

# ARAZİ KULLANIMININ BELİRLENMESİ AMACIYLA ORTOFOTO GÖRÜNTÜLERİN DERİN ÖĞRENME YÖNTEMLERİYLE SINIFLANDIRILMASI

Berfin Pelin GÜNER<sup>1</sup>, Onur Can BAYRAK<sup>2</sup>, Alper ŞEN<sup>3</sup>

<sup>1</sup>Berfin Pelin GÜNER, Yıldız Teknik Üniversitesi, Harita Mühendisliği Bölümü, Esenler, İstanbul, pelin.guner@std.yildiz.edu.tr

<sup>2</sup>Arş. Gör. Onur Can BAYRAK, Yıldız Teknik Üniversitesi, Harita Mühendisliği Bölümü, Esenler, İstanbul, onurcb@yildiz.edu.tr

<sup>3</sup>Doç. Dr. Alper ŞEN, Yıldız Teknik Üniversitesi, Harita Mühendisliği Bölümü, Esenler, İstanbul, alpersen@yildiz.edu.tr

## ÖZET

Arazi kullanımı, bir bölgenin insan faaliyetleri için nasıl ve hangi amaçlarla kullanıldığını tanımlar. Arazi kullanımının doğru ve verimli bir şekilde belirlenmesi, sürdürülebilir kalkınma ve planlama için büyük bir öneme sahiptir. Ortofotolar, karmaşık veri yapılarına sahip nokta bulutlarına göre daha uygun maliyetli, etkin, hızlı ve işleme süreci kolay olan verilerdir. Derin öğrenme, makine öğrenmesinin bir alt dalıdır ve yapay sinir ağlarını kullanarak daha karmaşık ve soyut temsil seviyelerine ulaşmayı hedefler. Derin öğrenme, özellikle büyük veri setleri ve güçlü hesaplama kaynakları ile etkili hale gelir. Bu yöntem, çok katmanlı sinir ağları (derin sinir ağları) kullanarak verilerdeki öznitelikleri ve desenleri otomatik olarak öğrenir ve çıkarır. Ortofotoların manuel olarak sınıflandırılması, zaman ve emek yoğunluğu, maliyet ve veri etiketlemedeki öznellik açısından dezavantajlara sahiptir. Derin öğrenme ile sınıflandırma, büyük veri setlerini hızlı ve tutarlı şekilde işleyebilmesi, karmaşık desenleri otomatik olarak öğrenebilmesi ve geniş çaplı projelere kolayca uyum sağlayabilmesi gibi avantajlar sunarken; çok sayıda etiketli veri gereksinimi, yüksek hesaplama maliyeti ve modelin karar alma süreçlerinin genellikle anlaşılabilir olması gibi dezavantajlar da taşır. Bu çalışmada Kırklareli'nin Kaynarca ilçesinde insansız hava aracı kullanılarak yapılan uçuş ile elde edilen hava fotoğrafları Pix4D yazılımı ile değerlendirilerek üretilen ortofoto veri seti kullanılmıştır. Görüntüler veri ön işleme sürecinde alt görüntülere ayrılmıştır. Veriler bina, yol, bitki örtüsü (ağaç/yeşil alan) ve arka plan olmak üzere 4 sınıfa ayrılarak etiketlenmiş ve sınıflandırılma gerçekleştirilmiştir. Sınıflandırma işlemi için DeepLabv3+, UNet++, SegFormer-MixVTB3 ve Attention UNet algoritmaları kullanılmıştır. Sonuçlar ilgili sınıflar karşılaştırılarak sunulmuştur. Genel olarak ortofoto görüntülerinin arazi kullanımı amacıyla sınıflandırılmasında DeepLabv3+ ve SegFormer-MixVTB3 algoritmalarının performansı diğer yöntemlere göre daha iyidir.

**Anahtar Sözcükler.** Arazi kullanımı, Attention UNet, DeepLabv3+, UNet++, SegFormer-MixVTB3, Ortofoto

## ABSTRACT

Land use defines how and for what purposes an area is used for human activities. Determining land use accurately and efficiently is of great importance for sustainable development and planning. Orthophotos are more cost-effective, efficient, fast and easy to process than point clouds with complex data structures. Deep learning is a sub-branch of machine learning and aims to reach more complex and abstract levels of representation using artificial neural networks. Deep learning becomes particularly effective with large data sets and powerful computational resources. This method automatically learns and extracts features and patterns from data using multilayer neural networks (deep neural networks). Manual classification of orthophotos has disadvantages in terms of time and labor intensity, cost, and subjectivity in data labeling. While deep learning classification has the advantages of being able to process large datasets quickly and consistently, automatically learn complex patterns, and easily adapt to large-scale projects, it also has the disadvantages of requiring a large amount of labeled data, high computational cost, and often incomprehensible model decision-making processes. In this study, the orthophoto dataset produced by evaluating the aerial photographs obtained from an unmanned aerial vehicle flight in Kaynarca district of Kırklareli with Pix4D software was used. The images were divided into sub-images in the data pre-processing process. The data were labeled and classified into 4 classes as building, road, vegetation (tree/green area) and background. DeepLabv3+, UNet++, SegFormer-MixVTB3 and Attention UNet algorithms were used for classification. The results are presented by comparing the relevant classes. In general, the performance of DeepLabv3+ and SegFormer-MixVTB3 algorithms is better than other methods in classifying orthophoto images for land use purposes.

