TEKNOFEST

HAVACILIK, UZAY VE TEKNOLOJİ FESTİVALİ

SAĞLIKTA YAPAY ZEKÂ YARIŞMALARI

(Bilgisayarlı Görüyle Abdomen (Karın) Bölgesi için Hastalık Tespiti Kategorisi)

PROJE DETAY RAPORU

TAKIM ADI: LENSAI

TAKIM ID: 347686

İçindekiler

1-Proje Mevcut Durum Değerlendirmesi……...3

2-Özgünlük……5

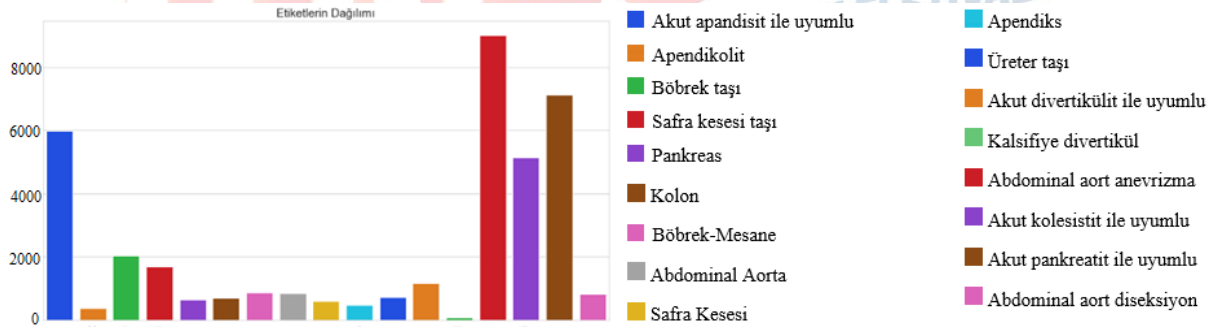
3-Sonuçlar ve İnceleme…….8

4-Deney ve Eğitim Aşamalarında Kullanılan Veri Setleri…….....10

5-Referanslar……...11

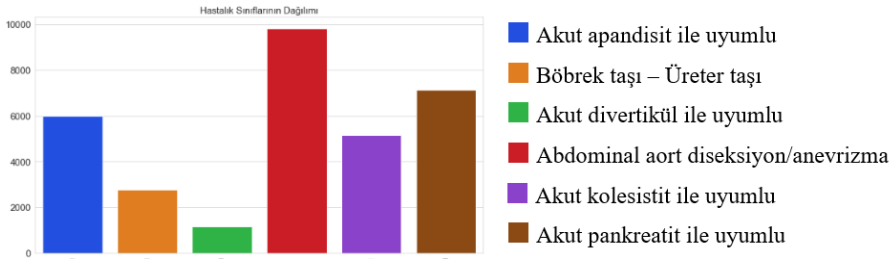
1. Proje Mevcut Durum Değerlendirmesi (20 puan)

