

## Mask R-CNN ile Mikroskopik İdrar Görüntüsü İçeriklerinin Tespiti

Yunus Emre YÖRÜK<sup>1\*</sup>, Hamdi Melih SARAOĞLU<sup>1</sup>, Ömer Faruk ÖZER<sup>2</sup>

<sup>1</sup> Kütahya Dumlupınar Üniversitesi, Lisansüstü Eğitim Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalı, Kütahya.

<sup>2</sup> Bezmialem Vakıf Üniversitesi, Tıp Fakültesi, Tıbbi Biyokimya Anabilim Dalı, İstanbul.

Sorumlu yazar e-posta\*: yeyoruk@gmail.com

hmelih.saraoglu@dpu.edu.tr

oozer@bezmialem.edu.tr

ORCID ID: <http://orcid.org/0000-0002-4455-0667>

ORCID ID: <http://orcid.org/0000-0002-5075-9504>

ORCID ID: <http://orcid.org/0000-0002-9034-4805>

Geliş Tarihi: 06.04.2023

Kabul Tarihi: 01.09.2023

### Öz

Mikroskopik idrar içerikleri doğru ve dikkatli bir şekilde analiz edildiğinde vücut hakkında önemli bilgiler verir. İdrar tahlilinin insan sağlığı için önemi nedeniyle mikroskopik idrar içeriklerinin tespit edilmesi amacıyla derin öğrenme görüntü işleme tekniği kullanılarak yapay zeka uygulamaları yapılmıştır. Literatürde yer alan çalışmaların çoğunda genel olarak semantik segmentasyon üzerine yoğunlaşmıştır. Bu çalışmada ise piksel düzeyinde segmentasyon yapabilen Mask R-CNN modeli ile mikroskopik idrar görüntülerindeki alyuvar, akyuvar, epitel, kristal, bakteri ve mantar içerikleri konum ve nesne türü bilgisiyle birlikte tespit edilmiştir. Mask R-CNN ile tespit edilen nesnelere maske ve çerçeve olmak üzere iki tip sınır çizilmektedir. Sistemin performansı her iki sınır tipi için ayrı ayrı incelenmiştir. Test için kullanılan 100 görüntüdeki toplam 1154 örüntüden maskelere göre 808 ve çerçevelere göre 843 nesne doğru şekilde tespit edilmiştir (IoU=0,5). En iyi tespit oranı akyuvarlar ve alyuvarlar için gerçekleşmiştir. Epiteller çerçevelere göre hesaplamada başarılı bir şekilde tespit edilmiştir fakat düzgün maske oluşturulamamıştır. Bakteriler diğerlerine göre çok küçük olduğu için doğru tespit oranı düşük kalmıştır. Kristallerin ve mantarların çoğu doğru şekilde tespit edilmiştir. Ayrıca, nesne tespitinde sıklıkla kullanılan değerlendirme metriği mAP de hesaplanmıştır. Sistem için hesaplanan mAP değerleri maskelere göre 0,7842 ve çerçevelere göre 0,8343 olmuştur. Mask R-CNN sistemi iyi bir şekilde optimize edilip daha fazla idrar içeriğine ait görüntülerle eğitilmesi durumunda idrar analiz sistemlerinde kullanılabilir.

### Anahtar kelimeler

Mask R-CNN; İdrar Analizi; Derin Öğrenme; Piksel Düzeyinde Segmentasyon

## Detection of Microscopic Urine Image Contents with Mask R-CNN

### Abstract

Urinary particles in microscopic images provide important information about the body when they analyse carefully and correctly. Based on the importance of urinalysis for human health, artificial intelligence applications were made using deep learning image processing technique in order to detect microscopic urine contents. Most of the studies in the literature have generally focused on semantic segmentation. Unlike the others, in this study, the urinary contents of red blood cells, white blood cells, epithelium, crystals, bacteria and yeast in microscopic urine images were determined using Mask R-CNN, which can perform instance segmentation. In object detection with Mask R-CNN, two types of boundaries are drawn as mask and bounding box. The performance of the system is examined for both boundary types. From a total of 1154 patterns in 100 images used for the test, 808 with masks and 843 with bounding boxes were correctly identified (IoU=0.5). The best detection occurred for white and red blood cells. Epithelium has also been successfully identified according to bounding boxes, but there were problems creating masks. Bacteria detection success rate is low because bacteria are so small. Most of the crystals and yeast were correctly detected. In addition, mAP, a frequently used evaluation metric for object detection, was also calculated. Calculated mAP values are 0.7842 and 0.8343 for masks and bounding boxes respectively. Mask R-CNN can be used in urine analysis systems if it is well-optimized and trained with images of more urine contents.

### Keywords

Mask R-CNN; Urinalysis; Deep Learning; Instance Segmentation

## 1. Giriş

İdrar analizi, kimyasal ve mikroskopik olmak üzere iki aşamalıdır. Kimyasal analizde özel olarak üretilmiş şeritlere damlatılan idrar sonrası renk değişimlerine bakılır. Mikroskopik analizinde ise idrar içeriğindeki akyuvar, alyuvar, kristal, bakteri gibi nesnelerin tespiti ve sayımı yapılır (Simerville *et al.* 2005, Strasinger and Di Lorenzo 2014).

İdrar örnekleri uygun teknikler kullanılarak doğru şekilde incelendiğinde mesane, idrar yolu, böbrekler ve diğer organ sistemleri hakkında önemli bilgiler sağlar (Perazella 2015, Int Kyn. 1). Tıp alanında teşhis ve tedavide sıklıkla kullanılan idrar analizleri, eskiden eğitilmiş uzman personeller tarafından yapılmaktaydı. Fakat bu şekilde yapılan analizler, insan kaynaklı hatalara açıktır, öznel değerlendirmeler içerir ve emek yoğunudur (Liang *et al.* 2018, Zaman *et al.* 2010). Günümüzde ise idrar tahlillerinin çoğu hastanelerde bulunan otomatik idrar analizörleri tarafından yapılmaktadır ve daha standart çıktılar elde edilmektedir (Zhang *et al.* 2018). Fakat, bu cihazların çoğu geleneksel görüntü işleme tekniklerini kullanır ve otomatik olmalarına rağmen yine de operatör kontrolüne ihtiyaç duymaktadırlar. Çünkü, tanıyı yanlış yönlendirebilecek düzeyde hatalı analiz sonuçları da üretebilmektedirler (İnce *et al.* 2021, Hu *et al.* 2010, Zaman *et al.* 2010). Mikroskopik analizi iyileştirmek adına bu cihazların görüntü işleme birimlerinde yapay zeka temelli görüntü işleme algoritmaları kullanılırsa; öğrenebilen, hatalarını düzeltip kendisini geliştirebilen ve insan kontrolüne ihtiyacı azaltan ya da ortadan kaldıran sistemler geliştirmek mümkün olabilir.

