

# VGGNet ve CBAM Dikkat Mekanizmasının Endoskopik Mesane Doku Görüntülerine Uygulanması

## Application of VGGNet and CBAM Attention Mechanism to Endoscopic Bladder Tissue Images

Birkan Büyükarıkan<sup>1\*</sup>

<sup>1</sup>Isparta Uygulamalı Bilimler Üniversitesi, Uluborlu Selahattin Karasoy Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Isparta, Türkiye.

\*Sorumlu Yazar:  
birkanbuyukarikan@isparta.edu.tr

Geliş/Received: 12.11.2023;  
Kabul/Accepted: 01.02.2024

Atıf/Citation: Büyükarıkan, B. (2024). VGGNet ve CBAM Dikkat Mekanizmasının Endoskopik Mesane Doku Görüntülerine Uygulanması. *UMBD*, 7(1), 38-47

### Öz

Tıbbi hastalıkların tespiti, tanısı ve izlenmesi amacıyla gerçekleştirilecek görevlerde, ilgili bilgilerin öne çıkartılması ve ilgisiz bilgilerin bastırılmasında Evrişimsel sinir ağları (ESA) yaygın olarak kullanılmaktadır. Ancak ESA modellerinin hesaplama karmaşıklığı, özellik kalitesi sorunu ve artan özellik boyutu gibi nedenler hastalığın tespit performansını zorlamaktadır. Son zamanlarda, ESA modellerinin performansını artırmak için dikkat mekanizmaları kullanılmaktadır, bu da sorunların üstesinden gelmeye yardımcı olmaktadır. Evrişimsel Blok Dikkat Modülü (CBAM) dikkat mekanizması, içerisinde barındırdığı modüller yardımıyla ilgili karmaşık özellikleri çıkarmak için geliştirilmiş bir yöntemdir. ESA modeliyle bu mekanizmanın birleştirilmesi, modelin performansını önemli ölçüde iyileştirilir. Çalışmanın amacı VGGNet ve CBAM dikkat mekanizmasını birleştirerek oluşturulan modellerle mesane tümörünün sınıflandırılmasıdır. Çalışmada VGGNet ve VGGNet+CBAM modellerinin performanslarını karşılaştırmak için doğruluk, kesinlik, duyarlılık, F1-skor ve eğri altında kalan alan (AUC) metrikleri kullanılarak deneyler gerçekleştirilmiştir. Sonuçlara göre VGG19+CBAM modeli, doğruluk, kesinlik, duyarlılık, F1-skor ve AUC ölçütleri açısından en yüksek performans değerlerini göstermiştir. Bu modelin doğruluk, kesinlik, duyarlılık, F1-skor ve AUC ölçütlerinin değerleri sırasıyla 0,990, 0,992, 0,984, 0,986 ve 0,994'tür. VGGNet+CBAM modelleri, VGGNet modellerinden daha iyi performans göstermiştir. Elde edilen performans değerlerine dayanarak, önerilen yaklaşımın mesane tümörü teşhisinde etkili olduğu görülmektedir.

**Anahtar Kelimeler:** CBAM dikkat mekanizması, Endoskopik mesane dokusu, Evrişimsel sinir ağı

### Abstract

Convolutional neural networks (CNNs) are widely used to highlight relevant information and suppress irrelevant information in tasks to be performed for detecting, diagnosing, and monitoring medical diseases. However, reasons such as computational complexity, feature quality problems, and increasing feature size challenge the disease detection performance of CNN models. In recent times, attention mechanisms have been used to enhance the performance of CNN models, helping to overcome challenges. The Convolutional Block Attention Module (CBAM) attention mechanism is a method developed to extract relevant complex features with the help of the modules it contains. The integration of this mechanism with the CNN model significantly improves the performance of the model. This study aims to classify bladder tissue with models created by combining VGGNet and CBAM attention mechanism. Experiments were carried out using accuracy, precision, recall, F1-score, and area under the curve metrics (AUC) to compare the performances of VGGNet and VGGNet+CBAM models. According to the results, the VGG19+CBAM model has demonstrated the highest performance values in terms of accuracy, precision, recall, F1-score, and AUC criteria. The accuracy, precision, recall, F1-score, and AUC values of this model are 0.990, 0.992, 0.984, 0.986, and 0.994, respectively. VGGNet+CBAM models have shown better performance than VGGNet models. Based on the achieved performance values, it is evident that the proposed approach is effective in the diagnosis of bladder tumors.

**Keywords:** CBAM attention mechanism, Endoscopic bladder tissue, Convolutional neural network

## 1. Giriş

Dünyada yaygın olarak görülen kanser türlerinden biri de mesane kanseridir. Bu kanser türünde farklı lezyonların görülmesi, çok sayıdaki insanda teşhisin yapılmasını zorlaştırır. Bu durum hastalığın ilerlemesine yol açar (Siegel vd., 2021). Genellikle mesane kanserinin teşhisi sistoskopi görüntüleme tekniğiyle yapılır (DeGeorge vd., 2017). Geleneksel sistoskopi tekniğiyle yapılan muayenede küçük ve düz lezyonlar gözden kaçabilir. Ayrıca bu lezyonların tanısı uzman deneyimine de bağlıdır (Inoue vd., 2012). Son yıllarda geleneksel teşhis yöntemlerinin yerine görüntüleri analiz etme, yanlış teşhis riskini azaltma ve tanı doğrulamayı artırma konusunda evrimsel sinir ağlarının (ESA) tıbbi uygulamalarda kullanılması, sağlık alanında önemli bir etki yaratmıştır (Gao vd., 2018; Langlotz vd., 2019; Mukhtov vd., 2023). Dolayısıyla ESA yardımıyla sistoskopik görüntülerin objektif olarak değerlendirilmesi mesane kanserinin tanı ve tedavisinin doğruluğunun iyileştirilmesine katkı sağlayacaktır (Ikeda vd., 2020).

