

Derin Öğrenme Ağları Kullanılarak Düşük Kaliteli Sperm Görüntülerinin Morfolojik Olarak Sınıflandırması

Morphological Classification of Low Quality Sperm Images Using Deep Learning Networks

Mecit Yüzkat, Hamza Osman Ilhan, Nizamettin Aydın
Bilgisayar Mühendisliği Bölümü
Yıldız Teknik Üniversitesi, 34220 İstanbul, Türkiye
{myuzkat, hoilhan, naydin}@yildiz.edu.tr

Özetçe —İnsanlarda kısırlık teşhisi yapılırken kadın ve erkeğin ayrı üreme hücreleri incelenir. Klinik çalışmalar, erkek kaynaklı kısırlık oranının, genel tanı içerisinde ortalama % 25 - 30 gibi yüksek bir orana sahip olduğunu göstermiştir. Erkek kısırlığı araştırılırken sperm sayısı, hareketi ve morfolojik yapısı değerlendirilir. Morfolojik analizlerde, sağlıklı sonuçlar elde etmek için sperm görüntülerinin detaylı elde edilmiş olması gerekir. Ancak, kameranın titremesi, düşük kalite kamera kullanımı vb. nedenlerden dolayı bu görüntüler kaliteli olmayabilir. Bu çalışmada düşük kalitede elde edilmiş sperm görüntülerinin bulunduğu SCIAN-Morpho verisetinin sınıflandırma başarımını arttırmak amacıyla öncelikle interpolasyon yöntemleri uygulanarak veri kalitesi artırılmıştır. Ardından veri artırma yöntemleriyle, veri dengesizliği problemine çözüm üretilmeye çalışılmıştır. Sınıflandırma aşamasında, evrişimsel sinir ağları, oluşturulan veri setlerine uygulanmıştır. %62 doğruluk, %85 kesinlik ve %75 hassasiyet, son katmanı modifiye edilmiş VGG-19 ağı ile elde edilmiştir.

Anahtar Kelimeler—Kısırlık, Sperm Morfoloji Analizi, İnterpolasyon, Veri Artırma, Transfer Öğrenme, Sınıflandırma

Abstract—The fertility of men and women are examined separately in the diagnosis of infertility. Clinical studies have shown that male infertility rate has a high rate of 25-30% in general diagnosis. Sperm concentration, motility and morphological abnormality are evaluated in male based infertility. In morphological analysis, sperm images should be obtained in detail to obtain objective results. However, the usage of low quality video camera or vibrations occurred in camera module causes to obtain low quality images. In this study, in order to increase the classification performance of the SCIAN-Morpho dataset with low quality sperm images, firstly interpolation methods were applied to increase the data quality. Then, data augmentation techniques have been applied for the data imbalance problem. In the classification phase, pre-trained convolutional neural networks were applied. As a result of the classification, 62% accuracy, 85% precision and 75% sensitivity were obtained by using the VGG-19 networks with the data augmentation and interpolation techniques.

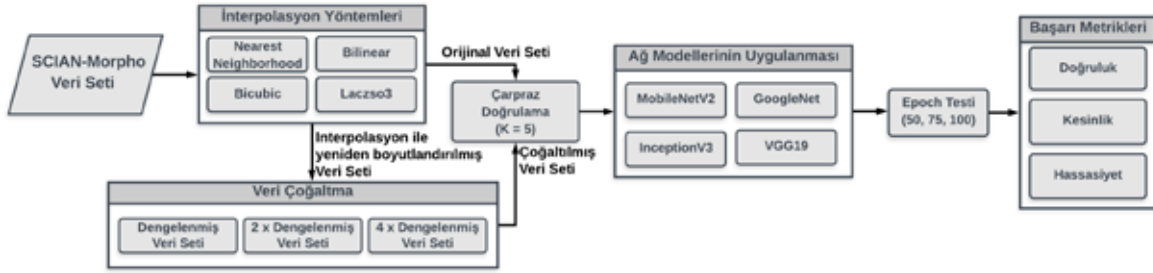
Keywords—Infertility, Sperm Morphology Analysis, Interpolation, Data Augmentation, Transfer Learning, Classification