Proje Sunuş Raporu tesliminden sonra veri seti paylaşılma tarihine kadar geçen süre zarfında eğitmek istediğimiz modellerin mimari detaylarına, veri seti için uygunluklarına, eğitim süreçleri, aldıkları girdiler ve sonuç olarak ürettikleri çıktılar üzerinde incelemelerde bulunduk. İlk raporumuzda da belirttiğimiz gibi yarışmada kullanılacak veri seti ile benzerliği sebebiyle gerçek veri seti gelene kadar 32.000'den fazla BT görüntüsü içeren DeepLesion [1] veri setinden yararlandık. ImageNet ağırlıkları ile model eğitimine başlanarak DeepLesion verisi ile eğitilen modellerimizi buradan elde ettiğimiz ağırlıkları kullanarak yarışma veri seti ile eğitmeyi planlamaktayız. Tabi bu sürecin modellerin performansına olan olumlu veya olumsuz etkisini incelemek adına modellerimizi ilk olarak ImageNet ağırlıkları ve yarışma veri seti ile eğittik. Hali hazırda DeepLesion verisi ile eğitilmiş model ağırlıklarımız ile yarışma veri setini bir araya getiremedik. İlerleyen süreçte bu işlemi gerçekleştirmeyi planlamaktayız. Bilgisayarlı görü problemlerinde girdi verisini anlamanın ve görselleştirmenin problemi çözmenin önemli bir parçası olduğuna inanmaktayız. Yarışma verisini alır almaz ilk olarak analizlere başladık. Yarışma verisini indirdiğimizde boyutunun yaklaşık 140 GB olduğunu gördük. Toplamda 1182 hastaya ait olan abdomen BT görüntüleri hastalara ait klasörler halinde verilmişti. Yaptığımız incelemeler sonucunda hasta klasörleri içerisinde birden fazla seri olduğunu da keşfettik. Yarışma için verilen BT görüntüleri dicom formatında axial düzlemde çekilmiş (slice) kesitlerden meydana gelmektedir. Kesit sayılarına baktığımızda her bir hasta için sabit olmadığını değişkenlik gösterdiğini fark ettik. Hastaların fiziksel özelliklerine, çekim yapan uzmana, çekim yapılan makineye bağlı olarak bunun değişiklik gösterebileceğini veriyi almadan önce tahmin ediyorduk. Sağlık bakanlığı tarafından sağlanan veri seti ile birlikte verilen kesit numaralarının ve etiketlerin yer aldığı dosyayı analiz ettik. Dosyada toplamda 38236 satır bulunmakla birlikte verinin 17 farklı sınıfa (Akut apandisit ile uyumlu, Apendikolit, Böbrek taşı,Safra kesesi taşı, Pankreas, Kolon, Böbrek-Mesane, Abdominal Aorta, Safra Kesesi, Apandiks, Üreter taşı, Akut divertikülit ile uyumlu, Kalsifiye divertikül, Abdominal aort anevrizma, Akut kolesistit ile uyumlu, Akut pankreatit ile uyumlu, Abdominal aort diseksiyon) dağıldığını ve dağılımın düzenli olmadığını gördük. Söz konusu 17 sınıfa ait etiket dağılımı Şekil 1'de verilmiştir.



Şekil 1: Etiketlerin Dağılımı

17 farklı etiket yani sınıf olmasına karşın yarışmada tespit edilmesi istenilen 6 farklı sınıf bulunmaktadır. Bunlar verilen sınıf numaraları ile Akut apandisit ile uyumlu(1), Akut kolesistit ile uyumlu(2), Akut pankreatit ile uyumlu(3), Böbrek taşı(4), Üreter taşı(4) Akut divertikülit ile uyumlu(5), Abdominal aort diseksiyon(6), Abdominal aort anevrizma(6). Sınıflar arasında dengesizlik olduğu Şekil 2'den açıkça görülmektedir.



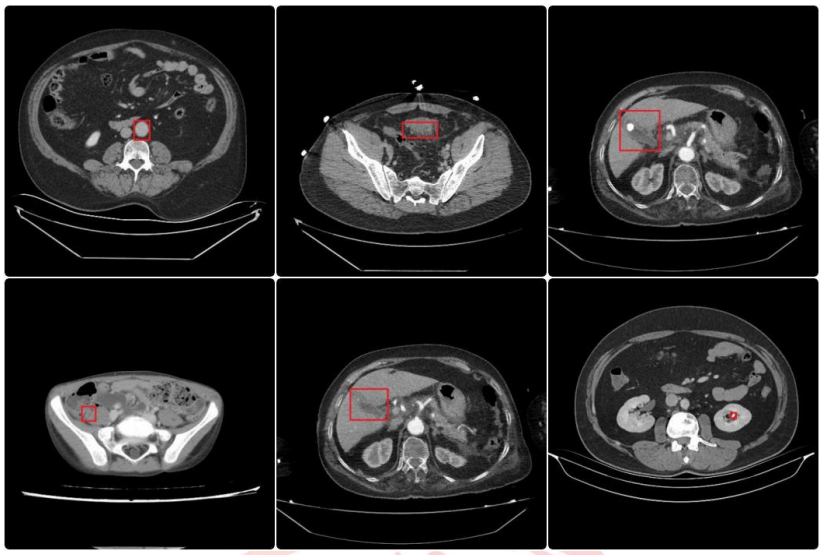
Şekil 2: Hastalık Sınıflarının Dağılımı

Modellerimizi eğitmek için hasta sayısını göz önünde bulundurarak veri setini %80 eğitim %20 test ve doğrulama olarak ayırdık. Toplamda 800 hastayı eğitim için 200 test için 100 doğrulama için kullanıldı. Bu ayırma işleminden sonra eğitim için bu kadar doğrulama için bu kadar test için bu kadar örnek elde edilmiştir. Bunların dağılımları Tablo 1'de verilmiştir.

