TEKNOFEST

HAVACILIK, UZAY VE TEKNOLOJİ FESTİVALİ

SAĞLIKTA YAPAY ZEKÂ YARIŞMALARI

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

PROJE DETAY RAPORU

TAKIM ADI: AbdoMate

TAKIM ID: 468888

MANİSA 2022

İÇİNDEKİLER

İÇİNDEKİLER i

ŞEKİL LİSTESİ ii

1. PROJE MEVCUT DURUM DEĞERLENDİRMESİ 1

2. ÖZGÜNLÜK 1

3. SONUÇLAR VE İNCELEME 2

4. DENEY VE EĞİTİM AŞAMALARINDA KULLANILAN VERİ SETLERİ 6

5. REFERANSLAR 7

ŞEKİL LİSTESİ

Şekil 3.1 mAP 0.5 ve mAP 0.5:0.99 Metrikleri 3
Şekil 3.2 Score Metriği 3
Şekil 3.3 Parametre Ayarlama Eğitimleri Sonuçları - Tüm Döngüler 4
Şekil 3.4 Parametre Ayarlama Eğitimleri Sonuçları – Nesiller 4
Şekil 3.5 Tam Eğitim Sonuçları 5
Şekil 3.6 Tespit Sonuçları - 1 5
Şekil 3.7 Tespit Sonuçları - 2 5
Şekil 3.8 Tespit Sonuçları - 3 6
Şekil 3.9 Tespit Sonuçları - 4 6

1. PROJE MEVCUT DURUM DEĞERLENDİRMESİ

Proje sunuş raporu (PSR) sonrası genel sistem mimarisinde bir değişiklik yapılmamıştır. Yarışmada kullanılacak nesne tespit algoritması ile veri arttırma teknikleri (data augmentation) nihai olarak belli olamadığı için, PSR sonrası ilk iş olarak yarışmada kullanılacak nihai algoritma ve veri arttırma teknikleri belirlenmiştir.

Algoritma seçiminin ilk aşamasında PSR içerisinde Bölüm 3'te bahsedilen algoritmalar yarışmada kullanılacak takım bilgisayarına kurulmuş ve değişik yapılandırmalarla (örneğin farklı görüntü çözünürlükleri ile) bu bilgisayar üzerindeki görüntü işleme hızları karşılaştırılmıştır. Bu testler sonucunda YOLOv5 algoritmasının l6 modelinin 512*512 çözünürlük ile yarışmadaki süre gereksinimlerini karşılayarak en yüksek doğruluk oranına ulaşacak algoritma olduğu görülmüştür. Bu nedenle nihai algoritma ve model olarak YOLOv5l6 512 seçilmiştir.

Nihai algoritma ve model seçildikten sonra Bölüm 4'teki veri seti içerindeki DICOM formatındaki tıbbi görüntülerin gerekli görülenleri JPEG formatına dönüştürülmüştür. Çalışmaların sonraki aşamasında ise farklı veri arttırma tekniklerinin parametrelerinin eğitilen modelin doğruluk değerlerine etkisini incelemek için JPEG formatına döndürülmüş veri setinin %10'luk bir kısmı kullanılarak eğitimler yapılmıştır. Bu eğitimlerin her 10 döngüsünde (epoch), Bölüm 2'de bahsedilen veri artırma tekniklerinin parametreleri değiştirilmiştir. Bu şekilde 210 döngü eğitim yapılmış ve veri arttırma tekniklerinin nihai parametrelerine ulaşılmıştır. Bu parametreler ile modelin ulaşabileceği en yüksek doğruluk oranına hızlı ve güvenli bir şekilde yakınsayacağı öngörülmektedir.

Projenin şu anki güncel aşamasında yukarıda belirtilen model ve karar kılınan parametreler ile eğitimler devam ettirilmektedir. Bu eğitimler Google Colab üzerinde Nvidia P100 ekran kartları kullanılarak yapılmaktadır. Çünkü takım bilgisayarları görüntü işleme tarafında oldukça yavaş kalmaktadır. Şu aşamada yapılan eğitimlerde takım açısından yarışmaya yönelik oldukça güzel ve daha iyisi için umut vadeden sonuçlar elde edilmektedir.

Bölüm 3'te burada anlatılan testlerin önemli görülen sonuçlarına ve grafiklerine yer verilmiştir.

