T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MİKROBİYOLOJİK GÖRÜNTÜLERİN İNCELENMESİ: BİLGİSAYAR DESTEKLİ SPERM TESPİTİ VE MORFOLOJİ ANALİZİ

Mecit YÜZKAT

DOKTORA TEZİ
Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Programı

Danışman
Prof. Dr. Nizamettin AYDIN

Eş-Danışman
Dr. Öğr. Üyesi Hamza Osman İLHAN

Haziran, 2022

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MİKROBİYOLOJİK GÖRÜNTÜLERİN İNCELENMESİ: BİLGİSAYAR DESTEKLİ SPERM TESPİTİ VE MORFOLOJİ ANALİZİ

Mecit YÜZKAT tarafından hazırlanan tez çalışması 29.06.2022 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği Programı **DOKTORA TEZİ** olarak kabul edilmiştir.

Prof. Dr. Nizamettin AYDIN
Yıldız Teknik Üniversitesi
Danışman

Dr. Öğr. Üyesi Hamza Osman İLHAN
Yıldız Teknik Üniversitesi
Eş-Danışman

Jüri Üyeleri

Prof. Dr. Nizamettin AYDIN, Danışman
Yıldız Teknik Üniversitesi _____

Prof. Dr. Zümray DOKUR ÖLMEZ, Üye
İstanbul Teknik Üniversitesi _____

Doç. Dr. Farzad KİANİ, Üye
İstinye Üniversitesi _____

Doç. Dr. Serkan AYVAZ, Üye
Yıldız Teknik Üniversitesi _____

Dr. Öğr. Üyesi Yunus Emre SELÇUK, Üye
Yıldız Teknik Üniversitesi _____

Danışmanım Prof. Dr. Nizamettin AYDIN sorumluluğunda tarafımca hazırlanan Mikrobiyolojik Görüntülerin İncelenmesi: Bilgisayar Destekli Sperm Tespiti ve Morfoloji Analizi başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Mecit YÜZKAT

İmza

Bu çalışma Yıldız Teknik Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü tarafından desteklenmiştir. Grant No: FKD-2021-4554

Rahmetli anneme ithafen

TEŞEKKÜR

Tez çalışmamın tüm süreci boyunca şahsıma gösterdiği hoşgörü ve sabrı için, aynı zamanda her türlü bilimsel konuda bilgisi, deneyimi, önerileriyle beni yönlendiren ve yardım eden saygıdeğer danışmanım Prof. Dr. Nizamettin AYDIN'a şükranlarımı sunmayı borç bilirim.

Tez aşamasında, anlayışlı tutumuyla, araştırma konusunda bilgi ve becerimin geliştirilmesinde rol alan, zamanını esirgemeyen, çalışmamın değer kazanmasını sağlayan, birlikte çalışma deneyimi yaşadığım değerli hocam Dr. Öğr. Üyesi Hamza Osman İLHAN'a en içten teşekkürlerimi sunarım.

Ayrıca katkılarından dolayı değerli hocalarım Prof. Dr. Zümray ÖLMEZ'e, merhum Doç. Dr. Veli HAKKOYMAZ'a, Dr. Öğr. Üyesi Yunus Emre SELÇUK'a ve Öğr. Gör. Dr. Ahmet ELBİR'e teşekkürlerimi sunarım.

Yıldız Teknik Üniversitesi Bilgisayar Mühendisliği Bölümü akademik ve idari personeline desteklerinden dolayı teşekkür ederim.

Bana daima destek ve yardımcı olan sevgili eşime, bu süreçte yeterince ilgilenemediğim biricik oğlum Miran Batın'a ve tüm aileme teşekkürlerimi sunarım.

Mecit YÜZKAT

İÇİNDEKİLER

SİMGE LİSTESİ ix

KISALTMA LİSTESİ x

ŞEKİL LİSTESİ xiii

TABLO LİSTESİ xv

ÖZET xvi

ABSTRACT xviii

1 GİRİŞ 1

1.1 Literatür Özeti 1

1.1.1 Sperm Morfolojisi ile İlgili Literatür Çalışmaları 2

1.1.2 Sperm Hareketi ile İlgili Literatür Çalışmaları 7

1.1.3 Sperm Sayımı ile İlgili Literatür Çalışmaları 11

1.2 Tezin Amacı 12

1.3 Hipotez 12

2 GÖRÜNTÜ İŞLEME 14

2.1 Görüntü İşlemenin Temel Adımları 14

2.1.1 Görüntünün Elde Edilmesi 14

2.1.2 Görüntünün Ön İşlenmesi 15

2.1.3 Görüntü Bölütleme 17

2.1.4 Öznitelik Çıkarılması 21

2.1.5 Görüntünün Sınıflandırılması 22

3 DERİN ÖĞRENME YÖNTEMLERİ İLE GÖRÜNTÜ İŞLEME 23

3.1 Yığılmış Otomatik Kodlayıcı 23

3.2 Derin İnanç Ağı 24

3.3 Tekrarlayan Sinir Ağı 24

3.4 Evrişimsel Sinir Ağı 24

3.5 Evrişimsel Sinir Ağı'nın Temel Bileşenleri 27

3.5.1 Toplu Normalleştirme - Batch Normalization 27
3.5.2 Aktivasyon Fonksiyonları 27
3.5.3 Optimizasyon Seçimi 28
3.5.4 Dropout 28
3.5.5 Veri Sayısını Çoğaltma 29
3.5.6 Transfer Öğrenme 30
3.6 Evrişimsel Sinir Ağlarında Parametre Seçimi 32
3.6.1 Model Oluşturulurken Kullanılan İlgili Hiper Parametreler . . . 32
3.6.2 Eğitim İle İlgili Hiper Parametreler 35
3.7 Sınıflandırma Sonuçlarının Değerlendirilmesi 38

4 TEZ KAPSAMINDA KULLANILAN MATERYALLER 40
4.1 Morfoloji Analizinde Kullanılan Veri Setleri 41
4.2 Hareket Analizinde Kullanılan Veri Setleri 44

5 CNN MODELİ OLUŞTURMA ve UYGULAMA 45
5.1 Veri Artırma 45
5.2 *k* - Çapraz Doğrulama 47
5.3 CNN Modellerinin Tasarımı 47
5.4 Modeller için Eğitim Parametrelerinin Belirlenmesi 48
5.5 Füzyon Yöntemleri 50
5.6 Performans Metrikleri 50
5.7 Deneysel Sonuçlar 50

6 TRANSFER ÖĞRENME KULLANILARAK SPERM GÖRÜNTÜLERİNİN SINIFLANDIRILMASI 59
6.1 Düşük Kaliteli Sperm Görüntülerinin Morfolojik Olarak Sınıflandırılması 59
6.1.1 İnterpolasyon Yöntemleri 59
6.1.2 Veri Setinin Dengelenmesi ve Artırılması 61
6.1.3 Transfer Öğrenme 61
6.2 Renk Uzay Dönüşümlerinin Sperm Sınıflandırmasına Etkisi 62
6.2.1 Mobil Platform Tabanlı Evrişimsel Sinir Modeli 62
6.2.2 Renk Uzayı Dönüşümü 63
6.3 MHSMA Veri setinin Sınıflandırılması 65

7 DERİN ÖĞRENME YÖNTEMLERİ İLE SPERM VİDEOLARINDA SPERM TESPİTİ 67
7.1 Veri Setinin Etiketlenmesi 68
7.2 Uygulanan Senaryolar 69
7.3 Nesne Tespitinde Kullanılan Derin Öğrenme Modelleri 70

7.3.1 R-CNN 71
7.3.2 Fast R-CNN 71
7.3.3 Faster R-CNN 72
7.3.4 Mask R-CNN 73
7.3.5 YOLOv1 73
7.3.6 YOLOv2 73
7.3.7 YOLOv3 74
7.3.8 YOLOv4 74
7.3.9 YOLOv5 74
7.3.10 Detectron2 74
7.3.11 EfficientDet 74
7.3.12 SSD 75
7.4 Performans Metrikleri 75
7.5 Deneysel Sonuçlar 76
7.5.1 Senaryo 1 Sonuçları 76
7.5.2 Senaryo 2 Sonuçları 80

8 SONUÇ VE ÖNERİLER **83**

KAYNAKÇA **85**

TEZDEN ÜRETİLMİŞ YAYINLAR **98**

SİMGE LİSTESİ

S	Adım sayısı
P	Dolgu
F	Filtre boyutu
K	Filtre sayısı
D	Görüntünün boyut sayısı
W	Görüntünün genişliği
x, y	Görüntünün merkez noktasının koordinatlarını temsil eder
H	Görüntünün yüksekliği
B_p	Tahminin sınırlayıcı kutusu
B_a	Temel gerçeğin sınırlayıcı kutusu

KISALTMA LİSTESİ

BiFPN	Bi-directional Feature Pyramid Network
BRIEF	Binary Robust Independent Elementary Features
CASA	Computer Aided Sperm Analysis
CCTA	Central Coordinate Tracking Algorithm
CED	Canny Edge Detector
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CSH	Coherency Sensitive Hashing
DBN	Deep Belief Network
EB-SAMC	Energy-Biased Stochastic Approximation Monte Carlo
EM	Expectation Maximization
FAST	Features from Accelerated Segment Test
FN	False Negative
FP	False Positive
FPS	Frame Per Second
GPU	Graphics Processing Unit
GT	Ground Truth
GVF	Gradient Vector Flow
HCD	Harris Corner Detection
HGDT	Hybrid Generative-Discriminative Tracker
HSI	Hue Saturation Intensity
HSMA	Human Sperm Morphology Analysis
HSV	Hue Saturation Value

HuSHeM	Human Sperm Head Morphology
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
IoU	Intersection over Union
IVOS	Integrated Visual Optical System
JPDAF	Joint Probabilistic Data Association Filter
kNN	k-Nearest Neighbors
LAB	Lightness, Green to Red and Blue to Yellow
LIN	Linearity
LMS	Least Mean Squares
LoG	Laplacian of Gaussian
mAP	Mean Average Precision
MHSMA	Modified Human Sperm Morphology Analysis dataset
MIR	Maximum Intensity Region
MIT	Massachusetts Institute of Technology
MLP	Multi Layer Perceptron
MRF	Markov Random Field
MSE	Mean Square Error
MSER	Maximally Stable Extreme Regions
PSNR	Peak Signal to Noise Ratio
R-CNN	Region-Based Convolutional Neural Networks
ReLU	Rectified Linear Unit Layer
RGB	Red Green Blue
RLVQ	Relevance Learning Vector Quantization
RNN	Recurrent Neural Network
ROI	Region of Interest
RPN	Region Proposal Network
SAE	Stacked Auto Encoder
SGD	Stochastic Gradient Descent
SIFT	Scale Invariant Feature Transform

SMIDS	Sperm Morphology Image Data Set
SSAMC	Smoothing Stochastic Approximate Monte Carlo
SSD	Single Shot Multibox Detector
SSIM	Structural Similarity Index
SURF	Speed-Up Robust Features
SVM	Support Vector Machines
TN	True Negative
TP	True Positive
VA	Visual Assessment
VCL	Curvilinear Velocity
VISEM	A Multimodal Video Dataset of Human Spermatozoa
VSL	Straight Line Velocity
WCF	Wavelet Co-occurrence Features
WHO	World Health Organization
YOLO	You Only Look Once

ŞEKİL LİSTESİ

Şekil 1.1 Örnek sperm görüntüsü 3
Şekil 1.2 Mikroskop görüntüleme yöntemleri: a) faz kontrastı b) floresan boyama 7
Şekil 2.1 Görüntü işlemenin temel adımları 14
Şekil 2.2 Örnek evrişim işlemi 16
Şekil 2.3 Görüntü bölütleme yöntemleri 17
Şekil 2.4 Farklı eşik değerlerine göre otsu metodu örneği 18
Şekil 3.1 Örnek CNN modeli 25
Şekil 3.2 Örnek havuzlama katmanı 26
Şekil 3.3 Örnek aktivasyon fonksiyonları 28
Şekil 3.4 Seyreltme örneği 29
Şekil 3.5 Veri artırma yöntemlerinin üç farklı veri seti için örnek gösterimi . . 30
Şekil 3.6 Örnek dolgu yapısı 34
Şekil 3.7 Model eğitimi ile ilgili örnek parametreler 35
Şekil 3.8 Örnek karışıklık matrisi 38
Şekil 4.1 Veri elde etme ve analiz yaklaşımı 40
Şekil 4.2 Veri setleri için örnek görüntüler 43
Şekil 4.3 Video veri setleri için örnek görüntüler 44
Şekil 5.1 Önerilen çalışmanın genel akışı 45
Şekil 5.2 Altı farklı veri artırma tekniği 46
Şekil 5.3 Sıfırdan oluşturulan altı modelin ayrıntılı gösterimi 49
Şekil 5.4 Örnek Soft-Voting yöntemi 51
Şekil 5.5 Örnek Hard-Voting yöntemi 51
Şekil 5.6 HuSHeM için veri artırma ve mini batch analizi 52
Şekil 5.7 SMIDS için veri artırma ve mini batch analizi 53
Şekil 5.8 SCIAN için veri artırma ve mini batch analizi 53
Şekil 5.9 3 veri seti için karışıklık matrisi a)SMIDS (Orijinal), b)HuSHeM (Orijinal), c)SCIAN-Morpho (Orijinal), d)SMIDS (Artırılmış), e)HuSHeM (Artırılmış), f)SCIAN-Morpho (Artırılmış) 55
Şekil 6.1 Çalışmanın genel akışı 60
Şekil 6.2 Çalışmanın genel akışı 62

Şekil 7.1 Senaryo 1'in akış şeması 68
Şekil 7.2 Senaryo 2'nin akış şeması 68
Şekil 7.3 Etiketleme örneği 69
Şekil 7.4 IoU örneği 76
Şekil 7.5 Her hasta için senaryo 1'e göre modellerin mAP sonuçları 78
Şekil 7.6 12 hastanın senaryo 1'e göre modellerin ortalama mAP ve eğitim süreleri 80
Şekil 7.7 Hasta 3 ve Hasta 5 için modellerin senaryo 1 ve senaryo 2'ye göre mAP farkları 81

TABLO LİSTESİ

Tablo 4.1 MHSMA veri setinin dağılımı 43
Tablo 4.2 SMIDS, HuSHeM ve SCIAN-Morpho veri setlerinin sınıflara göre dağılımları 44
Tablo 5.1 Veri setlerinin eğitimi için kullanılan parametreler 49
Tablo 5.2 Artırılmış veri setlerinin bireysel ve füzyon sonuçları 52
Tablo 5.3 Modellerin eğitim süreleri (saat) 55
Tablo 5.4 Şekil5.9'dan türetilen performans ölçütleri ve sonuçları 57
Tablo 5.5 Önerilen yöntem ile literatürdeki benzer çalışmaların performansının karşılaştırılması 58
Tablo 6.1 Modeller için parametre değerleri 61
Tablo 6.2 Modeller için parametre değerleri 63
Tablo 6.3 İki çalışmanın doğruluk değerlerinin karşılaştırılması (RGB->HSV) 64
Tablo 6.4 İki çalışmanın doğruluk değerlerinin karşılaştırılması (RGB->LAB) 64
Tablo 6.5 MHSMA veri setinin eğitimi için kullanılan parametre değerleri . . . 65
Tablo 6.6 MHSMA sperm veri setinin sınıflandırma başarısının literatürdeki diğer çalışmalarla karşılaştırılması 66
Tablo 7.1 Her bir video için etiket bilgisi 70
Tablo 7.2 Senaryo 1'in sonuçları 77
Tablo 7.3 Video 3, Video 5 için senaryo 1 ve senaryo 2 sonuçları 82

ÖZET

Mikrobiyolojik Görüntülerin İncelenmesi: Bilgisayar Destekli Sperm Tespiti ve Morfoloji Analizi

Mecit YÜZKAT

Bilgisayar Mühendisliği Anabilim Dalı
Doktora Tezi

Danışman: Prof. Dr. Nizamettin AYDIN
Eş-Danışman: Dr. Öğr. Üyesi Hamza Osman İLHAN

Dünya Sağlık Örgütü'ne göre kısırlık; çiftlerin herhangi bir koruma olmaksızın on iki ay ve daha fazla süre boyunca düzenli cinsel ilişkiye girmelerine rağmen gebeliğin oluşmama durumu olarak tanımlanır. Günümüzde kısırlık dünya nüfusunun yaklaşık %20'sini etkileyen yaygın bir problem haline gelmiştir. Kısırlığın teşhisi yapılırken kadın ve erkeğin ayrı ayrı üreme hücreleri incelenir. Erkeklere ilişkin kısırlık faktörlerinin teşhisinde, laboratuvar ortamında belirli koşullar altında sperm hücrelerinin analizi yapılır. Sperm hücrelerinin analizi yapılırken şekil (morfoloji), sayı (konsantrasyon) ve hareket (motilite) olmak üzere spermin üç önemli özelliğinden faydalanılır. Sperm analizi doktorlar tarafından görsel olarak yapılabileceği gibi bilgisayar destekli sperm analiz sistemleri kullanılarak da yapılabilmektedir. Görsel incelemenin kişiden kişiye farklı sonuçlar vermesi ve maliyetli olması nedeniyle bilgisayar destekli analizlerin önemi her geçen gün artmaktadır. Öte yandan, bilgisayar tabanlı uzman sistemler daha tutarlı ve güvenilirdir. Ancak çok pahalı sistemler oldukları için pek çok laboratuvarda bulunmamaktadır. Tez çalışmasında genel olarak, her iki yöntemin dezavantajlarını ortadan kaldırmak için bilgisayarlı analiz kısmı ile görsel değerlendirme ortamından yararlanan hibrit bir uzman sistem geliştirilmiştir. Geliştirilen sistem ile manuel yönlendirme ve kırpma işlemleri ortadan kaldırılmıştır. Tez kapsamında uzun eğitim sürelerinden kaçınmak adına geleneksel makine öğrenmesi modellerinin yerine derin öğrenme tabanlı yöntemler kullanılmıştır. Geliştirilen sistemde, sperm görüntülerinden sperm morfolojisi analizi ve sperm videolarından sperm tespiti gibi

iki modüler yaklaşım kullanılmıştır. İlk modül olan morfoloji analizinde, sağlıklı sonuçlar elde etmek için sperm görüntülerinin anlaşılabilir olması gerekir. Ancak, kameranın titremesi, düşük kalite kamera kullanımı vb. nedenlerden dolayı bu görüntüler yeterince anlaşılır olmayabilirler. Tez çalışması kapsamında sperm şekli için bilgisayar tabanlı bir analiz yaklaşımının sınıflandırma performansını artırmak için ön işleme adımı olarak farklı renk uzaylarının etkisi ve interpolasyon yöntemlerinin uygulanması araştırılmıştır. Ayrıca, hem en yüksek sınıflandırma doğruluğunu elde etmek, hem de veri dengesizliği ve veri yetersizliği problemlerini çözmek için çeşitli veri artırma işlemleri uygulanmıştır. Daha sonra sperm görüntülerinin morfolojik sınıflandırmasını tamamen otomatikleştirmek için derin öğrenme tabanlı altı farklı Evrişimsel Sinir Ağ (Convolutional Neural Network - CNN) modeli oluşturulmuştur. Ek olarak, bu modeller üzerinde çoğunluk oylaması (hard voting) ve ağırlıklı oylama (soft voting) olmak üzere iki karar düzeyinde birleştirme (fusion) yöntemi uygulanmıştır. Oluşturulan modellerin dışında transfer öğrenme yöntemleri ile önceden eğitilmiş modellerin bazı katmanları modifiye edilerek kullanılmıştır. Önerilen yaklaşımın performansını değerlendirmek için SMIDS, HuSHeM, SCIAN-Morpho ve MHSMA olarak kısaltılan dört sperm morfolojisi veri seti kullanılmıştır. Sperm hücrelerinin morfolojik olarak sınıflandırılması aşamasında doğruluk skorlarına ek olarak Kesinlik (Precision), Duyarlılık (Recall), Özgüllük (Specificity), F-Skor (F-score) ve Kappa gibi birçok performans metriği kullanılmıştır. İkinci modülde ise sperm hücrelerinin tespiti, konsantrasyon ve hareket parametrelerinin belirleme aşaması ele alınmıştır. Bu modülde, sperm örneklerinin konsantrasyon ve hareketlilik özelliklerini elde etmek için sperm tespitinde derin öğrenme tabanlı nesne tespit yöntemleri kullanılmıştır. İlk olarak farklı kısır hastalardan elde edilen örnek sperm videoları, mikroskobun oküler kısmına monte edilen bir akıllı telefon aracılığıyla kaydedilmiştir. Daha sonra videolar içindeki hücreler sperm ve sperm olmayan olarak ayrı ayrı etiketlenmiştir. Etiketlenmiş veri seti üzerinde farklı senaryolar test edilmiştir. Sonuçlar doğruluk (Mean Average Precision - mAP) ve hız açısından karşılaştırılmıştır.

Sonuç olarak hem sperm hücrelerinin morfolojik sınıflandırılması hem de videolarda sperm hücrelerinin tespit aşamalarında literatürdeki çalışmalara kıyasla daha başarılı sonuçlar elde edilmiştir.

Anahtar Kelimeler: Kısırlık, sperm analizi, görüntü işleme, video işleme, derin öğrenme yöntemleri

---

ABSTRACT

Examination of Microbiological Images: Computer Aided Sperm Detection and Morphology Analysis

Mecit YÜZKAT

Department of Computer Engineering
Doctor of Philosophy Thesis

Supervisor: Prof. Dr. Nizamettin AYDIN
Co-supervisor: Assist. Prof. Dr. Hamza Osman İLHAN

According to the World Health Organization, infertility is defined as the absence of pregnancy despite the fact that couples have regular sexual intercourse for twelve months or more without any protection. Today, infertility has become a common problem affecting approximately 20% of the world's population. When diagnosing infertility, reproductive cells of men and women are examined separately. In the diagnosis of male infertility factors, analysis of sperm cells is carried out under certain conditions in a laboratory setting. When analyzing sperm cells, three important characteristics of sperm (morphology, concentration, and motility) are utilized. Sperm analysis can be done visually by experts or by using computer aided sperm analysis systems. The importance of computer-assisted analysis is increasing day by day because visual examination gives different results from person to person and is costly. On the other hand, computer-based expert systems are more consistent and reliable. However, they are not available in many laboratories because they are very expensive systems. In this study, a hybrid expert system was developed using the computerized analysis part and the visual evaluation environment in order to eliminate the disadvantages of both methods. With the developed system, manual orientation and clipping processes have been eliminated. In the scope of the thesis, deep learning-based methods have been used instead of traditional machine learning models in order to avoid long training times. In the developed system, two modular approaches (sperm morphology analysis from sperm images and sperm detection from sperm videos) were used. In morphology analysis, sperm images must be

understandable for obtaining healthy results. However, due to some distorting effects such as camera shake, poor quality camera usage etc., these images may not be clear enough. Within the scope of the thesis, in order to improve the classification performance of a computer-based analysis approach for sperm shape, the effect of different color spaces and the application of interpolation methods as a pre-processing step were investigated. In addition, various data augmentation processes have been applied for obtaining the highest classification accuracy and solving the data imbalance and lack of data problems. Then, six different Convolutional Neural Network (CNN) models based on deep learning were created to fully automate the morphological classification of sperm images. In addition, fusion methods at two decision levels, hard voting, and soft voting, were applied on these models. Apart from the created models, some layers of the pre-trained models with transfer learning methods were modified and used. Four sperm morphology datasets abbreviated as SMIDS, HuSHeM, SCIAN-Morpho and MHSMA were used to evaluate the performance of the proposed approach. In addition to accuracy scores, many performance metrics such as Precision, Sensitivity (Recall), Specificity, F-score and Kappa were used in the morphological classification of sperm cells. In the second module, the detection of sperm cells, the determination of concentration and motility parameters were discussed. In this module, deep learning-based object detection methods were used for sperm detection to obtain the concentration and motility properties of sperm samples. First, sample sperm videos from different infertile patients were recorded via a smartphone mounted on the ocular part of the microscope. The cells within the videos were then labeled separately as sperm and non-sperm. Different scenarios were tested on the labeled data set. Results were compared for accuracy (Mean Average Precision - mAP) and training time.

As a result, more successful results were obtained in both the morphological classification of sperm cells and the detection of sperm cells in videos compared to the studies in the literature.

Keywords: Infertility, sperm analysis, image processing, video processing, deep learning methods

1
GİRİŞ

1.1 Literatür Özeti

Kısırlık bir çiftin bir yıl süre zarfında düzenli olarak hiçbir şekilde korunma olmaksızın cinsel ilişkiye girmelerine rağmen kadının gebe olmama durumuna denir [1]. Son 50 yılda kısırlık bilinen ya da bilinmeyen nedenlerden dolayı dünya çapındaki çiftlerin %15'ini etkileyen önemli bir sorun haline gelmiştir [2, 3]. Roser tarafından yapılan kapsamlı internet araştırmasına göre, küresel ortalama doğurganlık hızı 1950'lerde 4,97 iken 2005 ve 2010 yılları arasında 2,5 olarak kaydedilmiştir. Çalışmada ayrıca doğurganlık hızının 2050'li yıllarda yüzde 2,3'e, 2100'lü yıllarda ise yüzde 1,99'a düşeceği belirtilmektedir [4]. Kısırlığın %40'ı sadece erkekten, %40'ı sadece kadından ve %20'si ise hem erkeklerden hem de kadınlardan kaynaklanmaktadır [5]. Erkeklerde kısırlık vakalarının nedeni damar tıkanıklığı (varikosel), kanal tıkanıklığı veya boşalma sorunları gibi anatomik anormallikler olduğu gibi dış çevresel faktörler de olabilir [6]. Erkek kısırlığına neden olan dış çevresel faktörler zararlı gıdalar, sigara, ağır metaller, radyasyon, yüksek sıcaklık, bazı antibakteriyel ilaçlar, bazı bitkiler ve stres gibi birçok sebep örnek olarak verilebilir [7]. Düzenli beslenme ve stresten kurtulma gibi bazı yaşam tarzı değişiklikleri doğurganlığı olumlu yönde etkileyebilir [8]. Dünya Sağlık Örgütü'ne (World Health Organization - WHO, DSÖ) göre, sperm analizi en az 3-5 gün cinsel perhiz sonrası mastürbasyonla erkeklerden elde edilen sperm ile yapılır [9].

Erkeğin kısırlık değerlendirilmesi yapılırken ilk adım semen analizi olmalıdır. Daha önceki yıllarda erkeğin kısırlık teşhisi tamamen görsel inceleme ile yapılmaktaydı. Günümüzde ise laboratuvarlarda görsel değerlendirme (Visual Assessment - VA) veya Bilgisayar Destekli Sperm Analizi (Computer Aided Sperm Analysis - CASA) gibi iki yöntemle yapılmaktadır. VA, deneyimli laboratuvar asistanları veya teknisyenleri tarafından mikroskobik görüntülerin manuel olarak incelenmesidir. Sperm analizi yapılırken spermin üç özelliği üzerinden analiz yapılır. Spermin morfolojisi (morphology), sayısı (count) ve hareketi (motility) gibi özelliklerine bakılmalıdır [10–14]. Özellikle sperm hareketliliğinin görsel olarak değerlendirmesi çok zor, zaman alıcı ve gözlemciler arası değişkenliğe tabidir [15, 16]. Bu nedenle sperm

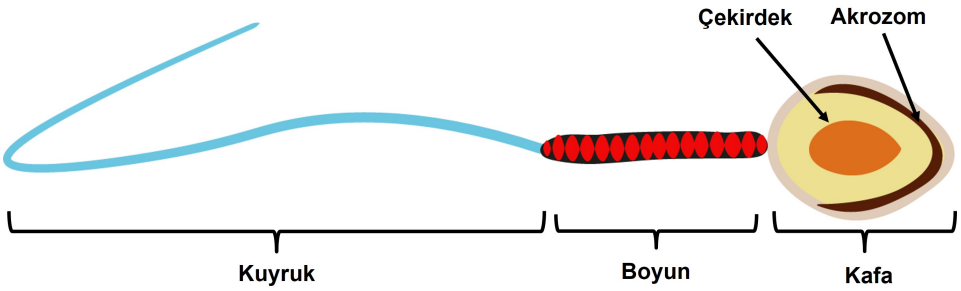
analizinin otomatik olarak yapılabilmesi için cihaz kullanılması gerekir. Makler [17], Hemositometre [18] vb. gibi farklı tıbbi sayım araçları, sayım ve motilite analiz süreçlerine yardımcı olmak için VA tekniğinde kullanılmaktadır. 1980'lerin sonlarından itibaren, CASA sistemleri sperm kalitesini analiz etmek için geliştirilmiştir. CASA sistemleri tıbbi tedavilerin daha doğru, tarafsız ve tutarlı analizini sağlamaktadır [19–21]. Sperm analizinde CASA sistemleri geliştirilmiş olsa da günümüzde geleneksel yöntemler hala yaygın olarak kullanılmaktadır [22]. CASA sistemlerinin en büyük dezavantajları nispeten yüksek maliyetleri, aşırı standardizasyona bağımlı olmaları, sistem kullanıcılarının eğitimi ve doğrulama ihtiyacıdır. Ayrıca farklı CASA sistemlerinin değişken sonuçlar üretmesi gibi dezavantajları da vardır [23–30]. CASA'da kullanılan binoküler kamera kalitesi, kameranın düzgün yerleşimi ve yazılım ile uyumluluğu sistem performansını doğrudan belirlemektedir. Bu nedenle sistemi performans açısından donanıma bağımlı hale getirir. CASA'nın bu dezavantajları, karmaşıklıkları ve belirsizliklerinden dolayı çoğu küçük ve orta ölçekli laboratuvar, bir CASA sistemi kullanmak yerine, analizleri hala gözlem yoluyla ve manuel olarak yürütmektedir [31–34].

Bu bölümde literatürde son yıllarda yapılan sperm morfoloji analizi, sperm tespiti ve hareketinin analizi, son olarak sperm sayısının ölçümü ile ilgili çalışmalar incelenmiştir ve sırasıyla belirtilmiştir. Ayrıca çalışmalarda kullanılan yöntemler ve bu yöntemlerin genel özellikleri incelenerek çalışmaların eksik yönlerine değinilmiştir.

1.1.1 Sperm Morfolojisi ile İlgili Literatür Çalışmaları

Sperm hücrelerinin şekil ve büyüklüğündeki kusurları ifade eden spermin morfolojik anormalliğinin erkek kısırlığı üzerinde önemli bir etkisi vardır [35]. Bundan dolayı, erkeklerde kısırlık teşhisinde ileri analizlerden önce ilk adım olarak, spermin şekilsel analizine bakılmalıdır. Sperm morfolojisi değerlendirilirken DSÖ'nün Kruger kıstasları ve Menkveld çalışması [36] göz önüne alınarak analiz yapılmaktadır. Kruger kıstaslarına göre normal bir sperm morfolojisi, oval, düzgün kenarlı ve sperm kafa bölgesinin ortalama %40-%70'ini kaplayan belirgin bir akrozoma sahip olmalıdır. Akrozom sperm kafasının ön kısmında bulunan, sperm hücresinin yumurtaya girmesini sağlayan sindirim enzimlerini içeren, golgi aygıtından oluşan bir yapıdır. Ayrıca kafa, orta kısım ve kuyruğuna ait herhangi bir şekil bozukluğu bulunmamalıdır. Bu açıklamaları göz önünde bulundurarak Şekil 1.1'de örnek bir sperm hücresinin çizimi gösterilmiştir.

Mortimer ve arkadaşları çalışmalarında CASA teknolojilerini incelemişlerdir. Yazarlar iki modern CASA platformu olan IVOS (Hamilton Thorne Research, Beverly, MA) ve Microptic SCA6 (Sperm Class Analyzer; Microptic S.L., Barcelona, Spain) sistemlerinin



Şekil 1.1 Örnek sperm görüntüsü

gelişmiş yeteneklerini incelemişlerdir. CASA'nın araştırma aracı olarak hayvan ve insan klinik analizlerinde muazzam bir potansiyele sahip olduğuna değinmişlerdir. Fakat özellikle insan sperminin analizi için teknolojilerin geliştirilip kesin olarak onaylanana kadar CASA'nın kullanılmaması gerektiğini savunmuşlardır [37].

Alegre ve arkadaşları çalışmalarında, yaban domuzlarına ait sperm akrozomlarının otomatik olarak sağlam ve hasarlı olarak kategorize edilmesini araştırmışlardır. Yaptıkları deneysel çalışmalara göre, Çok Katmanlı Algılayıcı (Multi Layer Perceptron - MLP) ve k-En Yakın Komşu (k-Nearest Neighbours - kNN) sınıflandırıcılarının moment bazlı tanımlayıcılardan (moment based descriptors) daha iyi performans gösterdiğini ve bu performansın %94'lük bir doğruluğa ulaştığını ifade etmişlerdir [38].

Bijar ve arkadaşları çalışmalarında Acrosome, Nucleus, Mid-piece gibi sperm morfoloji analizi yapmışlardır. Spermin bu parçalarının tespiti için entropi tabanlı (Expectation Maximization - EM) algoritması ve (Markov Random Field - MRF) modelini kullanan bir Bayesian sınıflandırıcısı önerilmiştir. Daha sonra spermin orta kısmının sonundaki bir piksel başlangıç noktası olarak seçilip spermin kuyruk kısmında daha önceden işaretlenmiş pikselleri bulmak için yinelemeli bir şemada Yapısal Benzerlik indeksi (Structural SIMilarity - SSIM) kullanılmıştır. Önerilen yaklaşımın performansını manuel bölütleme dahil önceki yaklaşımlarla karşılaştırmak için doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) hesaplanmıştır. Sperm kuyruğunun tespiti için önerilen yöntemin dezavantajı sperm kuyruğuna yerleştirilen nesnelerin önerilen yöntemin performansını etkilemesidir. Dolayısıyla gelecekteki çalışmalarında performansı artırmak için çeşitli geometrik özellikler kullanmayı amaçlamışlardır [39].

Alegre ve arkadaşları daha önceki çalışmalarına benzer yaban domuzu sperm kafalarının akrozom durumunu değerlendirmek için logaritmik mesafe fonksiyonu

(Logarithmic Distance Function) kullanmışlardır. Her bir sınıflandırıcı, ele alınan tanımlayıcılar (gradyan, gri seviye ortalaması ve yerel standart sapma) için elde edilen 280 özellik vektörünü karşılaştırmıştır. Sınıflandırma için; k-Nearest-Neighbor, Sınıf Koşullu Ortalama Yaklaşımı (Class Conditional Means Approach) ve İlişkili Vektör Ölçümü Öğrenme (Relevance Learning Vector Quantization) ile k-kat çapraz doğrulama (k-fold cross validation) gibi yöntemler kullanılmıştır. Çalışmalarında (Relevance Learning Vector Quantization - RLVQ) yöntemi ile %98 doğruluk elde ederek daha önce yayınlanmış çalışmalarından [38] daha yüksek doğruluk değerine ulaşmışlardır [40].

Chang ve arkadaşları insan spermlerine ait üç farklı renk uzayını kullanarak akrozom, çekirdeğin tespiti ve bölütlenmesi için iki aşamalı bir sistem (framework) hazırlamışlardır. İlk aşamada *k* means kullanılarak sperm kafa bölgeleri tespit etmişlerdir. İkinci aşamada kümeleme ve histogram kullanarak sperm kafasının yanı sıra çekirdek ve akrozomu doğru bir şekilde bölütlemişlerdir. Sadece sperm kafa tespitinde %98 oranında başarı elde etmişlerdir. Elde edilen sonuçların geneline bakıldığında ise elle bölütlenmiş altın standardına karşı %80 oranında başarımla kafa, akrozom ve çekirdek bölütlenmesinde örtüşmeyi sağlamışlardır [41].

Chang ve arkadaşları başka bir çalışmada, SCIAN-Morpho veri setinin sınıflandırılması için bir standart oluşturmuşlardır. Standardı oluşturmak için Hu momentleri, Zernike momentleri, Fourier gibi üç farklı şekil tabanlı tanımlayıcı ve kNN, Naive Bayes, Karar Ağaçları, Destek Vektör Makinesi (Support Vector Machine - SVM) gibi dört denetimli öğrenme yöntemi kullanmışlardır. Bu yaklaşımların hiçbiri sperm kafası sınıflandırma probleminin üstesinden gelememiştir. Fourier tanımlayıcısı ve SVM kullanılarak optimal sonuç %49 olarak elde edilmiştir [42]. Yazarlar bir diğer çalışmada, aynı veri seti için otomatik bir sınıflandırma sistemi sunduklarını öne sürmüşlerdir. Elde edilen sonuçlar uzmanlar ile %70 oranında uyumluluk göstermiştir. [43]. Çalışmalarında spermin kafa bölgelerini elde etmek için kırpma işlemi yaparak görüntüleri manipüle etmişlerdir. Bu nedenle bu çalışmada sperm görüntülerinin sınıflandırılması otomatik olmamıştır.

Ghasemian ve arkadaşları insan sperm hücrelerine ait görüntülerin morfolojik analizi için otomatik bir algoritma önermişlerdir. Öncelikle görüntülerdeki gürültüler giderilip görüntülerin kontrastı büyük ölçüde arttırılmıştır. Daha sonra spermin kafa, kuyruk gibi farklı kısımlarının morfolojik analizi yapılmıştır. Son olarak algoritma her spermi normal ya da normal olmayan olarak sınıflandırmıştır. Çalışmalarında 23 hastadan toplam 1457 sperm görüntüsüne sahip (Human Sperm Morphology Analysis - HSMA) veri seti kullanılmıştır. Kullandıkları algoritma ile normal olmayan spermlerin tespitinde %90'ın üzerinde doğruluğa ulaşılmıştır [44].

Garcia ve arkadaşları faz kontrast mikroskobu ile elde edilen gri ölçekli görüntülerde bulunan yaban domuzu spermatozoasını incelemişlerdir. Çalışmada sperm örnekleri sağlam akrozom veya hasarlı akrozom olarak sınıflandırılmıştır. Sınıflandırma yapılırken En Küçük Ortalama Kareler (Least Mean Squares - LMS) algoritması ve doğrusal çekirdek (linear kernel) tarafından desteklenen bir SVM kullanılmıştır. Performans değerlendirilmesi için; kesinlik, duyarlılık, F-skor ve doğruluk gibi ölçüm metrikleri kullanılmıştır. Analizler için 1851 adet görüntü kullanılarak %99 oranında doğruluk değerine ulaşılmıştır [45].

Shaker ve arkadaşları çalışmalarında otomatik sperm kafası tespiti ve bölütlenmesi için boyut kıstasına ek olarak HSV (Hue Saturation Value) renk uzayının Hue kanalının histogramını kullanmışlardır. Sperm kafalarını bölütlemek için Canny Kenar Dedektörü (Canny Edge Detector - CED) ile (Gradient Vector Flow - GVF) aktif sınırları birleştirmişlerdir. Önerilen yöntemde sperm kafaları, akrozomlar ve çekirdek için sırasıyla %92, %84, %87 doğruluk sağlamıştır. Ayrıca kuyruk algılama için %96 doğruluk oranına ulaşmışlardır [46].

Shaker ve arkadaşları başka bir çalışmada sperm görüntülerinin sözlük öğrenme yöntemi ve SVM kullanılarak morfolojik sınıflandırmayı hedeflemişlerdir. Çalışmada (Human Sperm Head Morphology - HuSHeM) ve SCIAN-Morpho olmak üzere iki farklı veri seti kullanmışlardır. HuSHeM ve SCIAN-Morpho veri setleri için sınıflandırma doğrulukları sırasıyla %92 ve %62 olarak elde edilmiştir [47]. Ancak, bu çalışmada tam bir otomatik sınıflandırma yapılmamıştır. Yazarlar HuSHeM veri setinde orijinal görüntüleri kullanmak yerine, kırpma ve yön değiştirme gibi ön işlemler uygulayarak görüntüleri manipüle etmişlerdir.

Witkowski sperm hücrelerinin morfolojik ve hareket özelliklerini analiz eden kapsamlı bir çalışma gerçekleştirmiştir. Çalışmada toplam 29 hastanın verileri kullanılmış, görüntülere medyan filtre ve normalizasyon uygulanmıştır. Hareketlilik ve morfoloji analizinde sırasıyla ortalama %72, %82 doğruluk elde edilmiştir [48]. Görüntü sayısının yetersizliği ve taranan alanın küçük olması nedeniyle iyi sonuçlar alınmamıştır.

Riordon ve arkadaşları VGG-16 ağını değiştirerek HuSHeM ve SCIAN-Morpho veri setlerinin sınıflandırılması için sırasıyla %94 ve %62 doğruluk elde etmişlerdir. [49]. Çalışmalarında, Shaker'in çalışmasındaki [47] gibi, sınıflandırmadan önce kırpma ve döndürme gibi ön işlemler kullanarak görüntüleri manipüle etmişlerdir. Bu nedenle, önerilen sınıflandırma yaklaşımı tam otomatik değildir.

