



ARAŞTIRMA MAKALESİ / RESEARCH ARTICLE

X-Işını Görüntüleri Kullanılarak Omuz İmplantlarının Tespiti ve Sınıflandırılmasında YOLOv5 Modellerinin Performanslarının İncelenmesi

Investigation of The Performance of YOLOv5 Models In The Detection and Classification of Shoulder Implants Using X-Ray Images

Elif Baykal Kablan 

Karadeniz Teknik Üniversitesi Mühendislik Fakültesi, Yazılım Mühendisliği, Trabzon, TÜRKİYE
 Sorumlu Yazar / Corresponding Author : ebaykal@ktu.edu.tr

Öz

Total Omuz Artroplastisi (TOA), omuz eklemindeki hasara bağlı ağrı ve hareket kısıtlanması yaşayan hastalarda omuzun hasar görmüş kısmının protez implant ile değiştirildiği cerrahi işlemdir. Zamanla protez implantların aşınması veya zarar görmesi durumunda, protez üreticisi ve model bilgisine ihtiyaç duyulur, ancak tıbbi kayıtlardaki eksiklikler nedeniyle bu bilgiler genellikle belirsiz olabilmektedir. Bu nedenle, uzman kişiye bağımlılığı azaltarak hızlı ve doğru bir şekilde protez bilgilerini sağlayacak otomatik sistemlere ihtiyaç duyulmaktadır. Bu çalışmada, omuz implantı üreticilerinin X-ışını görüntüleri kullanılarak tespiti ve sınıflandırılması için otomatik bir sistem önerilmiştir. Sistemde YOLOv5n, YOLOv5s, YOLOv5m ve YOLOv5l modelleri kullanılarak implant baş bölgesi tespiti ve sınıflandırma işlemleri gerçekleştirilmiş ve elde edilen sonuçlar karşılaştırılmıştır. Veri dengesizliği nedeniyle bazı sınıfların diğerlerinden daha iyi temsil edildiği sınıf ağırlıklandırma tekniği de kullanılmıştır. Bu tekniğin modelin sınıflandırma performansını önemli ölçüde artırdığı tespit edilmiştir. Karşılaştırma sonucunda sınıf ağırlıklandırmalı YOLOv5l modelinin mAP@0.5, mAP@0.5:0.95, kesinlik, duyarlılık ve F1-ölçütü değerlerinde en yüksek performansı gösterdiği görülmüştür ve bu değerler sırasıyla 98.3, 85.9, %97.0, %95.7 ve %96.3 olarak elde edilmiştir. Önerilen YOLOv5l modeli, literatürdeki diğer hibrit makine öğrenmesi yaklaşımlarına göre düşük maliyetli ve yüksek performanslı bir çözüm sunmaktadır. Bu sayede, YOLOv5 modellerinin omuz implantlarının tespit ve sınıflandırılmasında değerli bir araç olarak etkinliği gösterilmiştir. Ayrıca, önerilen sistem sayesinde radyologlar ve cerrahlar için verimli ve güvenilir analiz süreçleri sunulması ve iş yükünün azaltılması beklenmektedir.

Anahtar Kelimeler: Derin Öğrenme, Total Omuz Artroplastisi, X-ray, YOLOv5

Abstract

Total Shoulder Arthroplasty (TSA) is a surgical procedure in which the damaged part of the shoulder is replaced with a prosthetic implant in patients with pain and restriction of movement due to damage to the shoulder joint. If prosthetic implants wear out or become damaged over time, prosthesis manufacturer and model information is needed, but this information can often be unclear due to deficiencies in medical records. Therefore, there is a need for automated systems to provide prosthesis information quickly and accurately, reducing reliance on specialists. In this study, an automatic system is proposed for the detection and classification of shoulder implant manufacturers using X-ray images. In the system, implant head region detection and classification processes were performed using YOLOv5n, YOLOv5s, YOLOv5m, and YOLOv5l models and the results obtained were compared. The class weighting technique was also used, where some classes are better represented than others due to data imbalance. It was found that this technique significantly improves the classification performance of the model. As a result of the comparison, it was observed that the YOLOv5l model with class weighting showed the highest performance in mAP@0.5, mAP@0.5:0.95, precision, sensitivity, and F-measure values, and these values were obtained as 98.3, 85.9, 97.0%, 95.7% and 96.3% respectively. The proposed YOLOv5l model offers a low-cost and high-performance solution compared to other hybrid machine learning approaches in the literature. Thus, the effectiveness of YOLOv5 models as a valuable tool in the detection and classification of shoulder implants has been demonstrated. Furthermore, the proposed system is expected to provide efficient and reliable analysis processes and reduce the workload for radiologists and surgeons.

Keywords: Deep Learning, Total Shoulder Arthroplasty, X-ray, YOLOv5

EXTENDED ABSTRACT

Introduction

The Shoulder, the most mobile joint in the human body, often requires surgical intervention for conditions like osteoarthritis and rheumatoid arthritis, leading to Total Shoulder Arthroplasty (TSA). The objective of TSA is pain alleviation and restoration of

normal shoulder functions. Surgical decisions, including prosthesis choice, are commonly guided by X-ray images. However, despite advancements in prosthetics, wear and loosening may necessitate periodic reviews or replacements. This paper addresses the challenge of prosthetic wear or damage,

emphasizing the importance of knowing the prosthesis manufacturer and model for efficient replacement. Current gaps in medical records often lead to uncertainties, requiring additional X-ray imaging and expert examinations. Automatic diagnostic systems are crucial for early interventions, and recent advances in deep learning, especially YOLOv5, have shown promising results in medical image processing.

The study conducts a comprehensive comparison of YOLOv5's nano, small, medium, and large versions for both implant detection and classification, surpassing recent studies in classification accuracy. Utilizing a publicly shared dataset, this work stands as the most extensive comparative analysis in the literature. By introducing YOLOv5 for joint implant classification from X-ray images, this study contributes to the growing body of research in deep learning applications for medical image analysis.

Materials and Methods

In this study, the investigation revolves around a publicly available dataset introduced by Urban et al. in 2020, featuring 597 X-ray images specifically focused on shoulder implants from four distinct manufacturers—Cofield (83 images), Depuy (294 images), Tornier (71 images), and Zimmer (149 images). The dataset presents several challenges, including variable and relatively low image resolution, diverse aspect ratios, and imbalances in class distribution. These challenges underscore the complexity of the problem. The dataset was chosen due to its recent introduction, open accessibility, and widespread adoption in the literature. A systematic split allocated 90% of the data to the training set and the remaining 10% to the testing set. A 5-fold cross-validation strategy was implemented, ensuring each class had representation across training and testing sets. For labeling, the Labelimg annotation tool was used to mark implant head regions with rectangular bounding boxes, aligning with YOLO algorithms. YOLOv5, with its four variants (n, s, m, and l), served as the primary deep learning framework, each trained for 400 epochs with consistent hyperparameters. The study concludes with a comparative analysis of YOLOv5 variants' performance,

providing insights into their effectiveness in shoulder implant detection.

Results and Discussion

The proposed YOLOv5-based approach for shoulder implant detection and classification is evaluated against state-of-the-art methods, encompassing all its sub-versions. Utilizing transfer learning for efficient training, each YOLOv5 model is fine-tuned on a dataset of X-ray images containing shoulder implants, leveraging pre-trained weights from the COCO dataset. The models undergo 400 epochs with specific hyperparameters, including a batch size of 8, momentum of 0.9, decay of 0.0005, and a learning rate of 0.001. Quantitative assessments, employing metrics such as precision, recall, F1-score, and mAP, highlight the superior performance of the YOLOv5l model with detection metrics of 98.7% mAP@0.5, 85.7% mAP@0.5:0.95, 93.4% precision, 94.6% recall, and 93.9% F1-score, as presented in Table 3. The convergence of precision and recall to 100% during training and validation is visualized in Figure 5. Qualitatively, true positive detection results of the YOLOv5l model on selected test images are illustrated in Figure 6. A comprehensive comparison with previous studies on the same dataset is provided in Table 4, affirming the proposed YOLOv5 model's efficacy for accurate shoulder implant detection and classification.

