862	0.7920	0.8891	1317465	1663379	25000000	1317465	23336621	193515	152399

Ek Tablo 6. Viyana bölgesi eğitim verileri

Model-Bölütleme	Dosya	Doğruluk	IoU	Birleşik Ölçüt	Kesişim	Birleşim	Toplam Piksel	Gerçek Pozitif	Gerçek Negatif	Yanlış Pozitif	Yanlış Negatif
FPN-ResNeXt50	Viyana1	0.9601	0.7974	0.8788	3922067	4918545	25000000	3922067	20081455	407707	588771
VGG16- UNET	Viyana1	0.9551	0.7801	0.8676	3980795	5102873	25000000	3980795	19897127	592035	530043
FPN-ResNeXt50	Viyana2	0.9555	0.8492	0.9024	6259540	7370965	25000000	6259540	17629035	563513	547912
VGG16- UNET	Viyana2	0.9491	0.8311	0.8901	6266110	7539243	25000000	6266110	17460757	731791	541342
FPN-ResNeXt50	Viyana3	0.9265	0.8152	0.8708	8107155	9945019	25000000	8107155	15054981	723341	1114523
VGG16- UNET	Viyana3	0.9104	0.7824	0.8464	8050359	10289126	25000000	8050359	14710874	1067448	1171319
FPN-ResNeXt50	Viyana4	0.9547	0.8560	0.9053	6739276	7872733	25000000	6739276	17127267	714384	419073
VGG16- UNET	Viyana4	0.9388	0.8125	0.8756	6629786	8159896	25000000	6629786	16840104	1001547	528563
FPN-ResNeXt50	Viyana5	0.9582	0.8100	0.8841	4453111	5497713	25000000	4453111	19502287	563267	481335
VGG16- UNET	Viyana5	0.9487	0.7707	0.8597	4312309	5595596	25000000	4312309	19404404	661150	622137
FPN-ResNeXt50	Viyana6	0.9522	0.6672	0.8097	2393893	3587826	25000000	2393893	21412174	783249	410684
VGG16- UNET	Viyana6	0.9505	0.6574	0.8039	2376015	3614229	25000000	2376015	21385771	809652	428562

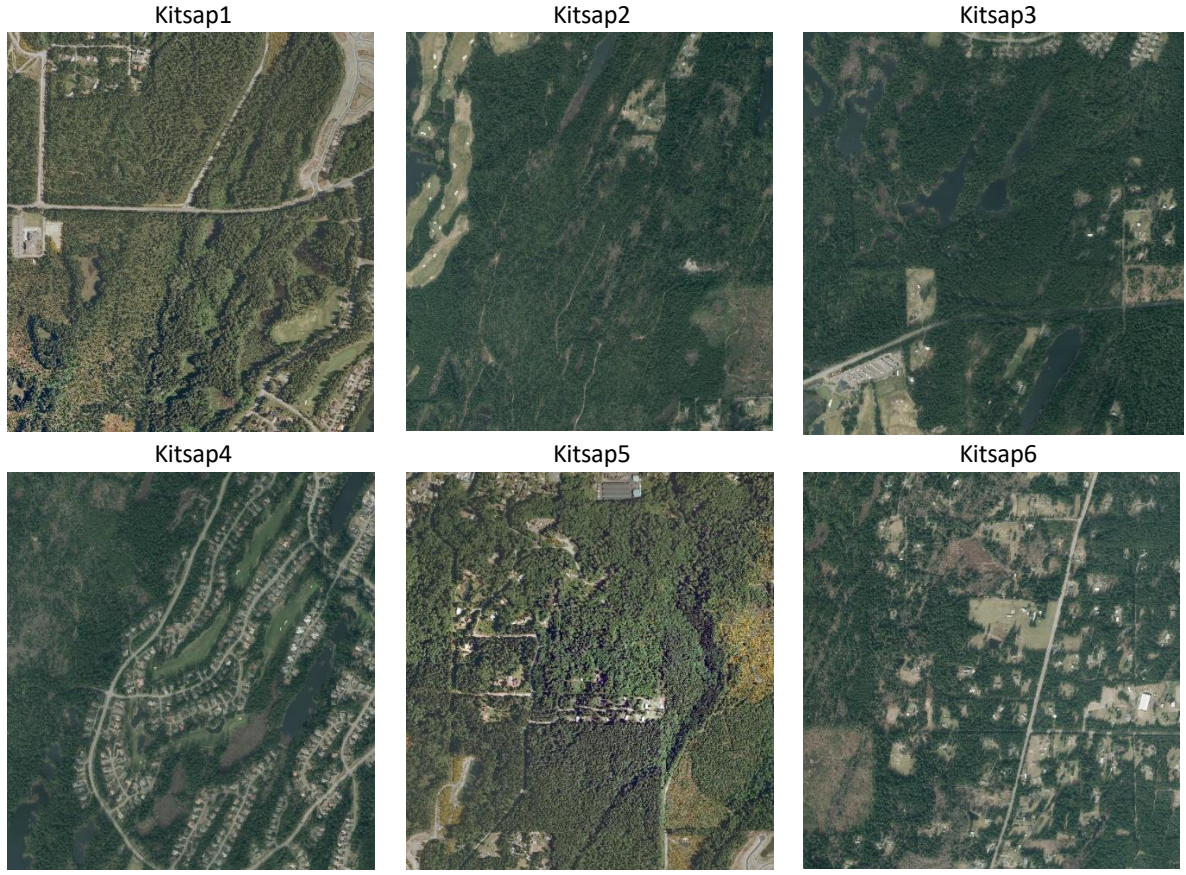
Her biri 6 bölgeye ayrılmış eğitim bölgelerinin görüntüleri Ek Şekil 1 – Ek Şekil 5’te verilmektedir.