Sınıf	**Eğitim**	**Doğrulama**	**Test**
Akut apandisit	4347	583	629
Akut kolesistit	3938	502	485
Akut pankreatit	5539	679	705
Böbrek/Üreter taşı	2021	296	239
Akut divertikülit	923	121	104
Aort anevrizma/dis	7934	929	954

Tablo 1: Veri Setinin Eğitim, Test ve Doğrulama İçin Dağılımı

Modelleri eğitmeden önce bizlerle paylaşılan veri seti üzerinde çeşitli ön işlemler gerçekleştirdik. Seri görüntüleri 'DICOM' formatı şeklinde paylaşılmıştı, 'DICOM' formatı içerdiği görüntü haricinde içerisinde ek bilgiler de barındırmaktadır, dolayısıyla görüntünün boyutunu oldukça arttırmaktaydı. Eğitim aşamasında bizlere sorun olacağını düşündüğümüz için görüntüleri uygun 'Hounsfield scale' değerlerine göre 'jpg' formatına çevirdik. Hounsfield scale değeri tomografi çekimlerinde kullanılan radyolojik ölçeklendirme birimidir[2]. Bu değer çekilen görüntülerin değerlendirilmesi ve yorumlanmasında kullanılmaktadır. Hounsfield ölçeklendirmesine göre 0 değeri su, 1000 kemik ve hava ise -1000 değerleriyle temsil edilmektedir. Bu ölçeği detaylı olarak kullanarak bizlerle paylaşılan görüntülerde bulunan dokuya uygun olarak bir hounsfield değeri belirledik. Bu ölçeğin aldığı bazı değerler organları dolayısıyla organlardaki hastalıkları daha iyi göstermektedir. Örneğin böbrekler için 20-50, pankreas için 30-50, karaciğer için 50-70 aralıkları bu organların tespitinin daha kolay olmasını sağlamaktadır. Biz de hastalıkların tespiti için denemelerimizin sonucunda 49 ile 450 aralığını veri seti için belirledik. DICOM formatında bizlerle paylaşılan görüntüleri hounsfield değeri ile jpg formatına çevirdik. Bu işlem sonucunda DICOM formatındaki bir görüntü jpg formatına çevrildiğinde hafızada kapladığı alan bakımından %60 oranında azalma oldu. Eğitim aşamasında kullanacağımız görüntüleri verilen excel dosyasını kullanarak seçtik. Veri setinin boyutu çok fazla olduğu için boyutu anlamlı şekilde küçültmek istedik. Excel dosyasında hastalığın bulunduğu (bbox içeren) görüntüleri seçerek sadece o dosyaları jpg formatına dönüştürdük. Görüntü formatını dönüştürme işlemi bize boyut avantajı sağladı. Ayrıca modellerin eğitiminde veri seti içerisindeki hastalıklı etiket içeren görüntüleri kullanarak kesit sayısını azalttık. Tüm ön işleme aşamalarından işleminden sonra kullanacağımız veri boyutu 1 GB' a düştü. Şekil 3'te veri setinde örnekler verilmiştir.

