

Derin öğrenme temelli nesne tespit yöntemleri kullanılarak insan sperm hücrelerinin tespiti

Detection of human sperm cells using deep learning-based object detection methods

Mecit YUZKAT^{1,2*}, Hamza Osman İLHAN³, Nizamettin AYDIN³

¹Türkiye Sağlık Enstitüleri Başkanlığı, İstanbul, Türkiye.

mecit.yuzkat@tuseb.gov.tr

²Yazılım Mühendisliği Bölümü, Mühendislik Mimarlık Fakültesi, Muş Alparslan Üniversitesi, Muş, Türkiye.

myuzkat@yildiz.edu.tr

³Bilgisayar Mühendisliği Bölümü, Elektrik-Elektronik Fakültesi, Yıldız Teknik Üniversitesi, İstanbul, Türkiye.

hoilhan@yildiz.edu.tr, naydin@yildiz.edu.tr

Geliş Tarihi/Received: 15.03.2023

Düzeltilme Tarihi/Revision: 17.08.2023

doi: 10.5505/pajes.2023.89346

Kabul Tarihi/Accepted: 13.09.2023

Araştırma Makalesi/Research Article

Öz

Kısırlık son 50 yılda dünya çapında önemli bir sağlık problemi haline gelmiştir. Dünyanın farklı bölgelerine göre oranı değişen bu problem ortalama olarak dünyadaki her on çiftten birini etkilemektedir. Erkek kaynaklı kısırlık teşhisi, sperm kalitesi değerlendirilerek yapılır. Sperm kalitesi araştırılırken sperm sayısı, hareketi ve morfolojik yapısı değerlendirilir. Sperm hareketi ve sayısının analizi öncesinde sperm tespiti önemli bir adımdır. Bu çalışmada, oluşturulan özgün yeni semen video veri kümesi üzerinde Faster R-CNN ve YOLOv3 derin öğrenme yöntemleri kullanılarak otonom sperm tespiti gerçekleştirilmiştir. Çalışma kapsamında oluşturulan bu özgün veri kümesi, mikroskop altında semen örneklerinin cep telefonu yardımı ile elde edilen 10 hastaya ait semen videolarını ve içeriğindeki nesnelerin sperm ve sperm olmayan şeklindeki etiket bilgisini içermektedir. Analiz için hazırlanmış etiketli videolar hasta odaklı ve hasta bağımsız olmak üzere iki senaryo ile değerlendirilmiştir. İlk senaryomuzda etiketli sekiz video birleştirilerek Faster R-CNN ve YOLOv3 modelleri 3 farklı oranda oluşturulmuş veriler ile eğitilmiş ve test edilmiştir. İkinci senaryoda ise eğitilmiş her bir modelimiz daha önce eğitime hiç katılmamış iki video ile test edilmiştir. İkinci senaryomuzda, eğitime hiç katılmamış videolar kullanılarak algılama performansları değerlendirilmiştir. Yapılan çalışmada YOLOv3 modeli ile bireysel videolarda %96, ortalama da ise %84,5 gibi mAP sperm tespit sonuçları elde edilmiştir. Sonuçlar nesne tespitinin doğruluğu ve eğitim süreleri gibi iki önemli kriter ile karşılaştırıldığında YOLOv3 yöntemi Faster R-CNN yönteminden daha başarılı olduğu gözlemlenmiştir.

Anahtar kelimeler: Kısırlık, Video analizi, Sperm tespiti, YOLOv3, Faster R-CNN.

Abstract

Infertility has become a significant health issue worldwide in the last 50 years. This issue, with varying rates across different regions of the world, affects approximately one out of every ten couples on average globally. Diagnosis of male-related infertility is conducted by evaluating sperm quality. When investigating sperm quality, factors such as sperm count, motility, and morphological structure are assessed. Detection of sperm before the analysis of sperm motility and count is an important step. In this study, autonomous sperm detection was carried out using deep learning methods, namely Faster R-CNN and YOLOv3, on a newly generated and unique semen video dataset. This distinct dataset, created within the scope of this study, includes semen videos from 10 patients obtained with the assistance of a mobile phone under a microscope. Videos contain label information that classifies objects as sperm and non-sperm. Labeled videos prepared for analysis were evaluated under two scenarios: patient-focused and patient-independent. In the first scenario, eight labeled videos were combined to train and test Faster R-CNN and YOLOv3 models in three different ratios. In the second scenario, each trained model was tested with two videos that had never been part of the training process. In this second scenario, detection performances were evaluated using videos that had not been involved in training. The study achieved sperm detection results of approximately 96% in individual videos using the YOLOv3 model and an average mAP of 84.5%. When compared against two significant criteria, object detection accuracy and training times, the YOLOv3 method was observed to be more successful than the Faster R-CNN method.

Keywords: Infertility, Video analysis, Sperm detection, YOLOv3, Faster R-CNN.

1 Giriş

Günümüzde kısırlık birincil veya ikincil kısırlık olarak sınıflandırılır. Birincil kısırlık; erkek ile kadının 12 ay veya daha fazla süre içerisinde düzenli olarak ve hiçbir cinsel korunma olmadan cinsel ilişkiye girmelerine rağmen kadının hamile kalamaması olarak tanımlanmaktadır [1]. İkincil kısırlık ise kadının en az bir hamilelikten sonra bir daha hamile kalamama durumuna denir. Açıklanabilen veya açıklanamayan olarak belirtilen kısırlık problemleri yaşayan bireylerin sayısı son 50 yılda hızla artarak bütün ülkelerde önemli, çok yönlü bir sorun

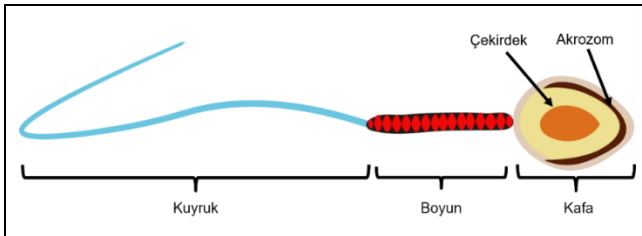
haline gelmiştir [2],[3]. Kısırlık teşhisinin görülme oranları bölgesel olarak farklılıklar göstermektedir [4]. Amerika Birleşik Devletleri'nde %6, Birleşik Krallık' ta %10-15 Afrika ülkelerinde ise %20-35 arasında kısırlık problemlerinin görüldüğü raporlanmıştır [5]. Ortalama olarak dünyadaki çiftlerin %15'ini etkileyen bir problem haline gelmiştir [4],[6]. Geçmişte yapılan çalışmalarda kısır vakaların %20'sinden erkek birey etmeni tek başına sorumluyken %40'na kadar ise katkısının olduğu raporlanmıştır [7]. Günümüzdeki yapılan çalışmalarda bu oran artarak kısır vakalarının %40'ında sadece erkek sorumluyken %50'sine kadar erkek faktörünün katkısı

*Yazışılan yazar/Corresponding author

bulunmaktadır [8],[9]. Erkeklerde kısırlık nedenleri; damar tıkanıklığı (varikosel), kanal tıkanıklığı veya boşalma sorunları gibi anatomik anormallikler olduğu gibi dış çevresel faktörler de olabilir [10]. Erkek kısırlığına neden olan dış çevresel faktörler, kötü beslenme alışkanlığı, sigara, radyasyon, ağır metaller, yüksek sıcaklık, bazı anti bakteriyel ilaçlar, bazı bitkiler ve stres gibi birçok faktör örnek olarak verilebilir [11]. Sperm analizi ise kısırlık vakasında erkek partneri değerlendirmek için tercih edilen test olmuştur. Sperm analizi ile ilgili çalışmalar 1929'lu yıllara dayanmaktadır [12]. Fakat çalışmaların büyük bir bölümü 1980'den sonra Dünya Sağlık Örgütü'nün (DSÖ) ayrıntılı semen analizine başlaması ile ortaya çıkmaktadır. Semen örneklerinin analiz edilmesi, belirli durumlarda erkeğin doğurganlığı hakkında değerli bilgiler sağlasa da bazı şartları ve zorlukları vardır. DSÖ'ne göre, sperm analizi en az 3-5 gün cinsel perhiz sonrası mastürbasyonla erkeklerden elde edilen semen örneği üzerinden yapılır [13]. Eski çalışmalarda kısırlık teşhisi tamamen mikroskop altı görsel inceleme ile yapılırken günümüzde ise görsel incelemenin yanında bilgisayar destekli sistemler ile yapılmaktadır. Görsel değerlendirme, deneyimli kişiler tarafından mikroskopik görüntülerin manuel olarak incelenmesidir. Makler [14], Hemositometre [15] vb. gibi farklı tıbbi sayım materyalleri, sayım ve hareket analiz süreçlerine yardımcı olmak için görsel değerlendirme tekniğinde kullanılmaktadır. Görsel analiz zor, zaman alıcı ve gözlemciler arası değişkenliğe tabidir [16]. Görsel analizin dezavantajlarından dolayı 1980'lerden itibaren Bilgisayar Destekli Sperm Analiz (CASA) sistemleri geliştirilmiştir. CASA sistemleri görsel yöntemlere göre objektif ve tutarlı analizler sağlamaktadır [17]-[19]. Fakat CASA sistemlerinin görsel sistemlere göre yüksek maliyet ve standardizasyona bağımlı olmaları, ayrıca sistem kullanıcılarının parametre ve konfigürasyon eğitim ihtiyaçları gibi birçok dezavantajları vardır [20].

1.1 Sperm hücresi ve hareketi

Sperm tespiti aşamasında incelenen hücrenin sperm olup olmadığı tespit edilirken spermin morfolojine bakılır. DSÖ'nün Kruger kriterleri ve Menkveld çalışması [21] göz önüne alınarak normal bir sperm oval, düzgün kenarlı ve kafa alanının ortalama %40-%70'ini kaplayan belirgin bir akrozoma sahip olmalıdır. Ayrıca baş, orta kısım ve kuyruğuna ait herhangi bir deformasyonu bulunmamalıdır. Bu açıklamalar göz önünde bulundurularak Şekil 1'de örnek bir sperm hücresinin çizimi görselleştirilmiştir.



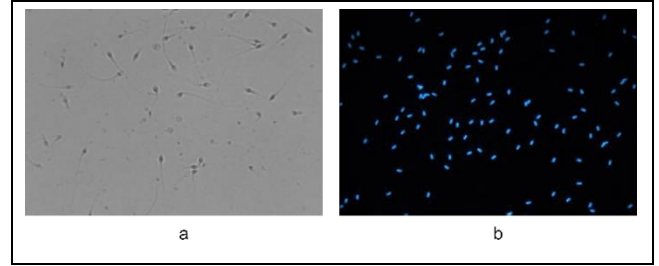
Şekil 1. Örnek sperm hücresi.