Conclusion

The YOLOv5 models achieve a low-cost and high-performance objective compared to other hybrid machine learning approaches in the literature, showcasing their effectiveness in medical imaging applications like shoulder implant detection and classification. The models are envisioned to offer an efficient and reliable analysis process for radiologists and surgeons, potentially alleviating their workload. However, a notable limitation is the limited diversity in implant manufacturer classes within the dataset, providing an opportunity for expansion by increasing this diversity. Moreover, exploring model performance enhancements through testing on larger datasets and diverse imaging conditions remains a promising avenue for future research.

1. Giriş

Omuz vücuttaki tüm eklemler arasında en yüksek hareketliliğe sahip karmaşık yapı bir eklemdir. Total Omuz Artroplastisi (TSA), omuz ekleminde özellikle kireçlenme ve romatoid artrit hastalıkları nedeniyle meydana gelen hasara bağlı ciddi ağrısı ve hareket kısıtlanması olan hastalarda uygulanan bir ameliyat türüdür. Bu ameliyatın amacı ağrıyı gidermek ve hastaya normal omuz fonksiyonlarını yeniden kazandırarak yaşam kalitesini artırmaktır [1-3]. TSA ameliyatı ile fonksiyonunu yitirmiş ve istenen potansiyelde çalışmayan eklem çıkarılarak yerine protez implant yerleştirilmektedir. Ameliyat sırasında hangi protezin yerleştirileceğine karar vermek için X-ışını görüntüleri kullanılmaktadır. Günümüzde birkaç farklı protez üreticisi olmakla birlikte her bir üretici vakaya ve hastaya özgü birkaç farklı model sunmaktadır. Mevcut kullanılan protez ürünleri, malzemeleri ve cerrahi teknikler oldukça gelişmiş olsa da protezler zamanla aşınabilmekte ya da gevşeyebilmektedir. Dolayısıyla cerrahi işlem uygulandıktan sonra protez durumunun belirli zaman aralıklarıyla tekrar gözden geçirilmesi ve gerektiği takdirde yeniden yerleştirilmesi gerekmektedir [4]. Ayrıca kaza gibi bazı durumlarda mevcut protez zarar görmüş ise değiştirilmektedir. Protezin iyileştirilmesi veya değiştirilmesi durumunda sürecin hızlı ve başarılı bir şekilde ilerlemesi

açısından protez üreticisi ve model bilgisi hasta veya doktor tarafından bilinmelidir. Ancak bu gibi durumlarda genellikle tıbbi kayıtlardaki eksiklikler sebebiyle belirsizlik durumu ortaya çıkmaktadır. Dolayısıyla hastadan tekrar X-ışını görüntüsü alınarak tıbbi uzman tarafından detaylı bir inceleme ile üretici ve model bilgisi doğrulanmaktadır. Bu incelemenin uzman kişiye bağlı, zaman alıcı ve yorucu olmasından dolayı otomatik sistemlere ihtiyaç duyulmaktadır.

Otomatik tanı sistemleri erken tedavi ve cerrahi işlem ile protezde daha fazla bozulmayı önlemek açısından kritik öneme sahiptir. Son zamanlarda, derin öğrenme yöntemleri geniş bir yelpazede tıbbi görüntü işleme görevlerinde yüksek performans sergilemektedir. Örneğin, derin sinir ağları, kanser teşhisi için meme kanseri, akciğer kanseri, beyin tümörü gibi çeşitli tıbbi görüntü sınıflandırma görevlerinde etkileyici sonuçlar elde etmiştir. Derin öğrenme yöntemleri, organ segmentasyonu, lezyon tespiti ve yoğunluk tahmini gibi tıbbi görüntü analizi alt alanlarında da başarılı olmuştur. Derin öğrenme modelleri, büyük miktarda veriyi etkili bir şekilde öğrenerek karmaşık ilişkileri ve özellikleri yakalayabilme yeteneğine sahiptir. Derin sinir ağlarının katmanlı yapısı, özelliklerin hiyerarşik olarak çıkarılmasını sağlar ve bu da tıbbi görüntülerdeki anatomik yapıların ve patolojik bölgelerin daha iyi temsil edilmesine

yardımcı olur. Ayrıca, derin öğrenme modelleri, veri artırma teknikleri ve transfer öğrenme gibi yöntemlerle de desteklendiğinde daha da güçlenir. Bu sayede, sınırlı veriye sahip tıbbi görüntü veri setleri üzerinde bile yüksek performans elde edilebilir. Sonuç olarak, derin öğrenme yöntemleri, tıbbi görüntü işleme alanında etkileyici sonuçlar veren güçlü bir araç haline gelmiştir. Daha detaylı bir tartışma ve karşılaştırma için okuyucular, araştırma makalelerine başvurabilirler [5-6]. Literatürde X-ışını görüntülerinden omuz implantlarının tespiti ve sınıflandırılması amacıyla önerilen derin öğrenme tabanlı birkaç çalışma mevcuttur [4, 7-12].

Urban vd. [7] omuz implantlarını sınıflandırmak amacıyla 6 farklı derin öğrenme tabanlı CNN mimarisini (VGG-16, VGG-19, ResNet-50, ResNet-152, NASNet, DenseNet-201) uygulamışlardır. Veri seti, dört farklı protez üreticisine ait 16 farklı modelin kullanıldığı 597 adet omuz implantı X-ışını tarama görüntüsünden oluşmaktadır. CNN mimarilerinin ImageNet [13] gibi büyük boyutlu veri seti üzerinde önceden eğitilmesinin ve daha sonra omuz implantlarının X-ışını görüntülerini içeren bir veri setinde ince ayar yapılmasının iyi sonuçlar sağladığını belirtmişlerdir. En yüksek sınıflandırma performansını 10-kat çapraz doğrulama uygulayarak %80,4 doğrulukla NASNet [14] mimarisi ile elde etmişlerdir. Yi vd. [8] beş farklı implant modeli içeren toplam 482 adet omuz implantı X-ışını tarama görüntüsü kullanmışlardır. Tüm modelleri tek bir sınıflandırıcı ile eğitmek yerine sırasıyla her implant modeli için beş ayrı ResNet-152 mimarisini ikili sınıflandırıcı olarak eğitmişlerdir. 20-kat çapraz doğrulama uygulayarak Solar için 0,86'dan Zimmer için 1,0'a kadar değişen AUC-ROC değeri elde etmişlerdir. Yılmaz [9] literatürden farklı olarak en öne çıkan özellik filtrelerini seçmek amacıyla yeni bir kanal seçim katmanı (channel selection layer) kullanan çok kanallı bir CNN modeli önermiştir. Model her görüntü için kanallar arasında uygulanan etkili özellik seçimi ile doğruluk oranını önemli ölçüde artırmıştır. Önerilen yöntem %97,2 doğruluk oranı ile diğer çalışmalara göre daha iyi performans göstermiştir. Sultan vd. [10] omuz implantlarını sınıflandırmak amacıyla CNN ağ topluluğu mimarisine dayalı bir yöntem önermişlerdir. Rotasyonel olarak veri artırma uygulayarak eğitim veri setini 36 kat artırmışlardır. Modifiye edilmiş ResNet ve DenseNet ağ modellerini DRE-Net ağ topluluğu mimarisini oluşturmak için derinlemesine birleştirmişlerdir. DRE-Net mimarisi ile 10-kat çapraz doğrulama ile %85,92 doğruluk, %84,69 F1-ölçütü, %85,33 kesinlik ve %84,11 duyarlılık elde etmişlerdir. Efeoglu vd. [11] 3 farklı üreticiye ait omuz implantlarını sınıflandırmak amacıyla 12 farklı sınıflandırıcının performansını karşılaştırmıştır. 10-kat çapraz doğrulama uygulanan K-NN algoritmasının %74 doğruluk ile diğer algoritmalara göre daha iyi performans sağladığı görülmüştür. Bu sonuç derin öğrenme tabanlı yöntemlerin klasik makine öğrenmesi yöntemlerine göre oldukça başarılı olduğunu göstermiştir. Sivari vd. [4] derin öğrenme ve makine öğrenmesi algoritmalarının birleşiminden oluşan 10 farklı hibrit sınıflandırıcı model oluşturmuş ve istatistiksel olarak test etmişlerdir. Deneysel sonuçlara göre DenseNet201 + Lojistik Regresyon modeli kullanılarak %95,07 doğruluk sağlamışlardır. Karaci [12] diğer tüm yöntemlerden farklı olarak YOLOv3 nesne tespiti modeli ile omuz implantlarının baş bölgesini tespit ettikten sonra bu bölgeleri çeşitli CNN mimarilerine giriş olarak vermiştir. Sonuç olarak YOLOv3 nesne tespiti modeli ile implantın baş bölgesine odaklanmanın sınıflandırma doğruluğunu artırdığı görülmüştür. YOLOv3 ile DenseNet201 modelinin birlikte kullanımı ile %84,76 doğruluk elde etmişlerdir.