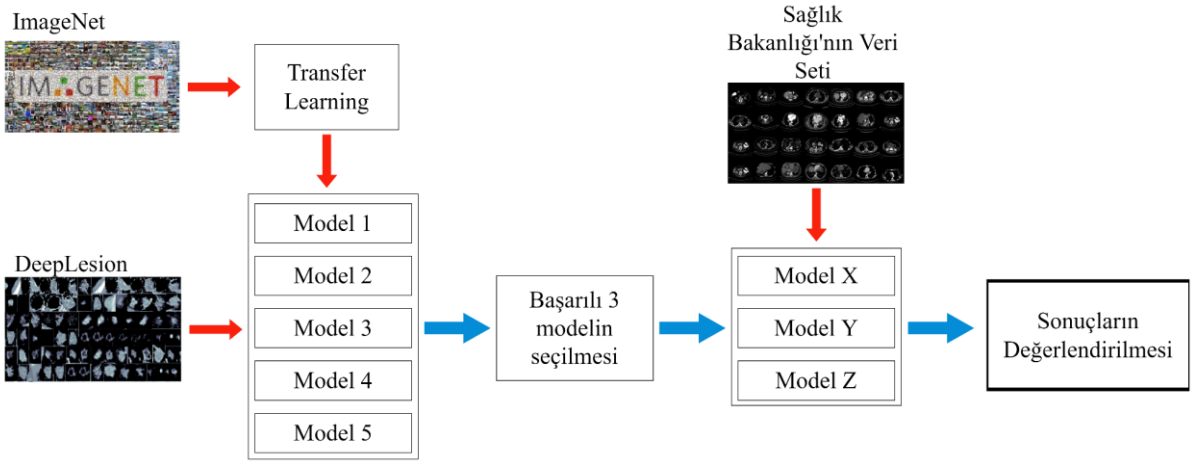


Şekil 3: Veri setinden Aort Anevrizma, Divertikülit, Pankreatit, Apandisit, Kolesistit, Böbrek Taşı Örnek Etiketlemeleri

Veri seti analizi ve ayırma işlemleri bittikten sonra ilk raporda belirtmiş olduğumuz modellerimizi eğitimine başladık. İlk aşamada 3 modelin eğitimini tamamladık. Bu modeller YOLOv5 [3], Faster R-CNN [4]. YOLOv5'in iki farklı versiyonunun eğitimini bitirdik. Bunlar YOLOv5s ve YOLOv5m'dir. Her bir model girdi görüntüsü 512x512x3 olacak şekilde 100 epok boyunca 32 batch size ile eğitilmiştir.
Eğitim aşamasında olduğumuz diğer bir model olan Faster R-CNN modelinin girdi görüntü boyutu 600x1024, batch size 1, öğrenme katsayısı 0.002 olarak ayarlanmış ve 70.000 step eğitilmiştir. İlk raporda belirttiğimiz SSD [5], RetinaNet [6] ve MULAN [7] modellerinin eğitimi için gerekli hazırlıkları tamamladık ancak henüz başlayamadık.

2. Özgünlük (30 puan)

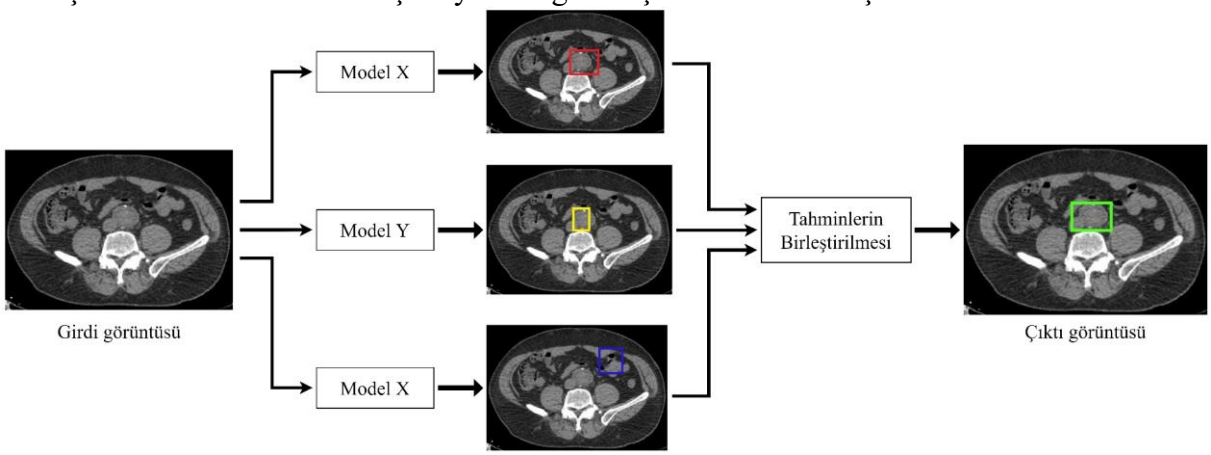
Bu çalışmanın iki özgün yönü bulunmaktadır. Bunlardan ilki direkt olarak ImageNet ağırlıkları kullanmak yerine yarışma veri seti ile benzer özellikleri taşıyan daha büyük bir veri seti ile modellerimizi eğittikten sonra elde ettiğimiz ağırlıkları kullanarak yarışma veri seti ile modellerimizi tekrar eğiteceğiz. Bu sayede modellerin daha başarılı tespitler yapacağına inanıyoruz. Şekil 4 'te bu durumu açıklayan bir görsele yer verilmiştir.