2. ÖZGÜNLÜK

Takımın özgünlüğe katkısı çoğunlukla; yarışmada kullanılacak modele sunulan verilerin ön işlemlerden geçirilmesi ve bu algoritmaların parametrelerinin ayarlanması şeklinde olacaktır. Çünkü yarışmadaki takımların çoğunluğunun benzer algoritmalar kullanacağı düşünüldüğünde geriye sadece veri setlerinin iyileştirilmesi ve algoritmaların en iyi performansı verecek şekilde ayarlanması kalmaktadır. Ayrıca yarışmanın takvimi, takım üyelerinin bilgi seviyesi ve yarışmadaki rekabetçi ortam düşünüldüğünde yarışmanın amacına yönelik yeni bir algoritma tasarlanması veya mevcut olanların düzenlenmesi gibi ağır çalışma isteyen girişimlerin çok uygulanabilir olmadığı düşünülmektedir.

Eğitim sırasında kullanılan görüntüler üzerinde farklı veri arttırma teknikleri kullanılmaktadır ve bu teknikler yarışma esnasında da kullanılacaktır. Eğitim sırasında [29]'daki kütüphane içerisinde bulunan veri arttırma tekniklerden faydalanılmaktadır. Kullanılan teknikler aşağıda sıralı bir şekilde verilmiş ve kısaca açıklanmıştır:

HueSaturationValue: Giriş görüntüsünün, rastgele olarak renk tonunu (Hue), renk doygunluğunu (Saturation) ve ışık değerini (Value) değiştirmektedir.

GaussianBlur: Giriş görüntüsüne rastgele bir Gauss filtresi uygulamaktadır.

GaussNoise: Giriş görüntüsüne rastgele olarak Gauss gürültüsü uygulamaktadır

RandomBrightnessContrast: Giriş görüntüsünün, rastgele olarak parlaklığını (Brightness) ve kontrastını değiştirmektedir.

CLAHE: Giriş görüntüsüne sınırlı uyarlanabilir histogram eşitleme (Limited Adaptive Histogram Equalization) uygulamaktadır.

Yukarıdaki veri arttırma teknikleri dışında YOLOv5 [19] içinde hazır bir şekilde bulunan veri arttırma teknikleri de vardır. Bunlar arasında eğitim sırasında kullanılanlar şöyledir:

hsv_h: Görüntünün renk tonu üzerinde değişiklik yapar.

hsv_s: Görüntünün renk doygunluğu üzerinde değişiklik yapar.

hsv_v: Görüntünün renk ışık değeri üzerinde değişiklik yapar.

degrees: Görüntünün açısını değiştirir.

translate: Görüntünün piksellerini kaydırır.

scale: Görüntüyü yeniden boyutlandırır.

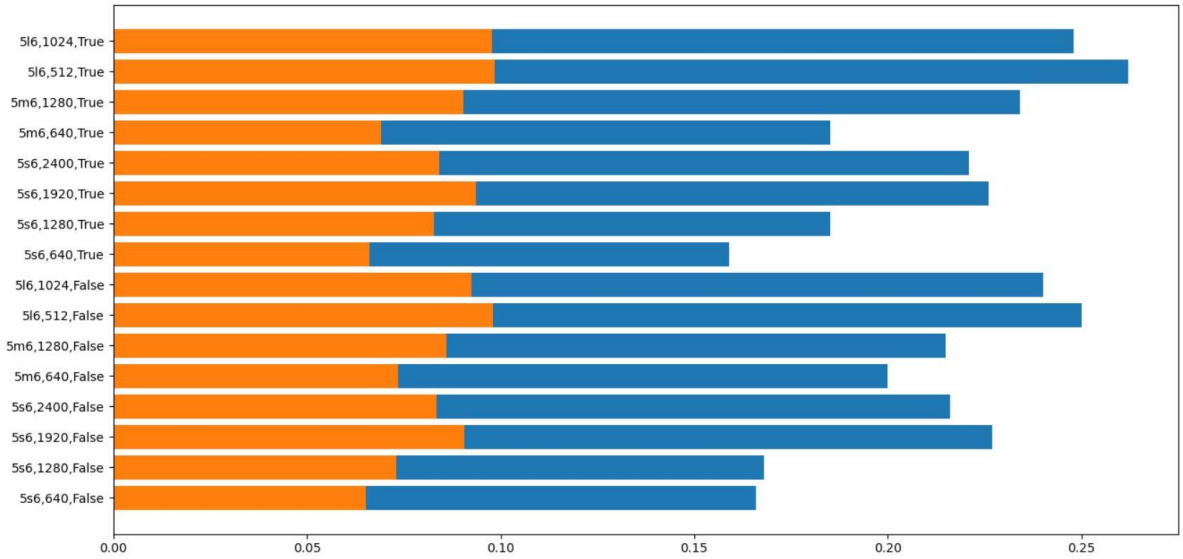
flipud: Görüntüyü yukarı-aşağı yönde çevirir.