Yapay zeka sistemlerinin kullanıldığı güncel çalışmalarda idrarın mikroskopik analizleri için konvolüsyonel sinir ağları ya da kısaca CNN (Convolutional Neural Network) olarak ifade edilen derin öğrenme modelleri tercih edilmiştir (Suhail and Brindha 2021, Zeb *et al.* 2020). Derin öğrenme sistemlerinin nesne tespitindeki başarısının klasik görüntü işleme yöntemlerine göre daha iyi olması tercih sebebi olmaktadır (Grenspan 2016). İdrarın mikroskopik içeriğinde yer alan nesnelere konum

bilgisiyle birlikte tespit etmek için en çok R-CNN (Regions with CNN) (Girshick *et al.* 2014), YOLO (You Only Look Once) (Redmon *et al.* 2016) ve SSD (Single Shot Multibox Detector) (Li *et al.* 2018) algoritmaları ile çalışılmıştır.

Li vd. (2020) CNN kullanarak mikroskopun farklı odaklanmalarında ve hücrenin farklı düzlemsel açılarında idrardaki alyuvarları tespit etmeyi amaçlamışlardır. Deneysel sonuçlar önerdikleri sistemin 6 farklı alyuvar tipi için %82,1 başarı sağladığını göstermiştir. Primas (2018) CNN kullanarak alyuvar, akyuvar ve diğer parçacıkları tespit etmek için düşük maliyetli bir sistem oluşturmuştur. Sistem görüntü kaydedici, ayırıcı ve sınıflandırıcı olmak üzere üç bölümden oluşur. Alyuvar ve akyuvarlar için duyarlılığı sırasıyla %88 ve %91 olarak; özgüllüğü ise sırasıyla %89 ve %97 olarak hesaplamıştır. Liang vd. (2018) tarafından yapılan çalışmada Faster R-CNN ve SSD sistemleri kullanılarak 7 farklı idrar içeriği (akyuvar, alyuvar, epitel çekirdeği, epitel hücresi, kristal, mantar ve silendir) tespit edilmeye çalışılmıştır. En iyi mAP (mean Average Precision) değeri 0,841 olarak hesaplanmıştır. Farklı bir Faster R-CNN uygulamasında Zhang vd. (2018) idrarda yer alan akyuvar ve alyuvar hücrelerini tespit etmişlerdir. 6000'den fazla etiketlenmiş idrar görüntüsü kullanılmıştır ve ağırlıklı ortalama F1 skoru 0,914 olarak hesaplanmıştır. Yang vd. (2017) tarafından yapılan ve 4000 görüntünün kullanıldığı farklı bir çalışmada yine Faster R-CNN kullanılmıştır. Görüntülerin yüzde 70'i eğitim ve yüzde 30'u test için kullanılmıştır. Akyuvar, alyuvar, mantar, kristal, silendir ve epitel olmak üzere 6 farklı nesne türü tespiti yapılmıştır. Hesaplanan mAP değeri 0,916 olmuştur. Wang vd. (2019) tarafından SSD ve YOLO kullanılarak yapılan farklı bir çalışmada akyuvar hücreleri tespit edilmiş ve yüzde 90,09 doğruluk sağlanmıştır.

Literatür taramasında idrar içeriğindeki bakteri gibi diğerlerine göre küçük nesnelerin tespit çalışmalarına rastlanmamıştır. Bakterilerin idrar görüntülerindeki boyutunun çok küçük olması nedeniyle başarıyı düşürmektedir ve bu nedenle idrar analizlerinde bakterilerle yapılan derin

öğrenme tespit çalışmalarından kaçınılmıştır. Ayrıca, çalışmalarda genel olarak az sayıda nesne türü tespiti yapıldığı görülmüştür. Nesne boyut farkı olan içeriklerin aynı anda tespiti zor olduğundan ve çok sayıda nesne türü ile çalışmanın tespit performansına olumsuz etkisi nedeniyle tercihlerin bu şekilde olduğu düşünülmektedir. Bu çalışmada ise mikroskopik idrar analizinde tıbbi anlamda önem arz eden bakteri, mantar, akyuvar, alyuvar, kristal ve epitel içerikleri aynı anda tespit edilmiştir. Yöntem olarak piksel düzeyinde (instance) segmentasyon yapabilen Mask R-CNN derin öğrenme modeli kullanılmıştır. Mask R-CNN sistemleri görüntüde tespit edilen nesne sınırlarını piksel düzeyinde ayrıntılı olarak işaretleyebilmektedir (He *et al.* 2017). Bu sayede idrar içeriğindeki nesne sınırları hassas bir şekilde belirlenerek segmentasyon yapılabilmektedir.

## 2. Materyal ve Metot

Mikroskopik idrar görüntüsü içeriklerinin tespit edildiği bu çalışmada Python programlama dili ile oluşturulmuş Mask R-CNN kütüphanesinden (Int Kyn. 2) faydalanılmıştır. İdrar görüntülerinin bir kısmı ile Mask R-CNN derin öğrenme sistemi eğitilmiştir ve kalan kısmı ile sistemin performansı test edilmiştir.

Derin öğrenme sistemlerinde kullanılan veri setinin görüntü analiz sonuçlarına etkisi çok büyüktür. Bu nedenle görüntülerin ön işleme doğru ve dikkatli bir şekilde yapılmalıdır. Görüntüler, CNN ağlarının eğitiminde girdi olarak kullanmadan önce içindeki nesnelere uygun araçlarla etiketlenmelidir. Sonrasında üzerinde çalışılan model eğitilip test edilebilir.