Şekil 4: Önerilen Transfer Learning Yöntemi

Çalışmamızın ikinci özgün yani ise oylama sistemine dayalı bir yöntem olan ensemble

yöntemidir. Her bir CNN modelinin mimari yapısına ve toplam parametre sayısına bağlı olarak güçlü ve zayıf yönleri bulunmaktadır. Bu çalışma en az 3 modelin tahmin sonuçlarını birleştirerek daha iyi tespit performansı elde etmeyi planlamaktayız. Bu amaçla K-Means benzeri bir algoritma tasarladık. Algoritmanın çalışma sistemini kısaca özetleyecek olursak, her bir modele örneğin böbrek taşı hastalığı şüphesini içeren görüntü ayrı ayrı verilecektir. Modeller hastalığın bulunup bulunmadığı, bulundu ise hastalığın bölgesini sınırlayıcı kutular ile işaretleyecekler. Modellerin işaretlemiş olduğu sınırlayıcı kutuların merkez noktarı göz önünde bulundurularak yeni bir merkez noktası belirlenecek ve 3 model için tek bir sınırlayıcı kutu çizilecektir. Bu durumu açıklayan bir görsel Şekil 5'te verilmiştir.



Şekil 5: Önerilen Ensemble Yöntemi

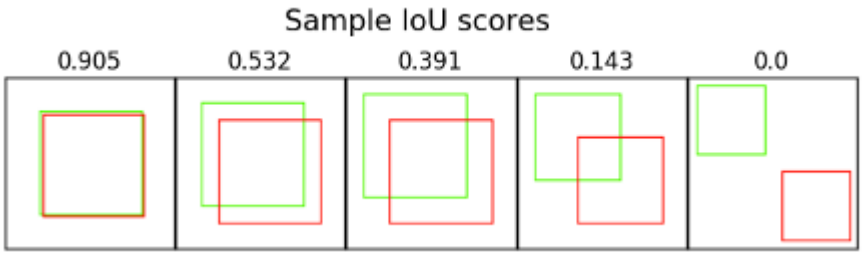
Şekil 5'te verilen yöntemin matematiksel hesabı anlatımı kısaca aşağıdaki gibidir. İlk olarak elimizde bulunan tüm modellerden görüntüye ait tahmin bilgileri alınır. Tahmin bilgileri şu bilgileri içermelidir; x1, y1, x2, y2 ve c değeri. Bu değerler sırası ile sınırlayıcı kutunun koordinat bilgilerini ve sınıf tahminini temsil etmektedir. Bu durumu bir örnek ile açıklayacak olursak bir görüntü için tahmin yapan 3 modelin x1, y1, x2, y2 ve sınıf tahmin oranı aşağıdaki gibidir.
Örneğin;

Model1= 90,30 160,90 Tahmin Değeri = 0,9

Model2= 120,40 160,80 Tahmin Değeri = 0,3

Model3= 200,50 220,90 Tahmin Değeri = 0,1

Tahmin Değeri en yüksek olan model ile diğer modeller arasında sırayla Intersection over Union (IoU) değeri hesaplanır. Şekil 6'da verilen yeşil renkli kare ile belirlenen alan tahmin değeri en yüksek olan modeli temsil etmektedir. IoU değeri 0 olarak hesaplanırsa diğer model ensemble learning yönteminden elenir.



Şekil 6: IoU Hesaplanma Örneği [8]

IoU değerleri hesaplaması aşamasından sonra geriye kalan modellerin orta noktaları hesaplanır. Hesaplanan orta noktalar modellerin tahmin değerlerindeki ağırlıklarına göre birleştirilir. Örneğin tahmin değeri 0.9 olan bir model tahmin değeri 0.5 modelin orta noktaları birleştirilirken oluşturulacak olan yeni nokta tahmin değeri yüksek olan yani 0.9 olan değere daha yakındır.