I. GİRİŞ

Kısırlık ya da infertilite, bir çiftin bir yıl içinde düzenli olarak hiçbir şekilde korunma olmaksızın cinsel ilişkiye girmesine rağmen gebe kalamama durumu olarak tanımlanır [1]. Son 50 yıldır dünyada birçok nedenden dolayı sperm kalitesi düşmekte, infertil birey sayısı artmaktadır. Erkeğin infertil değerlendirilmesi yapılırken ilk aşama sperm analizidir. Bu analizde, spermin morfolojik yapısı, sperm konsantrasyonu ve spermin hareket özellikleri olmak üzere üç teste odaklanılır. Analizler, gözlemsel veya bilgisayar destekli olarak iki şekilde yapılabilmektedir. Gözlemsel teknikte, uzman tarafından bireysel değerlendirmenin çok zor olması, zaman alması ve sonuçlarda gözlemciler arası değişkenliğin gözlemlenmesi nedeniyle bilgisayar destekli sistemlere ihtiyaç duyulmuştur [2]. Bu nedenle sperm kalitesini otomatik olarak ölçmek için cihaz kullanılması gerektiği yönünde güçlü algılar oluşmuştur. 1980'lerin sonlarından itibaren Bilgisayar Destekli Sperm Analiz (CASA) sistemleri sperm kalitesini analiz etmek için geliştirilmektedir [3].

Literatürde, sperm morfoloji analizleri için derin öğrenme tabanlı yaklaşımlar incelendiğinde, İqbal vd. yüksek ve düşük kaliteli sperm hücrelerini sınıflandırmak için özel bir evrişim sinir ağı (CNN) önermişlerdir. Bu modelde çeşitli katmanlar ve çoklu filtre boyutları kullanmışlardır. Önerilen ağda, verimliliği arttırmak için daha az filtre ve parametre uygulamışlardır. Sonuçlarda, diğer yöntemlerden daha iyi performans gösterdiği belirtilmiştir [4]. Fakat, derin öğrenme ağını kullanmadan önce veri seti üzerinde manuel düzenlemeler ve veri setinin bütünlüğünü bozacak manipülasyonlar uygulanması nedeniyle tam otomatik sperm morfoloji analiz sistemi olmadığı vurgulanmıştır.

Bir başka derin öğrenme tabanlı sınıflandırma çalışmasında ise, kaliteli görüntülere sahip sperm morfoloji veri seti üzerinde Mobil platform tabanlı derin öğrenme yaklaşımı olan MobileNetV2 kullanılarak %87 doğruluk başarıları raporlanmıştır [5]. Bu veri setinde yer alan görüntülerin yüksek kalitede olması nedeniyle başarımın doğru orantılı bir şekilde yüksek çıktığı belirtilmiştir.



Şekil 1: Çalışmanın Genel Akışı

CASA sistemlerinde sperm görüntüsü sınıflandırılması yapılırken doğru sonuçlara ulaşılması için genellikle kaliteli görüntüler üzerinde çalışmalar yapılır. Fakat, kameranın kaliteli olmaması, titremesi gibi türlü nedenlerden dolayı kaliteli görüntüler elde edilemez. Kalitesi düşük görüntülerin sınıflandırılması anlamında çalışmalar incelendiğinde, Chang vd. kademeli destek vektör makineleri (Cascade SVM) kullanarak, bu tip görüntülerin sınıflandırılması hakkında çalışmışlardır. Ayrıca, düşük çözünürlük ve kaliteye sahip öznel bir veri seti olan SCIAN-Morpho veri setini çalışmalarında tanıtmışlardır. Önerdikleri yaklaşım ile SCIAN-Morpho veri seti üzerinde %58 doğruluk başarılarını raporlamışlardır [6].

SCIAN-Morpho veri seti üzerinde gerçekleştirilen bir başka çalışmada ise, Shaker et al. sözlük öğrenme ve SVM tabanlı sınıflandırma yaklaşımını önermişlerdir. Bu yöntem ile daha önce %58 olarak raporlanan başarımların %62 doğruluk oranına çıkartılmıştır [7].