Nesne tespiti, birçok görüntü sınıflandırma tabanlı çalışmanın temel taşı olup bilgisayarlı görüntü alanında yoğun olarak

çalışılmaktadır. Nesne tespitinde temel amaç, verilen bir giriş görüntüsü içerisinde tespit edilmesi istenen nesnelerin konumunu sınırlayıcı kutu ile belirlemek ve hangi sınıfa ait olduklarını belirlemektir [15]. Literatürde bu amaçla önerilen yöntemler tek-adımlı ve iki-adımlı olmak üzere iki sınıf altında toplanmaktadır. Tek-adımlı yöntemler bölge önerilerinin oluşturulması ve bu bölgelerin sınıflandırılması olmak üzere iki adımda çalışmaktadır. SSD [16] ve YOLO [17] yöntemlerinde ise hem bölge önerme hem de sınıflandırma tek adımda gerçekleştirildiğinden tek-adımlı olarak adlandırılmaktadır. Bu özellik SSD ve YOLO yöntemlerinin hızlı çalışmasını ve böylece gerçek zamanlı uygulamalarda daha çok tercih edilmesini sağlamaktadır [18]. YOLO, R-CNN [19], Fast R-CNN [20] ve Faster R-CNN [21] gibi diğer popüler CNN tabanlı nesne algılama yönteminin algılama süresini azaltmayı amaçlamıştır. YOLO, hız ve doğruluk açısından güçlü performans sağlar ve gerçek zamanlı nesne algılama uygulamasında uygulanabilmektedir. Önerilen ilk YOLO yöntemi üzerinde yapılan gelişmeler sonucunda literatürde YOLOv2 (YOLO9000) [22], YOLOv3 [23], YOLOv4 [24] ve YOLOv5 [25] olmak üzere hızı ve doğruluğu yüksek gerçek zamanlı birçok nesne algılama yöntemi önerilmiştir. YOLO'nun tüm beş versiyonu birçok uygulamada yaygın olarak kullanılmış ve başarılı şekilde uygulanmıştır.

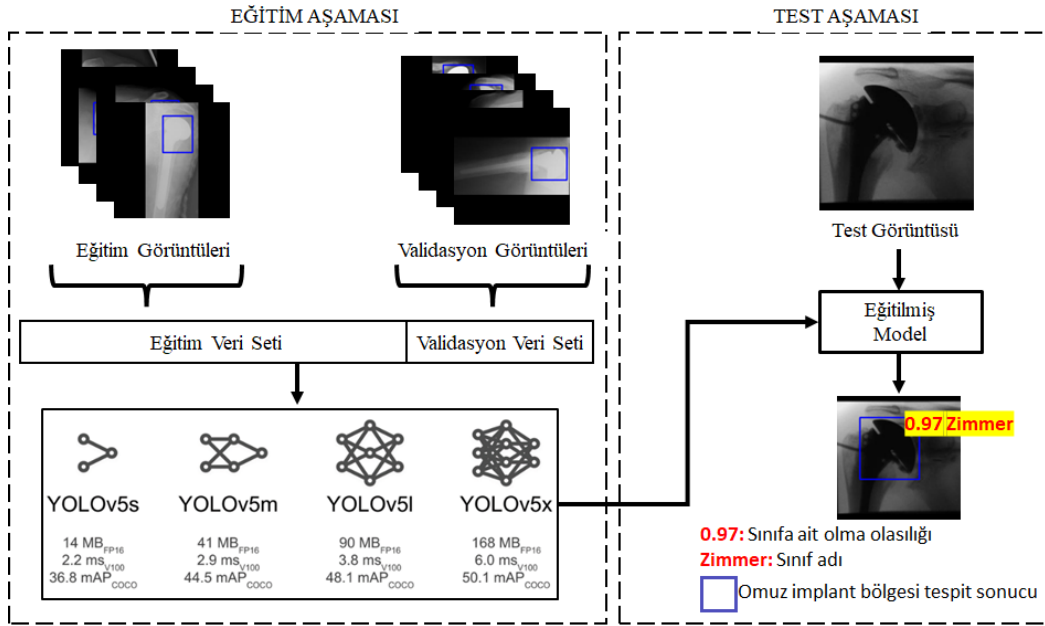
X-ışını görüntülerinden omuz implantı sınıflandırma alanında ise yalnızca Karaci vd. [12] YOLO'nun üçüncü sürümü olan YOLOv3 yöntemini implantın baş bölgesini tespit etmek amacıyla önermiştir. Sınıflandırma işlemi ise ayrıca tespit ettiği baş bölgelerini CNN mimarilerine giriş olarak vererek gerçekleştirmiştir. Bu çalışmada ise, hem implantın baş bölgesinin tespit edilmesi hem de sınıflandırılması amacıyla en güncel YOLOv5 yöntemi önerilmektedir. Önerilen çalışma kapsamında YOLOv5 modelinin YOLOv5n (nano), YOLOv5s (küçük), YOLOv5m (orta) ve YOLOv5l (büyük) olmak üzere dört alt versiyonu kullanılarak bu mimarilerin algılama ve sınıflandırma performansları karşılaştırılmıştır. YOLOv5 modelinin her versiyonu farklı model derinliklerine sahiptir ancak aynı ağ yapısı temel alınarak tasarlanmıştır. Çalışmada kullanılan omuz implant veri setinin dengesiz yapısı nedeniyle, sınıf ağırlıklandırma tekniği de kullanılmıştır. Sınıf ağırlıklandırma tekniği, veri dengesizliği nedeniyle bazı sınıfların diğerlerinden daha iyi temsil edilmesi problemini ele almayı amaçlamaktadır.

Çalışmanın ana katkıları aşağıdaki gibidir:

- X-ışını görüntülerinden omuz implantı sınıflandırmak amacıyla YOLO yaklaşımlarını hem algılama hem de sınıflandırma amacıyla öneren bilindiği kadarıyla ilk çalışmadır.
- Literatürde herkese açık paylaşılan veri setini kullanan diğer birçok güncel çalışmadan daha yüksek sınıflandırma performansı elde edilmiştir.
- YOLOv5n (nano), YOLOv5s (küçük), YOLOv5m (orta) ve YOLOv5l (büyük) olmak üzere dört farklı YOLOv5 versiyonu kullanılarak sonuçlar karşılaştırmalı olarak sunulmuş ve tartışılmıştır.
- Literatürde bu alanda yapılan diğer çalışmalarla bilindiği kadarıyla en kapsamlı karşılaştırmayı sağlayan çalışmadır.

