

**T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**HİSTOPATOLOJİK GÖRÜNTÜLERİN MAKİNE ÖĞRENMESİ
YÖNTEMLERİ İLE ANALİZİ**

ABDÜLKADİR ALBAYRAK

**DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ PROGRAMI**

**DANIŞMAN
DOÇ. DR. GÖKHAN BİLGİN**

İSTANBUL, 2018

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

HİSTOPATOLOJİK GÖRÜNTÜLERİN MAKİNE ÖĞRENMESİ
YÖNTEMLERİ İLE ANALİZİ

Abdülkadir ALBAYRAK tarafından hazırlanan tez çalışması 21.11.2018 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda **DOKTORA TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Doç. Dr. Gökhan BİLGİN
Yıldız Teknik Üniversitesi

Jüri Üyeleri

Doç. Dr. Gökhan BİLGİN
Yıldız Teknik Üniversitesi

Prof. Dr. Tamer ÖLMEZ
İstanbul Teknik Üniversitesi

Doç. Dr. Mehmet Sıddık AKTAŞ
Yıldız Teknik Üniversitesi

Prof. Dr. Lütfiye DURAK-ATA
İstanbul Teknik Üniversitesi

Doç. Dr. Mine Elif KARSLIGİL YAVUZ
Yıldız Teknik Üniversitesi

ÖNSÖZ

Tez çalışmamda benimle bilgi birikimini paylaşan, ileride mesleki hayatımda bana yön verecek tecrübelerini aktaran, motive eden sevgili danışman hocam Doç. Dr. Gökhan BİLGİN'e teşekkürlerimi sunarım. Ayrıca hem bilgisayar bilimleri hem de patoloji biliminin birleştiği bu çalışmada tüm özverileriyle yardımcı olan Dr. Aslı AKHAN ve Arş. Gör. Nurullah ÇALIK'a tüm özverilerinden dolayı teşekkür ederim. Akademik çalışmalarımda desteklerini hiçbir zaman esirgemeyen Prof. Dr. Lütfiye DURAK-ATA, Doç. Dr. Behçet Uğur TÖREYİN, Arş. Gör. Abdülkerim ÇAPAR'a teşekkür ve saygılarımı sunarım. Özellikle tez çalışmasının bir ayağında değerli zamanlarını ayırıp, katkılarını esirgemeyen Doç. Dr. İlknur TÜRKMEN ve Prof. Dr. Bahar MÜEZZİNOĞLU'na teşekkür ve saygılarımı sunarım.

Hayatımın her anında yanımda olan ve desteklerini hiçbir zaman esirgemeyen aileme sonsuz şükranlarımı sunuyorum. Son olarak en zor günlerimde manevi desteğini esirgemeyen sevgili eşim Gülbahar'a tüm kalbimle teşekkür ederim.

Kasım, 2018

Abdülkadir ALBAYRAK

İÇİNDEKİLER

Sayfa

SİMGE LİSTESİ	vii
KISALTMA LİSTESİ	viii
ŞEKİL LİSTESİ.....	ix
ÇİZELGE LİSTESİ	xii
ÖZET	xiv
ABSTRACT	xvi
BÖLÜM 1	
GİRİŞ	1
1.1 Literatür Özeti.....	1
1.2 Tezin Amacı.....	5
1.3 Hipotez.....	6
BÖLÜM 2	
SAYISAL PATOLOJİ	8
2.1 Patolojik Ön işlemler	8
2.1.1 Tespit (Fiksasyon)	8
2.1.2 Takip (Doku İşleme)	9
2.1.3 Blok lama	9
2.1.4 Kesme	9
2.1.5 Boyama	9
2.2 Sayısal Görüntü İşleme	10
2.2.1 Görüntünün Elde Edilmesi	10
2.2.2 Ön işleme - Görüntünün İyileştirilmesi	10
2.2.3 Bölütleme	11
2.2.4 Öznitelik Çıkarımı	11
2.2.5 Sınıflandırma	11

BÖLÜM 3

KULLANILAN VERİ KÜMELERİ	13
3.1 Böbrek Hücre Karsinomu (Renall Cell Carcinoma) Veri Kümesi	13
3.2 ICPR2014 - MITOS ATYPIA 14 Veri Kümesi	14
3.3 Camelyon Veri Kümesi - Lenf Düğümünde Metastaz Tespiti	15
3.4 Çalışmada Kullanılmak Üzere Oluşturulan Serviks (Rahim Ağzı Kanseri) Veri Kümesi	17

BÖLÜM 4

HİSTOPATOLOJİK GÖRÜNTÜLERİN BÖLÜTLENMESİ	19
4.1 Histopatolojik Görüntülerde Renk Uzaylarının Hücre Bölütlenmesine Etkileri	19
4.1.1 Görüntülerin Farklı Renk Uzaylarına Dönüştürülüp Bölütlenmesi	19
4.1.1.1 Renk Uzayları	19
4.1.1.2 Kullanılan Bölütleme Yöntemleri	20
4.1.2 Farklı Renk Uzaylarının Bölütlenmesi ile Elde Edilen Deneysel Sonuçların Değerlendirilmesi	22
4.1.3 Sonuç ve Gelecek Çalışmalar	26
4.2 Yüksek Çözünürlüklü Histopatolojik Görüntülerinin Bölütlenmesinde Süperpiksel Yaklaşımı	26
4.2.1 SLIC (Simple Linear Iterative Clustering) Süperpiksel Yaklaşımı	27
4.2.2 Süperpiksel Bölütleme Yaklaşımı ile Elde Edilen Deneysel Sonuçlar	28
4.2.3 Sonuç ve Çıkarımlar	32
4.3 Superpixel Bölütleme ve Derin Öğrenme Algoritmaları ile Histopatolojik Görüntülerin Segmentasyonu	32
4.3.1 SLIC ve Evrimsel Sinir Ağları Yöntemlerinin Birleştirilmesi	32
4.3.2 Önerilen Yaklaşım ile Elde Edilen Sonuçların Değerlendirilmesi	34
4.3.3 Sonuç ve Gelecek Çalışmalar	37
4.4 Süperpiksel-Tabanlı İki Aşamalı Bölütleme Algoritması ile Histopatolojik Görüntülerde Otomatik Hücre Segmentasyonu	37
4.4.1 Süperpiksel Algoritmasının Ön-Bölütleme olarak Kullanılması	38
4.4.2 Sonuç	42
4.5 Histopatolojik Görüntülerde Kodlayıcı-Çözücü Tabanlı Derin Öğrenme Algoritması ile Hücresel Yapıların Bölütlenmesi	43
4.5.1 Semantik Segmentasyon ile Bölütleme	43
4.5.2 Semantik Segmentasyonun Başarımının Değerlendirilmesi	45
4.5.3 Sonuç ve Çıkarımlar	48

BÖLÜM 5

HİSTOPATOLOJİK GÖRÜNTÜLERİN SINIFLANDIRILMASI	50
5.1 Histopatolojik Görüntülerde Mitozlu Hücrelerin Tamamlanmış Yerel İkili Örüntüler Kullanılarak Belirlenmesi	50
5.1.1 Yerel İkili Örüntü (YİÖ, Local Binary Pattern)	51
5.1.2 Yerel İkili Örüntülerin Sınıflandırma Başarısının Değerlendirilmesi ...	53

5.1.3	Sonuç	57
5.2	Mitozlu Hücrelerin Derin Öğrenme Tabanlı Öznitelikler Kullanılarak Belirlenmesi	58
5.2.1	Evrişimsel Sinir Ağları Tabanlı Mitozlu Hücre Tespiti Yaklaşımı	58
5.2.2	Önerilen ESA Yönteminin Başarısının Değerlendirilmesi	61
5.2.3	Sonuçların Değerlendirilmesi	62
5.3	Mitozlu Hücrelerin Tespitinde Derin Öğrenme ve Geleneksel Özellik Çıkarımı Algoritmalarının Karşılaştırılması.....	62
5.3.1	Önerilen ESA Yöntemi	63
5.3.2	Önerilen Yöntemin Konvansiyonel Yöntemlerle Karşılaştırılması	66
5.4	Büyük Boyutlu Medikal Görüntülerin Evrişimsel Sinir Ağları ile Sınıflandırılması	67
5.4.1	Görüntü Kesitlerinin ESA Yöntemi ile Sınıflandırılması	68
5.4.2	Elde Edilen Sonuçların Değerlendirilmesi.....	69
5.5	Histopatolojik Görüntülerde Tümörlü Bölgelerin Evrişimsel Sinir Ağlarıyla Sınıflandırılması	72
5.5.1	Lenf Nodlarındaki Tümörün Tespiti için Önerilen ESA Modeli	73
5.5.2	Önerilen Yöntemin Başarımının Değerlendirilmesi.....	74
5.5.3	Sonuç	76
5.6	Rahim Ağzı (Serviks) Kanseri Öncüsü Lezyonlarının Derecelendirilmesinde Kullanılmak üzere Oluşturulan Veri Kümesi ve Önerilen Yöntem	76
5.6.1	Materyal ve Metod	77
5.6.1.1	Verinin Elde Edilmesi ve Saklanması	77
5.6.1.2	Gözlemciler arası uyum/uyumsuzluk.....	80
5.6.1.3	Morfometrik Öznitelikler Kullanılarak Rahim Ağzı Kanseri Öncüsü Lezyonlarının Sınıflandırılması	81
5.6.2	DeneySEL Sonuçlar	84
5.6.3	Sonuç	87
BÖLÜM 6		
SONUÇ VE ÖNERİLER		88
KAYNAKLAR.....		95
ÖZGEÇMİŞ.....		102