Figure 1. Example sperm cell.

DSÖ'nün yayınladığı kılavuza göre sperm sayısı için 15 milyon/ml normal sınır olarak kabul edilmektedir. Ayrıca sperm hareketi dört kategori ile tanımlanır: Hızlı hareket (Grade A), Normal hareket (Grade B), tutarsız hareket (Grade C) ve titreşimli ya da hareketsiz (Grade D). Aynı kılavuza göre, tüm sperm konsantrasyonunun %40'ı hareketli olmalı ve

ölçülen tüm hareketlilerin %32'si ilerleyici tipte olmalıdır (Grade A+B) [1].

Literatürde yapılan çalışmalar incelendiğinde sperm hareketi analizinde tıkanma, aydınlatma değişiklikleri, kamera efektleri ve gürültü gibi zorlu konular, aşılması gereken önemli konular olarak sunulmaktadır. Sperm videolarının görüntülenmesi aşamasında temelde faz-zıtlık (phase-contrast) veya florasan boyama (fluorescence dyeing) gibi teknikler kullanılmaktadır [19]. Florasan boyama tekniğinde ışığın meni üzerine yansımaları nedeniyle hareketli sperm sayısı kolaylıkla yapılabilmektedir. Fakat meni örneğinde kan hücreleri, kalıntılar veya lökositler de bulunur. Bu durumda bu hücreler hareketsiz sperm örnekleri ile karıştırılabilir. Yani hareketsiz sperm ve bu parçacıkların ayırt edilebilirliği tam olarak sağlanamayabilir. Bu açıdan literatürde sperm analizi ile ilgili yapılan çalışmaların çoğunluğu faz-zıtlık mikroskopik görüntüler üzerinden gerçekleştirilmiştir [22]. Şekil 2'de faz-zıtlık ve florasan boyama tekniklerinden elde edilmiş örnek sperm görüntüleri göstermektedir [23].



Şekil 2(a): Faz-zıtlık ve (b): Florasan boyama tekniklerinden elde edilmiş örnek sperm görüntüleri.

Figure 2. Sample sperm images obtained from.

(a): Phase-contrast and (b): Fluorescence dyeing techniques.

1.2 Nesne tespiti alanında incelenen çalışmalar

Son yıllarda derin öğrenme yöntemleri nesne algılama gibi birçok yapay zekâ alanında kullanılmaktadır. Literatürde iki aşamalı ve tek aşamalı yöntemler olmak üzere iki ana yaklaşım kapsamında derin öğrenme tabanlı nesne tespit (dedektör) yöntemleri mevcuttur. İki aşamalı CNN serisi modeller [24], [25]-[27] tek aşamalı YOLO serisi modeller [28]-[31], SSD [32] ve RetinaNet [33] gibi birçok nesne algılama modelleri önerilmektedir [34]. Tek aşamalı dedektörler regresyon fikrine dayanır. Bu yöntemler ile önceden önerilen bölgeler oluşturmadan doğrudan giriş görüntüsünden nihai tahmin sonuçları çıkarılır. Genelde nesne algılama probleminde tek aşamalı dedektörler tarafından çok daha hızlı işlem sürelerinde yeterli doğruluk düzeyi sağlanır. YOLO ve ailesi, RetinaNet ve SSD gibi yöntemler tek aşamalı nesne dedektörlerinin en önemli örnekleridir. İki aşamalı dedektörler seçici arama algoritması aracılığıyla teklif bölgeleri oluşturur. Oluşturulan teklif bölgelerinin öznitelikleri CNN'ler kullanılarak çıkarılır. Son aşamada ise her bölgedeki nesnelere tahmin etmek ve nesnelere kategorisini belirlemek için SVM sınıflandırıcısı kullanılır. Bu modeller tek aşamalı yöntemlere göre bazen daha yüksek doğruluk elde edilebilir, ancak daha uzun eğitim süreleri alır. R-CNN ailesi iki aşamalı mimariye örnek verilebilir.

Benjdire ve diğ. çalışmalarında insansız hava araçlarıyla trafikteki araçların tespiti ve sayımıyla ilgili çalışma yapmışlardır [35]. Çalışmalarında insansız hava aracı tarafından elde edilen veri setinin [36] eğitimi için R-CNN ve YOLOv3 kullanılarak performansları karşılaştırılmıştır.

Adarsh ve diğ. tek aşamalı ve iki aşamalı nesne tespit yöntemleriyle ilgili genel bir araştırma sunmuşlardır. İki aşamalı nesne algılayıcı algoritmalar R-CNN ailesi ile tek aşamalı nesne algılayıcı algoritmaları olan YOLO ailesi ve SSD'ye değinmişlerdir. İki aşamalı algılayıcıların daha çok sonucun doğruluğuna odaklandığını, tek aşamalı algılayıcıların ise sonuca hızlı erişilmesinde bahsetmişlerdir. Çalışmalarında YOLOv3'ün başka versiyonu olan YOLO v3-Tiny ile hem doğruluk hem de hız açısından diğer yöntemlere göre daha iyi sonuç elde etmişlerdir. Fakat çalışmalarında küçük hedeflerin tespiti için yeterli doğruluğa ulaşamamışlardır [37].

Baygın ve diğ. çalışmalarında AlexNet ve MobileNetV2 birleştirerek yüz ifadesi tanıma hibrit derin öğrenme modelini kullanmışlardır. Çalışmalarında farklı veri setleri kullanarak otomatik yüz ifadesi tanıma yüksek başarı oranları elde etmişlerdir [38]. Müezzinoğlu ve diğ. çalışmalarında bölgesel (patch) tabanlı derin özellik modeller kullanarak beyin tümörü görüntü veri setini sınıflandırma işlemini yüksek doğrulukta gerçekleştirmişlerdir [39]. Süt ve diğ. araştırmalarında BT (Bilgisayarlı Tomografi) görüntülerini kullanarak Addison hastalığını yüksek başarımla sınıflandırmışlardır [40].

Cai ve Vasconcelos çalışmalarında [41] nesne tespiti için çok aşamalı nesne algılama mimarisine Cascade R-CNN ile eğitilmiş bir dizi algılayıcı kullanmışlardır. Algılayıcı çıktısı bir sonraki yüksek kaliteli algılayıcı ile aşamalı eğitilmiştir. Yazarlar önerdikleri yöntemi COCO [42] ve PASCAL VOC [43] veri setleri üzerinde test etmişlerdir. Sonuçlarının tek aşamalı nesne algılayıcılardan daha iyi olduğunu ifade etmişlerdir. Fakat yazarlar yöntemlerini hız açısından tek aşamalı yöntemler ile karşılaştırmamaları çalışmanın eksikliğidir.

1.3 Sperm tespiti alanında incelenen çalışmalar

Sperm analizi yapılırken morfolojisi (morphology), sayısı (count) ve hareketi (motility) gibi spermin üç önemli özelliği üzerine odaklanılır [44]. Sperm analizinde önemli bir değerlendirme kriteri olan hareket analizi öncesinde sperm tespiti yapılır.

Daha önce sperm tespiti ile ilgili yapılan çalışmalarda [45],[46] eşik değeri, kenar saptama gibi birçok geleneksel görüntü işleme tekniklerine dayanmaktadır. Sperm videolarındaki sperm yoğunluğuna göre sperm tespitinin başarısı değişmektedir. Örneğin videoda 10 ve 20 sperm olan bir çalışmada [47] sperm tespit başarısı çok yüksek olmaktadır. Fakat yoğunluk arttıkça sperm hücrelerinin anlık değişim gösteren hareketleri nedeniyle oluşan çarpışmalarından dolayı sperm tespitinde ölçüm hatalarına neden olmaktadır [48].

Hidayetullah ve diğ. çalışmalarında boğa sperm hücrelerine ait videolarda sperm tespiti için YOLO (You Only Look Once) tabanlı DeepSperm adlı bir yöntem önermişlerdir. Önerdikleri yöntemin performansını artırmak için veri artırma ve bazı hiper parametreleri değiştirerek kullanmışlardır. Önerilen yöntemi YOLO v3 ve Mask R-CNN ile karşılaştırmışlardır. Test veri setlerinde %86.91 mAP (mean average precisions) değerine ulaştıklarını belirtmişlerdir [49].

Aggarwal ve diğ. kanser hastalarında sperm tespiti için veri artırma kullanarak derin öğrenme tabanlı YOLOv3 ve Faster R-CNN uygulamışlardır. Eğitim öncesi sperm hücrelerini normal ve anormal olarak etiketlemişlerdir. YOLOv3 ve Faster R-CNN ile sırasıyla 0.37 ve 0.56 mAP sonuçları elde etmişlerdir. Ayrıca çalışmalarında YOLOv3 modelinin eğitim süresi açısından

Faster R-CNN modeline göre çok hızlı olduğunu sonucuna varmışlar [50].

Valiuškaite ve diğ. VISEM (A Multimodal Video Dataset of Human Spermatozoa) videolarında sperm kafasının hareketliliğini değerlendirmek için R-CNN mimarisini kullanmışlardır. Sperm kafasının segmentasyonu için merkezi bir koordinat izleme algoritması (CCTA) önerilmiştir. Önerdikleri yöntemle spermin kafa bölgesini %91,77 doğrulukla belirlemişlerdir [51]. Çalışmanın dezavantajlarından biri, sonuçlarını YOLO gibi diğer popüler algoritmalarla karşılaştırmamış olmasıdır.

Rahimzadeh ve diğ. çalışmalarında [52] sperm video veri setinde [51] sperm tespiti için derin öğrenme tabanlı RetinaNet'i kullanmışlar. Çalışmanın performansını artırmak için Retina ağına girdi olarak verilen birkaç ardışık çerçeveyi (frame) birleştirmişlerdir. Tespit aşamasından sonra izleme için ayırt edici korelasyon filtreleri (DCF) algoritmasını değiştirerek kullanmışlardır. Tespit aşamasında ortalama %96 doğruluk elde etmişler. Çalışmanın dezavantajı sperm tespit sonuçlarının performansını güncel çalışmalarla karşılaştırabilir olmamasıdır.

Hicks ve diğ. çalışmalarında [53] insan sperm hareketliliğinin otomatik olarak tahmini için klasik makine öğrenmesi yöntemleri ile derin öğrenme yöntemlerini karşılaştırmışlardır. Çalışmalarında 85 video veri seti [54] üzerinde eğitim yaparak derin öğrenme modellerinin klasik modellere kıyasla daha hızlı ve tutarlı olduğunu ifade etmişlerdir. Thambawita ve diğ. çalışmalarında [55], sperm başlarının hareket ve morfoloji analizi için otomatik kodlayıcı ile önceden eğitilmiş Resnet-34 CNN gibi iki aşamalı bir mimari önermişlerdir.