Bu çalışmada ise, düşük kaliteli sperm görüntülerinin öncelikle otomatik olarak düzeltilmesi ve uygun derin öğrenme ağlarını kullanarak morfolojik olarak sınıflandırılabilmesi hedeflenmiştir. Başarım metriği olarak doğruluk (accuracy), kesinlik (precision) ve hassasiyet (recall) metrikleri dikkate alınmıştır. Bu bağlamda, çeşitli transfer öğrenme algoritmaları incelenmiş ve test sonuçlarına göre en yüksek başarımları veren ağ üzerinde geliştirmeler gerçekleştirilmiştir.

II. YÖNTEM

Bu çalışmada, düşük kalite ve çözünürlükte sperm görüntüleri içeren SCIAN-Morpho veri setinin yüksek doğruluk, hassasiyetlik ve kesinlik değerleri doğrultusunda sınıflandırılması hedeflenmektedir. Bu anlamda dört derin öğrenme ağı, transfer öğrenme anlamında test edilmiş ve en yüksek sınıflandırma performansını veren VGG-19 CNN üzerinde geliştirmeler yapılarak başarımların artırılması sağlanmıştır. Ayrıca, derin öğrenme mimarilerinde yüksek başarımların elde edilmesi veri seti içerisindeki örnek sayısı ve dağılımı ile ilişkili olduğundan, farklı boyutlara veri artırımı testleri de uygulanarak, sınıflandırma başarımlarının veri boyutu ile olan ilişkisi de araştırılmıştır. Çalışmanın genel akışı Şekil 1'de görülmektedir.

A. Veri Seti

Sperm görüntülerinin veri setleri oluşturulurken genellikle üç farklı uzman her görüntü için bağımsız olarak değerlendirme yapar. Bu çalışma kapsamında düşük çözünürlüğe

sahip olan ve kısmi anlaşmanın olduğu veri seti SCIAN-Morpho kullanılmıştır. Kısmi anlaşmalı veri seti, minimum iki uzmanın aynı sınıfa ait olduğu kararın verildiği örnekleri ifade eder. SCIAN-Morpho veri seti Şili Üniversitesi Tıp Fakültesi Spermogram Laboratuvarı Anatomi ve Gelişim Biyolojisi Programından alınan insan sperm görüntülerinden oluşmaktadır. Veri setindeki görüntüler 63X büyütme ile çekilmiş gri renkli, boyutları 35X35 pikselden ve 1132 adet görüntüden oluşmaktadır. SCIAN veri seti Normal, Tapered, Pyriiform, Amorphous ve Small olmak üzere beş sınıf yapısında düzenlenmiştir. 1132 sperm görüntüsünden 100'ü Normal, 228'i Tapered, 76'sı Pyriiform, 656'sı Amorphous, ve 72'si Small görüntüden oluşur. Örnek sınıf dağılımı incelendiğinde ise düşük kalite ve çözünürlük dışında, veri dengesizliği probleminde olduğu görülmektedir. Bu dengesizliğin ortadan kaldırılması için veri çoğaltma tekniklerinin etkisi de bu çalışma kapsamında incelenmiştir.

B. İnterpolasyon Yöntemleri

Görüntü interpolasyonu, bilinen piksel değerlerinden bilinmeyen piksel değerlerini hesaplama işlemi olarak tanımlanır. Literatürde görüntü interpolasyonu ile ilgili birçok yöntem mevcuttur. Bu yöntemler temel olarak adaptif ve adaptif olmayan görüntü interpolasyonu olmak üzere iki gruba ayrılır. Adaptif olmayan interpolasyon yöntemi görüntünün eksik bileşenlerini kurtarmak için her piksele uygulanan sabit bir hesaplama modeli kullanır [8]. Adaptif olmayan yöntemlerin uygulanması adaptif yöntemlere göre kolay ve hesaplama gereksinimleri çok daha düşüktür. Bu çalışma kapsamında da adaptif olmayan interpolasyon yöntemlerinden, En Yakın Komşu (Nearest Neighborhood), bilinear, bicubic ve lanczos teknikleri uygulanmıştır.