Model1 orta nokta = 30,30 Tahmin Değeri = 0,9

Model2 orta nokta = 30,30 Tahmin Değeri = 0,3

$$Yeni\ merkez\ nokta\ \mathrm{x} = \frac{model1\ x*model1\ tahmin\ \ +model2\ x*model2\ tahmin}{model1\ tahmin+model2\ tahmin}$$

$$Yeni\ merkez\ nokta\ y = \frac{model1\ y*model1\ tahmin\ \ +model2\ y*model2\ tahmin}{model1\ tahmin+model2\ tahmin}$$

$$Yeni\ merkez\ nokta\ x = \frac{(30*0{,}9)+(10*0{,}3)}{0{,}9+0{,}3} = \frac{30}{1{,}2} = 25$$

$$Yeni\ merkez\ nokta\ y = \frac{(30*0{,}9)+(10*0{,}3)}{0{,}9+0{,}3} = \frac{30}{1{,}2} = 25$$

Hesaplamalar sonucunda yeni merkez nokta 25,25 noktası olmuştur. Hesaplanan noktanın boyut değeri ise modellerin boyutlarının ortalaması olarak güncellenecektir.

3. Sonuçlar ve İnceleme (30 puan)

Çalışma kapsamında farklı CNN modellerini bireysel olarak eğitip elde etmiş oldukları tahmin sonuçlarını birleştirmeyi hedeflemekteyiz. Bu amaçla şimdiye kadar YOLOv5s, YOLOv5m ve Faster R-CNN modellerinin eğitimini tamamladık. Tablo 2'de YOLOv5s ve YOLOv5m modellerine ait test sonuçlarına yer verilmiştir.

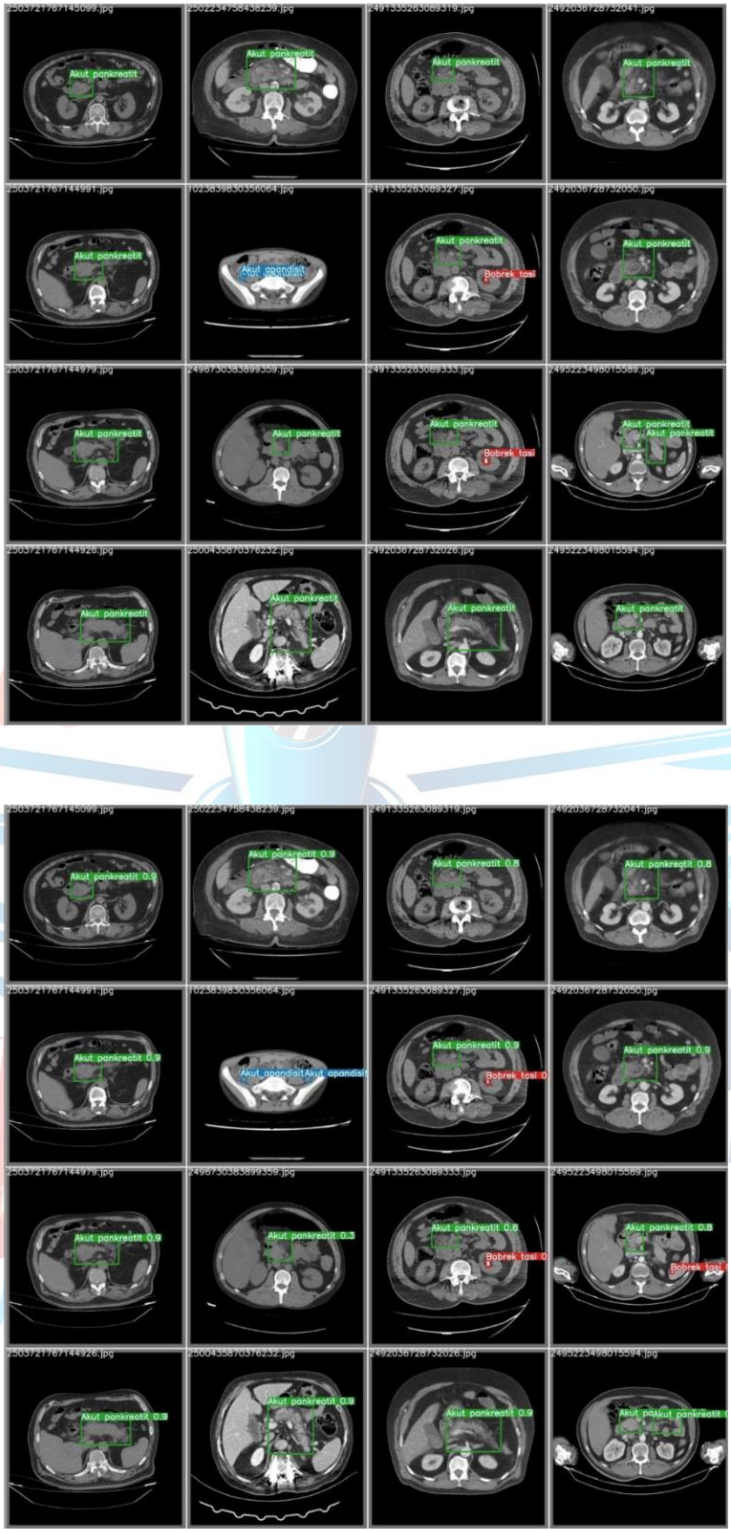
Sınıflar	Görüntü Sayısı	Hedef Sayısı	Precision-YOLOv5s	Precision-YOLOv5m	Recall-YOLOv5s	Recall-YOLOv5m	mAP-YOLOv5s	mAP-YOLOv5m
Tüm etiketler	2910	3100	0.83	**0.84**	**0.95**	0.94	**0.95**	0.94
Akut apandisit	2910	583	0.69	**0.72**	**0.91**	0.90	**0.92**	0.91
Akut kolesistit	2910	502	0.91	**0.92**	0.98	**0.99**	0.98	**0.99**
Akut pankreatit	2910	677	0.83	**0.85**	0.97	0.97	0.97	**0.98**
Böbrek taşı	2910	296	**0.73**	0.72	0.89	**0.90**	0.87	0.87
Akut divertikulit	2910	121	0.84	0.84	**0.94**	0.91	**0.94**	0.91
Abdominal aort diseksiyonu	2910	925	0.96	**0.97**	0.99	0.99	0.99	0.99