Zou ve diğ. çalışmalarında [56] 111 tane kaliteli sperm videosunda [57] Tiny Object Detection-CNN (TOD-CNN) modelini kullanarak %85 mAP başarısı elde etmişlerdir. Chen ve diğ. [57] sperm görüntü ve videolarından oluşan SVIA isimli bir veri seti üzerinde derin öğrenme modelleri kullanarak sperm tespiti, sınıflandırılması ve spermlerin izlenmesi ile ilgili bir çalışma yapmışlardır. Spermlerin tespiti için SSD, RetinaNet, Faster-RCNN ve YOLO gibi yöntemleri karşılaştırarak en yüksek %40 mAP ile YOLO yöntemi tarafından elde edildiği raporlanmıştır. Çalışmanın dezavantajı %40 mAP değeri literatürdeki diğer çalışmalara göre düşük bir oranın olmasıdır.

Gümüşçü ve Tenekeci çalışmalarında [58] İsfahan Doğurganlık ve Kısırlık Merkezinden elde edilen sperm videolarında Gauss Karışım Modeli kullanarak videodaki spermlerin hareketli bölgelerini tespit etmişlerdir. Hareketli bölge tespitinden sonra mevcut konumda bulunan hücrenin sperm hücresi olup olmadığını sınıflandırmışlardır. Çalışmalarında %82 gibi sınıflandırma doğruluğuna ulaşmışlardır. Çalışmalarını aynı veri setinde yapılan diğer çalışmalarla [59], [60], [61] kıyaslayarak en iyi sınıflandırma doğruluğuna ulaştıklarını, işlem sürelerinde ise diğer çalışmalara göre daha uzun sürdüğünü ifade etmişlerdir.

Mifune ve diğ. [62] çalışmalarında micro-TESE isimli sperm veri setinde sperm tespiti için YOLOv5 yöntemi kullanarak %81 mAP değeri elde etmişlerdir. Çalışmalarında 12 tanesini eğitim, 2 tanesini doğrulama diğer 5 tanesini de test olmak üzere toplam 19 video veri seti kullanmışlardır. Çalışmalarını daha önce aynı veri setinde yapılan başka bir çalışma [63] ile kıyaslayarak hem başarı ve hem de zaman açısından daha iyi olduğunu ifade etmişlerdir. Hung ve Carpenter çalışmalarında [64] spermler gibi çıplak gözle görünmeyen parazitlerin tespitini Faster R-

CNN kullanarak gerçekleştirmişlerdir. Çalışmanın en önemli dezavantajı veri setinde yeterli sayıda görüntü olmamasıdır.

Sperm hareketliliğinin doğru analizi için sperm tespiti tartışmasız çok önemlidir. Sperm gibi küçük hücrelerin veya nesnelere bulanık, düzensiz ve hassas konumlarından dolayı mevcut yöntemler ile algılanması zordur. Ayrıca sperm hareket analizi için CASA sistemlerini kullanmak oldukça maliyetli olduğundan dolayı bu çalışmada sperm hareket analizi öncesi sperm görüntülerinin tespitini geleneksel yöntemler yerine derin öğrenme yöntemleri kullanılmıştır. Bu çalışmada daha önce yapılan sperm hücre tespiti ile ilgili çalışmaların doğruluk ve yüksek hesaplama sürelerindeki sorunları çözmeye odaklandık. Sperm tespit ve hareket analizine ilişkin çalışmaların yetersiz olması, literatürde sperm veri setlerinin eksikliğinden kaynaklandığı görülmektedir. Önerilen çalışma kapsamında oluşturulan sperm video veri kümesi ile belirtilen eksikliğin giderilmesi temelde hedeflenmektedir. Ayrıca, oluşturulan veri kümesi üzerinden çalışmalar gerçekleştirilerek, düşük maliyetli ve yüksek başarımlı sperm tespit yöntemi geliştirilmesi hedeflenmiştir. Yapılan çalışmada derin öğrenme tabanlı olan iki aşamalı ve tek aşamalı nesne tespit yöntemleri ile sperm tespit çalışması gerçekleştirildi. Bu modellerin diğer klasik yöntemlere göre avantajı daha hızlı ve makul sonuç alınabilmesidir. Bu yöntemler klasik yöntemleri gibi görüntüyü parçalara ayırarak her bir parçayı kayan pencere yöntemi ile bakmak yerine her pikseli modele bir kez vererek öğrenmesidir. Tek aşamalı yöntem olarak YOLOv3, iki aşamalı yöntem için ise Faster R-CNN yöntemi kullanılarak sperm tespiti noktasında mAP ve eğitim süresi olarak performans karşılaştırmaları gerçekleştirilmiştir.

2 Materyaller ve yöntemler

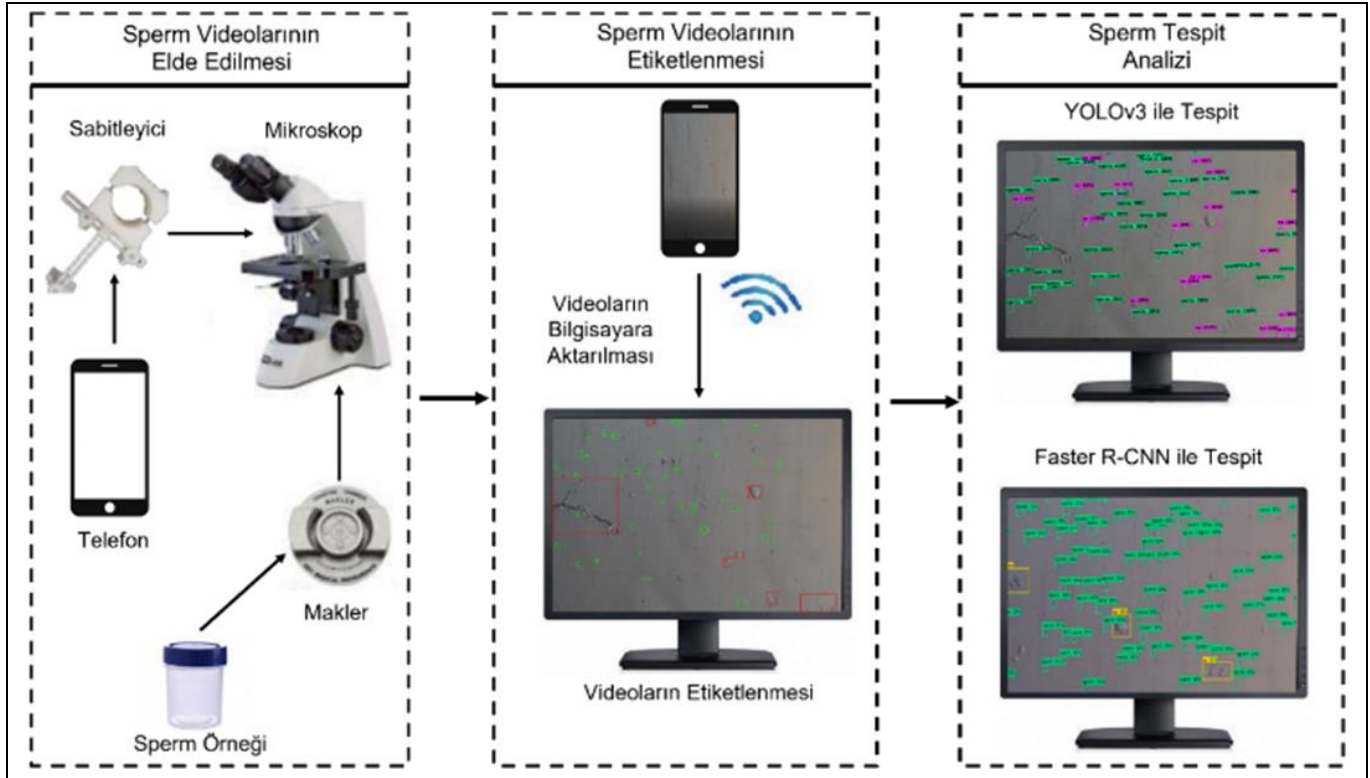
Çalışmada oluşturulan etiketli sperm video veri seti, Recep Tayyip Erdoğan Üniversitesi Tıp Fakültesinden Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu alınarak Şekil 3'te belirtilen veri toplama yöntemi ile elde edilmiştir. Sperm örnekleri 10 farklı hastadan yaşları 19 ile 37 arasında değişen ve minimum 3, maksimum 7 günlük cinsel perhiz sonrasında toplanmıştır. Sperm örnekleri hastalardan masturbasyon yoluyla steril, geniş ağızlı 60 ml'lik örnek kabında muhafaza edildi. Sperm örnekleri analiz öncesi vücut sıcaklığında karıştırıldıktan sonra 5 µL olacak şekilde Makler sayım çemberi üzerine yerleştirilerek analizi sağlandı.

2.1 Önerilen yaklaşım

Yapılan çalışmanın metodoloji Şekil 3'te gösterildiği gibi üç ana kategoride incelenebilir:

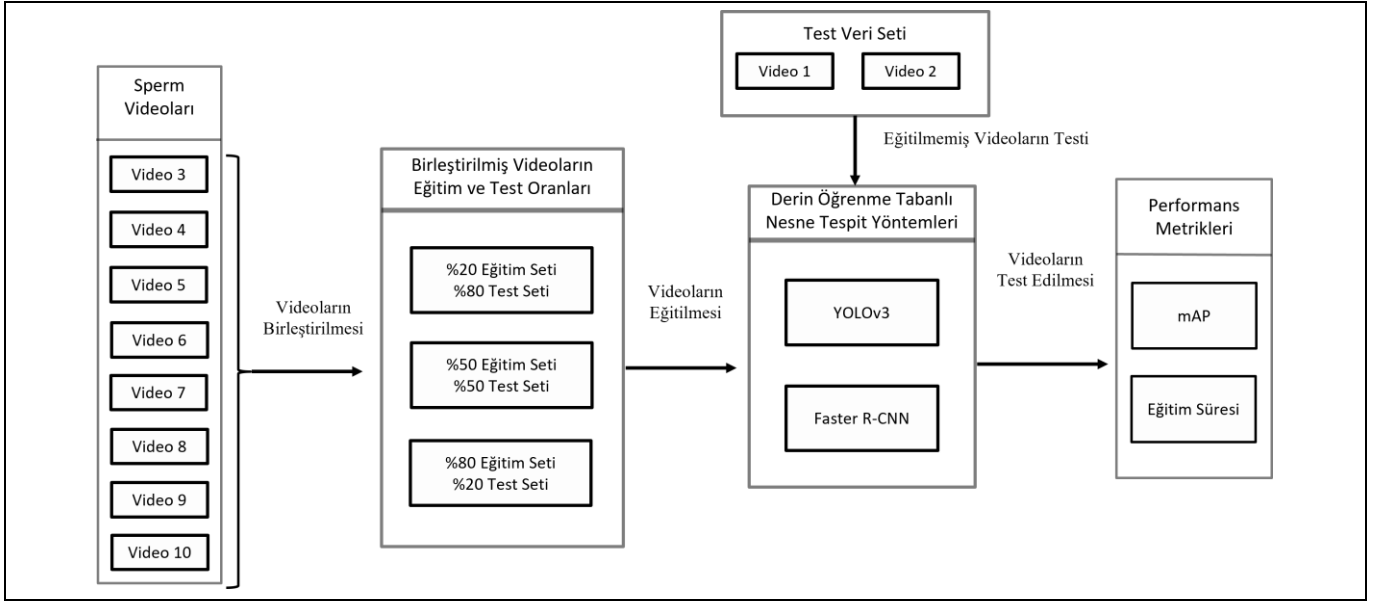
- 1) Verilerin elde edilmesi,
- 2) Elde edilen video görüntülerinin etiketlenmesi,
- 3) Etiketli videoların derin öğrenme tabanlı yöntemler ile eğitilmesi ve test edilmesi.