fliplr: Görüntüyü sağ-sol yönde çevirir.

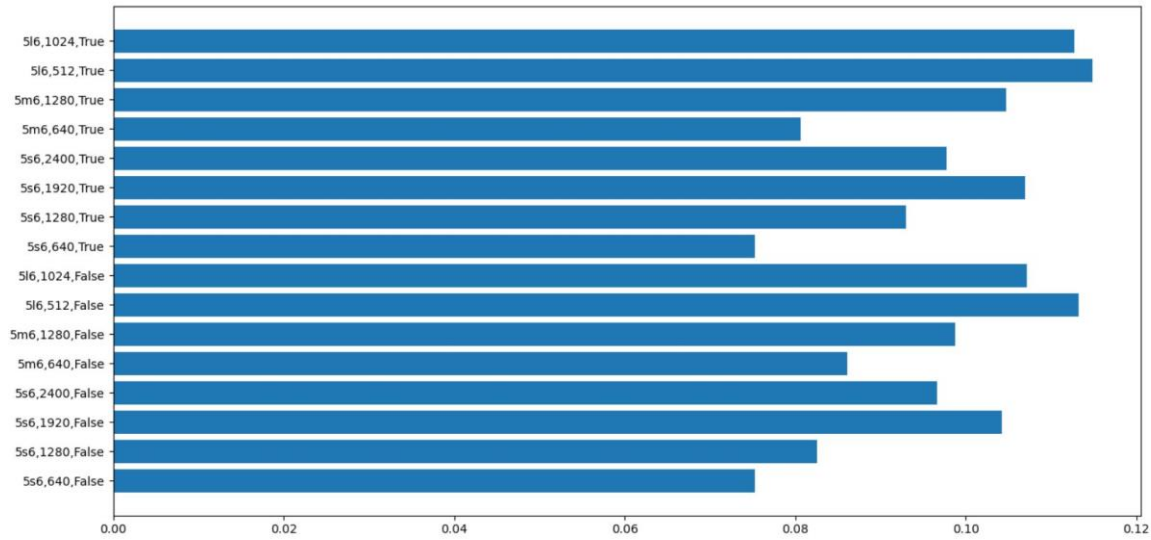
mosaic: Görüntü parçalara ayırıp bu parçaları tekrar rastgele bir şekilde birleştirir.

3. SONUÇLAR VE İNCELEME

Bölüm 1'de bahsedilen FPS testleri sonucunda, yarışma süresini verimli bir şekilde kullanarak takım bilgisayarında çalıştırılabilecek YOLOv5 modelleri ve yapılandırmaları sınırlıdır. Nihai modele ve yapılandırmaya karar verebilmek için bu seçenekler kısa süreli eğitimlere sokulmuştur. Eğitimler sırasında da Bölüm 4'te verilen veri setinin küçük bir kısmı kullanılmıştır. Bu eğitimlerin sonucunda Şekil 3.1 ve Şekil 3.2'deki grafikler elde edilmiştir. Şekil 3.1'de mavi yatay sütunlar "mAP 0.5" metriğinin ve turuncu olanlar "mAP 0.5:0.95" metriğinin değerini göstermektedir. Şekil 3.2'deki mavi yatay sütunlar ise "(0.1 * mAP 0.5 + 0.9 * mAP 0.5:0.95)" şeklinde hesaplanan "score" metriğinin değerlerini göstermektedir. Grafiklerin sol tarafında ise ilgili satırdaki sonucun hangi modele ve yapılandırmaya ait olduğu görülmektedir. Örneğin Şekil 3.1'de birinci satırdaki "5l6, 1024, True", o satırdaki sonuçların YOLOv5l6 modeli ile 1024 *1024 çözünürlükte ve veri arttırma ile yapılan ("True" yazması veri arttırma kullanıldığını göstermektedir. "False" yazanlarda veri arttırma yoktur.) eğitime ait olduğunu göstermektedir. İki grafikten de açıkça görüldüğü üzere YOLOv5l6 modeli ile 512*512 çözünürlük ve veri arttırma kullanılarak yapılan eğitimlerin elde ettiği doğruluk değerleri daha yüksektir. Bu nedenle yarışmada bu model ve yapılandırma kullanılacaktır.