Mesane doku görüntülerini ESA ile sınıflandıran araştırmalarda, bilinen ESA modelleri kullanıldığı gibi ince ayarlı ve yeni önerilen modellerde kullanılmıştır. Shkolyar vd. (2019) sistoskopik görüntülerde tümör tespiti için kullandıkları derin öğrenme modeliyle %90,6'lık duyarlılık ve %98,6'lık özgüllük gösterdiğini belirtmişlerdir. Ikeda vd., (2020) çalışmalarında tümörlü ve normal görüntüyü ince ayarlı GoogLeNet modelini kullanarak sınıflandırmışlardır. Çalışmalarında sınıfların kolaylıkla ayrılabilirdiğini gösteren ROC ölçütünü 0,98 olarak bulmuşlardır. Yang vd., (2021) kanserli ve kansersiz sistoskopi görüntülerini LeNet, AlexNet, GoogLeNet ve EasyDL modelleriyle sınıflandırmışlar. Onlar en yüksek doğruluğu EasyDL modelinde %96,9 oranıyla elde etmişlerdir. Başka bir çalışmada Du vd., (2021) 175 hastadan topladıkları mesane görüntülerini sınıflandırmak için geliştirdikleri EasyDL modelini kullanmışlar ve geliştirdikleri modelin etkinliğini araştırmışlardır. Tüm bu çalışmalarda mesane görüntüleri, kanserli ve kansersiz olacak şekilde kategorize edilmiş ve bu çalışmalarda ikili sınıflandırma görevi gerçekleştirilmiştir.

Lazo vd., (2023) çalışmalarında mesane görüntülerini arttırmak için yarı tabanlı Çekişmeli üretici ağ (Generative Adversarial Networks, GAN) tekniğini kullanmışlardır. Arttırılmış bu yeni veri setinde önceden eğitilmiş ResNet-101 ve SeCSi-GAN modellerinde doğruluk değerlerini sırasıyla 0,865 ve 0,905 olarak hesaplamışlardır. Çalışmalarında mesane doku görüntülerini Yüksek Dereceli Karsinom (High-Grade Carcinoma, HGC), Düşük Dereceli Karsinom (Low-Grade Carcinoma, LGC), Tümör Lezyonu Yok (No Tumor Lesion, NTL) ve Şüpheli Olmayan Doku (Non-Suspicious Tissue, NST) olarak kategorize eden Endoscopic Bladder Tissue Classification veri setini kullanmışlardır. Özetlenen çalışmalardan açıkça görülüyor ki araştırmacılar, mesane tümörünün etkili bir şekilde tespit edebilmek ve tespit güvenilirliğini arttırabilmek için halen daha bir model arayışı içerisindeylerdir. Nitekim modellerin hesaplama karmaşıklığı, özellik kalitesi ve artan özellik boyut sorunları tümör tespit performansını sınırlayabilir. Bu sınırlılığın üstesinden gelebilmek için yakın zamandaki sağlık alanında yapılan araştırmalarda Evrimsel Blok Dikkat Modülü (CBAM) mekanizması kullanılmaya başladığı görülmektedir.

CBAM mekanizması, modelin performansını arttırmak amacıyla kullanılan bir yöntemdir. CBAM içerisinde barındırdığı kanal ve uzamsal dikkat modelleri yardımıyla görüntünün özellik bilgisine ve ilgili görüntüdeki hastalığın konumuna odaklanır. Bu mekanizma, görüntülerden hastalıkların sınıflandırılması işlemi için yüksek hassasiyet gösterir. Ayrıca CBAM mekanizmasının içerdiği kanallar sayesinde zaman tasarrufu da sağlanır (Woo vd., 2018; Chao vd., 2019).

Mesane kanserinde tümör tespiti için önceden yapılan çalışmalara rağmen, hala geliştirilebilecek önemli bir alan bulunmaktadır. ESA modeliyle dikkat mekanizmasının kullanılması yukarıda belirten dezavantajları azaltabilir. Dolayısıyla ESA ve CBAM mekanizmasıyla geliştirilen modeller, uzmanlara olumlu yardımların yanı sıra bilinçli karar desteği için büyük fırsatlar sunar. Özetle bu mekanizma, ESA modellerinin ilgili tümör bölgesine odaklanarak tümör tespitinin doğruluğunun iyileştirilmesinde katkı sağlar.

Çalışmanın amacı, VGGNet ve CBAM dikkat mekanizmasının birleştirilerek oluşturulan bir evrimsel sinir ağ modeliyle mesane tümörünün sınıflandırılmasıdır. Deney, önceden eğitilmiş VGG16 ve VGG19 modelleriyle gerçekleştirilmiştir. Çalışmada Endoscopic Bladder Tissue Classification veri seti kullanılmıştır.