Önerilen çalışmanın analiz aşamasına ait genel akış şeması Şekil 4'te sunulmaktadır. Elde edilen sperm videolarının sekiz tanesi %20/%50/%80 oranlarında eğitim ve test setlerine ayrıldıktan sonra YOLOv3 ve Faster R-CNN yöntemleri ile ayrı ayrı eğitilmiş ve test edilmiştir. Sonrasında daha önce eğitime hiç katılmamış iki video, yöntemler üzerinde test işlemine tabi tutulmuştur. Sonuçlar mAP ve eğitim süreleri gibi performans metrikleri kullanılarak değerlendirilmiştir.



Şekil 3. Çalışmanın genel metodolojisi (sperm videolarının elde edilmesi, etiketlenmesi ve tespiti).

Figure-3. General methodology of the study (obtaining, labeling and detecting sperm videos).



Şekil 4. Çalışmanın genel akışı.

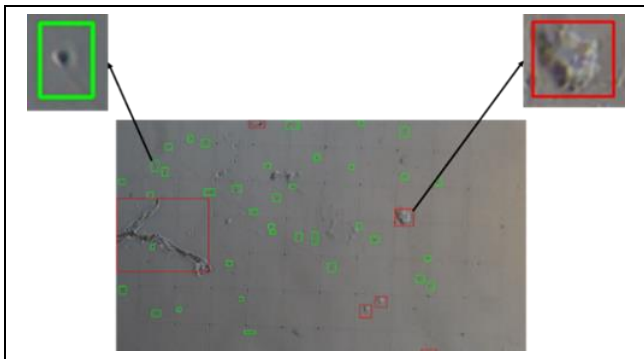
Figure 4. General flowchart of the work.

2.2 Sperm videolarının elde edilmesi

Sperm örnekleri hastalardan elde edildikten sonra Makler odasına aktarıldı ve Olympus CX31 mikroskobunun X100 büyütmesinde incelendi. Akıllı telefon aracılığı ile mikroskopta görüntülenen farklı uzunlukta 10 video Şekil 3'te görüldüğü gibi elde edilmiştir. Böylece analiz öncesinde daha az maliyetli ve kolay taşınabilir veri toplama sistemi tasarlandı.

2.3 Sperm video görüntülerinin etiketlenmesi

Çalışmada 10 farklı hastadan elde edilen 30 FPS (frame per second - saniye başına çerçeve) kaydedilen videoların her biri ön işleme anlamında derin öğrenme modellerinin eğitim sürelerini kısaltmak için 5 FPS ve 640x640 çözünürlükte olacak şekilde yeniden düzenlendi. Ardından uzmanlar yardımı ile LabelIMG [65] programı kullanılarak her bir karedeki (çerçeve) ilgi alanları sperm ve non-sperm (sperm olmayan) olarak etiketlendi. Etiketleme işlemi sonucunda örnek bir çerçeve Şekil 5'te verilmiştir.



Şekil 5. Sperm videolarının etiketleme örneği.

Figure 5. Labeling example of sperm videos.

Örnek görüntü üzerindeki yeşil kutular spermi, kırmızı kutular ise sperm olmayan (non-sperm) yapıları göstermektedir. Böylece her bir çerçevede ortalama 15-25 etiketli sperm görüntüsü bulunmaktadır. Analiz öncesinde kullanılan etiketli

sperm videoları ile ilgili ayrıntılı etiket bilgisi Tablo 1'de verilmiştir.

Tablo 1. Etiketli sperm videolarının özellikleri.

Table 1. Features of labelled sperm videos.

Video ID	Sperm Sayısı	Non-Sperm Sayısı	Toplam Etiket Sayısı	Toplam Çerçeve Sayısı
1	2510	361	2871	60
2	3095	267	3362	51
3	3895	191	4086	54
4	1224	45	1269	13
5	3443	163	3606	47
6	2004	157	2161	150
7	2652	238	2890	150
8	2831	380	3211	150
9	6679	4325	11004	150
10	2579	285	2864	150

2.4 Eğitim ve test oranlarının belirlenmesi

Çalışmada kullanılan 8 video, modellerin eğitimi ve testi için %20 eğitim %80 test, %50 eğitim %50 test ve %80 eğitim %20 test olmak üzere üç farklı oranda kullanılmıştır. Video 1 ve 2 ise hiçbir şekilde model eğitimi kullanılmamış, sadece diğer 8 video tarafından eğitimi tamamlanan model üzerinde test için kullanılmıştır. Böylece çalışmada her bir videonun eğitime katılan oranın mAP performansına etkisi incelenmiştir.

2.5 Derin öğrenme tabanlı nesne tespit yöntemleri

Bu çalışmada sperm tespit probleminde hem Faster R-CNN gibi iki aşamalı hem de YOLOv3 gibi tek aşamalı derin öğrenme tabanlı nesne tespit tekniği kullanılmıştır. Teknikler, üç farklı veri kümesi bölümlenme oranı ile değerlendirilmiştir.

2.5.1 YOLOv3

YOLOv3 kullanılarak sadece bir evrişim sinir ağı ile görüntülerden sınıf olasılıkları ve sınırlayıcı kutunun ofsetleri doğrudan tahmin edilir. Böylece bölge önerilerini tamamen

ortadan kaldırır ve uçtan uca algılama sistemi önerir. Özellik çıkarma için Darknet-53 olarak adlandırılan yeni bir ağ kullanır. YOLOv3 öngörülen sınırlayıcı kutuların başlangıç genişliğini ve yüksekliğini tahmin etmek için k-ortalama küme yöntemini kullanır. YOLOv3 [30] yöntemi ile sınıf tahminleri için ikili çapraz entropi fonksiyonu önerilmiştir. Böylece küçük nesnelere ilgili algılama doğruluğunu önceki yöntemlere göre geliştirilmiştir. YOLOv3 genel yapısı Şekil 6'da verilmiştir.

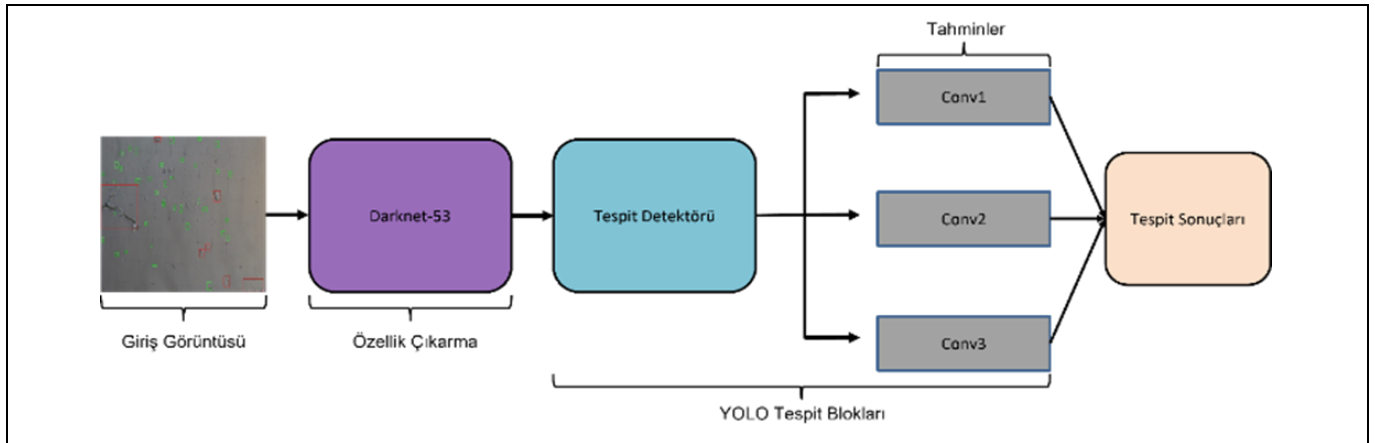
2.5.2 Faster R-CNN

Faster R-CNN, nesne tespiti için VGG-16 isimli bir CNN modelini kullanarak özellik keşfi gerçekleştirir. Daha sonra özellik haritasından bölge önerileri çıkarmak için bir bölgesel teklif ağı kullanır. Faster R-CNN R-CNN ailesinin gelişmiş bir versiyonudur. Teklif bölgeleri oluşturulurken önceki versiyonlar gibi seçici bir arama tekniği yerine bölgesel öneri ağı uygulanır. Faster R-CNN modeli iki modüle ayrılmıştır: RPN ve Fast R-CNN dedektörü. RPN, ikinci modül için girdi görevi gören çoklu ölçekler ve en boy oranları ile bölge önerileri sağlayan tamamen evrişimli bir ağıdır. Ardından nesne önerileri koleksiyonunu oluşturur ve çıktıda öneri başına bir hedef puanı oluşturur. Fast R-CNN dedektörü, giriş olarak çok sayıda ROI kullanır. Alt örnekleme katmanı daha sonra her bir ROI için bir özellik vektörü çıkarır. Bu özellik vektörü, bir dizi tam bağlı

(FC) katmana girer. Son olarak iki çıktı elde edilir. İlk çıktı, farklı nesnelere tahmin edilen bir olasılık dizisidir. İkinci çıktı ise sınırlayıcı kutunun koordinatlarıdır. Fast R-CNN dedektörü bölgesel teklifleri, orijinal görüntüden değil, ROI havuzuna girecek olan son özellik görüntüsünden alır. ROI havuzuna girecek olan özellik görüntüsünün çözünürlüğü, orijinal görüntünün çözünürlüğünden çok daha düşüktür. Bu nedenle, Faster R-CNN'nin hesaplanması, tüm R-CNN ailesinden daha kısadır. Sonuç olarak, Faster R-CNN, RPN ve Fast R-CNN cihazı aynı evrişim katmanlarını paylaştığından, nesne tespiti için tek ve birleşik bir ağ olarak düşünülebilir. Şekil 7'de, Faster R-CNN modelinin ayrıntılı yapısı gösterilmektedir.