Tablo 2: YOLOv5s ve YOLOv5m Modellerine Ait Test Sonuçları

Tablo 2 incelendiğinde tüm sınıflar içi YOLOv5s modelinin mAP değerinin, Recall değerinin 0.95 ile en daha başarılı olduğu görülmektedir. Bunun yanında tüm sınıflar için Precision değerinin YOLOv5m de daha başarılı olduğu görülmüştür. Bunun yanında sınıf bazında Recall değerleri incelendiğinde YOLOv5s'in Akut apandisit, Akut divertikülit tespit etmede daha başarılı olurken Yolov5-m modeli Akut kolesistit, Böbrek taşı tespit etmede daha başarılı olduğu görülmüştür. Abdominal aort diseksiyon ve Akut pankreatit tespitinde ise iki modelin eşit oranda Recall değeri elde ettiği görülmektedir. Buradan yola çıkarak modellerin farklı hastalık türlerinde göstermiş olduğu başarı oranlarının farklı olması önerdiğimiz yöntemin başarılı olacağı tezinin güçlendirmektedir. Şekil 7'de YOLOv5m modeline ait tahminler ve gerçek etiketler paylaşılmıştır. Eğitim esnasında karşılaştığımız en önemli sorunlardan biri donanım kısıtlarıydı. Bu yüzden eğitim sürelerimiz çok uzadı ve eğitmeyi planladığımız modellerden bazıları için planlanan plandan geri kaldık.

(a)

(b)