### 2.1 Veri Seti

Bu çalışmada kullanılan idrar görüntüleri Bezmialem Vakıf Üniversitesi Hastanesi laboratuvarlarında gerekli izinler alınarak toplanmıştır. İnsan idrar örnekleri santrifüj edilmiş ve mikroskop altına yerleştirilmiştir. Mikroskoba entegre edilmiş olan kamera yardımıyla çok sayıda idrar görüntüsü bilgisayar ortamına aktarılmıştır. Bu işlemler idrar tahlili standartları dikkate alınarak

gerçekleştirilmiştir (Kouri *et al.* 2000). Veri kümesi toplam 600 adet idrar görüntüsünden oluşur; bunlardan 500 tanesi eğitim ve 100 tanesi de test için kullanılmıştır. Eğitim setinde 4201 adet konum ve sınıf bilgisi ile etiketlenmiş idrar içeriğine ait nesnelere görüntüsü yer alırken; test setinde ise 1154 adet konum ve sınıf bilgisi ile etiketlenmiş idrar nesnelere görüntüsü bulunmaktadır. İdrar içerikleri bakteri, epitel, kristal, eritrosit, lökosit ve mantar olmak üzere 6 farklı sınıf etiketine sahiptir. Eğitim ve test için kullanılan etiketlenmiş idrar görüntülerinde yer alan içeriklerin sınıfları ve sayıları sırasıyla Çizelge 1 ve Çizelge 2'de verilmiştir.

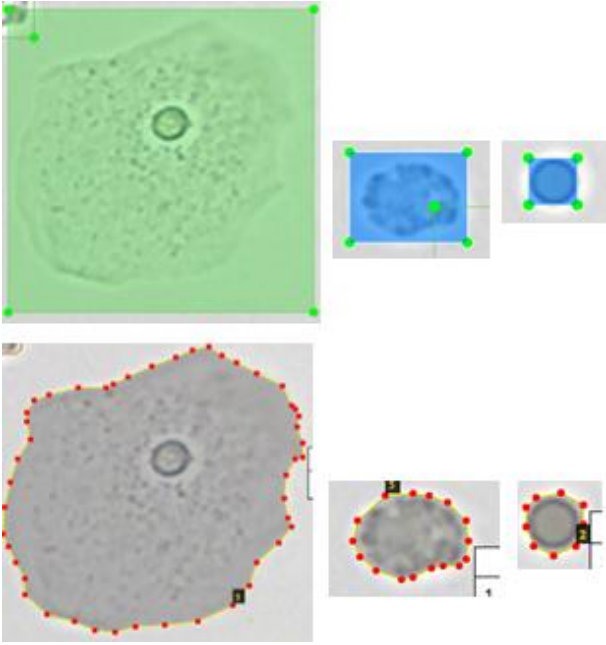
**Çizelge 1.** Eğitim setinde yer alan idrar görüntülerindeki nesnelere sınıflara göre dağılımı.

Sınıfı	Akyuvar	Alyuvar	Bakteri	Epitel	Kristal	Mantar
Sayısı	1164	1008	1228	148	401	252

**Çizelge 2.** Test setinde yer alan idrar görüntülerindeki nesnelere sınıflara göre dağılımı.

Sınıfı	Akyuvar	Alyuvar	Bakteri	Epitel	Kristal	Mantar
Sayısı	307	255	363	40	98	91

İdrardaki içerikler VGG Image Annotator (VIA) isimli uygulama (Int Kyn. 3) kullanılarak alanında uzman doktor ile etiketlenmiştir. VIA uygulaması web tarayıcısında çalışmaktadır ve herhangi bir kurulum gerektirmez. Görüntülerin hem sınıf hem de konum bilgilerini json formatında bir dosyada tutar. Derin öğrenme sistemlerinde tercih edilen dikdörtgen biçimli etiketlemenin yanında, görüntülerdeki içerikleri poligon bölge biçiminde etiketlemeye de olanak tanır. Poligon biçimli etiketleme hem zordur hem de zaman anlamında uzun sürmektedir. Fakat, görüntüdeki nesne sınırlarının daha doğru belirlenmesine imkân sağladığı için derin öğrenme ağının eğitiminde avantaj sağlamaktadır (Pande *et al.* 2022). Dikdörtgen biçimli etiketleme ile poligon biçimli etiketlemenin farkı Şekil 1'de görülebilir.

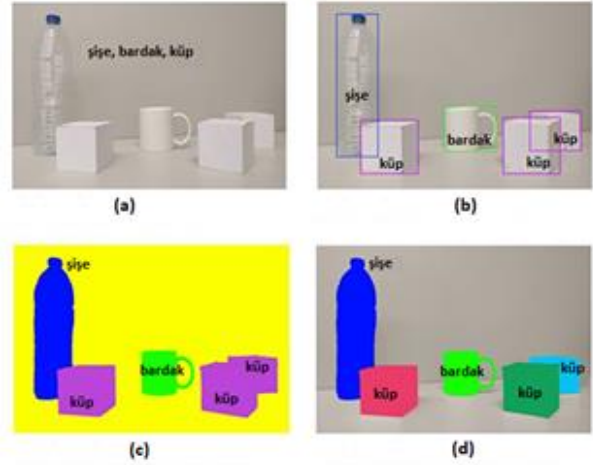


Şekil 1. Üstte dikdörtgen biçimli etiketlemeler ve altta poligon biçimli etiketlemeler.

## 2.2 Mask R-CNN

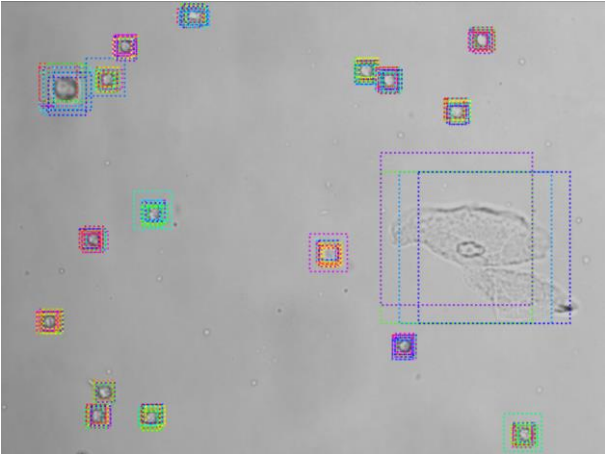
R-CNN modeli ilk olarak Ross Girshick vd. (2014) tarafından duyurulmuştur. R-CNN, görüntülerdeki örüntü konumlarını bulmak ve sınıf bilgisini tespit etmek için geliştirilmiş nesne tanıma modelidir. İhtiyaçlar doğrultusunda R-CNN yapısı sonraki yıllarda sırasıyla Fast R-CNN (Girshick 2015), Faster R-CNN (Ren *et al.* 2015) ve Mask R-CNN (He *et al.* 2017) olarak geliştirilmiştir.