SİMGE LİSTESİ

ε	Epsilon
θ	Teta
μ	Ortalama
U	Üyelik Matrisi

KISALTMA LİSTESİ

BCO	Bulanık C Ortalamaları
BDT	Bilgisayar Destekli Teşhis
BEP	Büyük Epitel Parçası
BM	Basal Membrane
CAD	Computer Aided Diagnosis
CIN	Cervical Intraepithelial Neoplasia
CNN	Convolutional Neural Network
DSÖ	Dünya Sağlık Örgütü
EM	Expectation Maximization
ERS	Entropy rate Superpixel
ESA	Evrişimsel Sinir Ağları
GNO	Gerçek Negatif Oranı
GPO	Gerçek Pozitif Oranı
HE	Hematoksilen-Eosin
HPV	Human Papilloma Virus
İHK	immünohistokimyasal
KEP	Küçük Epitel Parçası
Kİ	Kesin İsalet
OR	Overlap Ratio
PCA	Principal Components Analysis
SIL	Skumöz İntraepitelyal Lezyon
SLIC	Simple Linear Iterative Clustering
SP	Süperpiksel
TCGA	The Cancer Genome Atlas

ŞEKİL LİSTESİ

Sayfa

Şekil 3. 1	Veri kümesinden elde edilen üç adet yüksek çözünürlüklü histopatolojik görüntü ve bu iki görüntüye ait uzman patologlar tarafından işaretlenen kesin referans (ground truth) görüntüsü	14
Şekil 3. 2	ICPR2014 veri kümesindeki görüntülerden elde edilen çeşitli sayıda (a) mitozlu hücre ve (b) mitozsuz hücre örnekleri	15
Şekil 3. 3	Camelyon16 veri kümesindeki görüntülerden elde edilen (a) metastaz bölgeleri ve (b) metastaz olmayan bölgelerden elde edilen görüntü kesitleri.....	16
Şekil 3. 4	Rahim ağzı kanser öncüsü lezyonlarından oluşturulan veri kümesindeki görüntülerden elde edilen (a) yüksek çözünürlüklü HE ile boyanmış histopatolojik lezyon örneği (b) Bu örneğin Ki67 İHK boyası ile boyanmış örneği ve (c) aynı lezyonun p16 İHK boyası ile boyanmış örneği.....	18
Şekil 4. 1	Çalışmada takip edilen işlem adımları	23
Şekil 4. 2	Veri kümesinden alınan bir görüntüye ait (a) KYM renk uzayı (b) Kesin Referans (c) La*b* renk uzayının b* bandı (d) HSV renk uzayının v bandı (e) XYZ renk uzayı (f) Bölütleme sonucu	25
Şekil 4. 3	Çalışmada önerilen yöntemin akış diyagramı	28
Şekil 4. 4	(a) Veri kümesinden alınan histopatolojik görüntü ve (b) bu görüntüye ait kesin referans (ground truth), (c) görüntüye k-ortalamlar uygulanması sonucu, (d) k-ortalamlar sonucuna morfolojik işlem uygulanması, (e) bulanık c-ortalamlar sonucu, (f) bulanık c-ortalamlar sonucuna morfolojik işlem uygulanması	29
Şekil 4. 5	Veri kümesinden alınan (a) örnek bir görüntü, (b) bu görüntüye uygulanan süperpiksel yöntemi sonucu elde edilen piksel bölgeleri ve (c) çıktı görüntüsü	31
Şekil 4. 6	Veri kümesinden alınan (a) örnek bir görüntü, (b) bu görüntüye uygulanan süperpiksel yöntemi sonucu elde edilen piksel bölgeleri ve (c) çıktı görüntüsü	34
Şekil 4. 7	Önerilen bu çalışmada takip edilen işlem adımları. İlk aşamada elde edilen görüntüler yumuşatılmıştır. Daha sonra yumuşatılan görüntülerden süperpikseller elde edilmiştir. Son olarak bu süperpikseller kümeleme tabanlı yöntemlerle bölütlenmiş başarımları karşılaştırılmıştır.....	39

Şekil 4. 8	Örnek bir görüntüye uygulanan (a) süperpiksel bölütleme çıktısını, (b) bu süperpiksellere uygulanan k-ortalamar yönteminin çıktısını ve (c) nihai olarak elde edilen bölütleme sonucunu ifade etmektedir.	40
Şekil 4. 9	Belirli süperpiksel sayıları ve küme merkezlerine karşılık önerilen yöntemin elde ettiği bölütleme başarısı. (a) Kesinlik, (b) Özgünlük, (c) F-ölçütü ve (çakışma) oranları sırasıyla gösterilmektedir.	41
Şekil 4. 10	Veri kümesinden elde edilen bir görüntüye uygulanan algoritmalarından elde edilen sonuç.	42
Şekil 4. 11	Çalışmada takip edilen işlem adımları özetle şu şekildedir: Elde edilen histopatolojik görüntü ilk olarak histogram eşitleme işleminden geçirilmektedir. Daha sonra SegNet yapısına girdi olarak verilen derin öğrenme tabanlı yöntem ile bölütlenmektedir. Son olarak elde edilen bölütleme çıktısına morfolojik işlem uygulanmaktadır.	44
Şekil 4. 12	SegNet algoritmasının temel yapısı. Algoritma enkoder (kodlayıcı) ve dekoder (çözücü) olmak üzere iki aşamadan oluşmaktadır. Kodlayıcıda sınıflandırma amacıyla kullanılan standart vgg16 algoritması yer almaktadır. Çözücüde ise sonucun girdi görüntüsüne benzetilmesi için boyut arttırımı işlemi de gerçekleştirilmektedir.	46
Şekil 4. 13	Sırasıyla, (a) veri kümesinden elde edilen histopatolojik görüntüyü, (b) bu görüntüye ait uzman patolog tarafından işaretlenen kesin referans (ground truth) görüntüsünü (c) Segnet algoritmasının çalışması için kesin referans görüntülerinin dönüştürüldüğü görüntü tipini ve (d) SegNet algoritması sonucu elde edilen çıktı görüntüsünü ifade etmektedir.	48
Şekil 5. 1	Örnek 3x3'lük blok (b) LBP kodu (c) Merkeze olan uzaklıkların (fark) mutlak değerleri.	52
Şekil 5. 2	Örnek 3x3'lük blok (b) LBP kodu (c) Merkeze olan uzaklıkların (fark) mutlak değerleri.	53
Şekil 5. 3	(a) Veri kümesinden alınmış taranmış histopatolojik görüntü örneği (b)- (c) mitozlu hücre örneği (d)-(e) mitozsuz hücre örneği	54
Şekil 5. 4	Çalışmada takip edilen işlem adımları	54
Şekil 5. 5	(a) Orijinal görüntü (b) Bölütlenmiş görüntü (c) bölütlenmiş görüntüde gereksiz alanların temizlenmiş hali	56
Şekil 5. 6	Çalışma kapsamında takip edilen işlem adımları	59
Şekil 5. 7	ESA modelinin iç yapısı.	60
Şekil 5. 8	Çalışmada takip edilen işlem adımları	64
Şekil 5. 9	Çalışma kapsamında uygulanan ESA modelinin iç yapısı.	65
Şekil 5. 10	Önerilen çalışma kapsamında uygulanan ESA modeli	68
Şekil 5. 11	Veri kümesinden alınan görüntülerden elde edilen (a) metastaz bölgelerine ait görüntü örnekleri (b) Lam, yağ doku ve sağlıklı dokuları temsil eden görüntü örnekleri	70
Şekil 5. 12	İlgi bölgesinin işaretlenmesi (a) Veri kümesinden elde edilen bir görüntü ve, (b) daha detaylı analiz edilmek üzere ilgi bölgesinin işaretlenmesi. Kırmızı ile ifade edilen çizgi bazal membran, yeşil ile ifade edilen çizgi ise üst membranı göstermektedir. Epitel doku bu iki çizgi arasında yer almaktadır.	78
Şekil 5. 13	Önerilen bu çalışmada takip edilen işlem adımları..	80

Şekil 5. 14	Çalışma kapsamında veri kümesinden alınan ve belirli büyüklükte kesilen epitel parçalarının gösterimi.	81
Şekil 5. 15	Çalışma kapsamında veri kümesinden alınan ve belirli büyüklükte kesilen küçük epitel parçası. Kırmızı bazal membranı, yeşil çizgi üst membranı ve sarı çizgi ise papilla bölgelerini ifade etmektedir.	81
Şekil 5. 16	Veri kümesinden alınan örnek bir epitele SLIC superpiksel bölütleme algoritması uygulanması (a) Örnek KEP, (b) 3000 süperpikselin görüntü üzerine yansıtılması, (c) her bir süperpikselin ortalaması hesaplandıktan sonra elde edilen sonuç görüntüsü.	82
Şekil 5. 17	Önerilen yöntemin veri kümesinden elde edilen (a). Örnek girdi görüntüsü (b) bölütleme sonucu ve (c) morfolojik işlemlerden sonra oluşan çıktı görüntüsü.	83
Şekil 5. 18	Çakışık olarak elde edilen hücresel yapılar. (a) çakışık hücreleri temsil eden ikili görüntü (b) bu ikili görüntüye ait mesafe dönüşümü (Distance Transform) (c) Lokal maksimumların elde edilmesi.	84
Şekil 5. 19	Merkezi konumların kestiriminden sonra düzgün ve normal dağılım temelli elipslerin üretilmesi.	84