1) *En Yakın Komşu*: İnterpolasyonlu çıkış pikseli hesaplanırken giriş görüntüsündeki en yakın komşunun değeri alınarak hesaplanır. Ölçeklendirmeden sonra yumuşatma yapılmaz. Genellikle düşük kaliteli görüntü verir.

2) *Bilinear*: Bilinear interpolasyon yöntemi ile son interpolasyon değeri hesaplanırken en yakın 4 komşu pikselinin ağırlıklı bir ortalaması kullanılır. Sonuç olarak orijinal görüntüden çok daha düzgün bir görüntü elde edilir. Bu tekniğin avantajı 4 yakın komşu pikseli kullandığından Nearest Neighborhood yöntemine kıyasla daha doğru bir görüntü ortaya çıkarır. Dezavantajı ise 4 piksel kullandığından işlem karmaşıklığı daha fazladır.

3) *Bicubic*: Bicubic interpolasyon yöntemi ile görüntü boyutu değiştirilirken en yakın 16 komşu pikseli bir ağırlıklı ortalama kullanır. Bicubic, Nearest Neighborhood ve Bilinear

yöntemlerinden daha keskin görüntüler verir. Ayrıca bicubic yöntemi diğer tüm interpolasyon yöntemlerine göre oluşan bulanıklığı ve merdiven şeklindeki kenarları ortadan kaldırır. Ancak, daha fazla hesaplama süresi alır.

4) *Lanczos*: Lanczos, kaynak görüntüdeki noktanın yakınındaki 36 pikselin yoğunluklarını kullanır. Karmaşık bir yaklaşımdır.

C. Veri Çoğaltma

Aynı görüntü için farklı değişiklikler yapılarak bir görüntüden birden fazla sentetik görüntü elde etme yöntemine veri çoğaltma denir. Geleneksel makine öğrenmesi yöntemlerinde sınıflandırma performansı verinin miktarı ile belirli bir değere kadar doğru orantılıdır. Oluşturulan fazla sentetik veri ise performans düşümüne neden olur. Veri artırma yöntemiyle farklı tür veriler sisteme öğretilerek sürekli aynı veri verilmediğinden ezberlemenin de önüne geçilmiş olunur. Görüntü verileri çoğaltılırken birçok yöntem mevcuttur. Bu yöntemler görüntünün boyutunu büyütmeye ya da küçültmeye, görüntüyü belirli açılarla çevirmeye, görüntünün kontrast değişikliği, görüntüye gürültü eklemeye veya kaldırma gibi sayısız türleri deneyerek veri seti artırılabilir.

Çalışmada kullanılan veri setindeki görüntülerin sınıf dağılımları dengesiz bir dağılıma sahiptir. En fazla görüntü olan Amorphous sınıfında 656 görüntü varken en az görüntü olan Small sınıfında ise 72 görüntü vardır. Genellikle, daha iyi eğitim aşaması için sınıflardaki verilerin eşit olması tercih edilir. Bu nedenle, veri çoğaltma yöntemleri kullanılarak veri kümesindeki her sınıfın görüntü sayısı Amorphous sınıfının verilerinin sayısına eşitlenmiş ve farklı bir veri seti olarak kaydedilmiştir. Ayrıca dengelenmiş veri seti, 2 ve 4 katlarına çıkartılarak performansa etkisi de incelenmiştir.