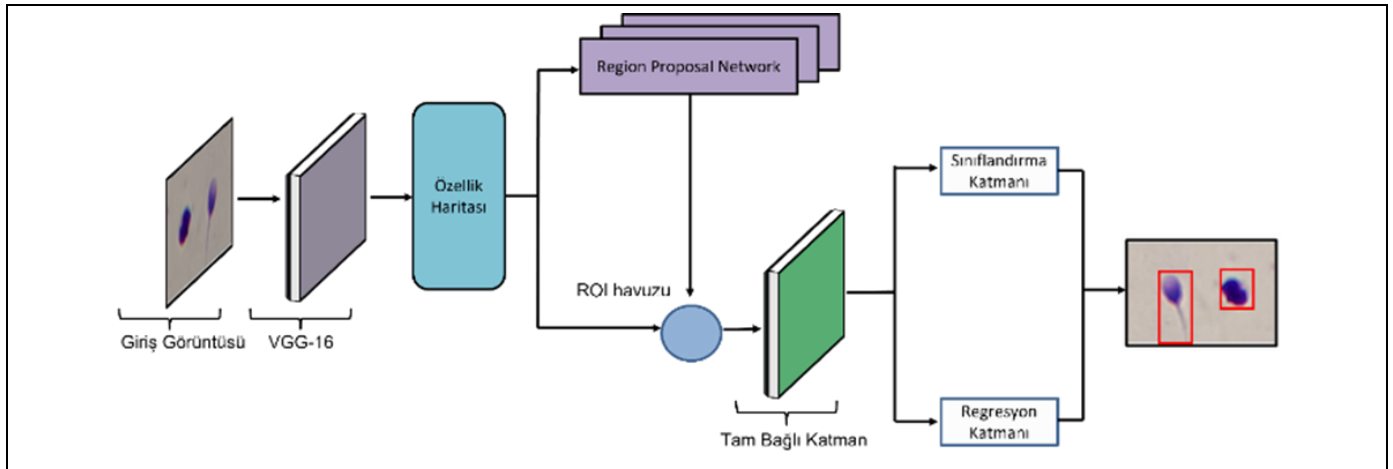
2.6 Nesne tespit algoritmalarının değerlendirilmesi için performans metriği

Nesne algılama problemleri için ortalama kesinlik (mAP) en çok kullanılan popüler bir değerlendirme metriğidir. Model tespitlerinin hem lokalizasyon hem de sınıflandırma performansları hakkında bilgi verir. Bu çalışmada derin öğrenme tabanlı YOLOv3 ile Faster R-CNN nesne tespit tekniklerinin sperm tespit problemi üzerindeki performansı, mAP puanları ve eğitim süreleri olmak üzere iki kriter açısından ölçülmüştür.



Şekil 6. YOLOv3 yönteminin genel yapısı [30].

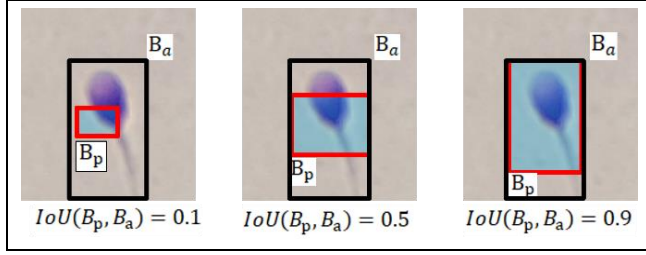
Figure 6. General structure of the YOLOv3 method [30].



Şekil 7. Faster R-CNN modelinin genel yapısı [26].

Figure 7. General structure of the Faster R-CNN model [26].

IoU (Intersection over Union), görüntüde bir nesne dedektörünün doğruluğunu ölçmek için kullanılan bir değerlendirme metriğidir [30]. IoU hesabı, iki sınırlayıcı kutunun kesiştikleri alanın bileşimi alanına bölünmesi ile hesaplanır. IoU hesaplanırken ground truth (kesin referans) kavramı kullanılır. Ground truth olması gereken değerlere denir. IoU hesabı iki bölgenin kesiştikleri alanın (intersection) bu iki dikdörtgenin bileşiminin (union) alanına bölümü olarak hesaplanır. IoU kullanımına üç farklı örnek Şekil 8'de gösterilmektedir. mAP puanı, tüm sınıflar üzerinden veya IoU tarafından tanımlanan bölge içindeki ortalama kesinlik (precision) alınarak hesaplanır.



Şekil 8. Sperm hücresi için üç farklı IoU hesaplama örneği.
Figure 8. Three different IoU calculation examples for sperm cell.

IoU hesaplanırken B_p modelimizin tahmin edilen bölgesini ifade ederken B_a ise ground truth yani hedeflenen bölgeyi ifade eder. Denklem Eş. 1'deki gibi tanımlanır:

$$IoU(B_p, B_a) = \frac{(B_p \cap B_a)}{(B_p \cup B_a)} \quad (1)$$

3 Deneysel sonuçlar

Çalışmada gerçekleştirilen deneyler, GPU (Graphics Processing Unit) hızlandırmalı bir donanım üzerinden Python programı kullanılarak gerçekleştirilmiştir. Anaconda yazılım geliştirme platformu üzerinde Jupyter Notebook kullanılarak yazılım geliştirilmiştir. Derin öğrenme ağ mimarisinde olan nesne tespit algoritmalarının eğitim aşamasında NVidia Geforce 2060 (6 gb) ekran kartı kullanılmıştır. Çalışmada sperm tespiti için videolar, iki farklı derin öğrenme tabanlı nesne tespit tekniği, üç farklı veri kümesi bölümlenme oranı ile iki farklı senaryo altında analiz edilmiş ve model performansları değerlendirilmiştir. Birinci senaryoda video #1 ve #2 modellerin eğitiminde kullanılmamış, diğer videolar ise üç farklı bölünme oranlarında modellerin eğitiminde ve performanslarının değerlendirilmesinde rol almıştır. Birinci senaryoya ait sonuçlar Tablo 2'de gösterilmektedir. İkinci senaryoda ise üç farklı oranda 8 video seti ile birinci senaryoda eğitilen modellerin eğitimde kullanılmayan veriler üzerindeki performanslarının değerlendirilmesi ve hasta bağımsız analizlerin gerçekleştirebildiğinin analizi için video #1 ve #2 sadece test seti olarak kullanılmıştır. İkinci senaryo kapsamındaki sonuçlar ise Tablo 3'te verilmiştir.

Faster R-CNN ve YOLOv3 modellerinin test performanslarında en iyi mAP değerinin elde edildiği görüntü karesi, sonuçların görselleştirilmesi ve analizlerin kıyaslanabilmesi için sırasıyla Şekil 9 ve Şekil 10'da sunulmuştur.

Tablo 2. Senaryo 1 için mAP sonuçları.

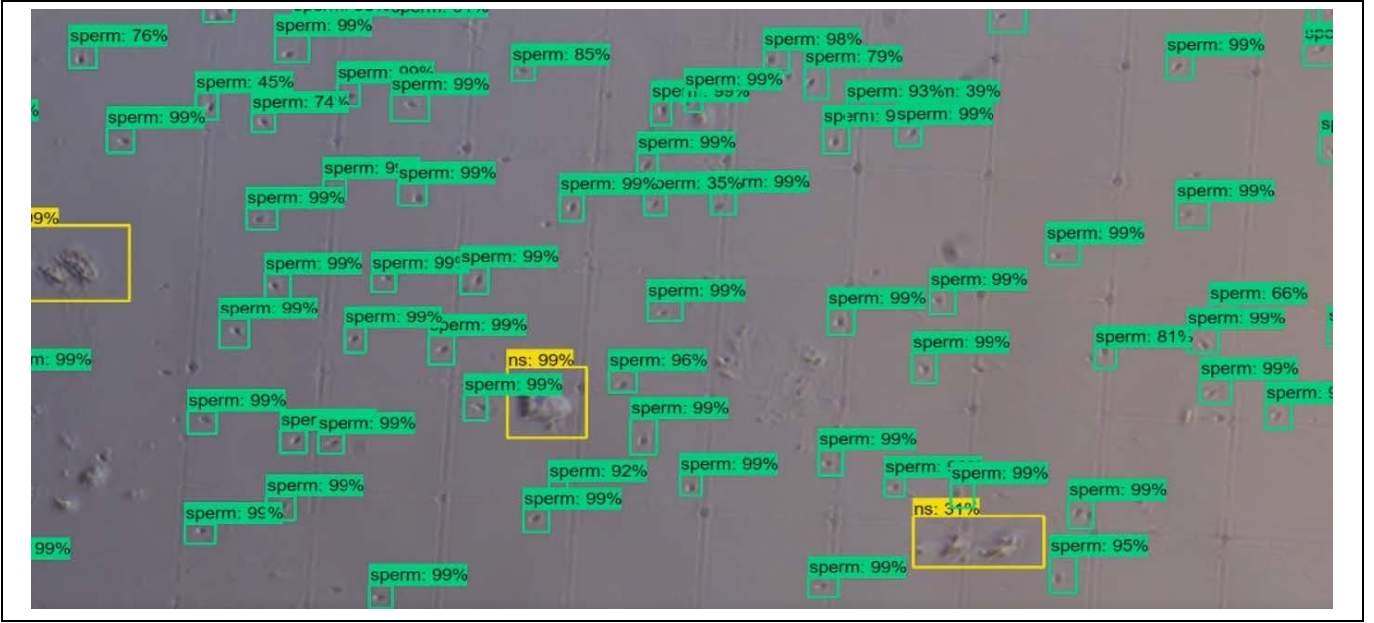
Table 2. mAP results for scenario 1

Senaryo 1	Videolar	Faster R-CNN	YOLOv3
%20 Eğitim %80 Test	3	0.77	0.83
	4	0.69	0.71
	5	0.76	0.82
	6	0.79	0.84
	7	0.80	0.81
	8	0.45	0.47
	9	0.66	0.66
	10	0.80	0.81
	Ortalama mAP	0.785	0.82
	Eğitim Süresi	29h	17h
Ortalama Eğitim Süresi	3.40h	2.10h	
%50 Eğitim %50 Test	3	0.82	0.84
	4	0.73	0.75
	5	0.79	0.85
	6	0.87	0.90
	7	0.81	0.82
	8	0.46	0.48
	9	0.73	0.76
	10	0.82	0.83
	Ortalama mAP	0.82	0.835
	Eğitim Süresi	34h	22h
Ortalama Eğitim Süresi	4.10h	2.50h	
%80 Eğitim %20 Test	3	0.83	0.85
	4	0.79	0.81
	5	0.80	0.85
	6	0.95	0.95
	7	0.93	0.96
	8	0.51	0.53
	9	0.77	0.92
	10	0.82	0.84
	Ortalama mAP	0.825	0.845
	Eğitim Süresi	37h	26h
Ortalama Eğitim Süresi	4.35h	3.10h	

Tablo 3. Senaryo 2 sonuçları.