ÇİZELGE LİSTESİ

Sayfa

Çizelge 4. 1	XYZ Renk uzayındaki bütün görüntüler için ortalama doğruluk, F/F1-ölçümü, Kesinlik ve Hassasiyet değerleri.....	24
Çizelge 4. 2	KYM, La*b*, HSV renk uzaylarına ait başarımlar değerleri	25
Çizelge 4. 3	KYM, La*b*, HSV renk uzaylarına ait kesinlik (K,precision), hassasiyet(H,recall) ve F-ölçütü (F,F-measure) başarımlar değerleri.....	25
Çizelge 4. 4	KO ve BCO algoritmalarının veri kümesindeki görüntülere uygulandıktan sonra elde edilen başarımlar.....	30
Çizelge 4. 5	KO algoritmasının değişen süperpiksellere göre kümeleme başarımları	31
Çizelge 4. 6	Bütün Veri kümesine ait K-ortalamlar, Bulanık C-ortalamlar Bölütleme Başarısı (GPO ve GNO) ve Çalışma Zamanı	36
Çizelge 4. 7	ESA modelinin veri kümesinde yer alan bütün süperpiksellerin doğru ve yanlış sınıflandırma sayısı	37
Çizelge 4. 8	K-ortalamları (KO), Bulanık C-Ortalamları (BCO), SLIC+KO, SLIC+BCO, SLIC+DBSCAN, ERS, TPRS ve veri kümesinin paylaşıldığı çalışmada yer alan yöntemle ait Gerçek pozitif Oranı (GPO), Kesin İsalet (KI), F-Ölçütü (F-Ö), Gerçek Negatif Oranı (GNO), Çakışma Oranı (ÇO) ve Çalışma Zamanı (Ç.Zamanı) oranlarının karşılaştırılması	40
Çizelge 4. 9	Böbrek karsinomu veri kümesinde dört algoritmanın bölütleme performansının karşılaştırılması.....	47
Çizelge 5. 1	Algoritmaya ait sınıflandırma başarımları.....	56
Çizelge 5. 2	Boyut indirgeme öncesinde algoritmaya ait sınıflandırma başarımları	61
Çizelge 5. 3	Boyut indirgeme sonrasında algoritmaya ait sınıflandırma başarımları.....	62
Çizelge 5. 4	Önerilen yöntemin diğer geleneksel özellik çıkarma algoritmaları ile karşılaştırılarak elde edilen başarımlar oranları	66
Çizelge 5. 5	Çalışma kapsamında önerilen ESA modelinin 5-kat çapraz geçerliliğe sahip karışım matrisi	70
Çizelge 5. 6	ESA modelinin doğruluk, hassasiyet, F-Ölçütü ve başarımlar değerleri	71
Çizelge 5. 7	ESA modelinin evrişim katmanında yer alan filtrelerin farklı pencere boyutlarına sahip olması durumunda ESA'nın sınıflandırma performansı	75
Çizelge 5. 8	Eğitim aşamasından sonra kaydedilen ESA modelinin farklı 1000 tümörlü ve 1000 normal doku olmak üzere toplamda 2000 adet görüntü kesitindeki sınıflandırma başarımları.....	75
Çizelge 5. 9	CIN tabanlı derecelendirmede büyük epitel parçası (BEP) ve küçük epitel parçası (KEP) sayıları.....	79

Çizelge 5. 10	SIL tabanlı derecelendirmede büyük epitel parçası (BEP) ve küçük epitel parçası (KEP) sayıları.....	79
Çizelge 5. 11	Patolog1 ve patolog2'nin birbirlerine göre uyum ve uyumsuzlukları.....	84
Çizelge 5. 12	Kesin Tanının Patolog1, Patolog2 ve Önerilen Yöntem ile uyum ve uyumsuzlukları.....	85
Çizelge 5. 13	Kesin tanının Patolog1 (P1) ve Patolog2 (P2)'ye göre uyum ve uyumsuzluk durumu.....	86
Çizelge 5. 14	Önerilen yöntemin literatürde diğer çalışmalarda kullanılan Delaunay Triangulation yöntemi ile SIL tabanlı derecelendirmede karşılaştırılması sonucu elde edilen sonuçlar.....	86
Çizelge 5. 15	Çalışma kapsamında önerilen bölütleme yöntemleri, kullanılan süperpiksel yöntemlerinin literatürdeki bilinen algoritmaların performans analizi.....	89
Çizelge 5. 16	MITOS-ATYPIA-14 veri kümesinde test verisi üzerindeki sınıflandırma başarısı.....	91
Çizelge 5. 17	MITOS-ATYPIA-14 veri kümesinde 5-kat çapraz geçerlilik ile elde edilen sınıflandırma başarısı.....	92

HİSTOPATOLOJİK GÖRÜNTÜLERİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE ANALİZİ

Abdülkadir ALBAYRAK

Bilgisayar Mühendisliği Anabilim Dalı

Doktora Tezi

Tez Danışmanı: Doç. Dr. Gökhan BİLGİN

Son yıllarda kanser ve ona bağlı hastalıkların neden olduğu ölümler diğer hastalıkların sebep olduğu ölümlerden daha fazla ön plana çıkmaktadır. Kanser ve ona bağlı hastalıkların erken teşhisi bu hastalığın tedavi edilebilmesi bakımından oldukça önemlidir. Görüntüleme cihazlarının gelişmesi ile birlikte hastalığın görüntülenmesi, takibi ve Bilgisayar Destekli Teşhis (BDT) sistemlerinin de yardımıyla tedavi edilebilmesi mümkün hale gelmiştir. Özellikle yüksek çözünürlüğe sahip tarayıcılar yardımıyla doku ve organlardaki değişimlerin BDT tabanlı sistemler tarafından otomatik olarak tespiti mümkündür.

Bu tez çalışması histopatolojik görüntülerin bölütlenmesi ve sınıflandırılması olmak üzere iki ana bölümden oluşmaktadır. Çalışmanın bölütleme başlığında hücresel yapıların bölütlenmesinde süperpiksel yaklaşımı ve derin öğrenme tabanlı semantik bölütleme algoritmaları kullanılmıştır. Özellikle son yıllarda bilgisayarla görü alanında sıkça kullanılan süperpiksel bölütleme yöntemlerinden SLIC, SLIC-DBSCAN, ERS ve TPRS algoritmalarının yüksek çözünürlüklü histopatolojik görüntülerde hücresel yapıların bölütlenmesindeki başarımlarını performansları elde edilmeye çalışılmıştır. Bölütleme amacıyla SLIC süperpiksel bölütleme algoritması ve kümeleme tabanlı algoritmalarla birleştirilerek histopatolojik görüntülerde hücresel yapıların bölütlenmesi amacıyla yeni bir bölütleme algoritması önerilmiştir. Elde edilen bölütleme başarımları literatürde sıkça kullanılan yöntemlerle karşılaştırılmıştır. Buna ek olarak literatürde hareketli objelerin takibi, dış ortamdaki nesnelerin

bölütlenmesinde oldukça başarılı bir yöntem olan derin öğrenme tabanlı semantik bölütleme (SEGNET) yöntemi hücresel yapıların bölütlenmesinde kullanılmıştır. Bölütleme başarısı incelendiğinde literatürde sıkça kullanılan geleneksel yöntemlerden daha iyi sonuç verdiği gözlemlenmiştir. Histopatolojik görüntülerin sınıflandırılması bölümünde ise hem mitozlu hücrelerin hem de lenf nodlarında yer alan tümörlü bölgelerin özellikle son dönemde oldukça popüler olan evrimsel sinir ağları yöntemi ile tespiti gerçekleştirilmiştir. Özellikle mitozlu hücrelerin tespitinde geleneksel şekil, renk, doku ve istatistiksel tabanlı özellik çıkarma yöntemleri ile ESA yöntemi karşılaştırılarak performans analizi gerçekleştirilmiştir. Elde edilen sonuçlar incelendiğinde ESA modelinin geleneksel yöntemlerden daha başarılı sonuç verdiği gözlemlenmiştir. Ayrıca verilerin dengesiz olduğu durumlarda dengesiz verilere karşı gürbüz olan sınıflandırma yöntemleri (RusBoost) ile başarımın arttırılabileceği de görülmüştür.

Son olarak, İstanbul Medipol Üniversitesi Hastanesi, İstanbul Teknik Üniversitesi ve Yıldız Teknik Üniversitesi işbirliğiyle rahim ağzı kanseri öncü lezyonlarının derecelendirilmesi ile ilgili literatürde yer alan en geniş veri kümelerinden biri oluşturulmuş ve sınıflandırılmıştır. Elde edilen sınıflandırma başarısı iki patoloğ tarafından oluşturulan referans verilerle karşılaştırılarak analiz edilmiştir.

Anahtar Kelimeler: Histopatolojik görüntü analizi, makine öğrenmesi, evrimsel sinir ağları, semantik segmentasyon, serviks

**ANALYSIS OF HISTOPATHOLOGICAL IMAGES
VIA MACHINE LEARNING METHODS**

Abdülkadir ALBAYRAK

Department of Computer Engineering

Ph.D. Thesis

Adviser: Assoc. Prof. Dr. Gökhan BİLGİN

In recent years, deaths caused by cancer and diseases are more prominent than deaths caused by other diseases. Early diagnosis of cancer is very crucial for the treatment of this disease. With the development of imaging devices, it has become possible to treat the disease, follow-up and treat with the help of Computer Aided Diagnostic (CAD) systems. Particularly with high resolution scanners, it is possible to automatically detect changes in tissues and organs by CAD systems.

This thesis consists of two main sections, namely the segmentation and classification of high resolution histopathological images. Superpixels based and deep learning based semantic segmentation algorithms are used for the segmentation of cellular structures. The performance of SLIC, SLIC-DBSCAN, ERS and TPRS superpixels segmentation algorithms, which are frequently used in computer vision, have been tried to be obtained in the segmentation of cellular structures. For segmentation purpose, a new segmentation algorithm is proposed in order to segment the cellular structures in the histopathological images by combining the SLIC superpixels segmentation algorithm and clustering-based algorithms such as k-means and fuzzy c-means. The performance of the proposed method was compared with other well known methods used in the literature. In addition, deep learning-based semantic segmentation (SEGNET) method, which is a very successful method

in object detection, tracking of moving objects, segmentation of objects in outdoor, is used in the segmentation of cellular structures. In the “classification of histopathological images” section, the detection of mitotic cells and tumor regions located in the lymph nodes was carried out by using the convolutional neural network method which is very popular in recent years. In the detection of mitotic cells, performance analysis is performed by comparing the conventional shape, color, texture and statistical based feature extraction methods with the deep learning based methods. When the results are examined, it is observed that the ESA model perform better classification accuracy than the conventional methods. In addition, it is observed that the performance can be increased by using sampling based classification methods which are robust against unbalanced data such as RusBoost.