Mask R-CNN bir önceki sürüm olan Faster R-CNN'den farklı olarak tespit edilen nesnelere maskeler oluşturan katmana sahiptir. Böylece piksel bazında (instance) segmentasyon gerçekleştirebilir. Bu şekilde yapılan segmentasyon ile görüntüdeki örüntülerin sınırları piksel bazında tespit edildiğinden daha detaylıdır. Bilgisayarlı görüde en zor görevlerden biridir (Garcia-Garcia *et al.* 2017). Şekil 2'de diğer nesne algılama türleriyle farklılıkları gösterilmiştir.

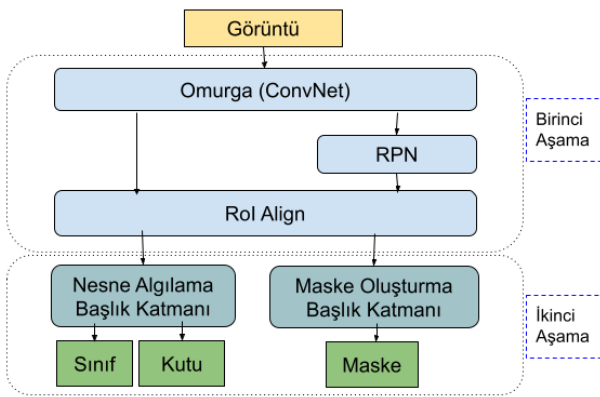


Şekil 2. a) Örüntünün sınıflandırılması b) Sınıflandırma ve nesne algılama c) Semantik segmentasyon d) Piksel bazlı (instance) segmentasyon. Kaynak: (Garcia-Garcia *et al.* 2017)

Mask R-CNN'de görüntü iki ana aşamada işlenir. İlk aşamada omurga olarak kullanılan CNN ağı tarafından görsel taranır ve konvolüsyonel öznelikler (özellik haritası) çıkarılır. RPN (Region Proposal Network) olarak isimlendirilen farklı bir ağ öznelikleri değerlendirir ve görselde nesne içermesi olasılığı bulunan ilgi alanlarını (RoI: Region of Interest) belirler. RoI olarak tanımlanan alanlara çerçeveler oluşturulur ve bunlar çapalar (anchors) olarak isimlendirilir. Oluşturulan her çapaya skor atanır ve yüksek skorlu olanlar pozitif çapalar olarak ifade edilir. Şekil 3'teki görselde örnek olarak veri tabanında yer alan resimlerden birinde en yüksek skorlu 200 çapanın gösterimi yapılmıştır. Çok fazla ve tekrarlı şekilde oluşturulan çerçeveler nedeniyle maksimum olmayı bastırma (non maximum suppression) algoritması da uygulanır. Veri işlemeyi kolaylaştırmak adına öznelikler ile ilgi alanları eşleştirilir ve bilgiler ikinci aşamaya devredilir. Mask R-CNN'in ikinci aşamasında ise belirlenen güven değeri üzerinde kalan ilgi alanları için görselde nihai çerçeveler en uygun olacak şekilde çizilerek sınıflandırma yapılır. Sonrasında tespit edilen nesnelere için maskeler oluşturularak işlem tamamlanır. Şekil 4'te Mask R-CNN'in genel akış şeması verilmiştir.



Şekil 3. Mask R-CNN ile görüntü işleme en yüksek skora sahip 200 çapanın gösterimi.



Şekil 4. Mask R-CNN aşamaları şematik gösterimi.

Mask R-CNN omurgası olarak ResNet101 (He *et al.* 2016) ağı kullanılmıştır. Ayrıca, özellik haritası çıkarmada özellik piramit ağından (FPN: Feature Pyramid Network) (Lin *et al.* 2017) da faydalanılmıştır.

### 2.3 Nesne Tespitinde Değerlendirme Metrikleri

Oluşturulan Mask R-CNN sisteminin performansı standart doğruluk (accuracy) ve genel ortalama hassasiyet (mAP: mean Average Precision) ile değerlendirilmiştir. R-CNN, YOLO gibi nesne tespiti modellerinin performansını ölçmede standart doğruluk hesabı yeterli görülmemiş ve mAP değerlendirme metriği oluşturulmuştur. Makine öğreniminde ve derin öğrenme ağlarının performans değerlendirmesinde başarı göstergesi olarak kullanılmaktadır (Int Kyn. 4). mAP hesaplamada ihtiyaç duyulan diğer metrikler bu bölümde verilmiş ve açıklanmıştır.

Kesişim bölgesi (IoU: Intersection of Union) ifadesi, görüntü işlemedeki nesne tespitinde model tahminlerinin veri kümesi etiketlemesi ile ne kadar uyumlu olduğunu göstermek için kullanılan bir metriktir (Rahman and Wang 2016). Mask R-CNN'in performans değerlendirmesinde eşik değeri gibi kullanılır. Belirlenen IoU değerinin altında kalan örtüşme hatalı tespit olarak; IoU değerini aşanlar ise doğru tespit olarak değerlendirilmektedir. Bu çalışmadaki performans değerlendirmesi hesaplamalarında veri kümesi ile tahminlerin minimum örtüşme oranı %50 (IoU=0,5) olarak alınmıştır. Mask R-CNN ile nesne tespitinde görüntüler için maskeler ve çerçeveler olmak üzere iki tip sınır çizilmektedir. Bu nedenle, IoU değerleri maskelere ve çerçevelere göre ayrı ayrı hesaplanmıştır. Hesaplama formülü Şekil 5'te görselleştirilerek verilmiştir.



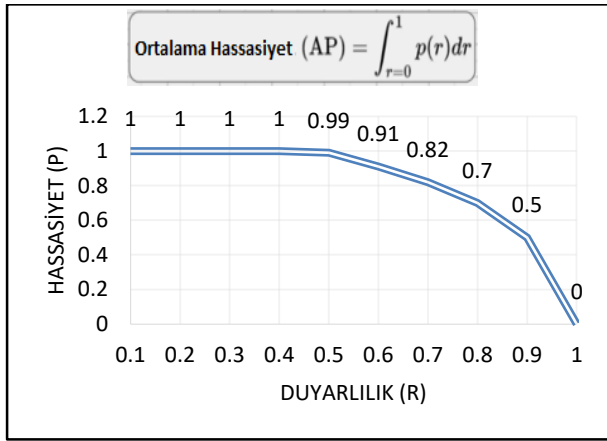
Şekil 5. Görselleştirilmiş IoU hesaplama formülü.