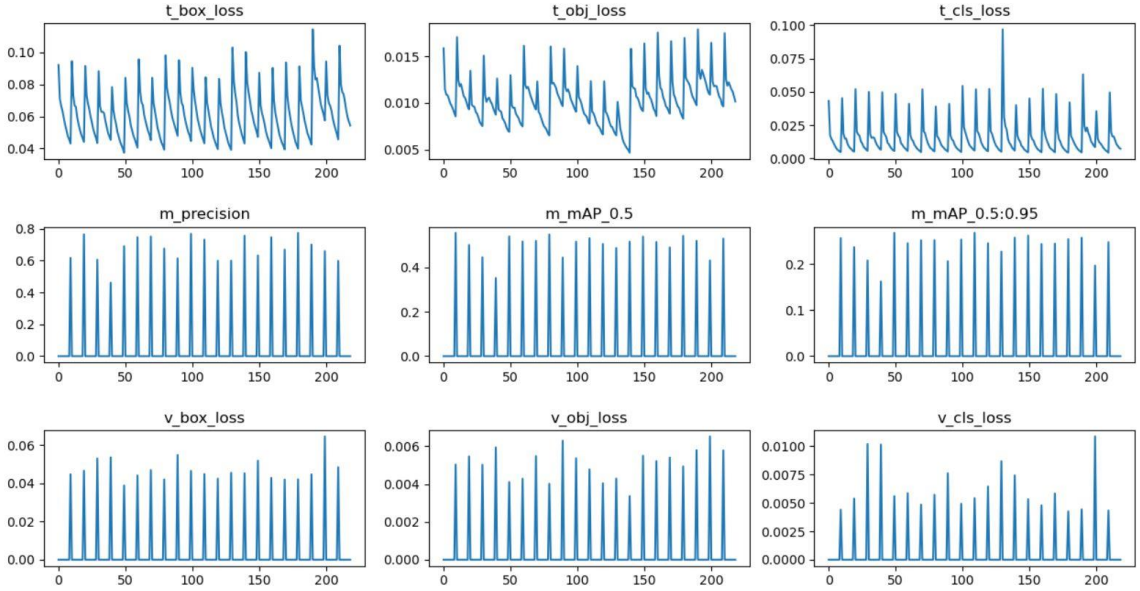


Şekil 3.1 mAP 0.5 ve mAP 0.5:0.99 Metrikleri

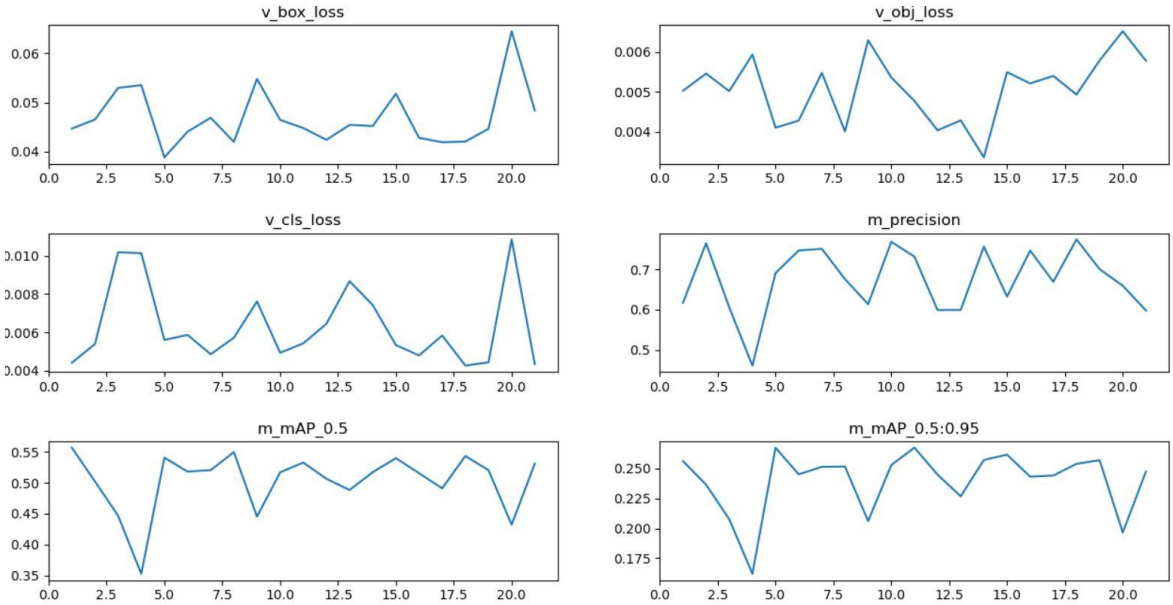


Şekil 3.2 Score Metriği

Nihai model ve yapılandırmaya (model) karar verildikten sonra, bu model üzerinde Bölüm 1'de de bahsedilen farklı veri arttırma tekniklerinin parametreleri her 10 döngüde bir değiştirilerek 210 döngü yapılmıştır. Şekil 3.3 ve 3.4'te bu eğitimin sonuçları gösterilmiştir. Şekil 3.3'te yatay eksen döngü (epoch) sayısını göstermektedir. Şekil 