Finally, in cooperation with Istanbul Medipol University Hospital, Istanbul Technical University and Yildiz Technical University, one of the largest whole slide image data set in the literature regarding the grading of cervical cancer precursor lesions has been created and classified.

Keywords: Histopathological image analysis, machine learning, convolutional neural networks, semantic segmentation, cervical cancer

GİRİŞ

Ülkemizde ve dünyada son yıllarda kanser hastalıklarına bağlı ölümler önemli ölçüde artış göstermiştir. Dünyada her yıl kanser ve kansere bağlı hastalıklara yaklaşık olarak 14 milyon insan yakalanırken yaklaşık 8 milyon insan hayatını kaybetmektedir. Ülkemizde ise her yıl ortalama 100 bin erkek ve 70 bin kadın olmak üzere 170 bin insan kanser ve kansere bağlı hastalıklara yakalanmaktadır [1-2]. Bu hastalığın erken teşhis edilmesi hastalığın tedavi edilmesi açısından oldukça büyük bir öneme sahiptir. Kanserın erken teşhis edilmesinde önemli parametrelerden bazıları ülkelerin gelişmişlikleri ve bireylerin farkındalıkları olsa da teknolojinin gelişmesi ile birlikte görüntüleme cihazları ile kanser ve diğer hastalıkların tespiti, gelişiminin seyri ve müdahale edilmesi mümkün hale gelmiştir. Özellikle, son yıllarda manyetik rezonans (MR), bilgisayarlı tomografi (BT), ultrason cihazlarının geliştirilmesi sayesinde olası şüpheli durumlar görüntülenebilmektedir [3-5]. Bu görüntüleme cihazlarına alternatif olarak özellikle son yıllarda yüksek çözünürlüklü histopatolojik görüntüleme cihazları ve patoloji laboratuvarlarında gerçekleştirilen incelemelerle kanserin erken teşhis edilmesi mümkün hale gelmiştir. Bu noktada patoloji biliminin katkısı oldukça önemlidir.

Patoloji bilimi, kanserin erken teşhis edilmesi için uğraş veren bir bilim dalıdır. Patoloji kelimesi, hastalık anlamına gelen 'pathos' ve bilim anlamına gelen 'logos' kelimelerinin birleşiminden meydana gelmiştir. Hastalıklar sonucu organlarda, hücre düzeyinde meydana gelen değişimleri incelemektedir. Uzman patolog mikroskop yardımıyla hücrelerin sahip olduğu görünüme göre hastalıklara tanı koymaktadır. Patolog, hastalık şüphesi taşıyan dokudan biyopsi ile aldığı örneği tespit, takip, bloklama, kesme ve

boyama gibi çeşitli patolojik önışlemden geçirdikten sonra mikroskop yardımıyla incelemektedir. Hastalıđa ait belirtilerin bulunup bulunmadığını tecrübe ve konsültasyonla belirlemeye çalışmaktadır. Son yıllarda yüksek çözünürlüklü görüntüleme cihazlarının geliştirilmesi uzman patoloğun dokuları daha az çaba ile incelemesini sağlamıştır. Mikroskoplarla karşılaştırıldığında sayısal görüntüleme cihazlarının pek çok avantajı mevcuttur. Mikroskopların aynı anda belirli bir sayıda kişinin gözlem yapmasına olanak tanınması ve uzaktaki bir uzmanın dokuyu inceleme şansının olmaması mikroskoplar açısından dezavantaj olarak sayılabilir. Buna karşın sayısal görüntüleme cihazları ile aynı anda yüzlerce kişinin dokuyu incelemesi sağlanabilir. Bunun yanında dokunun sayısal ortama aktarılmasından sonra uzak mesafede bulunan uzmanların da dokuyu inceleme şansı yüksek çözünürlüklü görüntüleme cihazlarının patolojiye katkısı olarak ifade edilebilir[6-9].

1.1 Literatür Özeti

Çeşitli patolojik önışlemlerden geçirilen ve sayısal ortama aktarılan görüntüler histopatolojik görüntü olarak adlandırılmaktadır. Histopatolojik görüntüler insan veya farklı bir canlıdan alınan dokuların büyütölmüş görüntüleri olarak ifade edilebilir. Bu görüntüler organizmanın farklı biyolojik koşullarına bağılı çeşitli karmaşık yapıları ve bilgileri içeren tıbbi görüntüler olarak ifade edilebilir. Patologlar hastalıđı teşhis etmeye çalışırken kanserin türüne bağılı olarak yüksek çözünürlüğe sahip görüntüde hücresel yapıların morfolojik özelliklerini, mitozlu hücre sayılarını veya hastalıktan dolayı dokuda gerçekleşen değışimi incelemektedirler. Patologlar tarafından yapılan bu incelemeler hücrelerin dokuların ve organların yapısal ve işlevsel olarak incelenip hastalıđa ait tanının konmasını içermektedir. Tanı koyma başarısı incelenen doku veya organın büyüklüğüne, hastalıđın sahip olduđu risk seviyesine ve uzmanın tecrübesine bağılı olarak değışkenlik göstermektedir. Gelişen görüntüleme teknolojisi ve görüntü işleme teknikleri ile birlikte patologlar tarafından zor ve zaman alıcı olan bu işlemin makineler tarafından otomatikleştirilmesi amaçlanmaktadır. Hücrelerin teker teker incelenmesi patologlar açısından zor ve zaman alıcı olabilmektedir. Bilgisayar destekli teşhis (BDT) sistemleri yardımıyla patologlar açısından zor ve zaman alıcı olan bu işlemin teknolojik cihazlarda yazılımsal olarak gerçekleştirilmesi mümkündür. Bu işleme Bilgisayar

Destekli Teşhis (BDT, Computer Aided Diagnosis-CAD) adı verilmektedir. BDT’de amaç sayısallaştırılmış histopatolojik görüntünün bilgisayar ortamında analizi yapılarak erken tanıya katkı sağlayan ikincil karar destek sistemleri oluşturmaktır [10-15].

Sayısal görüntü işleme, görüntüleme teknolojisine sahip cihazlardan elde edilen yüksek boyutlu görüntülerin bilgisayar ortamına aktarıldıktan sonra çeşitli işlem ve analizlerin gerçekleştirilmesi olarak ifade edilmektedir [16]. Bu işlem ve analizler düşük seviye olarak ifade edilebilen, görüntülerin taranması ile ilgili çeşitli problemlerin giderilmesi olabildiği gibi, orta ve yüksek seviyeli işlemler olan hedef bir nesnenin (hücreSEL yapı veya yağ dokunun tespiti gibi) veya hedef bölgenin (tümörlü bölgeler gibi) tespiti ve takibini amaçlayan işlemler de olabilmektedir [17-19].

Görüntü işleme tekniklerinin hastalık şüphesi bulunan doku veya organların tanısında yardımcı yöntem olarak kullanılması amaçlanan çalışmalarda, ilk olarak ilgi bölgesinin belirlenmesi gerekmektedir. Günümüzde kullanılan mevcut sistemlerde ilgi bölgesinin belirlenmesi patolog tarafından çoğu zaman manuel olarak gerçekleştirilmektedir. Daha sonra şüpheli bölgelerde yer alan hücre ve dokular analiz edilerek hastalığa ait teşhisin konması amaçlanmaktadır. Görüntülerin yüksek boyutlu olması ve sayıca fazlalığı düşünüldüğünde bu işlem patologlar için iş yükü oluşturmaktadır. Bu işlem BDT sistemler tarafından otomatik olarak gerçekleştirildiğinde uzmana büyük kolaylık sağlanacaktır [20-21].

İlgi bölgesinin belirlenmesi işleminden sonra, patologlar tarafından tanı konarken dikkat edilen en önemli hususlardan biri de hücreSEL yapılarıdır. Kanseri olduğu düşünülen bölgelerdeki hücreSEL yapılar sağlıklı dokulardaki hücreSEL yapılardan morfolojik olarak farklılık göstermektedir. Bu hücreSEL yapıların patologlar tarafından teker teker incelenmesi zor ve zaman alıcı bir işlemdir. Bunun için görüntü işleme tekniklerinden bölütleme (segmentasyon) işlemi BDT sistemlerinin hedeflediği temel işlemlerden biridir (Bölüm 4). HücreSEL yapı olduğu düşünülen piksel gruplarının yüksek doğruluk oranı ile bölütlendikten sonra tanı aşamasında bu bilgilerin kullanılması hedeflenmektedir.

Literatürde hücreSEL yapıların bölütlenmesi ile ilgili pek çok farklı yaklaşıma sahip çalışma gerçekleştirilmiştir [22-30]. Cheng Lu vd. çekirdek alanlarındaki görüntü

bölgeleri bastıran ve parlaklık değişkenliklerini dengeleyen gürbüz bir çekirdek bölütleme algoritması geliştirmişlerdir [31]. Çalışmalarında ilgili bölgeye adaptif eşikleme uygulayarak çekirdek yapılarının birbirinden ayrıştırılması işlemini gerçekleştirmişlerdir. Bir başka çalışmada Hongming Xy vd. Lu vd. benzer olarak adaptif bir eşikleme yöntemi ile ilk bölütleme işlemlerini gerçekleştirmektedirler. Daha sonra bölütlenmiş görüntüye eliptik eğri uydurarak dairesele yakın çekirdekler tespit edilmeye çalışılmıştır [32]. Farklı bir çalışmada Al-Lahham vd. Ki-67 ile boyanan göğüs kanseri görüntülerinin RGB renk uzayından $L^*a^*b^*$ renk uzayına dönüştürüldükten sonra a^*b^* bantlarına k-ortalamar algoritmasını uygulamışlardır. Son olarak görüntüde global eşikleme ve bir takım morfolojik işlemler gerçekleştirerek hücresel yapıların bölütlenmesi işlemini gerçekleştirmişlerdir [33]. Belirtilen çalışmalar renk temelli gerçekleştirilen işlemler olduğundan her görüntüde yüksek oranlarda başarı sağlamaları güç olabilmektedir. Özellikle hücrelerin üst üste gelmesi durumunda yukarıda bahsedilen yöntemlerin başarılı olma ihtimali oldukça düşüktür. Bu yüzden literatürde hücrelerin üst üste gelmesi konusunda watershed ve graf tabanlı yöntemler uygulanmıştır [34-37].