Bu makalenin geri kalanı şu şekilde organize edilmiştir. Bölüm 2, ilk olarak çalışmada yararlanılan herkese açık X-ışını veri setini tanıtmaktadır. İkinci olarak X-ışını görüntülerinden omuz implantı tespiti ve sınıflandırılması için kullanılan YOLOv5 yöntemini alt versiyonlarıyla birlikte ayrıntılı olarak açıklamaktadır. Son olarak değerlendirme metriklerini tanıtmaktadır. Deneysel sonuçlar ve sonuçların ayrıntılı analizi

Bölüm 3'te verilmektedir. Son olarak, Bölüm 4'te sonuçlar ve gelecek çalışmalar için öneriler tartışılmaktadır.



Şekil 1. Önerilen sistemin iş akışı.

Figure 1. Workflow of the proposed system.

2. Materyal ve Metod

Önerilen sistemin iş akışı Şekil 1'de gösterilmektedir. İş akışı eğitim (eğitim ve validasyon) ve test olmak üzere iki aşamadan oluşmaktadır. Tüm deneyler, Intel(R) Core(TM) i9-11900K 3.50 GHz CPU ve NVIDIA GeForce RTX 3080 12GB GPU donanımlı bir bilgisayarda gerçekleştirilmiştir. Çalışmada yararlanılan herkese açık X-ışını veri seti, önerilen YOLOv5 yöntemi ve alt versiyonlarının mimarileri, bu mimarilerin eğitim, validasyon ve test aşamalarının ayrıntıları, sınıflandırıcı modifikasyonu ve performans değerlendirmesinde kullanılan ölçütler alt başlıklarda sunulmaktadır.

2.1. Veri kümesi tanıtımı

Çalışmada kullanılan herkese açık veri seti Urban vd. [7], omuz implantlarını içeren 597 X-ışını görüntüsü içermektedir. Veri setinde 83 Cofield, 294 Depuy, 71 Tornier, 149 Zimmer olmak üzere 4 farklı omuz implant üreticisine ait X-ışını görüntüleri bulunmaktadır. Şekil 2'de veri setindeki 4 sınıfa ait X-ışını görüntüsü örnekleri verilmiştir.

Veri setinin içerdiği çeşitli zorluklardan biri, değişken ve nispeten düşük görüntü çözünürlüğe sahip olmasıdır. Görüntülerin çoğunun en uzun boyutu 250 pikseli geçmemekle birlikte en-boy oranları da farklıdır. Diğer zorluklar ise, değişken ve bazen çok düşük görüntü kontrastı ve her bir sınıfa ait örnek sayısı dengesizliğidir. Tüm bu zorluklar önerilen problemin zorluğunu açıkça ortaya koymaktadır. Bu çalışmada, bu veri setini seçmemizin sebebi 2020 yılında önerilen herkese açık güncel bir veri seti olması ve literatürde en güncel çalışmaların bu veri setini kullanıyor olmasıdır. Bu veri setini kullanan tüm çalışmalar Bölüm 3'te sunulacaktır.

Veri seti %90 eğitim ve %10 test olmak üzere iki parçaya ayrılmıştır. Ayrıca, eğitim seti beş gruba ayrılarak 5-kat çapraz doğrulama gerçekleştirilmiştir. Dört set alt eğitim için, bir set ise validasyon için kullanılmıştır. Veri seti bölünürken her sınıftan rastgele görüntüler seçilmiştir. Ayrıca her bir çapraz doğrulama katında, doğrulama için seçilen tüm örneklerin tamamen bir

sınıfa ait olması veya bir sınıftan hiç örnek olmaması gibi durumlar engellenmiştir. Veri setinin kullanımına göre her sınıftaki görüntü sayısı Tablo 1'de verilmiştir.



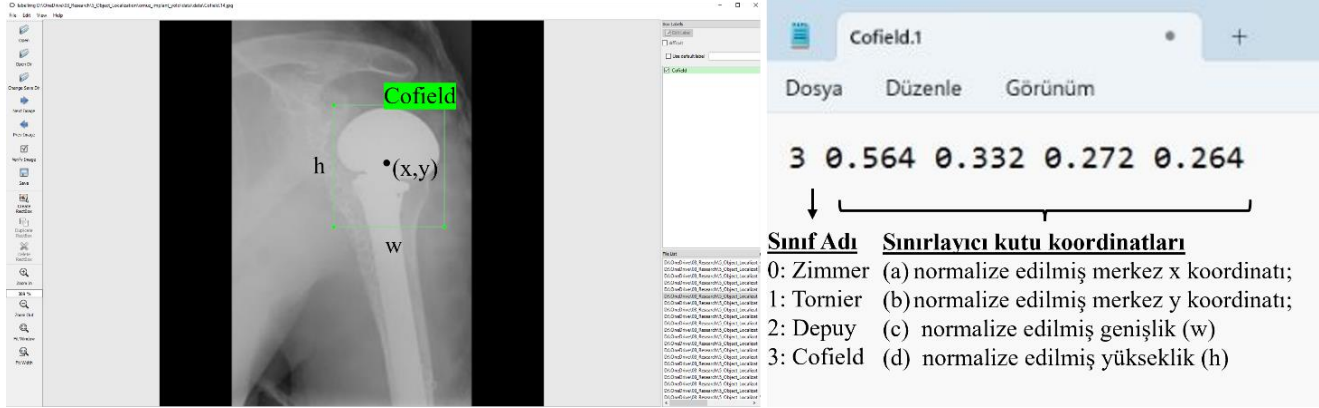
Şekil 2. Veri setindeki farklı üreticilerden bazı örnek görüntüler (a) Cofield (b) Depuy (c) Tornier (d) Zimmer

Figure 2. Some sample images from different manufacturers in the dataset (a) Cofield (b) Depuy (c) Tornier (d) Zimmer

Yukarıda bahsedilen zorluklardan biri olan veri dengesizliğini çözmek amacıyla Sivari vd. [4] çalışmalarında veri artırma uygulanmış ancak daha düşük sonuçlar elde etmişlerdir. Daha düşük sonuçların ise yapay örneklerden elde edilen özelliklerin gelişigüzel olması ve yapay örnek sayısının orijinal örnek sayısından fazla olması gibi nedenlerden kaynaklanıyor olabilir olduğunu belirtmişlerdir. Bu çalışmada, veri dengesizliği

sorununu ele almak için sınıf ağırlıklandırma tekniği uygulanmıştır. Bu yöntem, veri setindeki sınıfların frekansına dayalı olarak farklı ağırlıkların atanmasını ve böylece az temsil

edilen sınıflara daha fazla vurgu yapmayı amaçlamaktadır. Matematiksel olarak, w_i ile ifade edilen sınıf i 'ye atanmış ağırlık şu şekilde hesaplanmıştır.



Şekil 3. Veri kümesi içinde etiketlenmiş görüntülere bir örnek.

Figure 3. An example of labelled images within the dataset.

$$w_i = \frac{N}{n_i \times C} \quad (1)$$

Burada N toplam örnek sayısını, n_i sınıf i 'deki örnek sayısını ve C toplam sınıf sayısını temsil etmektedir. Elde edilen sınıf ağırlıkları, eğitim sırasında hata fonksiyonuna dahil edilerek modelin azınlık sınıflara daha fazla dikkat etmesi teşvik edilmiş ve böylece genel performans artırılmıştır.

Tablo 1. Veri setinin kullanımına göre her sınıftaki görüntü sayısı.

Table 1. Number of images in each class according to the use of the dataset.

Üretici	Train	Test
Cofield	75	8
Depuy	264	30
Tornier	64	7
Zimmer	134	15
Toplam	537	60

Karaci vd. [12] çalışmasında önerildiği gibi implantın baş bölgesini tespit etmek amacıyla veri setindeki tüm görüntüler etiketlenirken LabelImg etiketleme aracı kullanılmıştır. Etiketleme yöntemi olarak dikkörtgen sınırlayıcı kutu tekniği kullanılmıştır. Çünkü YOLO algoritmaları bu tür dikkörtgen etiketli verilerle çalışmaktadır. Her çizilen dikkörtgen sınırlayıcı kutuya ilgili implant sınıf adı verilerek etiketleme tamamlanmıştır.