Veri setinde yer alan bir sınıf için; ağ tarafından yapılan sınıf bilgisi tahmini doğru ise doğru pozitif (DP), yanlış ise yanlış pozitif (YP) olarak ifade edilir. Veri setindeki aynı sınıf için; veri seti etiketlemesinde bu sınıfta yer almayan bir nesneyi ağ bu sınıfta tahmin ediyorsa yanlış negatif (YN), farklı sınıfta tahmin ediyorsa doğru negatif (DN) olarak ifade edilir (Schwenke and Schering 2014). Kısaca DP ve DN oluşturulan derin öğrenme ağının doğru çalıştığını YP ve YN ise hatalı tahmin yaptığını göstermektedir. Hassasiyet (P: precision) doğru tahminlerin ne kadarının gerçekten doğru olduğunu ve duyarlılık (R: recall) ise doğru yapılması gereken tahminlerin gerçekten ne kadarının doğru yapıldığını ifade eder. P ve R sırasıyla Eşitlik 1 ve Eşitlik 2'de verilen formüllerle hesaplanır (Flach and Kull 2015).

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

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

Hesaplanan P ve R değerlerinden P-R grafiği çizilir. Çizilen P-R grafiği altında kalan alana ortalama hassasiyet (AP: average precision) adı verilir. Şekil 6'da örnek P-R grafiği ve AP hesaplama formülü verilmiştir. AP'ler her sınıf için hesaplanıp tüm görüntüler için ortalamaları alınır nesne tespitinde önemli bir ölçüm metriği olan genel ortalama hassasiyet (mAP) değeri elde edilir (Flach and Kull 2015).



Şekil 6. P-R grafiği ve AP hesaplama formülü.

### 2.3 Çalışma Ortamı ve Sistem Konfigürasyonu

Mask R-CNN sisteminin eğitimi ve doğrulaması kişisel kullanım için oyunculara özel üretilmiş bilgisayarda yapılmıştır. Bu bilgisayar bir adet 10. nesil Intel i7 10750H 6 çekirdekli işlemci (CPU) ve bir adet Nvidia Geforce RTX2070 Max-Q dizayn 8GB ekran kartı (GPU) içermektedir. Sistem, eğitim setindeki görüntülerle eğitilmiştir ve test setindeki görüntüler ile Mask R-CNN sisteminin performans değerlendirmesi yapılmıştır. Model ağının eğitiminde öğrenme aktarımından (transfer learning) (Zhuang *et al.* 2020) yararlanılmıştır ve ImageNet (Deng *et al.* 2009) ile ön-eğitilmiş ağırlık dosyası kullanılmıştır. Nesne tespiti için minimum güven değeri 0,8 olarak alınmıştır. Eğitimde devir (epoch) sayısı 50 olarak ayarlanmış ve devir başına adım sayısı da 500 olarak hesaplanmıştır (eğitim

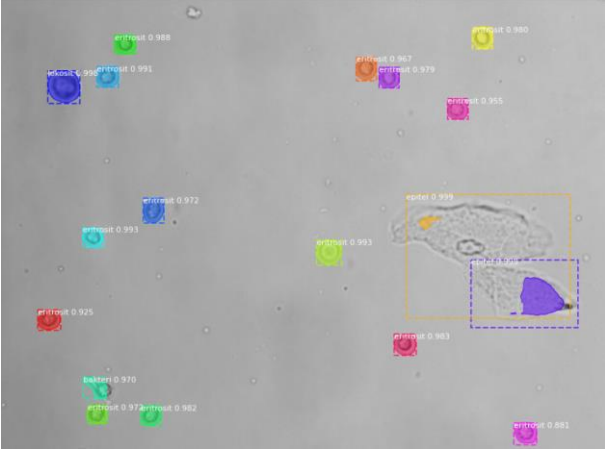
kümesindeki görüntü sayısı / gpu sayısı). Genel parametre yapılandırmaları Çizelge 3'te verilmiştir.

Çizelge 3. Test setinde yer alan idrar görüntülerindeki nesnelerin sınıflara göre dağılımı.

PARAMETRE	DEĞER
IMAGES_PER_GPU	1
STEPS_PER_EPOCH	500
VALIDATION_STEPS	100
BACKBONE	"resnet101"
COMPUTE_BACKBONE_SHAPE	None
BACKBONE_STRIDES	[4, 8, 16, 32, 64]
FPN_CLASSIF_FC_LAYERS_SIZE	1024
TOP_DOWN_PYRAMID_SIZE	256
NUM_CLASSES	6
RPN_ANCHOR_RATIOS	[0.5, 1, 2]
RPN_ANCHOR_STRIDE	1
RPN_NMS_THRESHOLD	0.7
RPN_TRAIN_ANCHORS_PER_IMAGE	256
POST_NMS_ROIS_TRAINING	2000
POST_NMS_ROIS_INFERENCE	1000
USE_MINI_MASK	True
MINI_MASK_SHAPE	(56, 56)
IMAGE_RESIZE_MODE	"square"
IMAGE_MIN_DIM	800
IMAGE_MAX_DIM	1024
IMAGE_MIN_SCALE	0
MEAN_PIXEL	np.array([123.7, 116.8, 103.9])
TRAIN_ROIS_PER_IMAGE	200
ROI_POSITIVE_RATIO	0.33
POOL_SIZE	7
MASK_POOL_SIZE	14
MASK_SHAPE	[28, 28]
MAX_GT_INSTANCES	200
RPN_BBOX_STD_DEV	np.array([0.1, 0.1, 0.2, 0.2])
BBOX_STD_DEV	np.array([0.1, 0.1, 0.2, 0.2])
DETECTION_MAX_INSTANCES	150
DETECTION_MIN_CONFIDENCE	0.8
DETECTION_NMS_THRESHOLD	0.3
LEARNING_RATE	0.001
LEARNING_MOMENTUM	0.9
USE_RPN_ROIS	True
TRAIN_BN	False
GRADIENT_CLIP_NORM	5

### 3. Bulgular

Eğitim sonrası çıktı olarak elde edilen ağırlık dosyası kullanılarak test için ayrılan 100 adet görüntüde Mask R-CNN ile nesne tespiti yaptırılmıştır. Şekil 7'de nesne tespiti sonrası görüntülerden birine ait çıktı paylaşılmıştır. Görsel çıktıda tespit edilen nesnelere çerçeve içine alınmaktadır ve nesnelere için oluşturulan maskeler renklendirilmektedir. Ayrıca, görüntülerin üst kısımlarına nesne türünü ifade eden sınıf bilgisi ve sistem tarafından tespit edilen güven değeri yazdırılmaktadır.