Bir önceki paragrafta vurgulanan hücresel yapıların analizi BDT sistemleri tarafından gerçekleştirilen bir yöntem olsa da son yıllarda, kanserli olduğu düşünülen dokuların bölgesel olarak incelendiği birçok çalışma da gerçekleştirilmiştir. Çünkü özellikle yüksek çözünürlüklü tarayıcıların geliştirilmesinden sonra her bir dokuda binlerce hücre yer alabilmekte ve bu hücrelerin teker teker analizi ve teşhise yönelik analizi (verilerin anlamlandırılması) zorlaştırabilmektedir. Kanser türünde hücresel yapılar bireysel olarak analiz edildiğinde kanserli bölgelerde yer alan hücrelerin diğer bölgelerde yer alan hücrelerden morfometrik olarak farklı olduğu gözlemlenebilir. Fakat görüntünün yüksek çözünürlüğe sahip olması hücrelerin teker teker analiz edilip bir çıktı üretmeyi zorlaştırabilmektedir. Bu yüzden literatürde histopatolojik görüntü analizinde hücresel yapıların incelenmesini gerçekleştiren çalışmalardan farklı olarak bölgesel bilgilerin de işin içine dâhil edilip tümörlü veya kanserli bölgelerin tespitini gerçekleştiren yöntemler (Bölüm 5.4 ve Bölüm 5.5) de yer almaktadır [38-43].

1.2 Tezin Amacı

Dünya Sağlık Örgütü (DSÖ) verilerine göre kanser dünya çapında ölümlerin ikinci önde gelen nedeni olarak gösterilmektedir. 2018 yılında kansere bağlı ölüm sayısının 9.6 milyon olması beklenmektedir. Dünyada yaklaşık olarak 6 ölümden 1'i kanserden kaynaklanmaktadır. Erken teşhis edilir ve gerekli tedavi gerçekleştirilirse birçok kanserin tedavi şansı yüksektir. Bu noktada patoloji uzmanlarının kanserin teşhis edilmesindeki rolü yadsınamayacak bir gerçektir.

Günümüzde yüksek çözünürlüklü görüntüleme teknolojilerinin gelişmesi ile birlikte kanserin erken teşhis edilmesi daha olası hale gelmiştir. Görüntüleme cihazları hastalığa ait değişimlerin daha detaylı incelenmesini mümkün kıldığından kanserin erken teşhis edilmesinde hayati bir öneme sahiptir. Özellikle bilgisayar destekli teşhis (BDT) sistemleri ve yapay zekâ uygulamalarının bu alanda kullanılmaya başlanması hastalıkların tespiti, seyrinin izlenmesi ve sonuç olarak gerekli ve uygun tedavinin başlatılabilmesini sağlamıştır. Sayısal patoloji olarak genel bir isimlendirme yapılabilmekle beraber, tanı, eğitim, konsültasyon, arşivleme ve morfometrik analizler patolojiye katkı sağlamaktadır.

Çalışmamızda yüksek çözünürlüklü tarayıcılar yardımıyla elde edilen çeşitli kanser türlerine ait veri kümelerinin (Bölüm 3) BDT sistemleri ve makine öğrenmesi yöntemleri ile analizinin gerçekleştirilmesi amaçlanmaktadır. Böylelikle kanserin erken teşhis edilmesinde patoloji uzmanlarına ikincil karar destek sistemlerinin geliştirilmesine katkı sağlanması hedeflenmektedir. Çalışma kapsamında literatürde çıkıntı (saliency detection) tespiti, nesne bölütlenmesi gibi çalışmalarda yaygın olarak kullanılan görüntü işleme yöntemlerinden biri olan süperpiksel bölütleme yönteminin hücresel yapıların bölütlenmesindeki katkısının incelenmesi amaçlanmaktadır. Süperpiksel, birbirleriyle anlamlı bütünlük oluşturan piksel grubuna verilen isim olarak ifade edilebilir. Hücresel yapıların çevrelendikleri yağ ve bağ dokulardan daha koyu olduğu düşünüldüğünde, süperpiksel bölütleme yönteminin hücresel yapıların ayırte edilmesine önemli bir katkısının olacağı düşünülmektedir. Bölütleme işleminin sağlıklı yapılabilmesi daha sonra gerçekleştirilecek sınıflandırma aşaması öncesi hayati bir öneme sahiptir. Yüksek başarı ile gerçekleştirilen bölütleme işleminin sınıflandırma

başarısını da arttıracaktır. Histopatolojik görüntülerde hücresel yapıların bölütlenmesinden sonra kanserin agresifliğinin belirlenmesi açısından mitozlu hücrelerin tespiti hayati bir öneme sahiptir. Bu bağlamda literatürde son dönemde oldukça popüler olan evrişimsel sinir ağları (ESA) yöntemi kullanılarak mitozlu hücrelerin tespit edilmesi amaçlanmaktadır. Önerilecek ESA yönteminin literatürde en çok kullanılan geleneksel özellik çıkarma algoritmaları ile karşılaştırılması da amaçlanmaktadır.

Çalışmanın bir diğer kısmında bölütlenen hücresel yapıların morfometrik analizlerinin yapılabilmesi için çeşitli istatistiksel özelliklerinin çıkarılması ve sınıflandırma işleminin gerçekleştirilmesi hedeflenmektedir. Bunun için çalışma kapsamında hazırlanan rahim ağzı kanserine (serviks) ait yüksek çözünürlüklü görüntü veri kümesi bu amaçla kullanılmıştır. Patoloji uzmanlarının belirli bir kanser türünde (Ör: Serviks) benzer dokuya farklı tanıyı koyması söz konusu olabilmektedir. Bu durum gözlemciler arası uyum/uyumsuzluk olarak ifade edilmektedir. Çalışma kapsamında hazırlanan veri kümesinin iki farklı patoloji uzmanı tarafından derecelendirilmesi sonucu patologlar arası uyum/uyumsuzluk durumu incelenecektir. Gerçekleştirilecek yöntemin patologlarla uyum/uyumsuzluğu da araştırılması amaçlanan bir diğer konudur.

1.3 Hipotez

Kanserlerin %30-50'si risk faktörlerinden kaçınarak ve mevcut kanıta dayalı önleme stratejileri uygulanarak önlenebilmektedir. Kanser yükü, kanserin erken teşhisi ve kanser geliştiren hastaların tedavisi yoluyla da azaltılabilir. Kanser en temelde hücrede meydana geldiği ve hücrede çeşitli değişimlere sebep olduğu düşünüldüğünde hücrelerin analizi kanserin tespiti noktasında oldukça önemli bir yere sahiptir.

Patoloji uzmanları mevcut durumda mikroskop kullanarak hastalığa ait tanıyı koymaya çalışmaktadır. Sayısal patolojide ise doku örneklerinin sayısal ortama aktarılması ve burada hem konsültasyon yapılması hem de gerekirse eğitimin yapılması amaçlanmaktadır. Bu durum tecrübenin daha fazla kişiye aktarılmasının önünü açacaktır. Mikroskop kullanılarak gerçekleştirilen tanı yaklaşımında uzmanların fiziksel olarak mikroskopu kullanması gerekirken uzakta olan uzmanın veriyi inceleyebilmesi

mümkün değildir. Sayısal patolojide ise farklı bir yerde olan uzmanın çevrimiçi ortamda dokuyu incelemesi mümkündür. Sayısal patoloji birçok açıdan avantaj sağlamaktadır.

Çalışmada hücrel yapıların bölütlenmesi amacıyla kullanılacak kümeleme tabanlı bölütleme algoritmalarına ek olarak süperpiksel tabanlı bölütleme algoritmaları ve derin öğrenme tabanlı algoritmalar kullanılacaktır. Süperpiksel bölütleme algoritmaları teoride parlaklık bilgilerinin yanında, koordinat bilgilerini de özellik olarak değerlendirdiğinden birbirlerine yakın piksel gruplarının anlamlı bütünlük halinde aynı süperpiksele dâhil olması beklenmektedir. Bu durum süperpiksellerin hücrel yapıların bölütlenmesinde literatürde kullanılan yöntemlere nispeten daha iyi sonuç verebileceğini göstermektedir. Süperpikselin ön bölütleme yöntemi olarak kullanılması hücrel yapıların bölütlenmesinde literatürdeki mevcut bölütleme algoritmalarının bölütleme başarımına katkıda bulunacağı düşünülmektedir.

Buna ek olarak özellikle son dönemde sıkça kullanılan derin öğrenme tabanlı algoritmaların hem bölütleme hem de sınıflandırma aşamalarında kullanılması amaçlanmaktadır. Derin öğrenme tabanlı yöntemlerin birçok katman ve birçok parametreye bağlı kendine has yapısı onu literatürde mevcut kullanılan şekil tabanlı, doku tabanlı veya istatistik tabanlı geleneksel yöntemlerden ayırmaktadır. Bu çalışmada gerçekleştirilecek derin öğrenme tabanlı yöntemlerin veriyi sezgisel olarak öğrenmesinden dolayı literatürdeki geleneksel yöntemlerden daha iyi sonuç vereceği düşünülmektedir. Ayrıca histopatolojik görüntü veri kümelerinde kullanılan yaygın bir veri kümesi olmadığından elde edilen başarılar çalışmaya özgü olmaktadır. Bu çalışmada rahim ağzı kanseri ile ilgili literatürdeki en geniş veri kümelerinden biri oluşturulmaya çalışılmıştır. Bu veri kümesinin paylaşılması düşünülmektedir. Böylece literatürde bilim insanları tarafından yaygın olarak kullanılan bir veri kümesi oluşturulması düşünülmektedir.