**Keywords:** Land use, Attention UNet, DeepLabv3+, UNet++, SegFormer\_MixVTB3, Ortofoto

## 1. GİRİŞ

Arazi kullanımı, bir bölgedeki toprakların insan faaliyetleri için nasıl ve hangi amaçlarla kullanıldığını tanımlar. Arazi kullanımının doğru ve verimli bir şekilde belirlenmesi, sürdürülebilir kalkınma ve planlama için büyük bir öneme sahiptir. Arazi kullanımının belirlenmesinde uydu görüntüleri, hava fotoğrafları, insansız hava aracı (İHA) görüntüleri ve ortofoto gibi farklı görüntü çeşitleri kullanılmaktadır. Ortofotolar, hava fotoğraflarının fotogrametrik yöntemlerle geometrik olarak düzeltilmesiyle elde edilir ve bu sayede yüksek konum doğruluğu sunar. Ortofotolar, coğrafi bilgi sistemleri (CBS) ile entegre edilerek farklı veri setlerinin bir arada kullanılması ve analiz edilmesi açısından avantaj sağlar. Ortofotoların sınıflandırılması işlemi, arazi kullanımının doğru ve detaylı bir şekilde belirlenmesi için kritik bir adımdır. Bu süreç, operatör tarafından görsel kontrolle yapılabileceği gibi, makine öğrenmesi ve derin öğrenme yöntemleri kullanılarak otomatikleştirilebilir.

Zhou vd. (2019) çalışmalarında vurgulandığı gibi, bu yöntemler, özellikle uzman görüşü bilgisi gerektiren karmaşık veya nadir sınıflar için yüksek doğruluk sağlar. Uzmanlar, arazi örtüsünü ve kullanımlarını ince detaylarıyla tanımlayabilir; bu

da yerel bilgiyi ve sahaya özgü koşulları hesaba katmaya olanak tanır. Ancak, manuel sınıflandırma oldukça zaman alıcı ve iş gücü yoğunudur. Li vd. (2020) manuel sınıflandırmanın geniş ölçekli projeler için sürdürülebilir olmadığını ve öznel dayalı olarak farklı kişiler arasında tutarsızlıkların oluşabileceğini belirtmiştir.

1940'lı yıllarda başlayan yapay zekâ çalışmaları temel olarak insan becerilerinin farklı yollarla makinelere aktarımını hedefler. Makine öğrenmesi karmaşık veriyi kullanma, anlamsal bölütleme ve sınıflandırma için kuvvetli bir matematiksel araçtır. Makine öğrenmesi veriyi ve beklenen çıktıları kullanarak yeni verilere özgün cevaplar üretebilecek kuralları öğrenir. Bilgiyi verilen kurallar çerçevesinde işlemekten, bu bilgiyi nasıl işleyeceğine dair öğrenme görevlerini yerine getirmeyi, duruma özgü yeni kurallar keşfetmeyi hedefler. Makine öğrenmesi modelleri genel olarak girdi veriyi beklenen çıktılara dönüştürmeyi başaracak fonksiyon ve ilişkili parametrelerden oluşmaktadır. Bu fonksiyon ve parametre setlerinin seçilmesi, ilgili parametrelerin otomatik biçimde öğrenilmesinin sağlanması ana araştırma alanları olarak ön plana çıkmıştır (Goodfellow vd., 2015).

Derin öğrenme, makine öğrenmesinin bir alt dalıdır ve yapay sinir ağlarını kullanarak daha karmaşık ve soyut temsil seviyelerine ulaşmayı hedefler. Özellikle büyük veri setleri ve güçlü hesaplama kaynakları ile etkili hale gelir. Bu yöntem, çok katmanlı sinir ağları kullanarak verilerdeki öznelikleri ve örüntüleri otomatik olarak öğrenir ve çıkarır (LeCun, vd., 2015). Konvülsiyonel sinir ağları (KSA- Convolutional Neural Networks, CNN) gibi mimariler aracılığıyla, hava fotoğraflarının sınıflandırılmasında derin öğrenme yaygın olarak kullanılmaktadır. Chen vd. (2020) KSA tabanlı bir yaklaşım kullanarak tarım arazileri, kentsel alanlar ve ormanlar gibi farklı arazi kullanım türlerinin başarılı bir şekilde sınıflandırılabilirliğini göstermiştir. Araştırmalar, görüntülerin sınıflandırılmasında derin öğrenme modellerinin, geleneksel makine öğrenme yöntemlerine göre daha yüksek doğruluk oranlarına ulaştığını ve bu nedenle büyük ölçekli coğrafi analizlerde tercih edildiğini ortaya koymaktadır. Hu vd. (2020) tarafından yapılan bir çalışmada, Sentinel-2 uydu görüntülerinin sınıflandırılmasında derin öğrenme modelleri ile geleneksel makine öğrenme yöntemleri karşılaştırılmıştır. Bu çalışmada, özellikle KSA kullanılarak yapılan sınıflandırmaların doğruluğu, geleneksel yöntemlere kıyasla belirgin şekilde daha yüksek bulunmuştur. KSA modellerinin görüntülerdeki mekânsal öznelikleri daha iyi öğrenebildiği ve büyük veri setlerinde yüksek doğrulukla çalışabildiği vurgulanmıştır.