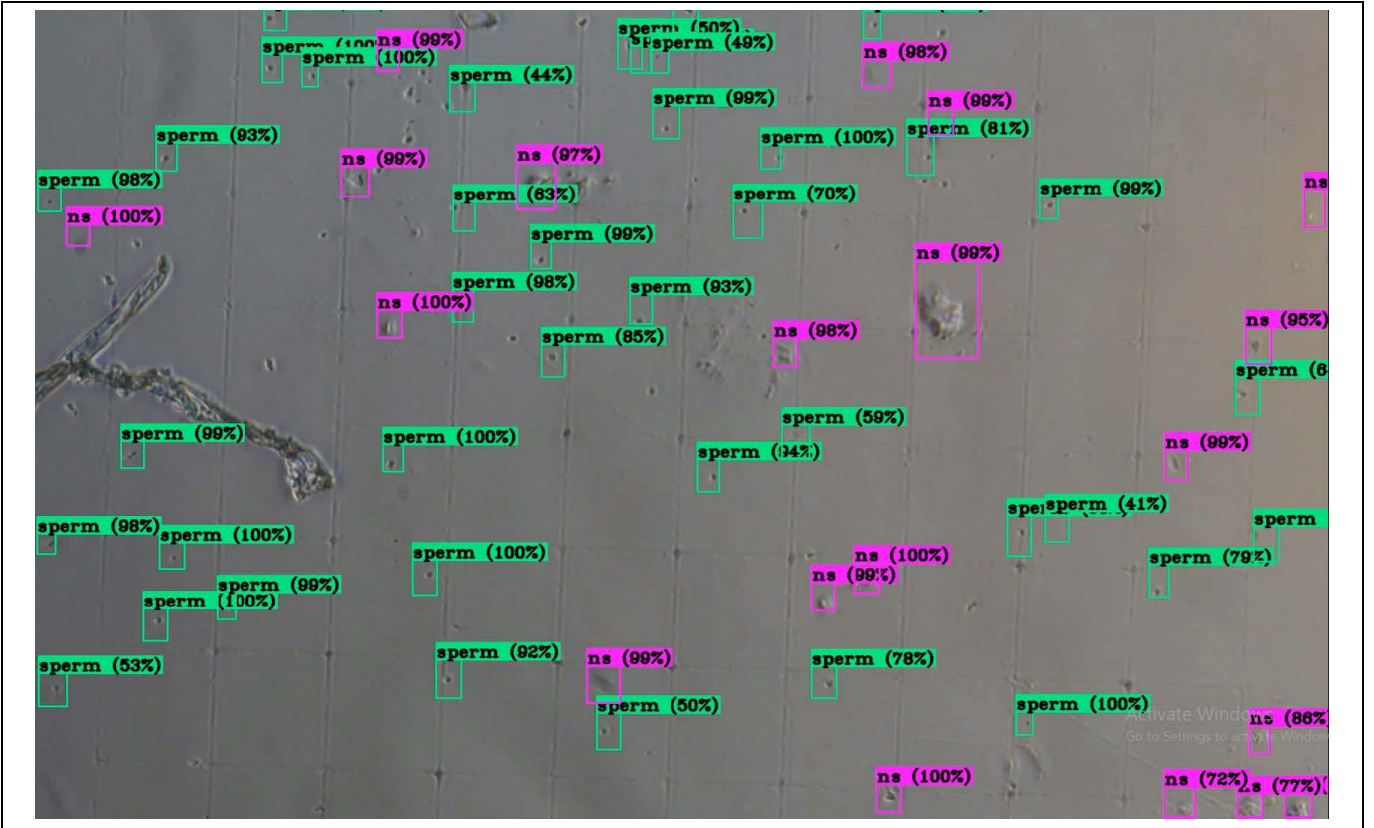
Table 3. Scenario2 results.

Senaryo 2	Videolar	Faster R-CNN	YOLOv3
%20 Eğitim %80 Test	1	0.62	0.84
	2	0.80	0.85
%50 Eğitim %50 Test	1	0.73	0.62
	2	0.83	0.80
%80 Eğitim %20 Test	1	0.81	0.60
	2	0.85	0.78



Şekil 9. Faster R-CNN yöntemi ile elde edilen en iyi sonuç.

Figure 9. The best result obtained with Faster R-CNN method.



Şekil 10. YOLOv3 yöntemi ile elde edilen en iyi sonuç.

Figure 10. The best result obtained with YOLOv3 method.

Birinci analiz senaryosunun sonuçları veri seti eğitim / test bölümlendirme oranları kapsamında incelendiğinde eğitime katılan veri miktarı arttıkça her iki model içinde eğitim süreleri artarken mAP değerleri daha iyi sonuçlar vermiştir. Genel olarak en yüksek mAP değerlerini %80 eğitim %20 test

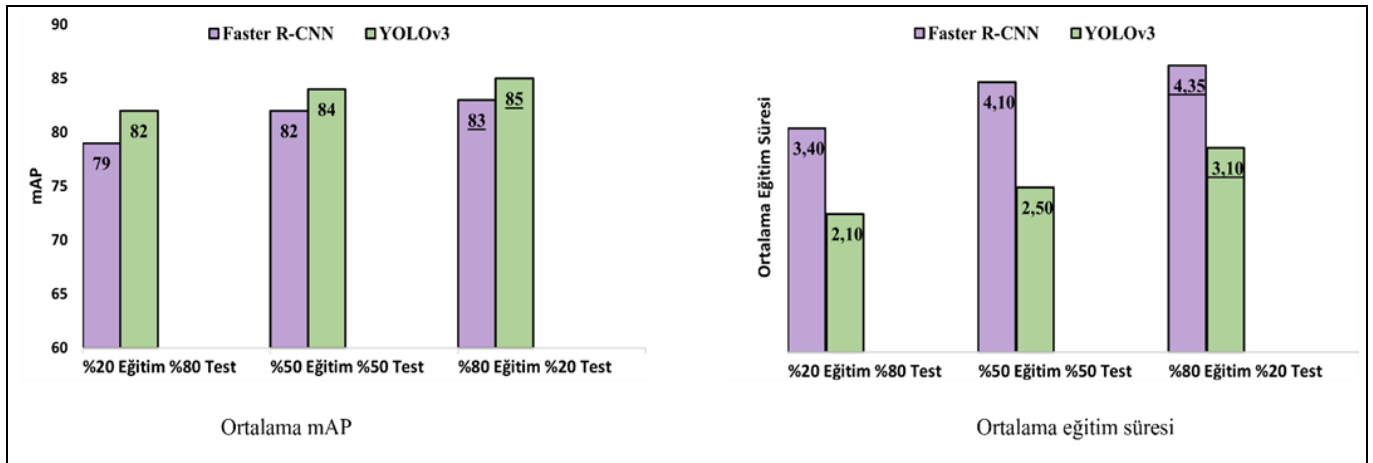
ayarlaması ile elde edilmiştir. YOLOv3 test kümesi dahilinde kullanılan bütün bireysel video analizleri içerisinde Faster R-CNN modeline göre daha yüksek mAP değerine ulaşmıştır. Videoların ayrı ayrı bireysel analiz sonuçları incelendiğinde, video #8 üzerinde her iki model için elde edilen performans

sonuçları diğer videoların analizlerinden daha düşük olmuştur. Bunun nedeni, kameranın titreşiminden dolayı bulanıklık ve Halo etkisidir. Halo etkisi mikroskop ışığı altında oluşan efektler olarak tanımlanır ve biyomedikal mikroskop görüntülerde nesnelerin tanımlanmasında büyük engeldir [66]. Önerilen çalışmada da bu etki, sperm hücrelerinin tam olarak ayırt edilmemesine neden olmaktadır. Yapılan çalışmada modellerin süre bazlı performans kıyaslamasında ise YOLOv3 yöntemi Faster R-CNN yöntemine göre ortalama 1,5 kat daha hızlı eğitim sürecini tamamlamıştır.

Çalışmada gerçekleştirilen ikinci senaryoda, Video #1 ve #2 dışında kalan videoların %20, %50 ve %80 gibi 3 farklı bölümleri eğitime dahil edilmiş ve modellerin daha önceden hiçbir şekilde eğitime katılmamış Video #1 ve #2 üzerinde performansları ölçülmüştür. Sonuçlar Tablo 3'te raporlanmıştır. Elde edilen sonuçlara göre, YOLOv3 modelinin birinci senaryonun %20 eğitilmiş hali Video #1 ve #2 için en yüksek mAP değerine ile sperm tespiti sonucunu verdiği gözlemlenmiştir. YOLOv3 modeli eğitiminde kullanılan test görüntülerinden bağımsız veri oranı arttıkça eğitimde hiç kullanılmamış videoların testinde başarısız olmuştur. Faster R-

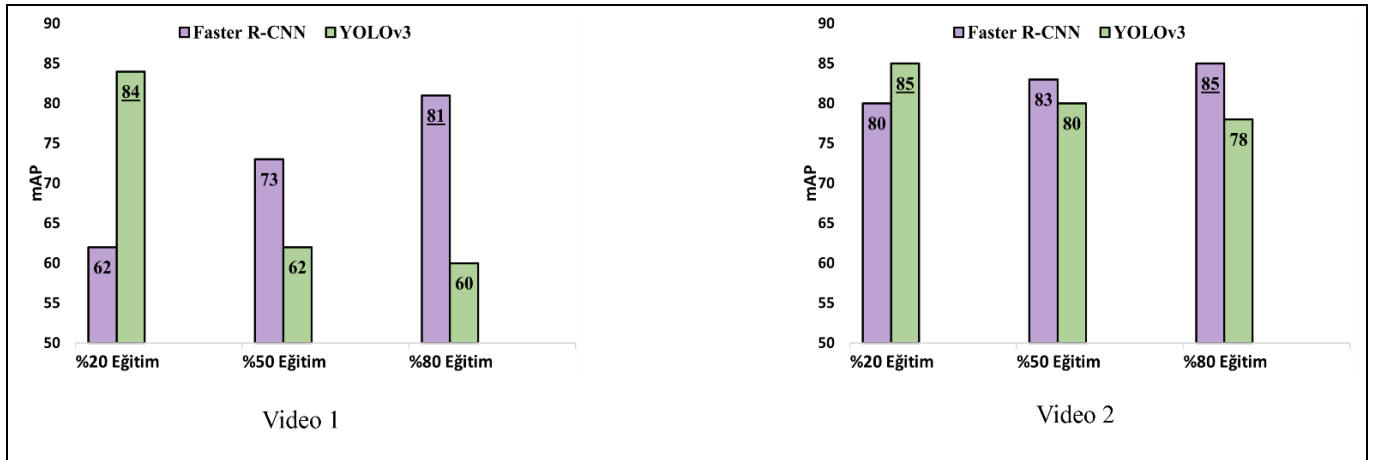
CNN ise YOLOv3'ün aksine eğitime katılan oran arttıkça daha önce hiç görmediği videoların testinde daha yüksek mAP değerlerine ulaşılmıştır. Faster R-CNN ikinci senaryoda Video #1 ve #2 için en yüksek başarısını %80 eğitime katılan videolarla elde etmiştir.