BÖLÜM 2

SAYISAL PATOLOJİ

Patoloji, hastalıklara yol açan temel sebepleri araştırırken doku ve organlarda meydana gelen morfolojik değişimleri de gözlemlemektedir. Bu anlamda tıbbın temelini oluşturmaktadır. Patoloji, klinik bir dal olarak görülmesi de işleyişi bakımından tanı ve tedaviye doğrudan etki etmesi bakımından klinik dallara katkı sağlamaktadır.

2.1 Patolojik Önlemler

Uzman, hastalardan tarama ve tanı amacıyla alınan doku veya organların öncelikle makroskopik (dış görünüm) incelemesini gerçekleştirmektedir. Hastalıklı olduğu düşünülen doku veya organlardan mikroskopik incelemesi yapılmak üzere hastalığı en iyi yansıttığı düşünülen kesitler alınır. Alınan bu kesitler daha sonra mikroskopik inceleme yapılması amacıyla çeşitli patolojik ön işlemlerden geçirilmektedir [44].

2.1.1 Tespit (Fiksasyon)

Canlı vücudundan alınan dokularda yeralan hücreler belirli bir süre geçtikten sonra bakteriler veya içerisinde bulunan sindirici enzimler sebebiyle otolize uğramakta ve sahip olduğu morfolojik özellikleri kaybetmektedir. Bu durum hastalığın sebeplerinin belirlenmesini engellemektedir. Tespit, hücrelerin canlı vücudundan ayrıldıkları andaki morfolojik özelliklerini korumaları amacıyla gerçekleştirilen işleme denir. Tespit işleminde doku ve organlar, yapılarını korumaları amacıyla çeşitli özel sıvıların içerisine bırakılırlar. Bu amaç için en çok kullanılan sıvılardan biri formalindir. Dokular

kendilerinden 10-20 kat büyüklükte formalin sıvısının içerisinde dokunun hangi organa ait olduğuna bağlı olarak birkaç saat ile birkaç hafta bekletilmektedir.

2.1.2 Takip (Doku İşleme)

Tespit işleminden sonra elde edilen kesitler dehidratasyon işleminden geçirilmektedir. Dehidratasyon, kesitin tespit işleminde içine bırakıldıkları formalin sıvısından uzaklaştırılması işlemi olarak ifade edilebilir. Bu gerçekleştirilen işlemle doku kesitinin daha sert bir yapıya kavuşması sağlanır ve sonraki işlemlerde daha ince kesilebilmesini sağlar. Ksilol yardımıyla alkolün kesitten uzaklaşması sağlanmaktadır. Kesitler daha sonra aralıklarla ısıtılarak sıvılaştırılan parafine konmaktadır. Parafin oda sıcaklığında katışan bir madde olduğundan kesitler son halde oda sıcaklığında katı halde bulunmaktadır. Bu işlemlerden geçirilen bütün kesitler numaralanarak daha sonra yapılacak bütün işlemlerde raporlara işlenmesi sağlanmaktadır. Belirtilen bu işlemler oda sıcaklığı ile 60 C aralıklarında yapılmaktadır.

2.1.3 Bloklama

Bu aşamada kesitler dikdörtgen şeklindeki kalıplara konarak üzerlerine ısıtılmış parafin dökülür ve soğumaya bırakılır. Böylelikle dokular daha ince kesilebilir hale gelirler.

2.1.4 Kesme

Soğuyan parafin blokları mikrotom adı verilen alet ile yaklaşık olarak 4-5 mikron kalınlıklarında kesilmektedir. Kesilen dokular ılık su daldırılıp daha sonra lamaların üzerine konmaktadır.

2.1.5 Boyama

Patolojik işlemlerde rutin olarak kullanılan boyalar Hematoksilen ve Eosin (H&E) boyalarıdır. Bu boyalardan geçirilen kesitlerde çekirdekler mavi, çekirdeği saran sitoplazma ise kırmızı-pembe rengini almaktadır.

2.2 Sayısal Görüntü İşleme

Görüntü; insanın, başka bir canlı veya bir objenin güneş ışınlarını yansıtıktan sonra görme merkezinde oluşturduğu temsil olarak ifade edilebilir. Güneş ışınları normal şartlar altında 0.000001nm ila 100000000nm dalga boylarına sahiptir. İnsan gözü 400nm ile 700nm dalga boylarına sahip güneş ışığını algılayabilmektedir. Farklı canlılar farklı dalga boylarındaki ışığı algılama özelliğine sahiptir.

2.2.1 Görüntünün Elde Edilmesi

Özellikle son 50 yılda bilgisayar ve elektronik dünyasında yaşanan teknolojik gelişmelerden sonra dış ortamda analog olarak bulunan ışığın sayısallaştırılarak bilgisayar ortamına aktarılması mümkün hale gelmiştir. Bunun gerçekleştirilmesinde sensörlerin (algılayıcı) önemi oldukça fazladır. Işığın farklı dalga boylarını algılayabilen sensörler sayesinde nesnelere ait görüntüler sayısal ortamda saklanabilmektedir. Sayısal bir görüntü analog işaretin belirli aralıklarla örneklenmesi ile elde edilmektedir. $f(x,y)$ sayısal ortama aktarılan bir görüntüyü ifade etsin. x ve y iki boyutlu düzlemde yatay ve dikey koordinat bilgisini, f ise x ve y 'nin değişen değerlerine karşılık gelen genliği (parlaklık) ifade etmektedir. Sayısal bir görüntüde parlaklık değeri 0-255 arasındadır. 0 değeri siyahı, 255 değeri ise beyazı ifade eder. Kırmızı-Yeşil-Mavi(KYM) formatında saklanan renkli bir görüntü 3 boyutlu bir matris ile ifade edilmektedir. $f(x,y,z)$ 3 boyutlu matrisi ifade edecek olursa x ve y matriste yer alan koordinat bilgisini z ise bant bilgisini (kırmızı, yeşil veya mavi) ifade etmektedir. Örneğin $f(1,1,1) = 255$, $f(1,1,2) = 0$, $f(1,1,3) = 0$ değerlerinde ilgili pikselin kırmızı renkte görünmesi anlamına gelecektir.

2.2.2 Önileme - Görüntünün İyileştirilmesi

Kamera, mikroskop veya tarayıcı yardımıyla elde edilen görüntüler kimi zaman görüntüleme cihazından kimi zaman da iletim sırasında fiziksel ortamdan kaynaklı çeşitli artefaktlar içermektedir. Görüntünün cinsinden bağımsız olarak kalitenin artırılması veya gürültünün giderilmesi amacıyla görüntülere uygulanan çeşitli önileme algoritmaları vardır. Artefaktın çeşidine göre önilemede uygulanan süzgecin cinsi de değişiklik göstermektedir. Gürültüde yüksek frekanslı gürültüler yer alıyorsa bunu

çeşitli yumuşatma süzgeçleriyle giderek mümkündür. Yumuşatma süzgeçlerine örnek olarak ortalama, ortanca, gauss örnek verilebilir. Bunun dışında ortamın ışık yoğunluğundan kaynaklı olarak karanlık veya aşırı parlak görüntülerde ters işlem uygulamak da söz konusu olabilmektedir. Böylece görüntü işlemede gerçekleştirilmek istenen işlem için görüntü önceden hazırlanmış olmaktadır.

2.2.3 Bölütleme

Görüntü işleme tekniklerinden biri olan bölütleme işleminde hedef objenin düzlemde veya görüntüde yer alan diğer nesnelerden ayırtedilmesidir. Bu durum elde edilmek istenen hedef objenin dışındaki bilgilerin elimine edilmesi anlamına da gelebilmektedir. Ayırtedilmek istenen nesnenin (veya piksel grubunun) bir sonraki işlemde daha detaylı analizinin yapılabilmesi için öncelikle elde edilmesi gerekmektedir. Kümeleme tabanlı, süperpiksel-tabanlı, çizge-tabanlı ve son dönemde oldukça popüler bir yaklaşım olan semantik bölütleme tabanlı yöntemler literatürde sıkça kullanılagelen yaklaşımlardır.

2.2.4 Öznitelik Çıkarımı

Bir önceki başlıkta analiz edilmek istenen hedef objenin elde edilmesi için çeşitli işlemler uygulanmasından bahsedilmişti. Bu objenin elde edilmesi için objeyi görüntüdeki diğer objelerden ayıran belirli özellikleri mevcuttur. Objenin türüne bağlı olarak sadece sahip olduğu renkten ayırt edilebileceği gibi, sahip olduğu desen veya şekilden de ayırt edilebilir. Bu durum objeye ait olan özelliklerin çıkarılması ile sağlanabilir. Renk, şekil, doku(desen) öznitelikleri kullanılarak objeler ayırt edilebilmektedir. Literatürde son dönemde oldukça popüler olan derin öğrenme tabanlı öznitelik çıkarma yöntemleri de uygulanagelmıştır.