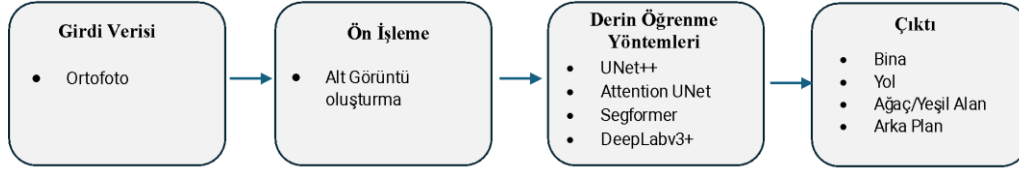
Derin öğrenme, yüksek doğruluk ve otomasyon kapasitesi ile dikkat çekmektedir. Liu vd. (2020) derin öğrenme modellerinin özellikle geniş veri setlerinde yüksek doğruluk sağladığını ve manuel işlemlerle karşılaştırıldığında daha tutarlı sonuçlar verdiğini göstermiştir. Derin öğrenme, karmaşık örüntüleri tanıma kabiliyeti ile manuel sınıflandırmanın zorluklarını aşabilir. Zhu vd. (2019) bu modellerin büyük miktarda etiketlenmiş veri gerektirdiğini ve bu verilerin elde edilmesinin maliyetli olabileceğini vurgulamıştır. Ayrıca, derin öğrenme modelleri yüksek hesaplama gücü gerektirir ve eğitim süreci zaman alıcı olabilir. Ancak, bu modeller geniş veri setlerinde yüksek doğruluk sağlayabilir. Zhu vd. (2021), derin öğrenmenin karmaşık örüntüleri tanıma yeteneği sayesinde, özellikle karmaşık arazi sınıfları ve değişken görüntü koşullarında üstün performans sergilediğini belirtmektedir. Ayrıca, derin öğrenme modelleri, ham veriden anlamlı öznelikleri otomatik olarak öğrenebilir. Kussul vd. (2019), bu özelliğin, farklı coğrafi bölgelerden elde edilen büyük veri setlerinde daha esnek ve güçlü bir çözüm sunduğunu vurgulamaktadır.

Çalışmanın amacı, ortofoto görüntülerden yararlanarak derin öğrenme yöntemlerini kullanmak suretiyle arazi kullanımının belirlenmesidir. Bu bağlamda, ortofoto görüntüler üzerinde uygulanan derin öğrenme algoritmaları, farklı arazi kullanım türlerini (örneğin, tarım, sanayi, yerleşim alanları ve ulaşım ağları) tanımlamak ve sınıflandırmak için kullanılacaktır. Elde edilen verilerin analizi, arazi kullanım değişimlerinin izlenmesi ve yönetimi açısından önemli bilgiler sağlayacaktır. Bu süreç, sürdürülebilir planlama ve çevre yönetimi çalışmalarına katkıda bulunmayı amaçlamaktadır.

Bu çalışmada, etiketlenen ortofoto görüntüleri kullanılarak arazi kullanımının belirlenmesinde farklı derin öğrenme yöntemlerinin (DeepLabv3+, UNet++, SegFormer-MixVTB3 ve Attention UNet algoritmaları) performansı karşılaştırılmıştır. Karşılaştırmada kesişim-birleşim oranı (Intersection over Union, IoU) ve F1 Skoru metrikleri dikkate alınmıştır. Genel olarak ortofoto görüntülerin arazi kullanımı amacıyla sınıflandırılmasında DeepLabv3+ ve Segformer-MixVTB3 algoritmalarının performansı diğer yöntemlere göre daha iyidir.

## 2. MATERYAL VE YÖNTEM

Bu çalışmada, veri seti olarak 1/1000 ölçekli ortofotolar kullanılmıştır. İlk olarak, ön işleme adımı ile veriler modellerin işleyebileceği uygun bir formata dönüştürülmüştür. Ön işleme yapılmış görüntüler etiketlenmiş ve UNet++, Attention UNet, Segformer\_MixVTB3 ve DeepLabv3+ gibi derin öğrenme yöntemleri uygulanarak ortofoto verileri üzerinde sınıf bazlı bölütleme yapılmıştır. Bu süreç sonucunda bina, yol, yeşil alan ve arka plan sınıflarına ait bölgeler belirlenerek çıktılar oluşturulmuştur. Şekil 1'de iş akış diyagramı gösterilmiştir.