Son olarak, sperm tespiti için modellerin başarısı ve genel olarak eğitim süreleri incelendiğinde, YOLOv3 modeli Faster R-CNN modeline göre eğitim süresi ve mAP değerleri kapsamında hasta odaklı analizlerde (birinci senaryo) daha yüksek performansa sahip olduğu gözlemlenmiştir. Fakat Faster R-CNN modelinin eğitim süresinin YOLO'ya kıyasla daha uzun sürmesine rağmen, daha önce eğitime katılmamış videoların tespitinde (ikinci senaryo - hasta bağımsız analiz yaklaşımı) daha yüksek doğrulukta sperm tespitini gerçekleştirdiği belirlenmiştir. Şekil 11, Tablo 2 üzerinden elde edilmiş ortalama değerlerini (eğitim zamanı ve mAP), Şekil 12 ise Tablo 3'ten elde edilmiş Video1 ve Video2 için mAP performansının grafiksel olarak sunumunu göstermektedir. Faster R-CNN ve YOLO, gerçek zamanlı algılamaya çok yakındır. Fakat tamamen gerçek zamanlı algılama için yüksek hesaplamalı güçlü GPU'lar gerekmektedir.



Şekil 11. Senaryo 1 için sekiz videonun ortalama mAP ve eğitim süresi sonuçları.

Figure 11. Average mAP rates and training times of eight videos analyzed in scenario 1.



Şekil 12. Senaryo 2'de eğitime katılmadan sadece test için kullanılan her iki videonun mAP sonuçları.

Figure 12. mAP results of both videos used only for testing without participating in the training in scenario 2.

4 Tartışmalar ve sonuçlar

Günümüzde kısırlık oranı tüm dünyada giderek artmaktadır. Kısırlıkta erkek faktörleri kritik öneme sahiptir çünkü vakaların yaklaşık %50'si erkek kaynaklı problemlerden kaynaklanmaktadır. Erkek kısırlığı tanısında spermin hareketliliği en bilgilendirici özelliklerden biridir. Hareket analizi öncesi sistemin performansı sperm tespiti doğru yapılması ile doğrudan ilişkilidir.

Bu çalışmada veri toplamanın görsel değerlendirmede olduğu gibi gerçekleştirildiği, ancak analizin bilgisayar tarafından yapıldığı bir sistem önerildi. Bu çalışma kapsamında hem konsantrasyon hem de hareket analizi için kritik bir aşamaya sahip olan sperm tespiti problemi üzerinde derin öğrenme tabanlı nesne tespit tekniklerinin performans değerlendirmesi sunulmuştur. Etiketlenen videolar ile derin öğrenme tabanlı nesne tespit teknikleri hasta odaklı ve hasta bağımsız olarak isimlendirilebilecek iki farklı senaryoda eğitilmiş ve performansları ölçülmüştür. Model performansları, zaman ve doğruluk ölçütleri açısından değerlendirilmiştir.

Sperm tespitinde çalışma kapsamında hazırlanan video veri seti üzerinde elde edilen sonuçlar incelendiğinde en yüksek başarımlar YOLOv3 ile ortalama %84.5 mAP değerleri raporlanmıştır. Bireysel video analizlerinde ise %96 mAP değeri özel hasta odaklı video analizlerinde önerilen çalışma kapsamında elde edilmiştir. Hasta bağımsız olarak nitelendirilen Senaryo 2 çalışması incelendiğinde ise, Faster R-CNN yönteminin daha yüksek bir performans ile %85'e kadar eğitimde kullanılmamış veri setleri üzerinde sperm tespitini sağladığı gözlemlenmiştir. Elde edilen sonuçlar kapsamında, YOLOv3 veya Faster R-CNN gibi derin öğrenme tabanlı nesne tespit tekniklerinin gerçek hayatta uygulanabilir olduğu kanıtlanmıştır.

Gelecekteki araştırmalarda Vision Transformer (ViT), YOLO v8 veya v9 karşılaştırılması, U-Net tarzı simetrik mimarilerin kullanılması ve bu yöntemlerden hibrit bir yöntem geliştirmek hedeflenmektedir. Geliştireceğimiz yöntemi literatürdeki diğer yöntemler ile, aynı veri setleri ve aynı eğitim koşullarında süre ve mAP açısından karşılaştıracacağız. Ayrıca detaylı sperm analizi için sperm tespiti probleminin ardından LSTM gibi yöntemleri kullanarak sperm izlemeye problemine odaklanarak yeni bir sperm izleme ve yörünge analiz yöntemi geliştirmeyi hedefliyoruz.

5 Discussions and conclusions

Infertility rates are progressively increasing worldwide today. Male factors play a crucial role in infertility, as approximately 50% of cases stem from male-related issues. Sperm mobility stands as one of the most informative characteristics in diagnosing male infertility. The system's performance prior to motion analysis is directly linked to accurate sperm detection.

In this study, a system was proposed where data collection is conducted similarly to visual assessment, but the analysis is computer-based. The performance evaluation of deep learning-based object detection techniques for sperm detection, a critical step for both concentration and motion analysis, is presented. The deep learning-based object detection techniques were trained, and their performance was measured on labeled videos under two different scenarios, which can be termed patient-focused and patient-independent. Model performances were evaluated in terms of time and accuracy metrics.

When examining the results obtained on the video dataset prepared within the scope of sperm detection, the highest achievement was reported with YOLOv3, showing an average mAP value of 84.5%. In individual video analyses, a special patient-focused video analysis yielded a mAP value of 96% according to the proposed study. When considering the patient-independent Scenario 2, it was observed that the Faster R-CNN method performed better, achieving up to 85% on previously unused datasets for sperm detection. Based on the achieved results, it is evidenced that deep learning-based object detection techniques such as YOLOv3 or Faster R-CNN are applicable in real-life scenarios.

In future research, we aim to compare Vision Transformer (ViT), YOLO v8 or v9, and U-Net style symmetric architectures, as well as to develop a hybrid approach from these methods. We intend to compare the method we develop with other existing methods in the literature in terms of time and mAP under the same data sets and training conditions. Furthermore, after addressing the sperm detection issue, we aim to focus on sperm tracking by employing methods like LSTM for detailed sperm analysis. Our goal is to develop a novel sperm tracking and trajectory analysis method, concentrating on the challenge of sperm tracking, using techniques such as LSTM, following the sperm detection problem.

6 Teşekkür

Bu çalışma, 2022-B-03 çağrı kodlu ve 27698 Proje No.lu Türkiye Sağlık Enstitüleri Başkanlığı tarafından desteklenmiştir.

7 Yazar katkı beyanı

Gerçekleştirilen çalışmada Mecit YUZKAT, fikrin oluşması, araştırılması, literatür taraması, metodolojisi, deney ve yazma, başlıklarında, Hamza Osman İLHAN, elde edilen sonuçların değerlendirilmesi, incelemesi ve tartışma başlıklarında, Nizamettin AYDIN, yazım denetimi ve içerik açısından makalenin kontrol edilmesi başlıklarında katkı sunmuşlardır.

8 Etik kurul onayı ve çıkar çatışması beyanı

Yapılan çalışmada Recep Tayyip Erdoğan Üniversitesi, Tıp Fakültesi'nden "Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu"ndan alınarak veriler elde edilmiştir. Hazırlanan makalede herhangi bir kişi/kurum ile çıkar çatışması bulunmamaktadır.

9 Kaynaklar

- [1] World Health Organization. *World Health Statistics*. Geneva, Switzerland, World Health Organization, 2010.
- [2] Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. "Evidence for decreasing quality of semen during past 50 years". *British Medical Journal*, 305, 609-613, 1992.
- [3] Petraglia F, Serour GI, Chapron C. "The changing prevalence of infertility". *International Journal of Gynecology & Obstetrics*, 123, 4-8, 2013.
- [4] Agarwal A, Mulgund A, Hamada A, Chyatte MR. "A unique view on male infertility around the globe". *Reproductive Biology and Endocrinology*, 13(1), 1-9, 2015.
- [5] Eze UA, Okonofua FE. "High prevalence of male infertility in Africa: are Mycotoxins to blame?". *African Journal of Reproductive Health*, 19(3), 9-17, 2015.