3.4'teki yatay eksen ise nesil (her 10 döngü eğitim sonucunda ulaşılan model ağırlıkları, generation) numarasını göstermektedir. Grafikte de görüldüğü üzere sonuçlar birbiriyle ne kadar benzer olsa da 18. nesildeki yapılandırma daha kararlı ve doğru sonuç vermektedir. Bu neden şu anki eğitimlere bu yapılandırmadaki veri arttırma parametreleri ile devam edilmektedir.



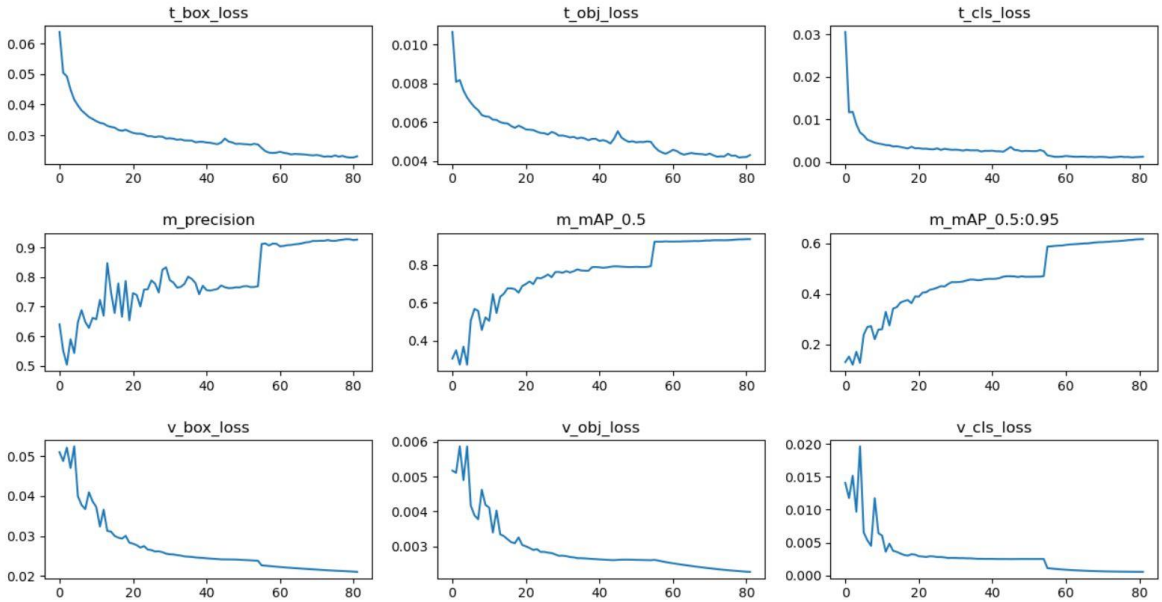
Şekil 3.3 Parametre Ayarlama Eğitimleri Sonuçları - Tüm Döngüler