Şekil 1. İş akış diyagramı

## 2.1 Veri Seti

Bu çalışmada, Kırklareli Kaynarca Belediyesi'nin talep ve yetkisi doğrultusunda İller Bankası Anonim Şirketi kontrolörlüğünde üretilmiş olan ortofotolar veri kaynağı olarak kullanılmıştır. Şekil 2'de çalışma alanı gösterilmiştir.



Şekil 2. Çalışma alanı

DJI Matrice 300 İHA ile Kaynarca beldesinde gerçekleştirilen uçuşlar, Pix4D yazılımıyla değerlendirilerek 5.45 cm yer örnekleme aralığı (GSD) ile yatay (XY) ve dikey (Z) eksenleri için sırasıyla 7 ve 9 mm karesel ortalama hata (RMSE) ile elde edilen ortofoto veri seti kullanılmıştır.

## 2.2 Ön İşleme

Veriler bina, yol, bitki örtüsü (ağaç/yeşil alan) ve arka plan olmak üzere 4 sınıfa ayrılarak Anylabeling yazılımında etiketlenmiştir. Etiketleme işlemini takiben, derin öğrenme ağlarının eğitimi ve testi için görüntüler sol üst köşelerinden itibaren 512x512 piksellik alt görüntülere ayrılmıştır ve toplam 1247 alt görüntü üretilmiştir. Bu görüntülerin %60'ı eğitim (747), %20'si doğrulama (250) ve %20'si test (250) verisi olarak kullanılmıştır. Görüntülere maksimum-minimum normalizasyonu uygulanmış olup eğitim setindeki görüntülere 45, 135, 225 ve 315 derecelik döndürme, 0,5 ve 0,2 ölçekleme ve kontrast germe işlemleri uygulanmıştır.

## 2.3 Derin Öğrenme Yöntemleri

### 2.3.1 UNet++

UNet mimarisi, tıbbi görüntü bölütlemesi gibi biyomedikal görüntü analizinde yaygın olarak kullanılan derin öğrenme modellerinden biridir (Ronneberger vd., 2015). Geleneksel UNet mimarisi, bir kodlayıcı-çözücü yapısına sahiptir. Kodlayıcı kısmı, girdi görüntüsünü giderek daha küçük ve daha yüksek düzeyde özelliklere ayırırken, çözücü kısmı bu özellikleri kullanarak orijinal görüntü boyutunda bir çıktı üretir. Kodlayıcı ve çözücü arasında, aynı seviyede olan katmanlar arasında atlamalı bağlantılar vardır; bu sayede orijinal çözünürlükteki bilgilerin bir kısmı korunur.

UNet++ mimarisi, UNet'in temel yapısını geliştirmektedir. UNet++'ta, atlamalı bağlantılar doğrudan değil, bir dizi yoğun katmandan geçirilerek yapılır. Bu sayede, ara katmanlarda daha fazla bilgi işlenir ve daha zengin özellikler çıkarılır. Her bir atlamalı bağlantı, birden fazla aşamalı yoğun katmandan geçer ve bu şekilde bilgi daha etkili bir şekilde taşınır.

### 2.3.2 Attention UNet

Attention UNet mimarisinde, klasik UNet yapısına ek olarak dikkat mekanizmaları entegre edilmiştir. Bu mekanizmalar, modelin belirli bölgelerdeki özelliklere daha fazla odaklanmasını sağlayarak, özellikle karmaşık ve zor bölütleme görevlerinde performansı artırmayı hedefler. Attention UNet, temel olarak UNet'in kodlayıcı-çözücü yapısını korur.

Ancak, dikkat mekanizmaları kod çözücü kısmına entegre edilmiştir. Bu mekanizmalar, her bir atlamalı bağlantı üzerine yerleştirilir ve böylece modelin, orijinal görüntüdeki hangi alanların daha önemli olduğuna karar vermesini sağlar. Dikkat mekanizması, öznitelik haritalarını dinamik olarak ağırlıklandırarak, yalnızca ilgili bölgeleri vurgular ve bu bölgelerin bölütlenmesinde daha yüksek doğruluk sağlar (Oktay vd., 2018).