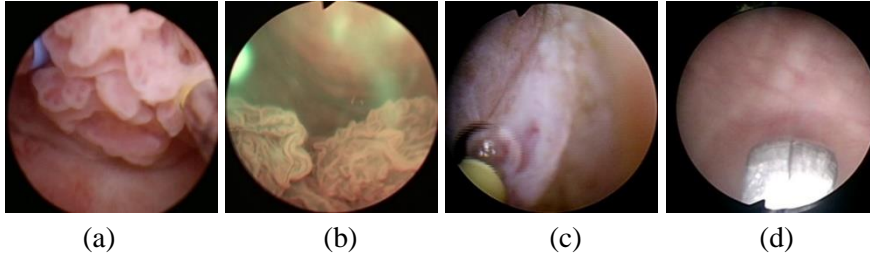
Bu veri setinde görüntüler HGC, LGC, NTL ve NST şeklinde kategorizeye sahiptir. Veri setindeki görüntülerin %80'i eğitim verisi ve %20'si test verisi olacak şekilde rastgele olarak ayrılmıştır. Önerilen ESA modelleri beş kez çalıştırılmış ve modellerin performans sonuçlarının ortalama ile standart sapma değerleri verilmiştir. Bu çalışmanın katkıları şu şekilde özetlenebilir: Çalışmada hastalıklı bölgenin özellik bilgisine odaklanmak için yenilikçi bir şekilde, VGGNet ve CBAM dikkat mekanizmasını birleştirilmesine dayanan bir yaklaşım önerilmiş ve bu yaklaşımın etkinliği gösterilmiştir. Ayrıca modellerin nispeten daha düşük öğrenilebilir parametreleri sağlamak ve aktarım öğrenmeyle fazladan ince ayar yapmamak için katmanlar dondurularak kullanılmıştır.

Çalışmanın diğer bölümleri şu şekilde organize edilmiştir. İkinci bölümde; endoskopik mesane görüntüleri içeren veri seti, VGGNet modeli, CBAM dikkat mekanizması, önerilen yaklaşım ve değerlendirme ölçütleri tanıtılmıştır. Üçüncü bölümde; deneysel sonuçlar ve sonuçların literatürdeki diğer çalışmalarla karşılaştırılması açıklanmıştır. Son bölümde ise çalışmanın sonuçları sunulmuştur.

## 2. Materyal ve Yöntem

### 2.1. Veri seti

Çalışmada açık ve ücretsiz olan Endoscopic Bladder Tissue Classification veri seti kullanılmıştır. Bu veri seti mesane ve üst idrar yolundan alınan görüntüleri içerir. Veri seti beyaz ışık görüntüleme (White Light Imaging, WLI) ve dar bant görüntüleme (Narrow Band Imaging, NBI) yöntemlerini kullanarak sistoskopi ve/veya üreteroskopi teknikleri yardımıyla 24 hastanın video görüntülerinden kareler çıkartılarak oluşturulmuştur. Mesane doku görüntüleri; HGC, LGC, NST ve NTL şeklinde dört sınıfta kategorize edilmiştir (Lazo vd., 2021; Lazo vd. 2023). Şekil 1'de veri setinden örnek bazı görüntüler verilmiştir.



Şekil 1. Mesane doku görüntüleri, (a) HGC sınıfı, (b) LGC sınıfı, (c) NTL sınıfı, (d) NST sınıfı

Tablo 1'de Endoscopic Bladder Tissue Classification veri setinin detayları ve her sınıfta bulunan görüntü sayıları listelenmiştir. Veri setindeki görüntüler RGB renk uzayında ve 350 x 350 piksel boyutlarındadır. Veri seti toplam 1754 görüntüden oluşmaktadır. HGC sınıfında 469, LGC sınıfında 647, NST sınıfında 504 ve NTL sınıfında 134 görüntü bulunmaktadır.

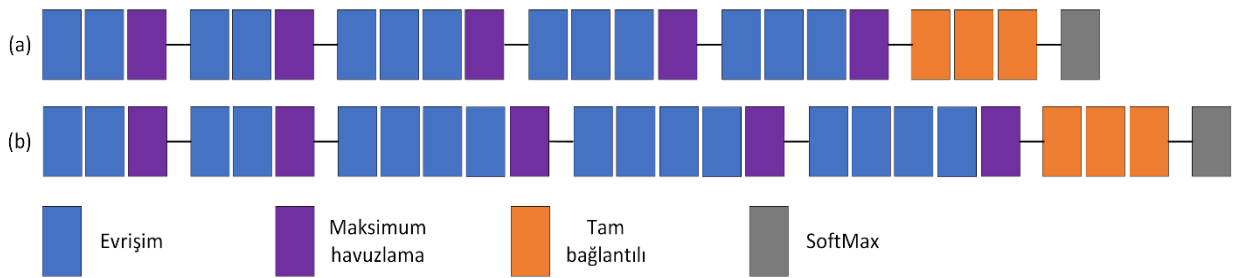
Tablo 1. Veri seti detayları (Lazo vd., 2023)

Sınıf/Görüntüleme yöntemleri	WLI	NBI	Toplam
HGC	386	64	469
LGC	454	145	647
NST	439	75	504
NTL	97	37	134
Toplam	<b>1433</b>	<b>321</b>	<b>1754</b>

## 2.2. VGGNet

Bu bölümde VGG16 ve VGG19 mimarilerinin giriş görüntü boyutundan, evrişim, havuzlama ve tam bağlantılı katmanlarının genel yapısından bahsedilmektedir. Şekil 2’de VGG16 ve VGG19 mimarilerinin şematik diyagramı gösterilmiştir. Şekil 2’ye göre; mavi dikdörtgen yerel özellikleri çıkararak evrişim katmanını, mor dikdörtgen görüntü boyutunu azaltan ve süreci hızlandıran havuzlama katmanını, turuncu dikdörtgen özellikleri olasılıklara göre seçen tam bağlantılı katmanı ve gri dikdörtgen ise sınıflandırma işlemini gerçekleştiren softmax sınıflandırıcı fonksiyonunu gösterir.