Yüzkat ve arkadaşları SCIAN-Morpho veri setinin sınıflandırma performansını artırmak için çeşitli interpolasyon ve veri artırma (data augmentation) yöntemleri

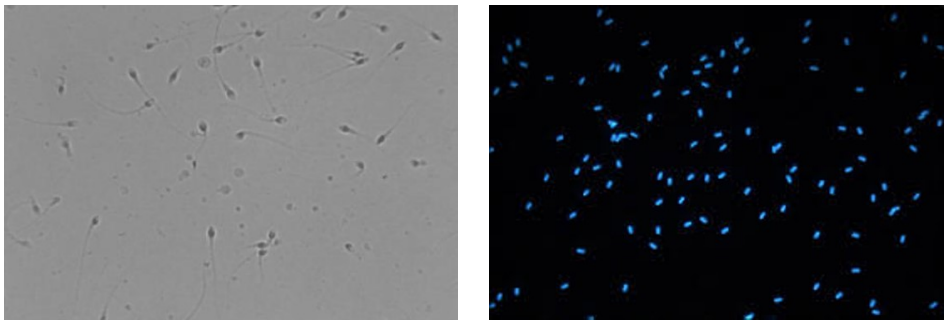
uygulamışlardır. Daha sonra değiştirilmiş bir VGG-19 ağı kullanılarak %62 sınıflandırma doğruluğu elde edilmişlerdir [50]. Daha yakın tarihli bir çalışmalarında, sperm görüntülerini sınıflandırmak için tam otomatik bir analiz yaklaşımı sunmuşlardır. Çalışmalarında altı CNN modeli oluşturmuşlar ve modellerin sonuçlarını hard ve soft füzyon yöntemleriyle birleştirmişlerdir. (Sperm Morphology Image Data Set - SMIDS), HuSHeM ve SCIAN-Morpho veri setleri için sırasıyla %91, %85 ve %72 doğruluk elde etmişlerdir [51]. Başka bir çalışmalarında ise, daha önce yapılmış bir çalışmadaki sınıflandırma başarısını karşılaştırmak için bir ön işleme adımı olarak farklı renk uzaylarını kullanmışladır. Çalışmada SMIDS, HuSHeM ve SCIAN-Morpho veri setlerini kullanılmışlardır. Veri setlerindeki dengesiz dağılım ve yetersiz veri nedeniyle veri artırma uygulamışlardır. Daha sonra veri setleri popüler iki renk uzayına, (Lightness, Green to Red and Blue to Yellow - LAB) ve HSV'ye dönüştürülmüştür. Sınıflandırma modeli olarak MobileNetV2 kullanarak SMIDS, HuSHeM, SCIAN-Morpho veri setleri için sırasıyla %89, %85 ve %68 doğruluk değerlerine ulaşmışlardır [52].

İlhan ve arkadaşları SMIDS veri setinin ilk versiyonu için %82 sınıflandırma başarısı elde etmişlerdir [53]. Sınıflandırma yaparken makine öğrenmesi ve dalgacık tabanlı yöntemleri kullanmışlardır. Diğer bir çalışmalarında [54] aynı veri seti üzerinde uyarlamalı gürültü giderme yaklaşımı ve tanımlayıcı tabanlı öznitelikler kullanarak doğruluk oranını %82,33'ten %85,42'ye yükseltmişlerdir.

İlhan ve arkadaşları bir diğer çalışmalarında sperm görüntülerinin tam otomatik sınıflandırma yaklaşımı için SMIDS ve HuSHeM veri setlerine özel bir ön işleme yaklaşımı uygulamışlardır. Bunun için daha önce yayınlamış oldukları yönlü maskeleme adı verilen otomatik bir ön işleme tekniği kullanmışlardır [55]. Yönlü maskeleme tekniğini görüntülere uyguladıktan sonra, önceden işlenmiş görüntülerin (Speed-Up Robust Features - SURF) ve (Maximally Stable Extreme Regions - MSER) gibi tanımlayıcı özelliklerini çıkarmışlardır. Daha sonra çıkarılan öznitelikleri SVM ile sınıflandırarak SMIDS için önceki başarı oranını %84'ten %86'ya çıkarmışlardır. HuSHeM veri seti için ise %85 doğruluk değeri elde etmişlerdir [56].

İlhan ve arkadaşları başka çalışmalarında sperm sınıflandırılması için MobileNetV2 ağını kullanmışlardır. Fakat MobileNetV2'yi önceden eğitilmiş bir ağ olarak kullanmak yerine, SMIDS kullanarak ağı yeniden eğitmişlerdir. Farklı veri artırma yöntemleri uygulayarak %87 doğruluğa ulaşmışlardır [57]. Çalışmanın dezavantajı, bu kadar büyük bir ağın baştan sona eğitilmesinin çok zaman almasıdır.

Iqbal ve arkadaşları CNN tabanlı derin öğrenme yöntemi ile kaliteli ve kalitesiz sperm hücrelerini sınıflandırmışlardır. Önerdikleri yöntemle her katmanda çoklu filtre



a *b*

Şekil 1.2 Mikroskop görüntüleme yöntemleri: a) faz kontrastı b) floresan boyama

boyutları kullanmışlardır. Önerdikleri modelin etkinliğini artırmak için daha az sayıda filtre ve parametre kullanılmışlardır [58]. Fakat araştırmacılar tam otomatik bir sistem geliştirmemişlerdir. Model eğitimi öncesinde veri setine müdahale etmişlerdir.

1.1.2 Sperm Hareketi ile İlgili Literatür Çalışmaları

DSÖ'nün 2010 yılında yayınlanan kılavuzuna göre, konsantrasyondaki sperm sayısı için 15 milyon/ml normal sınır olarak kabul edilmektedir. Kılavuza göre sperm hareketi dört kategori ile tanımlanır: Hızlı hareket (Grade A), Normal hareket (Grade B), Yavaş hareket (Grade C) ve titreşimli ya da hareketsiz (Grade D). Aynı kılavuza göre, tüm sperm konsantrasyonunun %40'ı hareketli olmalı ve ölçülen tüm motilitenin %32'si ilerleyici tipte olmalıdır (Grade A+B) [1].

Sperm hareketi analizinde tıkanma, aydınlatma değişiklikleri, kamera efektleri ve gürültü gibi zorluklar, aşılması gereken önemli konular olarak sunulmaktadır. Sperm analizinde faz kontrastı (phase-contrast) ve floresan boyama (fluorescence dying) olmak üzere iki temel mikroskop görüntüleme tekniğini kullanır [20, 59]. Floresan boyama görüntüleme tekniği, nesneleri arka plandan çıkarmak ve gürültüsüz bir sistem oluşturmak için yararlıdır. Floresan boyama tekniğinde ışığın meni üzerine yansıması nedeniyle hareketli sperm sayımı kolaylıkla yapılabilmektedir [60]. 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 bölütlenmesi tam olarak sağlanamayabilir. Bu açıdan literatürde sperm morfolojisi ile ilgili yapılan çalışmaların çoğunluğu faz kontrast mikroskobik görüntüler üzerinden gerçekleştirilir. Şekil 1.2'de faz kontrastı ve floresan boyama yöntemlerinden üretilmiş örnek sperm görüntüleri gösterilmektedir. Literatürde sperm hareketi ile ilgili çeşitli çalışmalar yapılmıştır. Bu çalışmalara aşağıda yer verilmiştir.

Liu ve arkadaşları çalışmalarında sperm hareketini izleyen birçok sperm izleme algoritmasını sunmuşlardır. Algoritmaları üç adımdan oluşmaktadır. İlk adımda sperm kafasının eğrisel hızını (Curvilinear Velocity - VCL), düz çizgi hızını (Straight Line Velocity - VSL) ve spermin izlediği yolun doğrusallığını hesaplamışlardır. İkinci aşamada spermin kuyruk bölgesi çıkarılmıştır. Ardından maksimum yoğunluk bölgesi (Maximum Intensity Region - MIR) algoritması kullanılarak sperm kuyruğundaki bir noktayı bulmuşlardır. Son olarak sperm kuyruğunda bulunan noktanın doğruluğunu artırmak için Kalman filtresi kullanmışlardır. Sonuçlarının sperm kafa hızı ile kuyruk atma genliği arasında anlamlı bir ilişki olduğu doğrulanmıştır [61].

Witkowski çalışmasında sperm hücrelerinin yoğunluğunu ve seçilen hareket parametrelerini incelemek için sağlıklı bir kişinin semen görüntü dizisini kaydetmiştir. Veriler 400x400 piksel boyutunda 50 adet görüntüden oluşmaktadır. Öncelikle görüntüler üzerinde medyan filtre gibi ön işlemler uygulamıştır. Örneklerinde 12 yörüngenin 10 tanesini uygun çalışma 2 tanesini ise kötü çalışma olarak sınıflandırmıştır [62]. Çalışmasının avantajı analizin daha kısa süre olması, dezavantajı ise örnek sayısının yetersiz olmasıdır.

Ravanfar ve arkadaşları çalışmalarında sperm tıkanma sahnelerinde yanlış yapılan hataların üstesinden gelebilecek sperm izleyen yöntem önermişlerdir. Tek bir sperm takibi için tasarlanmış parçacık filtresi kullanılarak bölütleme yöntemi kullanılmıştır. Çalışmada her karede spermin durumu üç aşamaya ayrılmıştır. Bu aşamalar tıkanmadan önceki aşama, tıkanıklık anı ve tıkanmanın ardından diye ayrılmıştır. Yazarlar ön işleme için bir uzaysal - zamansal Gaussian filtresi kullanmışlardır. Böylece düşük örnekleme hızının etkisi azaltılarak sperm hareketleri yumuşatılmıştır. Yazarlar *gray gatershed* algoritmasının aşırı bölütlemeye neden olacağından bunun yerine satranç tahtası mesafe dönüşümüne dayanan *binary watershed* algoritması kullanmışlardır. Algoritmayı değerlendirmek için iki örnek grup üzerinde çalışılmıştır. Birinci grupta algoritmaların sperm izleme ile genel uyumluluğunu göstermek için, algoritmalar yörüngeleri boyunca tıkanmayan spermlere uygulanmıştır. İkinci grup için geliştirilen algoritma ise sperm hareketi sırasında en az bir tıkanma bulunan spermlere uygulanmıştır. Algoritma birinci grupta %95, ikinci grupta ise %87 doğruluk göstermiştir [63].

Zhou ve arkadaşları tek bir spermin takibi için hibrit üretken ayrımcı bir izleyici (Hybrid Generative Discriminative Tracker - HGDT) önermişlerdir. Sperm hücresinin hareket belirsizliği ile başa çıkmak için enerji kaynaklı stokastik yaklaşım olan Monte Carlo (Energy-Biased Stochastic Approximation Monte Carlo, EB-SAMC) örnekleme algoritması önermişlerdir. PASCAL örtüşme hızını kullanarak önerilen izleyici ile diğer beş izleyici arasında nicel karşılaştırmalar yapılmıştır. Birkaç videodaki deney

sonuçlarının etkinliğini izleme doğruluğu ve hesaplama verimliliği açısından son teknoloji yöntemlere göre üstün olduğunu ifade etmişlerdir. Gelecekteki çalışmalar için çok-görevli öğrenmenin (multi-task learning) ve seyrek öğrenmenin (sparse learning) sperm izlemesinde kullanılmasını hedeflemişlerdir [64].

Gunawan ve arkadaşları çalışmalarında sperm hareketinin ani hareketlerinden dolayı izleme problemini ortadan kaldırmak için sperm videolarını yaklaşık 20 (Frame Per Second - fps) gibi düşük hızlara sahip kameralarla elde etmişlerdir. Ani hareketlerin görsel olarak izlenmesi için stokastik araştırmaya dayalı bir çerçeve önermişlerdir. Arama yapmak için (Coherency Sensitive Hashing - CSH) yöntemi kullanmışlardır. Daha sonra görüntü üzerindeki gürültüleri kaldırmak için (Smoothing Stochastic Approximate Monte Carlo - SSAMC) yaklaşımını uygulamışlardır. Geçici tıkanmaları önlemek için çerçevelerin boyutunu ayarlarken çok dikkat edilmesinden bahsetmişlerdir. Çerçeveler çok küçük seçildiğinde algoritma ani hareketleri fark edemeyeceğinden, çok büyük seçildiğinde ise aynı çerçevede birkaç sperm olacağından algoritmanın iyi sonuçlar vermeyeceği yargısına varmışlardır [65]. Sadece tek bir nesneyi izlemekle sınırlı ve birden fazla nesneyi aynı anda izleyemeyeceğinden geliştirdikleri algoritma çok etkin değildir.

Yang ve arkadaşları görsel inceleme ile elde edilemeyecek bir istatistiksel yöntem kullanarak analiz sürecini otomatikleştiren bir sistem sunmuşlardır. Bu sistemde hem sperm kafasını hem de sperm kuyruğunu izlemişlerdir. Sistem sperm kafası algılama, kafa izleme ve kuyruk izleme gibi üç ana bölümden oluşmaktadır. Kuyruk hareketi desenlerini elde etmek için kuyruk hareketi izleme algoritmasını geliştirmişlerdir. Sperm kuyruklarının farklı koşullar altında farklı hızlarda hareket ettikleri gözlemlenmiştir. Yüksek frekanslı kuyruklara sahip spermlerin, hızlı ve çapraz hareket edeceklerinden, verimlilik açısından daha düşük frekanslı kuyruklara sahip spermlere göre daha az verimli oldukları kanaatine varmışlardır [15].

Urbano ve arkadaşları çalışmalarında otomatik çoklu sperm izleme algoritması sunmuşlardır. Çalışmalarında Gaussian filtresi ve (Laplacian-of-Gaussian, LoG) filtresinin ardından Otsu metodunu kullanmışlardır. Ortak Olasılıksal Veri İlişkilendirme Filtresi (Joint Probabilistic Data Association Filter - JPDAF) olarak bilinen yöntem ile aynı anda yüzlerce insan spermini otomatik olarak takip etmişlerdir. Mevcut CASA cihazlarından farklı olarak algoritmalarının birbirine yakın belirgin sperm çarpışmaları sırasında hareketi izleyebilme özelliğine sahip olduğunu öne sürmüşlerdir [66].

Hidayatullah ve arkadaşları çalışmalarında alternatif bir CASA sistemini geliştirmişlerdir. Çalışmalarında VCL, VSL ve (Linearity - LIN) gibi üç CASA

parametresini kullanarak multimedya mikroskobundan çekilen bir videodan sperm hareketliliğini otomatik olarak ölçmeyi amaçlamışlardır. Geliştirdikleri sistem manuel değerlendirmeye göre %7 oranında bir başarım artışı elde etmişlerdir [67].

Hidayetullah ve arkadaşları başka çalışmalarında yoğun ve kısmi tıkanıklıklara sahip boğa sperm hücrelerine ait videolarda sperm tespiti için YOLO (You Only Look Once) tabanlı DeepSperm adlı bir yöntem önermişlerdir. Yöntemlerinde, küçük nesneleri algılamak için ayrıca bir algılama katmanı kullanmışlardır. Önerdikleri yöntemin başarısını artırmak için veri artırma uygulamışlardır ve bazı hiper parametreleri değiştirerek kullanmışlardır. Yöntemlerini YOLOv3 ve Mask R-CNN ile karşılaştırmışlardır. Test veri setlerinde %87 mAP değerine ulaştıklarını belirtmişlerdir. Sonuçlara göre önerdikleri yöntemin doğruluk ve hız açısından YOLOv3 ve Mask R-CNN'ye göre daha başarılı olduğunu belirtmişlerdir [68].

Aggarwal ve arkadaşları çalışmalarında kanser hastalarında sperm tespiti için derin öğrenme tabanlı YOLOv3 ve Faster R-CNN kullanmışlardır. Veri yetersizliğinden dolayı veri artırma kullanarak veri sayısını artırmışlardır. Etiketleme yaparken sperm hücrelerini normal ve normal olmayan olarak etiketlemişlerdir. YOLOv3 ve Faster R-CNN ile sırasıyla %37 ve %56 mAP sonuçları elde etmişlerdir. YOLO modelinin eğitim süresi açısından Faster R-CNN modeline göre çok hızlı olduğu sonucuna varmışlardır [69]. Fakat sperm tespiti için elde edilen mAP değerleri kabul edilebilir sonuçlar değildir.

Valiuškaite ve arkadaşları insan sperm videolarında sperm kafasının hareketliliğini değerlendirmek için (Region-Based Convolutional Neural Networks, R-CNN) mimarisini kullanmışlardır. Sperm kafasının bölütlenmesi için merkezi bir koordinat izleme algoritması (Central Coordinate Tracking Algorithm - CCTA) önerilmiştir. Bu algoritma ile sperm kafalarının hareket hızını hesaplamışlardır. Çalışmalarında VISEM (A Multimodal Video Dataset of Human Spermatozoa) sperm video örneğini kullanmışlardır. Bu veri setinde spermin kafa bölgesini %92 doğrulukla belirlemişlerdir [70]. Çalışmanın dezavantajlarından biri, sonuçlarının YOLO gibi diğer popüler algoritmalarla karşılaştırılmamış olmasıdır.

Rahimzadeh ve arkadaşları çalışmalarında [71] sperm video veri setinde [72] sperm tespiti için derin öğrenme tabanlı RetinaNet'i kullanmışlardır. Çalışmalarında, performansı artırmak için Retina ağına girdi olarak verilen birkaç ardışık çerçeveyi birleştirmişlerdir. Sperm izleme aşamasında CSR-DCF algoritması değiştirilerek kullanılmıştır. Tespit aşamasında ortalama %99 kesinlik elde edilirken, takip aşamasında ortalama %96 F-skor elde edilmiştir.

1.1.3 Sperm Sayımı ile İlgili Literatür Çalışmaları

Sperm analizi yapılırken spermin morfolojik ve hareket değerlendirilmesinin yanında konsantrasyon değerlendirilmesi de yapılır. Kısırlık düşük sperm sayısı ile ilişkili olduğundan, semen değerlendirmesi sırasında konsantrasyon önemli bir parametredir. Erkeklerdeki üreme biyolojisindeki farklılıklar, sperm konsantrasyonunun farklılığına neden olabilir. Günümüzde sperm konsantrasyonunun belirlenmesi mevcut CASA sistemleri tarafından otomatikleştirilmiştir [73]. Birçok CASA'nın sperm sayım ölçümü, numune hacmine, sulandırıcı tipine, analiz süresine ve çözülme sıcaklığına bağlıdır [74].

Iguer-ouada ve Verstegen [75] çalışmalarında CASA sistemlerinin sperm sayımı sırasında sperm sayısını fazla tahmin ettiği sonucuna varmışlardır. Bunun sperm olmayan diğer hücrelerden veya sperm çarpışmalarından kaynaklandığını bildirmişlerdir. Sperm çarpışmalarından dolayı fazla sayıda sperm tahminini engellemek için meni seyreltme fikri ortaya konmuştur [76]. Bailey ve arkadaşları Makler odası 10μm ve Leja 20μm bölümünü karşılaştırmışlardır. Çalışmalarında Makler odasının sperm konsantrasyonunu belirlemek için oldukça yüksek değerler vermesine rağmen, Leja'nın makul sonuçlar verdiğini fark etmişlerdir [77].

Literatürde manuel sperm sayımı için Makler benzeri birçok sayım odası mevcuttur. Bunlardan biri olan Hemositometre hematoloji uygulamalarında kullanılmasına rağmen, sperm sayımı için yaygın olarak kullanılan eski bir yöntemdir. Fakat çok fazla emek ve zaman gerektirir [78, 79]. Ayrıca yeterlilik testlerinde büyük farklılıklara neden olabilir. Günümüzde ise birçok sperm analiz laboratuvarı Makler odaları kullanmaktadır. Hemositometre ve Makler odaları tek kullanımlık olmadığı için kullanılmadan önce gerekli bakımı yapılmalıdır. Fakat her bakım, odacığın yüzeyini aşındırabileceği için yanlış sperm sayımına neden olabilir. Bu yüzden belli kullanım sayısından sonra değiştirilmelidir. Fakat günümüzde bazı laboratuvarlarda tek kullanımlık sperm sayma odaları ile analiz öncesi bakımı ortadan kaldırarak hem işçilikten tasarruf sağlanır hem de daha doğru analiz elde edilir.

Kanakasabapathy ve arkadaşları telefon yardımı ile evde sperm sayımı analizi için bir cihaz geliştirmişlerdir [80]. Çalışmalarında mikro akışkanları ölçen bir cihaz ve modüler bir kablosuz tartı kullanarak semen numunesi üzerinde görüntü analizi yapmışlardır. Sonuçlarını CASA sistemleri ile karşılaştırarak %97 doğruluk elde etmişlerdir.

Bompart ve arkadaşları çeşitli memeli türlerinde sayım odası tipi ve derinliğinin CASA-Mot sonuçları üzerindeki etkisini gözden geçirmişlerdir [81].

İlhan ve arkadaşları sperm görüntülerinin toplanması aşamasında akıllı telefon tabanlı yeni yaklaşım kullanarak sperm sayımı ve analizi yapmışlardır [82, 83]. İlhan ve arkadaşları güncel çalışmalarında ise sperm sayma, takip ve sonuç raporlama sistemi önermişlerdir. Önerilen sistem ucuz, doğru ve kullanılan platformdan bağımsızdır. Çalışmalarını başka bir CASA sistemi olan SQA-Vision ile karşılaştırmışlardır [84].

1.2 Tezin Amacı

Günümüz dünyasında farklı sebeplerden kaynaklı kısırlıktan etkilenen bireylerin sayısı her geçen gün artmaktadır. Son 50 yılda kısırlık önemli bir sorun haline gelmiştir ve dünyadaki her on çiftten birini etkilemektedir [2, 3]. Çiftlerde kısırlık teşhisi yapılırken kadın ve erkeğin ayrı ayrı üreme hücreleri incelenir. Klinik çalışmalar, erkek kaynaklı kısırlık oranının her geçen gün arttığını ortaya koymaktadır. Sperm analizi, sperm kalitesinin değerlendirilmesinde anahtar adımdır. Sperm kalitesi araştırılırken sperm sayısı, hareketi ve morfolojik yapısı değerlendirilir. CASA ve VA, analizlerde kullanılan iki değerlendirme tekniğidir. VA sayım odalarında sperm gözlemlenerek gerçekleştirilir. Bu nedenle, teşhis büyük ölçüde beceri ve deneyimlere bağlıdır. CASA, bilgisayar tabanlı yöntemleri kullanarak insan faktörünü ortadan kaldırır. Ancak, VA'dan daha maliyetlidir ve detaylı parametre ayarları gerektirir. Bu nedenle pek çok laboratuvarda tercih edilmemektedir. Bu tez çalışmasında hem VA analizinde gözlemci değişkenliği hem de CASA sistemlerinin maliyetleri, karmaşıklıkları ve belirsizlikleri gibi dezavantajları ortadan kaldıracak veya azaltacak, birçok küçük ve orta ölçekli laboratuvarda objektif analizler yapmayı sağlayacak otonom bir sistemin gerçeklenmesi hedeflenmiştir. Genel anlamda, hedeflenen sistemdeki amaç, VA analizinde kullanılan veri toplama adımı ile CASA sistemlerinin bilgisayarlı analiz adımlarının birlikte kullanılmasıdır. Önerilen sistemde daha önce kullanılan geleneksel makine öğrenmesi yöntemleri yerine derin öğrenme yöntemleri kullanılarak görüntülerin otonom morfolojik sınıflandırılması ve sperm videolarında sperm tespitinin yapılması hedeflenmiştir. Böylece VA yöntemini kullanan ve CASA tabanlı sperm analizinin mevcut dezavantajlarını dikkate alan daha geniş kullanım alanlarına sahip bir sistemin geliştirilmesi amaçlanmıştır.

1.3 Hipotez

Bu tez çalışmasında sperm analizi için geleneksel değerlendirme ile bilgisayar destekli değerlendirme yöntemlerinin hibrit bir versiyonu tanıtılmıştır. Tez çalışmasının hipotezleri aşağıda listelenmiştir.

- Sperm analizinde telefon ve bilgisayar sistemi birlikte kullanılabilir. Akıllı

telefon aracılığı ile verilerin toplanması uygun maliyet sağlayabilir. Bu sayede, bilgisayar sistemi ile sperm morfoloji analizi ve sperm videolarında sperm tespiti doğru, kolay ve ucuz olarak yapılabilir.

- Önerilen yaklaşım, VA analizinde gözlemci değişkenliğini ortadan kaldırabilir ve analizler bilgisayar sistemi ile daha objektif bir şekilde yapılabilir.
- Maliyetli olması ve detaylı parametre ayarları gerektirdiğinden dolayı, CASA sistemleri pek çok küçük ve orta ölçekli laboratuvarda tercih edilmemektedir. Bu çalışmada önerilen sistemin ürettiği sonuçların VA ve CASA sistemlerinin sonuçlarıyla yüksek oranda örtüşmesi nedeniyle önerilen sistem birçok laboratuvar tarafından gelecekte kullanılabilir.
- Geleneksel çok aşamalı makine öğrenmesi yöntemleri yerine derin öğrenme yöntemlerinin kullanılması ile daha yüksek başarımlı sperm analiz sistemi elde edilebilir.
- Sperm morfoloji analizinde, başarılı sonuçlar elde etmek için kaydedilen sperm görüntülerinin detaylı ve kaliteli olması gerekir. Ancak, kameranın titremesi, düşük kalite kamera kullanımı vb. nedenlerden dolayı bu görüntüler istenen kalitede olmayabilir. Bu çalışmada düşük kalitede elde edilmiş sperm görüntülerinin sınıflandırma başarımını artırmak amacıyla bazı ön işlemler uygulanabilir.
- Ayrıca sperm görüntüleri üzerinde manuel kırpma, döndürme gibi işlemleri kullanan diğer çalışmaların aksine, herhangi bir manuel müdahale olmaksızın otonom bir sistem (framework) sunulabilir.
- Kullanılan veri setlerinin veri dengesizliği ve yetersizliği problemleri nedeniyle, veri artırma yapılarak yüksek performanslara sahip sonuçların elde edilebilmesi mümkün olabilir.
- İyi bilinen dört sperm morfolojisi veri setine ait sınıflandırma sonuçları literatürdeki diğer yöntemlerle karşılaştırılarak değerlendirilebilir.
- Sperm morfolojisi veri setlerinin sınıflandırılması için yeni CNN modellerinin oluşturulması ve bu modellerin sonuçlarının füzyon fikrini kullanan bir sınıflandırma yaklaşımı önerilebilir.

2
GÖRÜNTÜ İŞLEME

İnsanların görme potansiyeli sınırlı olduğundan belirli elektromanyetik aralıktaki bantları görür. Fakat gelişmiş makineler daha geniş elektromanyetik aralıktaki bantları görebilir. Bir görüntünün x, y koordinatındaki fonksiyonun değeri görüntünün o nokta için yoğunluğu ya da gri seviyesini belirtir [85]. Buradaki x, y ve yoğunluk değerleri sonlu ise sayısal görüntü olarak adlandırılır. Günümüzde yüz tanıma [86–88], mimari yapıların tasarımı [89, 90], parmak izi okuma [91], trafikte araç plakalarının tespiti [92–94], gıda ürünlerinin sınıflandırılması [95] gibi birçok alanda görüntü işleme sistemleri kullanılmaktadır [96]. Bunların yanı sıra, biyomedikal alanda görüntü işleme çok büyük bir öneme sahiptir [97–100].

2.1 Görüntü İşlemenin Temel Adımları

Görüntü işlemenin temel adımları görüntünün elde edilmesi, ön işlemler, görüntüde aranan alanın tespiti, görüntünün sınıflandırılması olarak incelenmektedir. Görüntü işlemenin başarısı bu adımların iyi uygulanmasına bağlıdır. Görüntü işlemenin ilk adımı görüntünün sayısallaştırılması ile başlar ve görüntünün onarılması, görüntünün bölümlendirilmesi ve öznitelik çıkarımı ile sonuçlanır [101, 102]. Bu adımlar Şekil 2.1'de gösterilmiştir.



Şekil 2.1 Görüntü işlemenin temel adımları

2.1.1 Görüntünün Elde Edilmesi

Analiz edilecek görüntüler kameralar aracılığıyla elde edilir. Görüntüler elde edilirken kameraların kalitesi ve ortam şartlarına dikkat edilmesi görüntü analizi için önemlidir. Bu şartların iyi olmadığı durumlarda hedeflenen çıktılara ulaşmak için görüntü işleme algoritmalarının önemi bir hayli artmaktadır.

2.1.2 Görüntünün Ön İşlenmesi

Bu adımda görüntünün bilgisayar tarafından daha iyi yorumlanması için kalitesinin iyileştirilmesi ve istenilen formatlara dönüşümü yapılır. Genelde görüntünün piksel parlaklığı, yoğunluğu, kenar belirginleştirme, gürültü giderme, ölçeklendirme ve filtreleme gibi işlemlerle ilgilenilir [103, 104]. Bu bölümde görüntü ön işleme adımlarının önemli işlemleri hakkında açıklamalar verilmiştir.

2.1.2.1 Renk Uzayları Arasında Dönüşümler

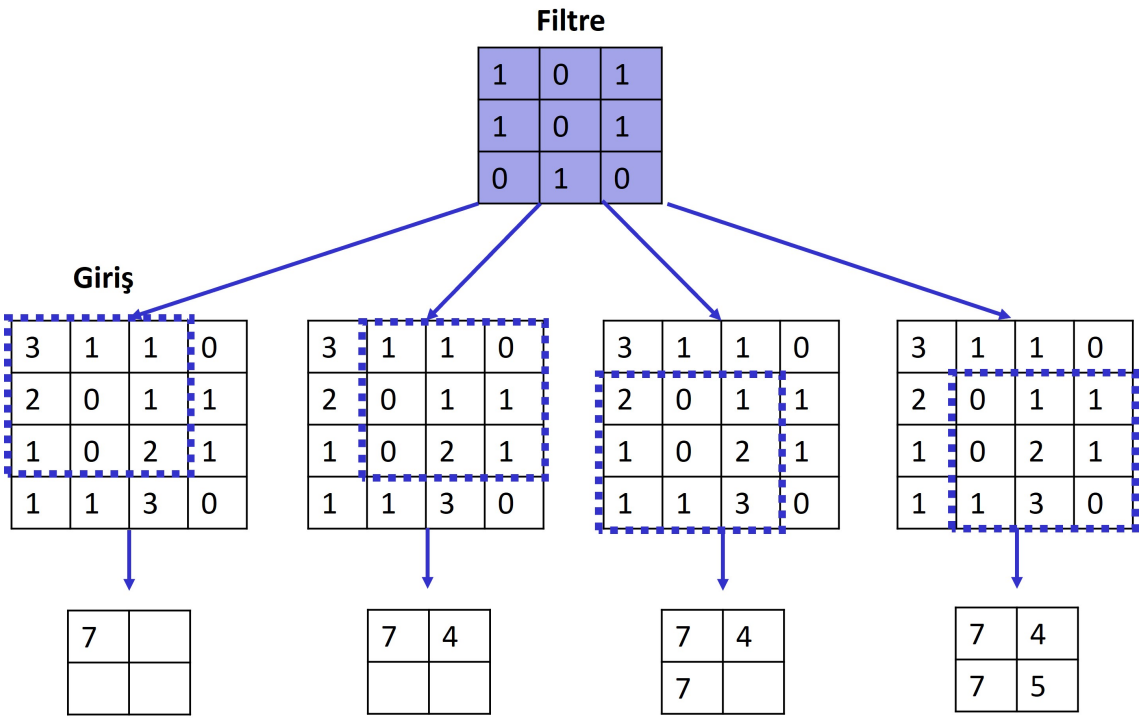
Görüntü işleme alanında renkli görüntüler genellikle bir dijital kamera ile çekilir ve üç boyutlu RGB (Red Green Blue) renk uzayına kaydedilir. Renkli görüntüler ile çalışmanın avantajı görüntüdeki çok çeşitli durumlarla başa çıkma yeteneğidir. Bazı durumlarda görüntü sınıflandırılması işlemelerini iyileştirmek için renk uzayı dönüşümleri gerekebilir. Bazı çalışmalarda [105] RGB, HSI (Hue Saturation Intensity), HSV ve Lab gibi yaygın olarak kullanılan renk uzayları, görüntü tespiti ve sınıflandırmasında sıklıkla kullanılır. Ayrıca farklı renk uzayı dönüşümleri yapılarak çalışmanın başarısı karşılaştırılabilir. Böylece çalışmaya göre kullanılması gereken renk uzayları tespit edilir.

2.1.2.2 Evrişim

Evrişim yaygın olarak kullanılan özel bir doğrusal işlem türüdür. Giriş görüntüsünü temsil eden iki boyutlu matrise, daha küçük iki boyutlu çekirdek matrisinin karşılıklı piksellerinin çarpımlarının toplanması işlemidir. Evrişim sürecinde iki boyutlu çekirdek matrisi, giriş görüntüsü boyunca yukarıdan aşağıya ve soldan sağa doğru çalışır. Böylece evrişim işlemi ile bir pikselin çıkış değeri kendisinin ve komşu piksellerin değerlerinin ağırlıklı toplamı olarak bulunur. Şekil 2.2'de $3x3$ filtre ile evrişim örneği verilmiştir.

2.1.2.3 Filtreler

Filtreleme işlemi birçok görüntü işleme algoritması için önemli bir adımdır. Filtreler kullanılarak gürültü azaltma, gürültü ekleme, görüntünün kalitesini artırma, kenar belirleme, bazı ayrıntıların ayıklanması vb. gibi görüntü üzerinde istenilen işlemler yapılabilir. Filtreleme yapılırken aslında evrişim işlemi uygulanır. Genel olarak işlem süresi filtrenin boyutuna bağlıdır. Literatürde amacına uygun çok sayıda filtre kullanılır ve bunlardan bazıları aşağıda açıklanmıştır.



Şekil 2.2 Örnek evrişim işlemi

- **Ortalama Filtre**

  Görüntü işleme alanında yaygın olarak kullanılan en temel filtredir. Görüntü üzerinde kayan pencere yöntemi ile ortalama değer bulunarak hesaplanır. Filtrenin temel amacı bir piksel ile diğer pikseller arasındaki farklılığı en aza indirmektir. Çoğu zaman resimdeki gürültüyü azaltmak için kullanılır. Bu filtre kullanılırken resimdeki herhangi bir piksel değerini komşu piksellerin değerlerinin ortalaması ile değiştirir. Böylece o piksellin etrafındaki aykırı değerlere sahip pikseller değiştirilir. Genelde $3x3$ kare matris kullanılır. Bu filtre ile gürültü azaltılırken aynı zamanda görüntü yumuşatılmış olur. Filtre olarak kullanılan matrisin boyutu artırılırsa yumuşatma artar.

- **Ortanca Filtre**

  Görüntü içerisindeki her bir piksel değeri o pikselin çevresindeki piksellerin ortanca (medyan) değeri ile değiştiren bir yöntemdir. Ortanca filtresi ortalama filtresi gibi gürültü azaltmak için kullanılan filtrelerden biridir. Ortanca filtresinin ortalama filtresinden en büyük farklı doğrusal olmayan bir filtre olmasıdır.

- **Gauss Filtre**

  Gauss dağılımı fonksiyonunu kullanarak filtreleme işlemi yapar. Gauss filtresi ile görüntünün yumuşatılması ve bulanıklaştırılması gibi işlemler yapılarak

görüntüdeki gürültüler azaltılır. Gauss yumuşatma filtresi çalışması açısından ortalama filtreye benzer. Gauss filtresi görüntüyü düzgünleştirirken bu düzeltme derecesi gauss standart sapması ile belirlenir. Standart sapma değeri ne kadar büyük ise kullanılacak matris boyutu da o kadar büyük olur. Böylece elde edilen çan eğrisi grafiği daha geniş olur. Gauss ile her bölgenin kendi içerisinden ağırlıklı piksel ortalaması elde edilir. Ortalama filtrede olduğu gibi her piksel ortalama değer gibi aynı değeri almaz. Bundan dolayı kenarların korunmasında ortalama filtreden daha iyi sonuç elde edilir.

2.1.3 Görüntü Bölütleme

Bölütleme görüntüyü renk, doku, parlaklık ve kontrast gibi benzer özelliklere sahip parçalara veya görüntü içindeki nesnelere bölen süreçtir [106, 107]. Diğer bir ifadeyle, ana görüntüden aranan görüntünün ayrılma işlemine bölütleme denir. Görüntü bölütleme ile ilgili literatürde birçok yöntem kullanılmaktadır. Bölütleme yöntemleri günümüzde aktif bir araştırma alanıdır ve henüz tüm sorunların çözümü için genel bir bölütleme algoritması bulunamamıştır. Literatürde sık kullanılan genel görüntü bölütleme yöntemleri Şekil 2.3'te verilmiştir [108–110].



Şekil 2.3 Görüntü bölütleme yöntemleri

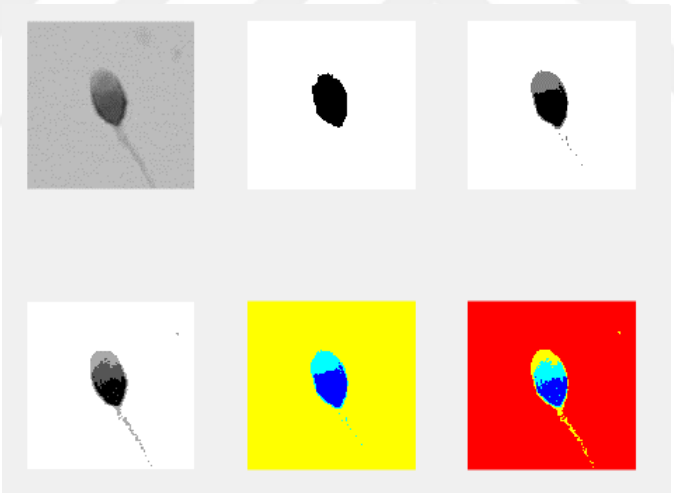
2.1.3.1 Eşik Tabanlı Yöntemler

En önemli ve yaygın kullanılan görüntü bölütleme yöntemlerinden biridir. Eşikleme yönteminde arka plan ile aranan görüntü farklı gri seviye piksellerine ayrılır. Çoğu zaman aranan görüntü siyah piksellerden oluşurken, arka plan ise beyaz piksel

değerlerine denk gelir. Bu farklılıktan dolayı görüntü arka plandan ayrılmış olur. Bu yöntem ön bilgi gerektirmediğinden diğer yöntemlere göre basit ve hızlıdır. Literatürde global ve yerel eşikleme yöntemlerini vardır. Yaygın olarak kullanılan yöntemlerden biri *Otsu* algoritmasıdır.

- **Otsu Algoritması**

  Nobuyuki Otsu [111] tarafından ortaya konulmuştur. Görüntü işlemede nesneleri arka planlarından çıkarmak için uygun bir gri seviye eşiği seçilmesi önemlidir. *Otsu* otomatik bir eşik seçim bölge tabanlı bölütleme yöntemidir. Öncelikle görüntü ikili formata dönüştürülür. Renklerin piksel değerleri görüntü histogramı ile hesaplanır ve eşik değeri tespit edilir. Bu işlem yapılırken en yüksek sınıflar arası değişinti (varyans) ve en düşük sınıf içi değişinti göz önüne alınır. [112]. Şekil 2.4 farklı seviyelerde *Otsu* yöntemi uygulanmış sperm hücre örneklerini göstermektedir. İlk görüntü orijinal görüntüden elde edilmiş gri seviyeli görüntüyken, diğerleri orijinal görüntüye *Otsu* yöntemi farklı eşik değerlerinde uygulanarak elde edilmiştir.



Şekil 2.4 Farklı eşik değerlerine göre otsu metodu örneği

2.1.3.2 Kenar Tespiti Yöntemleri

Kenar algılama görüntüdeki bir kenarın veya hattın varlığını belirleyen ve bunları uygun bir şekilde ana hatlarıyla belirten bir tür görüntü bölütleme tekniğidir. Kenar algılama yapılırken sınır bölgeler arasındaki yoğunluk farkı belirlenerek görüntü bölümlere ayrılır. Bunun amacı işlenecek veri miktarını en aza indirgeyerek görüntü verisini basitleştirmektir. Görüntü içindeki nesnelerin arasındaki zıtlık (kontrast/contrast) yüksekse kolay bölütlenebilir. Fakat çok sayıda kenarın olduğu

durumda uygun değildir. Literatürde temelde gradyan tabanlı ve Laplace tabanlı olmak üzere iki tür kenar belirleme algoritmaları vardır [113].

- **Gradyan Tabanlı Yöntemler**

  Gradyan tabanlı yöntemler görüntünün birinci mertebeden türevde en büyük ve en küçük değerleri kullanarak kenarları algılar. En önemli yöntemleri *sobel*, *prewitt* ve *robert* operatörleridir.

  - **Sobel Operatörü**

    Sobel kenarı detektörü satır ve sütunlar arasındaki ayrık farkları kullanarak gradyanı hesaplar. Sobel operatörü görüntünün küçük, ayrılabilir ve tam sayı değerli bir filtreyle evrişilmesine dayanır [114].

  - **Prewitt Operatörü**

    Prewitt kenar algılayıcısı bir kenarın büyüklüğünü ve yönünü tahmin etmenin uygun bir yoludur. Prewitt operatörü 8 olası yönlendirme ile sınırlıdır. Bu gradyan tabanlı kenar dedektörü, 8 yön için $3x3$ matrisin komşularını tahmin etmektedir. 8 evrişim maskesinin tümü hesaplanır. Daha sonra en büyük modüle sahip evrişim maskesi seçilir.