Şekil 7. Mask R-CNN ile nesne tespiti.

Oluşturulan Mask R-CNN sisteminin performans değerlendirmesi test setindeki görüntülerle yapılarak IoU değerinin maskelere göre hesaplandığı durumda, doğrulama setindeki 1154 idrar

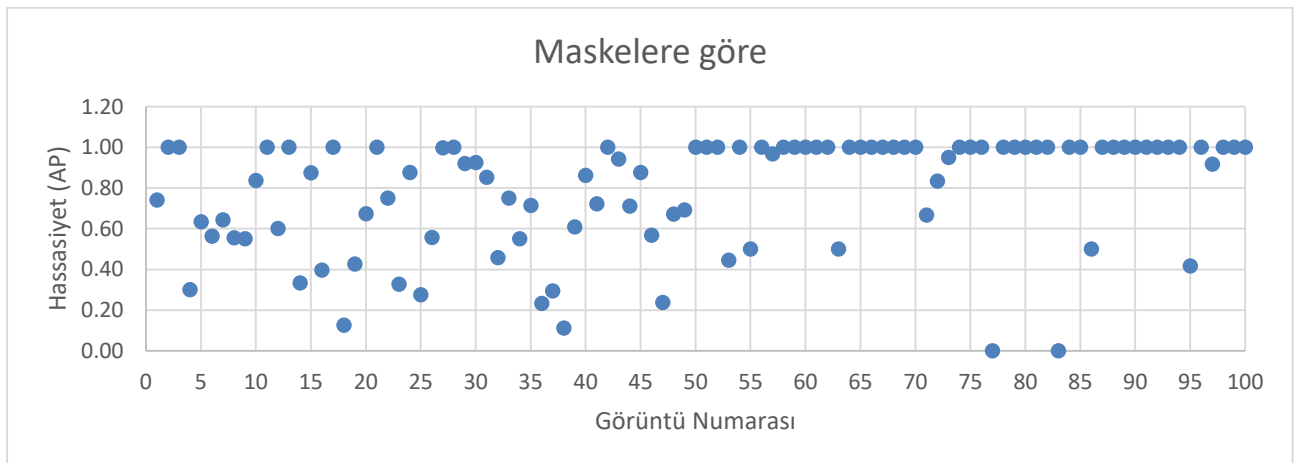
içeriğinden 808'i doğru şekilde tahmin edilmiştir. Diğer yandan, IoU değerlerinin çerçeveler ile hesaplandığı durumda 843 idrar içeriği doğru bir şekilde tespit edilmiştir. Hem maskelere göre hem çerçevelere göre yapılan değerlendirmede; en iyi tespit oranı akyuvarlar ( $\approx 93$ ) ve alyuvarda ( $\approx 89$ ) elde edilmişken en kötü tespit oranı bakterilerde ( $\approx 42$ ) gerçekleşmiştir. Epitel hücreleri sınırlayıcı kutulara göre hesaplamada başarılı bir şekilde tespit edilmişken ( $92,5$ ) maskelere göre olan hesaplamada başarı oranı çok düşmüştür ( $37,5$ ). Kristallerin ve mantarların yarıdan fazlası doğru tespit edilmiştir fakat tespit oranı idrar analizi için istenen düzeyde değildir. Çizelge 4'te sınıf bazında doğru ve yanlış tahminler ile tespit edilemeyen içeriklerin sayıları detaylı verilmiştir.

Çizelge 4. Veri setindeki test görüntülerinin tahmin sonuçları.

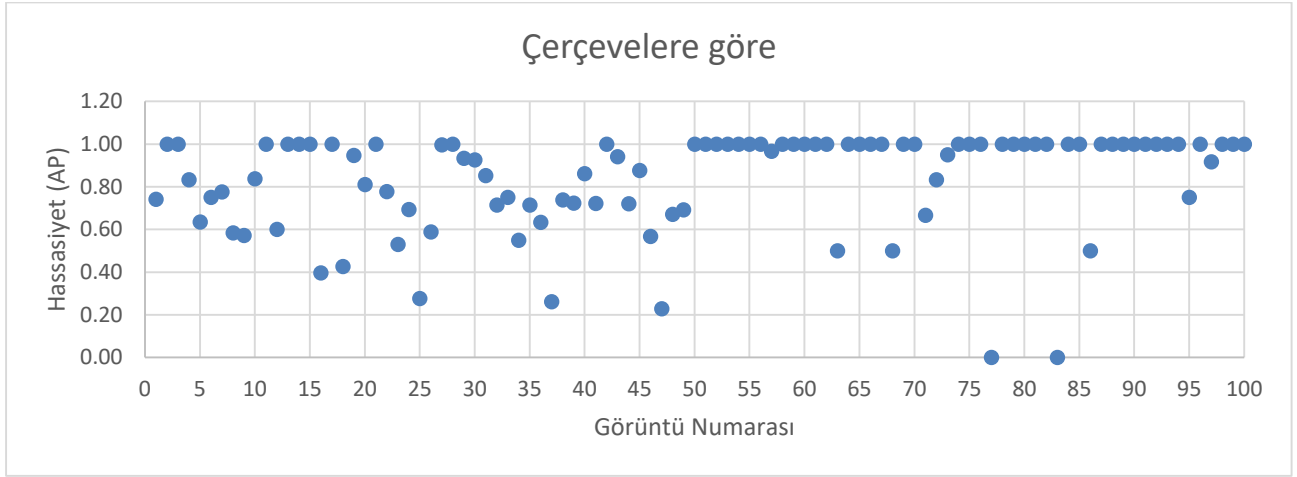
@ IoU=0,5		alyuvar	akyuvar	bakteri	epitel	kristal	mantar	Toplam
Maskelere göre	Doğru	226	285	153	15	71	58	808
	Yanlış	11	11	0	0	0	1	23
	Tespit Yok	18	11	210	25	27	32	323
	Doğruluk Oranı (%)	88,63	92,83	42,15	37,5	72,45	63,74	70,02
Çerçevelere göre	Doğru	226	288	158	37	75	59	843
	Yanlış	11	11	0	0	0	1	23
	Tespit Yok	18	8	205	3	23	31	288
	Doğruluk Oranı (%)	88,63	93,81	43,53	92,5	76,53	64,84	73,05