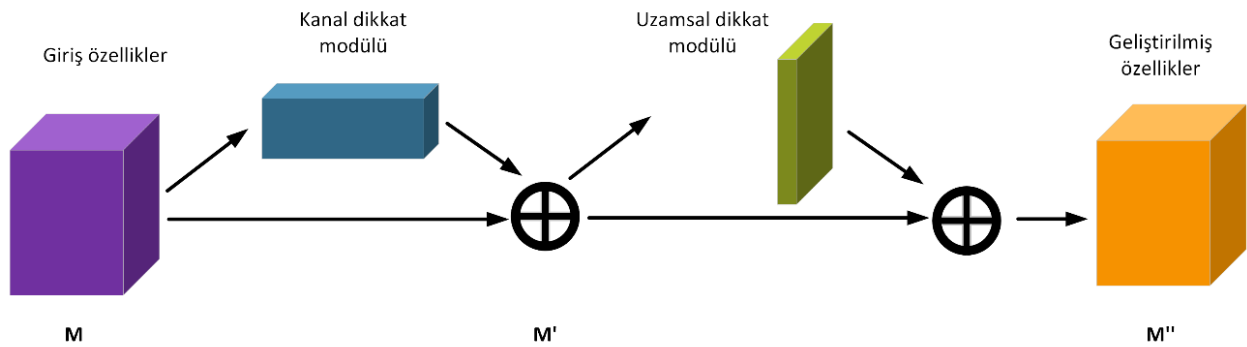
VGG16, giriş boyutu 224 x 224 piksel boyutuna sahip, on üç evrişim, beş havuzlama ve üç tam bağlantılı katmanlardan oluşan bir ağ yapısına sahip modeldir. Bu modelin tam bağlantı katmanları sırasıyla 4096, 4096 ve 1000 nöronu içerir (Sertkaya vd., 2019). VGG19 modelinin hesaplama parametresi VGG16’ya göre daha fazladır ve giriş boyutu 224 x 224 piksel boyutunda olan daha derin bir modeldir. Bu model, on altı evrişim, beş havuzlama ve üç tam bağlantılı katmanı içerir. Bu modelin tam bağlantılı katmanlarının yapısı VGG16 ile aynıdır (Mateen vd., 2018).



Şekil 2. ESA modellerinin şematik diyagramı, (a) VGG16, (b) VGG19

## 2.3. CBAM Dikkat Mekanizması

Büyük boyutlu ESA modelleri, birbirinden farklı veya benzer veri kümeleri üzerinde, gereksiz ve alakasız özellikleri çıkarabilirler. Dolayısıyla bu durum, hesaplama açısından verimli bir süreci engeller. Bu sorunu çözebilmek amacıyla Woo vd., (2018) kanal ve uzamsal dikkat modüllerini içeren CBAM mekanizmasını önermişlerdir. CBAM dikkat mekanizması, ileri beslemeli evrişimli bir ağ modülüdür ve bu mekanizma iki sıralı alt modülü içerir. Görüntülerden çıkarılan özellikler bu iki alt modülde incelenir. CBAM dikkat mekanizmasının temel yapısı Şekil 3’te görülmektedir (Wang vd., 2019; Alirezazadeh vd., 2023).



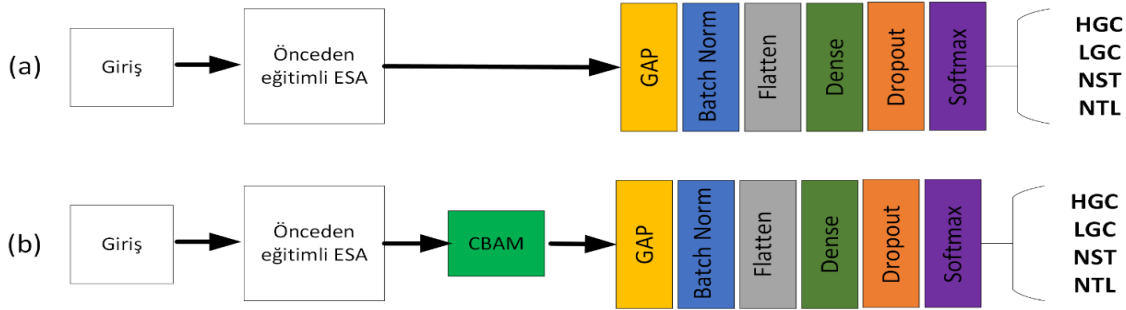
Şekil 3. CBAM dikkat mekanizması

Kanal dikkat modülü, sinir ağını özellik haritalarında önemli bölgelere yönlendirmek için kullanılır. Diğer bir ifadeyle önemli kanalları vurgulamak için kanalların önemliliklerine göre ağırlıkları uygulamaya çalışır. Ayrıca görüntülerden çıkarılan özellik haritası bu modülün yardımıyla sıkıştırılır. Uzamsal dikkat modülü ilgili özellikleri seçerken, yüksek belirginliğe sahip nesneyi vurgular. Bu modül, kanal modülüne benzer bir çıktı havuzu kullanır. Elde edilen değerleri de evrişim katmanına aktarır (Woo vd., 2018). Tüm bu bilgiler ışığında kanal dikkat modülü girdi görüntüsündeki “neyin” anlamlı olduğuna ve uzamsal dikkat modülü, kanal dikkat

modülünü yönlendirerek tamamlayıcı bilginin “nerede” olduğuna odaklanır (Shirasuna ve Gradwohl, 2023). Özetle CBAM dikkat mekanizması ile görüntülerde önemli alanların seçilmesine yardımcı olarak performansa katkı sağlar (Woo vd., 2018; Toğaçar vd., 2020).