  - **Robert Operatörü**

    Dikey ve yatay kenarları ayrı ayrı ortaya çıkarır ve daha sonra kenar algılaması için bir araya getirir. Operatör $2x2$ evrişim matrisine sahiptir.

- **Laplace Tabanlı Yöntemler**

  Bu yöntemde ise görüntünün ikinci mertebeden türevi yoluyla sıfır geçişi için kenarlar aranır. LoG ve Canny en çok kullanılan Laplace yöntemleridir.

  - **LoG (Laplacian Of Gaussian)**

    LoG görüntüdeki kenarları tespit etmeye yarayan filtredir. LoG filtreleme gürültü giderme işlemlerini ve ikinci türevi hesaplamayı içerir. LOG filtreleme gerçek zamanlı bir kenar algılama yöntemi olarak kullanılmaktadır.

  - **Canny**

    Canny algoritması türev işlemini temel alan doğrusal ve sürekli bir filtreme yöntemidir. Böylece kenar tespiti, gürültü giderme ve kenar belirginleştirme gibi özellikleri barındıran bir diğer kenar belirleme yöntemidir. Algoritmanın amacı, kenar değerlerini gürültü değerlerinden daha güçlü yaparak sonuca ulaşmaktır [115]. Canny yöntemi gürültüye karşı daha az hassastır ve diğer yöntemlerden daha iyi performans gösterir.

2.1.3.3 Bölge Tabanlı Yöntem

Bölgeye dayalı bölütleme yöntemleri görüntüyü benzer özelliklere sahip çeşitli bölgelere ayıran yöntemlerdir. Aynı türdeki pikseller tanımlanır ve aynı tür bölgeler halinde gruplanırlar. Aslında bölge tabanlı yöntemlerde, homojenlik ilkesine dayanarak benzer özellikli pikseller aynı bölge olarak düşünülür [107]. Gürültülü görüntülerin bölütlenmesinde oldukça başarılıdırlar. Fakat fazla işlem gerektiğinden zaman açısından maliyetlidirler. Bölgeye dayalı bölütlemenin ana türleri bölge büyütme, bölge bölme ve bölge birleştirmedir.

2.1.3.4 Kümeleme Tabanlı Yöntem

Görüntünün histograma dönüştürüldüğü ve görüntüyü benzer özelliklere sahip pikselleri kümeleyerek bölen yöntemdir. En çok kullanılan kümeleme algoritmalarından biri, *k*-ortalamalar (*k*-means) algoritmasıdır. Gürültülü durumlarda iyi çalışır. Bölümlenecek küme sayısının az olduğu durumlarda hesaplama hızlıdır. Fakat bölümlenecek küme sayısını tahmin etmek bazen zor olabilir.

2.1.3.5 Sınır Tabanlı Yöntemler

Su havzasına dayanan yöntemler olarak bilinir. Su havza sınırına ulaştığında bitişik havzalar bir araya getirilir. Havza yöntemleri görüntünün gradyanını topografik yüzey olarak kabul eder. Daha fazla gradyana sahip pikseller sürekli olan sınırlar olarak temsil edilir. Tespit edilen sonuçlar sürekli olduğundan genellikle kararlıdır. Ancak gradyanların hesaplanması karmaşıktır.

2.1.3.6 Kısmi Diferansiyel Denklem Tabanlı Yöntemler

Hızlı bölümlendirme yöntemlerinden biridir. Zaman açısından kritik uygulamalar için uygundur. Bu yöntem görüntüdeki gürültüyü azaltmak, kenarları ve sınırları daha iyi tespit etmek için kullanılır.

2.1.3.7 Yapay Sinir Ağı Tabanlı Yöntemler

Yapay sinir ağı tabanlı bölütleme yöntemleri karar alma aşamasında insan beyninin öğrenme stratejilerini taklit eder. Günümüzde bu yöntem çoğunlukla tıbbi görüntülerin bölütlenmesinde kullanılmaktadır. Gerekli görüntüyü arka plandan ayırmak için kullanılır. CNN ve derin otomatik kodlayıcılar gibi çeşitli derin öğrenme algoritmalarının geliştirilmesiyle görüntü bölütleme başarısı da artmıştır [116].

2.1.4 Öznitelik Çıkarılması

Bir görüntüden önemli olabilecek bileşenleri diğer bileşenlerden çıkartma işlemine denir. Uygun bileşenlerin seçilmesi, bir görüntü sınıflandırmasını başarılı bir şekilde uygulamak için kritik bir adımdır. Görüntü bölütleme yapıldıktan sonra görüntünün boyutu, şekli ve konumu gibi özellikleri kullanarak aranan görüntüye ait bazı nitelikleri elde etmektir. Çıkarılan önemli özniteliklerin sayısı genellikle giriş penceresindeki piksel sayısından daha küçüktür. Böylece veri sayısı ilk görüntüye göre azaltılmış olur. Özellik çıkarma aşaması görüntünün ikili formatından elde edilir. Bu çalışmada kullanılan veri setindeki sperm görüntüleri için bazı çalışmalarda özellik çıkarımı konusunda renk, şekil gibi özellikler dikkate alınmıştır. Şekil olarak yuvarlaklık değeri hesaplanır, bu değer belli bir eşik değerinin üstündeyse sperm değil yani sperm kalıntısı ya da lökosit olabilir yorumu yapılabilir. Özellik tespiti için literatürde birçok algoritma kullanılır. Bu yöntemler Harris Köşe Tespiti (Harris Corner Detection - HCD), Ölçekten Bağımsız Özellik Dönüşümü (Scale Invariant Feature Transform - SIFT), Hızlandırılmış Güçlü Özellikler (Speeded-Up Robust Features - SURF), Hızlandırılmış Bölütleme Testi (Features from Accelerated Segment Test - FAST) ve İkili Dayanıklı Bağımsız Temel Özellikler (Binary Robust Independent Elementary Features - BRIEF) içerir. Genel anlamda özellik çıkarımı yapılırken üç özelliğe bakılır.

2.1.4.1 Renk Özellikleri

Renk görüntülerin en önemli özelliklerinden biridir. Renk özellikleri belirli bir renk uzayına veya modele bağlı olarak tanımlanır. Literatürde RGB, LUV, HSV ve HMMD gibi renk uzayları kullanılmıştır [117].

2.1.4.2 Doku Özellikleri

Genel olarak renk bir pikselin özelliğini temsil ederken doku ise bir piksel grubunun özelliğini temsil eder.

2.1.4.3 Şekil Özellikleri

Şekil özelliği çıkarma yöntemleri genel olarak kontur tabanlı ve bölge tabanlı olmak üzere iki gruba ayrılır. Birincisi şekil özelliklerini yalnızca şeklin sınırından hesaplar, ikincisi ise tüm bölgedeki özellikleri ayıklar [118].

2.1.5 Görüntünün Sınıflandırılması

Görüntüdeki nesnelerin belli özniteliklerine göre ayrıştırılmasına görüntü sınıflandırma denir. Görüntülerin sınıflandırması birçok faktörden etkilenebilecek karmaşık bir problemdir. Yapay sinir ağları, karar ağacı sınıflandırıcıları ve bilgi tabanlı sınıflandırma gibi parametrik olmayan sınıflandırıcılar, çok kaynaklı veri sınıflandırması için giderek daha önemli yaklaşımlar haline gelmiştir. Genel anlamda sınıflandırma yapılırken özniteliklerin iyi seçilmesi performans açısından çok önemlidir. Belirli bir çalışma için hangi sınıflandırma yaklaşımının uygun olduğu sorusunu yanıtlamak kolay değildir. Genel anlamda görüntü sınıflandırma yaklaşımları denetimli ve denetimsiz, parametrik ve parametrik olmayan, hard ve soft sınıflandırma olarak gruplandırılabilir. Lu ve Weng [119] literatürde kullanılan geleneksel sınıflandırma yaklaşımlarına genel bir araştırma bakışı ile açıklık getirmişlerdir. Sınıflandırma sonuçlarının değerlendirilmesi uzman bilgisine dayalı olduğu gibi nicel değerlendirmelerle de yapılabilir. Bir sınıflandırma yönteminin performansı değerlendirilirken doğruluk, sağlamlık, uygulanabilirlik ve objektiflik kıstasları göz önünde bulundurulmalıdır.

3
DERİN ÖĞRENME YÖNTEMLERİ İLE GÖRÜNTÜ İŞLEME

Günümüzde Grafik İşleme Birimi (Graphics Processing Unit - GPU) ve Merkezi İşlem Birimi (Central Processing Unit - CPU) gibi işlemcilerin büyük ölçüde artan işlem yetenekleri, bilgisayar donanımının maliyetinin önemli ölçüde düşürülmüş olması, makine öğrenmesi ve veri işleme araştırmalarındaki son gelişmelerle beraber derin öğrenme yöntemleri yaygınlaşmıştır. Bu alanda kişisel çalışmaların dışında Google, Facebook, MIT (Massachusetts Institute of Technology) ve daha pek çok önemli kuruluş aktif olarak çalışmaktadır. Bu kuruluşlardaki araştırmacılar bilgisayarla görme, ses işleme, ses tanıma, sesli arama, konuşma ve görüntü özelliği kodlama, el yazısı tanıma, görsel nesne tanıma, bilgi alma, sınıflandırma gibi çeşitli uygulamalarda derin öğrenmenin başarılarını göstermişlerdir. Günümüzde derin öğrenme yöntemlerinin tıp dünyasında her geçen gün kullanımı artmaktadır. Hatta son zamanlarda yeni ilaçların elde edilmesi aşamasında moleküllerin analizinde bile kullanılmaktadır [120].

Literatürde çeşitli derin öğrenme modelleri geliştirilmiştir. En yaygın kullanılanlar Yığılmış Otomatik Kodlayıcı (Stacked Auto-Encoder - SAE), Derin İnanç Ağı (Deep Belief Network - DBN), Evrişimsel Sinir Ağı ve Tekrarlayan Sinir Ağı (Recurrent Neural Network - RNN) modelleridir. Diğer derin öğrenme modellerinin çoğu bu dört temel modelin türevleridir [121].

3.1 Yığılmış Otomatik Kodlayıcı

Yığılmış bir otomatik kodlayıcı modeli genellikle ileri beslemeli sinir ağları olan birkaç otomatik kodlayıcının yığılmasıyla oluşturulur. Temel bir otomatik kodlayıcının kodlama aşaması ve kod çözme aşaması olmak üzere iki aşaması vardır [122].

3.2 Derin İnanç Ağı

Başarılı bir şekilde eğitilen ilk derin öğrenme modeli derin inanç ağıdır. Yığılmış otomatik kodlayıcıdan farklı olarak derin inanç ağı birkaç sınırlandırılmış Boltzmann makinesi tarafından yığılır. Sınırlandırılmış Boltzmann makinesi görünür katman ve gizli katman olmak üzere iki katmandan oluşur. Temel bir sınırlandırılmış Boltzmann makinesi, parametrelerini eğitmek için Gibbs örneklemesini kullanır [123, 124]. Derin inanç ağı yığılmış otomatik kodlayıcıya benzer şekilde iki aşamalı katmandan oluşur. Eğitim öncesi aşamada başlangıç parametreleri denetlenmemiş bir şekilde kullanılırken, ince ayar aşamasında ise en son katmana softmax katmanı eklenerek denetleme stratejisi kullanır.

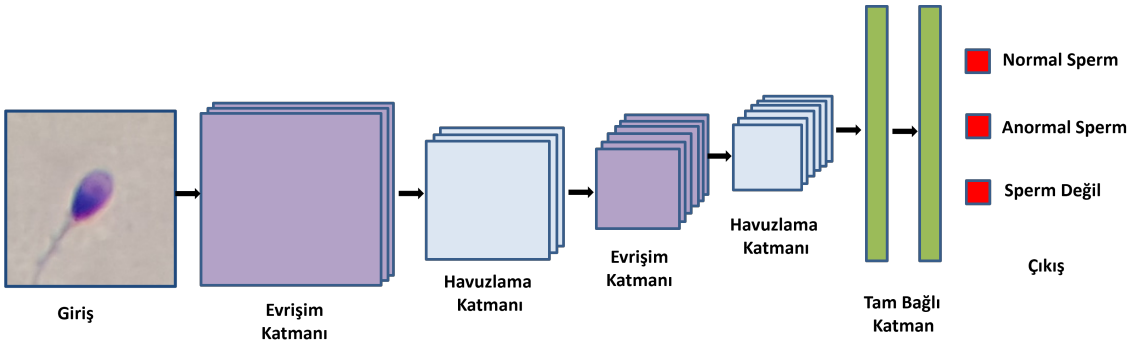
3.3 Tekrarlayan Sinir Ağı

Yığılmış otomatik kodlayıcılar, derin inanç ağları ve evrişimsel sinir ağları gibi geleneksel derin öğrenme modelleri zaman serilerini dikkate almaz. Bu nedenle bu ağlar zaman serisi verilerinin özelliklerini öğrenmeye uygun değildir. Çünkü önceden verilmiş girişlerin bilgilerini saklayamazlar. Tekrarlayan sinir ağları tipik bir sıralı öğrenme modelidir. Seri halindeki veri için özellikleri önceki hafızanın hafızasına göre öğrenir [125]. Tekrarlayan sinir ağı birimler arasındaki bağlantıların tek yönlü bir döngü oluşturduğu konuşma ve el yazısı da dahil olmak üzere geçici veriler için uzmanlaşmış bir ağdır [126].

3.4 Evrişimsel Sinir Ağı

CNN büyük ölçekli görüntü sınıflandırma ve tanıma için özellik öğrenmede en yaygın kullanılan derin öğrenme modelidir. Birçok çalışmada CNN'ler nesne algılama ve sınıflandırma gibi birçok görüntü işleme alanında üstün performans elde etmişlerdir. Bir CNN Şekil 3.1'de gösterildiği gibi evrişim katmanı, havuzlama katmanı ve tam bağlı katman olmak üzere üç katmandan oluşur [127].

Derin CNN mimarisi genellikle birçok katman içerdiğinden tahmin edilmesi gereken milyonlarca ağırlık parametresi olabilir. Bu nedenle modelin eğitimi ve parametre ayarı için çok fazla veri örneği gerektirir. Örneğin sınıflandırma probleminde kullanılmak üzere bir derin öğrenme mimarisini sıfırdan eğitmek için sınıf başına 1000'den fazla veriye ihtiyaç vardır. Veri sayısının az olduğu durumlarda ise veri artırma ya da önceden belirlenmiş ağın yeniden kullanımı ile başarı sağlanabilir [126]. Ağı eğitmek için ön ve geri aşama olmak üzere iki aşama vardır. Ön aşamada giriş görüntüsü her katmandaki geçerli ağırlıklar ve bias (sapma) değerleri eklenerek ifade edilir. Geri aşamada ise kayıp maliyetine dayanarak geriye doğru her parametrenin



Şekil 3.1 Örnek CNN modeli

gradyan zincir kuralları ile hesaplanır. İleri ve geri aşamaların yeterli yinelemesinden sonra ağın öğrenmesi durdurulur.

- **Evrişim Katmanı**

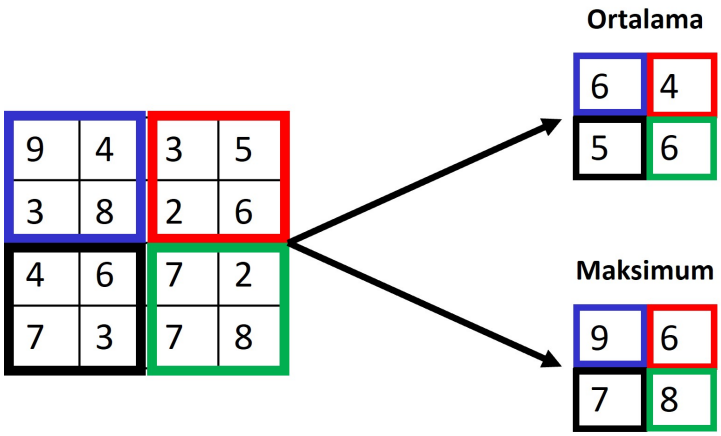
  Evrişim katmanı yoğun hesaplama işlemlerinin çoğunun yapıldığı evrişimsel sinir ağının temel yapı birimidir. Bu katmanda çeşitli özellik haritalarını oluşturan ve ara görüntü haritalarının yanı sıra bütün görüntüyü elde etmek için filtreler kullanılır. Bu filtreler sayesinde giriş görüntüsünden öznitelikler elde edilir. Her filtre boyutsal olarak, yani genişlik ve yükseklik açısından, giriş değerine göre daha küçük olmalıdır.

  Örneğin tipik bir filtrenin boyutu $7x7x3$ olabilir. Buradaki 7 değerleri genişlik ve yüksekliği belirtirken 3 değeri ise görüntülerin derinliğini yani kanal sayısını belirtir. İlk olarak bir giriş dizisinin bir alt dizisi ile çekirdekler arasında elementer çarpımları gerçekleştirilir. Daha sonra çarpılan değerler toplanır ve toplanan değerlere bias eklenir [128]. Bu işlem her kanal için ayrı ayrı yapılır [129]. Böylece herhangi bir konumdaki filtre girişleri ile giriş arasındaki nokta ürünler hesaplanır.

- **Havuz Katmanı**

  Bu katman evrişim katmanını izler ve bir giriş verisinin boyutu azaltılır. Bu işlem genellikle alt örnekleme olarak tanımlanır. Şekil 3.2'de gösterildiği gibi ortalama ve maksimum havuz en çok kullanılan örnekleme çeşitleridir. Maksimum havuzlama maksimum değerleri bir giriş dizisinin alt dizilerinden alırken ortalama havuzlama ortalama değerleri alır [128]. Son yıllarda yapılan çalışmalarda maksimum havuzlamanın ortalama havuzlamaya göre daha hızlı genellemeyi geliştirdiği keşfedilmiş böylece daha iyi performans gösterdiği doğrulanmıştır [130].

  Bu katmanda matrisin boyutu azaltıldığından ağdaki parametre ve hesaplama miktarı azalır böylece aşırı öğrenme engellenmiş olur. Alt örnekleme



Şekil 3.2 Örnek havuzlama katmanı

katmanında genellikle iki parametre kullanılır. Bunlar filtre boyutu F, adım sayısı S. Alt örnekleme sonucunda derinlik boyutu değişmez. Bu katmanda P dolgu kullanımı pek yaygın değildir. Literatürde birçok uygulamada filtre sayısı 2, adım sayısı 2 kullanılır. Ayrıca bu katman genellikle geniş ağlarda uygulanır ve genişliği az olan bir ağda uygulanması onun bilgi kaybetme oranının artırılmasına neden olur [129].

Giriş görüntüsü $W1xH1xD1$ boyutunda olsun, çıkış görüntüsü $W2xH2xD2$ boyutunda olsun. Çıkış görüntüsünün boyutlarının hesaplanması için $W2$, $H2$ ve $D2$ sırasıyla Eşitlik 3.1, 3.2 ve 3.3'te verilmiştir.

$$\mathrm{W2} = \frac{W1 - F}{S + 1} \tag{3.1}$$

$$\mathrm{H2} = \frac{H1 - F}{S + 1} \tag{3.2}$$

$$\mathrm{D2} = D1 \tag{3.3}$$

- **Tam Bağlı Katman - Fully Connected Layer**

  Evrişim ve alt örnekleme katmanlarından sonra bir veya birkaç tam bağlı katman gelir. İsminden de anlaşılacağı gibi nöronların önceki katmandaki tüm nöronlara bağlandığı bir katmandır. Görüntü boyunca önceki katmanlar tarafından öğrenilen tüm özellikleri birleştirir. Böylece iki boyutlu özellik haritaları bir boyutlu özellik vektörüne dönüştürülür. Tam bağlı katmanlar bir CNN'deki parametrelerin yaklaşık % 90'ını içerir.

  Örneğin tam bağlı katmanın $1x1x5$ gibi bir boyutla sonuçlanması 5 sınıflı bir yapıya karşılık geldiğini ifade eder. Tam bağlantılı katmanda sadece

giriş değerleri için değil aynı zamanda nöronlar, ağırlıklar ve sapmalar gibi parametrelerin de aktivasyon fonksiyonu dönüşümleri gerçekleştirilir. Bu işlem sadece evrişim katmanı ve tam bağlı katmanda gerçekleşir. Fakat alt örnekleme katmanlarında ve aktivasyonlarda sadece sabit bir fonksiyon uygulanır. Evrişim katmanı ve tam bağlı katmanlardaki parametreler gradyan iniş (gradient descent) ile eğitilir. Böylece evrişimsel sinir ağının hesapladığı sınıf puanları her görüntü için eğitim setindeki etiketlerle tutarlı olur [131]. Ayrıca evrişim, tam bağlı, alt örnekleme katmanları ek hiper parametreler içerebilirken *ReLU* (Rectified Linear Unit) gibi aktivasyon fonksiyonları ek hiper parametreler içermez.

3.5 Evrişimsel Sinir Ağı'nın Temel Bileşenleri

Bir önceki bölümde CNN'nin temel katmanlarından bahsedildi. Bu bölümde ise CNN'nin, katman tasarımı, aktivasyon işlevi, kayıp işlevi, düzenlileştirme ve optimizasyon gibi farklı yönleri açıklandı.

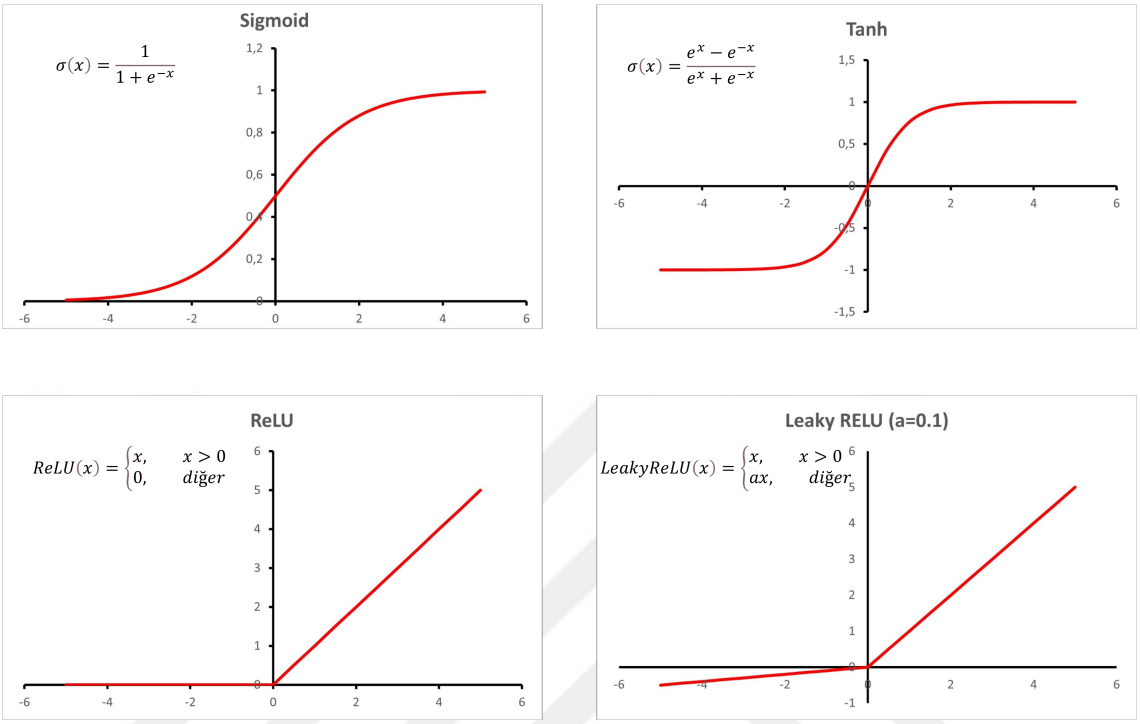
3.5.1 Toplu Normalleştirme - Batch Normalization

Toplu normalleştirme, katmanları bir ağ üzerinden aktivasyonlar ve gradyanlar ile normalleştirir. Böylece ağ eğitimini daha kolay bir optimizasyon problemi haline getirir [132]. Bu katman evrişimsel katman ile *ReLU* arasında uygulanır. Her bir giriş kanalını küçük topluluklar (mini-batch) halinde normalleştirir. Katmanda ilk olarak her bir kanalın küçük topluluklarının ortalaması elde edilir ve standart sapmasına bölerek normalleştirir. Bu katmanın asıl amacı ağın başarısını artırmaktır [133].

3.5.2 Aktivasyon Fonksiyonları

Aktivasyon fonksiyonları genellikle doğrusallıktan (linerlikten) doğrusal olmayan duruma (non-linerlik) geçişi sağlar. Bunun nedeni geri yayılımda türev alınmasını mümkün kılmaktır. Aktivasyon fonksiyonları birimin boyutunu değiştirmez. Yapay Sinir Ağlarında doğrusal olmayan durumlar için en çok sigmoid fonksiyonları kullanılır. Fakat sigmoid fonksiyonlarının eğitim hesaplamalarını yavaşlattığı ortaya konulmuştur. Son dönemlerde *ReLU* doğrusal olmayan bir aktivasyon işlevi olarak tanıtılmıştır [134]. Şekil 3.3'te doğrusal olmayan fonksiyonlardan *Sigmoid*, *Tanh*, *ReLU* ve *Leaky ReLU* gibi örnekler gösterilmektedir [128]. *ReLU* diğer aktivasyon fonksiyonlarıyla karşılaştırıldığında negatif çıktı değerleri 0 ile sınırlanmış olsa da, pozitif girdi değerleri için sınırlı bir çıkışı yoktur. Böylece *ReLU* sigmoidal fonksiyonlardan daha hızlı ve daha iyi doğruluk değerleri sağlar. *Leaky ReLU*

fonksiyonu ise *ReLU* fonksiyonun geliştirilmiş bir sürümüdür. *ReLU* fonksiyonu negatif değerler için nöronları devre dışı bırakırken *Leaky ReLU* ise bu sorunu çözmek için tanımlanmıştır. *Leaky ReLU* fonksiyonu x'in negatif değerlerini 0 tanımlamak yerine, onu x'in son derece küçük bir doğrusal bileşeni olarak tanımlar.



Şekil 3.3 Örnek aktivasyon fonksiyonları

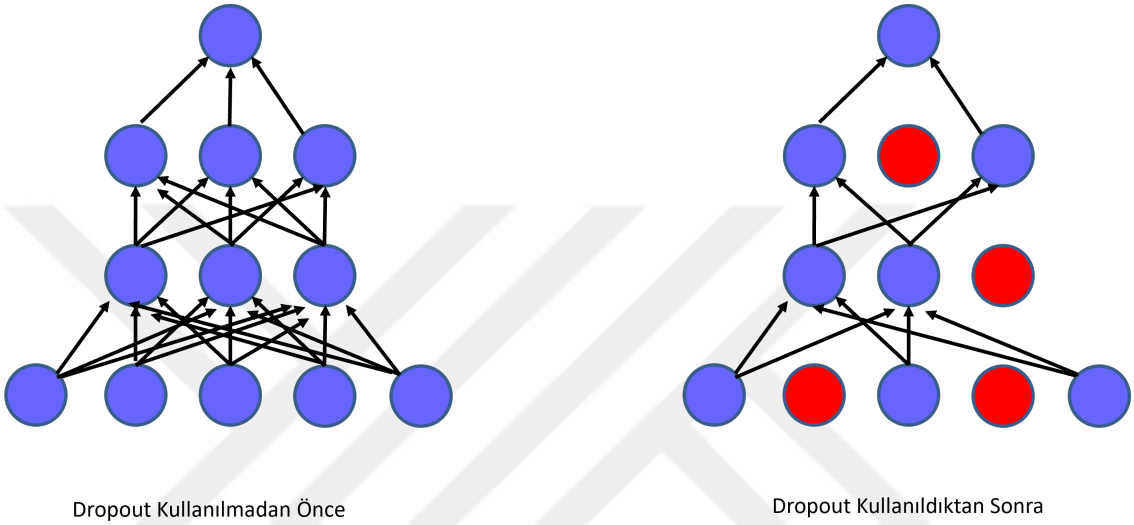
3.5.3 Optimizasyon Seçimi

Derin öğrenme başlı başına optimizasyon problemleri ile ilgilenen yöntemlerden oluşur. Doğrusal olmayan problemlerin çözümü için uygun değerlerin belirlenmesi işlemidir. Derin öğrenmede kullanılan en yaygın optimizasyon yöntemleri *Stochastic Gradient Descent (sgd)* , *adagrad*, *adadelta*, *adam*, *adamax* vb gibi algoritmalardır. *sgd* bir fonksiyonun en küçük değerini bulan yöntemdir. Yani kısaca herhangi bir problemin hata oranını en aza indirmek ve doğruluğu en yüksek bulmak için tekrarlı uygulanan optimizasyon yöntemidir. Asıl amaç en küçük hatayı elde etmektir. Bu algoritmalar zaman ve hız açısından farklı veri setleri üzerinde farklı sonuçlar verirler [135].

3.5.4 Dropout

Aşırı öğrenme (overfitting) makine öğrenmesi alanında uzun süredir devam eden bir problem haline gelmiştir. Bu problem modelin tatmin edici doğrulama sonuçları sağlamadığı bir olgudur. Bu sorunu çözmek için dropout (seyreltme) yöntemi

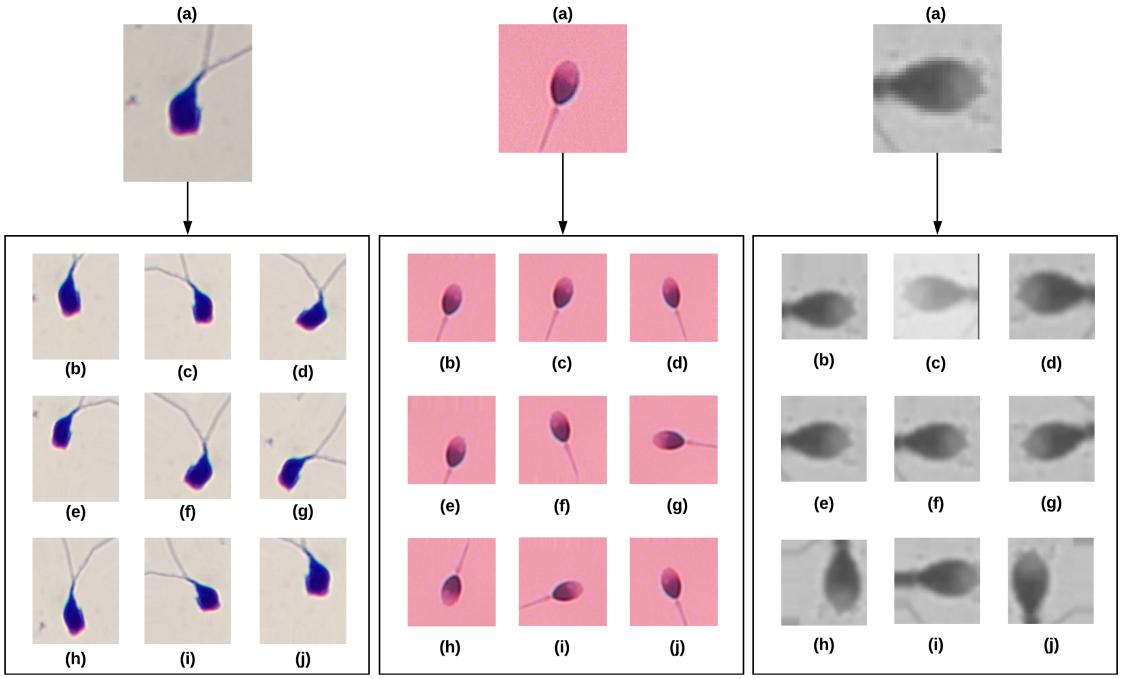
kullanılmaktadır. Burada kullanılan yöntem bağlı katmanların nöronları arasındaki bağlantıları belirli bir dropout oranıyla rastgele kesmektir. Seyreltme değeri rastgele seçildiği gibi modeli oluşturan kişi tarafından da 0 ve 1 arasında bir değer seçilebilir. Genellikle 0,5 gibi bir değer kullanılır. Ortalama 0,5'ten fazla olmayacak 0,2'den az olmayacak şekilde nöronlar kaldırılabilir. Seyreltme oranı ağ içinde her katmanda farklı değer alabilir. Şekil 3.4'te seyreltme örneği kullanılmayan ve kullanılan örnek ağlar birlikte gösterilmiştir. Böylece bir ağ için aşırı uyum azaltılarak eğitim örnekleri daha verimli bir şekilde kullanılabilir.



Şekil 3.4 Seyreltme örneği

3.5.5 Veri Sayısını Çoğaltma

Görüntü sınıflandırmada veri setinin büyüklüğü hayati öneme sahiptir. Veri setinin büyüklüğünün verinin eğitim süresini uzatma, depolama alanı sorunu oluşturma gibi farklı dezavantajları da vardır. Problem çözümünde depolama sorunu yoksa ve zaman açısından sorun oluşturmuyorsa başarının önemi için veri setinin büyüklüğü artırılmalıdır. Derin öğrenme modelleri kullanılırken daha fazla ve çeşitli veriyi eğiterek kullanılan yöntemin daha başarılı sonuçlar vermesi sağlanabilir. Görüntü verileri çoğaltılırken birçok yöntem vardır. Bu yöntemler görüntünün boyutunu büyütme-küçültme, görüntüyü çevirme (rotation), görüntünün kontrast değişikliği, görüntüye gürültü ekleme-azaltma, bulanıklık ekleme-azaltma gibi görüntü sayıları artırılabilir. Sistemin başarısı için bu yöntemlerin iyi seçilmesi gerekir. Şekil 3.5'te gösterildiği gibi üç sperm veri seti için her bir sperm görüntüsünden dokuz farklı sperm görüntüsü elde edilmiştir.



Şekil 3.5 Veri artırma yöntemlerinin üç farklı veri seti için örnek gösterimi

3.5.6 Transfer Öğrenme

Transfer öğrenme bir problemi çözerken kazandığı bilgiyi benzer problemlerin çözümünde kullanan yöntemdir. Diğer bir ifadeyle, önceden öğrenilmiş olan bilginin transfer yoluyla yeni bir görev için öğrenmenin geliştirilmesidir. Örneğin arabaları tanımayı öğrenirken edinilen bilgiler arabaya benzeyen diğer araçların tanınmasında kullanılır. Amaç daha fazla eğitim verisine veya nesne tespiti için daha iyi modellere ihtiyaç duyduğumuzda transfer öğrenme ile gerçekleştirmektir. CNN'leri kullanarak öğrenme aktarımı genellikle farklı alanlarda kullanılır. Shin ve arkadaşları tıbbi görüntülerde daha önceden eğitilmiş CNN modellerinde ince ayarlar yaparak algılama problemini çözmüşlerdir [136].

Bazı problemlerde transfer öğrenme sıfırdan eğitilmiş bir ağdan daha iyi sonuç verebilir. Önceden eğitilmiş ağlara örnek olarak AlexNet [137], Clarifai [138], GoogLeNet [139] ve VGG [140] gibi birçok ağ vardır. Bu ağlar giriş olarak farklı piksel değerleri kabul ederler. VGG16, VGG19 ve ResNet, $224x224$ giriş görüntülerini kabul ederken, Inception V3 ve Xception ise $299x299$ piksel girişi gerektirir. Bu ağlardan popüler olanlarının bir kısmına aşağıda detaylı yer verilmiştir.

- **AlexNet**

  AlexNet 2012'deki ImageNet Büyük Ölçekli Görsel Tanıma Yarışmasını (ILSVRC) kazanmıştır. Bu yarışmada ImageNet [141] veri tabanına ait bir seti kullanılmıştır. Kullanılan veri seti 1000 sınıfa ait 1,2 milyon $256x256$

görüntüden oluşur. Bu ağ beş evrişim katmanı ve üç adet tam bağlantılı katmandan oluşur. Bu ağda *ReLU* aktivasyon fonksiyonu kullanılırken, havuzlama katmanlarında da maksimum havuzlama kullanılır. Aşırı öğrenmeyi önlemek için dropout ilk defa AlexNet ile kullanılmıştır. Ortalama 60 milyon parametre hesaplanmaktadır.

- **VGGNet**

  VGGNet (ImageNet Large Scale Visual Recognition Challenge - ILSVRC) 2014 görüntü sınıflandırma yarışmasını kazanmıştır [140]. VGGNet'tin VGGNet-16 ve VGGNet-19 diye sırasıyla 16 ve 19 katman derinliğine sahip iki önemli mimarisi vardır. VGGNet-16 ağının toplam 41 katmanı vardır. 13 evrişimsel katman ve 3 tam bağlantılı katmanı olmak üzere toplam 16 öğrenilebilir katmanı vardır. VGGNet-19 ağının toplam 47 katmanı vardır. 16 evrişimsel katman ve 3 tam bağlantılı katman olmak üzere toplam 19 öğrenilebilir katmanı vardır. VGGNet basit mimarisi nedeniyle yaygın olarak kullanılır. Tüm derin öğrenme modelleri gibi girişten çıkışa doğru matrislerin boyutu azalırken derinliği artar [142].

- **ResNet**

  ResNet çok derin yapıya sahip sinir ağlarının eğitilmesini kolaylaştırmak için önerilmiştir [143]. Son derece mükemmel genelleme performansına sahiptir [142]. ResNet mimarisinde derinlik önemli ölçüde artmasına rağmen, VGG mimarisine göre daha düşük karmaşıklığa sahip olduğu görülmüştür. Bu bilgiden hareket edilerek ResNet modeli oluşturulmuştur.

- **GoogLeNet (Inception)**

  ImageNet verisetinde 2014 yılında sınıflandırma ve algılama için en iyi başarıya GoogleNet ulaşmıştır [139]. Bu ağın ismi aslında Inception'dir. Bu ağ çok çeşitli görüntüler için zengin özellik gösterimleri öğrenmiştir. Ağın giriş katmanı $299x299$ boyutlu bir resim için ayarlanmıştır. Bu ağ farklı boyutlu evrişim ve maksimum havuzlama işlemlerinden oluşur. Ayrıca aynı katmandaki farklı boyutlardaki filtre çekirdeklerinin kullanılmasına izin verir. Aşırı eğitmeye karşı daha hassas ve parametre sayısının daha az olduğu farklı bir CNN mimarisidir.

- **Xception**

  Derinlikli ayrılabilir evrişimler içeren bir CNN mimarisidir. Google araştırmacıları tarafından geliştirilmiştir. Xception, standart Inception modüllerini derinden ayrılabilen evrişimlerle değiştiren Inception mimarisinin bir uzantısıdır.

- **MobileNetV2**

  Google tarafından yenilenen farklı bir CNN ağıdır. Mobil cihazlar veya düşük işlem gücüne sahip tüm cihazlara uygun olan ağın karmaşıklık maliyetini ve model boyutunu önemli ölçüde azaltan bir modeldir. Özellik çıkarma, nesne algılama ve anlamsal bölümleme için kayda değer sonuçlar elde edilir.

- **NASNet**

  Küçük bir veri setinde bir mimari yapı bloğunu daha büyük bir veri setine transfer etmeyi önermektedir. Önce CIFAR-10'da daha sonra ImageNet'e uygulanmıştır. Sonuç olarak NASNet modeli daha küçük model boyutu ve daha düşük karmaşıklık ile önemli sonuçlar elde edebilir.

3.6 Evrişimsel Sinir Ağlarında Parametre Seçimi

Derin öğrenme modeli tasarlanırken modelin başarısını etkileyen parametreleri modeli tasarlayan kişi hesaplar ve tasarlanan modelin her zaman iyi bir performans vereceği kesin değildir. Model oluşturulurken kullanılan hiper parametreler ve model eğitimi için hiper-parametreler olmak üzere ikiye ayrılır [144]. Problemin başarısında bu hiper parametreler hayati önem taşımaktadır.

3.6.1 Model Oluşturulurken Kullanılan İlgili Hiper Parametreler

Model oluşturulurken kullanılan hiper parametreler filtre boyutu, filtre sayısı, adım sayısı, dolgu (padding) ve aktivasyon fonksiyonudur.

- **Filtre Boyutu**

  Filtre boyutu problemin başarısını ve çıkış görüntüsünün boyutunu etkiler. Gereğinden büyük boyutlu bir filtre bazı özellikleri göz ardı edebilir ve görüntülerdeki önemli ayrıntıları ortadan kaldırabilir, daha küçük boyutlu filtreler de zamansal karmaşıklığa neden olabilir [144]. Derin öğrenmede genelde $3x3$, $5x5$, $7x7$ boyutlu filtreler kullanılır.

- **Filtre sayısı**

  Filtre sayısı girişteki aynı bölgeye bağlanan evrişimsel katmandaki nöronların sayısına karşılık gelir. Bu parametre, evrişimsel katmanın çıkışındaki kanal sayısını belirler. Genelde modelde kullanılan filtre sayısı artıkça öğrenilen özellik sayısı da doğru orantılı olarak artar. Böylece ağın karmaşıklığı artmış olur. Fakat gereğinden fazla filtre olması giriş görüntüsü üzerinde yineleme yapacağından

gereksiz yere hafıza kullanılmasına neden olur. Toplam parametre sayısı Eşitlik 3.4'teki gibi bulunur. K değeri filtre sayısını ifade eder.

$$\text{Toplam Parametre sayısı} = ((W \times H \times D) + 1) \times K \tag{3.4}$$

Burada 1 bias değeridir.