### 2.3.3 SegFormer-MixVTB3

Yama birleştirme ve hiyerarşik dönüştürücülere sahip SegFormer-MixVTB3, çok seviyeli öznitelikleri çıkararak çeşitli çözünürlüklere uyarlanabilir hale getirir. SegFormer-MixVTB3, iki önemli bileşene dayanan verimli bir bölütleme çerçevesi sunar: Bunlar, birden fazla çözünürlükte öznitelikler üretmek için hiyerarşik bir dönüştürücü kodlayıcı ve nihai bölütleme sonucuna ulaşmak için bu çeşitli öznitelikleri entegre eden modern bir Tüm-ÇKA (All-MLP) kod çözücüdür. Başlangıçta görüntü tahmin görevinin hassasiyetini arttırmak için daha küçük  $4 \times 4$  yamalara bölünür. Bu yamalar daha sonra çeşitli çözünürlüklerde öznitelikler elde etmek için dönüştürücü kodlayıcı tarafından işlenir. Daha sonra Tüm-ÇKA kod çözücü, bölütleme maskesini küçültülmüş bir çözünürlükte tahmin etmek için bu özelliklerden yararlanır (Xie vd., 2021). Geliştirilmiş öznitelik temsili ve hesaplama verimliliği yama birleştirme ve kendi kendine dikkat stratejileri ile elde edilir. Karma-İBA (İleri Beslemeli Ağ, Mix-FFN) modülü, İBA içine  $3 \times 3$  konvolüsyon ekleyerek ve geleneksel konumsal kodlamayı atlayarak bölütleme doğruluğunu artırır. Çok Katmanlı Algılayıcı (ÇKA- Multi-Layer Perceptron, MLP) tabanlı bir kod çözücü kullanarak etkinliği korurken mimariyi basitleştirmek için kodlayıcının geniş alıcı alanından yararlanır. Bu tasarım, kodlayıcıdan gelen çok seviyeli çıktıların sorunsuz bir şekilde birleştirilmesini ve iyileştirilmesini sağlar, Böylece bölütleme parametrelerinin tahmin edilmesi sürecini kolaylaştırır (Xie vd., 2021).

### 2.3.4 DeepLabv3+

DeepLabv3+ derin öğrenme algoritması, semantik bölütleme görevleri için etkili bir modeldir ve özellikle arazi kullanım ve örtüsü analizi gibi yüksek çözünürlüklü görüntü işleme problemlerinde kullanılmaktadır. Model, geniş alanlardan bilgi toplamak için boşluklu (atrous) konvülsiyonu kullanarak görüntü çözünürlüğünü düşürmeden geniş bağlam bilgisini işler. Bunun yanı sıra, kodlayıcı-çözücü mimarisi sayesinde düşük çözünürlükte çıkarılan öznitelik haritaları yüksek çözünürlüğe dönüştürülerek ayrıntılı bölütleme maskeleri üretilir. DeepLabv3+'ün Boşluklu Mekânsal Piramit Havuzu (Atrous Spatial Pyramid Pooling, ASPP) modülü, farklı ölçeklerdeki öznitelikleri yakalayarak modelin hem büyük hem de küçük nesnelere ayırt etmesini sağlar (Chen vd., 2018). Bu özellikler, modelin arazi kullanımı ve örtüsü analizleri gibi çeşitli uzaktan algılama görevlerinde yüksek doğrulukta sonuçlar vermesini sağlamaktadır.

## 2.4 Model Değerlendirme Metrikleri

Seçilen modellerin eğitilmesi için PyTorch, doğruluk analizi işlemleri için ise scikit-learn kütüphanesi kullanılmıştır. Eğitimler Nvidia RTX 4090 ekran kartlı bir bilgisayarda yapılmış olup her bir yöntem için minimum parça sayısı (batch size) iki, Adam optimizasyon algoritması ve 0.0001 öğrenme hızı seçilmiştir. Derin öğrenme yöntemlerinde model değerlendirme ölçütleri, modelin performansını değerlendirmek için kullanılır. Bu çalışmada kullanılan metrikler aşağıda belirtilmiştir (Sokolova vd., 2006).

**Kesinlik (Precision):** Modelin pozitif olarak tahmin ettiği örneklerin ne kadarının gerçekten pozitif olduğunu ölçer.

$$Kesinlik = \frac{DP}{DP+YP} \quad (1)$$

**Duyarlılık (Recall):** Gerçek pozitif örneklerin ne kadarının doğru bir şekilde tahmin edildiğini gösterir.

$$Duyarlılık = \frac{DP}{DP+YN} \quad (2)$$

Burada DP, doğru pozitif, modelin pozitif olarak tahmin ettiği ve gerçekte de pozitif olan örneklerdir. DN, doğru negatif, modelin negatif olarak tahmin ettiği ve gerçekte de negatif olan örneklerdir. YP, yanlış pozitif, modelin pozitif olarak tahmin ettiği fakat gerçekte negatif olan örneklerdir. YN, yanlış negatif, modelin negatif olarak tahmin ettiği fakat gerçekte pozitif olan örneklerdir.

**F1 skoru:** Kesinlik ve duyarlılık arasında bir denge sağlar. İki metriğin harmonik ortalaması olarak hesaplanır ve dengesiz veri setlerinde daha iyi bir değerlendirme sağlar.