- [6] Sharlip ID, Jarow JP, Belker AM, Lipshultz, LI, Sigman M, Thomas AJ, Sadvovsky R. "Best practice policies for male infertility". *Fertility and Sterility*, 77(5), 873-882, 2022.
- [7] Thonneau P, Marchand S, Tallec A, Ferial ML, Ducot B, Lansac J, Spira A. "Incidence and main causes of infertility in a resident population (1 850 000) of three French regions (1988-1989)". *Human Reproduction*, 6(6), 811-816, 1991.
- [8] Fainberg J, Kashanian JA. "Recent advances in understanding and managing male infertility". *F1000Research*, 8, 1-8, 2019.
- [9] Alahmar AT. "Role of oxidative stress in male infertility: an updated review". *Journal of Human Reproductive Sciences*, 12(1), 4-18, 2019.
- [10] Sinclair, S. "Male infertility: nutritional and environmental considerations". *Alternative Medicine Review: A Journal of Clinical Therapeutic*, 5(1), 28-38, 2000.
- [11] Olayemi FO. "Review on some causes of male infertility". *African Journal of Biotechnology*, 9(20), 1-12, 2010.
- [12] Macomber D, Sanders, MB. "The spermatozoa count". *New England Journal of Medicine*, 200(19), 981-984, 1929.
- [13] World Health Organisation. *Who Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction*. Cambridge, UK, Cambridge University Press, 1999
- [14] Makler A. "The improved ten-micrometer chamber for rapid sperm count and motility evaluation". *Fertility and Sterility*, 33, 337-338, 1980.
- [15] Smith JT, Mayer DT. "Evaluation of sperm concentration by the hemacytometer method: Comparison of four counting fluids". *Fertility and Sterility*, 6(3), 271-275, 1955.
- [16] Centola, GM. "Semen assessment". *Urologic Clinics*, 41(1), 163-167, 2014.
- [17] Yániz J, Alquézar-Baeta C, et al. "Expanding the limits of computer-assisted sperm analysis through the development of open software". *Biology*, 9(8), 1-16, 2020.
- [18] Urbano LF. *Robust Automatic Multi-Sperm Tracking in Time-Lapse Images*. PhD Dissertation, Drexel University, Philadelphia, USA, 2014.
- [19] Amann RP, Waberski D. "Computer-assisted sperm analysis (CASA): Capabilities and potential developments". *Theriogenology*, 81(1), 5-17, 2014.
- [20] Engel KM, Grunewald S, Schiller J, Paasch U. "Automated semen analysis by SQA Vision® versus the manual approach-A prospective double-blind study". *Andrologia*, 51(1), 1-10, 2019.
- [21] Menkveld R, Stander FS, Kotze TJV, Kruger TF, Zyl JAV. "The evaluation of morphological characteristics of human spermatozoa according to stricter criteria". *Human Reproduction*, 5(5), 586-592, 1990.
- [22] Yeung CH, Nieschlag E. "Performance and comparison of CASA systems equipped with different phase-contrast optics". *Journal of Andrology*, 14(3), 222-228, 1993.
- [23] Padilla R, Netto SL, Da Silva EA. "A survey on performance metrics for object-detection algorithms". *IEEE In 2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*. Niteroi, Brazil, 1-3 July 2020.
- [24] Girshick R, Donahue J, Darrell T, Malik J. "Rich feature hierarchies for accurate object detection and semantic segmentation". In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Ohio, USA, 23-28 June 2014.
- [25] Girshick R. "Fast r-cnn". In *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, 7-13 December 2015.
- [26] Ren S, He K, Girshick R, Sun J. "Faster r-cnn: Towards real-time object detection with region proposal networks". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137-1149, 2016.
- [27] He K, Gkioxari G, Dollár P, Girshick R. "Mask r-cnn". In *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 22-29 October 2017.
- [28] Redmon J, Divvala S, Girshick R, Farhadi A. "You only look once: Unified, real-time object detection". In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Nevada, USA, 26 June-1 July 2016.
- [29] Redmon J, Farhadi A. "YOLO9000: better, faster, stronger". In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Hawaii, USA, 21-26 July 2017.
- [30] Redmon J, Farhadi A. "Yolov3: an incremental improvement". *ArXiv*, 2018. <https://arxiv.org/pdf/1804.02767>
- [31] Bochkovskiy A, Wang CY, Liao HYM. "Yolov4: optimal speed and accuracy of object detection". *ArXiv*, 2020. <https://arxiv.org/pdf/2004.10934>
- [32] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC. "SSD: single shot multibox detector". In *European Conference on Computer Vision*, Amsterdam, Netherlands, 11-14 October 2016.
- [33] Lin TY, Goyal P, Girshick R, He K, Dollár P. "Focal loss for dense object detection". In *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 22-29 October 2017.
- [34] Zou Z, Shi Z, Guo Y, Ye J. "Object detection in 20 years: A survey". *ArXiv*, 2019. <https://arxiv.org/pdf/1905.05055>
- [35] Benjdira B, Khursheed T, Koubaa A, Ammar A, Ouni K. "Car detection using unmanned aerial vehicles: Comparison between faster r-cnn and yolov3". In *2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS)*, Muscat, Oman, 5-7 February 2019.
- [36] Jekhor. "Aerial-Cars-Dataset". <https://github.com/jekhor/aerial-cars-dataset>. (20.08 2022).
- [37] Adarsh P, Rathi P, Kumar M. "YOLO v3-Tiny: Object Detection and Recognition using one stage improved model". In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 6-7 March 2020.
- [38] Baygin M, Tuncer I, Dogan S, Barua PD, Tuncer T, Cheong KH, Acharya UR. "Automated facial expression recognition using exemplar hybrid deep feature generation technique". *Soft Computing*, 27(13), 8721-8737, 2023.
- [39] Muezzinoglu T, Baygin N, et al. "PatchResNet: multiple patch division-based deep feature fusion framework for brain tumor classification using MRI images". *Journal of Digital Imaging*, 36(3), 973-987, 2023.
- [40] Sut SK, Koc M, et al. "Automated adrenal gland disease classes using patch-based center symmetric local binary pattern technique with CT images". *Journal of Digital Imaging*, 36(3), 879-892, 2023.
- [41] Cai Z, Vasconcelos N. "Cascade r-cnn: Delving into high quality object detection". In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Utah, USA, 18-22 June 2018.

- [42] Lin TY, Maire M, et al. "Microsoft coco: Common objects in context". *Computer Vision-ECCV 2014: 13th European Conference*, Zurich, Switzerland, 6-12 September 2014.
- [43] Everingham M, Van Gool L, Williams CK, Winn J, Zisserman A. "The pascal visual object classes (voc) challenge". *International Journal of Computer Vision*, 88(2), 303-338, 2010.
- [44] Lammers J, Chtourou S, Reignier A, Loubersac S, Barrière P, Fréour T. "Comparison of two automated sperm analyzers using 2 different detection methods versus manual semen assessment". *Journal of Gynecology Obstetrics and Human Reproduction*, 50(8), 1-7, 2021.
- [45] Urbano LF, Masson P, VerMilyea M, Kam M. "Automatic tracking and motility analysis of human sperm in time-lapse images". *IEEE Transactions on Medical Imaging*, 36(3), 792-801, 2016.
- [46] İlhan HO, Yuzkat M, Aydın N. "Sperm Motility Analysis by using Recursive Kalman Filters with the smartphone based data acquisition and reporting approach". *Expert Systems with Applications*, 186, 1-12, 2021.
- [47] Hidayatullah P, Mengko TLER, Munir R. "A survey on multisperm tracking for sperm motility measurement". *International Journal of Machine Learning and Computing*, 7(5), 144-151, 2017.
- [48] Broekhuijse MLWJ, Šoštarić E, Feitsma H, Gadella, BM. "Additional value of computer assisted semen analysis (CASA) compared to conventional motility assessments in pig artificial insemination". *Theriogenology*, 76(8), 1473-1486, 2011.
- [49] Hidayatullah P, Wang X, et al. "DeepSperm: A robust and real-time bull sperm-cell detection in densely populated semen videos". *Computer Methods and Programs in Biomedicine*, 209, 1-12, 2021.
- [50] Aggarwal M, Nair VS, Sun T. "Using deep learning to streamline intracytoplasmic sperm injection in cancer Patients". <http://web.stanford.edu/~manav/CS%20231N%20Final%20Paper.pdf> (10.08.2022).
- [51] Valiuškaitė V, Raudonis V, Maskeliūnas R, Damaševičius R, Krilavičius T. "Deep learning based evaluation of spermatozoid motility for artificial insemination". *Sensors*, 21(1), 1-14, 2020.
- [52] Rahimzadeh M, Attar A. "Sperm detection and tracking in phase-contrast microscopy image sequences using deep learning and modified csr-dcf". *ArXiv*, 2020. <https://arxiv.org/pdf/2002.04034>
- [53] Hicks SA, Andersen JM, et al. "Machine learning-based analysis of sperm videos and participant data for male fertility prediction". *Scientific Reports*, 9(1), 1-10, 2019.
- [54] Haugen TB, Hicks SA, et al. "Visem: A multimodal video dataset of human spermatozoa". In *Proceedings of the 10th ACM Multimedia Systems Conference*, Massachusetts, USA, 18-21 June 2019.
- [55] Thambawita V, Halvorsen P, Hammer H, Riegler M, Haugen TB. "Extracting temporal features into a spatial domain using autoencoders for sperm video analysis". *ArXiv*, 2019. <https://arxiv.org/pdf/1911.03100>
- [56] Zou S, Li C, et al. "TOD-CNN: An effective convolutional neural network for tiny object detection in sperm videos". *Computers in Biology and Medicine*, 146, 1-12, 2022.
- [57] Chen A, Li C, et al. "SVIA dataset: A new dataset of microscopic videos and images for computer-aided sperm analysis". *Biocybernetics and Biomedical Engineering*, 42(1), 204-214, 2022.
- [58] Gümüşçü A, Tenekeci ME. "Estimation of active sperm count in spermogram using motion detection methods". *Journal of the Faculty of Engineering and Architecture of Gazi University*, 34(3), 1274-1280, 2019.
- [59] Ravanfar MR, Moradi MH. "Low contrast sperm detection and tracking by watershed algorithm and particle filter". In *2011 18th Iranian Conference of Biomedical Engineering (ICBME)*, Teheran, Iran, 20-21 December 2011.
- [60] Abbiramy VS, Shanthi V, Allidurai C. "Spermatozoa detection, counting and tracking in video streams to detect asthenozoospermia". In *2010 International Conference on Signal and Image Processing*, Chennai, India, 17-19 December 2010.
- [61] Nurhadiyah A, Latifah AL, Fryantoni D, Wirahman T, Wijayanti R, Muttaqien FH. "Comparison and implementation of motion detection methods for sperm detection and tracking". In *2014 International Symposium on Micro-NanoMechatronics and Human Science (MHS)*, Surabaya, Indonesia, 8-10 November 2014.
- [62] Mifune A, Ezaki Y, Saito D, Uto K, Ikeuchi M. "Image-Based Cell Sorting System Using Light-Actuated Microvalve Array". In *2022 IEEE 35th International Conference on Micro Electro Mechanical Systems Conference (MEMS)*, Sapporo, Japan, 23-27 January 2022.
- [63] Sasaki H, Nakata M, Yamamoto M, Takeshima T, Yumura Y, Hamagami T. "Investigation about control of false positive rate for automatic sperm detection in assisted reproductive technology". In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Miyazaki, Japan, 7-10 October 2018.
- [64] Hung J, Carpenter A. "Applying faster R-CNN for object detection on malaria images". In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Hawaii, USA, 21-26 July 2017.
- [65] D. Tzatalin. "Labellmg is a Graphical Image Annotation Tool and Label Object Bounding Boxes in Images". <https://github.com/tzatalin/labellmg> (20.05.2022).
- [66] Nguyen TH, Kandel M, Shakir HM, Best-Popescu C, Arikath J, Do MN, Popescu G. "Halo-free phase contrast microscopy". *Scientific Reports*, 7(1), 1-10, 2017.