- **Adım**

  Görüntü matrisi üzerinde ağırlık matrisi olan filtrenin her defasında kaç adet piksel kaydırılması işlemine adım (stride) denir [129]. Filtre işlem yaptığımız görüntü üzerinde gezdirilirken adım sayısı ile kontrol edilir. Adım 1'e eşit olduğunda filtre her defasında bir piksel hareket ettirilir. Bazen adım 2 veya nadiren 3 olarak kullanıldığında filtre adımlara bağlı olarak her seferinde 2 veya 3 piksel hareket ettirilir. Adımın boyutu artırıldığı zaman elde edilen çıktı o kadar küçük olur.

- **Dolgu**

  Adım uygulanırken çıkış görüntüsünün boyutu azalabilir. Bundan dolayı çıkış görüntüsünün boyutunu giriş görüntü boyutu ile aynı tutmak için dolgu yapılabilir. Bunun dışında çıkış görüntüsünün boyutunu istediğimiz gibi artırmak için giriş görüntüsüne 0 piksel değerleri eklenebilir. Dolgu aynı zamanda görüntünün çıktı hacminin boyutunu kontrol etmesini sağlar. Örneğin dolgu eklenmezse ve adım değeri 1 olarak kullanılırsa çıktı görüntüsünün uzamsal boyutu küçülür. Bir başka örnek ise $7x7$ girişli görüntü örneğinde adım 1 ve dolgu 1 kullanılırsa çıkış hacmi de $7x7$'ye eşit olur [131]. Bu eklenen değerin boyutu kullanıcı tarafından el ile istenildiği kadar eklenebilir ya da 'same' parametresiyle çıkış görüntüsü giriş görüntüsüyle aynı boyutta ayarlanabilir.

  Adım S = 1 olduğunda dolgunun Eşitlik 3.5'teki gibi ayarlanması, giriş hacminin ve çıkış hacminin uzamsal olarak eşit boyuta sahip olmasını sağlar. Genel olarak sıfır dolgunun bu şekilde kullanımı çok yaygındır. Formülden anlaşıldığı gibi dolgu boyutu filtre boyutuna bağlı olarak doğru orantılı seçilir. Örneğin; filtre boyutu 3 iken dolgu 1 piksel, filtre boyutu 7 iken dolgu 3 seçilebilir. Faklı bir dolgu ekleme işlemi ise Şekil 3.6'da görüldüğü gibi Pstart ve Pend ile ifade edilir. Fakat genellikle Pstart ile Pend eşit kabul edilir ve 2P olarak hesaplanır [145].

$$\text{P} = \frac{F-1}{2} \tag{3.5}$$

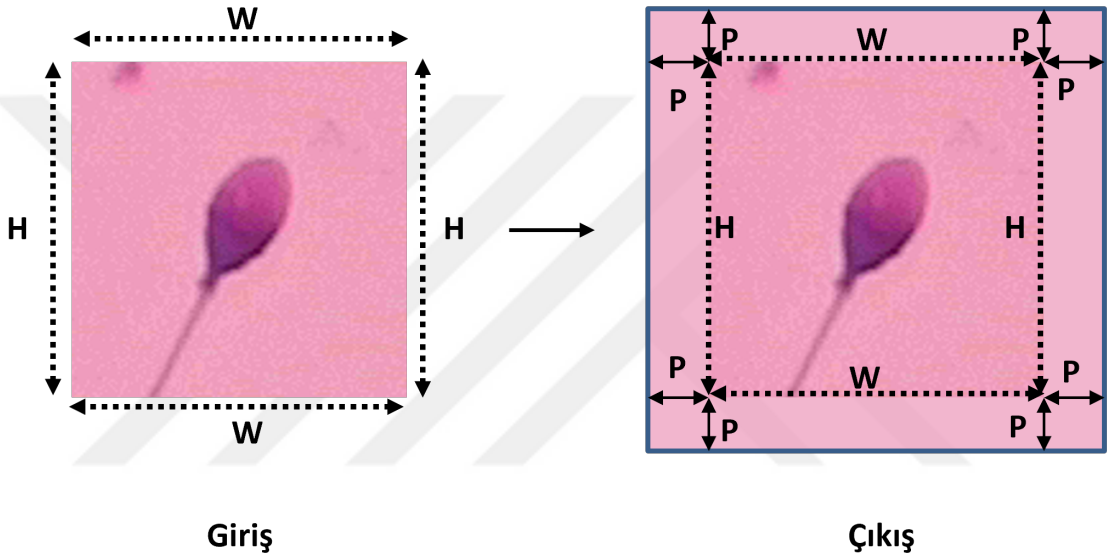
Aşağıdaki Eşitlik 3.6 ve 3.7'deki formül görüntünün boyutunu hesaplamaya

yarar. Formülde W giriş görüntüsünün genişliğini, H yüksekliği, F filtre boyutu, P dolgudur ve S adımdır.

$$\text{Çıkış Görüntüsünün Genişliği} = \frac{W - F_w + 2P}{S_w} + 1 \tag{3.6}$$

$$\text{Çıkış Görüntüsünün Yüksekliği} = \frac{H - F_h + 2P}{S_h} + 1 \tag{3.7}$$

Yukarıdaki formüle göre örneğin $7x7$ giriş görüntüsü ve adım=1 ve dolgu=0 kullanıldığında, $3x3$ filtre için sonuç $5x5$ çıkış elde edilir. Aynı görüntü için adım=2 dolgu=0 ise $3x3$'lük bir çıktı elde edilir.



Şekil 3.6 Örnek dolgu yapısı

- **Genişleyen Evrişimler - Dilated Convolution**

  Evrişim katmanına uygulanan bir hiper parametre olarak bilinir. Bu yöntemde bitişik filtre kullanmak yerine filtrenin komşuları arasında dilatasyon adı verilen boşluklar olan filtreler kullanılır. 'DilationFactor' özelliğini kullanarak dilatasyon faktörünün değeri belirlenir. Burada amaç parametre veya hesaplama sayısını arttırmadan katmanın gördüğü alanı artırmak için kullanılır. Formülü Eşitlik 3.8'de belirtilmiştir.

$$\text{Yeni Filtre Boyutu} = (F - 1) \times \text{Dilatasyon Faktörü} + 1 \tag{3.8}$$

Örneğin dilatasyon faktörü [2 2] olan $3x3$'lük filtre için elemanlar arasında sıfır bulunan $5x5$'lik filtreye eşdeğerdir. Bir başka örnekte ise [2 3 1] gibi bir dilatasyon uygulandığında filtreyi dikey olarak 2 kat, yatay olarak 3 kat ve derinlik boyunca 1 kat büyütür.

3.6.2 Eğitim İle İlgili Hiper Parametreler

Bu aşamada, eğitim aşamasında optimizasyonu sağlayan öğrenme oranı, maksimum iterasyon sayısı, mini batch değeri gibi hiper parametreler kullanılır. Eğitimin optimizasyonu ile ilgili açıklanan kavramlar ile ilgili Matlab'ta oluşturulan örnek parametreler ve değerleri Şekil 3.7'de belirtilmiştir.

```
options = trainingOptions('sgdm', ...
    'MiniBatchSize',64, ...
    'MaxEpochs',30, ...
    'InitialLearnRate',0.001, ...
    'LearnRateSchedule','piecewise',...
    'LearnRateDropFactor',0.1,...
    'LearnRateDropPeriod',8,...
    'L2Regularization', 0.004,...
    'Shuffle','every-epoch', ...
    'ValidationData',imdsValidation, ...
    'ValidationFrequency',4, ...
    'Verbose',false, ...
    'Plots','training-progress');
```

Şekil 3.7 Model eğitimi ile ilgili örnek parametreler

- **Katman Sayısı**

  Katman sayısının artması şüphesiz başarı oranını artırırken belli bir sayıdan sonra katman sayısının artırımı problemin başarısını çok etkilemez. Ayrıca, çok sayıda katman olması sinir ağının algoritmik karmaşıklığının artmasına veya ezberleme durumuna düşülmesine neden olabilir. Bunun yerine katman sayısı artırılırken ona paralel olarak diğer hiper parametrelerle, problem çözüme uygun hale getirilmelidir.

- **Nöron Sayısı**

  Modelin öğrenme kapasitesinin en önemli parametrelerinden bir diğeri katmanlarda kullanılacak nöron sayısıdır. Nöron sayısının artması başarıyı artırırken depolama ve zaman anlamında sorun teşkil etmektedir. Nöron sayısının az olması uyumsuzluğa neden olur. İşlemin basitliği ve karmaşıklığına göre nöron sayısı da doğru orantılı olarak belirlenir. Tabi çok sayıda nöron kullanılması ezberleme problemini de beraberinde getirecektir.

- **Tur Sayısı**

  Bir tur (epoch) tüm veri setinin yalnızca bir kez ağ üzerinden ileri ve geri aktarılmasıdır. Bunun sonucunda ağırlıklar güncellenir böylece her seferinde yeni eğitim setiyle eğitime devam edilir. Bu eğitimin her bir turuna epoch denir. Doğal olarak ilk epoch'ta modelin başarısı düşük çıkar ve her epoch'ta artar. Bu

artış belli bir seviyeye gelince değişmeyecektir. Epoch sayısı büyüdükçe genelde başarı da artar. Fakat epoch'lar arası artış hızında değişim azalınca eğitim durdurulmalıdır. Modelimizin eğitimi ne zaman durdurulmalıdır sorusuna doğrulama hatası son 10 ya da 20 epoch'ta düzelmezse eğitim durdurulmalıdır denebilir [146].

- **Mini-Batch Boyutu**

  Verilerin birlikte aynı anda derin öğrenme ağlarında işlenmesi zaman ve hafıza açısından problem yarattığı için verilerin parça parça derin öğrenme ağlarında işlenmesine mini-batch denir. Derin Öğrenme modeli tasarlanırken bu mini-batch boyutu problem çözücü tarafından belirlenir. Mini-Batch çok küçük seçildiği durumlarda öğrenme esnasında gürültüye neden olabilir. Mini-Batch değeri çok büyük seçildiğinde ise model iyi öğrenmemiş olur. Bu yüzden Mini-batch değeri belirlenirken çok küçük ya da çok büyük değerlerden kaçınılmalıdır. Genellikle literatürde 2'nin katları olmak üzere 8 ile 256 arasında değerler kullanılır. Yeterli veriye sahip bir veri setinde en sık kullanılan mini batch boyutu 32 olarak kullanılır.

- **Veri Seti Karıştırma - Shuffle**

  Aşırı öğrenmeyi yani ezberlemeyi önlemek için modelin eğitiminden önce veri setinin karıştırılmasına denir.

- **Öğrenme Hızı**

  Öğrenme hızı (learning rate) modelin başarısını önemli ölçüde etkiler. Öğrenme hızının küçük olması öğrenmeyi geciktirirken büyük olması gürültüye ve salınıma neden olur [146]. Eğitim süreci boyunca küçük ya da sabit bir değer kullanmak yerine eğitimin başlangıcında daha büyük bir öğrenme hızı seçilmesi ve optimizasyon sırasında bu değeri kademeli olarak azaltmak daha uygun olur [147]. Böylece eğitim süresi kısalır ve eğitim ilerledikçe hatanın en aza indirgenmesi sağlanır.

- **Momentum Katsayısı**

  Gürültü ve grafikteki salınımları azaltmak için kullanılan normalizasyon yöntemidir. Literatürde momentum katsayısı için 0,9 gibi bir değer kullanılır. Aldığı genel aralık 0,8 - 0,99'dir.

- **Başlangıç Ağırlık Değerlerinin Belirlenmesi**

  Ağırlık değerleri evrişim katmanı ve tam bağlantılı katmanda kullanılır. Model başlatılırken sıfır gibi ağırlık değerleri seçilebildiği gibi farklı değerler de seçilebilir.

- **Ağırlıkların Düzenlenmesi**

  Kullanılan modelin daha iyi genelleştirme yapması için öğrenme algoritmasında başvurulan değişikliğe denir. Aşırı öğrenmeyi (overfitting) önlemek için ağırlıklar ve sapmaların küçük değerler almasını sağlamaktır [148]. Bu kaybı engellemek için ağırlıklarla ilişkili bir maliyet eklenir. *L1*, *L2* olmak üzere iki tür ağırlık düzenleyici yöntem bulunur. *L1* Regularization eklenen maliyet ağırlık katsayılarının mutlak değeri ile orantılıyken, *L2* Regularization ise eklenen maliyet ağırlık katsayılarının değerinin karesiyle orantılıdır. Böylece regularization ile modelin başarısının artması sağlanabilir.

- **Adımlardaki Sınırlamalar**

  Bazı durumlarda çıkış görüntüsü boyutu hesaplanırken parametreler farklı değerlerde kullanıldığı için bazı hatalara neden olur. Bundan dolayı parametrelerin kullanımı ile ilgili bazı kısıtlamaları vardır. Örneğin giriş W = 14 boyutuna sahip olduğunda, dolgu kullanılmadığını yani P = 0 ve filtre boyutu F = 3 olduğu zaman S = 1 adımını kullanmak imkânsız olacaktır.

$$\text{Sonuç} = \frac{W - F + 2P}{S + 1} \tag{3.9}$$

  Çünkü Eşitlik 3.9'da Sonuç = 5,5 gibi tamsayı olmayan bir değer elde edilir. Bu da nöronların giriş boyunca düzgün ve simetrik olarak uygun olmadığını gösterir. Bu gibi durumlarda evrişim katmanında uyum sağlamak için dolgu ekleme, giriş görüntüsünün kırpılması gibi yöntemler uygulanabilir.

- **Parametre Paylaşımı**

  Parametre paylaşımı ile evrişim katmanlarında parametre sayıları kontrol edilir. Örneğin $228x228x3$ boyutundaki görüntü için ilk evrişim katmanında filtre boyutu F = 13, adım sayısı S = 4 olan ve dolgu olmayan yani P = 0 için çıkış görüntüsünün boyutu hesaplanır. Çıkış görüntüsü hesaplanırken Eşitlik 3.10'da verildiği gibi 43 olarak hesaplanır.

$$\text{Çıkış görüntüsünün boyutu} = \frac{W - F + 2P}{S + 1} = 43 \tag{3.10}$$

  Evrişim katmanının filtre sayısı yani derinliği K = 64 verildiğinde evrişim katmanının çıktı hacmi $43x43x64$ boyutuna sahip olur.

  Filtre sayısı K, filtre boyutu F, adım S, dolgu P,

  Giriş görüntüsü $W1 \times H1 \times D1$ boyutunda olsun,

  Çıkış görüntüsü de $W2 \times H2 \times D2$ boyutunda olsun.

Burada $W2$, $H2$ ve $D2$ değerlerinin hesaplanması sırasıyla Eşitlik 3.11, 3.12, 3.13'te verilmiştir.

$$\mathrm{W2} = \frac{W1 - F + 2P}{S + 1} = 43 \tag{3.11}$$

$$\mathrm{H2} = \frac{H1 - F + 2P}{S + 1} = 43 \tag{3.12}$$

$$\mathrm{D2} = K = 64 \tag{3.13}$$

3.7 Sınıflandırma Sonuçlarının Değerlendirilmesi

Sınıflandırma yapılırken sonuçların değerlendirilmesi aşamasında sadece doğruluğa bakarak karar vermek yeterli değildir. Bunun yanında kesinlik, duyarlılık, özgüllük ve F-skor değerlerine de bakmak gerekir [149]. Bu parametrelere bakılırken bir karışıklık (confusion) matrisi oluşturulur [150]. Bu matris Şekil 3.8'de görüldüğü gibi oluşturulan model ile gerçekte var olan sınıfların doğruluğunu tahmin eder. Sütün değerleri verilerin olması gereken gerçek sınıflarını ifade ederken yatay değerleri geliştirilmiş sistemin tahmin değerlerini ifade eder.

		GERÇEK DEĞERLER		
		Pozitif	Negatif	Toplam
TAHMİN DEĞERLERİ	Pozitif	TP	FP	Tahmin Pozitif Sayısı
	Negatif	FN	TN	Tahmin Negatif Sayısı
	Toplam	Gerçek Pozitif Sayısı	Gerçek Negatif Sayısı	

Şekil 3.8 Örnek karışıklık matrisi

Karışıklık matrisi tablosunda; True Positive (TP) ve True Negative (TN) önerilen sistemin doğru tahminlerini belirtirken, False Positive (FP) ve False Negative (FN) ise sistemin hatalı tahminlerini belirtir. Doğruluk değeri, Eşitlik 3.14'te gösterildiği gibi TP ve TN sayılarının toplamı yani matrisin köşegenlerinin toplamının tüm değerlere bölünmesiyle elde edilir. Farklı bir ifadeyle doğruluk değeri, tüm veri seti üzerinden doğru şekilde sınıflandırılan pozitif ve negatif örneklerin yüzdesini ifade eder [151]. Genel olarak doğruluk değerleri, dengeli veri setleri için sınıflandırma problemlerinde kullanılan en önemli performans ölçütüdür. Ancak, özel bir sınıf için sistem tutarlılığını analiz etmek gibi belirli görevlerde doğruluk bilgilendirici değildir.

$$\text{Doğruluk} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.14}$$

Kesinlik; sistemin pozitif olarak tahmin ettiği ve gerçekte de ne kadarının doğru tahmin edildiğini ifade eder. Kesinlik puanı, sınıfa dayalı, doğru tahmin edilen örneklerin performansını gösterir ve Eşitlik 3.15'te verilmiştir.

$$\text{Kesinlik} = \frac{TP}{TP + FP} \tag{3.15}$$

Eşitlik 3.16'da verildiği gibi; duyarlılık, sistemin pozitif olarak tahmin etmesi gereken verilerin ne kadarını pozitif olarak tahmin ettiğini ifade eder.

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \tag{3.16}$$

Özgüllük, verilerdeki negatif olarak tahmin edilenler arasında gerçek negatiflerin oranı olarak tanımlanır. Bir diğer ifadeyle gerçek negatiflerin toplam negatiflere oranı olarak tanımlanır. Eşitlik 3.17'de özgüllük hesaplaması verilmiştir.

$$\text{Özgüllük} = \frac{TN}{TN + FP} \tag{3.17}$$

F-skor değeri hesaplanırken duyarlılık ve kesinlik değerlerinin aritmetik ortalama değeri yerine harmonik ortalaması alınır. Harmonik ortalama kullanılmasının nedeni aykırı değerlerin ortalamaya etkisinin azaltılmasını sağlamaktır [152]. F-skor, her iki metriği de hesaba kattığı için daha objektif bir karşılaştırma sağlar [153]. F-skor hesaplama Eşitlik 3.18'de verilmiştir.

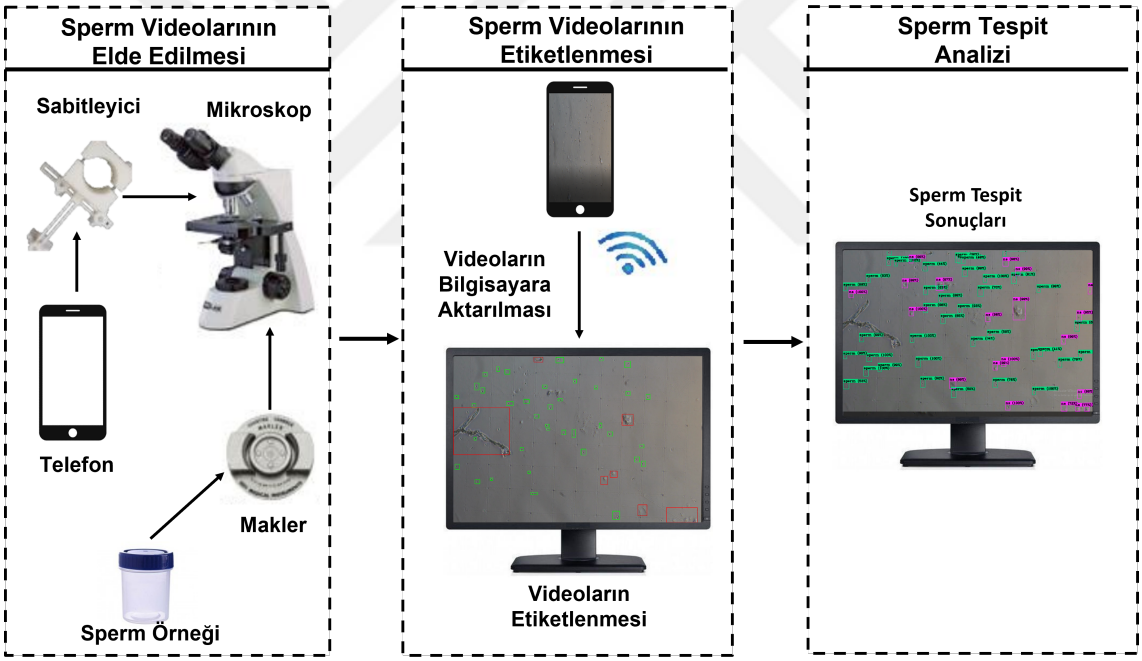
$$\text{F-skor} = \frac{2 \times \text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \tag{3.18}$$

Fleiss' kappa katsayısı [154] literatürde en sık kullanılan istatistiksel yöntemdir ve ikiden fazla yorumlayıcı arasındaki anlaşmanın güvenilirliğini ölçer. Bu yöntem, basit bir yüzde oranı yerine daha güçlü bir katsayı kullanır. Böylece elde edilen sonuçların tesadüf olmayacağına işaret edebilir. Fleiss' kappa katsayısı Eşitlik 3.19'da verilmiştir.

$$\text{Kappa} = \frac{\text{Toplam Doğruluk} - \text{Rastgele Doğruluk}}{1 - \text{Rastgele Doğruluk}} \tag{3.19}$$

4
TEZ KAPSAMINDA KULLANILAN MATERYALLER

Tez çalışması kapsamında hem sperm hücrelerine ait görüntüler hem de videolar üzerinde çalışmalar yapılmıştır. Sperm görüntüleri ile morfoloji analizi yapılırken sperm videoları üzerinde ise sperm tespiti yapılmıştır. Örnek veriler mikroskobun oküler kısmından akıllı telefon ile elde edilmiştir. Veri toplama düzeni ve işlenmesi ile ilgili süreç Şekil 4.1'de gösterilmiştir.



Şekil 4.1 Veri elde etme ve analiz yaklaşımı

Şekil 4.1'de görüldüğü gibi sperm örnekleri hastalardan elde edildikten sonra Makler odasına aktarılmış ve mikroskop aracılığı ile incelenmiştir. Telefon sabitleyici ile akıllı telefon sabitlenmiş, ardından mikroskopta görüntülenen sperm resimleri ve farklı uzunlukta sperm videoları telefona kaydedilmiştir. Kaydedilen görüntüler ve videolar etiketlenerek analiz edilmiştir.

4.1 Morfoloji Analizinde Kullanılan Veri Setleri

Tez kapsamında sperm morfolojisinin analizi için SMIDS, HuSHeM, SCIAN-Morpho ve Modified Human Sperm Morphology Analysis (MHSMA) olmak üzere dört farklı veri seti kullanıldı.

1. **SMIDS**

   Bu veri seti İstanbul Üniversitesi Tıp Fakültesinde İstanbul Üniversitesi Etik Kurulu tarafından onaylanarak elde edilmiştir. Örnek veren erkeklerin yaşları 19 – 37 aralığındadır. Semen örneği en az 2 ile 7 gün arasında cinsel perhiz sonrası elde edilmiştir. Örnekler sabah 9 - 11 saatleri arasında mastürbasyon yoluyla sağlanan steril numune kabında muhafaza edilmiştir. Görüntüler mikroskobun oküler kısmına monte edilen akıllı telefon ile elde edilmiştir. Toplamda 3000 görüntü bir uzman tarafından normal (1021), normal olmayan (1005) ve sperm olmayan (974) şeklinde etiketlenmiştir. Bu veri seti derin öğrenme algoritmaları için yeterli sayıda görüntüyü içeren ve sınıflar arası dengeli bir dağılıma sahiptir [51].

2. **HuSHeM**

   HuSHeM veri seti, İsfahan Doğurganlık ve Kısırlık Merkezi'ndeki 15 hastanın semen örneklerinden oluşturulmuştur. Normal, Tapered, Pyriform ve Amorphous olmak üzere dört sınıfı vardır [47]. Bu veri seti SMIDS'e benzer şekilde, her sınıf için eşit olarak dağıtılmış renkli görüntülerden oluşur. Her sınıf için ortalama olarak sadece 50 adet görüntünün bulunması, derin öğrenme yönteminin kullanması açısından bir dezavantajdır.

3. **SCIAN-Morpho**

   SCIAN-Morpho veri seti Şili Üniversitesi Tıp Fakültesi Spermiogram Laboratuvarından alınan manuel olarak kırpılmış ve döndürülmüş insan semen örneklerinden oluşur. Görüntüler, $35x35$ gibi çok düşük bir çözünürlükte ve gri tonlamalı bir renk uzayında toplanmıştır. Etiketleme yapan uzmanlar arasındaki anlaşmaya göre SCIAN-Morpho veri setinin üç versiyonu vardır.

   Üç uzman, etiketlemeyi manuel ve bağımsız olarak yapmıştır. İlk versiyonda, uzmanlar arasında etiketlerle ilgili herhangi bir görüş birliği yoktur. Bu veri seti için etiketler her uzman için farklılık gösterebilir. Bu veri seti 175 Normal, 420 Tapered, 188 Pyriform, 919 Amorphous ve 152 Small olmak üzere toplam 1854 sperm görüntüsü içermektedir. İkinci versiyonda, iki uzman arasında kısmi görüş birliğine dayalı bir karar durumu vardır. Bu versiyon, 100 Normal, 228 Tapered, 76 Pyriform, 656 Amorphous ve 72 Small olmak üzere toplam

1132 sperm görüntüsünden oluşmaktadır. Üçüncü versiyonda, görüntülerin seçiminde etiketler üzerinde tam bir görüş birliği vardır. 35 Normal, 69 Tapered, 7 Pyriform, 262 Amorphous ve 11 Small olmak üzere 384 sperm görüntüsü içermektedir. SCIAN-Morpho veri setinin ilk versiyonu, Amorphous ve Normal sınıflar için büyük bir veri dengesizliği sorunu içermektedir. Öte yandan, üçüncü versiyondaki görüntü sayısı, iyi eğitilmiş bir CNN modeli sağlamak için yetersizdir. Bu nedenle, tez kapsamında, sınıflardaki uzmanlar arasında kısmen anlaşma bulunan SCIAN-Morpho veri setinin ikinci versiyonu kullanılmıştır.

Bu veri seti üç zorluk içermektedir. İlk olarak, düşük çözünürlüklü ve düşük kaliteli bir veri setidir. İkincisi, ilk sürümden daha iyi olsa bile, örneklerin sınıf başına farklı şekilde düzenlenmesi ve veri dengesizliği sorunları vardır. Üçüncüsü, uzmanlar arasındaki kısmi görüş birliği nedeniyle bu veri seti için görüntülerin etiket bilgileri hakkında bazı faklı görüşler rapor edilmiştir. Veri setinin sınıfları arasındaki görsel farklılıklar uzmanlar tarafından bile kesin olarak ayırt edilememektedir.

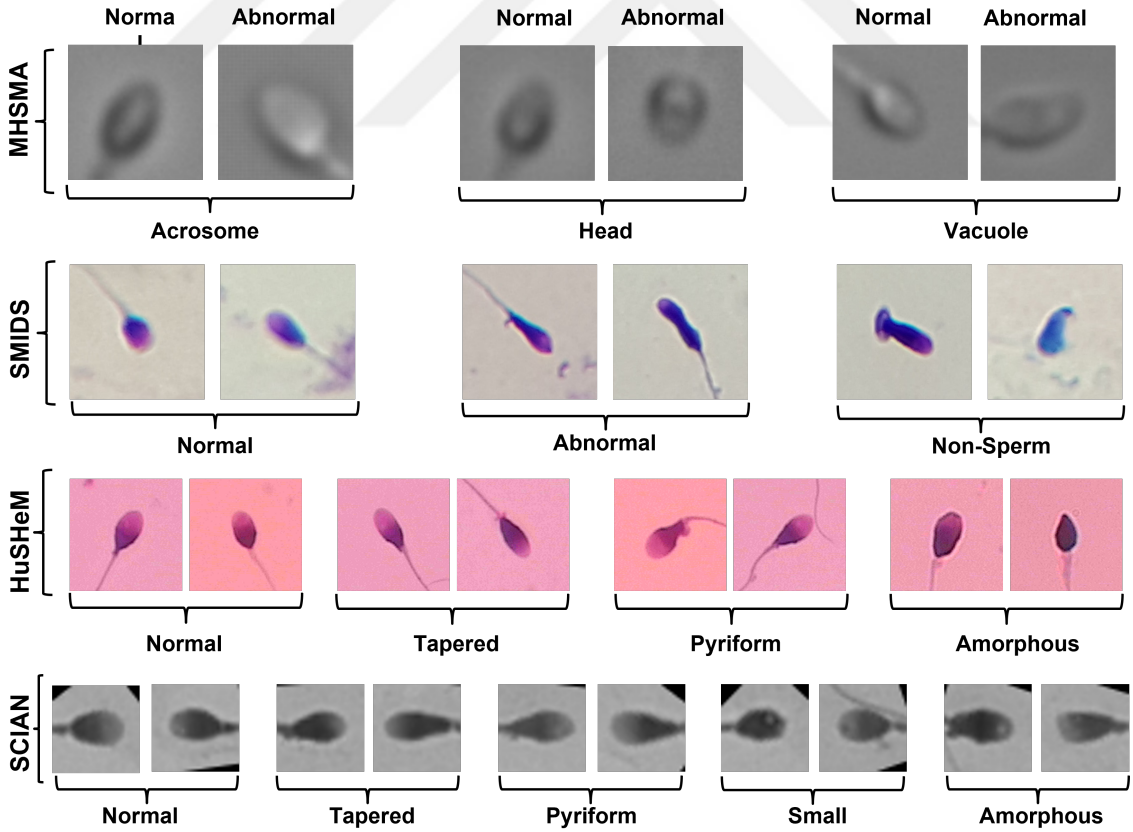
4. **MHSMA**

   MHSMA [155] açık erişimli ve düşük çözünürlüklü bir veri setidir. MHSMA veri seti aslında HSMA-DS'nin [44] yenilenmiş halidir. MHSMA sperm başı (Head) ile Akrozom, Vacuole ve Tail/Neck gibi bölümlerden oluşan bir veri setidir. Her bir bölümde toplam 1540 adet $64x64$ boyutunda, gri renge benzer RGB görüntüleri bulunur. Her bölümdeki görüntüler uzmanlar tarafından normal ve normal olmayan olarak etiketlemiştir. Ayrıca Vacuole için normal ve normal olmayan sınıflar, sırasıyla vakuolün yokluğu ve varlığı olarak kabul edilir. MHSMA veri setinin tamamı ve eğitim seti, validasyon seti ve test seti şeklinde üç farklı veri setine bölünmüştür. Bu veri seti, veri eksikliği ve sınıf bazında dengesiz dağılımlardan dolayı derin öğrenme açısından dezavantajlıdır. Tail/Neck sınıflandırılması Akrozom, Head ve Vacuole'nin sınıflandırılmasına göre oldukça kolaydır. Ayrıca daha önce yapılan çalışmalar incelendiğinde Tail/Neck sınıflandırılması yapılmadığından tez kapsamında akrozom, head ve vakuol gibi üç farklı problemin çözülmesine odaklanılmıştır.

Dört veri seti için her sınıftan örnek görüntüler Şekil 4.2'de sunulmuştur. Ayrıca, SMIDS, HuSHeM, SCIAN-Morpho veri setlerinde her sınıf için örneklerin dağılımı ve veri setlerinin toplam örneklem büyüklüğü Tablo 4.2'de sunulmuştur. MHSMA veri seti diğer veri setlerinden farklı olarak kendi içerisinden Acrosome, Head Vacuole, Tail/Neck gibi alt veri setlerine ayrıldığından Tablo 4.1'de ayrıntılı verilmiştir.

Tablo 4.1 MHSMA veri setinin dağılımı

Küme	Sınıf	# Normal	# Anormal
Veri Setinin Tamamı	Acrosome	1086	454
	Head	1122	418
	Vacuole	1301	239
	Tail and Neck	1471	69
Eğitim Seti	Acrosome	699	301
	Head	727	273
	Vacuole	830	170
	Tail and Neck	954	46
Validasyon Seti	Acrosome	174	66
	Head	176	64
	Vacuole	209	31
	Tail and Neck	233	7
Test Seti	Acrosome	213	87
	Head	219	81
	Vacuole	262	38
	Tail and Neck	284	16



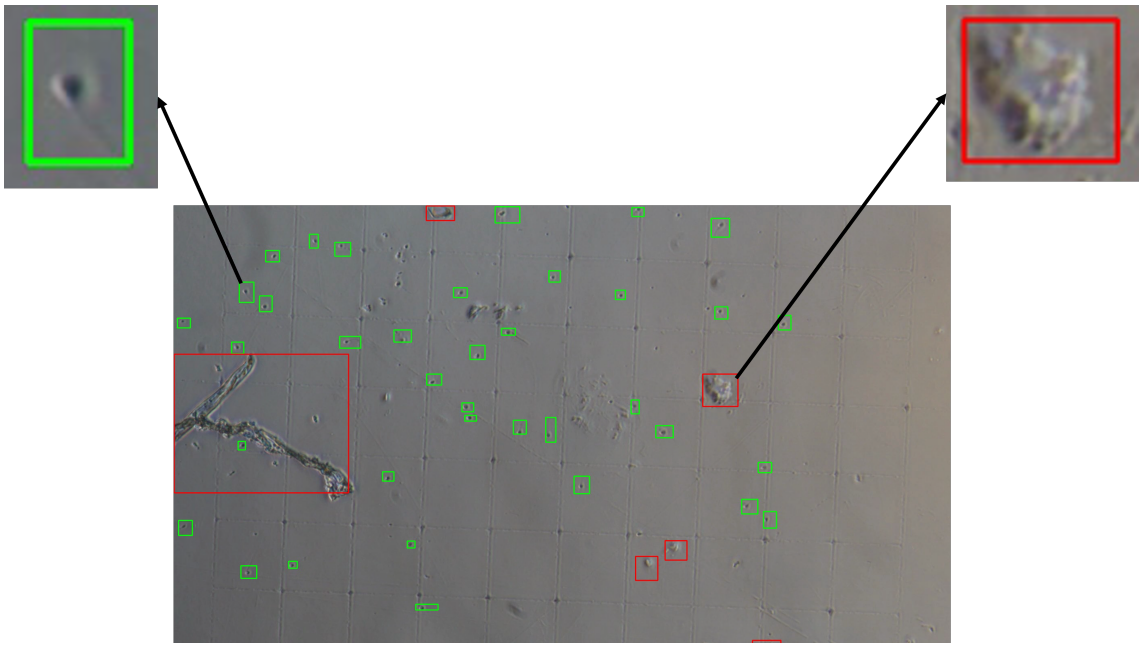
Şekil 4.2 Veri setleri için örnek görüntüler

Tablo 4.2 SMIDS, HuSHeM ve SCIAN-Morpho veri setlerinin sınıflara göre dağılımları

SMIDS		HuSHeM		SCIAN-Morpho	
Sınıf	Görüntü Sayısı	Sınıf	Görüntü Sayısı	Sınıf	Görüntü Sayısı
Normal	1021	Normal	54	Normal	100
Abnormal	1005	Tapered	53	Tapered	228
Non-Sperm	974	Pyriform	57	Pyriform	76
		Amorphous	52	Small	72
				Amorphous	656
Toplam	3000	Toplam	216	Toplam	1132

4.2 Hareket Analizinde Kullanılan Veri Setleri

Tez kapsamında sperm hareketi analiz çalışmaları için kullanılacak veri seti İstanbul Üniversitesinden elde edilen 250 kişiye ait 1000 videodan oluşmaktadır. Videolar $1920x1200$ çözünürlüğe sahip ve 30 saniye uzunluğundadır. Daha sonra her videodaki bölgeler uzmanlar tarafından sperm ve sperm olmayan bölgeler olarak etiketlenerek veri seti oluşturulmuştur. Hastalardan 30 fps ile elde edilen videoların her biri 5 fps olacak şekilde kaydedilmiştir. Yani her videonun her bir saniyesinde beş kare olacak şekilde ayarlama yapılmıştır. Ayrıca videolar $640x640$ çözünürlükte olacak şekilde düzenlenmiştir. Çalışmada kullanılan veri setinden örnek bir gösterim Şekil 4.3'te sunulmuştur. Şekilde yeşil kutucuklar spermleri, kırmızı kutucuklar ise sperm olmayan nesneleri ifade edecek şekilde etiketlenmiştir.

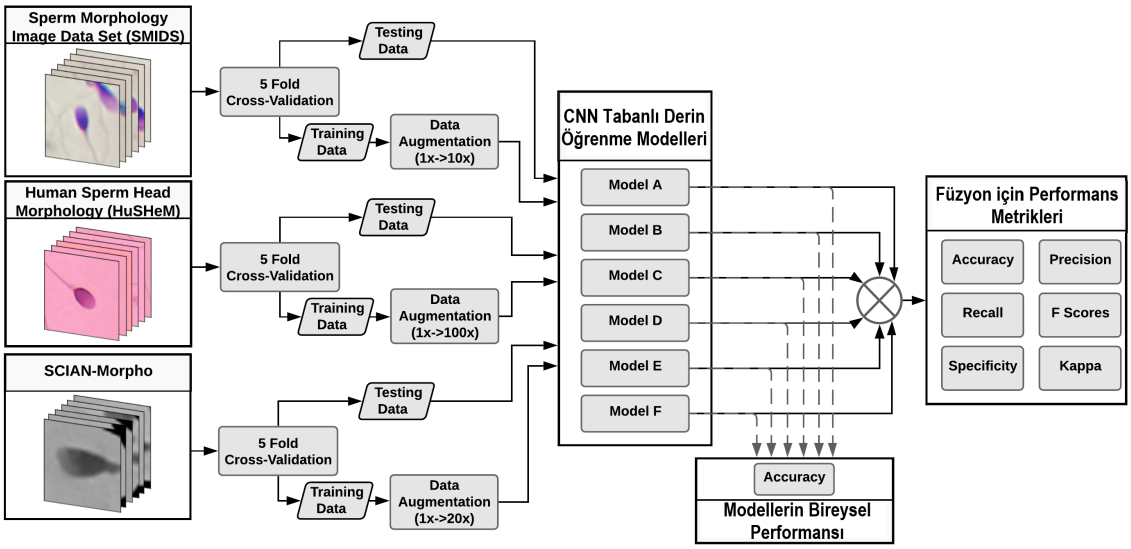


Şekil 4.3 Video veri setleri için örnek görüntüler

5
CNN MODELİ OLUŞTURMA ve UYGULAMA

Bu bölümde, tez çalışması kapsamında oluşturulan özgün altı CNN modeli incelenmiştir. Modellerin eğitimi ve testi için SMIDS, HuSHeM ve SCIAN veri setleri kullanılmıştır. Bu veri setleri farklı veri artırma yöntemleri ile farklı oranlarda artırılmıştır. Önerilen 6 farklı model veri artırma yöntemleri ile oluşturulan farklı oranlardaki yeni veri setleri ile eğitilmiş ve test edilmiştir. 6 farklı modelin sonuçlarına füzyon yöntemleri uygulanmıştır. Veri setlerinde sınıflandırma yapılırken herhangi bir ön işleme uygulanmamıştır.

Bu bölüm veri artırma, *k*-kat çapraz doğrulama, CNN modellerinin tasarımı, füzyon yöntemleri ve deneysel sonuçlar olmak üzere Şekil 5.1'de gösterildiği gibi organize edilmiştir.