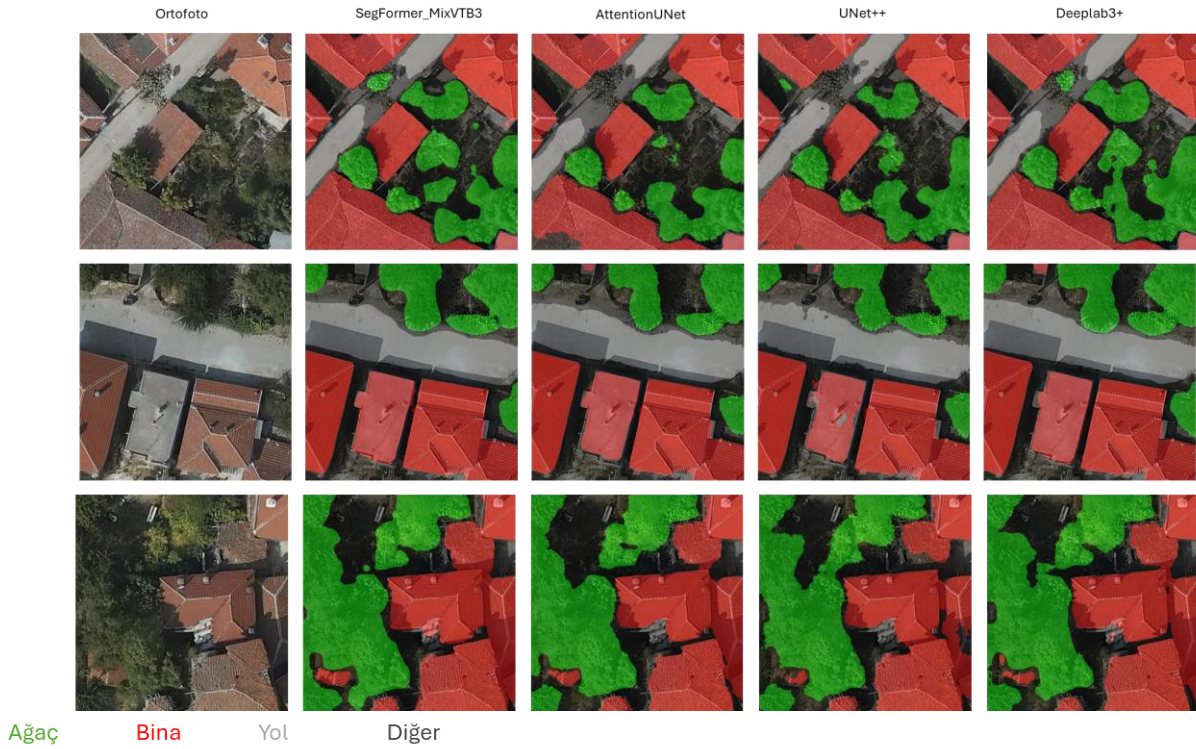
$$F1\ skoru = 2x \frac{Kesinlik \times Duyarlılık}{Kesinlik + Duyarlılık} \quad (3)$$

**Kesişim-Birleşim Oranı (Intersection over Union, IoU):** Bölütleme görevlerinde kullanılır ve tahmin edilen bölüt ile gerçek bölütün kesişimini toplam alana bölerek hesaplanır. Bölütlemenin ne kadar doğru olduğunu gösterir.

$$IoU = \frac{(Tahmin\ bölütü \cap Gerçek\ bölüt)}{(Tahmin\ bölütü \cup Gerçek\ bölüt)} \quad (4)$$