Mask R-CNN'in performans değerlendirmesinde doğru sınıflandırma ile birlikte doğru konum tahmini de önemli olduğundan AP ve mAP değerleri de hesaplanmıştır. Test kümesindeki her bir görüntü için P-R grafikleri çizdirilerek AP değerleri bulunmuştur. Test setindeki tüm görüntüler için maskelere ve çerçevelere göre ayrı ayrı hesaplanan

AP değerleri dağılım grafikleri sırasıyla Şekil 8 ve Şekil 9'da verilmiştir. Bu grafiklerde yer alan noktalar test setindeki her bir görüntü için elde edilmiş AP değerini ifade etmektedir. AP değerlerinin aritmetik ortalaması alınarak maskelerle genel ortalama hassasiyet  $mAP=0,7842$  ve çerçevelerle  $mAP=0,8343$  olarak hesaplanmıştır.



Şekil 8. Test setindeki 100 görüntünün maskelere göre AP değerleri dağılımı.



Şekil 9. Test setindeki 100 görüntünün çerçevelere göre AP değerleri dağılımı.

#### 4. Tartışma ve Sonuç

Literatürde yer alan çalışmaların çoğu semantik segmentasyon ve dikkörtgen biçimli etiketleme yapmaktadır. Bu nedenle, üst üste binen veya bitişik olan hücreleri/nesneleri tespit etmede sorunlar yaşanabilmektedir. Ayrıca, sadece birkaç tip idrar içeriğinin tespiti yapılmış çalışmalar literatürde yaygındır ve bu çalışmaların sonuçları az sayıda nesne sınıfı tespitinden dolayı başarılı görünmektedir. Daha fazla idrar içerik türü ile çalışılırsa (özellikle bakteri gibi küçük nesnelere) çalışmalarda tespit performanslarının kötüleşmesi kaçınılmaz olacaktır.

Mask R-CNN ile yapılan bu çalışmada ise tespit performansına olumsuz etkisine rağmen tıbbi önemi olan idrar nesnelere çalışmaya dahil edilmiştir. Küçük nesnelere tespit performansını olumsuz etkilemesine rağmen Mask R-CNN ile kurgulanan sistemin nesne tespit performansı tatminkardır. Özellikle akyuvar ve alyuvar hücrelerinin tespitinde yaklaşık %90 doğrulukta başarılı sonuçlar alınmıştır. Epitel hücreleri başarılı bir şekilde tespit edilmiş ve çerçevelere göre %92,5'lik doğruluk yüzdesi yakalanmıştır. Fakat epitel hücreleri için maskeler düzgün oluşturulamamış ve bu nedenle maskelere göre doğru tespit oranı %37,5'e düşmüştür. Epitellerin şekilleri çok farklılık gösterdiği için ve eğitim setinde epitel hücre görüntüsü diğerlerine göre az olduğu için bu durum yaşanmıştır. Mask R-CNN'in maske oluşturma katmanı özellikle epiteller için yeniden optimize edilebilir. Bakteriler %40

civarında başarı ile tespit edilebilmiştir. Bakteriler diğer idrar içeriklerine göre boyut olarak çok küçük olduğundan tespit etmek güçtür. Bakterileri ayrı bir katmanda ya da farklı döngüde tespit etmek doğru bir yaklaşım olacaktır. Kristaller ve mantarların çoğu doğru şekilde tespit edilmiştir fakat idrar analizinde daha yüksek başarı oranı gereklidir. Eğitim setinde yer alan kristaller ve mantarlar diğerlerine göre daha azdır. Bu sayıyı arttırma gerekliliği açıkça görülmüştür.

Genel ortalama hassasiyet sonuçları incelendiğinde ilk 50 görüntü için AP değerlerinin dağılımı çok farklılık gösterirken son 50 görüntü daha stabildir. Bunun nedeni ilk 50 görüntüde idrar nesnelere ait görüntü sayısının fazla; son 50 görüntüdeki görüntü sayısının ise az olmasından kaynaklanmaktadır. Bazı görüntülerde sadece bir tane görüntü olup, onun da tespit edilememesi sonucu sıfır AP değerleri görülmüştür. Bakterilerde kötü tespit performansı ve epitellerde maske oluşturulamaması nedeniyle mAP değeri aşağı çekilmiştir. Buna rağmen başarılı sayılabilecek sonuçlar alınmıştır.

Derin öğrenme sistemlerinde eğitim setindeki görüntü sayısının fazla olması daha iyi öğrenmeyi sağlar. Bu nedenle, kullanılan eğitim setindeki görüntülerin arttırılması sistem performansına olumlu yönde katkı sağlayacaktır. Oluşturulan Mask R-CNN sistemi daha fazla görüntü ile eğitilip iyi bir şekilde optimize edildiğinde idrar analizinde kullanılabilir. Ayrıca, piksel düzeyinde



segmentasyon, idrarda küçük nesnelere tespiti ve polinom bazlı etiketleme gibi ekstra özelliklere de sahiptir.

Bu çalışma, idrar analizörlerinin derin öğrenme ile geliştirilmesine rehberlik edebilir. Öğrenebilen, hatalarını düzeltebilen ve insan kontrolüne ihtiyaç duymayan yapay zeka temelli tıbbi cihazlar üretilebilir.

### Teşekkür

Kütahya Dumlupınar Üniversitesi Bilimsel Araştırma Projeleri (BAP) ve TÜBİTAK 1071 – Uluslararası Araştırma Fonlarından Yararlanma Kapasitesinin ve Uluslararası Ar-Ge İşbirliklerine Katılımın Arttırılmasına Yönelik Destek Programı kapsamında desteklenmiştir.

Bu çalışmada kullanılan idrar görüntülerinin temini konusunda Bezmialem Vakıf Üniversitesi Hastanesi destek vermiştir.