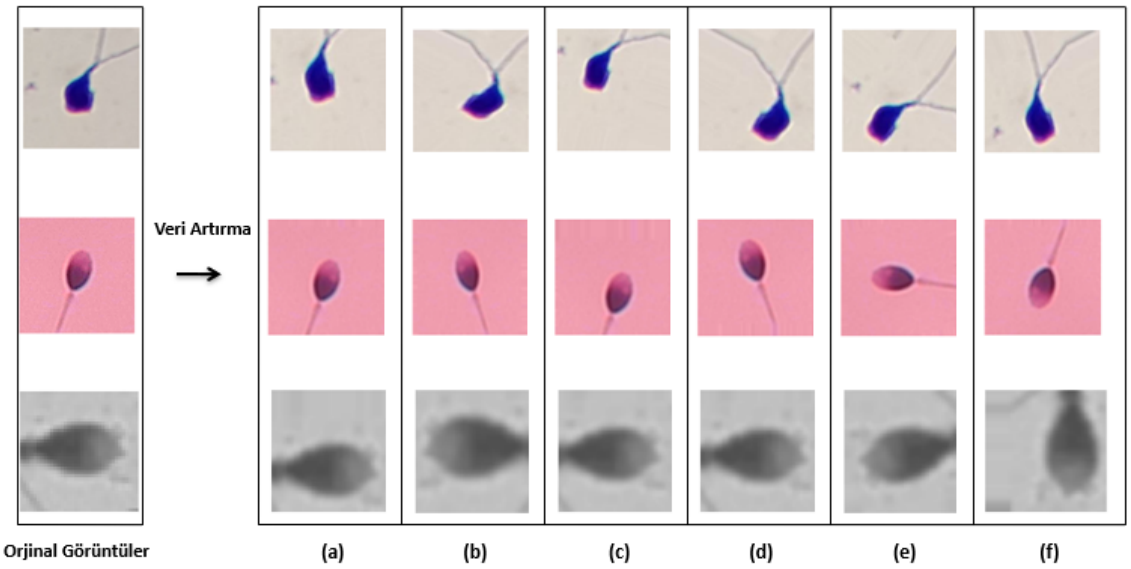
Şekil 5.1 Önerilen çalışmanın genel akışı

5.1 Veri Artırma

Derin öğrenme modellerinin eğitimi ve parametre ayarı için birçok veri örneğine ihtiyaç vardır. Eğitim için yeterli veri olmaması durumunda, veri öğelerinin

sayısı yapay olarak artırılarak modellerin performansı iyileştirilebilir. Tez kapsamında kullandığımız veri setlerindeki görüntü sayısı derin öğrenme tabanlı sınıflandırmada yetersiz kaldığı için aynı görüntü üzerinde farklı yöntemler uygulanarak eğitim görüntülerinin sayısının yapay olarak arttırılması hedeflenmiştir. Böylece modellerin eğitim aşamasında farklı şekilde oluşturulan yeni veriler kullanılmıştır. Artırılmış veriler farklı [136] olduğundan ezberleme etkisi en aza indirgenmiştir. Artırma yöntemleri, görüntünün boyutunu büyütme/küçültme, görüntüyü döndürme/çevirme, kontrastı değiştirme, gürültü ekleme/çıkarma, bulanıklık ekleme/azaltma gibi basit işlemlerdir. Görüntüden elde edilen özgün bilgiyi bozmamak için veri artırmada kullanılan işlem türü, görüntülerin karakteristik özelliklerine göre düşünülmelidir. Sperm görüntüleri için, önceden tanımlanmış küçük aralıklarda yalnızca yakınlaştırma (zooming), kesme (shearing), kaydırma (shifting) ve çevirme (flipping) gibi dört ana işlem uygulanmıştır. Daha önceki çalışmalarda bulanıklaştırma ve gürültü ekleme gibi diğer işlem türleri test edilmiştir. Ancak bu yöntemlerin sınıflandırma performansına fazla bir katkısı olmamıştır [156–159].

Üç veri seti için uygulanan veri artırma yöntemlerinin nasıl yapıldığı Şekil 5.2'de sunulmaktadır. Bu işlemler (a) dikey kaydırma + kesme, (b) yatay çevirme + yakınlaştırma, (c) yatay kaydırma + dikey kaydırma, (d) yatay çevirme + dikey kaydırma, (e) döndürme ve (f) dikey kaydırmadır. Görüntüleri bozmamak için, yakınlaştırma, kesme ve kaydırma işlemleri, üç veri setinin tamamı için orijinal görüntü boyutlarının SMIDS, HuSHeM ve SCIAN için sırasıyla %10, %5 ve %15 ile sınırlandırılmıştır.



Şekil 5.2 Altı farklı veri artırma tekniği

5.2 *k* - Çapraz Doğrulama

Çapraz doğrulama, bir modelin gerçek test verileriyle nasıl sonuç üreteceğini tahmin etmek için kullanılan istatistiksel bir doğrulama yöntemidir. Bu yöntemde, veri seti *k* adet bölüme ayrılır. Ardından, test için yalnızca bir bölüm kullanılırken, kalan *k-1* bölümleri eğitim için kullanılır. Bu işlem, her bölümün test seti olarak kullanılması tamamlanana kadar tekrarlanır. Sonunda, sonuçların ortalaması alınır. Bu yöntem, bir veri setindeki tüm görüntülerin hem eğitim hem de test için kullanılmasını ve her görüntünün tam olarak bir kez test için kullanılmasını sağlar [160, 161]. Tez çalışması kapsamında önerilen sınıflandırma yaklaşımının doğrulanmasında 5 katlı çapraz doğrulama şeması kullanılmıştır.

5.3 CNN Modellerinin Tasarımı

Tez kapsamında isimleri Model A, B, C, D, E ve F olarak kısaltılan altı farklı CNN oluşturulmuştur. Bu modeller sırasıyla 21, 24, 46, 28, 44 ve 32 katmandan oluşmaktadırlar. Modellerde kullanılan katmanların ayrıntılı bir gösterimi Şekil 5.3'teki blok diyagramda sunulmuştur.

Model A, $224x224x3$ görüntü boyutuna sahip giriş katmanından oluşmaktadır. Sonrasında Blok 1 ve 2 sırayla iki ve üç kez kullanılmaktadır. Blok 1, evrişimsel, toplu normalleştirme, ReLU ve maksimum havuzlama katmanlarını içermektedir. Blok 2'de ise Blok 1'den farklı olarak maksimum havuzlama katmanı kullanılmıştır. Blok 1 ve 2'nin ilk evrişim katmanlarında, filtre boyutu $7x7$ olarak ayarlanmıştır. Sonraki evrişim katmanında, filtre boyutu $3x3$'a düşürülmüştür. Sırasıyla birinci ve ikinci evrişim katmanlarında herhangi bir doldurma olmadan 64 ve 192 adet filtre kullanılmıştır.

Model B başlangıç katmanındaki giriş görüntüsü $227x227x3$ boyutundadır. Daha sonra Blok 3 ve 4 arka arkaya dört kez kullanılmıştır. Blok 3, evrişimsel ve ReLU katmanlarından oluşurken, Blok 4 Blok 3'e ek olarak maksimum havuzlama katmanını içerir. Model A ve B arasındaki farklardan biri, toplu normalleştirme katmanının Model B'de bulunmamasıdır. Model B, sınıflandırma problemlerinde bu katmanın etkisini karşılaştırmak için bir toplu normalleştirme katmanı olmadan oluşturulmuştur. Filtre boyutları Blok 3 ve 4 için ilk evrişim katmanlarında $11x11$ olarak düzenlenmiştir. Ardından, bir sonraki katmanda $3x3$ filtreler kullanılmıştır. Toplamda, sırasıyla birinci ve ikinci evrişim katmanlarında herhangi bir doldurma olmadan 96 ve 512 adet filtre kullanılmıştır.

C Modeli, diğer modellerden daha farklı ve daha karmaşık bloklar içermektedir.

Ayrıca bazı katmanlar paralel olarak uygulanmaktadır. Model A'ya benzer şekilde $224x224x3$ boyutunda girişleri kabul etmektedir. Bu modelde farklı olarak Blok 5 içinde çapraz kanal normalleştirme kullanılmıştır. Tam bağlı katmandan önce özellik vektörünün boyutunu azaltmak için Blok 7'de tanımlanan bir seyreltme katmanı kullanılmıştır. İlk evrişim katmanındaki filtreler, $11x11$ olarak belirlenmiştir. Daha sonra, diğer modellerde olduğu gibi kademeli olarak azaltılmıştır. Son katmanda $1x1$ filtreler uygulanmıştır. Her evrişim katmanında kullanılan toplam filtre sayısı 96'dan 128'e yükseltilmiştir. Model C, diğer modellerden daha derin ve karmaşık olarak tasarlanmıştır.

D Modeli, $150x150x3$ boyutunda girdileri kabul eden bir modeldir. Aslında D Modeli A Modelinin daha fazla katman içeren farklı bir versiyonu olarak tasarlanmıştır. Blok 1, art arda altı kez kullanılmıştır. Evrişim katmanlarında, $3x3$ filtre boyutu kullanılmıştır. Her evrişim katmanında kullanılan filtrelerin sayısı yinelemeli olarak 8'den 128'e çıkarılmıştır.

E Modeli D Modeli'ne benzer şekilde $150x150x3$ çözünürlükteki görüntüleri giriş olarak kabul eden bir modeldir. Bu modelde; Blok 1, 10 kez kullanılmıştır. Model E Model D'den daha derin olarak tasarlanmıştır. Model E'de, ilk ve son evrişim katmanlarında $3x3$ filtre boyutu kullanılmıştır. Orta katmanlarda $1x1$, $3x3$ ve $7x7$ gibi çeşitli filtre boyutları kullanılmıştır. İlk katmanda filtre sayısı 32 ile başlamıştır. Sonraki katmanlarda filtre sayısı önce azaltılmış sonra artırılmıştır. Son katmanda toplam 88 filtre kullanılmıştır. Ayrıca farklı dolgu konfigürasyonları da gerçekleştirilmiştir.

Son olarak tasarlanan F Modeli, Model D ve Model E'ye benzer şekilde $150x150x3$ girişleri kabul eden bir ağdır. Bu modelin derinliği Model E'den daha az olarak tasarlanmıştır. Model F'de Blok 1 yedi kez arka arkaya kullanılmıştır. Katmanlarda kullanılan filtre pencere boyutları Model E ile aynı olmakla birlikte filtre sayısı önce 48 adet ile başlamış, hemen sonrasında azaltılmış ve sonlara doğru 84 filtreye kadar artırılmıştır.

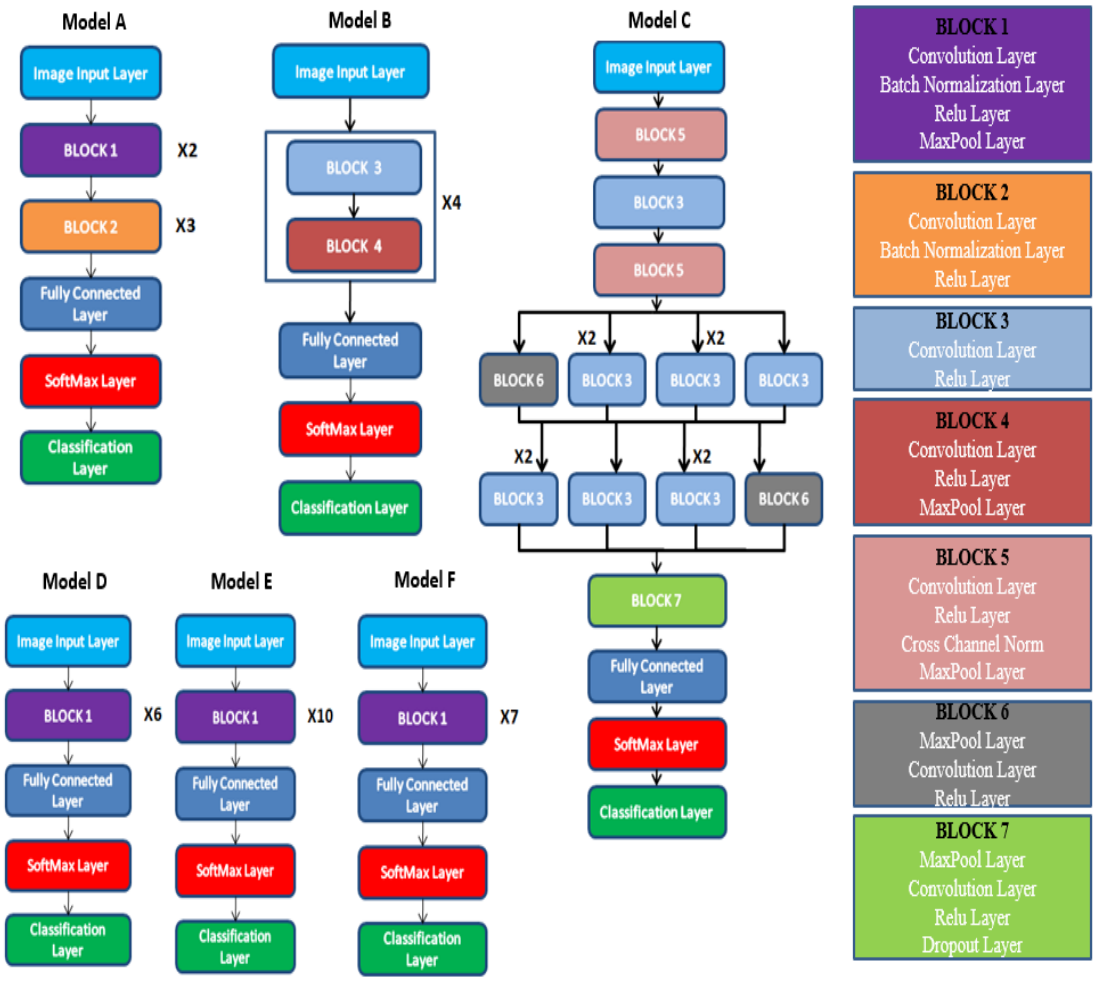
5.4 Modeller için Eğitim Parametrelerinin Belirlenmesi

Oluşturulan altı model, SMIDS, HuSHeM ve SCIAN veri setleri için ayrı ayrı sıfırdan eğitilmiştir. Her bir modelin giriş boyutunu $N \times M \times 3$ temsil etmektedir. SCIAN-Morpho veri seti tek kanallı olan gri tonlamalı görüntülerden oluştuğu için her görüntü yeni bir boyutta üç kez tekrarlanmıştır. Böylece gri görüntüler önerilen ağlarda kullanılabilecek üç kanallı görüntülere dönüştürülmüştür. Optimize edici, öğrenme hızı, mini batch boyutu ve farklı epoch boyutlarındaki çeşitli parametreler test edildi.

Tablo 5.1 Veri setlerinin eğitimi için kullanılan parametreler

Parametreler	SMIDS	HuSHeM	SCIAN-Morpho
Optimizer	sgdm	sgdm	adam
Epoch	50	50	20
Mini Batch Size	8	8	8
Initial Learning Rate	1e-4	1e-4	1e-4

Öğrenme hızı parametreleri için 10^{-2} ile 10^{-5} arasındaki değerler test edilmiştir. Epoch değerleri ise 20'den 70'e kadar 10'ar artırılarak test edilmiştir. Optimize edici olarak *sgdm* ve *adam* gibi iki önemli optimizasyon yöntemi kullanılmıştır. Her bir veri setinin sınıflandırılması en iyi sonuçlara sahip parametrelerin detayları Tablo 5.1'de gösterilmiştir.



Şekil 5.3 Sıfırdan oluşturulan altı modelin ayrıntılı gösterimi

Tablo 5.1'de gösterildiği gibi SMIDS ve HuSHeM veri setleri için en etkili sınıflandırma sonuçları *sgdm* ile sağlanırken, SCIAN-Morpho veri seti için *adam* optimizasyonu kullanılmıştır. SMIDS ve HuSHeM veri setleri için Epoch boyutu 50 ile en iyi sonuçlara ulaşırken, SCIAN-Morpho için ise 20 ile en iyi sonuçlara ulaşılmıştır. Mini batch analizi

açısından, 8 başlangıç değeri olarak tanımlanmıştır. Bu parametrelere dayalı olarak en etkili artırma ölçeği belirlendikten sonra farklı mini batch boyutu (4, 16 ve 32) analizleri yapılmıştır.

5.5 Füzyon Yöntemleri

Oluşturulan modeller farklı ağ mimarilerine sahip oldukları için aynı görüntü için farklı sınıf tahminlerinde bulunabilirler. Modellerden elde edilen tahmin kombinasyonları, sınıflandırma performansını artırabilir. Tahminlerin kombinasyonları karar düzeyi birleştirme tekniği (decision-level fusion) olarak adlandırılır [162, 163]. Tez çalışmasında üç veri seti için altı modelin soft-voting ve hard-voting olmak üzere iki füzyon yaklaşımı test edilmiştir. Soft-voting yöntemi kullanılan her model için Softmax katmanındaki çıktılardan elde edilir. Her bir tahminin olasılıklarının birleştirilmesini ve en yüksek toplam olasılığa sahip tahminin seçilmesini gerektirir [164]. Soft-voting yöntemi yanlış sınıflandırmış bireysel ağların yanlış sınıflandırma etkisini azaltmak için kullanılır. Böylece çelişkili sınıflandırmalar olması durumunda, olasılığa dayalı füzyon şeması ile daha doğru bir şekilde sınıflandırma yapılabilir.

Bir başka füzyon yaklaşımı olan hard-voting, çoğunluk kuralı oylama şeması [165] ile tahmin edilen sınıf etiketlerini kullanır. Bu yöntem ile diğer tahminlerin olasılıkları göz ardı edilir ve en yüksek oy sayısına sahip tahminin seçilmesi sağlanır.

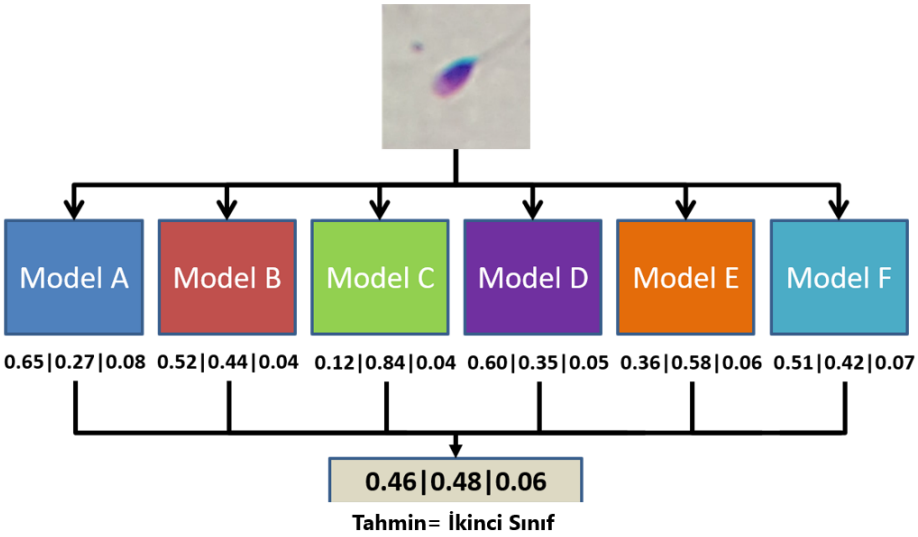
Örnek bir görüntüye uygulanan soft ve hard voting, sırasıyla Şekil 5.4 ve Şekil 5.5'te gösterilmektedir. Bu şekillerden de anlaşılacağı üzere füzyon yaklaşımında kullanılan oylama şemasına göre farklı sınıf atamaları elde edilebilmektedir.

5.6 Performans Metrikleri

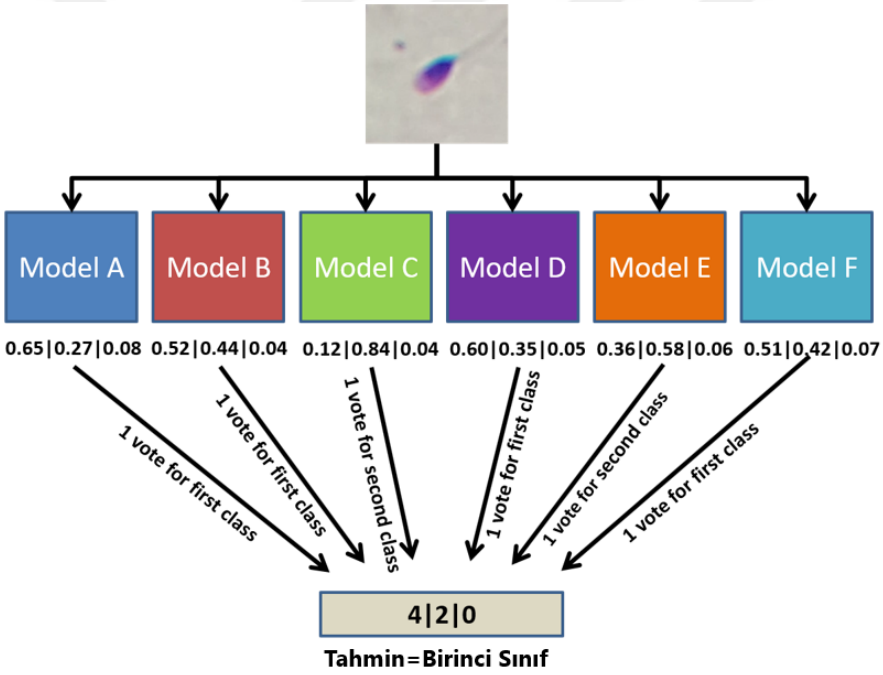
Oluşturulan altı CNN modelinin her bir veri seti için doğruluk değerleri ayrı ayrı ölçülmüştür. Daha sonra doğruluk skorlarına ek olarak altı modelin füzyon sonuçlarının kesinlik, duyarlılık, özgüllük, F-skor ve Kappa skorları hesaplanmıştır. Karışıklık matrisinden bu metriklerin hesaplanması Eşitlik 3.14 - Eşitlik 3.19 ile verilmiştir.

5.7 Deneysel Sonuçlar

Modellerin oluşturulması ve eğitilmesi aşamasında MATLAB 2020a [166] kullanılmıştır. Deneylerde kullanılan tüm veri setlerine Şekil 5.2'de gösterildiği



Şekil 5.4 Örnek Soft-Voting yöntemi



Şekil 5.5 Örnek Hard-Voting yöntemi

gibi farklı veri artırma yöntemleri uygulanmıştır. SMIDS, 3, 5, 8 ve 10 kat, HuSHeM veri setindeki görüntüler ise diğer veri setlerine göre en az örneğe sahip olduğu için 10, 20, 40 ve 100 kat artırılmıştır. SCIAN-Morpho veri seti ise dengesiz dağılıma sahip olduğu için önce dengeli hale getirilmiştir. Dengeleme adımında Amorphous sınıfı ile sayıları eşitlemek için Normal, Tapered, Pyriform ve Small olarak etiketlenen sınıf görüntüleri üzerinde artırma yöntemleri uygulanmıştır. Bu işlem öncesinde SCIAN-Morpho veri setinin sınıf dağılımı Tablo 4.2'de gösterildiği gibi toplam 1132 sperm görüntüsünden oluşmaktadır. Dengeleme işleminden sonra veri seti 580 Normal, 594 Tapered, 564 Pyriform, 536 Small ve 656 Amorphous dağılımına sahip

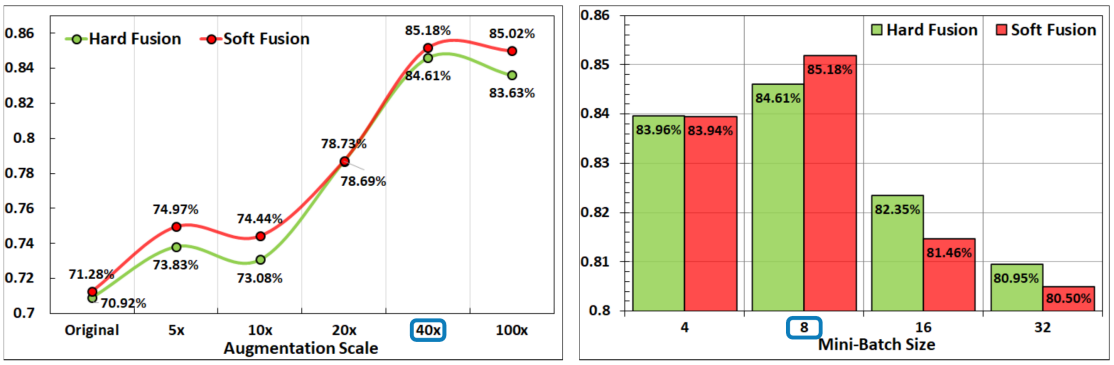
Tablo 5.2 Artırılmış veri setlerinin bireysel ve füzyon sonuçları

Sınıflandırma Doğruluğu (%)			Modellerin Bireysel Performansları						Modellerin Füzyonu	
			A	B	C	D	E	F	Hard Füzyon	Soft Füzyon
HuSHeM		Orijinal	67,39%	65,73%	57,80%	61,67%	60,56%	69,97%	70,92%	71,28%
	Artırma	5x	69,31%	68,09%	68,41%	65,12%	67,99%	73,68%	73,83%	74,97%
		10x	68,61%	69,65%	69,43%	66,92%	69,80%	72,05%	73,08%	74,44%
		20x	68,56%	72,14%	72,28%	74,36%	73,91%	77,15%	78,69%	78,73%
		40x	76,31%	80,04%	76,33%	80,31%	81,76%	80,30%	84,61%	**85,18%**
		100x	78,61%	75,47%	77,66%	78,60%	80,77%	83,04%	83,63%	85,02%
SMIDS		Orijinal	82,17%	84,20%	84,10%	82,80%	82,60%	83,83%	85,73%	86,23%
	Artırma	3x	84,00%	86,03%	87,07%	84,63%	84,97%	85,07%	87,83%	88,13%
		5x	85,07%	87,00%	87,00%	85,50%	84,80%	85,70%	88,53%	88,60%
		8x	87,50%	89,23%	88,53%	88,03%	87,57%	87,67%	89,80%	**90,20%**
		10x	85,93%	87,20%	86,80%	86,43%	85,23%	85,70%	88,77%	89,30%
SCIAN		Orijinal	58,76%	62,39%	60,70%	62,37%	60,45%	60,06%	65,65%	66,45%
	Artırma	Dengeli	59,11%	64,05%	61,68%	61,66%	62,29%	61,66%	65,30%	66,09%
		2xB	54,32%	62,11%	62,12%	60,98%	61,84%	59,39%	65,83%	68,12%
		5xB	60,16%	63,08%	63,87%	65,29%	64,84%	63,06%	68,73%	69,44%
		10xB	64,57%	66,17%	64,59%	65,72%	64,40%	66,95%	71,02%	**71,82%**
		20xB	66,17%	67,14%	64,84%	64,32%	67,84%	67,05%	70,84%	71,38%

olmuştur. Daha sonra dengeli veri seti 2, 5, 10 ve 20 katı kadar artırılmıştır. Ayrıca orijinal veri setleri başlangıçta test ve eğitim setlerine bölünmüştür. Veri artırma yöntemleri sadece eğitim setindeki görüntülere uygulanmıştır.

Modellerin sınıflandırma başarısını objektif olarak karşılaştırmak için, veri setlerinin her bir artırılmış versiyonu, her bir artırma işleminden sonra ayrı ayrı kaydedilmiştir. Daha sonra, artırılmış her bir veri seti üzerinden her bir modelin eğitimi gerçekleştirilmiştir. Eğitim gerçekleşirken tüm modeller için Tablo 5.1'de listelenen parametre değerleri kullanılmıştır. Modellerin sınıflandırma tahminleri hard ve soft voting yöntemleri kullanılarak birleştirilmiştir. Daha sonra optimal sınıflandırma performansı için mini-batch analizleri yapılmıştır.

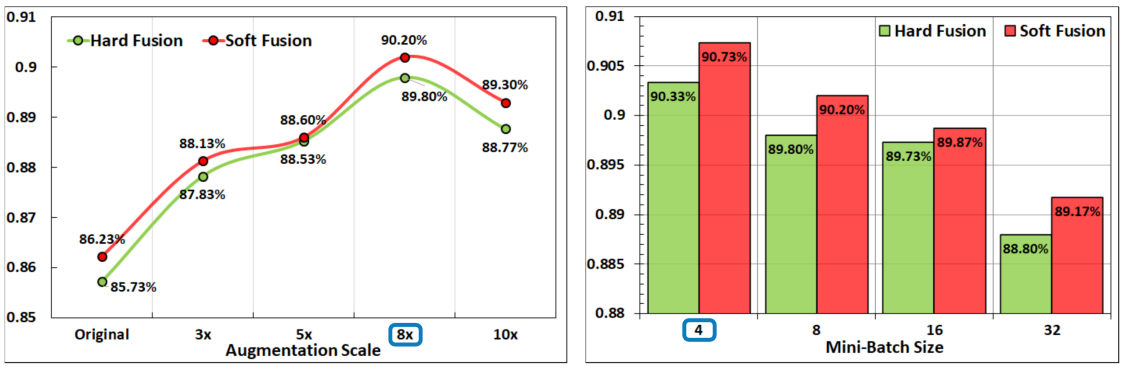
Altı modelin orijinal ve artırılmış veri setleri üzerindeki bireysel ve füzyon sınıflandırma performansları, Tablo 5.2'de sunulmaktadır. Tanımlanan en iyi artırma ölçeği üzerinden her bir veri seti için mini batch analizlerinin sonuçları Şekil 5.6, Şekil 5.7 ve Şekil 5.8'de ayrı ayrı olarak sunulmuştur.



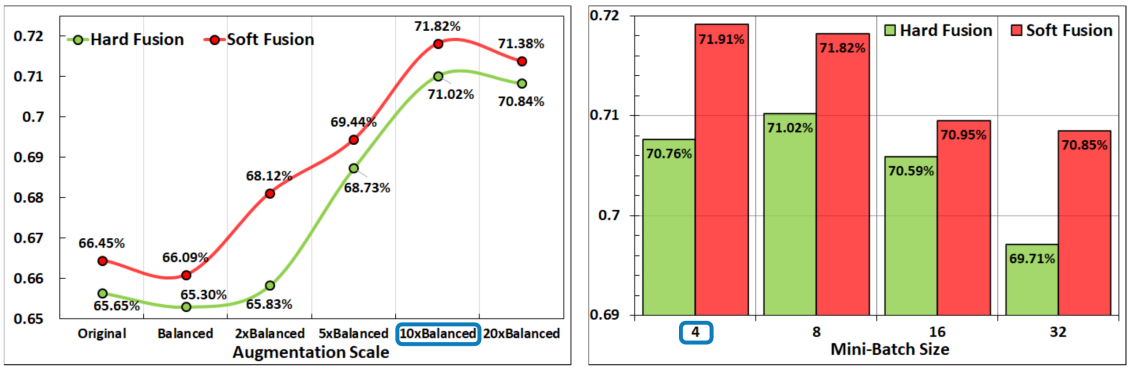
Şekil 5.6 HuSHeM için veri artırma ve mini batch analizi

Tablo 5.2'ye göre, orijinal HuSHeM veri setinin sınıflandırılmasında en iyi bireysel model performansı %69,97 doğrulukla F Modeli ile elde edilmiştir. HuSHeM veri setinin orijinal versiyonu için soft voting tekniği, sınıflandırma doğruluğunu %71,28'e yükseltmiştir. Artırılmış veri ölçeği analizinde, 100 kat artırılmış veri seti kullanıldığında Model F ile bireysel ağlar için %83,04'lik optimal sınıflandırma doğruluğu elde edilmiştir. Ancak füzyon yaklaşımında en yüksek sınıflandırma performansı, 40 kat artırılmış veri seti için %85,18 doğrulukla soft voting şeması tabanlı füzyon tekniği ile elde edilmiştir. Artırma ölçeğine göre soft ve hard voting yöntemleri açısından ağ topluluklarının sınıflandırma sonuçları, Şekil 5.6'nın sol kısmında grafiksel olarak sunulmuştur. İkinci aşama analizi olarak, en yüksek sınıflandırma performansıyla sonuçlanan 40 kat artırılmış veri seti üzerinde mini batch analizi yapılmıştır. Sınıflandırma performansını artırmak için özel olarak 4, 8, 16 ve 32 mini batch boyutlarıyla deneyler yapılmıştır. Sonuçlar, Şekil 5.6'nın sağ tarafında sunulmaktadır. %85,18'lik en yüksek performans, soft voting füzyon tekniği kullanılarak 8 mini batch boyutuyla elde edilmiştir.

SMIDS orijinal veri seti için yapılan sınıflandırma analizinde, bireysel model performansı açısından Model B kullanılarak %84,20 doğrulukla optimal performans elde edilmiştir. Orijinal SMIDS için füzyon sonuçları incelendiğinde, soft-voting



Şekil 5.7 SMIDS için veri artırma ve mini batch analizi

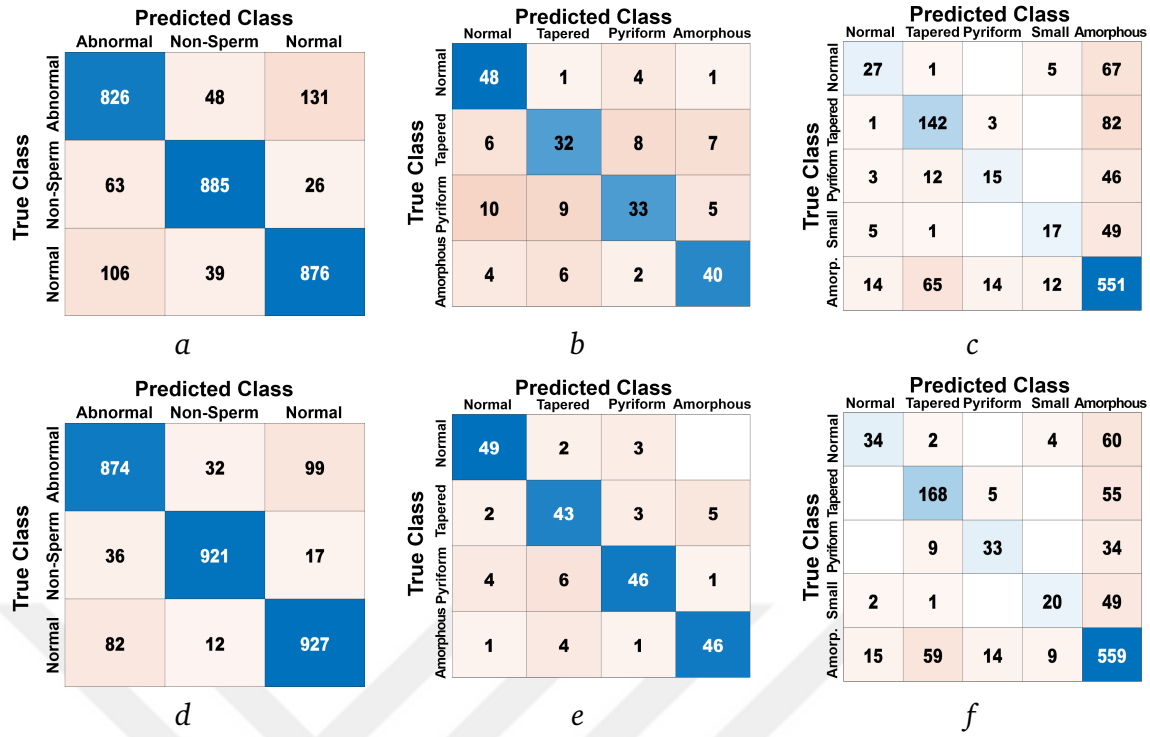


Şekil 5.8 SCIAN için veri artırma ve mini batch analizi

tabanlı füzyon yöntemi kullanılarak %86,23 doğrulukla optimal sonuç elde edilmiştir. Sınıflandırmada artırılmış veri setleri kullanıldığında, bireysel olarak en yüksek doğruluk Model B ile 8 kat artırılmış veri setinde %89,23 olarak elde edilmiştir. Aynı zamanda aynı veri setinde altı modelin soft voting füzyon yöntemi ile %90,20'lik doğruluğa ulaşılmıştır. Ağ füzyonu tabanlı sınıflandırma sonuçlarında artırma ölçeğinin etkileri Şekil 5.7'de sunulmaktadır. Mini batch analizleri, önceden tanımlanmış parametreler için en yüksek sınıflandırma performanslarıyla sonuçlanan 8 kez artırılmış veri seti kullanılarak gerçekleştirilmiştir. 4'lük bir mini batch boyutu kullanılarak %90,20 doğruluk oranı %90,73'e yükseltilmiş ve Şekil 5.7'nin sağ kısmında görülebileceği gibi bu başarıya soft füzyon yöntemiyle ulaşılmıştır.

Son olarak, orijinal SCIAN-Morpho veri seti için bireysel ağların optimal sonucu, Model B ile %62,39 doğruluk oranı olarak elde edilmiştir. CNN modellerinin füzyonu sayesinde doğruluk oranı %66,45'e ulaşmıştır. Orijinal SCIAN-Morpho veri seti, her sınıfta eşit olmayan örneklem büyüklüklerine sahip dengesiz bir veri setidir. Daha objektif sonuçlar için öncelikle dengeli bir veri seti oluşturulmalı, ardından performans değerlendirilmelidir. Bunu yapmak için dengeli bir SCIAN-Morpho veri seti oluşturulmuş ve yine Model B, bireysel ağ performansı açısından %64,05'lik en yüksek sınıflandırma doğruluğu elde edilmiştir. Soft voting yöntemiyle ile ağların birleştirilmesi, aynı dengeli veri seti üzerinde %66,09'luk bir doğruluk oranıyla sonuçlanmıştır. Bireysel ağlar için optimal sınıflandırma doğruluğu, Model E ile 20 kat artırılmış dengeli veri seti üzerinden %67,84 doğrulukla elde edilmiştir. Füzyon yaklaşımında, 10 kat artırılmış dengeli veri seti için bu doğruluk soft füzyon ile %71,82'ye yükseltilmiştir. Artırma ölçeklerine göre füzyon yöntemlerinin sınıflandırma performansı, Şekil 5.8'in sol kısmında sunulmaktadır. SCIAN-Morpho veri seti için en yüksek sınıflandırma doğruluğu 10 kat artırılmış dengeli veri setinde elde edildiğinden mini-batch analizi bu veri seti üzerinden gerçekleştirilmiştir. Mini-batch analizine göre, Şekil 5.8'de gösterildiği gibi, 4'lük bir mini-batch boyutu kullanılarak %71,82 doğruluk oranı %71,91'e yükseltilmiştir.

Modellerin üç sperm morfolojisi veri seti üzerindeki sınıflandırma performansına ek olarak, bireysel ağların eğitim süreleri değerlendirilmiş ve Tablo 5.3'te listelenmiştir. Sadece orijinal veri setlerini kullanan ağların eğitim süreleri ölçülmüştür. Eğitimdeki en hızlı model, eğitim aşamasını HuSHeM için 1 saatte, SCIAN-Morpho için 4 saatte ve en büyük veri seti SMIDS için 11 saatte tamamlayan Model A olmuştur. En uzun eğitim süreleri, HuSHeM, SMIDS ve SCIAN-Morpho veri setleri için sırasıyla 2,1 saat, 23,1 saat ve 8,4 saat ile Model C için ölçülmüştür. Ancak, en uzun eğitim süreleri, en yüksek sınıflandırma doğruluğunu ifade etmez. Bireysel bir ağ olarak Model C, herhangi bir veri seti için en yüksek sınıflandırma performansıyla sonuçlanmamıştır. Bunun yerine Model F, diğer ağlarla karşılaştırıldığında ortalama eğitim süresiyle HuSHeM veri seti



Şekil 5.9 3 veri seti için karışıklık matrisi a)SMIDS (Orijinal), b)HuSHeM (Orijinal), c)SCIAN-Morpho (Orijinal), d)SMIDS (Artırılmış), e)HuSHeM (Artırılmış), f)SCIAN-Morpho (Artırılmış)

için en yüksek doğruluk oranıyla sonuçlanmıştır. SMIDS ve SCIAN-Morpho veri setleri için Model B en yüksek performansı vermiştir. Modellerin eğitim hızı; modelin boyutu, kullanılan filtre sayısı, giriş görüntüsünün boyutu ve giriş görüntüsünün kalitesi gibi birçok faktöre göre değişiklik göstermiştir. Bu nedenle, optimal sınıflandırma doğruluğu için model seçimi, ağ tasarımının yanı sıra eğitim verilerine de bağlıdır.

Tablo 5.3 Modellerin eğitim süreleri (saat)

Veri Setleri	A	B	C	D	E	F
HuSHeM	1	1,3	2,1	1,4	2	1,5
SMIDS	11	14,3	23,1	15,4	22	16,5
SCIAN-Morpho	4	5,2	8,4	5,6	8	6

Şekil 5.9'da orijinal görüntüler ve artırılmış görüntüler için modellerin soft voting tabanlı füzyon tekniği kullanılarak elde edilen optimal sınıflandırma performansının karışıklık matrisleri verilmiştir. Doğruluk, kesinlik, duyarlılık, özgüllük, F-Skoru ve Kappa gibi birçok performans metriği bu karışıklık matrislerinden türetilmiştir ve Tablo 5.4 içinde ifade edilmiştir. Amaç, sınıf odaklı sınıflandırma performansları açısından artırma yöntemlerinin katkısını ortaya koymaktır.

Tablo 5.4'te, kesinlik değerleri, sınıf tabanlı tahminler açısından modellerin

sağlamlığına atıfta bulunur. Artırılmış veya orijinal veri setlerini kullanan yaklaşım sayesinde, SMIDS için Non-Sperm sınıfında en yüksek kesinlik değerine ulaşılmıştır. Artırma yöntemleri, SMIDS'deki her sınıf için yaklaşık ∼4% kesinlik puanı artışları sağlamıştır. HuSHeM veri seti için en zorlu sınıf, görüntülerin çoğunlukla Pyriform görüntülerine benzer olan Tapered sınıfıydı. En kusurlu sperm şekillerini içeren HuSHeM veri setinde Amorphous sınıfı için en yüksek kesinlik değeri elde edilmiştir. Veri artırma etkisi en çok Normal ve Pyriform sınıfları için gözlenmiştir (yaklaşık ∼15%). SCIAN-Morpho veri setinde, Normal, Tapered ve Amorphous sınıfları için en yüksek puanlar alınırken, Pyriform ve Small sınıf tabanlı görüntüler, düşük çözünürlüklü görüntüler olmaları nedeniyle modeller için tam olarak bilgilendirici değildir. Ancak, bu sınıflar üzerinde veri artırma yöntemlerinin kayda değer bir etkisi %17 ve %10 artışlarla gözlemlenmiştir.