Şekil 3.4 Parametre Ayarlama Eğitimleri Sonuçları – Nesiller

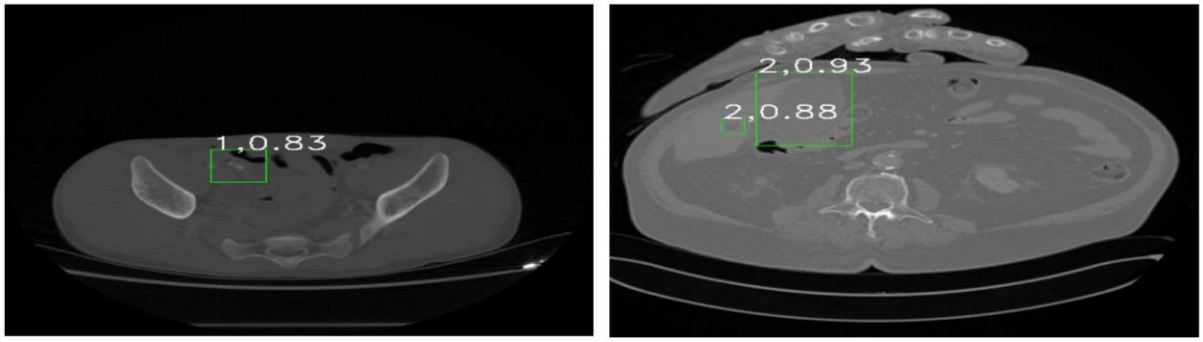
Şekil 3.5'te şu ana kadar yapılan eğitimlerin sonuçları grafik şeklinde verilmiştir. Veri setindeki etiketler arasında ve yarışmada tespit edilmesi beklenilen hastalıkların bazı vakalarında görülen fazladan 3 adet farklı klinik durum bulunmaktadır. Bunlar "Apendikolit", "Kalsifiye divertikül" ve "Safra kesesi taşı" şeklindedir. Yarışma esnasında bu durumların tespiti beklenmemektedir. Eğitimin 55. döngüsüne kadar bu etiketler de eğitime dahil edilmiştir. Fakat bu klinik durumlara ait örnek sayısı az olduğu için modelin bunları öğrenmesi zor olmaktadır ve bu durum test veri seti üzerinden hesaplanan doğrululuk değerlerinin yarışmaya yönelik hatalı, olması gerekenden daha düşük çıkmasına sebep olmaktadır. Bu nedenle eğitimin 56. ve devamındaki döngülerinde yukarıda verilen 3 adet klinik duruma ait etiket verileri veri setinden çıkartılmıştır. Eğitimlere

bu etiketler kullanılmadan devam edilmektedir. Şekil 3.5'teki grafik üzerindeki 55. döngüden 56. döngüye geçişteki ani sıçrayışların nedeni budur. Şu anki aşamada mAP0.5 metriğinde %93.75, mAP0.5:0.95 metriğinde ise %61.6 doğruluk oranına ulaşılmıştır. Grafikte de görüldüğü üzere bu değerler şu an için bir doyum noktasına ulaşmamıştır ve hatalar (loss fonksiyonları) düşmeye, doğruluk değerleri artmaya devam etmektedir.