Ek Şekil 1. Austin Kenti 6 bölgeye ayrılmış eğitim veri seti görüntüsü



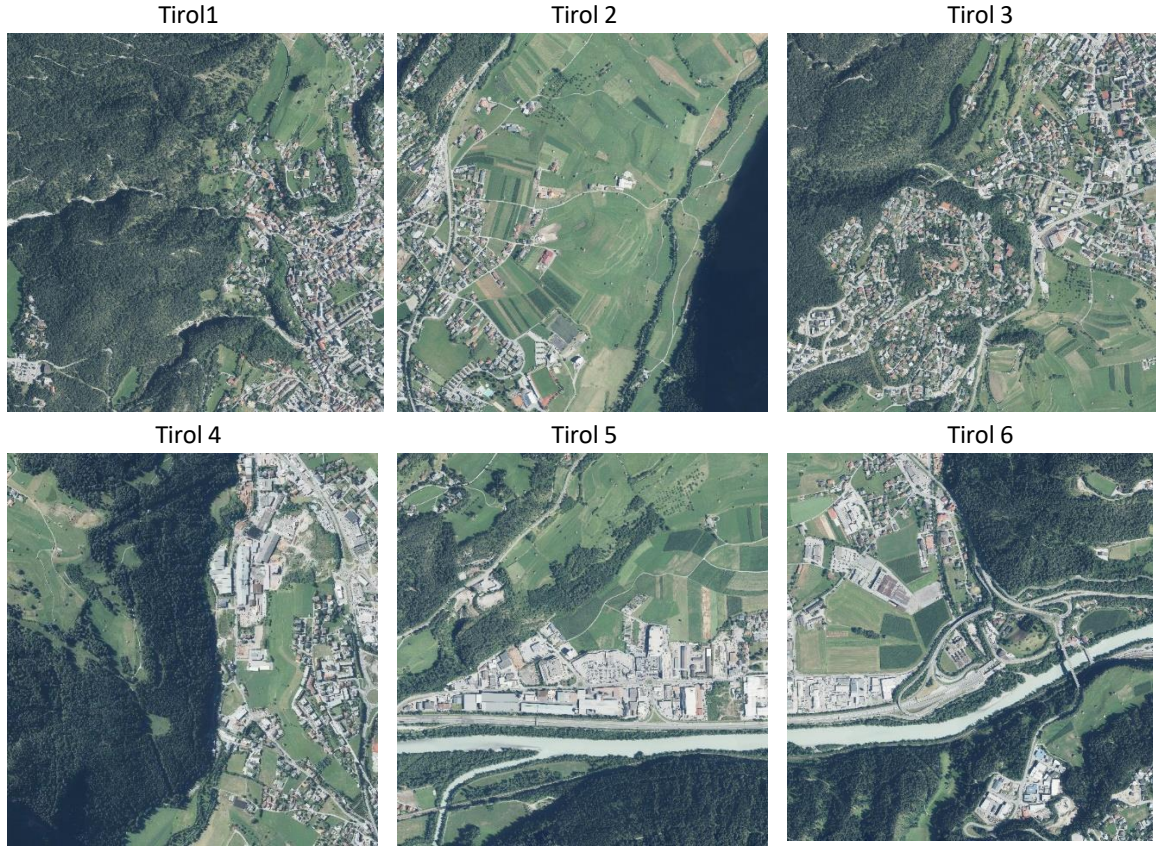
Ek Şekil 2. Şikago Kenti 6 bölgeye ayrılmış eğitim veri seti görüntüsü



Ek Şekil 3. Kitsap Kenti 6 bölgeye ayrılmış eğitim veri seti görüntüsü



Ek Şekil 4. Viyana Kenti 6 bölgeye ayrılmış eğitim veri seti görüntüsü



Ek Şekil 5. Tirol Kenti 6 bölgeye ayrılmış eğitim veri seti görüntüsü