Tablo 5.4'te görüldüğü gibi, SMIDS için Non-Sperm sınıfının duyarlılık değeri, herhangi bir sperm şeklinden en farklı şekil özelliklerine sahip olduğundan Normal ve normal olmayan sınıflardan daha yüksektir. İyi çözünürlüklü görüntülerden oluşan HuSHeM veri seti için Normal sınıf görüntüler için, duyarlılık değerleri daha bilgilendirici özelliklerden dolayı daha doğru tahmin edilmiştir. Veri artırma yöntemleri, çoğunlukla Tapered ve Pyriform sınıfları için duyarlılık puanlarını artırdı. Diğer sınıflardaki görüntülerden tamamen farklı şekillere sahip olduğu için, SCIAN-Morpho veri setindeki en yüksek duyarlılık değerine sahip sınıf Amorphous olmuştur. Ek olarak, veri artırma yöntemleri ile, Pyriform sınıfları için duyarlılık puanları açısından önemli bir artış sağlanmıştır.

Kesinlik ve duyarlılık değerleri birleştirilerek F-Skor ve Kappa skorları hesaplandı. F-Skoru, kesinlik ve duyarlılık metriklerinin harmonik ortalamasıdır. Böylece kesinlik ve duyarlılık arasındaki uyumsuzluğu ortadan kaldırır. Bu metriklere dayalı olarak SMIDS, HuSHeM ve SCIAN-Morpho veri setlerinin en belirgin sınıfları sırasıyla Non Sperm, Normal ve Amorphous sınıfları olarak belirlenmiştir. Benzer şekilde, veri artırma yöntemlerinin etkileri de bu puanlar üzerinden gözlemlenebilmektedir.

Tablo 5.5'te, önerilen yaklaşımın sonuçları, aynı veri setlerini ele alan literatürdeki diğer çalışmalarla karşılaştırılmıştır. Önerilen yaklaşım, herhangi bir ön işleme aşaması olmaksızın, iyi bilinen üç sperm morfometrisi veri setinde sperm morfolojisini otomatik olarak sınıflandırabilmektedir. SCIAN-Morpho, HuSHeM ve SMIDS veri setleri için sırasıyla tasarlanmış altı CNN modelinin önerilen soft voting tabanlı füzyon yaklaşımı kullanılarak %71,91, %85,18 ve %90,73 doğruluk elde edilmiştir. Tablo 5.5'te görüldüğü gibi, önerilen yaklaşım, SMIDS ve SCIAN-Morpho veri setleri için diğer çalışmalarla karşılaştırıldığında, daha yüksek sınıflandırma performansları elde edilmiştir. HuSHeM veri seti için sonuçları karşılaştırırken,

Tablo 5.4 Şekil5.9'dan türetilen performans ölçütleri ve sonuçları

Modellerin Füzyonlarının Ayrıntılı Sonuçları			SMIDS Veriseti			HuSHeM Veriseti				SCIAN-Morpho Veriseti				
			Normal	Abnormal	Non-Sperm	Normal	Tapered	Pyriform	Amorphous	Normal	Tapered	Pyriform	Small	Amorphous
	Orijinal	Precision	84,80	83,01	91,04	70,58	66,66	70,21	75,47	54	64,25	46,87	50	69,30
		Recall	85,79	82,18	90,86	88,88	60,37	57,89	76,92	27	62,28	19,73	23,61	83,99
		Specificity	92,06	91,52	95,70	87,65	90,18	91,19	92,07	97,77	91,26	98,39	98,39	48,73
		F-Score	85	83	91	79	63	63	76	36	63	28	32	76
		Accuracy		86,23				71,28				66,45		
		Kappa		0,793				0,611				0,385		
			SMIDS Veriseti			HuSHeM Veriseti				SCIAN-Morpho Veriseti				
			Normal	Abnormal	Non-Sperm	Normal	Tapered	Pyriform	Amorphous	Normal	Tapered	Pyriform	Small	Amorphous
	Artırılmış	Precision	88,88	88,11	95,44	87,50	78,18	86,79	88,46	66,67	70,29	63,46	60,61	73,84
		Recall	90,79	86,97	94,56	90,74	81,13	80,70	88,46	34	73,68	43,42	27,78	85,21
		Specificity	94,13	94,08	97,82	95,67	92,63	95,59	96,34	98,35	92,14	98,20	98,77	58,40
		F-Score	90	88	95	89	80	84	88	45	72	52	38	79
		Accuracy		90,73				85,18				71,91		
		Kappa		0,861				0,802				0,499		

önerilen yöntem daha düşük sınıflandırma doğruluğu ile sonuçlandırılmıştır. Ancak daha önceki çalışmalarda görüntüleri gürültüden arındırmak için kullanılmış olan insan faktörünü ortadan kaldırma avantajına sahiptir. Önerilen çok modelli füzyon tabanlı sınıflandırma yaklaşımı, görüntü ön işleme gibi herhangi bir manuel müdahale gerektirmez. Baştan sona tam otomatik bir sınıflandırma sağlar.

Diğer çalışmalarla karşılaştırıldığında, önerilen yaklaşımın bir diğer kritik avantajı, çok katmanlı modeller kullanmak yerine çok daha hızlı sonuç almak için özel düşük katmanlı modeller ile tasarlanması ve kullanılmasıdır.

Tablo 5.5 Önerilen yöntem ile literatürdeki benzer çalışmaların performansının karşılaştırılması

Veriseti Adı	Veriseti Bilgisi	Yöntem	Makale	Metot	Doğruluk (%)
SCIAN-Morpho	1132 Sperm 35x35 Çözünürlük Not Overlapped 5 sınıf	Already Oriented and Cropped	Chang et al. 2017 [42]	Feature Extraction by Fourier Descriptor SVM Classification	49
			Chang et al. 2017 [43]	Feature Extraction by Ensemble of Fourier, Zernike and Shape Descriptors Ensemble of SVM Classification	58
			Shaker et al. 2017 [47]	Color Space Converting Dictionary Learning	62
			Proposed Approach	10×Data Augmentation over Balanced Dataset Design of Six CNN-Based Deep Learning Models Decision Level Soft Fusion	71,91
HuSHeM	216 Sperm 131x131 Çözünürlük Overlapped Dört Sınıf	Manual	Shaker et al. 2017 [47]	Manually Sperm Head Rotation and cropping into 50 × 76 pixels Color Space Converting Dictionary Learning	92,2
		Automatic	Ilhan et al. 2019 [55]	Automatic Directional Masking *k*-NN classification	57,4
			Ilhan et al. 2020 [56]	Adaptive De-Noising Overlapping Group Shrinkage Image Gradient Directional Masking Feature Extraction by MSER descriptor SVM Classification (non-linear Kernel)	86,6
		Not Applied	Proposed Approach	40×Data Augmentation over Dataset Design of Six CNN-Based Deep Learning Models Decision Level Soft Fusion	85,18
SMIDS	536 Sperm 190x170 Çözünürlük Overlapped Üç Sınıf	Not Applied	Ilhan et al. 2018 [53]	Dual Tree Complex Wavelet Trans. SVM classification	82,3
		Not Applied	Ilhan et al. 2018 [54]	Adaptive De-Noising Feature Extraction by SURF Descriptor SVM classification	83,4
	3000 Sperm 190x170 Çözünürlük Overlapped Üç sınıf	Not Applied	Ilhan et al. 2020 [57]	Overlapping Group Shrinkage Fuzzy-C Means Clustering Data Augmentation Classification by MobileNet	87
		Automatic	Ilhan et al. 2020 [56]	Adaptive De-Noising Overlapping Group Shrinkage Image Gradient Directional Masking Feature Extraction by MSER descriptor SVM Classification (non-linear Kernel)	85,7
		Not Applied	Proposed Approach	8×Data Augmentation over Dataset Design of Six CNN-Based Deep Learning Models Decision Level Soft Fusion	90,73

6
TRANSFER ÖĞRENME KULLANILARAK SPERM GÖRÜNTÜLERİNİN SINIFLANDIRILMASI

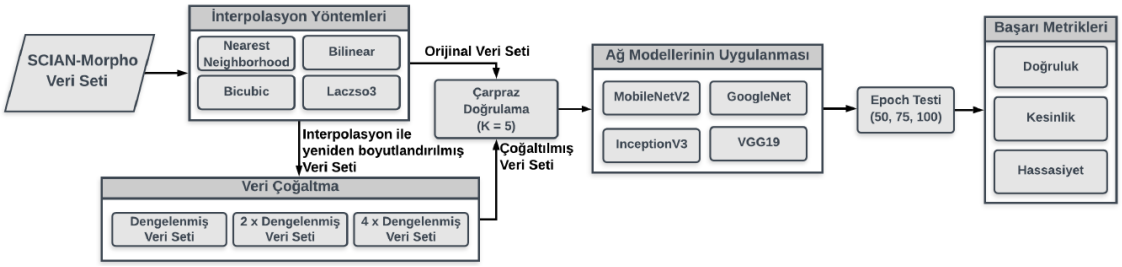
Biyomedikal görüntülerin sınıflandırılması derin öğrenme modelleri oluşturularak yapılabildiği gibi (Bölüm 5) daha önceden eğitilmiş bir modelin kullanıldığı transfer öğrenme yöntemiyle de yapılabilir. Bu bölümde sperm görüntülerinin sınıflandırılması için veri setlerine ön işlemler uygulanarak elde edilen yeni veri setleri transfer öğrenme ile sınıflandırılmıştır. Ön işleme olarak interpolasyon ve renk dönüşümleri gibi farklı yöntemler uygulanmıştır.

6.1 Düşük Kaliteli Sperm Görüntülerinin Morfolojik Olarak Sınıflandırılması

Düşük kalitede elde edilmiş sperm görüntülerinin bulunduğu SCIAN-Morpho veri setinin sınıflandırma başarımını artırmak amacıyla öncelikle interpolasyon yöntemleri uygulanarak veri kalitesi artırılmıştır. Böylece sperm görüntülerinin otomatik olarak düzeltilmesi ve uygun derin öğrenme ağlarını kullanarak morfolojik olarak sınıflandırılabilmesi hedeflenmiştir. Ayrıca veri artırma yöntemleriyle, veri dengesizliği ve veri eksikliği problemlerine çözüm üretilmeye çalışılmıştır. Sınıflandırma aşamasında MobilNetV2, GoogleNet, InceptionV3 ve VGG19 gibi CNN modelleri uygulanmıştır. Başarım metriği olarak doğruluk, kesinlik ve duyarlılık metrikleri dikkate alınmıştır. En yüksek sınıflandırma başarısını veren VGG19 modeli üzerinde geliştirmeler yapılarak başarımın arttırılması sağlanmıştır. Yapılan tüm işlemlerin genel akışı Şekil 6.1'de görülmektedir.

6.1.1 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



Şekil 6.1 Çalışmanın genel akışı

görüntü interpolasyonu olmak üzere iki guruba ayrılır. Adaptif olmayan interpolasyon yöntemi görüntünün eksik bileşenlerini tahmin etmek için her piksele uygulanan sabit bir hesaplama modeli kullanır. Adaptif olmayan yöntemlerin uygulanması adaptif yöntemlere göre kolay ve hesaplama karmaşıklığı çok daha düşüktür. Bu çalışma kapsamında adaptif olmayan interpolasyon yöntemlerinden, Nearest Neighborhood, bilinear, bicubic ve lanczos yöntemleri uygulanmıştır. Bu yöntemlere aşağıda yer verilmiştir.

1. **Nearest Neighborhood**

   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 pikselli bir ağırlıklı ortalama kullanılır. 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**

   Yöntem adını mucidi olan Cornelius Lanczos'dan almıştır. Bu yöntem, Lanczos filtresiyle görüntü boyutunu her yönde küçültülebilir veya büyütebilir. Bu

yöntem, bir enterpolasyon işlevi olarak 3 loblu Lanczos pencere işlevine dayanmaktadır. Lanczos, kaynak görüntüdeki noktanın yakınındaki 36 pikselin yoğunluklarını kullanır. Karmaşık bir yaklaşımdır.

6.1.2 Veri Setinin Dengelenmesi ve Artırılması

SCIAN-Morpho veri setindeki görüntüler dengesiz sınıf dağılımına ve yetersiz görüntü sayısına 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 setindeki 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.

6.1.3 Transfer Öğrenme

Transfer öğrenme 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 sonuç verdiği çalışmalarda kanıtlanmıştır [167]. Eğitilmiş modelde, ağırlıkları ile bir başka problem uzayında sadece belirlenen katman sayısının ağırlıklarının güncellenmesi üzerine ince ayar (fine-tuning) işlemi yapılır. Böylece modelin transfer öğrenimi şeklinde kullanılması mümkün hale gelir. Çalışmada transfer öğrenme kapsamında SCIAN-Morpho veri seti sınıflandırmasındaki ImageNet görüntü veri seti ile önceden eğitilmiş GoogLeNet, VGG19, Inceptionv3 ve MobileNetv2 ağlarının performans değerlendirilmesi yapılmıştır. Ayrıca çalışmada kullanılan veri seti çapraz doğrulama ile 5 fold yapısına dönüştürülmüştür. Sonrasında bütün modeller Tablo 6.1'de belirtilen parametreler ile test edilmiştir.

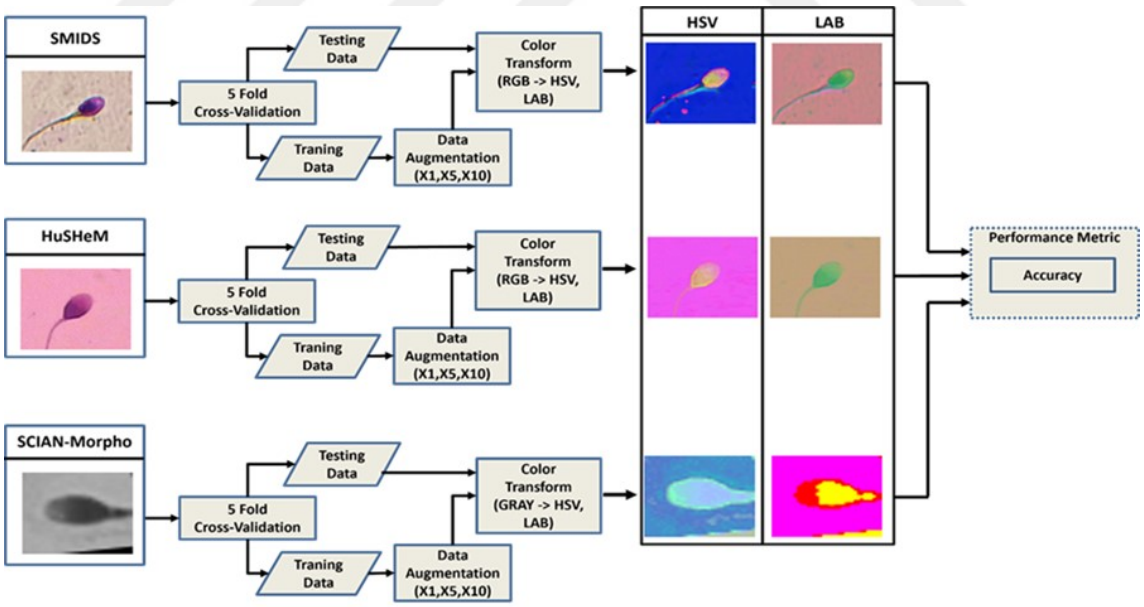
Tablo 6.1 Modeller için parametre değerleri

Parametre	Değerler
Mini Batch Boyutu	8
Aktivasyon Fonksiyonu	SoftMax
Optimizer	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

Test edilen modeller arasıda en iyi sonuçları VGG19 modeli vermiştir. Bu aşamadan sonra epoch analizi yapılmıştır. Böylece bazı katmanları değiştirilmiş VGG19 modeli ile %62 doğruluk, %85 kesinlik ve %75 duyarlılık başarı oranları elde edilmiştir.

6.2 Renk Uzay Dönüşümlerinin Sperm Sınıflandırmasına Etkisi

Sperm morfolojisinde sınıflandırma performansını artırmak için bir ön işleme adımı olarak farklı renk uzaylarının etkisi araştırılmıştır. Deneysel testlerde SMIDS, HuSHeM ve SCIAN-Morpho veri setleri kullanılmıştır. Sperm görüntülerinin sınıflar arasındaki dengesiz dağılımı ve yetersiz veri nedeniyle veri setleri üzerinde veri artırma işlemi uygulanmıştır. Daha sonra renk uzayının sınıflandırmadaki etkisini gözlemlemek için veri setleri çok iyi bilinen iki renk uzayı olan LAB ve HSV formatlarına dönüştürülmüştür. Sınıflandırma modeli olarak MobileNetV2 kullanılmıştır. Renk uzaylarının etkilerini göstermek için sonuçlar, renk dönüşümünün uygulanmadığı daha önce yayınlanmış bir çalışma [168] ile karşılaştırılmıştır. Daha önce yayınlanan çalışmada [168], MobileNetV1 ve MobileNetV2 gibi mobil platform tabanlı ağlar üç orijinal veri setine uygulanmıştır. Elde edilen sonuçlara göre MobileNetV2 en başarılı ağ olarak tanımlanmıştır. Bu bağlamda, renk uzayı dönüştürme ön işleminin sperm görüntülerini sınıflandırma performansı üzerindeki etkisini araştırmak için bir sınıflandırıcı olarak MobileNetV2 kullanılmıştır. MobileNetV2'nin ince ayar prosedürü, üç veri setinin RGB formatları yerine LAB ve HSV formatlarında gerçekleştirilmiştir. Bu çalışma kapsamında; Epoch sayısı, veri artırma oranı ve diğer parametreler daha önce yayınlanan çalışmada [168] olduğu gibi seçilmiştir. Sunulan çalışmanın genel akış şeması Şekil 6.2'de gösterilmiştir.



Şekil 6.2 Çalışmanın genel akışı

6.2.1 Mobil Platform Tabanlı Evrişimsel Sinir Modeli

Modelin eğitim aşamasında önce veri setleri 5 fold yapısına dönüştürülmüştür. Görüntülerin farklı varyasyonlarının kullanılması test setinin sınıflandırılmasında

ezberleme etkisini ortadan kaldırdığından veri artırma yöntemleri kullanılmıştır. Ayrıca, farklı aktivasyon fonksiyonları ile farklı epoch ve mini batch büyüklükleri uygulanarak ön testler yapılmıştır. Bu testler sonucunda en iyi doğruluk değerleri veren parametreler belirlenmiştir. Bu parametreleri belirlerken literatürde sıklıkla kullanılan değerler test edilmiştir. Yüksek epoch büyüklükleri, daha fazla işlem süresi gereksinimlerinin yanı sıra fazla uyum sorununa da neden olabilir. Düşük epoch büyüklükleri ise ağın tam olarak eğitilmemesine neden olabilir. Bu nedenle epoch büyüklüklerinin sınıflandırma performanslarına etkisini gözlemlemek için 10, 30 ve 50 epoch büyüklükleri ile eğitim gerçekleştirilmiştir.

MobileNetV2 hafif bir modeldir ve çalışmada transfer öğrenme fikri açısından kullanılmıştır. Sınıflandırma katmanı kullanılan veri setlerinin sınıf yapısına göre değiştirilmiş ve ardından *adamax* optimizasyon yöntemi uygulanmıştır. MobileNetV2'nin kullanılmasının temel nedeni, bilgisayar dışı mobil uygulamalar için daha uygun olması ve MobileNetV1, InceptionV3, DenseNet121 gibi diğer modellere göre daha iyi sonuçlar vermesidir [169, 170].

Değerlendirmede ana kıstas, sınıflandırmada en yüksek doğruluğu ve en az zaman karmaşıklığını sağlamaktır. LAB ve HSV renk uzayı dönüşümü sınıflandırma performanslarında farklı renk uzaylarının etkilerini gözlemlemek için bir ön işleme adımı olarak gerçekleştirilmiştir. MobileNetV2 ağının eğitim parametreleri Tablo 6.2'ye göre ayarlanmıştır.

Tablo 6.2 Modeller için parametre değerleri

Parametre İsimleri	**Değerleri**
Batch Size	8
Optimizer	adamax
Learning Rate	0,0001
Augmentation Sizes	X1 (Original Data), X5, X10
Epoch Sizes	10, 30, 50

6.2.2 Renk Uzayı Dönüşümü

Görüntü işleme alanında özellikle görüntü bölütlemede farklı renk uzaylarının kullanılması farklı doğrulukta sonuçlar üretmiştir. Bu nedenle görüntü işlemede uygun renk uzayının seçimi konusunda birçok araştırma yapılmıştır. Genel olarak, LAB ve HSV, seçilen en önemli iki renk uzayıdır [171]. Görüntü bölütlemesi için bu iki renk alanı arasında performans karşılaştırması yapılmıştır. Performans ölçümü için MeanSquare Error (MSE) ve Tepe İşaret Gürültü Oranı (Peak Signal to Noise Ratio - PSNR) gibi parametreler kullanılmıştır. RGB formatlı SMIDS, HuSHeM veri setleri LAB ve HSV formatlarına dönüştürülmüştür. Ayrıca gri seviye formatındaki SCIAN-Morpho veri seti LAB ve HSV formatlarına dönüştürülmüştür. Ardından bu veri setleri daha

Tablo 6.3 İki çalışmanın doğruluk değerlerinin karşılaştırılması (RGB->HSV)

(%)	MobileNetV2 [168] (RGB Format)									MobileNetV2 (Sonuçlar) (HSV Format)								
Aug/Epo	HuSHeM			SMIDS			SCIAN			HuSHeM			SMIDS			SCIAN		
	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30
X1	38	44	39	84	85	84	48	54	59	67	74	72	85	86	85	65	64	65
X5	62	73	75	85	87	87	54	62	62	78	80	71	88	87	88	67	68	67
X10	72	77	77	86	86	87	65	66	66	81	81	78	88	88	87	66	65	66

Tablo 6.4 İki çalışmanın doğruluk değerlerinin karşılaştırılması (RGB->LAB)

(%)	MobileNetV2 [168] (RGB Format)									MobileNetV2 (Sonuçlar) (LAB Format)								
Aug/Epo	HuSHeM			SMIDS			SCIAN			HuSHeM			SMIDS			SCIAN		
	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30	10	20	30
X1	38	44	39	84	85	84	48	54	59	76	76	75	87	88	86	60	62	63
X5	62	73	75	85	87	87	54	62	62	80	85	85	89	89	88	65	64	62
X10	72	77	77	86	86	87	65	66	66	82	84	83	89	89	88	64	65	62

önce yayınlanan çalışma [168] ile aynı koşullar altında MobileNetV2 kullanılarak sınıflandırılmıştır. MobileNetV2'nin sınıflandırma performansı, orijinal veri setlerinde ve farklı veri artırma ölçekleri üzerinden ölçülmüştür. Ayrıca maksimum doğruluğu elde etmek için ağın eğitim aşamasında farklı epoch boyutları test edilmiştir.

Daha önce yayınlanan çalışmada [168] her üç veri seti için mobil ağlara girdi olarak 3 kanallı RGB renk uzayı formatı kullanılmıştır. Çalışmanın bu kısmında HSV ve LAB renk uzayına dönüştürülmüş veri setleri, daha önce yayınlanmış çalışmada kullanılan RGB renk uzayının sonuçlarıyla karşılaştırılmıştır. RGB'den HSV ve LAB'ye dönüştürülen görüntüler için sınıflandırma performans sonuçları sırasıyla Tablo 6.3 ve Tablo 6.4'te verilmiştir.

Tablo 6.3 ve Tablo 6.4'e göre, artırılmış veri setleri (X5 ve X10) için elde edilen sınıflandırma performansları, hem HSV hem de LAB dönüşümü açısından orijinal (X1) veri setlerinden daha yüksektir.

HSV dönüşümünden sonra en yüksek sınıflandırma başarısı HuSHeM, SMIDS ve SCIAN-Morpho veri setleri için sırasıyla %81, %88 ve %68 doğruluk oranları olarak elde edilmiştir. LAB ile dönüştürülmüş görüntülerin sınıflandırılmasında HuSHeM, SMIDS ve SCIAN-Morpho veri setleri için sırasıyla %85, %89 ve %65 doğruluk oranlarında maksimum performans elde edilmiştir. Her iki tablo incelendiğinde en iyi sonuçlar eğitimde 30 epoch kullanıldığında elde edilmiştir. Diğer bir deyişle, 50 epoch kullanmak yerine 30 epoch kullanılarak daha hızlı ve daha yüksek sonuçlar alınabilir.

Dönüştürülen veri setlerinin sınıflandırma performansı orijinal formlarla (RGB) karşılaştırıldığında hem HSV hem de LAB renk uzaylarında RGB veri setlerinden daha yüksek doğruluk oranları elde edilmiştir. HuSHeM ve SMIDS için en yüksek doğruluk

değerleri LAB renk uzayı dönüştürme ön işleminden sonra sırasıyla %85 ve %89 olarak elde edilmiştir. SCIAN-Morpho veri seti için HSV'ye dönüştürülmüş görüntülerin sınıflandırılmasında maksimum doğruluk %68 olarak ölçülmüştür. Böylece, HuSHeM ve SMIDS, LAB renk uzayında daha iyi sınıflandırılırken, SCIAN-Morpho ise HSV renk uzayında daha iyi sınıflandırılmıştır.

6.3 MHSMA Veri setinin Sınıflandırılması

Tez kapsamında SMIDS, HuSHeM ve SCIAN-Morpho veri setlerinin sınıflandırılması için birçok çalışma gerçekleştirilmiştir. Çalışmanın bu bölümünde Modified Human Sperm Morphology Analysis (MHSMA) isimli açık erişimli düşük çözünürlüklü veri seti kullanılmıştır. Elde edilen sonuçlar MHSMA veri seti ile ilgili literatürde yapılan çalışmalarla karşılaştırılmıştır. [172, 173]. Veri seti ile ilgili kapsamlı bilgiler Tablo 4.1'de verilmiştir.

Çalışmada Resim 5.3'te gösterilen sıfırdan oluşturulan altı CNN modeli kullanılmıştır. Modellerin eğitiminden önce orijinal MHSMA veri seti 10, 20, 30, 40 ve 50 katına kadar artırılarak kaydedilmiştir. Veri setleri artırılırken Resim 5.2'de gösterilen yöntemler kullanılmıştır. Bu aşamada sadece eğitim veri setleri artırılmıştır. Test ve Validation setleri artırılmamıştır. Çalışmada modellerin eğitimi için kullanılan parametreler Tablo 6.5'te verilmiştir.

Tablo 6.5 MHSMA veri setinin eğitimi için kullanılan parametre değerleri

Parametre İsimleri	**Değerleri**
Batch Size	8
Optimizer	sgdm
Learning Rate	0.0001
Augmentation Sizes	X1 (Original Data), X10, X20, X30, X40, X50
Epoch Sizes	10, 30, 50

MHSMA veri setinin her bir artırılmış versiyonu için her bir modelin eğitimi gerçekleştirilmiştir. Daha sonra modellerin sınıflandırma çıkışları hard ve soft voting yöntemleri kullanılarak birleştirilmiştir. Doğruluk değerleri, 40 katına çıkarılmış veri setinde acrosome, head ve vacuole için sırasıyla %80,45, %84,07, %94,66 olarak elde edilmiştir.

Çalışmanın sonuçları Tablo 6.6'da görüldüğü gibi literatürdeki diğer çalışmalarla karşılaştırılmıştır. Elde edilen sonuçlar diğer yöntemlerle karşılaştırıldığında kayda değer sınıflandırma başarısı elde edilmiştir. Ayrıca önerilen yöntemde sınıflandırma tamamen otomatik olması ve daha basit modeller kullandığı için daha hızlı sonuçlar alınmıştır.

Tablo 6.6 MHSMA sperm veri setinin sınıflandırma başarısının literatürdeki diğer çalışmalarla karşılaştırılması

Veriseti Adı	Veriseti Bilgisi	Makale	Metot	Doğruluk (%)
MHSMA Acrosome	1540 Görüntü 64x64 Çözünürlük İki Sınıf (Normal - Anormal)	Miahi et al. 2022 [172]	Genetic Neural Architecture Search (GeNAS) Weighting Factor Random Search	77,66
		Javadi et al. 2019 [155]	Deep Learning Method Five-Fold Cross-Validation Sampling Data Augmentation Techniques	76,67
		Ghasemian et al. 2015 [44]	N/A	N/A
		Abbasi et al. 2021 [173]	Deep Multi-task Transfer Learning (DMTL) VGG 19	80,66
		Proposed Approach	40×Data Augmentation Design of Six CNN-Based Deep Learning Models Decision Level Soft Fusion	80,45
MHSMA Head	1540 Görüntü 64x64 Çözünürlük İki sınıf (Normal - Anormal)	Miahi et al. 2022 [172]	Genetic Neural Architecture Search (GeNAS) Weighting Factor Random Search	77,33
		Javadi et al. 2019 [155]	Deep Learning Method Five-Fold Cross-Validation Sampling Data Augmentation Techniques	77,00
		Ghasemian et al. 2015 [44]	Gaussian Filter Wavelet Transform Sobel Edge Median Filter	61,00
		Abbasi et al. 2021 [173]	Deep Multi-task Transfer Learning (DMTL) VGG 19	84,00
		Proposed Approach	40×Data Augmentation Design of Six CNN-Based Deep Learning Models Decision Level Soft Fusion	84,07
MHSMA Vacuole	1540 Görüntü 64x64 Çözünürlük İki Sınıf (Normal - Anormal)	Miahi et al. 2022 [172]	Genetic Neural Architecture Search (GeNAS) Weighting Factor Random Search	91,66
		Javadi et al. 2019 [155]	Deep Learning Method Five-Fold Cross-Validation Sampling Data Augmentation Techniques	91,33
		Ghasemian et al. 2015 [44]	Gaussian Filter Wavelet Transform Sobel Edge Median Filter	80,33
		Abbasi et al. 2021 [173]	Deep Multi-task Transfer Learning (DMTL) VGG 19	94,00
		Proposed Approach	40×Data Augmentation Design of Six CNN-Based Deep Learning Models Decision Level Soft Fusion	94,66

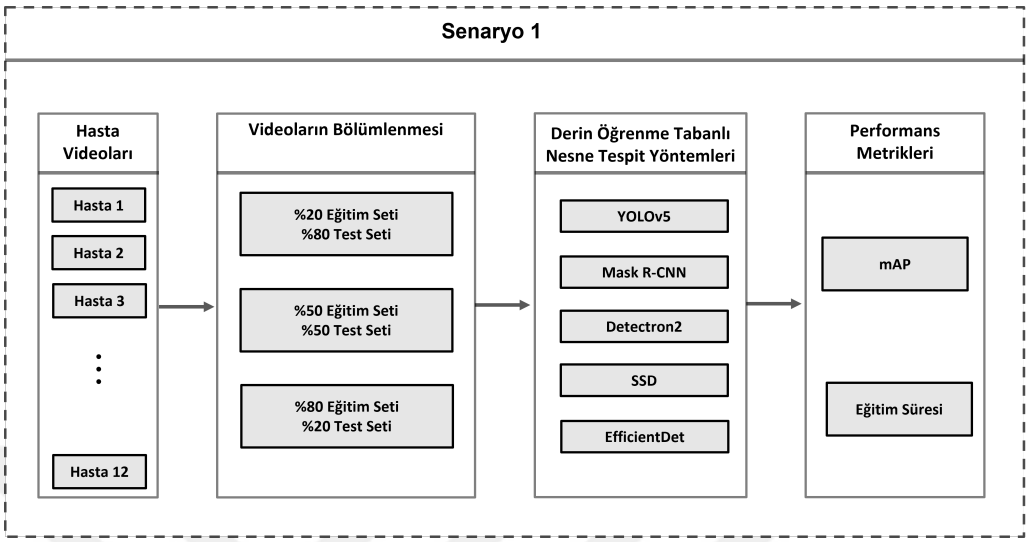
7
DERİN ÖĞRENME YÖNTEMLERİ İLE SPERM VİDEOLARINDA SPERM TESPİTİ

Sperm analizi yapılırken spermin morfolojik yapısı, sperm konsantrasyonu ve spermin hareket özellikleri olmak üzere üç teste odaklanılır. Daha önceki bölümlerde sperm görüntülerinin morfolojik sınıflandırılması üzerine çalışmalar yapılmıştır. Bu bölümde ise video veri setlerinde sperm örneklerinin konsantrasyon ve hareketlilik özelliklerini elde etmek için derin öğrenme tabanlı nesne tespit yöntemleri kullanılmıştır. İlk olarak farklı hastalardan elde edilen 12 adet sperm örneği videosu, Şekil 4.1'de mikroskobun oküler kısmına monte edilen bir akıllı telefon aracılığıyla kaydedilmiştir. Daha sonra videolar sperm ve sperm olmayan olarak ayrı ayrı etiketlenmiştir. Etiketlenmiş veri seti üzerinde iki farklı senaryo test edilmiştir. Ayrıca her senaryo için farklı düzenlenmiş eğitim ve test veri oranlarıyla test edilmiştir. İlk senaryoda, derin öğrenme tabanlı nesne algılama algoritmaları, farklı eğitim/test bölme oranları açısından etiketlenen videolar üzerinden doğrudan test edilmiştir. İkinci senaryoda, birinci senaryo ile sonuçlanan en düşük mAP (Mean Average Precision) değerlerine sahip videolar seçilmiştir. Videoların geri kalanı modellerin eğitimi için kullanılmıştır. İkinci senaryoda, model eğitiminde aynı videonun küçük bir bölümünün kullanılması yerine farklı videolar kullanılarak daha zorlu videolar için algılama performansının artırılması hedeflenmiştir. Sonuçlar doğruluk ve hız açısından karşılaştırılmıştır.

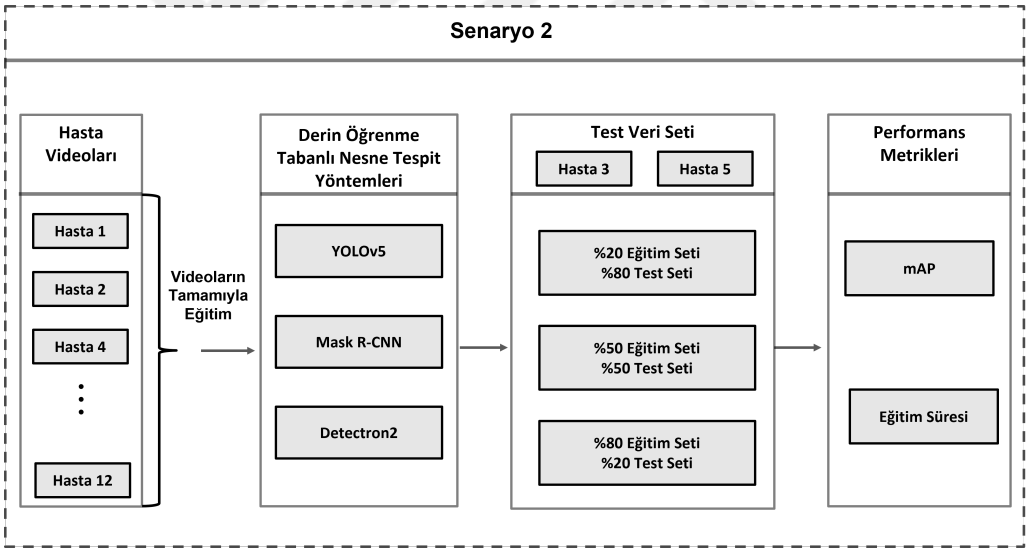
Bu bölümde literatürde görüntü işlemede etkili olan YOLOv5, Mask R-CNN, Detectron2, Single Shot Multibox Detector (SSD) ve EfficientDet olmak üzere beş yaygın derin öğrenme tabanlı nesne algılama yöntemlerinin performansları, sperm algılama problemi üzerinden değerlendirilmiştir.

Bölüm 4'te elde edilen sperm videolarının her bir bölgesi uzmanlar tarafından sperm ve sperm olmayan bölgeler olarak etiketlenerek veri seti oluşturulmuştur. Derin öğrenme tabanlı nesne algılama yöntemleri, oluşturulan veri seti kullanılarak mAP değeri ve model eğitim süreleri olmak üzere iki performans metriği üzerinden iki senaryo altında değerlendirilmiştir. Çalışmada kullanılan Senaryo 1 ve 2'nin akış

şemaları sırasıyla Şekil 7.1 ve Şekil 7.2'de verilmiştir. Oluşturulan veri seti, deneysel senaryolar, kullanılan yöntemler ve performans ölçütleri hakkında daha detaylı bilgi aşağıdaki alt bölümlerde verilmiştir.



Şekil 7.1 Senaryo 1'in akış şeması

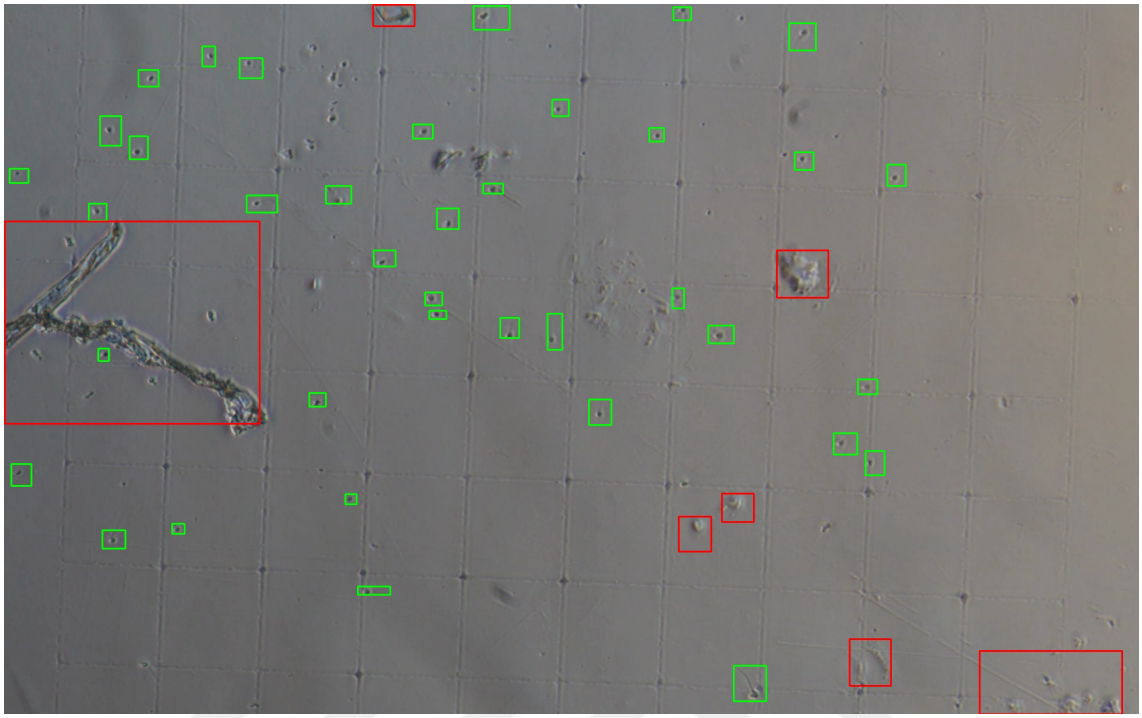


Şekil 7.2 Senaryo 2'nin akış şeması

7.1 Veri Setinin Etiketlenmesi

Sperm tespit problemi kapsamında derin öğrenme tabanlı nesne tespit yöntemlerinin değerlendirilmesi için farklı hastalardan elde edilen videolar 30 karede (fps) kaydedilmiştir. Daha sonra etiketleme işlemini kolaylaştırmak için etiketleme adımında kare hızı 5 fps'ye düşürüldü. Ayrıca görüntü boyutu 1920×1200 çözünürlük yerine 640×640 olacak şekilde yeniden boyutlandırılmıştır. Daha sonra LabelIMG [174] programı kullanılarak her çerçevedeki nesneler sperm ve sperm olmayan olarak

etiketlenmiştir. Etiketleme sonrası örnek bir çerçeve Şekil 7.3'te verilmiştir. Yeşil kutular sperm bölgelerini, kırmızı kutular ise uzmanlar tarafından sperm dışı olarak etiketlenen yapıları temsil etmektedir.



Şekil 7.3 Etiketleme örneği

Çalışmanın bu bölümünde veri seti üzerinde uzamsal veya başka herhangi bir veri artırma tekniği uygulanmamıştır. Her video için sperm ve sperm olmayan olarak etiketlenen bölgelerin ayrıntılı sayısı ve videoların süreleri ile ilgili bilgiler Tablo 7.1'de verilmiştir.