### 3. SONUÇLAR VE TARTIŞMA

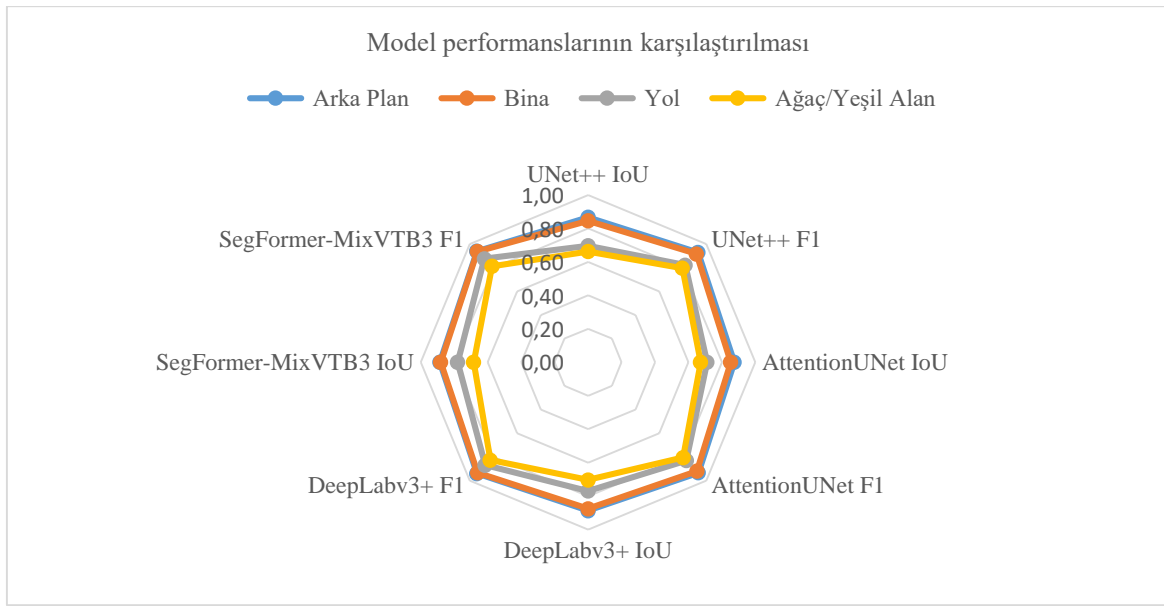
Bu çalışmada, derin öğrenme tabanlı çeşitli bölütleme algoritmaları (DeepLabv3+, UNet++, SegFormer-MixVTB3, ve Attention UNet) kullanılarak Kırklareli'nin Kaynarca ilçesinden elde edilen ortofoto görüntüleri üzerinde arazi kullanımının sınıflandırılması gerçekleştirilmiştir. Elde edilen sonuçlar, kesişim-birleşim oranı (Intersection over Union, IoU) ve F1 skoru gibi yaygın kullanılan değerlendirme metriklerine göre analiz edilmiştir. Görüntüler 512x512 piksellik alt görüntüye ayrılmıştır ve toplam 1247 alt görüntü üretilmiştir. Bu görüntülerin %60'ı eğitim (747), %20'si doğrulama (250) ve %20'si test (250) verisi olarak kullanılmıştır.



Şekil 3. Sınıflandırılmış görüntülerin karşılaştırılması

Çizelge 1. Değerlendirme metrikleri (IoU: Kesişim-birleşim oranı)

Sınıf	UNet++		AttentionUNet		DeepLabv3+		SegFormer-MixVTB3	
	IoU	F1 skor	IoU	F1 skor	IoU	F1 skor	IoU	F1 skor
Arka Plan	0,8670	0,9287	0,8737	0,9326	0,8863	0,9397	0,8860	0,9396
Bina	0,8457	0,9164	0,8521	0,9201	0,8747	0,9332	0,8805	0,9364
Yol	0,6965	0,8211	0,7103	0,8306	0,7685	0,8691	0,7795	0,8761
Ağaç/Yeşil alan	0,6616	0,7963	0,6725	0,8042	0,7031	0,8257	0,6834	0,8120
Ortalama	0,7677	0,8656	0,7772	0,8719	0,8082	0,8919	0,8074	0,8910



Şekil 4. Derin öğrenme yöntemleri ile elde edilen sınıflandırma doğruluk metriklerinin karşılaştırılması

### 3.1 F1 Skoru ve IoU Analizi

Ağların genel performansı incelendiğinde IoU ve F1 skorları, UNet++, Attention UNet, DeepLabv3+ ve Segformer-MixVTB3 için sırasıyla 0,7677 ve 0,8656; 0,7772 ve 0,8719; 0,8082 ve 0,8919; 0,8074 ve 0,8910 şeklinde olup DeepLabv3+'ın daha başarılı sonuçlara ulaştığı gözlemlenmiştir. Sınıf bazındaki en yüksek IoU ve F1 skoru'na ulaşan ağlar incelendiğinde ise arka plan sınıfı için DeepLabv3+'ın 0,8863 ve 0,9397, bina sınıfı için Segformer-MixVTB3'ün 0,8805 ve 0,9364, yol sınıfı için Segformer-MixVTB3'ün 0,7795 ve 0,8761, ağaç/yeşil alan sınıfı için DeepLabv3+'ın 0,7031 ve 0,8257 değerleri ile en yüksek sonuca ulaştığı görülmüştür. Genel olarak ortofoto görüntülerin arazi kullanımı amacıyla sınıflandırılmasında DeepLabv3+ ve Segformer-MixVTB3 algoritmalarının performansı en yüksektir (Çizelge 1 ve Şekil 4).

Bu çalışmada, SegFormer-MixVTB3 algoritmasının bina sınıfında 0,9364 F1 Skor ve 0,8805 IoU değeri ile en yüksek başarıyı elde ettiği gözlemlenirken, yeşil alan sınıfı tüm algoritmalarda en düşük sonuçlara sahip sınıf olmuştur.. Bu durum, bitki örtüsü gibi homojen olmayan sınıfların, diğer sınıflara kıyasla daha fazla çeşitlilik göstermesi ve bu yüzden daha zor sınıflandırılabilmesinden kaynaklanmaktadır. Ayrıca, bitki örtüsü gibi doğrudan belirgin sınırları olmayan sınıflar, derin öğrenme modellerinin nesne çıkarma mekanizmalarını zorlayabilir. Düşük IoU ve F1 skorları, bu sınıfın daha karmaşık örüntülere sahip olduğuna işaret etmektedir.