D. Transfer Öğrenme

Transfer öğrenme, farklı problem çözümünde kazanılan bilginin benzer problemlerde kullanılması olarak açıklanabilir [9]. Örneğin böcek ailesi içerisinde karıncaları tanımayı öğrenirken edinilen bilgiler karıncalara benzeyen diğer böceklerin tanınmasında transfer öğrenme kapsamında kullanılabilir. Transfer öğrenme, nesne tanıma ve sınıflandırma problemlerinde, daha fazla eğitim verisine veya daha iyi modellere ihtiyaç duyulduğu durumlarda sıfırdan öğrenme yerine kullanılır. Eğitimde ağlar rastgele ayarlanmış parametreler yerine önceden eğitilmiş parametrelerle başlatılır. Önceden eğitilmiş ağların, geleneksel bir şekilde sıfırdan eğitilmiş ağlardan daha iyi çalıştığı çalışmalarda kanıtlanmıştır [10]. Bu ağlar sabit katmanlara sahiptir. Herhangi bir ağ modeli ImageNet gibi büyük bir veri setinde eğitildiğinde sinir ağının tüm parametreleri eğitilir ve böylece model öğrenme yeteneğini kazanmış olur. Eğitilmiş model, ağırlıkları ile beraber bir başka problem uzayında sadece belirlenen katman sayısının ağırlıklarının güncellenmesi üzerine ince ayar (fine-tuning) işlemine tabi tutulur. Böylece modelin transfer öğrenimi şeklinde kullanılması mümkün hale gelir.

Bu çalışmada transfer öğrenme kapsamında SCIAN-Morpho veri seti sınıflandırmasındaki ImageNet görüntü veri seti ile önceden eğitilmiş GoogLeNet, VGG-19, InceptionV3 ve MobileNetV2 ağlarının performans değerlendirilmesi yapılacaktır.

E. Eğitim-Test Verilerinin Hazırlanması

Çapraz doğrulama bir modelin gerçek test verileriyle nasıl sonuç üreteceğini tahmin etmek için kullanılan istatistiksel bir doğrulama yöntemidir. Bu teknik veri setindeki tüm görüntülerin hem eğitim hem de test için kullanılmasını sağlar [11]. Bu çalışmada kullanılacak veri seti öncelikle çapraz doğrulama ile $K = 5$ yapısına dönüştürülmüştür. toplam 1132 görüntünün bulunduğu dosyanın her tekrarlama % 80'ni eğitim için %20'si test için kullanılmıştır.

III. DENEYSEL SONUÇLAR

A. Derin Öğrenme Ağlarının Transfer Öğrenme Analizleri

Çalışma kapsamında MobileNetV2, InceptionV3, GoogleNet ve VGG-19 derin öğrenme ağları transfer öğrenme anlamında SCIAN-Morpho veri setinin orijinal versiyonunun sınıflandırılması üzerine test edilmiştir. Sonrasında belirlenen model üzerinde ince ayarlar yapılarak en iyi sonuçları elde etmek hedeflenmiştir. Performans ölçütü olarak karışıklık matrisi (confusion matrix) tabanlı doğruluk, kesinlik ve hassasiyet ölçütleri kullanılmıştır. Bütün ağlar Tablo I'de belirtilen parametreler ile test edilmiştir.

Tablo I: Kullanılan Ağların Parametre Değerleri

Parametre	Değerler
Batch Boyutu	8
Aktivasyon Fonksiyonu	SoftMax
Optimizör	Adam
Loss Function	Categorical Crossentropy
Öğrenme Katsayısı	0.001
Veri Çoğaltım Boyutları	Orijinal Veri Seti Dengelenmiş Veri Seti 2 x Dengelenmiş Veri Seti 4 x Dengelenmiş Veri Seti
Epoch Sayısı	50, 75, 100

Tablo II'de görüldüğü gibi test edilen ağ modellerinden en iyi doğruluk başarısını InceptionV3 kullanılarak elde edildiği görüne de karışıklık matrisinden elde edilen diğer performans metriklerine bakıldığında, VGG-19 ağı daha dengeli sonuçlar vermiştir. Bu yüzden doğruluk, kesinlik ve karışık matrisi açısından değerlendirildiğinde en iyi sonucu veren ağ modeli VGG-19 olarak kararlaştırılmıştır. Bu aşamadan sonraki interpolasyon etkisi, epoch sayısı ve veri çoğaltma anlamındaki tüm deneysel sonuçlar VGG-19 ağı kullanılarak elde edilmiştir.