7.2 Uygulanan Senaryolar

Çalışmada iki senaryo test edilmiştir. Şekil 7.1'de gösterildiği gibi ilk senaryoda aynı hasta videosundan eğitim ve test setlerini kullanan hasta odaklı şema üzerinden değerlendirilmiştir. Bu şekilde modeller videoya özgü yönler açısından ayrı ayrı değerlendirilmiştir. Modeller üzerinde örneklem büyüklüğü etkisini gözlemlemek için ilk senaryo deneylerinde eğitim - test seti düzenlemeleri için üç oran belirlenmiştir. İlk olarak her videonun %20'si eğitim seti olarak ayrılmış, %80'i ise model değerlendirmesinde test olarak kullanılmıştır. Daha sonra eğitim verisi örneklerinin sayısının artışının model performansları üzerindeki etkisini göstermek için eğitim seti %50 ve %80'e yükseltilmiş ve test seti sırasıyla %50 ve %20'ye düşürülmüştür. Model eğitiminde az sayıda numune kullanılması ve bunun sonucunda yüksek performans elde edilmesi, daha az etiketleme işlemi gereksinimini gösterir. İlk senaryoda, üç

Tablo 7.1 Her bir video için etiket bilgisi

Video ID	Sperm Etiketi	Non-Sperm Etiketi	Toplam Etiket	Frame Sayısı
1	2510	361	2871	60
2	3095	267	3362	51
3	3798	151	3949	54
4	1224	45	1269	13
5	3443	163	3606	47
6	2974	1950	4924	150
7	15544	3597	19141	150
8	13191	4201	17392	150
9	7913	4800	12713	150
10	4516	3000	7516	150
11	5782	3000	8782	150
12	4554	2250	6804	150

farklı eğitim/test seti düzenlemesi açısından modellerin bireysel performansları 12 video üzerinden ölçülmüştür.

Şekil 7.2'de gösterildiği gibi ikinci senaryoda, ilk senaryoda elde edilen mAP değerleri en düşük olan iki video test olarak seçilirken, kalan tüm videolar modellerin eğitimi için kullanılmıştır. Ayrıca bu aşamada seçilen 2 videonun %20, %50 ve %80 bölümleri diğer videolarla beraber eğitime dahil edilerek farklı yaklaşımlar test edilmiştir. Bu senaryonun amacı, daha düşük mAP'li bu 2 videonun mAP değerlerinin iyileşip iyileşmediğini kontrol etmektir.

Çalışmanın ilk senaryosunda sperm tespiti ve sınıflandırması için YOLOv5, Mask R-CNN, Detectron2, SSD ve EfficientDet gibi beş nesne tespit yöntemi belirlenmiştir. İkinci senaryoda bu beş yöntemden YOLOv5, Mask R-CNN ve Detectron2 senaryo 1'de elde edilen en yüksek mAP puanlarına göre seçilmiştir. Sunulan çalışmada kullanılan tüm yöntemler, nesne algılama ile ilgili problemlerde kullanımları yaygın oldukları için seçilmiştir. Ayrıca sperm tespiti alanında bu yöntemlerle yapılmış yeterli çalışma bulunmamaktadır.

7.3 Nesne Tespitinde Kullanılan Derin Öğrenme Modelleri

Günümüzde literatürde birçok derin öğrenme tabanlı nesne algılama yöntemleri kullanılmaktadır. R-CNN [175], Fast R-CNN [176], Faster R-CNN [177] ve Mask R-CNN gibi modeller bölgesel tespitlere dayanan yöntemlerdir. Bu yöntemler, bölge teklif ağları (Region Proposal Networks - RPN) oluşturur ve ardından bu bölge tekliflerini kategorilere ayırır. Bölgesel hedef tespit yöntemi, her türlü potansiyel

bölgeyi ve çeşitli özellik katmanlarını içerdiğinden gerçek zamanlı nesne tespitini garanti etmeyebilir. Öte yandan, Regresyon tabanlı SSD, YOLO [178] ve türevleri [179] [180] [181] gibi yöntemler RPN'nin farklı versiyonlarını oluşturur ve bu önerileri anında sınıflandırır. Bu bölümde yüksek performanslı ve hızlı yöntemlerin belirlenmesi için literatürdeki yöntemler incelenmiş ve bu yöntemlerden daha başarılı olan YOLOv5, SSD, EfficientDet, Detectron2, Mask R-CNN modelleri seçilmiştir.

7.3.1 R-CNN

Bölge tabanlı evrişimsel sinir ağları nesne algılamaya derin modeller uygulayan öncü bir yaklaşımdır. Genel anlamda, R-CNN'ler dört ana bölümden oluşur:

1. Birden çok önerilen bölgeyi seçmek için giriş görüntüsü üzerinde seçici arama yapılır. Önerilen bu bölgeler genellikle birden çok ölçekte seçilir ve farklı şekil ve boyutlara sahiptir. Önerilen her bölgenin kategorisi ve sınırlayıcı kutusu etiketlenmiştir.
2. Önceden eğitilmiş bir CNN seçilir. Önerilen her bölgeyi ağın gerektirdiği girdi boyutlarına dönüştürür ve önerilen bölgelerden çıkarılan özelliklerin çıktısını almak için ileri hesaplamayı kullanır.
3. Önerilen her bölgenin özellikleri ve etiket kategorisi, nesne sınıflandırmasında birden çok destek vektör makinesini eğitmek için bir örnek olarak birleştirilir. Burada her bir destek vektör makinesi, bir örneğin belirli bir sınıfa ait olup olmadığını belirlemek için kullanılır.
4. Önerilen her bölgenin özellikleri ve etiketlenmiş sınırlayıcı kutusu doğrusal bir regresyon modeli eğitmek için bir örnek olarak birleştirilir.

R-CNN modelleri, görüntü özelliklerini etkin bir şekilde çıkarmak için önceden eğitilmiş CNN'ler kullansa da temel dezavantajı yavaş olmasıdır. Ayrıca nesne algılamayı gerçekleştirmek için CNN'den çok sayıda ileri hesaplama gerektiren tek bir görüntüden binlerce önerilen bölge seçilir. Bu işlem yükü, R-CNN'lerin gerçek uygulamalarda yaygın olarak kullanılmasında genel problem anlamına gelir.

7.3.2 Fast R-CNN

Bir girdi görüntüsü ve bir dizi nesne önerisi alır. Model bir özellik haritası oluşturmak için önce tüm görüntüyü birkaç evrişimsel ve maksimum havuzlama katmanıyla işler. Ardından, her nesne teklifinin ilgi bölgesi için havuzlama katmanı, özellik haritasından

sabit uzunlukta bir özellik vektörü çıkarır. Her özellik vektörü, sonunda softmax olasılıklarına ve çıktı olarak sınıf başına sınırlayıcı kutu regresyon ofsetlerine dallanan tam bağlı katmanların bir dizisiyle beslenir.

Fast R-CNN, önceki sürümün sahip olduğu birkaç dezavantajı çözerek R-CNN'nin çalışma süresini azaltmış ve buna ek olarak hızını ve doğruluğunu geliştirmiştir [176]. Ayrıca aşağıdaki avantajlara sahiptir;

1. R-CNN'den daha yüksek algılama kalitesi (mAP).
2. Eğitim, çoklu görev kaybı kullanılarak tek aşamalı olacak şekilde azaltılmıştır.
3. Tüm ağ katmanları güncellendiğinden eğitim süresi azalır.
4. Önbelleğe alma için disk depolaması gerekmez, bu bellek tasarrufu için iyidir.

7.3.3 Faster R-CNN

Faster R-CNN, Fast R-CNN'nin geliştirilmiş bir versiyonudur. Faster R-CNN modeli iki modüle ayrılmıştır: RPN ve Fast R-CNN sensörü. RPN, girişte görüntünün özellik haritalarını kabul eder. Ardından nesne önerileri koleksiyonunu oluşturur ve çıktıda öneri başına bir hedef puanı oluşturur. RPN'ler, giriş görüntüsünde aday nesneleri içerebilen sınırlayıcı kutulardır.

Fast R-CNN sensörü, giriş olarak çok sayıda İlgi Bölgesi (Region of Interest - ROI) kullanır. Alt örnekleme katmanı daha sonra her bir ROI için bir özellik vektörü çıkarır. Bu özellik vektörü, bir dizi tam bağlı katmanına girecektir. Son olarak iki çıktı elde edilir. İlk çıktı, farklı nesneler üzerinde tahmin edilen bir olasılık dizisidir. İkinci çıktı, sınırlayıcı kutunun koordinatlarıdır.

Fast R-CNN, 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 CNN ailesinden daha kısadır. Sonuç olarak, Faster R-CNN, RPN ve Fast R-CNN sensörü aynı evrişim katmanlarını paylaştığından, nesne tespiti için tek ve birleşik bir ağ olarak düşünülebilir. Faster R-CNN, kullandığı RPN sayesinde Fast R-CNN yaklaşımından daha az zaman alır. Dezavantajı, belirli bir resimdeki tüm hedefleri çıkarmak için birden fazla geçişe ihtiyaç duymasıdır. Ayrıca sıralı bir çalışma mantığı olduğu için bir sonraki işlemin başarısı bir öncekinin başarısına bağlıdır.

7.3.4 Mask R-CNN

Mask R-CNN, ilk aşamada R-CNN'lerin diğer sürümlerine benzer bir RPN'yi içerir. RPN aşamasında, aranan nesneyi içeren alanın önerileri tanımlanır. İkinci aşamada, RPN'nin çıktıları sınıflandırılır. Son olarak, sınıflandırma şemasına göre bir sınırlayıcı kutu oluşturulur ve çevresi [182] maskelenir. Mask R-CNN'nin ilk katmanlarında genellikle görüntünün kenar ve köşe gibi basit özellikleri algılanır. Sonraki katmanlarda, aranan nesnelerin özellikleri türetilir.

7.3.5 YOLOv1

YOLOv1, Joseph Redmon ve diğerleri tarafından yayınlanan bir makalede duyurulmuştur [178]. "Sadece Bir Kez Bakarsınız" anlamına gelir. Birçok nesne algılama algoritması, girdi görüntüsünün nesne algılama ağından birden çok kez geçirilmesini gerektirir veya nesne algılama için iki aşama gerekir. YOLO algoritmasında, giriş görüntüyü nesne algılama ağından yalnızca bir kez geçirilir. Bu, YOLO'nun diğer ağlardan temel farkıdır ve giriş görüntüsünün tek işlenmesi nedeniyle diğer modellerden daha hızlıdır.

YOLO, giriş görüntüsünü $7x7$'lik bir ızgaraya (grid) böler. YOLO, 49 ızgara hücrelerinin her biri için aynı anda bir sınıflandırma ve lokalizasyon problemi çalıştırır. Buradaki temel sorun, bir ızgara hücresinin yalnızca bir nesneyi algılayabilmesidir. Diğer bir problem ise 49 ızgara olduğu için bir resimdeki en fazla 49 nesneyi algılayabilmesidir. Ayrıca, eğer algılanacak nesne birden fazla ızgarayı kapsıyorsa bu nesne algılanmayabilir. Her hücre 2 sınırlayıcı kutu öngörür ve her kutu için model bir güven puanı verir. Bu puan, algılanan nesnenin ne kadar doğru olduğunu gösterir. Güven puanı düşük olan kutular dikkate alınmaz.

7.3.6 YOLOv2

Joseph Redmon ve meslektaşları, "YOLO9000: Better, Faster, Stronger" makalesinde YOLOv2'yi tanıtmıştır [179]. YOLOv2 ile duyarlılık ve lokasyonun geliştirilmesine odaklanılmıştır. Tüm evrişimsel katmanlara toplu normalleştirme eklenerek, mAP'de yüzde 2'lik bir artış elde edilmiştir. Daha önceki versiyonun tek bir ızgarada yalnızca bir nesneyi algılama gibi bir problemi vardı. Birden çok nesneyi algılamak için k sınırlayıcı kutu kullanmışlardır. YOLOv2'de, k sınırlayıcı kutuyu tahmin etmek için Bağlantı kutuları (Anchor boxes) fikri kullanılmış ve Darknet-19 adlı yeni bir sınıflandırma modeli önerilmiştir.

7.3.7 YOLOv3

Joseph Redmon ve Ali Farhadi, YOLOv2'yi güncelleyerek YOLOv3 modelini [180] tanıtmışlardır. YOLOv3, lojistik regresyon kullanarak her sınırlayıcı kutu için bir nesnellik puanı tahmin eder. YOLOv3 softmax yerine, herhangi bir sınıf için bağımsız lojistik sınıflandırıcılar kullanır. Eğitim sırasında, sınıf tahminleri için ikili çapraz entropi kaybını (binary cross-entropy loss) kullanmışlardır. Orta ve büyük boyutlu nesnelerde performansı daha kötüyken, küçük nesneleri algılamada daha başarılıdır. Özellik çıkarma için Darknet-53 ağı kullanır.

7.3.8 YOLOv4

Alexey Bochkovskiy ve meslektaşları tarafından [181] 2020 yılında geliştirilmiştir. YOLOv4'ün YOLOv3'ten belirgin farkı kullandığı temel yapıdır. Çünkü YOLOv3, Darknet53 çerçevesini kullanırken YOLOv4, CSPDarknet53 çerçevesini kullanır.

7.3.9 YOLOv5

YOLOv5, 2020 yılında Glenn Jocher [183] tarafından GitHub'da yayınlanmıştır. Önceki sürümlerden farklı olarak PyTorch'ta tasarlanmıştır. Jocher tarafından YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x olmak üzere dört farklı YOLOv5 modelinin alt versiyonları geliştirilmiştir. YOLOv5s, küçük ağ kullanması nedeniyle en az doğruluğu sağlarken, en yeni model olan YOLOv5x, en iyi doğruluğu veren ancak daha fazla eğitim süresi alan ve daha fazla veri gerektiren son derece büyük ağ kullanır. YOLOv5'in alt dalları arasında daha ayrıntılı karşılaştırma [183]'de verilmiştir. YOLOv5 daha önceki tüm sürümlerinden çok daha küçük ve daha hızlıdır. YOLOv4, belirli bir durumda 50 fps'ye ulaşırken, YOLOv5, aynı durum için 140 fps'ye kadar işlem yapma yeteneğine sahiptir. Ayrıca, YOLOv5x dışındaki tüm sürümlerin boyutları YOLOv4 [184] sürümünden daha küçüktür.

7.3.10 Detectron2

Detectron2 [185] panoptik ve öncü Mask R-CNN modelleriyle piramit ağlarının temel biçimini içerir. Bu yönüyle Detectron2, iki aşamalı mimarilerle benzerliklere sahiptir ancak eğitim sürelerini kısaltmak için daha özetlenmiş versiyonlardır.

7.3.11 EfficientDet

EfficientDet mimarisi, Google tarafından 2020 [186]'de tanıtılmıştır. Mimaride omurga ağı olarak kullanılan EfficientDet evrişimsel bir sinir ağı üzerinde

tasarlanmıştır. Ek olarak, kolay ve hızlı özellik çıkarımı sağlayan iki yönlü bir özellik ağı olan Çift Yönlü Özellikli Piramit Ağı (Bi-directional Feature Pyramid Network - BiFPN) burada önerilmiştir. EfficientDet görüntünün bölümlerini havuzlar ve karıştırır. Ardından, BiFPN ile füzyon katmanından geçen özellikleri çıkarır. BiFPN, çeşitli özellikleri çeşitli ayrıntılarda birleştirir, sınırlayıcı kutuları ve sınıf etiketlerini tahmin etmeyi sağlar.

7.3.12 SSD

SSD, 2016 yılında Liu ve arkadaşları [187] tarafından tanıtılmıştır. Bu yöntem, YOLOv1'e göre daha hızlı ayrıca bölge önerileri üreten diğer yöntemler kadar doğru sonuçlar üretir. Yaklaşım sınırlayıcı kutuların çıktı alanını, özellik haritası konumu başına farklı en boy oranları ve ölçekler üzerinden bir dizi varsayılan kutuya ayırır. Tahmin anında, her varsayılan kutudaki her nesne kategorisinin varlığı için puanlar üretir ve nesne şekline daha iyi uyması için kutuda ayarlamalar yapar. Ek olarak; model, çeşitli boyutlardaki nesneleri doğal olarak işlemek için farklı çözünürlüklere sahip birden çok özellik haritasından gelen tahminleri birleştirir. SSD, nesne önerileri gerektiren iki aşamalı yöntemlere benzer şekilde çalışır. Ancak teklif oluşturma adımını ortadan kaldırarak tekniği daha hızlı hale getirir. Ayrıca bir algılama bileşeni gerektiren sistemlere entegre etmek daha kolaydır.

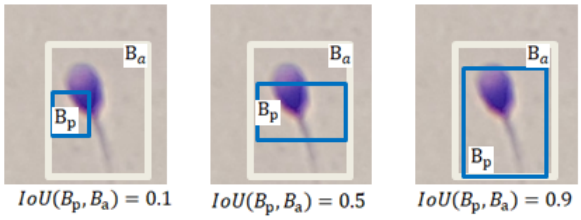
7.4 Performans Metrikleri

Derin öğrenme tabanlı nesne tespit yöntemlerinin sperm tespit problemi üzerindeki performansı *mAP* skorları ve eğitim süreleri olmak üzere iki kıstas açısından ölçülür. Nesne algılama problemlerinde en önemli değerlendirme metriği mAP'dır. Model tespitlerinin hem lokalizasyon hem de sınıflandırma performansları hakkında bilgi verir. Lokalizasyon performansı, bir örneğin tahmin edilen konum doğruluğunu belirler ve sınıflandırma performansını gösterir.

Intersection over Union (IoU) tahmin edilen bir sınırlayıcı kutunun (B_p) gerçek sınırlayıcı (Ground Truth - GT) kutu (B_a) üzerinde ne kadar doğru konumlandırıldığını ölçen bir işlevdir. IoU kullanımına bir örnek Şekil 7.4'te gösterilmektedir. Eşitlik 7.1'deki gibi tanımlanır.

$$IoU(B_p, B_a) = \frac{(B_p \cap B_a)}{(B_p \cup B_a)} \tag{7.1}$$

Literatürde genellikle $IoU(B_p, B_a) \geq 0,5$ olması durumunda makul olarak kabul edilir.



Şekil 7.4 IoU örneği

Tez çalışmasının genelinde, *mAP* skoru hesaplamalarında eşik değeri 0,5 olarak kabul edilen Pascal VOC mAP metriği kullanılmıştır.

Derin öğrenme modellerinin eğitim süreleri performans açısından önemli bir ölçüdür. Eğitim süreleri daha kısa olan ve daha yüksek *mAP* puanları veren modeller, tıp gibi zamanın kritik öneme sahip olduğu gerçek yaşam uygulamalarında kullanılabilecek en yararlı yöntemlerdir. Hastalardan veriler alındıktan sonra sonuçların en kısa sürede raporlanması tanıda önemli bir faktördür. Bu nedenle derin öğrenme tabanlı tespit yöntemleri, tıbbi teşhis problemlerinden biri olan sperm tespitinde kullanılması için eğitim süreleri açısından da değerlendirilmiştir. Sunulan tablolarda eğitim süreleri saat olarak gösterilmiştir.

7.5 Deneysel Sonuçlar

Sperm tespiti için YOLOv5, Mask R-CNN, Detectron2, SSD ve EfficientDet olmak üzere beş derin öğrenme tabanlı nesne tespit tekniği kullanılmıştır. Aşağıda Senaryo 1 ve Senaryo 2 sonuçları verilmiştir.

7.5.1 Senaryo 1 Sonuçları

Çalışmanın ilk senaryosunda, tüm hastaların videolarında sperm tespiti için kullanılan 5 yöntemin sonuçları Tablo 7.2'de verilmiştir. Ayrıca Tablo 7.2 ile, her hastanın videosu için farklı bölüm oranlarına sahip alt senaryoların sonuçları ayrıntılı olarak gösterilmektedir.

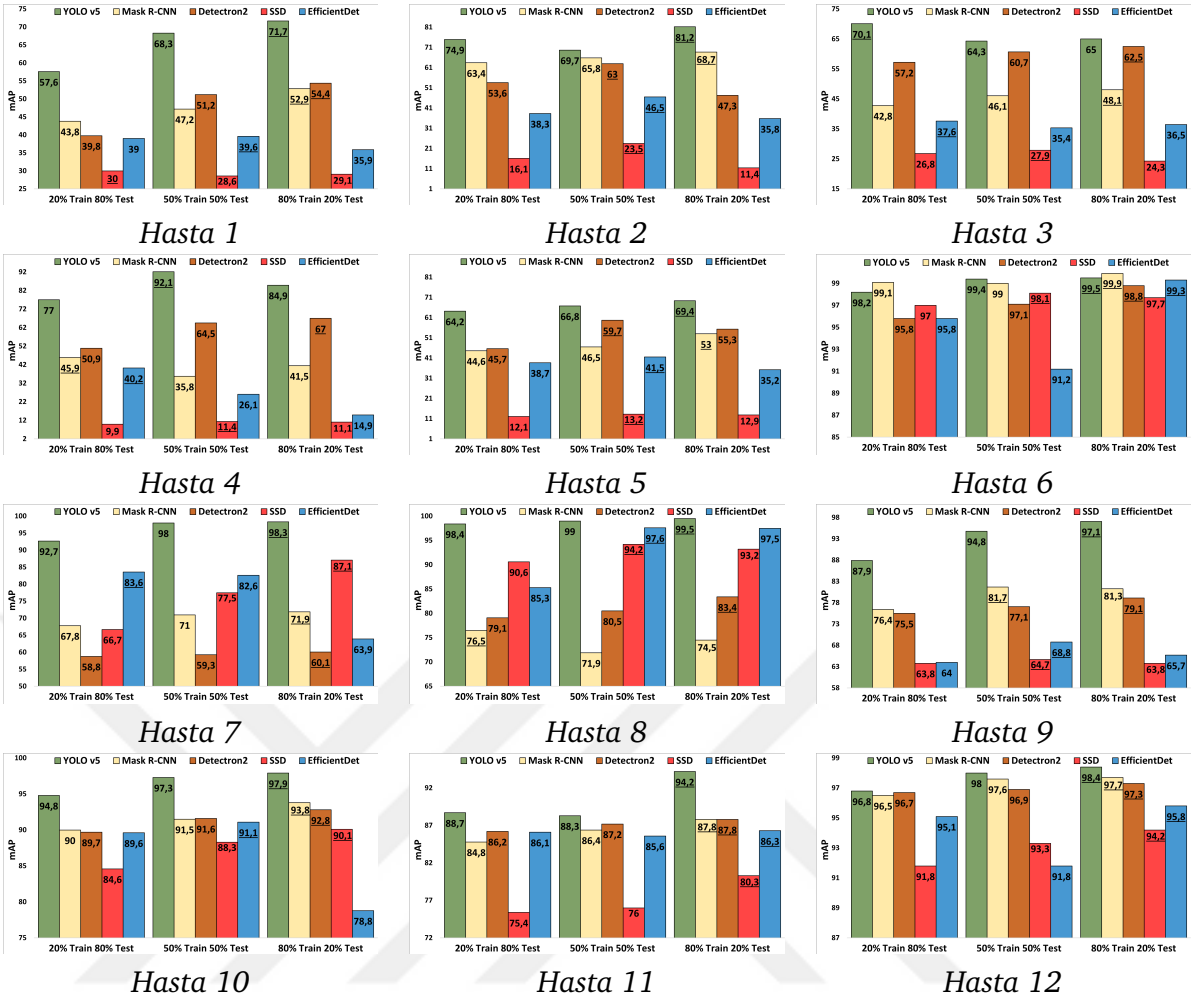
7.5.1.1 Senaryo 1 %20 Eğitim %80 Test Sonuçları

1. senaryonun ilk alt senaryosunda, veri setinin %20'si eğitim için, %80'i test için ayrılmıştır. Bu durumda sonuçlar Tablo 7.2'de ayrıntılı olarak verilmiştir. Sonuçlar incelendiğinde eğitim süreleri diğer alt senaryolara göre kısadır. Ancak *mAP* değerleri yeterli seviyede değildir.

Tablo 7.2 Senaryo 1'in sonuçları

Senaryo 1	Video	YOLOv5	Mask R-CNN	Detectron2	SSD	EfficientDet
20% Eğitim 80% Test	1	0,576	0,438	0,398	0,300	0,390
	2	0,749	0,634	0,536	0,161	0,383
	3	0,701	0,428	0,572	0,268	0,376
	4	0,770	0,459	0,509	0,099	0,402
	5	0,642	0,446	0,457	0,121	0,387
	6	0,982	0,991	0,958	0,970	0,958
	7	0,927	0,678	0,588	0,667	0,836
	8	0,984	0,765	0,791	0,906	0,853
	9	0,879	0,764	0,755	0,638	0,640
	10	0,948	0,900	0,897	0,846	0,896
	11	0,887	0,848	0,862	0,754	0,861
	12	0,968	0,965	0,967	0,918	0,951
	Ortalama mAP	**0,834**	**0,693**	**0,690**	**0,554**	**0,661**
	Ortalama Eğitim Süresi	**2,07**	**3,46**	**2,25**	**4,13**	**6,32**
50% Eğitim 50% Test	1	0,683	0,472	0,512	0,286	0,396
	2	0,697	0,658	0,630	0,235	0,465
	3	0,643	0,461	0,607	0,279	0,354
	4	0,921	0,358	0,645	0,114	0,261
	5	0,668	0,465	0,597	0,132	0,415
	6	0,994	0,990	0,971	0,981	0,912
	7	0,980	0,710	0,593	0,775	0,826
	8	0,990	0,719	0,805	0,942	0,976
	9	0,948	0,817	0,771	0,647	0,688
	10	0,973	0,915	0,916	0,883	0,911
	11	0,883	0,864	0,872	0,760	0,856
	12	0,980	0,976	0,969	0,933	0,918
	Ortalama mAP	**0,863**	**0,700**	**0,740**	**0,580**	**0,664**
	Ortalama Eğitim Süresi	**2,56**	**3,52**	**2,00**	**4,38**	**6,43**
80% Eğitim 20% Test	1	0,717	0,529	0,544	0,291	0,359
	2	0,812	0,687	0,473	0,114	0,358
	3	0,650	0,481	0,625	0,243	0,365
	4	0,849	0,415	0,670	0,111	0,149
	5	0,694	0,530	0,553	0,129	0,352
	6	0,995	0,999	0,988	0,977	0,993
	7	0,983	0,719	0,601	0,871	0,639
	8	0,995	0,745	0,834	0,932	0,975
	9	0,971	0,813	0,791	0,638	0,657
	10	0,979	0,938	0,928	0,901	0,788
	11	0,942	0,878	0,878	0,803	0,863
	12	0,984	0,977	0,973	0,942	0,958
	Ortalama mAP	**0,880**	**0,725**	**0,738**	**0,579**	**0,621**
	Ortalama Eğitim Süresi	**3,13**	**3,57**	**2,26**	**4,41**	**6,50**

Tablo 7.2 ile görüldüğü gibi, bu alt senaryoda YOLOv5 modeli en yüksek ortalama *mAP* değerine ulaşırken SSD modeli en düşük *mAP* değerini elde etmiştir. Mask R-CNN, Detectron2 ve EfficientDet modelleri ortalama olarak benzer sonuçlar elde etmiştir. Hastaların videoları ayrı ayrı incelendiğinde 6-12 arası hastalarda tüm modeller için sperm tespit başarısı 1-5 arası hastalara göre daha yüksektir. Bunun nedeni, çok sayıda frame olması ve etiketlenmiş veridir. 4. hastanın videosunda çok az etiketli veri



Şekil 7.5 Her hasta için senaryo 1'e göre modellerin mAP sonuçları

olduğundan SSD için bu çalışmada 0,099 gibi çok düşük bir *mAP* değeri elde edilmiştir. Bu bölümde her modelin ortalama eğitim süresi incelendiğinde en hızlı yöntem 2,07 saat ile YOLOv5 ve 2,25 saat ile Detectron2 yöntemleridir. En yavaş model 6,32 saat ile EfficientDet yöntemi olmuştur.

7.5.1.2 Senaryo 1 %50 Eğitim %50 Test Sonuçları

1. senaryonun ikinci alt senaryosunda, veri seti eşit miktarlarda bölünmüştür. Diğer bir ifadeyle, %50 eğitim için ve kalan %50 test için ayrılmıştır. Bu durumda sonuçlar Tablo 7.2'de ayrıntılı olarak verilmiştir.

Çalışmanın bu alt senaryosunda, Tablo 7.2'de görüldüğü gibi, bir önceki alt senaryoya göre 5 modelde ortalama olarak daha yüksek bir *mAP* değeri elde edilmiştir. Bir önceki alt senaryoda olduğu gibi bu senaryoda da sperm tespitinde en başarılı model YOLOv5 olmuştur. Şekil 7.5 incelendiğinde, Detectron2 ile modeli önceki alt

senaryoya göre her hasta için daha yüksek bir *mAP* değeri elde edilmiştir. YOLOv5, Mask R-CNN ve SSD modelleri, hastaların çoğunda daha yüksek *mAP* değerleri elde etmiştir. EfficientDet bir önceki duruma kıyasla en az *mAP* artış değerini elde etmiştir. Bir önceki alt senaryo ile karşılaştırıldığında, *mAP* değerindeki genel artışın nedeni, eğitime katılan veri sayısının %20'den %50'ye yükselmesinden kaynaklanmıştır. Şekil 7.6 ile görüldüğü gibi *mAP* değerleri bir önceki duruma göre artarken, Detectron2 dışındaki modellerin eğitim süreleri uzamıştır.

7.5.1.3 Senaryo 1 %80 Eğitim %20 Test Sonuçları

Son olarak, ilk senaryonun üçüncü alt senaryosunda, veri setinin %80'i eğitime, %20'si ise teste ayrıldı. Bu alt senaryoda da diğer iki alt senaryoda olduğu gibi YOLOv5 modeli en yüksek *mAP* değerine sahip en başarılı yöntem olmuştur. Tablo 7.2'deki ortalama *mAP* sonuçları incelendiğinde, YOLOv5 ve Mask R-CNN'nin önceki 2 alt senaryoya göre en yüksek *mAP* değerlerine ulaştığı görülmüştür. Detectron2 ve SSD ise önceki alt senaryo ile aynı ortalama sonucunu elde etmiştir. Ancak EfficientDet modeli bir önceki alt senaryoya göre daha düşük *mAP* elde etmiştir.

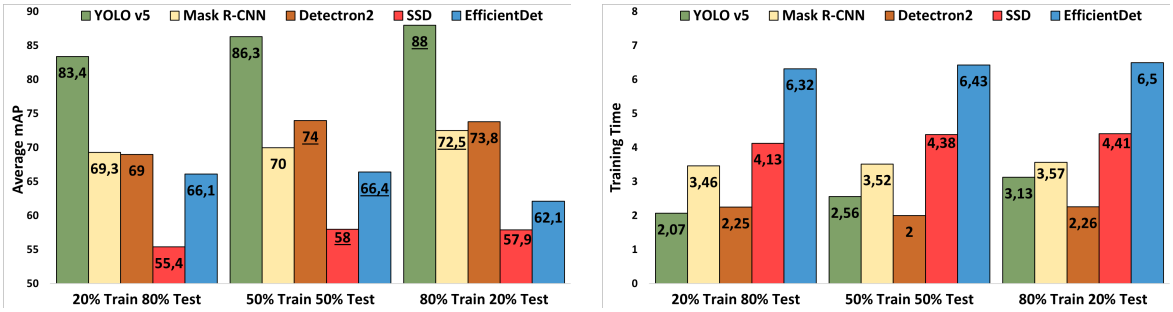
Genel anlamda sonuçlar incelendiğinde eğitime katılan veri miktarı arttıkça YOLOv5 ve Mask R-CNN modellerinin daha iyi öğrendiği ve daha yüksek *mAP* değerlerine ulaştığı görülmektedir. Detectron2 ve SSD modelleri ise bir önceki senaryoda olduğu gibi benzer sonuçlar alındığından çok fazla eğitilmesine gerek duyulmadığı ortaya çıkmıştır. Hatta bir önceki alt senaryoda yeterli öğrenme durumuna ulaştığı görülmektedir. Son olarak EfficientDet modelinde ise eğitim veri miktarı arttıkça ezberlemeye başlamakta ve daha kötü sonuçların elde edildiği görülmektedir.

Ek olarak, Şekil 7.5'te görüldüğü gibi, YOLOv5 modeli %50 eğitim seviyesine kıyasla 4. hasta hariç tüm hastalar için daha yüksek bir *mAP* değerine ulaşmıştır. 4. hasta için daha düşük *mAP* değerinin elde edilmesinin nedeni en az sayıda etiketli verinin bu videoda bulunmasıdır. Mask R-CNN, önceki alt senaryoya göre 11 hasta için daha yüksek başarı elde etmiştir. Bu alt senaryoda, diğer iki alt senaryoya göre daha fazla veri ile eğitim yapıldığında, 5 modelde diğer alt senaryolara göre daha fazla zaman harcanmıştır.

7.5.1.4 Senaryo 1 Eğitim Süresi

Tablo 7.2 incelendiğinde Senaryo 1'in tüm alt senaryoları için kullanılan 5 modelin ortalama eğitim süreleri gösterilmektedir.

Şekil 7.6'da ortalama *mAP* ve ortalama eğitim süresi bölümleri birlikte



Ortalama mAP *Ortalama eğitim süresi*

Şekil 7.6 12 hastanın senaryo 1'e göre modellerin ortalama mAP ve eğitim süreleri

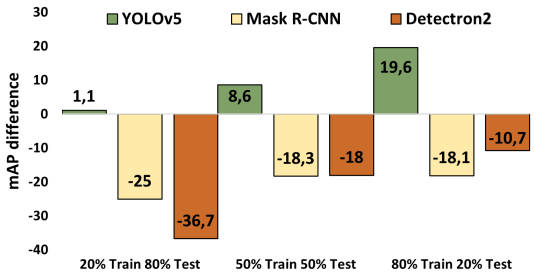
değerlendirildiğinde, 1. alt senaryoda YOLOv5 yönteminin hem eğitimi en kısa sürede tamamladığı hem de en yüksek ortalama *mAP* değerine sahip olduğu görülmektedir. EfficientDet yöntemi ise ortalama *mAP* değeri ve ortalama eğitim süresi açısından diğer yöntemlere göre en başarısız yöntem olarak görülmektedir. Detectron2 yönteminin diğer yöntemlere göre YOLOv5 yönteminden sonra en hızlı eğitim modeli olduğu görülmektedir. 2. ve 3. alt senaryolarda Detectron2 modeli videoların eğitimini YOLOv5 modeline göre daha hızlı tamamlamıştır. Ancak ortalama *mAP* açısından YOLOv5 modeline kıyasla çok vasattır.

7.5.2 Senaryo 2 Sonuçları

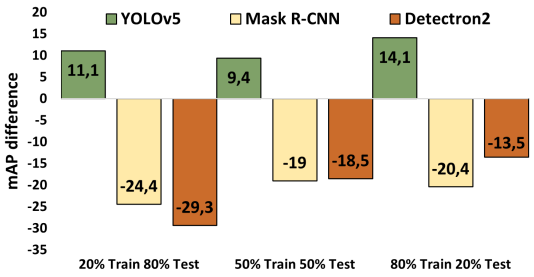
İlk senaryoda genel olarak en yüksek *mAP* değerleri 3. alt senaryoda yani %80 eğitim ve %20 test olarak elde edilmiştir. Tablo 7.2'de 1. senaryonun 3. alt senaryosu incelendiğinde en kötü *mAP*'ye sahip videolar tespit edilmiştir. Bu videolar tespit edilirken çalışmanın 1. senaryosunun tüm alt senaryolarında en iyi sonuçları veren YOLOv5, Mask R-CNN ve Detectron2 sonuçlarına yüksek *mAP* değerleri dikkate alınarak karar verilmiştir.

Tablo 7.2'de görüldüğü gibi, bu üç algoritma için en kötü *mAP* değeri Video 3 ve Video 5'te elde edilmiştir. Video 3'te yeterli etiket verisi yoktur ve bu videoda sperm görüntüleri kameranın titreşiminden dolayı bulanıktır. Bu nedenlerden dolayı, YOLOv5'in bu videodaki başarımı diğer videolardan biraz daha düşüktür. Video 5 ise Video 3 gibi yeterli etiketli veriye sahip değildir. Ayrıca Video 5'te, mikroskop ışığı altında oluşan efektler olarak tanımlanan Halo etkisi görülmektedir. Böylece sperm hücreleri olası Halo etkisi ile tam olarak ayırt edilemediği için diğer videolara göre yüksek bir *mAP* değeri elde edilememiştir.

İki problemli videonun eğitimi için diğer videoları eğitimde kullanarak daha yüksek bir *mAP* değeri elde edilmesi amaçlanmıştır. Bu yüzden YOLOV5, Mask R-CNN ve



Hasta 3 için mAP farkı



Hasta 5 için mAP farkı

Şekil 7.7 Hasta 3 ve Hasta 5 için modellerin senaryo 1 ve senaryo 2'ye göre mAP farkları

Detectron2 kullanarak 2. senaryo oluşturulmuştur. Böylece bu modellerin başarısı etiket sayısı az, gürültü ve Halo etkisi olan videolarda analiz edilmiştir. Ayrıca bu aşamada problemli videoların %20, %50 ve %80 gibi bölümleri eğitime dahil edilmiş ve 3 farklı alt senaryo yaklaşımı test edilmiştir.

Tablo 7.3'te Video 3 ve Video 5 için YOLOv5, Mask R-CNN ve Detectron2 yöntemlerinin *mAP* ve eğitim süreleri hakkında detaylı bilgi verilmektedir. Sonuçlar her iki senaryonun 3 alt senaryosuna göre verilmiştir. Şekil 7.7'de 3. ve 5. videolar için 2. senaryonun *mAP* değerleri ile 1. senaryonun *mAP* değerlerinin farkı 3 alt senaryoya göre ayrı ayrı verilmiştir.

Bu sonuçlar incelendiğinde YOLOv5'in *mAP* değeri birinci senaryoya göre artarak daha iyi sonuçlar elde edilirken, Mask R-CNN ve Detectron2 modelleri birinci senaryoya göre daha düşük *mAP* değerleri elde etmiştir. Bu sonuçlar göz önüne alındığında, eğitimde daha fazla video veri setinin bir arada kullanılmasıyla YOLOv5 yetersiz etiket, gürültü ve Halo efektlerine sahip videolarda başarılı sonuçlar verirken, Mask R-CNN ve Detectron2 ise tam tersine daha düşük bir *mAP* değeri elde etmiştir.

Son olarak, modellerin başarısı ve genel olarak eğitim süreleri incelendiğinde, YOLOv5 açık ara en iyi dedektördür. İkinci sırada Detectron2 yer almıştır, YOLOv5'ten daha hızlı bir eğitim süresine sahiptir ancak *mAP* ortalaması daha düşüktür. Faster R-CNN'nin türetilmiş bir versiyonu olan Mask RCNN, Detectron2 kadar iyi skor elde etmiştir. Fakat eğitim süresi ortalaması Detectron2'nin neredeyse iki katı kadar olmuştur. İyi sonuçların aksine, SSD, küçük nesneler üzerinde kötü algılama performansı nedeniyle sperm algılama sorunu için en düşük *mAP* puanlarını almıştır. Bu, nesne algılama yöntemi için iyi bilinen bir sorundur.

Tablo 7.3 Video 3, Video 5 için senaryo 1 ve senaryo 2 sonuçları

VİDEO	SENARYO		YOLOv5		Mask R-CNN		Detectron2	
			mAP	Time	mAP	Time	mAP	Time
Video 3	Senaryo 1	20% Eğitim 80% Test	0,701	1,43	0,428	3,40	0,572	02,29
		50% Eğitim 50% Test	0,643	1,41	0,461	4,25	0,607	02,25
		80% Eğitim 20% Test	0,650	1,44	0,481	4,28	0,625	02,23
	Senaryo 2	20% Eğitim 80% Test	0,712	2,40	0,178	3,52	0,205	02,21
		50% Eğitim 50% Test	0,729	2,43	0,278	4,05	0,427	02,25
		80% Eğitim 20% Test	0,846	2,51	0,300	4,18	0,518	02,16
Video 5	Senaryo 1	20% Eğitim 80% Test	0,642	1,43	0,446	4,00	0,457	2,21
		50% Eğitim 50% Test	0,668	2,05	0,465	4,12	0,597	2,22
		80% Eğitim 20% Test	0,694	1,49	0,530	4,35	0,553	2,21
	Senaryo 2	20% Eğitim 80% Test	0,753	2,18	0,202	4,16	0,164	2,19
		50% Eğitim 50% Test	0,762	2,20	0,275	4,25	0,412	2,22
		80% Eğitim 20% Test	0,835	2,25	0,326	4,18	0,418	2,21

8
SONUÇ VE ÖNERİLER

Kısırlık oranı tüm dünyada giderek artmaktadır. Kısırlık problemi dünya nüfusunun %20'sini etkilemektedir. Sağlıksız ebeveynlerden aktarılan genetik problemlerin yanı sıra; stres, beslenme alışkanlıkları, yüksek radyasyon ve ısıya maruz kalma gibi birçok çevresel etki, sperm kalitesini etkileyen sorunların en önemli nedenleridir. Bu tür sorunlarla karşılaşan birçok kişi çare aramakta ve hastanelere başvurmaktadır. Bu noktada tedavi için doğru teşhis koymak çok önemlidir. Vakaların %50'ye yakınını erkek kaynaklı problemler oluşturduğu için, erkek kaynaklı faktörler kısırlık teşhisinde kritik öneme sahiptir. Erkek kısırlığı tanısında spermin morfolojik yapısı, konsantrasyonu ve hareketliliği en ayırt edici ve bilgilendirici özelliklerdir. Bu özelliklerin analizi spermiogram testleri ile günümüzde laboratuvarlarda VA ve CASA sistemleri ile yapılmaktadır. VA'nın uygulanması kolaydır ve ucuzdur, ancak gözlemci değişkenlik sorunu olarak adlandırılan kullanıcı deneyimlerine ve uzmanlığına büyük ölçüde bağlıdır. Ancak meni örneğinde çok sayıda sperm bulunduğundan, uzmanlar tarafından mikroskopla gözlemlemeye dayalı VA yöntem ile doğru ve kesin analiz yapmak zordur. Özellikle sayım ve hareket analizinde uzmanlar; ani hareket ve tıkanıklıklar nedeniyle analiz edilmesi gereken asgari sperm sayılarını saptayamamaktadırlar. VA'nın dezavantajlarını ortadan kaldırmak için CASA sistemlerinden yararlanılmıştır. CASA sistemlerinin VA'ya göre avantajları vardır. Ancak daha karmaşık ve maliyetlidirler. Ayrıca CASA sistemlerinin çıktıları sistemden sisteme değişiklikler göstermektedir ve bazı sınırlamaları vardır.