Etiketleme aracıyla gerçekleştirilen bir görüntü etiketleme örneği Şekil 3'te gösterilmiştir. Etiketleme işlemi tamamlandıktan sonra araç çıktı olarak her bir görüntü için o görüntü üzerinde etiketlenen bölgelerin sınırlayıcı kutu koordinatlarını ve sınıf bilgilerini vermektedir. Bu etiketler PASCAL VOC veya YOLO formatında txt dosyaları olarak kaydedilmektedir. ImageNet ve MS COCO veri setine ait etiketler PASCAL VOC formatında kaydedildiği için, bu çalışmada LabelImg aracı ile elde edilen etiketler PASCAL VOC formatında

kaydedilmiştir. Görüntüdeki sınırlayıcı kutu ilgili .txt dosyasında depolanan veri satırına karşılık gelmektedir. Satırdaki beş adet sayısal değer sırasıyla sınıf adını ve görüntü üzerindeki sınırlayıcı kutu bilgilerini temsil etmektedir. Sınırlayıcı kutu koordinatları ise sırasıyla (a) normalize edilmiş merkez x koordinatı, (b) normalize edilmiş merkez y koordinatı, (c) normalize edilmiş genişlik ve (d) normalize edilmiş yüksekliği temsil etmektedir. Sınırlayıcı kutunun sol alt köşe koordinatı (x_1, y_1) , sağ üst köşe koordinatı (x_2, y_2) , görüntü genişlik w ve yüksekliği h olmak üzere değerler aşağıdaki gibi hesaplanmıştır.

$$a = \frac{(x_1 + x_2)/2.0}{w} \quad (2)$$

$$b = \frac{(y_1 + y_2)/2.0}{h} \quad (3)$$

$$c = \frac{x_2 - x_1}{w} \quad (4)$$

$$d = \frac{y_2 - y_1}{h} \quad (5)$$

2.2. YOLOv5

Çalışmada YOLOv5 ağının YOLOv5n (nano), YOLOv5s (küçük), YOLOv5m (orta) ve YOLOv5l (büyük) olmak üzere dört alt versiyonu kullanılmıştır. Alt versiyonlar adlandırılırken s, m, l ve x, ağ mimarisinin artan derinliğini göstermektedir. YOLOv5 ağının her bir versiyonu farklı model derinliklerine sahiptir, ancak genel olarak omurga, boyun ve kafa olmak üzere üç ana bölümden oluşan aynı ağ yapısı temel alınarak tasarlanmıştır. YOLOv5 model omurgası, belirli bir girdi görüntüsünden temel özellikleri çıkarmak için kullanılmaktadır. Elde edilen özellikler kullanılarak özellik haritası oluşturulur. Yüksek doğruluk sağlarken ve modellerin işlem süresini azaltırken üst düzey özellikleri çıkarmayı amaçlayan Çapraz Aşamalı Kısmi Ağ (Cross Stage Partial Network, CSPNet) [26] mimarisine dayalı olarak tasarlanmıştır. Model boynu, özellik piramitleri oluşturmak amacıyla model omurgasının farklı aşamalarından elde edilen özellik haritalarını toplar. YOLOv5 özellik piramitlerini elde etmek için model boynu içinde Yol Toplama Ağ (PANet) [27]

gerekmektedir. Çıktı boyutu $3 \times (5 + 4) = 27$ olacaktır. Bu çalışmada, COCO veri seti üzerinde önceden eğitilmiş YOLOv5 versiyonları omuz implantı veri seti üzerinde ince ayarlanmıştır. Bunun sebebi, veri setinde bulunan örnek sayısının az olması ve modelin rastgele başlatılan ağırlıklarla sıfırdan eğitildiği takdirde, yakınsamasının uzun sürebilecek olmasıdır. Bu durumda bile, yüksek doğruluk elde edilmeyebilmektedir. Önceden eğitilmiş YOLOv5 modellerinin ince ayarlanması, transfer öğrenme olarak adlandırılmaktadır ve modelin transfer öğrenme ile eğitilmesi daha yüksek başarı sağlamakla birlikte daha az zaman almaktadır.

2.4. Değerlendirme metrikleri

Bu çalışmada YOLOv5 versiyonlarının omuz implantı tespit performansını nicel olarak değerlendirmek amacıyla doğruluk, ortalama kesinlik (mean Average Precision, mAP), kesinlik (precision), duyarlılık (recall) ve F1-ölçütü olmak üzere beş metrik kullanılmıştır. Doğruluk, tüm örnekler içerisinde modeller tarafından doğru olarak tespit edilen örneklerin oranını ölçer. Ortalama kesinlik (Average Precision, AP), her bir sınıf için kesinlik-duyarlılık eğrisi üzerinde farklı duyarlılık değerlerine karşılık gelen kesinlik değerlerinin ortalaması olarak hesaplanmakta, ortalama kesinlik değerlerinin ortalaması (mean Average Precision; mAP) ise, bütün sınıflara ait kesinlik değerlerinin ortalaması olarak ifade edilmektedir. mAP@0.5, nesne tespitlerinin kesinlik ve duyarlılık oranını Intersection over Union (IoU) eşiği 0.5'ten büyük olan tahminler üzerinde değerlendirir. Öte yandan, mAP@[0.5:0.95], 0.5 ile 0.95 arasındaki farklı IoU eşiklerini dikkate alarak modelin performansını daha kapsamlı bir şekilde değerlendirir. Kesinlik, modeller tarafından pozitif olarak tespit edilen örnekler içerisinde gerçek pozitif örneklerin oranını ölçerken, duyarlılık ise veri kümesindeki pozitif örneklerin içinde gerçek pozitif tespitlerin oranını ölçmektedir. F1-ölçütü, kesinlik ve duyarlılık başarı değerlendirme ölçütlerinin birlikte değerlendirilmesi olup, kesinlik ve duyarlılık değerlerinin harmonik ortalamasıdır. Bu ölçütler (6-11) eşitlikleri kullanılarak hesaplanmaktadır. Eşitliklerde DP, DN, YP ve YN sırasıyla, doğru pozitif, doğru negatif, yanlış pozitif ve yanlış negatif sayısını ifade etmektedir. P, R, N değerleri ise sırasıyla tüm kategorilerdeki kesinliği, duyarlılığı ve toplam sınıf sayısını temsil etmektedir.

$$\text{Doğruluk} = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (7)$$

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (8)$$

$$F - \text{ölçütü} = 2 \times \frac{\text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (9)$$

$$AP = \int_0^1 P(R)dR \quad (10)$$

$$AP = \frac{\sum_{i=1}^N AP_i}{N} \quad (11)$$

3. Bulgular ve Tartışma

Bu bölümde, omuz implantlarının tespiti ve sınıflandırılması için önerilen YOLOv5 tabanlı yaklaşımın (tüm alt versiyonlarıyla birlikte) performansı en son teknoloji yöntemlerin performansı ile karşılaştırılmıştır. Bir ağır eğitiminde yüksek doğruluk seviyesine ulaşmak için sıfırdan eğitim yapıldığı takdirde gereken zaman ve veriden tasarruf etmek amacıyla transfer öğrenme kullanılmaktadır. YOLO versiyonlarında transfer öğrenme COCO veri seti üzerinde önceden eğitilmiş ağırlıklar kullanılarak yapılmaktadır. Bu çalışmada, her bir YOLOv5 versiyonu, COCO veri seti üzerinde önceden eğitilmiş ağırlıklar kullanılarak omuz implantları içeren X-ışını görüntülerinden oluşan bir veri seti üzerinde ince ayarlanmıştır. YOLOv5 modelinin tüm alt versiyonları 400 epoch boyunca eğitilmiştir ve ağırların öğrenme sürecinde hiper-parametre set şöyle ayarlanmıştır: batch boyutu = 8, momentum = 0.9, decay = 0.0005 ve öğrenme hızı = 0.001. Tüm modeller için eğitim süreci, validasyon kayıpları belirli iterasyon boyunca değişmediğinde veya 400 epoch tamamlandığında durdurulmuştur ve adil karşılaştırma için aynı test veri kümesi üzerinde test edilmiştir ve hiper-parametreler sabit tutulmuştur. Performans hem niceliksel hem de niteliksel olarak karşılaştırılmıştır.