2.2.5 Sınıflandırma

Farklı sınıflara ait olduğu bilinen çeşitli öznitelik vektörlerinin ayırtedilmesini sağlayan algoritmalar. Benzer öznitelik vektörlerine sahip örnekler bu algoritmalar tarafından aynı gruptaymış gibi değerlendirilir. Farklı sınıflara ait birçok örnek ile eğitilen bir algoritmaya test olarak verilen bir obje yine ayırt edici olduğunu düşünülen öznitelik

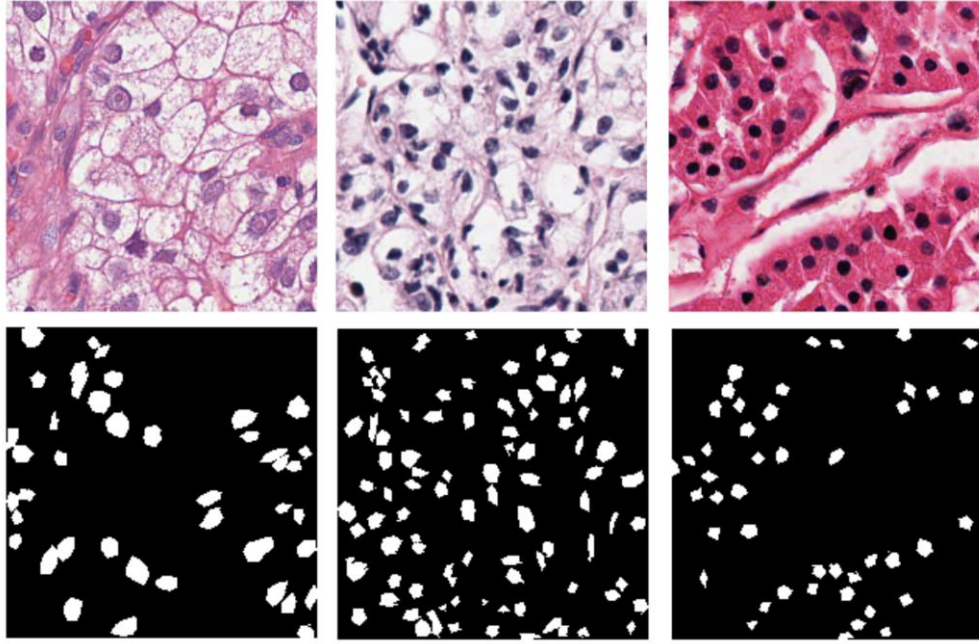
vektörüne göre bütün gruplara yakınlığı hesaplanır. Hangi gruba yakın ise sınıflandırma algoritması tarafından o gruba aitmiş gibi etiketlenir.

KULLANILAN VERİ KÜMELERİ

Çalışma kapsamında 4 farklı veri kümesi kullanılmıştır. İki adet veri kümesi makine öğrenmesi yöntemlerinin bölütleme işlemindeki başarılarının değerlendirilmesi amacıyla kullanılmıştır. Bir adet veri kümesi histopatolojik görüntülerde kanser derecelendirme işlemine çok önemli katkısı olan mitoz belirleme için oluşturulan veri kümesidir. Bu çalışma kapsamında oluşturulan veri kümesi ise rahim ağzı kanserinin yüksek çözünürlüklü histopatolojik görüntülerinde derecelendirme işlemi BDT ve makine öğrenmesi yöntemleri ile gerçekleştirmek amacıyla oluşturduğumuz veri kümesidir.

3.1 Böbrek Hücre Karsinomu (Renall Cell Carcinoma) Veri Kümesi

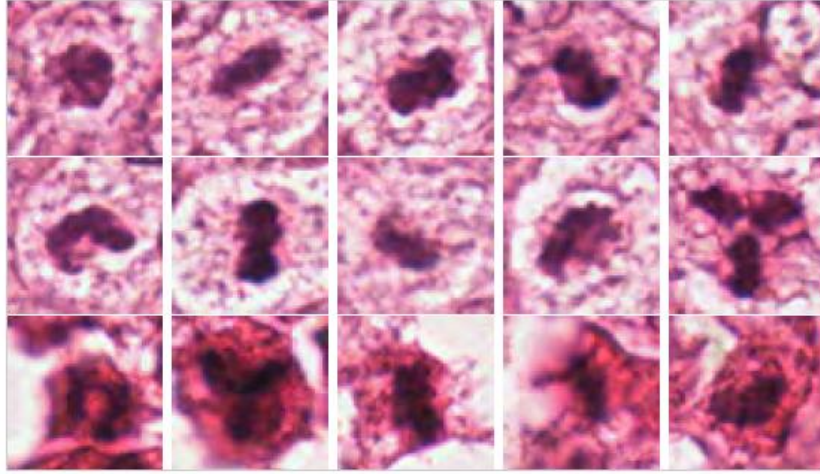
Bu çalışmada kullanılan veri seti, Harvard Üniversitesi'nde Beck Laboratuvarı'ndan alınmıştır. Veri kümesi, The Cancer Genome Atlas (TCGA) veri portalından seçilen böbrek karsinomunun yüksek çözünürlüklü histopatolojik görüntülerini içermektedir [45]. Veri kümesinde yer alan 10 adet bütün slayt (whole slide) 40x çözünürlükte taranmıştır. Görüntüler daha sonra 400x400 boyutlarında görüntü kesitlerine kırılmış ve her görüntüdeki her hücresel yapı patoloji uzmanları tarafından etiketlenmiştir. Bu çalışmada, algoritmaların değerlendirilmesi için patologlar tarafından etiketlenen 64 adet görüntü kullanılmıştır. Şekil 3.1, veri kümesinden elde edilen üç adet yüksek çözünürlüklü histopatolojik görüntü ve bu iki görüntüye ait uzman patologlar tarafından işaretlenen kesin referans (ground truth) görüntüsünü ifade etmektedir.



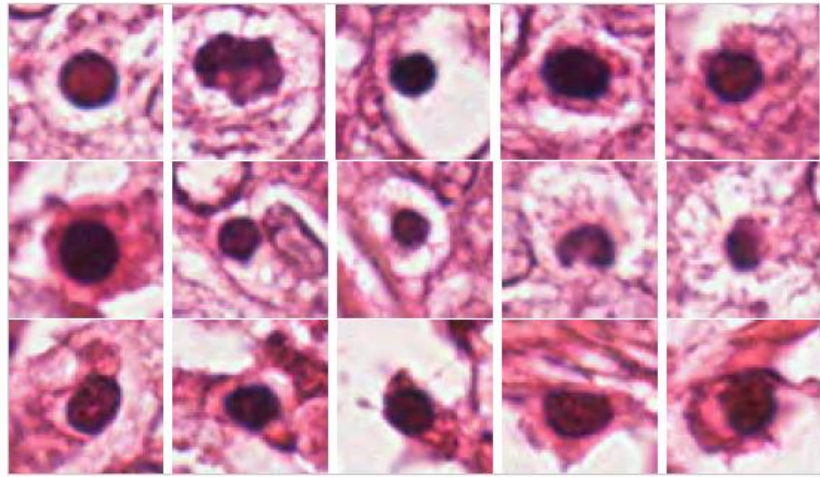
Şekil 3.1 Veri kümesinden elde edilen üç adet yüksek çözünürlüklü histopatolojik görüntü ve bu iki görüntüye ait uzman patoloğlar tarafından işaretlenen kesin referans (ground truth) görüntüsü

3.2 ICPR2014 - MITOS ATYPIA 14 Veri Kümesi

Çalışma kapsamında International Conference on Pattern Recognition (ICPR) tarafından 2014 yılında organize edilen ve meme kanseri teşhisi için önemli bir parametre olan mitozlu hücre tespitini amaçlayan veri kümesi kullanılmıştır [46]. Yarışma iki aşamadan oluşmaktadır: atipik çekirdeklerin tespiti ve mitozlu hücrelerin tespiti. Hematoxylin ve Eosin (H&E) boyaları ile boyanmış yüksek çözünürlüklü histopatolojik görüntüler Aperio Scanscope XT ve Hamamatsu Nanozoomer 2.0-HT tarayıcıları ile taranmıştır. Mitoz belirleme için kullanılan veri kümesi 40x çözünürlüğe sahiptir. Şekil 3.2 veri kümesinden elde edilen ve mitozlu ile mitozsuz hücreleri ifade eden görüntüleri göstermektedir. Veri kümesinde 749 adet mitozlu ve yaklaşık olarak 180000 adet mitozsuz hücre olmak üzere toplamda yaklaşık 181000 hücre bulunmaktadır.



(a)



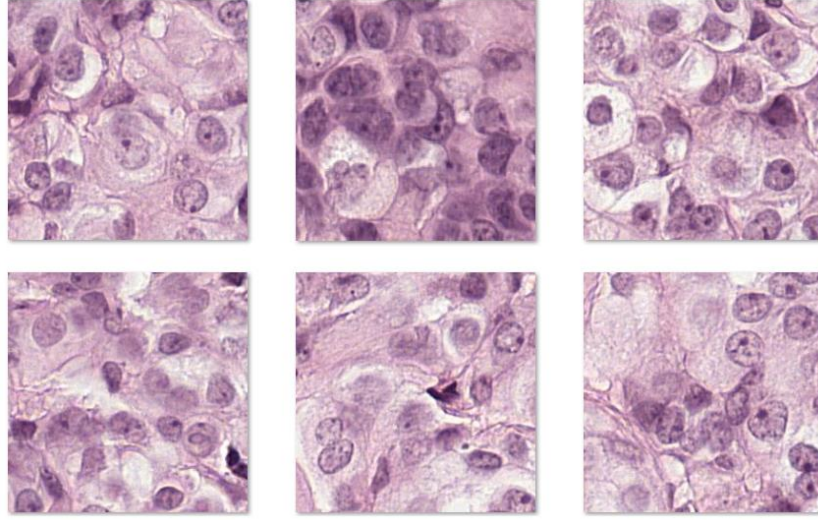
(b)

Şekil 3.2 ICPR2014 veri kümesindeki görüntülerden elde edilen çeşitli sayıda (a) mitozlu hücre ve (b) mitozsuz hücre örnekleri

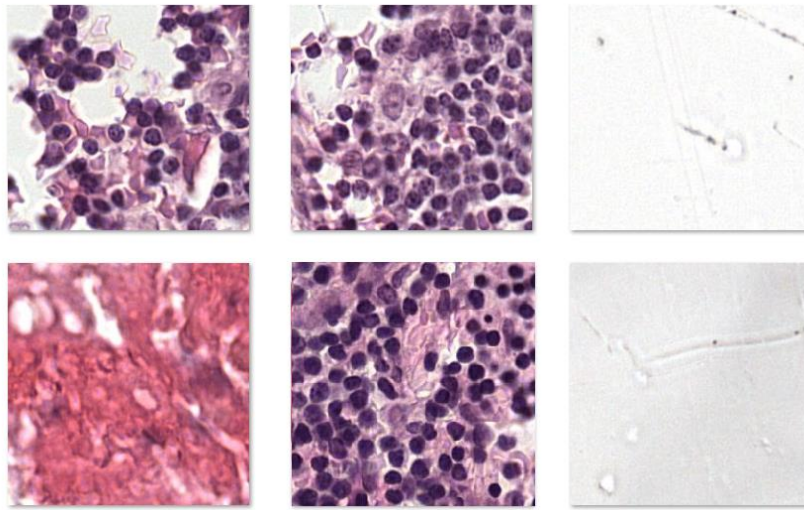
3.3 Camelyon Veri Kümesi - Lenf Düğümünde Metastaz Tespiti

Uygulamada kullanılan veri kümesi Radboud University Medical Center (Nijmegen, the Netherlands) ve University Medical Center Utrecht (Utrecht, the Netherlands) tarafından sentinel lenf düğümlerindeki metastaz bölgelerinin tespitinin yapılması amacıyla hazırlanmıştır. Veri kümesinde yer alan bütün görüntüler, patoloji uzmanları tarafından işaretlenmiş yüksek çözünürlüklü ve büyük boyutlu histopatolojik görüntülerden oluşmaktadır. Veri kümesinde yer alan görüntüler hematoxylin ve eosin

(H&E) kimyasal boyaları ile boyanmış ortalama 100.000x200.000 genişlik ve yüksekliğe sahiptir. Şekil 3.3 veri kümesinden alınan örnek bir görüntüyü ifade etmektedir [47].