Tez kapsamında, bilgisayarlı sistemler ile manuel değerlendirmeleri birleştiren bir sistem geliştirilmiştir. Veri toplamanın görsel değerlendirmede olduğu gibi gerçekleştirildiği, ancak analizin bir bilgisayar tarafından yapıldığı yeni bir bilgisayarlı yaklaşım önerilmiştir. Böylece karmaşık CASA sistemlerini kullanmak yerine daha makul çözümler üretmek için farklı görüntü işleme tabanlı yaklaşımlar önerilmiştir. Tez çalışması kapsamında sperm görüntülerinin morfoloji analizi ve sperm videolarında sperm tespiti olmak üzere iki modüler sistem geliştirilmiştir.

Birinci modülde öncelikle sperm görüntülerini içeren dört farklı veri setinin morfolojik

sınıflandırılması yapılmıştır. Sınıflandırma yapılmadan önce interpolasyon ve renk dönüşümü gibi işlemler gerçekleştirilmiştir. Ön işlemlerden sonra veri setlerinin veri eksikliği ve dengesizliğini çözmek için farklı veri artırma işlemleri uygulanmıştır. Ayrıca daha önceden büyük verilerle eğitilen yaygın modellerin ağırlıkları transfer öğrenme yöntemi ile veri setlerinin sınıflandırılmasında kullanılmıştır. Bunların yanı sıra, sperm morfolojisi analizini otomatikleştiren altı farklı derin öğrenme ağı oluşturulmuştur. Oluşturulan modellerin sınıflandırma sonuçları soft-voting ve hard-voting olmak üzere iki füzyon yöntemi ile birleştirilmiştir. Sonuç olarak, SMIDS, HuSHeM ve SCIAN-Morpho veri setleri için sırasıyla %90,73, %85,18 ve %71,91'lik en iyi sınıflandırma performansları elde edilmiştir. Bir başka morfoloji veri seti olan MHSMA için ise acrosome, head ve vacuole için sırasıyla %80,45, %84,07 ve %94,66 doğruluk değerlerine ulaşılmıştır. Elde edilen sonuçlar literatürde yapılan ilgili çalışmalarla karşılaştırıldığında kayda değer başarılı sonuçlara ulaşıldığı görülmüştür.

Tez çalışmasının ikinci modülünde ise hem konsantrasyon hem de hareket analizi için kritik bir aşamaya sahip olan sperm tespiti problemi üzerinde derin öğrenme tabanlı nesne tespit yöntemlerinin performans değerlendirmesi sunulmuştur. Performanslar, zaman ve doğruluk ölçütleri açısından değerlendirilmiştir. Deneylerde hastalardan elde edilen sperm video örnekleri kullanılmıştır. Deneylerde uygulanan yöntemlerin belirlenmesinde sperm benzeri küçük nesne algılama problemleri üzerinde literatür performanslarına göre derin öğrenme tabanlı nesne algılama yöntemleri araştırılmış ve seçilmiştir. Sonuçların adil bir şekilde değerlendirilebilmesi için tüm modeller belirli parametrelere sahip değerlerle eğitilmiştir. Video veri setinin doğru etiketlenmesi nesne tespitinde büyük rol oynadığından, etiketleme öncesi hastaların videoları 30 kareden 5 kareye düşürülmüştür. Etiketlenen videolar farklı senaryolarda eğitilmiş ve test edilmiştir. Sperm tespitinde %95 üzerinde kabul edilebilir sonuçlara ulaşan bu çalışmadaki önerilen yöntem ayrıca hızlı bir şekilde tamamlanan eğitime sahiptir.

Önerilen sistem esas olarak ucuz ve kolay erişilebilir bir sperm analiz çerçevesi sunmaktadır. Bu tez kapsamında yapılan çalışmanın özgün yanı veri seti üzerinde yapılan tüm değişikliklerin otomatik olarak yapılmasıdır. Yüksek performanslı bilgisayar kullanıldığı takdirde önerilen sistem ile sperm videolarında gerçek zamanlı sperm analizi kolaylıkla gerçekleştirilebilir.

Gelecek çalışmalarda farklı boyama-işlemleri uygulayan çeşitli laboratuvarlardan alınacak farklı sperm görüntülerinin analizleri hedeflenmektedir. Ayrıca sperm morfolojisi ve sperm tespitinin yanında sperm hareketi ve sayımı ile ilgili analizi sağlayan daha kapsamlı sistemler geliştirilecektir.

KAYNAKÇA

[1] W. H. Organization, *World health statistics 2010*. World Health Organization, 2010.

[2] E. Carlsen, A. Giwercman, N. Keiding, N. E. Skakkebæk, "Evidence for decreasing quality of semen during past 50 years.," *British medical journal*, vol. 305, no. 6854, pp. 609–613, 1992.

[3] A. Agarwal, A. Mulgund, A. Hamada, M. R. Chyatte, "A unique view on male infertility around the globe," *Reproductive biology and endocrinology*, vol. 13, no. 1, pp. 1–9, 2015.

[4] M. Roser, *Fertility rate*. [Online]. Available: `https://ourworldindata.org/fertilityrate` (visited on 2020).

[5] Ş. Başarslan, S. İçer, O. Ekmekçioğlu, "Image processing assisted sperm analysis," in *2017 Medical Technologies National Congress (TIPTEKNO)*, IEEE, 2017, pp. 1–4.

[6] S. Sinclair, "Male infertility: Nutritional and environmental considerations.," *Alternative medicine review: a journal of clinical therapeutic*, vol. 5, no. 1, pp. 28–38, 2000.

[7] F. Olayemi, "Review on some causes of male infertility," *African Journal of Biotechnology*, vol. 9, no. 20, 2010.

[8] E. H. Ruder, T. J. Hartman, M. B. Goldman, "Impact of oxidative stress on female fertility," *Current opinion in obstetrics & gynecology*, vol. 21, no. 3, p. 219, 2009.

[9] W. H. Organisation, *WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction*. Cambridge university press, 1999.

[10] S. T. Mortimer, "A critical review of the physiological importance and analysis of sperm movement in mammals," *Human reproduction update*, vol. 3, no. 5, pp. 403–439, 1997.

[11] S. Chan, C. Wang, S. Chan, P. Ho, W. So, Y. Chan, H. Ma, "Predictive value of sperm morphology and movement characteristics in the outcome of in vitro fertilization of human oocytes," *Journal of in vitro fertilization and embryo transfer*, vol. 6, no. 3, pp. 142–148, 1989.

[12] K. S. Murray, A. James, J. B. McGeady, M. L. Reed, W. W. Kuang, A. K. Nangia, "The effect of the new 2010 world health organization criteria for semen analyses on male infertility," *Fertility and sterility*, vol. 98, no. 6, pp. 1428–1431, 2012.

[13] J. Lu, Y. Huang, N. Lü, "Computer-aided sperm analysis: Past, present and future," *Andrologia*, vol. 46, no. 4, pp. 329–338, 2014.

[14] D. A. SATAR, S. GENÇDAL, "Sperm değerlendirmesi," *Arşiv Kaynak Tarama Dergisi*, vol. 22, no. 4, pp. 532–542, 2013.

[15] H.-F. Yang, X. Descombes, S. Prigent, G. Malandain, X. Druart, 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.

[16] G. M. Centola, "Semen assessment.," *The Urologic clinics of North America*, vol. 41, no. 1, pp. 163–167, 2013.

[17] A. Makler *et al.*, "The improved ten-micrometer chamber for rapid sperm count and motility evaluation," 1980.

[18] J. T. Smith, D. T. Mayer, "Evaluation of sperm concentration by the hemacytometer method: Comparison of four counting fluids," *Fertility and Sterility*, vol. 6, no. 3, pp. 271–275, 1955.

[19] L. F. Urbano, *Robust automatic multi-sperm tracking in time-lapse images*. Drexel University, 2014.

[20] R. P. Amann, D. Waberski, "Computer-assisted sperm analysis (casa): Capabilities and potential developments," *Theriogenology*, vol. 81, no. 1, pp. 5–17, 2014.

[21] C. L. R. Barratt, M. J. Tomlinson, I. D. Cooke, "Prognostic significance of computerized motility analysis for in vivo fertility," *Fertility and sterility*, vol. 60, no. 3, pp. 520–525, 1993.

[22] S. Baskaran, R. Finelli, A. Agarwal, R. Henkel, "Diagnostic value of routine semen analysis in clinical andrology," *Andrologia*, vol. 53, no. 2, e13614, 2021.

[23] L. Björndahl, C. Barratt, L. Fraser, U. Kvist, D. Mortimer, "Eshre basic semen analysis courses 1995–1999: İmmediate beneficial effects of standardized training," *Human Reproduction*, vol. 17, no. 5, pp. 1299–1305, 2002.

[24] C.-H. Yeung, T. G. Cooper, E. Nieschlag, "A technique for standardization and quality control of subjective sperm motility assessments in semen analysis," *Fertility and sterility*, vol. 67, no. 6, pp. 1156–1158, 1997.

[25] W. Ombelet, H. Pollet, E. Bosmans, A. Vereecken, "Results of a questionnaire on sperm morphology assessment.," *Human reproduction (Oxford, England)*, vol. 12, no. 5, pp. 1015–1020, 1997.

[26] A. M. JEQUIER, E. Ukombe, "Errors inherent in the performance of a routine semen analysis," *British Journal of Urology*, vol. 55, no. 4, pp. 434–436, 1983.

[27] D. Mortimer, M. Shu, R. Tan, "Standardization and quality control of sperm concentration and sperm motility counts in semen analysis," *Human Reproduction*, vol. 1, no. 5, pp. 299–303, 1986.

[28] B. C. Dunphy, L. M. Neal, I. D. Cooke, "The clinical value of conventional semen analysis," *Fertility and sterility*, vol. 51, no. 2, pp. 324–329, 1989.

[29] T. Rijsselaere, A. Van Soom, D. Maes, W. Nizanski, "Computer-assisted sperm analysis in dogs and cats: An update after 20 years," *Reproduction in Domestic Animals*, vol. 47, pp. 204–207, 2012.

[30] K. M. Engel, S. Grunewald, J. Schiller, U. Paasch, "Automated semen analysis by sqa vision® versus the manual approach—a prospective double-blind study," *Andrologia*, vol. 51, no. 1, e13149, 2019.

[31] S.-Y. Wei, H.-H. Chao, H.-P. Huang, C. F. Hsu, S.-H. Li, L. Hsu, "A collective tracking method for preliminary sperm analysis," *Biomedical engineering online*, vol. 18, no. 1, pp. 1–17, 2019.

[32] A. Agarwal, R. Henkel, C.-C. Huang, M.-S. Lee, "Automation of human semen analysis using a novel artificial intelligence optical microscopic technology," *Andrologia*, vol. 51, no. 11, e13440, 2019.

[33] T. R. Dias, C.-L. Cho, A. Agarwal, "Sperm assessment: Traditional approaches and their indicative value," in *In vitro fertilization*, Springer, 2019, pp. 249–263.

[34] E. Giaretta, M. Munerato, M. Yeste, G. Galeati, M. Spinaci, C. Tamanini, G. Mari, D. Bucci, "Implementing an open-access casa software for the assessment of stallion sperm motility: Relationship with other sperm quality parameters," *Animal reproduction science*, vol. 176, pp. 11–19, 2017.

[35] D. S. Guzick, J. W. Overstreet, P. Factor-Litvak, C. K. Brazil, S. T. Nakajima, C. Coutifaris, S. A. Carson, P. Cisneros, M. P. Steinkampf, J. A. Hill, *et al.*, "Sperm morphology, motility, and concentration in fertile and infertile men," *New England Journal of Medicine*, vol. 345, no. 19, pp. 1388–1393, 2001.

[36] R. Menkveld, F. S. Stander, T. J. v. Kotze, T. F. Kruger, J. A. v. Zyl, "The evaluation of morphological characteristics of human spermatozoa according to stricter criteria," *Human Reproduction*, vol. 5, no. 5, pp. 586–592, 1990.

[37] S. T. Mortimer, G. van der Horst, D. Mortimer, "The future of computer-aided sperm analysis," *Asian journal of andrology*, vol. 17, no. 4, p. 545, 2015.

[38] E. Alegre, V. González-Castro, R. Alaiz-Rodrıéguez, M. T. Garcıéa-Ordás, "Texture and moments-based classification of the acrosome integrity of boar spermatozoa images," *Computer Methods and Programs in Biomedicine*, vol. 108, no. 2, pp. 873–881, 2012.

[39] A. Bijar, M. Mikaeili, R. Khayati, *et al.*, "Fully automatic identification and discrimination of sperm's parts in microscopic images of stained human semen smear," *Journal of Biomedical Science and Engineering*, vol. 5, no. 2012, 2012.

[40] E. Alegre, M. Biehl, N. Petkov, L. Sanchez, "Assessment of acrosome state in boar spermatozoa heads using n-contours descriptor and rlvq," *Computer methods and programs in biomedicine*, vol. 111, no. 3, pp. 525–536, 2013.

[41] V. Chang, J. M. Saavedra, V. Castañeda, L. Sarabia, N. Hitschfeld, S. Härtel, "Gold-standard and improved framework for sperm head segmentation," *Computer methods and programs in biomedicine*, vol. 117, no. 2, pp. 225–237, 2014.

[42] V. Chang, A. Garcia, N. Hitschfeld, S. Härtel, "Gold-standard for computer-assisted morphological sperm analysis," *Computers in biology and medicine*, vol. 83, pp. 143–150, 2017.

[43] V. Chang, L. Heutte, C. Petitjean, S. Härtel, N. Hitschfeld, "Automatic classification of human sperm head morphology," *Computers in biology and medicine*, vol. 84, pp. 205–216, 2017.

[44] F. Ghasemian, S. A. Mirroshandel, S. Monji-Azad, M. Azarnia, Z. Zahiri, "An efficient method for automatic morphological abnormality detection from human sperm images," *Computer methods and programs in biomedicine*, vol. 122, no. 3, pp. 409–420, 2015.

[45] O. Garcıéa-Olalla, E. Alegre, L. Fernández-Robles, P. Malm, E. Bengtsson, "Acrosome integrity assessment of boar spermatozoa images using an early fusion of texture and contour descriptors," *Computer methods and programs in biomedicine*, vol. 120, no. 1, pp. 49–64, 2015.

[46] F. Shaker, S. A. Monadjemi, A. R. Naghsh-Nilchi, "Automatic detection and segmentation of sperm head, acrosome and nucleus in microscopic images of human semen smears," *Computer methods and programs in biomedicine*, vol. 132, pp. 11–20, 2016.

[47] F. Shaker, S. A. Monadjemi, J. Alirezaie, A. R. Naghsh-Nilchi, "A dictionary learning approach for human sperm heads classification," *Computers in biology and medicine*, vol. 91, pp. 181–190, 2017.

[48] L. Witkowski, "A computer system for a human semen quality assessment," *Biocybernetics and Biomedical Engineering*, vol. 33, no. 3, pp. 179–186, 2013.

[49] J. Riordon, C. McCallum, D. Sinton, "Deep learning for the classification of human sperm," *Computers in biology and medicine*, vol. 111, p. 103 342, 2019.

[50] M. Yüzkat, H. O. Ilhan, N. Aydın, "Morphological classification of low quality sperm ımages using deep learning networks," in *2020 Medical Technologies Congress (TIPTEKNO)*, IEEE, 2020, pp. 1–4.

[51] M. Yüzkat, H. O. Ilhan, N. Aydin, "Multi-model cnn fusion for sperm morphology analysis," *Computers in Biology and Medicine*, vol. 137, p. 104 790, 2021.

[52] M. Yüzkat, H. O. Ilhan, N. Aydin, "Effects of color space transformations on classification performance of sperm morphology," *Avrupa Bilim ve Teknoloji Dergisi*, no. 29, pp. 70–75, 2021.

[53] H. O. Ilhan, G. Serbes, N. Aydin, "Dual tree complex wavelet transform based sperm abnormality classification," in *2018 41st International Conference on Telecommunications and Signal Processing (TSP)*, IEEE, 2018, pp. 1–5.

[54] H. O. Ilhan, I. O. Sigirci, G. Serbes, N. Aydin, "The effect of nonlinear wavelet transform based de-noising in sperm abnormality classification," in *2018 3rd International Conference on Computer Science and Engineering (UBMK)*, IEEE, 2018, pp. 658–661.

[55] H. Ilhan, G. Serbes, N. Aydin, "Automatic directional masking technique for better sperm morphology segmentation and classification analysis," *Electronics Letters*, vol. 55, no. 5, pp. 256–258, 2019.

[56] H. O. Ilhan, G. Serbes, N. Aydin, "Automated sperm morphology analysis approach using a directional masking technique," *Computers in Biology and Medicine*, vol. 122, p. 103 845, 2020.

[57] H. O. Ilhan, I. O. Sigirci, G. Serbes, 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*, vol. 58, no. 5, pp. 1047–1068, 2020.

[58] I. Iqbal, G. Mustafa, J. Ma, "Deep learning-based morphological classification of human sperm heads," *Diagnostics*, vol. 10, no. 5, p. 325, 2020.

[59] J. Rittscher, R. Machiraju, S. T. Wong, *Microscopic image analysis for life science applications*. Artech House, 2008.

[60] O. L. Şavkay, M. E. Yalçın, "Analysis of sperm motility with cnn architecture," in *2012 13th International Workshop on Cellular Nanoscale Networks and their Applications*, IEEE, 2012, pp. 1–4.

[61] J. Liu, C. Leung, Z. Lu, Y. Sun, "Quantitative analysis of locomotive behavior of human sperm head and tail," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 2, pp. 390–396, 2012.

[62] Ł. Witkowski, "Examination of the density of semen and analysis of sperm cell movement," *Journal of Medical Informatics & Technologies*, vol. 3, 2002.

[63] M. Ravanfar, L. Azinfar, M. H. Moradi, R. Fazel-Rezai, "Occlusion robust low-contrast sperm tracking using switchable weight particle filtering," *Advances in Sexual Medicine*, vol. 2014, 2014.

[64] X. Zhou, L. Ma, Y. Shang, M. Xu, X. Fu, H. Ding, "Hybrid generative-discriminative learning for online tracking of sperm cell," *Neurocomputing*, vol. 208, pp. 218–224, 2016.

[65] A. A. Gunawan, A. M. Arymurthy, "Visual tracking for abrupt motions of human sperm using smoothing stochastic approximate monte carlo," *Procedia Computer Science*, vol. 59, pp. 64–72, 2015.

[66] L. F. Urbano, P. Masson, M. VerMilyea, M. Kam, "Automatic tracking and motility analysis of human sperm in time-lapse images," *IEEE transactions on medical imaging*, vol. 36, no. 3, pp. 792–801, 2016.

[67] P. Hidayatullah, I. Awaludin, R. D. Kusumo, M. Nuriyadi, "Automatic sperm motility measurement," in *2015 International Conference on Information Technology Systems and Innovation (ICITSI)*, IEEE, 2015, pp. 1–5.

[68] P. Hidayatullah, X. Wang, T. Yamasaki, T. L. Mengko, R. Munir, A. Barlian, E. Sukmawati, S. Supraptono, "Deepsperm: A robust and real-time bull sperm-cell detection in densely populated semen videos," *Computer Methods and Programs in Biomedicine*, vol. 209, p. 106 302, 2021.

[69] M. Aggarwal, V. S. Nair, T. Sun, "Using deep learning to streamline ıntracytoplasmic sperm ınjection in cancer patients,"

[70] V. Valiuškaitė, V. Raudonis, R. Maskeliūnas, R. Damaševičius, T. Krilavičius, "Deep learning based evaluation of spermatozoid motility for artificial insemination," *Sensors*, vol. 21, no. 1, p. 72, 2021.

[71] M. Rahimzadeh, A. Attar, *et al.*, "Sperm detection and tracking in phase-contrast microscopy image sequences using deep learning and modified csr-dcf," *arXiv preprint arXiv:2002.04034*, 2020.

[72] A. Arasteh, B. V. Vahdat, R. S. Yazdi, "Multi-target tracking of human spermatozoa in phase-contrast microscopy image sequences using a hybrid dynamic bayesian network," *Scientific reports*, vol. 8, no. 1, pp. 1–19, 2018.

[73] G. van der Horst, “Computer aided sperm analysis (casa) in domestic animals: Current status, three d tracking and flagellar analysis,” *Animal reproduction science*, vol. 220, p. 106 350, 2020.

[74] A. Singh, A. Kumar, A. Bisla, “Computer-assisted sperm analysis (casa) in veterinary science: A review,” *The Indian Journal of Animal Sciences*, vol. 91, no. 6, 2021.

[75] M. Iguer-Ouada, J. Verstegen, “Evaluation of the “hamilton thorn computer-based automated system” for dog semen analysis,” *Theriogenology*, vol. 55, no. 3, pp. 733–749, 2001.

[76] J. Verstegen, M. Iguer-Ouada, K. Onclin, “Computer assisted semen analyzers in andrology research and veterinary practice,” *Theriogenology*, vol. 57, no. 1, pp. 149–179, 2002.

[77] E. Bailey, N. Fenning, S. Chamberlain, L. Devlin, J. Hopkisson, M. Tomlinson, “Andrology lab corner* validation of sperm counting methods using limits of agreement,” 2007.

[78] T. G. Cooper, J. Neuwinger, S. Bahrs, E. Nieschlag, “Internal quality control of semen analysis,” *Fertility and sterility*, vol. 58, no. 1, pp. 172–178, 1992.

[79] M. Freund, B. Carol, “Factors affecting haemocytometer counts of sperm concentration in human semen,” *Reproduction*, vol. 8, no. 2, pp. 149–155, 1964.

[80] M. K. Kanakasabapathy, M. Sadasivam, A. Singh, C. Preston, P. Thirumalaraju, M. Venkataraman, C. L. Bormann, M. S. Draz, J. C. Petrozza, H. Shafiee, “An automated smartphone-based diagnostic assay for point-of-care semen analysis,” *Science translational medicine*, vol. 9, no. 382, eaai7863, 2017.

[81] D. Bompart, A. Garcıéa-Molina, A. Valverde, C. Caldeira, J. Yániz, M. N. de Murga, C. Soler, “Casa-mot technology: How results are affected by the frame rate and counting chamber,” *Reproduction, Fertility and Development*, vol. 30, no. 6, pp. 810–819, 2018.

[82] H. O. Ilhan, N. Aydin, “A novel data acquisition and analyzing approach to spermiogram tests,” *Biomedical Signal Processing and Control*, vol. 41, pp. 129–139, 2018.

[83] H. O. Ilhan, N. Aydin, “Smartphone based sperm counting-an alternative way to the visual assessment technique in sperm concentration analysis,” *Multimedia Tools and Applications*, vol. 79, no. 9, pp. 6409–6435, 2020.

[84] H. O. Ilhan, M. Yuzkat, N. Aydin, “Sperm motility analysis by using recursive kalman filters with the smartphone based data acquisition and reporting approach,” *Expert Systems with Applications*, vol. 186, p. 115 774, 2021.

[85] R. C. Gonzalez, R. E. Woods, S. L. Eddins, “Digital ımage processing, publishing house of electronics ındustry,” *Beijing*, pp. 295–300, 2002.

[86] B. Kamgar-Parsi, W. Lawson, B. Kamgar-Parsi, “Toward development of a face recognition system for watchlist surveillance,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, pp. 1925–1937, 2011.

[87] E. Hjelmås, B. K. Low, “Face detection: A survey,” *Computer vision and image understanding*, vol. 83, no. 3, pp. 236–274, 2001.

[88] G. Shakhnarovich, J. W. Fisher, T. Darrell, "Face recognition from long-term observations," in *European Conference on Computer Vision*, Springer, 2002, pp. 851–865.

[89] H. Li, W. Wu, "A new approach to image-based realistic architecture modeling with featured solid library," *Automation in construction*, vol. 13, no. 5, pp. 555–564, 2004.

[90] T. Zischinsky, L. Dorffner, F. Rottensteiner, "Application of a new model helicopter system in architectural photogrammetry," *International Archives of Photogrammetry and Remote Sensing*, vol. 33, no. B5/2; PART 5, pp. 959–959, 2000.

[91] N. R. Council, W. B. Committee, *et al.*, "Biometric recognition: Challenges and opportunities," 2010.

[92] C.-N. E. Anagnostopoulos, I. E. Anagnostopoulos, I. D. Psoroulas, V. Loumos, E. Kayafas, "License plate recognition from still images and video sequences: A survey," *IEEE Transactions on intelligent transportation systems*, vol. 9, no. 3, pp. 377–391, 2008.

[93] J.-W. Hsieh, S.-H. Yu, Y.-S. Chen, "Morphology-based license plate detection from complex scenes," in *Object recognition supported by user interaction for service robots*, IEEE, vol. 3, 2002, pp. 176–179.

[94] L. Xie, T. Ahmad, L. Jin, Y. Liu, S. Zhang, "A new cnn-based method for multi-directional car license plate detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 2, pp. 507–517, 2018.

[95] M. Sun, Q. Liu, K. Schmidt, J. Yang, N. Yao, J. D. Fernstrom, M. H. Fernstrom, J. P. DeLany, R. J. Sclabassi, "Determination of food portion size by image processing," in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2008, pp. 871–874.

[96] C. Bhabatosh *et al.*, *Digital image processing and analysis*. PHI Learning Pvt. Ltd., 1977.

[97] J. Schindelin, C. T. Rueden, M. C. Hiner, K. W. Eliceiri, "The ımagej ecosystem: An open platform for biomedical image analysis," *Molecular reproduction and development*, vol. 82, no. 7-8, pp. 518–529, 2015.

[98] K. Najarian, R. Splinter, *Biomedical signal and image processing*. Taylor & Francis, 2012.

[99] M. Haidekker, *Advanced biomedical image analysis*. John Wiley & Sons, 2010.

[100] J. Rajeswari, M. Jagannath, "Advances in biomedical signal and image processing–a systematic review," *Informatics in Medicine Unlocked*, vol. 8, pp. 13–19, 2017.

[101] G. Fernandez, M. Kunt, J.-P. Zryd, "Multi-spectral based cell segmentation and analysis," in *Proceedings of the Workshop on Physics-based Modeling in Computer Vision*, IEEE Computer Society, 1995, pp. 166–166.

[102] J. Sklansky, "Image segmentation and feature extraction," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 8, no. 4, pp. 237–247, 1978.

[103] S. K. Bandyopadhyay, "Pre-processing of mammogram images," *international journal of engineering science and technology*, vol. 2, no. 11, pp. 6753–6758, 2010.

[104] D. N. Ponraj, M. E. Jenifer, P. Poongodi, J. S. Manoharan, "A survey on the preprocessing techniques of mammogram for the detection of breast cancer," *Journal of Emerging Trends in Computing and Information Sciences*, vol. 2, no. 12, pp. 656–664, 2011.

[105] I. Philipp, T. Rath, "Improving plant discrimination in image processing by use of different colour space transformations," *Computers and electronics in agriculture*, vol. 35, no. 1, pp. 1–15, 2002.

[106] R. C. Gonzalez, *Digital image processing*. Pearson education india, 2009.

[107] N. Sharma, L. M. Aggarwal, "Automated medical image segmentation techniques," *Journal of medical physics/Association of Medical Physicists of India*, vol. 35, no. 1, p. 3, 2010.

[108] A. Eklund, P. Dufort, D. Forsberg, S. M. LaConte, "Medical image processing on the gpu–past, present and future," *Medical image analysis*, vol. 17, no. 8, pp. 1073–1094, 2013.

[109] D. Kaur, Y. Kaur, "Various image segmentation techniques: A review," *International Journal of Computer Science and Mobile Computing*, vol. 3, no. 5, pp. 809–814, 2014.

[110] S. Lomte, A. Janwale, "Plant leaves image segmentation techniques: A review," *Artic Int J Comput Sci Eng*, vol. 5, pp. 147–150, 2017.

[111] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[112] A. B. Patil, J. Shaikh, "Otsu thresholding method for flower image segmentation," *International Journal of Computational Engineering Research (IJCER)*, vol. 6, no. 05, 2016.

[113] S. Kaur, I. Singh, "Comparison between edge detection techniques," *International Journal of Computer Applications*, vol. 145, no. 15, pp. 15–18, 2016.

[114] P. P. Acharjya, R. Das, D. Ghoshal, "Study and comparison of different edge detectors for image segmentation," *Global Journal of Computer Science and Technology*, 2012.

[115] J. Canny, "A computation approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 670–700, 1986.

[116] X. Liu, Z. Deng, Y. Yang, "Recent progress in semantic image segmentation," *Artificial Intelligence Review*, vol. 52, no. 2, pp. 1089–1106, 2019.

[117] P. Stanchev, D. Green Jr, B. Dimitrov, "High level color similarity retrieval," 2003.

[118] D. Zhang, G. Lu, "Review of shape representation and description techniques," *Pattern recognition*, vol. 37, no. 1, pp. 1–19, 2004.

[119] D. Lu, Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International journal of Remote sensing*, vol. 28, no. 5, pp. 823–870, 2007.

[120] L. Deng, "A tutorial survey of architectures, algorithms, and applications for deep learning," *APSIPA transactions on Signal and Information Processing*, vol. 3, 2014.

[121] Q. Zhang, L. T. Yang, Z. Chen, P. Li, "A survey on deep learning for big data," *Information Fusion*, vol. 42, pp. 146–157, 2018.

[122] J. Gehring, Y. Miao, F. Metze, A. Waibel, "Extracting deep bottleneck features using stacked auto-encoders," in *2013 IEEE international conference on acoustics, speech and signal processing*, IEEE, 2013, pp. 3377–3381.

[123] G. E. Hinton, R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *science*, vol. 313, no. 5786, pp. 504–507, 2006.

[124] R. Salakhutdinov, A. Mnih, G. Hinton, "Restricted boltzmann machines for collaborative filtering," in *Proceedings of the 24th international conference on Machine learning*, 2007, pp. 791–798.

[125] X. Chen, X. Liu, Y. Qian, M. J. Gales, P. C. Woodland, "Cued-rnnlm—an open-source toolkit for efficient training and evaluation of recurrent neural network language models," in *2016 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, IEEE, 2016, pp. 6000–6004.

[126] J.-G. Lee, S. Jun, Y.-W. Cho, H. Lee, G. B. Kim, J. B. Seo, N. Kim, "Deep learning in medical imaging: General overview," *Korean journal of radiology*, vol. 18, no. 4, pp. 570–584, 2017.

[127] E. Maggiori, Y. Tarabalka, G. Charpiat, P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Transactions on geoscience and remote sensing*, vol. 55, no. 2, pp. 645–657, 2016.

[128] Y.-J. Cha, W. Choi, O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 5, pp. 361–378, 2017.

[129] A. Kızrak, *Derine daha derine: Evrişimli sinir ağları*. [Online]. Available: https://ayyucekizrak.medium.com/deri%CC%87ne-daha-deri%CC%87ne-evri%C5%9Fimli-sinir-a%C4%9Flar%C4%B1-2813a2c8b2a9 (visited on 2021).

[130] Y.-L. Boureau, J. Ponce, Y. LeCun, "A theoretical analysis of feature pooling in visual recognition," in *Proceedings of the 27th international conference on machine learning (ICML-10)*, 2010, pp. 111–118.

[131] S. Kim, *A beginner's guide to convolutional neural networks*. [Online]. Available: https://towardsdatascience.com/a-beginners-guide-to-convolutional-neural-networks-cnns-14649dbddce8 (visited on 2021).

[132] MathWorks, *Create simple deep learning network for classification*. [Online]. Available: https://www.mathworks.com/help/deeplearning/ug/create-simple-deep-learning-network-for-classification.html (visited on 2021).

[133] Mathworks, *Batchnormalizationlayer*. [Online]. Available: https://www.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.batchnormalizationlayer.html (visited on 2021).

[134] V. Nair, G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Icml*, 2010.

[135] Coursera, *Understanding mini-batch gradient descent*. [Online]. Available: https://www.coursera.org/lecture/deep-neural-network/understanding-mini-batch-gradient-descent-lBXu8 (visited on 2021).

[136] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R. M. Summers, "Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.

[137] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[138] M. D. Zeiler, R. Fergus, "Visualizing and understanding convolutional networks," in *European conference on computer vision*, Springer, 2014, pp. 818–833.

[139] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[140] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[141] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, Ieee, 2009, pp. 248–255.

[142] J. Han, D. Zhang, G. Cheng, N. Liu, D. Xu, "Advanced deep-learning techniques for salient and category-specific object detection: A survey," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 84–100, 2018.

[143] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[144] M. E. Sertkaya, *Derin öğrenme tekniklerinin biyomedikal imgeler üzerine uygulamaları*. [Online]. Available: https://openaccess.firat.edu.tr/xmlui/bitstream/handle/11508/18404/524890.pdf?sequence=1 (visited on 2021).

[145] S. Amidi, *Evrişimli sinir ağları*. [Online]. Available: https://stanford.edu/~shervine/l/tr/teaching/cs-230/cheatsheet-convolutional-neural-networks (visited on 2021).

[146] Coursera, *Parameters vs hyperparameters*. [Online]. Available: https://www.coursera.org/lecture/neural-networks-deep-learning/parameters-vs-hyperparameters-TBvb5 (visited on 2021).

[147] MathWorks, *Set up parameters and train convolutional neural network*. [Online]. Available: https://www.mathworks.com/help/deeplearning/ug/setting-up-parameters-and-training-of-a-convnet.html (visited on 2021).

[148] *Weight regularization*. [Online]. Available: https://medium.com/@mjbhobe/classifying-fashion-with-a-keras-cnn-achieving-94-accuracy-part-2-a5bd7a4e7e5a (visited on 2021).

[149] A. Koirala, K. B. Walsh, Z. Wang, C. McCarthy, "Deep learning–method overview and review of use for fruit detection and yield estimation," *Computers and electronics in agriculture*, vol. 162, pp. 219–234, 2019.

[150] F. Lussier, V. Thibault, B. Charron, G. Q. Wallace, J.-F. Masson, "Deep learning and artificial intelligence methods for raman and surface-enhanced raman scattering," *TrAC Trends in Analytical Chemistry*, vol. 124, p. 115 796, 2020.

[151] L. A. Jeni, J. F. Cohn, F. De La Torre, "Facing imbalanced data–recommendations for the use of performance metrics," in *2013 Humaine association conference on affective computing and intelligent interaction*, IEEE, 2013, pp. 245–251.

[152] E. Moen, D. Bannon, T. Kudo, W. Graf, M. Covert, D. Van Valen, "Deep learning for cellular image analysis," *Nature methods*, vol. 16, no. 12, pp. 1233–1246, 2019.

[153] D. M. Powers, "Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020.

[154] D. M. W. Powers, "The problem with kappa," in *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, 2012, pp. 345–355.

[155] S. Javadi, S. A. Mirroshandel, "A novel deep learning method for automatic assessment of human sperm images," *Computers in biology and medicine*, vol. 109, pp. 182–194, 2019.

[156] L. Perez, J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv preprint arXiv:1712.04621*, 2017.

[157] C. Shorten, T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, p. 60, 2019.

[158] Z. Hussain, F. Gimenez, D. Yi, D. Rubin, "Differential data augmentation techniques for medical imaging classification tasks," in *AMIA Annual Symposium Proceedings*, American Medical Informatics Association, vol. 2017, 2017, p. 979.

[159] Q. Wen, L. Sun, F. Yang, X. Song, J. Gao, X. Wang, H. Xu, "Time series data augmentation for deep learning: A survey," *arXiv preprint arXiv:2002.12478*, 2020.

[160] J. D. Rodriguez, A. Perez, 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.

[161] Y. Bengio, Y. Grandvalet, "No unbiased estimator of the variance of k-fold cross-validation," *Journal of machine learning research*, vol. 5, no. Sep, pp. 1089–1105, 2004.

[162] G. Aceto, D. Ciuonzo, A. Montieri, A. Pescapé, "Multi-classification approaches for classifying mobile app traffic," *Journal of Network and Computer Applications*, vol. 103, pp. 131–145, 2018.

[163] A. J. Pinar, J. Rice, L. Hu, D. T. Anderson, T. C. Havens, "Efficient multiple kernel classification using feature and decision level fusion," *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 6, pp. 1403–1416, 2016.

[164] J. Cao, S. Kwong, R. Wang, X. Li, K. Li, X. Kong, "Class-specific soft voting based multiple extreme learning machines ensemble," *Neurocomputing*, vol. 149, pp. 275–284, 2015.

[165] M. Amin-Naji, A. Aghagolzadeh, M. Ezoji, "Cnns hard voting for multi-focus image fusion," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–21, 2019.

[166] T. Mathworks, *(r2020a), the mathworks*. [Online]. Available: `https://www.mathworks.com/products/new_products/release2020a.html` (visited on 11/10/2020).

[167] 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.

[168] O. L. Tortumlu, H. O. Ilhan, "The analysis of mobile platform based cnn networks in the classification of sperm morphology," in *2020 Medical Technologies Congress (TIPTEKNO)*, IEEE, 2020, pp. 1–4.

[169] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.

[170] Q. Xiang, X. Wang, R. Li, G. Zhang, J. Lai, Q. Hu, "Fruit image classification based on mobilenetv2 with transfer learning technique," in *Proceedings of the 3rd International Conference on Computer Science and Application Engineering*, 2019, pp. 1–7.

[171] D. J. Bora, A. K. Gupta, F. A. Khan, "Comparing the performance of l* a* b* and hsv color spaces with respect to color image segmentation," *arXiv preprint arXiv:1506.01472*, 2015.

[172] E. Miahi, S. A. Mirroshandel, A. Nasr, "Genetic neural architecture search for automatic assessment of human sperm images," *Expert Systems with Applications*, vol. 188, p. 115 937, 2022.

[173] A. Abbasi, E. Miahi, S. A. Mirroshandel, "Effect of deep transfer and multi-task learning on sperm abnormality detection," *Computers in Biology and Medicine*, vol. 128, p. 104 121, 2021.

[174] *Labelımg is a graphical image annotation tool and label object bounding boxes in images*, `https://github.com/tzutalin/labelImg`, Accessed: 2021-09-30.

[175] R. Girshick, J. Donahue, T. Darrell, J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.

[176] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.

[177] S. Ren, K. He, R. Girshick, J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *arXiv preprint arXiv:1506.01497*, 2015.

[178] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[179] J. Redmon, A. Farhadi, "Yolo9000: Better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.

[180] J. Redmon, A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[181] A. Bochkovskiy, C.-Y. Wang, H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.

[182] K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.

[183] G. Jocher *et al.* (). "Yolov5," [Online]. Available: `https://github.com/ultralytics/yolov5` (visited on 08/13/2020).

[184] U. Nepal, H. Eslamiat, "Comparing yolov3, yolov4 and yolov5 for autonomous landing spot detection in faulty uavs," *Sensors*, vol. 22, no. 2, p. 464, 2022.

[185] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, R. Girshick. (). "Detectron2," [Online]. Available: `https://github.com/facebookresearch/detectron2` (visited on 11/10/2020).

[186] M. Tan, R. Pang, Q. V. Le, "Efficientdet: Scalable and efficient object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 10 781–10 790.

[187] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*, Springer, 2016, pp. 21–37.

TEZDEN ÜRETİLMİŞ YAYINLAR

Makale

1. Ilhan, H. O., Yuzkat, M., Aydin, N. (2021). Sperm Motility Analysis by using Recursive Kalman Filters with the smartphone based data acquisition and reporting approach. Expert Systems with Applications, 186, 115774.

2. Yüzkat, M., Ilhan, H. O., Aydin, N. (2021). Multi-model CNN fusion for sperm morphology analysis. Computers in Biology and Medicine, 137, 104790.

3. Yüzkat, M., Ilhan, H. O., Aydin, N. (2021). Effects of Color Space Transformations on Classification Performance of Sperm Morphology. Avrupa Bilim ve Teknoloji Dergisi, (29), 70-75.

Konferans Bildirisi

1. Yüzkat, M., Ilhan, H. O., Aydın, N. (2020, November). Morphological Classification of Low Quality Sperm Images Using Deep Learning Networks. In 2020 Medical Technologies Congress (TIPTEKNO) (pp. 1-4). IEEE.

Proje

1. KARİYER DESTEK PROJELERİ (KADEP) "Bilgisayar Destekli Sperm Tespiti ve Hareketliliğinin İncelenmesi" 2021-Devam