## 2.4. Önerilen Yöntem

Çalışmada modellerin eğitimi için VGGNet (Şekil 4a) ve VGGNet+CBAM (Şekil 4b) olmak üzere iki yaklaşım kullanılmıştır. Her iki yaklaşımda da aktarım öğrenme yöntemi benimsenmiştir. İlk yaklaşımda VGGNet modellerinin son evrişim katmanına kadar tüm katmanlar dondurulmuş (modellerin evrişimli ve maksimum havuzlama katmanları, ağırlıkları değiştirilemeyecek şekilde dondurulmuş) ve ardından Global Average Pooling (GAP), Batch Norm, Düzgünleştirme (Flatten), Dense (1024 nöron), 0,5 değerinde Seyreltme (Dropout) ve sınıflandırma (softmax sınıflandırıcısı) katmanları eklenerek yeni modeller oluşturulmuştur. İkinci yaklaşımda dondurulan katmanlara ek olarak CBAM dikkat mekanizması ve ardından Global Average Pooling (GAP), Batch Norm, Düzgünleştirme (Flatten), Dense (1024 nöron), 0,5 değerinde Seyreltme (Dropout) ve sınıflandırma (softmax sınıflandırıcısı) katmanları eklenerek yeni modeller geliştirilmiştir (Alirezazadeh vd., 2023). Şekil 4’te modellerin yapıları görülmektedir.



Şekil 4. Sınıflandırıcı modellerin genel tasarımı

Veri setindeki görüntülere herhangi bir veri büyütme işlemi uygulanmamıştır. Deneylerde tüm görüntüler VGGNet modellerinin giriş görüntü boyutlarına uyarlanacak şekilde yeniden boyutlandırılmıştır. Daha sonra her bir görüntü 1/255 oranında ölçeklendirilmiştir. Deneylerde veri setindeki görüntüler %80 eğitim ve %20’si test verisi olarak rastgele ayrılmıştır. Çalışmada ESA modellerinin eğitiminde genellikle tercih edilen parametreler ve değerler kullanılmıştır. Eğitim sürecinde ESA mimarilerinin yığın boyutu 16 olarak seçilmiştir. Tüm modeller 100 devirde eğitilmiştir. Öğrenme oranı olarak 0,01 ve momentum değeri 0,9 değerleri seçilmiştir. Optimize edici algoritma olarak Stokastik Gradyan İnişi ve modeller için kayıp fonksiyonu olarak kategorik çapraz entropi (Categorical Cross-Entropy) tekniği kullanılmıştır. Her bir ESA modeli beş defa çalıştırılmıştır.

## 2.5. Performans Ölçütleri

Bu çalışmada modellerin performanslarının değerlendirilmesi için karışıklık matrisi kullanılmıştır. Bu matrise göre gerçek ve tahmin edilen sınıflar arasında doğru pozitif (DP), doğru negatif (DN), yanlış pozitif (YP) ve yanlış negatif (YN) olmak üzere dört gösterge oluşur. Bu dört göstergenin aralığı 0 ile 1 değerleri arasındadır. Bu göstergelere göre doğruluk, kesinlik, duyarlılık, F1-skoru ve eğri altında kalan alan (AUC) ölçütleri Denklem 1-5’teki gibi hesaplanır.

$$\text{doğruluk} = \frac{DP+DN}{DP+DN+YP+YN} \quad (1)$$

$$\text{kesinlik} = \frac{DP}{DP+YP} \quad (2)$$

$$\text{duyarlılık} = \frac{DP}{DP+YN} \quad (3)$$

$$F1 - \text{skor} = 2 * \frac{\text{kesinlik} * \text{duyarlılık}}{\text{kesinlik} + \text{duyarlılık}} \quad (4)$$

$$AUC = \frac{1}{2} * \frac{DP}{DP+YN} + \frac{DN}{DN+YP} \quad (5)$$

Çalışmalarda doğruluk, F1-skor ve AUC ölçütleri ne kadar yüksek olursa sınıflandırma amacıyla kullanılan modellerin performansları da o kadar iyidir. Buna ek olarak kesinlik ve duyarlılık ölçütleri birbirlerine ne kadar yakın olursa sınıflandırma için kullanılan modelin performansı o kadar dengeli olur (Zhou vd., 2022).

### 3. Deneysel Sonuçlar

Deneysel çalışmada tüm modeller aktarım öğrenme yaklaşıma göre eğitilmiş ve Google Colaboratory platformunda derlenmiştir. Deneysel olarak kullanılan modellerin parametre ve değerleri Bölüm 2.4'te açıklanmıştır.

#### 3.1. ESA Modelleriyle Mesane Dokularını Sınıflandırma

Endoskopik mesane görüntülerinin sınıflandırılması için VGG16 ve VGG19 modelleri kullanılmıştır. Performans sonuçlarının tutarlılığı için her bir ESA modeli beş defa çalıştırılmıştır. ESA modellerinin beş defa çalıştırılması sonucunda oluşan performans değerlerinin ortalama sonuçları Tablo 2'de yer almaktadır. Buna göre VGG19 neredeyse tüm ölçütlerde VGG16'ya göre daha yüksek performans sergilemiştir.

**Tablo 2.** ESA modellerin ortalama performans sonuçları (Ortalama ± Standard sapma), en yüksek değerler koyu fonttadır

Ölçüt / Model	VGG16	VGG19
Doğruluk (Eğitim)	0,996±0,002	<b>0,997±0,002</b>
Doğruluk (Test)	0,984±0,004	<b>0,986±0,002</b>
Kesinlik	<b>0,984±0,009</b>	0,978±0,011
Duyarlılık	0,964±0,009	<b>0,978±0,008</b>
F1-skor	<b>0,974±0,009</b>	<b>0,974±0,005</b>
AUC	0,989±0,003	<b>0,991±0,001</b>

#### 3.2. ESA+CBAM Modelleriyle Mesane Dokularını Sınıflandırma

VGG16 ve VGG19 modellerinin CBAM dikkat mekanizmasıyla birleştirilmesiyle oluşturulan VGG16+CBAM ve VGG19+CBAM modellerin beş defa çalıştırılmasıyla elde edilen ortalama performans sonuçları Tablo 3'te sunulmaktadır. Test sonuçlarına göre VGG19+CBAM modeli tüm performans ölçütlerinde ortalama olarak daha iyi performans sergilemiştir.