Nicel performans değerlendirmesi için, IoU metriği, her bir omuz implant bölgesinin tespit doğruluğunu belirlemek kullanılmıştır. IoU, tahmin edilen sınırlayıcı kutu (A_{pred}) ile kesin referans sınırlayıcı kutu (A_{gt}) arasındaki örtüşmeyi ölçer ve Eşitlik 12'deki gibi hesaplanır.

$$IoU = \frac{A_{pred} \cap A_{gt}}{A_{pred} \cup A_{gt}} \quad (12)$$

Bu çalışmada, 0.6 veya daha büyük IoU değeri ile sınıflandırılan nesne tespit sonuçları, IoU değeri 0.6'dan büyükse doğru pozitif (TP) olarak kabul edilmiştir. IoU değeri 0.6'dan küçükse sonuç yanlış pozitif (FP) olarak değerlendirilmiştir. IoU çıktısı kullanılarak kesinlik, duyarlılık, F1-ölçütü ve mAP metrikleri sınıflandırma performansını hesaplamak için kullanılmıştır.

YOLOv5 modellerinin alt versiyonlarının deneysel sonuçlarını araştırmak için 60 görüntü test seti olarak kullanılmıştır. Sonuçlar 5-kat çapraz doğrulama uygulanarak elde edilmiştir. 4 farklı modelin ve YOLOv5l modelinin sınıf ağırlıklandırma versiyonunun dayanıklılığını değerlendirmek için elde edilen kesinlik, duyarlılık, F1-ölçütü ve mAP değerleri Tablo 3'te özetlenmiştir. Modelin sınıflandırma performansını değerlendirmek için kesinlik, duyarlılık, F1-ölçütü ve mAP kullanılmıştır. Tabloda sunulan tüm değerler, sınıfların ayrı ayrı kesinlik, duyarlılık, F1-ölçütü ve mAP değerlerinin makro ortalamasını temsil etmektedir. Makro ortalama, her sınıfın performansını eşit olarak kabul etmektedir. Tüm modeller, 5-kat çapraz doğrulama ile elde edilen veri kümeleri üzerinde eğitilmiştir. Karşılaştırma sonuçları, YOLOv5n modeli diğer modellere göre en düşük performans göstermiştir. Bu modelin mAP@0.5, mAP@0.5:0.95, kesinlik, duyarlılık ve F1-ölçütü sırasıyla 94.8, 82.2, %91.3, %79.8 ve %85.1 olarak tespit edilmiştir. Diğer modeller arasında ise YOLOv5l modelinin sınıf ağırlıklandırma tekniği ile birlikte uygulandığında önemli ölçüde daha iyi bir performans sergilediği görülmüştür. Bu modelin mAP@0.5, mAP@0.5:0.95, kesinlik, duyarlılık ve F1-ölçütü sırasıyla 98.3, 85.9, %97.0, %95.7 ve %96.3 olarak tespit edilmiştir.

Tablo 3. YOLOv5n, YOLOv5s, YOLOv5m ve YOLOv5l modellerinin performans karşılaştırması.**Table 3.** Performance comparison of YOLOv5n, YOLOv5s, YOLOv5m and YOLOv5l models.

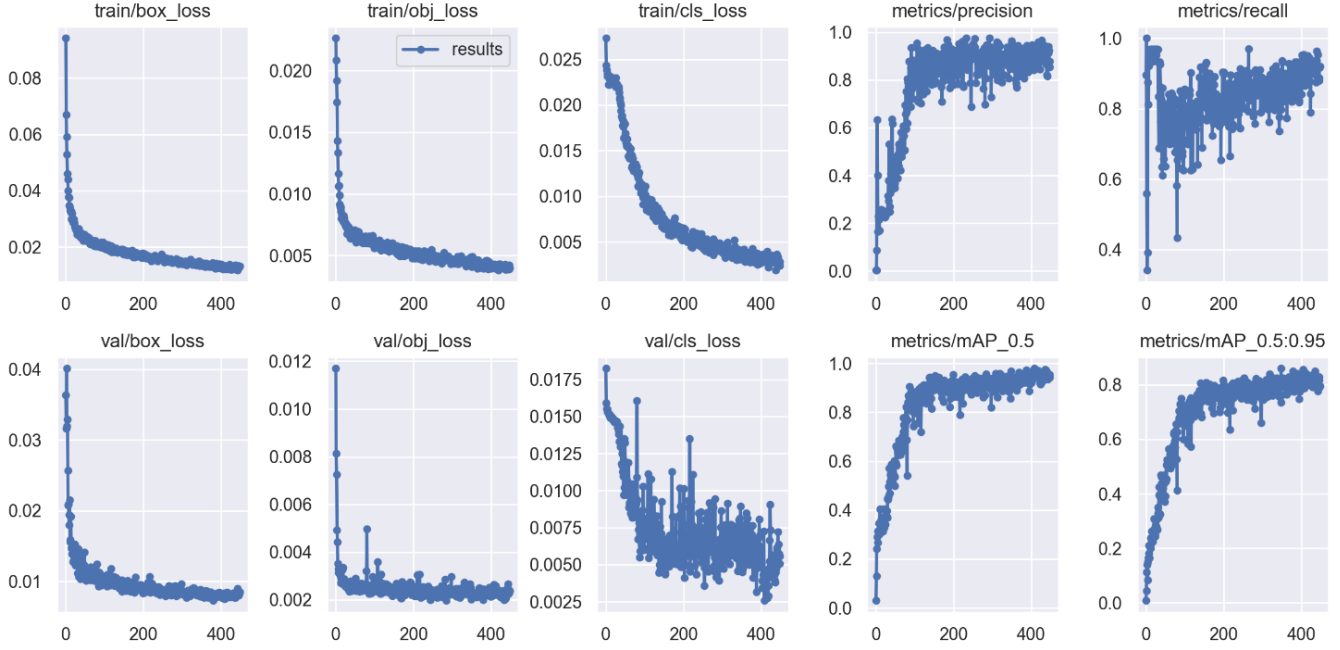
Model	Sınıf	Kesinlik(%)	Duyarlılık(%)	F1-ölçütü(%)	mAP (0.5%)	mAP (0.5:0.95%)
YOLOv5n	Cofield	84.1	66.7	74.4	94.0	77.5
	Depuy	96.3	86.7	91.2	96.8	84.2
	Tornier	94.7	85.7	89.9	94.4	84.9
	Zimmer	90.1	80.0	84.8	93.9	82.1
	Ortalama	91.3	79.8	85.1	94.8	82.2
YOLOv5s	Cofield	100.0	62.4	76.8	91.0	72.1
	Depuy	93.2	93.3	93.2	98.8	81.4
	Tornier	84.7	100.0	91.7	97.8	88.6
	Zimmer	78.2	86.7	82.2	95.8	82.2
	Ortalama	89.1	85.6	86.0	95.8	81.1
YOLOv5m	Cofield	100.0	75.3	85.9	97.1	84.6
	Depuy	90.6	93.3	91.9	96.8	83.7
	Tornier	90.1	100.0	94.8	99.5	91.3
	Zimmer	71.4	93.3	80.9	94.0	84.1
	Ortalama	88.0	90.5	88.4	96.9	85.9
YOLOv5l	Cofield	100.0	89.9	94.7	99.5	86.5
	Depuy	91.1	96.7	93.8	98.5	86.4
	Tornier	95.3	100.0	97.6	99.5	85.6
	Zimmer	87.3	92.0	89.6	97.1	84.1
	Ortalama	93.4	94.6	93.9	98.7	85.7
YOLOv5l + Class Weight	Cofield	96.5	87.5	91.77	96.2	80.9
	Depuy	99.4	96.7	98.03	98.9	85.3
	Tornier	100.0	98.8	99.39	99.5	93.7
	Zimmer	91.9	100.0	95.77	98.7	83.6
	Ortalama	97.0	95.7	96.34	98.3	85.9