Tablo II: Modellerinin Karşılaştırılması

Ağ Modelleri	Doğruluk	Kesinlik	Hassasiyet
MobileNetV2	0,38	0,65	0,17
InceptionV3	0,59	0,73	0,21
GoogleNet	0,31	0,85	0,26
VGG-19	0,42	0,74	0,37

B. Interpolasyon Tekniklerinin Sınıflandırmaya Etkisi

Veri setimizdeki görüntü boyutları 35X35 gibi çok küçük boyutlara sahiptir. Bu görüntüleri transfer öğrenmede kullanabilmek için görüntü boyutlarının büyütülmesi gerekir. Örneğin VGG-19 için 224X224, InceptionV3 için 299X299 gibi giriş boyutlarında olmalıdır. Görüntü boyutları büyütülürken çeşitli interpolasyon yöntemleri olan En Yakın Komşu, Bilinear, Bicubic ve Sinc fonksiyonunu temel

alan Laczso3 yöntemleri ayrı ayrı uygulanmıştır. Bütün testlerde 50 epoch üzerinden analiz gerçekleştirilmiştir. Uygulanan interpolasyon yöntemlerinin VGG-19 ağı kullanımında sınıflandırmaya olan etkisi Tablo III'de sunulmuştur. En iyi sonuç ise Sinc fonksiyonunu temel alan Laczso3 yöntemiyle elde edilmiştir. Epoch sayısı ve Veri çoğaltma yöntemlerinin sınıflandırmaya etkisi, Laczso3 interpolasyon tekniği ve VGG-19 ağı üzerinden analiz edilmiştir.

Tablo III: VGG-19 Sınıflandırma Başarımları

Interpolasyon Yön.	Doğruluk	Kesinlik	Hassasiyet
Orijinal Görüntüler	0,42	0,74	0,37
En Yakın Komşu Int.	0,43	0,74	0,38
Bilinear Int.	0,46	0,75	0,45
Bicubic Int.	0,47	0,76	0,45
Laczso3 Int.	0,55	0,80	0,64

C. Epoch Sayısının Sınıflandırma Performansına Etkisi

Derin öğrenme tabanlı sınıflandırma uygulamalarında epoch sayısı öğrenmenin tam olabilmesi veya öğrenme sırasındaki ezber probleminden kaçınabilmek için önem teşkil eder. Çalışmamızda, farklı epoch sayıları test edilerek başarımın artırılması hedeflenmiştir. Bu kapsamda, Laczso3 interpolasyon tekniği ile yeniden boyutlandırılmış resimlerin VGG-19 ağı ile sınıflandırılmasında 50, 75 ve 100 epoch sayıları ile testler tekrarlanmıştır. Sonuçlar Tablo IV'de sunulmuştur. Bu sonuçlara göre 75 epoch sayımı kesinlik ve hassasiyet değerleri nedeniyle en verimli epoch sayısı olarak belirlenmiştir. Veri artırma testleri bu epoch sayımı üzerinden gerçekleştirilmiştir.

Tablo IV: Epoch Sayılarına göre Performans Değişimi

Epoch Sayısı	Doğruluk	Kesinlik	Hassasiyet
50	0,55	0,80	0,64
75	0,53	0,86	0,70
100	0,45	0,84	0,68

D. Veri Kümesini Artırma

Derin öğrenme tabanlı çalışmalarda, yüksek başarıları elde etmek için yeterli veri sayısına ve aynı zamanda dengeli veriye sahip olunmalıdır. Bu çalışmada, SCIAN-Morpho veri setinin üç farklı versiyonu veri çoğaltma anlamında elde edilmiştir. Öncelikle veri setindeki sınıf tabanlı görüntü sayıları dengelendi. Sonrasında dengelenmiş veri seti iki ve dört katına çıkartılarak iki farklı versiyonda kaydedildi. Artırılmış veri setlerinde de Laczso3 interpolasyon yöntemi ve 75 epoch sayısı sabitlenerek VGG-19 derin öğrenme ağı üzerinden sınıflandırma başarımları ölçüldü. Elde edilen sonuçlarda veri kümesinin dengelendikten sonra ikiye katlanmasıyla en yüksek başarımlar elde edilmiştir.