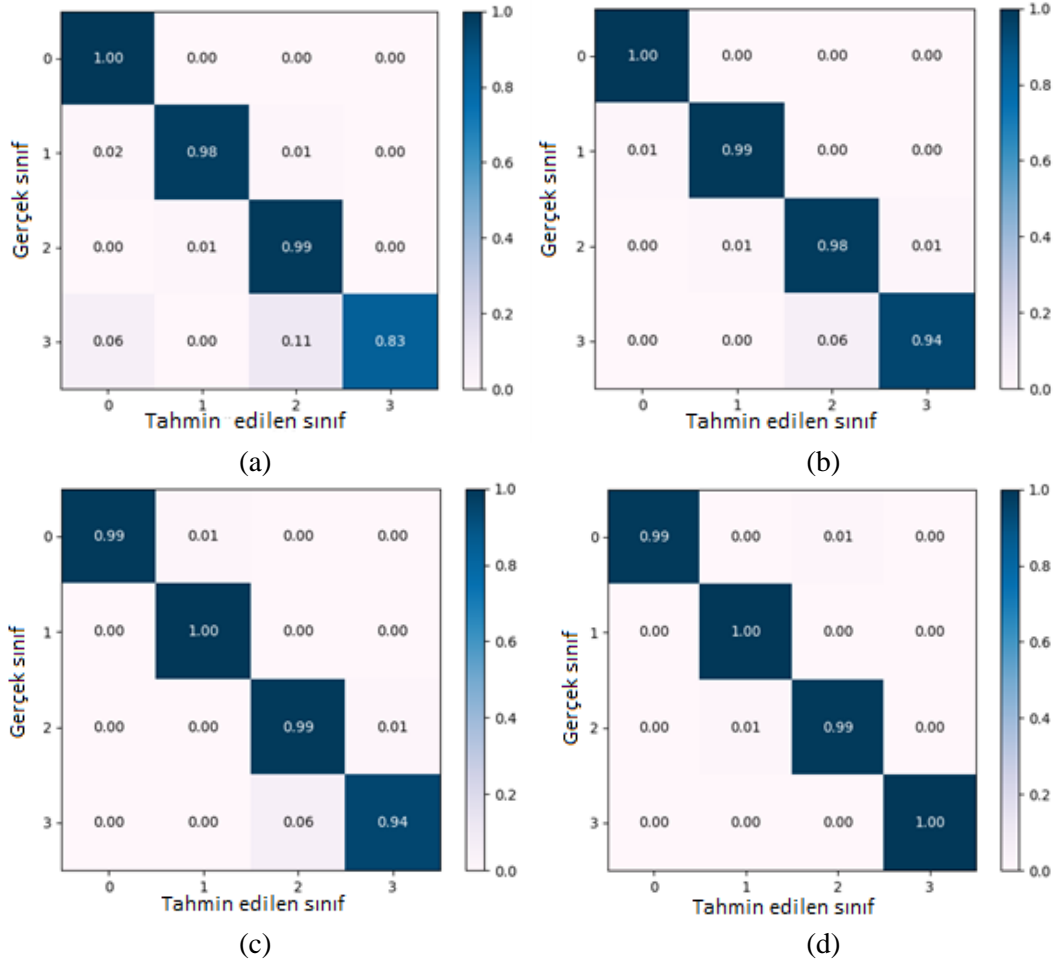
**Tablo 3.** ESA+CBAM modellerin ortalama performans sonuçları (Ortalama  $\pm$  Standard sapma), en yüksek değerler koyu fonttadır

Ölçüt / Model	VGG16+CBAM	VGG19+CBAM
Doğruluk (Eğitim)	<b>0,997<math>\pm</math>0,002</b>	0,995 $\pm$ 0,002
Doğruluk (Test)	0,987 $\pm$ 0,003	<b>0,990<math>\pm</math>0,003</b>
Kesinlik	0,976 $\pm$ 0,005	<b>0,992<math>\pm</math>0,004</b>
Duyarlılık	0,976 $\pm$ 0,005	<b>0,984<math>\pm</math>0,009</b>
F1-skor	0,974 $\pm$ 0,009	<b>0,986<math>\pm</math>0,011</b>
AUC	0,992 $\pm$ 0,002	<b>0,994<math>\pm</math>0,002</b>

### 3.3. Tartışma

#### 3.3.1. Modellerin performans sonuçlarının karşılaştırılması

Deneylerde en iyi performans sonucunu gösteren modellerin karışıklık matrisleri, Şekil 5'te sunulmuştur. Matrislerde Sınıf 0 HGC'yi, Sınıf 1 LGC'yi, Sınıf 2 NST'yi ve Sınıf 3 ise NTL kategorilerini ifade etmektedir. Elde edilen sonuç matrislerine göre Sınıf 3'ün tahmin edilme durumu, diğer sınıflara göre genellikle düşüktür. Buna karşın VGG19+CBAM modelinde Sınıf 3'ün tahminin iyi olduğu görülmektedir.



**Şekil 5.** ESA modellerin karışıklık matrisleri: (a) VGG16 (Deney 2), (b) VGG19 (Deney 3), (c) VGG16+CBAM (Deney 2), (d) VGG19+CBAM (Deney 1)

VGG19+CBAM (Tablo 3) modelinin doğruluk, F1-skor ve AUC ölçütlerinin değerleri diğer modellere göre daha yüksektir. Diğer bir ifadeyle mesane doku görüntülerini ayırmada kullanılan bu model diğer modellere göre daha iyi performans göstermektedir. Tablo 4'te VGGNet ve VGGNet+CBAM ile oluşturulan modellerin ortalama doğruluklarının artış performansları listelenmiştir. Buna göre VGG16+CBAM modelinin VGG16

modeline göre artış oranı 0,002'dir. VGG19+CBAM modelinin VGG19 modeline göre artış oranı 0,003'tür. CBAM mekanizmasının eklenmesinden sonra en yüksek artış VGG19'da olduğu görülmektedir.