Model performansları sınıf bazlı değerlendirildiğinde ise, Cofield sınıfında, YOLOv5s, YOLOv5m ve YOLOv5l modellerinin %100 kesinlikle en yüksek başarıya ulaştığı dikkat çekmiştir. Aksine, YOLOv5n modeli, Cofield sınıfında %84.1 ile en düşük kesinlik değerine sahiptir. Depuy sınıfında, YOLOv5l modeli sırasıyla %99.4 ve %96.7 olmak üzere en yüksek kesinlik ve duyarlılık değerlerine sahiptir. Aynı şekilde Tornier sınıfında YOLOv5l modeli sırasıyla %100.0 ve %98.8 ile en yüksek kesinlik ve duyarlılık değerlerine sahiptir. Son olarak, Zimmer sınıfında, YOLOv5l modelinin duyarlılık değeri diğer modellere kıyasla daha yüksektir. F-ölçütünde ise YOLOv5l modeli tüm sınıflarda en yüksek başarıyı sağlamıştır.

Genel olarak, bu sonuçlar sınıf ağırlıklandırmalı YOLOv5l modelinin en iyi performansı gösterdiğini ve daha küçük modellerin (YOLOv5n, YOLOv5s ve YOLOv5m) bazı sınıflarda daha düşük başarıya sahip olduğunu göstermektedir.

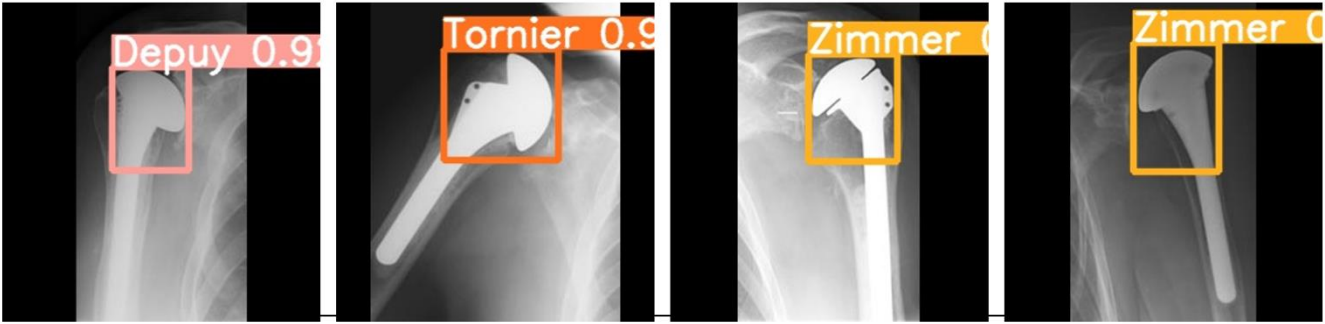
Şekil 5'te YOLOv5l modelinin eğitim ve validasyon grafiklerinin değişim sonuçları görülmektedir. Kesinlik ve duyarlılık değerlerinin hem eğitim hem de validasyon kümesinde hızlı bir şekilde %100'e, hata oranının da benzer hızda 0'a yakınsadığı görülmüştür.

Nitel performans değerlendirmesi için, test setinden rastgele seçilen birkaç görüntü üzerinde YOLOv5l modeli ile elde edilen tespit etme ve sınıflandırma sonuçları Şekil 6'da verilmiştir.



Şekil 5. YOLOv5l modelinin eğitim ve validasyon grafiklerinin değişim sonuçları.

Figure 5. Change results of the training and validation graphs of the YOLOv5l model.



Şekil 6. Veri setine ait bazı test görüntüleri üzerinde YOLOv5l modeli ile elde edilen doğru pozitif algılama ve sınıflandırma sonuçları.

Figure 6. True positive detection and classification results obtained with the YOLOv5l model on some test images of the dataset

YOLOv5l, omuz implantlarının farklı kategorileri için iyi bir tanıma performansına sahiptir çünkü konvolüsyonel sinir ağı algoritması manuel özellik çıkarma gerektirmemekle birlikte ve genelleme yeteneğini artırmaktadır.

Tablo 4, omuz implantı üreticilerinin tespiti ve sınıflandırılması amacıyla önerilen sınıf ağırlıklandırmalı YOLOv5 modelinin aynı veri kümesi üzerinde önerilen önceki çalışmalarla kapsamlı bir karşılaştırmasını sunmaktadır. Önceki çalışmalarda genellikle bireysel makine öğrenmesi, bireysel derin öğrenme yaklaşımları ve hibrid yaklaşımlar önerilmiştir. Bu çalışmalar oldukça güncel olup derin öğrenme ve makine öğrenmesi yaklaşımlarının doğru ve güvenilir sınıflandırma sonuçları elde etmek için etkili oldukları açıktır. Sonuçlar incelendiğinde en yüksek çapraz doğrulama doğruluk, kesinlik ve duyarlılık değerleri [9] çalışmasında sırasıyla %97.2, %98.6 ve %98.0 olarak elde edilmiştir. Öte yandan diğer en iyi sonuçların elde edildiği Sivari vd. [4] çalışmasında da belirtildiği gibi, [9] çalışmasında veri kümesi test ve doğrulama setine ayrılmamakla birlikte, yalnızca 5 kat çapraz doğrulama sonuçları verilmiştir. Doğrulama veri setinin amacı modelin hiper parametrelerini ayarlamak ve

modelin yeteneklerini ölçmektir. Dolayısıyla Sivari vd. [4] çalışmasında sonuç değerlendirmesinin tarafsız bir test setinde gerçekleştirilmesi önerilmiştir. Bu öneri doğrultusunda %95.0 doğruluk, %97.0 kesinlik ve %95.7 duyarlılık değerleri elde edilmiştir. Önerilen yöntem literatürdeki diğer çalışmalarla kıyaslandığında karşılaştırılabilir düzeyde oldukça başarılı sonuçlar elde edilmiştir.

4. Sonuçlar

Bu çalışmada, omuz implantı üreticilerinin otomatik tespiti ve sınıflandırılması için önerilen sistemde, YOLOv5n, YOLOv5s, YOLOv5m ve YOLOv5l modelleri kullanılarak model implant baş bölgesi tespiti ve sınıflandırma gerçekleştirilmiş ve sonuçlar karşılaştırılmıştır. Ayrıca, veri dengesizliği problemi çözmek amacıyla önerilen sınıf ağırlıklandırma tekniğinin modelin sınıflandırma performansını önemli ölçüde artırdığı tespit edilmiştir. Yapılan karşılaştırmada, en yüksek mAP@0.5, mAP@0.5:0.95, kesinlik, duyarlılık ve F1-ölçütü değerleri sınıf ağırlıklandırmalı YOLOv5l modeli ile sırasıyla 98.3, 85.9, %97.0, %95.7 ve %96.3 olarak elde edilmiştir.

Tablo 4. Aynı veri seti üzerinde yapılan literatürdeki diğer çalışmaların karşılaştırması.**Table 4.** Comparison of other studies in the literature on the same data set.