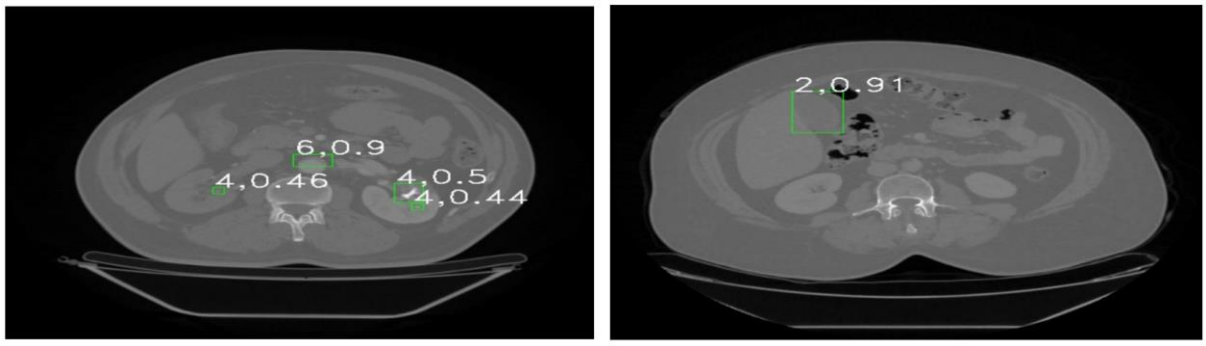


Şekil 3.5 Tam Eğitim Sonuçları

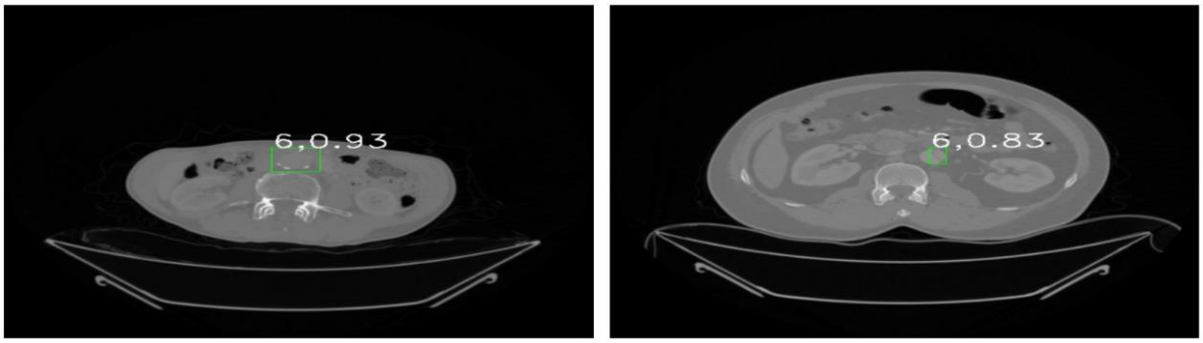
En son ulaşılan ağırlıklar ile modelin, test verisi içerisinden seçilmiş rastgele 8 görüntü üzerinde yaptığı tahminler Şekil 3.6, 3.7, 3.8 ve 3.9'da verilmiştir.



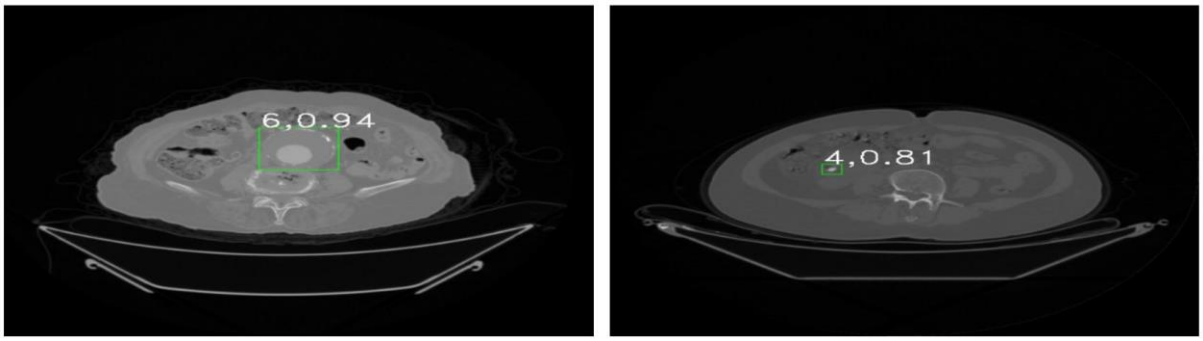
Şekil 3.6 Tespit Sonuçları - 1



Şekil 3.7 Tespit Sonuçları - 2



Şekil 3.8 Tespit Sonuçları - 3



Şekil 3.9 Tespit Sonuçları - 4

4. DENEY VE EĞİTİM AŞAMALARINDA KULLANILAN VERİ SETLERİ

Yarışmaya yönelik yapılan nesne tespit algoritması eğitimleri sadece T.C. Sağlık Bakanlığı tarafından sağlanan tıbbı görüntüler ve bu görüntülerin etiket verileri kullanılarak yapılmaktadır. Bu veri setinin içerisindeki etiketli görüntüler ayıklandığında 30 bin civarında farklı etiketlenmiş görüntü elde edilmektedir. Bu görüntü sayısının iyi bir eğitim için yeterli olduğu düşünülmüş ve bu nedenle farklı bir veri seti arayışına gidilmemiştir. Bölüm 3'te verilen eğitim sonuçları incelendiğinde 30 bin etiketli görüntünün iyi bir eğitim için yeterli olduğu görüşünün doğru olduğu görülebilir. Bu veri setinin %80'lik kısmı eğitim, %10'luk kısmı doğrulama (validation) ve kalan %10'luk kısmı da test verisi olarak kullanılmaktadır.