**Tablo 4.** VGGNet ve VGGNet+CBAM modellerinin karşılaştırılması, en yüksek değerler koyu fonttadır

Model	Test doğruluğu	Gelişme değeri
VGG16	0,984	
VGG16+CBAM	0,986	+0,002
VGG19	0,987	
VGG19+CBAM	0,990	<b>+0,003</b>

### 3.3.2. Literatürdeki diğer çalışmalarla performans değerlerinin karşılaştırılması

Bu alt başlıkta, literatürdeki çeşitli veri setleri üzerinde gerçekleştirilen mesane tümör görüntülerini sınıflandıran araştırmaların sonuçları gözden geçirilmiştir. Tablo 5, farklı veri setlerindeki en iyi sınıflandırma performansını gösteren modellerin değerlendirme sonuçlarını açıklamaktadır. Mesane tümör görüntülerini sınıflandıran bu çalışmalarda çeşitli derin öğrenme modelleri kullanılmıştır. Bu modellerin ağ yapıları birbirinden farklıdır ve çalışmalarda kullanılan veri setlerindeki görüntü sayıları da büyük ölçüde değişiklik göstermektedir. Bu nedenle modelin ağ yapısı ve görüntülerin sayısı sınıflandırma sonuçlarını etkileyebilir. Bu durum, literatürdeki diğer çalışmalarla adil bir karşılaştırma ve yorumlama yapmayı zorlaştırabilir. Bu bağlamda, karşılaştırmalar aynı veri setini kullanan araştırmaların modelleri açısından incelenmiştir.

Aynı veri setini kullanan Lazo vd., (2023) çalışmalarında, orijinal görüntülerin sayısını veri büyütme yöntemiyle arttırmış ve SeCSi-GAN modeliyle yeni mesane görüntülerini sınıflandırmışlardır. Lazo vd., (2023) önerdikleri model ile doğruluk değerini 0,90 olarak hesaplamışlardır. Bilindiği gibi veri büyütme uygulaması sınıflandırma uygulamalarındaki performansı arttırmak için kullanılmaktadır. Nitekim orijinal görüntülerin etkisi incelendiğinde, bu çalışmada kullanılan VGG19+CBAM modeli, performans ölçütleri açısından daha iyi sonuç göstermiştir.

**Tablo 5.** Literatürdeki diğer çalışmalarla karşılaştırma

Referans	Görüntü sayısı	Model	Değerlendirme sonuçları
Shkolyar vd., (2019)	2335	CystoNet	Duyarlılık: %90,6
Ikeda vd., (2020)	8728	GoogLeNet	ROC: 0,98
Yang vd. (2021)	2350	EasyDL	Doğruluk: %96,9
Du vd., (2021)	734	EasyDL	Doğruluk: %96,9
Lazo vd., (2023)	8368	SeCSi-GAN	Doğruluk: 0,90 Kesinlik: 0,88 Hatırlatma: 0,89
Bu çalışma	1754	VGG19+CBAM	Doğruluk: 0,990 Kesinlik: 0,992 Duyarlılık: 0,984 AUC: 0,986

## 4. Sonuç ve Öneriler

Bu çalışmada mesane tümör dokularının sınıflandırılması için VGGNet modeliyle CBAM dikkat mekanizmasının birleştirilmesine dayanan bir yaklaşım önerildi. Çalışmada VGG16 ve VGG19 modelleri kullanıldı. Ortalama olarak en iyi performans sonuçlarını gösteren model VGG19+CBAM olarak belirlendi. Bu modelin doğruluk, kesinlik, duyarlılık, F1-skor ve AUC değerleri sırasıyla 0,990, 0,992, 0,984, 0,986 ve 0,994'tür. Deneysel sonuçlar, önerilen yaklaşımın mesane tümör görüntülerini ayırma kararlılığını



değerlendirme ölçütleri açısından önemli avantajlara sahip olduğunu gösterdi. Mesane tümörü görüntülerini sınıflandırmak için sadece ESA modelleri yeterli olsa da CBAM dikkat mekanizmasıyla ESA modelleri birleştirildiğinde performansın artışının olduğu görülmüştür. Gelecek çalışmalarda mesane görüntülerinin sınıflandırılması için önerilen modeller Grad-CAM (Gradientweighted Class Activation Mapping) modeliyle birleştirilmesiyle görsel olarak da sonuçların belirlenmesi sağlanacaktır. Ayrıca derin öğrenme tabanlı oluşturulacak farklı modellerle karşılaştırılma çalışmalarına da odaklanılacaktır.

## Veri Seti

Endoscopic Bladder Tissue Classification veri seti: <https://www.kaggle.com/datasets/aryashah2k/endoscopic-bladder-tissue-classification-dataset/data>

## Çıkar Çatışması

Yazar, bilinen herhangi bir çıkar çatışması veya herhangi bir kurum/kuruluş ya da kişi ile ortak çıkar bulunmadığını onaylamaktadırlar.