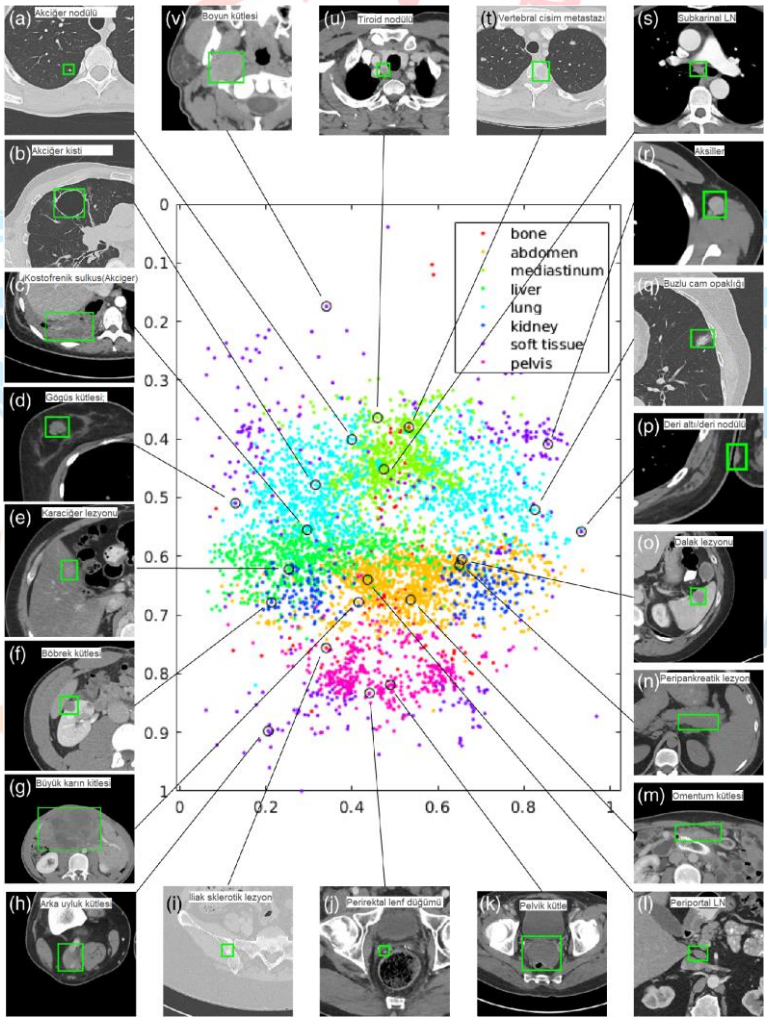


Şekil 7: Gerçek Etiketler (a) İlk 16 Görüntü, (b) YOLOv5m Tarafından etiketlenmiş görüntüler

4. Deney ve eğitim aşamalarında kullanılan veri setleri (15 puan)

DeepLesion veri seti

DeepLesion [1] veri seti tıbbı vakadan toplanmış kamu erişimine açık bir veri setidir. Bu veri seti 4,427 farklı vakadan 10.594 bilgisayarlı tomografi sonuçlarından elde edilen 32,120 yatay dilimde 32.735 lezyon içerir. Tüm yatay dilimler 1 ile 3 arasında lezyon barındırır. Bu veri seti içerisindeki görseller genellikle küçük boyutlara sahiptir(area ≤ 32x32 pixel). Bundan dolayı obje tespiti zorlaşır. DeepLesion veri setinin önemli bir özelliği, geniş çapta incelenen akciğer nodüllerinden karaciğer lezyonlarına, böbreklere, lenf düğümlerine kadar her türlü kritik radyoloji bulgusunu içermesidir. (Şekil 8) DeepLesion veri seti farklı görüntü modellerinden oluşur bu modellerin dağılımı şu şekildedir; %59,53 CT, %21,64 MR, %15,51 NM, %1,39 CR, %1,27 PT, %0,39 US, %0,28 diğer modellerdir. Modeller içerisinde yer alan görüntüler ağırlıklı olarak 512x512 resim boyutuna sahiptir. [1].



Şekil 8: DeepLesion Veri Setinin %15'lik dilimin görselleştirilmesi [1]

5. Referanslar (5 puan)

[1] Yan, K., Wang, X., Lu, L., & Summers, R. M. (2018). DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning. *Journal of medical imaging*, *5*(3), 036501.

[2] Carr, A., Weber, E. S., Murphy, C. J., & Zwingenberger, A. (2014). Computed tomographic and cross-sectional anatomy of the normal pacu (Colossoma macroponum). *Journal of Zoo and Wildlife Medicine*, *45*(1), 184-189.

[3] https://github.com/ultralytics/yolov5

[4] 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*.

[5] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In *European conference on computer vision* (pp. 21-37). Springer, Cham.

[6] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2980-2988).

[7] Yan, K., Tang, Y., Peng, Y., Sandfort, V., Bagheri, M., Lu, Z., & Summers, R. M. (2019, October). MULAN: multitask universal lesion analysis network for joint lesion detection, tagging, and segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 194-202). Springer, Cham.

[8]Restrepo,Ronny(T.Y).Intersect over Union (IoU),Ronny Restrepo, http://ronny.rest/tutorials/module/localization_001/iou/