### 3.2 Modeller Arası Karşılaştırma

DeepLabv3+ ve SegFormer-MixVTB3 modelleri, genellikle geniş veri setlerinde ve karmaşık desenlerde yüksek doğruluk sunan sınıflandırma algoritmaları olarak öne çıkmaktadır. DeepLabv3+, görüntülerdeki detayları etkin bir şekilde çıkarabilen çok katmanlı yapısı sayesinde ve SegFormer\_MixVTB3, dönüştürücü tabanlı yapısı ve yama birleştirme mekanizması ile üstün performans sergilemiştir. Dönüştürücü mimarisinin, farklı çözünürlüklerdeki görüntülerde çalışma performansı, SegFormer-MixVTB3'ün başarılı bölütleme yapabilmesinin temel nedenidir.

Attention UNet modeli, çalışmada incelenen ortofoto görüntülerin sınıflandırılmasında, dikkat mekanizmalarının potansiyelini gösterememiştir. UNet++ modeli ise bu çalışmada kullanılan ortofoto görüntülerinin yüksek çözünürlüğü ve karmaşık sınıf yapısı gibi zorluklarıyla başa çıkmakta diğer modellere göre daha zayıf performans göstermiştir.

Bu çalışma, derin öğrenme tabanlı dört farklı modelin ortofoto görüntülerinin sınıflandırılması görevlerinde nasıl performans gösterdiğini karşılaştırmalı olarak ortaya koymaktadır. DeepLabv3+ ve SegFormer-MixVTB3, yüksek doğruluk oranları ve çeşitli sınıflarda sağladıkları üstün performans ile ön plana çıkmıştır. Bu iki model, geniş ölçekli veri setleri ve karmaşık arazi sınıflandırma görevlerinde tercih edilebilir. Attention UNet ve UNet++, diğer modellere göre daha düşük performans göstermiştir, ancak bu modellerin tıbbi görüntü analizi gibi diğer alanlarda daha etkili olabileceği göz önünde bulundurulmalıdır.

## KAYNAKLAR

- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H.** (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proceedings of the European Conference on Computer Vision (ECCV)*, 833–851.
- Chen, Y., Lin, Z., & Zhao, X.** (2020). Convolutional neural networks for airborne image classification: A case study. *International Journal of Remote Sensing*, 41(5), 1695-1712.
- Goodfellow, I. J., Bengio, Y., & Courville, A. C.** (2015). Deep learning. *Nature*, 521(7553), 436–444.
- Hu, F., Xia, G.-S., Hu, J., & Zhang, L.** (2020). A comparative study of deep learning and traditional methods for remote sensing image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5438–5450.
- Kussul, N., Lavreniuk, M., & Skakun, S.** (2019). Deep learning models for land use and land cover classification: A case study on airborne imagery. *Remote Sensing of Environment*, 237, 111466.
- LeCun, Y., Bengio, Y., & Hinton, G.** (2015). Deep learning. *Nature*, 521(7553), 436–444.
- Li, H., Yang, W., & Chen, X.** (2020). Challenges and opportunities in land use classification using manual and automated methods. *Remote Sensing*, 12(16), 2614.
- Liu, S., Zhang, Q., & Shen, Y.** (2020). Deep learning approaches for orthophoto classification: A review. *International Journal of Remote Sensing*, 41(16), 6055–6078.
- Oktay, O., Schlemper, J., Le Folgoc, L., Lee, M., Heinrich, M., Misawa, K., ... & Rueckert, D.** (2018). Attention U-Net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*.
- Sokolova, M., Japkowicz, N., & Szpakowicz, S.** (2006). Beyond accuracy, F-score, and ROC: A family of discriminant measures for performance evaluation. In *AI 2006: Advances in Artificial Intelligence: 19th Australian Joint Conference on Artificial Intelligence, Hobart, Australia, December 4-8, 2006. Proceedings, 19*, 1015–1021.
- Ronneberger, O., Fischer, P., & Brox, T.** (2015). U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III*, Springer, 234–241.
- Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P.** (2021). SegFormer: Simple and efficient design for semantic segmentation with transformers. *Advances in Neural Information Processing Systems*, 34, 12077–12090.
- Zhou, X., Li, J., & Wang, Y.** (2019). Manual vs. automated classification of land use in orthophotos: Comparative analysis. *Journal of Geospatial Information Science*, 22(2), 79–91.
- Zhu, X., Wang, Z., & Liu, J.** (2019). Deep learning in remote sensing: A comprehensive review and case study. *Geoscience and Remote Sensing Letters*, 16(3), 370–374.
- Zhu, T., Wang, Z., & Liu, J.** (2021). Data-driven land use classification using deep learning: Efficiency and accuracy. *Geospatial Analysis*, 52(2), 207–221.