### 5. Kaynaklar

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 248-255). IEEE.

Flach, P., & Kull, M., 2015. Precision-recall-gain curves: PR analysis done right. *Advances in Neural Information Processing Systems*, **28(1)**, 838-846.

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J., 2017. A review on deep learning techniques applied to semantic segmentation. *arXiv preprint arXiv:1704.06857*.

Girshick, R., Donahue, J., Darrell, T., & Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 580-587). IEEE.

Girshick, R., 2015. Fast r-cnn. In *IEEE International Conference on Computer Vision* (pp. 1440-1448). IEEE.

Greenspan, H., Van Ginneken, B., & Summers, R. M., 2016. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, **35(5)**, 1153-1159.

He, K., Zhang, X., Ren, S., & Sun, J., 2016. Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770-778). IEEE.

He, K., Gkioxari, G., Dollár, P., & Girshick, R., 2017. Mask r-cnn. *IEEE International Conference on Computer Vision* (pp. 2961-2969). IEEE.

Hu, X., Zhang, J., & Zhang, X., 2010. Evaluation of the Sysmex UF-1000i urine analyzer as a screening test to reduce the need for urine cultures for urinary tract infection. *Laboratory Medicine*, **41(6)**, 349-352.

İnce, H., İmamoğlu, S. E., & İmamoğlu, S. Z., 2021. Yapay zeka uygulamalarının karar verme üzerine etkileri: Kavramsal bir çalışma. *International Review of Economics and Management*, **9(1)**, 50-63.

Kouri, T., Fogazzi, G., Gant, V., Hallander, H., Hofmann, W., & Guder, W. G., 2000. European urinalysis guidelines. *Scandinavian journal of clinical and laboratory investigation*, **60(sup231)**, 1-96.

Li, Y., Huang, H., Xie, Q., Yao, L., & Chen, Q., 2018. Research on a surface defect detection algorithm based on MobileNet-SSD. *Applied Sciences*, **8(9)**, 1678.

Li, X., Li, M., Wu, Y., Zhou, X., Hao, F., & Liu, X., 2020. An accurate classification method based on multi-focus videos and deep learning for urinary red blood cell. *Conference on Artificial Intelligence and Healthcare* (pp. 67-71). ACM Digital Library.

Liang, Y., Kang, R., Lian, C., & Mao, Y., 2018. An end-to-end system for automatic urinary particle recognition with convolutional neural network. *Journal of Medical Systems*, **42(9)**, 1-14.

Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S., 2017. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2117-2125). IEEE.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C., 2016. Ssd: Single shot multibox detector. In *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14* (pp. 21-37). Springer International Publishing.

- Pande, B., Padamwar, K., Bhattacharya, S., Roshan, S., & Bhamare, M., 2022. A Review of Image Annotation Tools for Object Detection. In *2022 International Conference on Applied Artificial Intelligence and Computing* (pp. 976-982). IEEE.
- Perazella, M. A., 2015. The urine sediment as a biomarker of kidney disease. *American Journal of Kidney Diseases*, **66**(5), 748-755.
- Primas, S. R., 2018. The AutoScope: an automated point-of-care urinalysis system. Unpublished Doctoral Dissertation, Massachusetts Institute of Technology, USA.
- Rahman, M. A., & Wang, Y., 2016. Optimizing intersection-over-union in deep neural networks for image segmentation. In *International symposium on visual computing* (pp. 234-244). Springer, Cham.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A., 2016. You only look once: Unified, real-time object detection. In *IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779-788). IEEE.
- Ren, S., He, K., Girshick, R., & Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, **28**, 91-99.
- Schwenke, C., & Schering, A. G., 2014. True positives, true negatives, false positives, false negatives. Wiley StatsRef: Statistics Reference Online.
- Simerville, J. A., Maxted, W. C., & Pahira, J. J., 2005. Urinalysis: a comprehensive review. *American family physician*, **71**(6), 1153-1162.
- Strasinger, S. K., & Di Lorenzo, M. S., 2014. Urinalysis and body fluids. FA Davis.
- Suhail, K., & Brindha, D., 2021. A review on various methods for recognition of urine particles using digital microscopic images of urine sediments. *Biomedical Signal Processing and Control*, **68**, 102806.
- Wang, Q., Bi, S., Sun, M., Wang, Y., Wang, D., & Yang, S., 2019. Deep learning approach to peripheral leukocyte recognition. *PLoS One*, **14**(6), e0218808.
- Yang, S., Fang, B., Tang, W., Wu, X., Qian, J., & Yang, W., 2017. Faster R-CNN based microscopic cell detection. In *2017 International Conference on Security Pattern Analysis and Cybernetics* (pp. 345-350). IEEE.
- Zaman, Z., Fogazzi, G. B., Garigali, G., Croci, M. D., Bayer, G., & Kránicz, T., 2010. Urine sediment analysis: Analytical and diagnostic performance of sediMAX®-a new automated microscopy image-based urine sediment analyser. *Clinica Chimica Acta*, **411**(3-4), 147-154.
- Zeb, B., Khan, A., Khan, Y., Masood, M. F., Tahir, I., & Asad, M., 2020. Towards the Selection of the Best Machine Learning Techniques and Methods for Urinalysis. In *Proceedings of the 2020 12th International Conference on Machine Learning and Computing* (pp. 127-133). ACM Digital Library.
- Zhang, X., Chen, G., Saruta, K., & Terata, Y., 2018. Detection and classification of RBCs and WBCs in urine analysis with deep network. In *ACHI 2018: The Eleventh International Conference on Advances in Computer-Human Interactions* (pp. 194-198). IARIA.
- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q., 2020. A comprehensive survey on transfer learning. *Proceedings of the IEEE*, **109**(1), 43-76.

#### İnternet kaynakları

- 1-<https://pdf.medicalexpo.com/pdf/roche/compendium-urinalysis-urine-test-strips-microscopy/71020-136212.html>, (30.03.2023)
- 2-[https://github.com/matterport/Mask\\_RCNN](https://github.com/matterport/Mask_RCNN), (30.03.2023)
- 3-<https://www.robots.ox.ac.uk/~vgg/software/via/>, (30.03.2023)
- 4-<https://jonathan-hui.medium.com/map-mean-average-precision-for-object-detection-45c121a31173>, (30.03.2023)