Çalışma	Veri seti, Sınıf bilgisi	Yöntem	Doğruluk (%)	Kesinlik(%)	Duyarlılık(%)
Urban vd. [7]	597: Cofield, Depuy, Tornier, Zimmer	NasNet	80.0	80.0	75.0
Sultan vd. [10]	597: Cofield, Depuy, Tornier, Zimmer	DRE-Net	85.92	85.33	84.11
Yılmaz [9]	597: Cofield, Depuy, Tornier, Zimmer	Çok kanallı CNN	97.2	98.6	98.0
Efeoglu vd. [11]	349: Cofield, Depuy, Zimmer	K-NN	74.0	64.16	56.1
Karaci [12]	597: Cofield, Depuy, Tornier, Zimmer	YOLOv3 + DenseNet201	84.76	82.37	80.35
Sivari vd. [4]	597: Cofield, Depuy, Tornier, Zimmer	DenseNet201 + Lojistik Regresyon	95.07	96.77	91.64
Önerilen Çalışma	597: Cofield, Depuy, Tornier, Zimmer	YOLOv5l	95.0	97.0	95.7

Önerilen YOLOv5 modelleri, literatürdeki diğer hibrit makine öğrenmesi yaklaşımlarına göre düşük maliyetli ve yüksek performanslı bir hedefe ulaşmaktadır. Dolayısıyla, omuz implantlarının tespiti ve sınıflandırılması gibi tıbbi görüntüleme uygulamalarında YOLOv5 modellerinin etkili bir araç olduğunu göstermiştir. Ayrıca, radyologlar ve cerrahlar için verimli ve güvenilir bir analiz süreci sağlayabileceği ve iş yükünü azaltabileceği öngörülmüştür. Bu çalışmanın önemli bir sınırlaması, veri kümesindeki implant üretici sınıflarının sayısının sınırlı olmasıdır. Veri kümesindeki implant üreticilerinin çeşitliliğinin artırılması, çalışmanın genişletilmesi için bir fırsat sunmaktadır. Öte yandan, daha geniş veri kümelerinde ve farklı görüntüleme koşullarında test edilerek, modellerin performanslarının daha da geliştirilmesi mümkündür.

Etik kurul onayı ve çıkar çatışması beyanı

Hazırlanan makalede etik kurul izni alınmasına gerek yoktur. Hazırlanan makalede herhangi bir kişi/kurum ile çıkar çatışması bulunmamaktadır.

Kaynaklar

- [1] Cofield, R. H., 1984. Total shoulder arthroplasty with the Neer prosthesis. *JBJS*, Cilt. 66(6), s. 899-906. DOI: 10.2106/00004623-198466060-00010.
- [2] Sanchez-Sotelo, J., 2011. Total shoulder arthroplasty. *The Open Orthopaedics Journal*, Cilt. 5, s. 106. DOI: 10.2174/1874325001105010106.
- [3] Sukjamsri, C., 2015. The effect of implant misalignment on shoulder replacement outcomes. Doctoral dissertation, Imperial College London. DOI: 10.25560/28581.
- [4] Sivari, E., Güzel, M. S., Bostanci, E., Mishra, A., 2022. A novel hybrid machine learning-based system to classify shoulder implant manufacturers. *Healthcare*, Cilt. 10(3), s. 580. DOI: 10.3390/healthcare10030580.
- [5] Sahoo, D. P., Rout, M., Mallick, P. K., Samanta, S. R., 2022. Comparative analysis of medical images using transfer learning based deep learning models. *International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)*, s. 1-8.
- [6] Sistaninejad, B., Rasi, H., Nayeri, P., 2023. A review paper about deep learning for medical image analysis. *Computational and Mathematical Methods in Medicine*, Cilt. 2023.
- [7] Urban, G., Porhemmat, S., Stark, M., Feeley, B., Okada, K., Baldi, P., 2020. Classifying shoulder implants in X-ray images using deep learning. *Computational and Structural Biotechnology Journal*, Cilt. 18, s. 967-972. DOI: 10.1016/j.csbj.2020.04.005.
- [8] Yi, P. H., Kim, T. K., Wei, J., Li, X., Hager, G. D., Sair, H. I., Fritz, J., 2020. Automated detection and classification of shoulder arthroplasty models using deep learning. *Skeletal Radiology*, Cilt. 49, s. 1623-1632. DOI: 10.1007/s00256-020-03463-3.
- [9] Yılmaz, A., 2021. Shoulder implant manufacturer detection by using deep learning: Proposed channel selection layer. *Coatings*, Cilt. 11(3), p. 346. DOI: 10.3390/coatings11030346.
- [10] Sultan, H., Owais, M., Park, C., Mahmood, T., Haider, A., Park, K.R., 2021. Artificial intelligence-based recognition of different types of shoulder implants in X-ray scans based on dense residual ensemble-network for personalized medicine. *Journal of Personalized Medicine*, Cilt. 11, p. 482.
- [11] Efeoglu, E., Gürkan, T. U. N. A., 2021. Radyografi görüntüleri ve sınıflandırma algoritmaları kullanılarak omuz protezlerinin üreticilerinin belirlenmesi. *Kırklareli Üniversitesi Mühendislik ve Fen Bilimleri Dergisi*, Cilt. 7(1), s. 57-73. DOI: 10.34186/klujes.906660.
- [12] Karaci, A., 2022. Detection and classification of shoulder implants from X-ray images: YOLO and pretrained convolution neural network-based approach. *Journal of Faculty of Engineering and Architecture of Gazi University*, Cilt. 37, s. 283-294. DOI: 10.17341/gazimmf.888202.
- [13] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition*, s. 248-255. DOI: 10.1109/CVPR.2009.5206848.
- [14] Zoph, B., Vasudevan, V., Shlens, J., Le, Q. V., 2018. Learning transferable architectures for scalable image recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, s. 8697-8710. DOI: 10.1109/CVPR.2018.00907.
- [15] Sahin, O., Ozer, S., 2022. YOLODrone+: Improved YOLO Architecture for Object Detection in UAV Images. *IEEE Signal Processing and Communications Applications Conference (SIU)*, s. 1-4. DOI: 10.1109/SIU55565.2022.9864746.
- [16] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., Berg, A. C., 2016. SSD: Single shot multibox detector. *ECCV 2016*, Cilt. 14, s. 21-37. DOI: 10.1007/978-3-319-46448-0_2.
- [17] Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection. *IEEE Conference on Computer Vision and Pattern Recognition*, s. 779-788.
- [18] Jintasuttisak, T., Edirisinghe, E., Elbattay, A., 2022. Deep neural network based date palm tree detection in drone imagery. *Computers and Electronics in Agriculture*, Cilt. 192, p. 106560. DOI: 10.1016/j.compag.2021.106560.
- [19] Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. *IEEE Conference on Computer Vision and Pattern Recognition*, s. 580-587.
- [20] Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (s. 1440-1448).
- [21] 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*, Cilt. 28.
- [22] Redmon, J., Farhadi, A., 2017. YOLO9000: better, faster, stronger. *IEEE Conference on Computer Vision and Pattern Recognition*, s. 7263-7271.
- [23] Redmon, J., Farhadi, A., 2018. YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
- [24] Bochkovskiy, A., Wang, C. Y., Liao, H. Y. M., 2020. YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
- [25] Nelson, J., Solawetz, J., 2020. YOLOv5 is here: State-of-the-art object detection at 140 fps. *Roboflow*.
- [26] Wang, C. Y., Liao, H. Y. M., Wu, Y. H., Chen, P. Y., Hsieh, J. W., Yeh, I. H., 2020. CSPNet: A new backbone that can enhance learning capability of CNN.

- IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, s. 390-391.
- [27] Liu, S., Qi, L., Qin, H., Shi, J., Jia, J., 2018. Path aggregation network for instance segmentation. IEEE Conference on Computer Vision and Pattern Recognition, s. 8759-8768.
- [28] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., 2014. Microsoft COCO: Common objects in context. ECCV 2014, Cilt. 13, s. 740-755. DOI: 10.1007/978-3-319-10602-1_48.