Tablo V: Veri Çoğaltmanın Sınıflandırmaya Etkisi

Veri Boyutu	Doğruluk	Kesinlik	Hassasiyet
Orijinal Veri	0,53	0,86	0,70
Dengelenmiş Veri	0,55	0,83	0,71
2X Dengelenmiş Veri	0,62	0,85	0,75
4X Dengelenmiş Veri	0,37	0,64	0,54

IV. SONUÇ

Erkek kısırlığı teşhisi yapılırken öncelikle sperm analizi yapılır. Sperm analizi spermin hareketi, sayısı ve morfolojik özellikleri dikkate alınarak gerçekleştirilir. Sperm görüntüsünün bilgisayar destekli analizinin yapılabilmesi için kaliteli görüntülerin elde edilmesi gerekir. Fakat bazı nedenlerden dolayı kaliteli görüntü elde edilemez. Bu çalışmada düşük kaliteli sperm veri setini sınıflandırmak için öncelikle interpolasyon yöntemleri uygulandı. Ardından önceden eğitilmiş bazı evrişimli sinir ağı modelleri uygulanarak başarı metrikleri ölçüldü. Veri seti görüntülerinin döndürülmesi, kırılması gibi yöntemlerle veri sayısı artırıldı. Son durumda iki katına artırılmış dengeli veri setinin Lanczos interpolasyon tekniği kullanılarak yeniden boyutlandırılması ve 75 epochlu VGG-19 uygulanarak %62 Doğruluk, %85 Kesinlik ve %75 Hassasiyet değerlerinde sınıflandırılması sağlanmıştır. Çalışmamızın avantajı veri seti üzerinde yapılan tüm değişikliklerin otonom yapılmasıdır. Çalışmanın dezavantajı ise çok yüksek başarı elde edilememesidir. Fakat elde ettiğimiz değerler düşük kaliteli sperm görüntüleri için kabul edilebilir bir sonuçtur.

Sonraki çalışmalarımızda düşük çözünürlüklü görüntüler için özel evrişimli sinir ağı geliştirilmesi hedeflenmektedir. Ayrıca bu çalışmada kullandığımız adaptif olmayan interpolasyon yöntemleri yerine adaptif interpolasyon yöntemleri uygulanarak başarımın artırılması amaçlanmaktadır.

KAYNAKÇA

- [1] W. H. Organisation, *WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction*. Cambridge university press, 1999.
- [2] H.-F. Yang, X. Descombes, S. Prigent, G. Malandain, X. Druart, and F. Plouraboué, "Head tracking and flagellum tracing for sperm motility analysis," in *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2014, pp. 310–313.
- [3] L. F. Urbano, *Robust automatic multi-sperm tracking in time-lapse images*. Drexel University, 2014.
- [4] I. Iqbal, G. Mustafa, and J. Ma, "Deep learning-based morphological classification of human sperm heads," *Diagnostics*, vol. 10, no. 5, p. 325, 2020.
- [5] H. O. Ilhan, I. O. Sigirci, G. Serbes, and N. Aydin, "A fully automated hybrid human sperm detection and classification system based on mobile-net and the performance comparison with conventional methods," *Medical & Biological Engineering & Computing*, pp. 1–22, 2020.
- [6] V. Chang, A. Garcia, N. Hirschfeld, and S. Härtel, "Gold-standard for computer-assisted morphological sperm analysis," *Computers in biology and medicine*, vol. 83, pp. 143–150, 2017.
- [7] F. Shaker, S. A. Monadjemi, J. Alirezaie, and A. R. Naghsh-Nilchi, "A dictionary learning approach for human sperm heads classification," *Computers in biology and medicine*, vol. 91, pp. 181–190, 2017.
- [8] A. N. Htwe and M. M. Nge, "Image resampling using non-adaptive interpolation." Fourth Local Conference on Parallel and Soft Computing, 2009.
- [9] L. Torrey and J. Shavlik, "Transfer learning," in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*. IGI global, 2010, pp. 242–264.
- [10] R. Mehra *et al.*, "Breast cancer histology images classification: Training from scratch or transfer learning?" *ICT Express*, vol. 4, no. 4, pp. 247–254, 2018.
- [11] J. D. Rodriguez, A. Perez, and J. A. Lozano, "Sensitivity analysis of k-fold cross validation in prediction error estimation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 3, pp. 569–575, 2009.