## Kaynakça

- Alirezazadeh, P., Schirrmann, M., & Stolzenburg, F., (2023). Improving Deep Learning-based Plant Disease Classification with Attention Mechanism. *Gesunde Pflanzen*, 75(1), 49-59. <https://doi.org/10.1007/s10343-022-00796-y>
- Chao, H., Fenhua, W., & Ran, Z., (2019). Sign language recognition based on cbam-resnet. Proceedings of the 2019 International Conference on Artificial Intelligence and Advanced Manufacturing, Ireland, 48, 1-6. <https://doi.org/10.1145/3358331.3358379>
- DeGeorge, K.C, Holt, H.R., & Hodges, S.C. (2017). Bladder cancer: diagnosis and treatment. *American family physician*, 96(8): 507-514.
- Du, Y., Yang, R., Chen, Z., Wang, L., Weng, X., & Liu, X., (2021). A deep learning network-assisted bladder tumour recognition under cystoscopy based on Caffe deep learning framework and EasyDL platform. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 17(1), 1-8. <https://doi.org/10.1002/rcs.2169>
- Gao, F., Wu, T., Li, J., Zheng, B., Ruan, L., Shang, D., & Patel, B., (2018). SD-CNN: A shallow-deep CNN for improved breast cancer diagnosis. *Computerized Medical Imaging and Graphics*, 70, 53-62. <https://doi.org/10.1016/j.compmedimag.2018.09.004>
- Ikeda, A., Nosato, H., Kochi, Y., Kojima, T., Kawai, K., Sakanashi, H., Murakawa, M., & Nishiyama, H., (2020). Support system of cystoscopic diagnosis for bladder cancer based on artificial intelligence. *Journal of endourology*, 34(3), 352-358. <https://doi.org/10.1089/end.2019.0509>
- Inoue, K., Fukuhara, H., Shimamoto, T., Kamada, M., Iiyama, T., Miyamura, M., Kurabayashi, A., Furihata, M., Tanimura, M., & Watanabe, H., (2012). Comparison between intravesical and oral administration of 5-aminolevulinic acid in the clinical benefit of photodynamic diagnosis for nonmuscle invasive bladder cancer. *Cancer*, 118(4), 1062-1074. <https://doi.org/10.1002/cncr.26378>
- Langlotz, C.P., Allen, B., Erickson, B.J., Kalpathy-Cramer, J., Bigelow, K., Cook, T.S., Flanders, A.E., Lungren, M.P., Mendelson, D.S., & Rudie, J.D., (2019). A roadmap for foundational research on artificial intelligence in medical imaging: from the 2018 NIH/RSNA/ACR/The Academy Workshop. *Radiology*, 291(3), 781-791. <https://doi.org/10.1148/radiol.2019190613>
- Lazo, J.F., Moccia, S., Marzullo, A., Catellani, M., De Cobelli, O., Rosa, B., de Mathelin, M., & De Momi, E., (2021). A transfer-learning approach for lesion detection in endoscopic images from the urinary tract. *arXiv preprint arXiv:210403927*.
- Lazo, J.F., Rosa, B., Catellani, M., Fontana, M., Mistretta, F.A., Musi, G., de Cobelli, O., de Mathelin, M., & De Momi, E., (2023). Semi-supervised Bladder Tissue Classification in Multi-Domain Endoscopic Images. *IEEE Transactions on Biomedical Engineering*, 70(10), 2822-2833. <https://doi.org/10.1109/TBME.2023.3265679>

- Mateen, M., Wen, J., Nasrullah, Song, S., & Huang Z., (2018). Fundus image classification using VGG-19 architecture with PCA and SVD. *Symmetry*, 11(1), 1. <https://doi.org/10.3390/sym11010001>
- Mukhtorov, D., Rakhmonova, M., Muksimova, S., & Cho, Y-I., (2023). Endoscopic image classification based on explainable deep learning. *Sensors*, 23(6), 3176. <https://doi.org/10.3390/s23063176>
- Sertkaya, M.E., Ergen, B., & Togacar, M., (2019). Diagnosis of eye retinal diseases based on convolutional neural networks using optical coherence images. 2019 23rd International conference electronics, Lithuania, 1-5. <https://doi.org/10.1109/ELECTRONICS.2019.8765579>
- Shirasuna, V.Y., & Gradwohl A., (2023). An optimized training approach for meteor detection with an attention mechanism to improve robustness on limited data. *Astronomy and Computing*, 45, 100753. <https://doi.org/10.1016/j.ascom.2023.100753>
- Shkolyar, E., Jia, X., Chang, T.C., Trivedi, D., Mach, K.E., Meng, MQ-H., Xing, L., & Liao, J.C., (2019). Augmented bladder tumor detection using deep learning. *European urology*, 76(6), 714-718. <https://doi.org/10.1016/j.eururo.2019.08.032>
- Siegel, R.L., Miller, K.D., Fuchs, H.E., & Jemal, A., (2021). Cancer statistics, 2021. *Ca Cancer J Clin*, 71(1), 7-33.
- Toğaçar, M., Özkurt, K.B., Ergen, B., & Cömert, Z., (2020). BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer. *Physica A: Statistical Mechanics and its Applications*, 545, 123592. <https://doi.org/10.1016/j.physa.2019.123592>
- Wang, D., Gao, F., Dong, J., & Wang, S., (2019). Change detection in synthetic aperture radar images based on convolutional block attention module. 2019 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), China, 1-4. <https://doi.org/10.1109/Multi-Temp.2019.8866962>
- Woo S., Park, J., Lee J-Y., & Kweon, I.S., (2018). Cbam: Convolutional block attention module. Proceedings of the European conference on computer vision (ECCV), 3-19. <https://doi.org/10.48550/arXiv.1807.06521>
- Yang, R., Du, Y., Weng, X., Chen, Z., Wang, S., & Liu, X., (2021). Automatic recognition of bladder tumours using deep learning technology and its clinical application. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 17(2), e2194. <https://doi.org/10.1002/rcs.2194>
- Zhou, F., Liu, X., Zhang, X., Liu, Y., Jia, C., & Wu, C., (2022). Keyhole status prediction based on voting ensemble convolutional neural networks and visualization by Grad-CAM in PAW. *Journal of Manufacturing Processes*, 80, 805-815. <https://doi.org/10.1016/j.jmapro.2022.06.034>