5. REFERANSLAR

[1] A. Robicquet, A. Sadeghian, A. Alahi, S. Savarese, “Learning Social Etiquette: Human Trajectory Prediction In Crowded Scenes” , European Conference on Computer Vision (ECCV), 2016.
[2] Emmanuel Maggiori, Yuliya Tarabalka, Guillaume Charpiat and Pierre Alliez. “Can Semantic Labeling Methods Generalize to Any City? The Inria Aerial Image Labeling Benchmark”. IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2017.
[3] Ding, Jian and Xue, Nan and Xia, Gui-Song and Bai, Xiang and Yang, Wen and Yang, Michael and Belongie, Serge and Luo, Jiebo and Datcu, Mihai and Pelillo, Marcello and Zhang, Liangpei, “Object Detection in Aerial Images: A Large-Scale Benchmark and Challenges”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021
[4] Xia, Gui-Song and Bai, Xiang and Ding, Jian and Zhu, Zhen and Belongie, Serge and Luo, Jiebo and Datcu, Mihai and Pelillo, Marcello and Zhang, Liangpei, “DOTA: A Large-Scale Dataset for Object Detection in Aerial Images”, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018
[5] Jian Ding, Nan Xue, Yang Long, Gui-Song Xia, Qikai Lu, “Learning RoI Transformer for Detecting Oriented Objects in Aerial Images”, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019
[6] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”, 2015
[7] Ross Girshick, “Fast R-CNN”, 2015
[8] Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation”, 2013
[9] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection”, 2020
[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, “Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition”, 2015
[11] https://it-qa.com/what-is-cspdarknet-53/ (Erişim 15 Haziran 2022)
[12]https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detectionalgorithms-36d53571365e (Erişim 1 Temmuz 2022)
[13] https://cvgl.stanford.edu/projects/uav_data/ (Erişim 1 Temmuz 2022)
[14] http://dronedataset.icg.tugraz.at/ (Erişim 1 Temmuz 2022)
[15] https://project.inria.fr/aerialimagelabeling/ (Erişim 1 Temmuz 2022)
[16] https://captain-whu.github.io/DOTA/dataset.html (Erişim 1 Temmuz 2022)
[17] https://github.com/rbgirshick/py-faster-rcnn (Erişim 1 Temmuz 2022)
[18] https://docs.ultralytics.com/ (Erişim 1 Temmuz 2022)
[19] https://github.com/ultralytics/yolov5 (Erişim 1 Temmuz 2022)
[20] https://github.com/facebookresearch/detectron2 (Erişim 1 Temmuz 2022)
[21] https://github.com/facebookresearch/Detectron/ (Erişim 1 Temmuz 2022)
[22] https://github.com/facebookresearch/maskrcnn-benchmark/ (Erişim 1 Temmuz 2022)
[23] https://www.kaggle.com/competitions/tensorflow-great-barrier-reef (Erişim 1 Temmuz 2022)
[24] https://www.kaggle.com/c/tensorflow-great-barrier-reef/discussion/307878(Erişim 1 Temmuz 2022)
[25] https://www.kaggle.com/c/tensorflow-great-barrier-reef/discussion/307760(Erişim 1 Temmuz 2022)
[26] https://www.kaggle.com/c/tensorflow-great-barrier-reef/discussion/307707 (Erişim 1 Temmuz 2022)

[27] https://chowdera.com/2021/07/20210724150724572z.html (Erişim 1 Temmuz 2022)
[28] https://ashutoshmakone.medium.com/faster-rcnn-502e4a2e1ec6 (Erişim 1 Temmuz 2022)
[29] https://github.com/albumentations-team/albumentations (Erişim 1 Temmuz 2022)