(a)



(b)

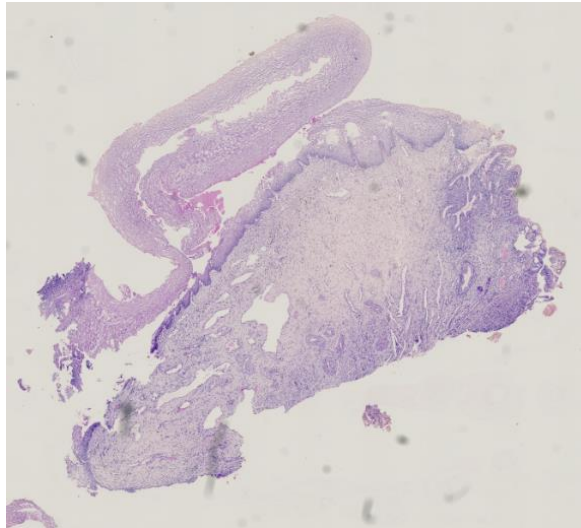
Şekil 3.3 Camelyon16 veri kümesindeki görüntülerden elde edilen (a) metastaz bölgeleri ve (b) metastaz olmayan(kan doku, normal hücreler ve arkaplan cam bölgesi) bölgelerden elde edilen görüntü kesitleri

3.4 Çalışmada Kullanılmak Üzere Oluşturulan Serviks (Rahim Ağız Kanseri) Veri Kümesi

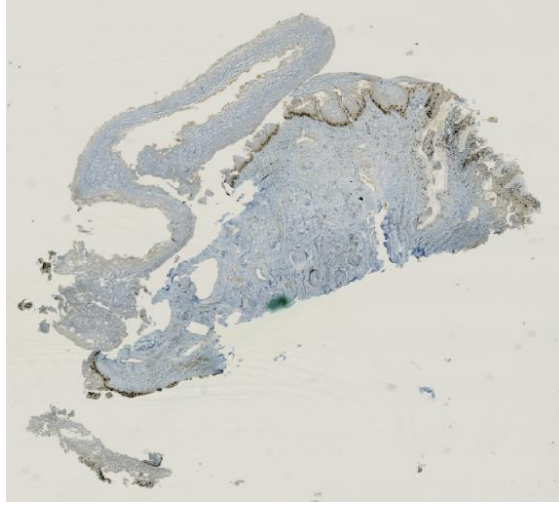
Yüksek çözünürlüklü büyük boyutlu rahim ağzı kanseri öncü lezyonlarından oluşan bu veri kümesi 54 hastaya ait 118 adet lam'dan meydana gelmektedir. Tüm lamlardan 127 adet yüksek çözünürlüklü görüntü, tüm görüntülerden 320 adet büyük epitel parçası(BEP), BEP'lerin daha küçük parçalara ayrılması ile de 957 adet küçük epitel parçaları (KEP) elde edilmiştir. Elde edilen 957 KEP her iki patolog tarafından ayrı ayrı incelenmiştir.

Yüksek çözünürlüklü görüntülerin, hematoksilin-eosin (HE) ve immünohistokimyasal (İHK) incelemeleri bir arada yorumlanarak kabul edilen kesin tanıların KEP'lere göre tanı dağılımı üçlü sistemde; 471 adet (%49,2) normal, 240 adet (%25,1) CIN1, 107 adet (%11,2) CIN2, 139 adet (%14,5) CIN3 ve BEP'lere göre dağılımı 150 adet (%46,9) normal, 79 adet (%24,7) CIN1, 34 adet (%10,6) CIN2, 57 adet (%17,8) CIN3'dür.

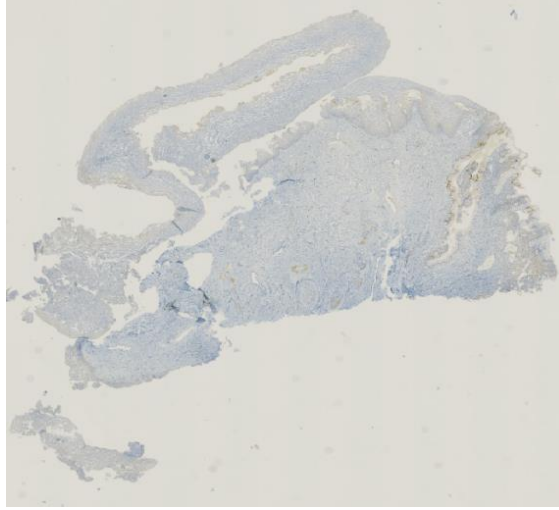
Aynı yaklaşım ile kesin tanıların ikili sisteme göre dağılımı 471 adet (%49,2) normal, 240 adet (%25,1) LSIL ve 246 adet (%25,7) HSIL; BEP'lere göre dağılımı 150 adet (%46,9) normal, 79 adet (%24,7) LSIL, 91 adet (%28,4) HSIL'dir. Veri kümesi akademik anlamda kullanılmak üzere [48] adresinden talep edilebilmektedir.



(a)



(b)



(c)

Şekil 3.4 Rahim ağzı kanser öncüsü lezyonlarından oluşturulan veri kümesindeki görüntülerden elde edilen (a) yüksek çözünürlüklü HE ile boyanmış histopatolojik lezyon örneği (b) Bu örneğin Ki67 İHK boyası ile boyanmış örneği ve (c) aynı lezyonun p16 İHK boyası ile boyanmış örneği

HİSTOPATOLOJİK GÖRÜNTÜLERİN BÖLÜTLENMESİ

4.1 Histopatolojik Görüntülerde Renk Uzaylarının Hücre Bölütlenmesine Etkileri

Bilgisayar Destekli Teşhis (BDT) sistemlerinin gelişmesiyle birlikte yüksek çözünürlüklü histopatolojik görüntülerin analizi daha kolay hale gelmektedir. Bu çalışma kapsamında literatürde farklı çalışmalarda kullanılan renk uzaylarının histopatolojik görüntülerde hücresel yapıların arka plandan ayırt edilerek belirlenmesindeki etkileri araştırılmıştır. Bu amaç doğrultusunda hücresel yapıların bölütlenmesinde K-Ortalamlar, bulanık c-ortalamlar ve beklenti-ençoklama algoritmalarının farklı renk uzaylarındaki bölütleme performansları karşılaştırılmıştır. Sonuç bölümünde çeşitli nesnel ölçütlere göre elde edilen bölütleme doğruluğu sonuçları karşılaştırılmalı olarak sunulmuştur.

4.1.1 Görüntülerin Farklı Renk Uzaylarına Dönüştürülüp Bölütlenmesi

Renk uzayları, sayısal ortamda insan gözünün algıladığı renklerin ifade edilmesi için kullanılan matematiksel modellerdir. Renk uzayları teknik olarak doğada bulunan bütün renkleri temsil edecek şekilde oluşturulmaktadır.

4.1.1.1 Renk Uzayları

Genel itibarıyla birçok renk uzayında üç temel renk kullanılmaktadır. Renk çeşitliliği de bu renklerin her birinin çeşitli oranlarda değiştirilmesi ile sağlanmaktadır.

KYM (Kırmızı-Yeşil-Mavi, RGB) Renk Uzayı: Monitörler tarayıcılar gibi cihazlarda renkleri ifade edebilmek için en çok kullanılan renk uzayıdır. Temelde Kırmızı, Yeşil ve

Mavi renklerini kübün bir köşegeni olarak düşünecek olunursa bu renklerin belirli oranlarda karışımı farklı renkleri meydana getirecektir. Her üç rengin %100 oranında karıştırılması Beyaz'ı, %0 oranında karıştırılması da Siyah'ı meydana getirmektedir.

HSV Renk Uzayı: HSV renk uzayı Hue(Renk özü), Saturation (doygunluk) ve Value (değer) değişkenlerinden meydana gelmektedir. Hue rengin çeşidine veya tonuna göre rengi diğer renklerden ayıran niteliktir. Saturation, bir rengin ne kadar doygunlukta kullanıldığını belirtmektedir. Value ise açık olan bir rengin koyu bir renkten ayırt edilmesi amacıyla kullanılmaktadır.

La*b* Renk Uzayı: La*b* renk uzayının bileşenleri L (lightness-ışıklılık), a (tonlama) ve b (doygunluk) olarak ifade edilmektedir. La*b* renk uzayının en büyük özelliği insan algılaması yönünden düzgün dağılım göstermesidir. Düzgün dağılım göstermesinden dolayı literatürde yer alan birçok histopatolojik görüntü analizi çalışmalarında La*b* renk uzayı kullanılmaktadır.

XYZ Renk Uzayı: XYZ renk uzayı 1931 yılında CIE tarafından tasarlanan ve tüm renklerin temel yapısını oluşturan bir renk uzayıdır. X, Y ve Z renkleri beyinde yorumlanan KYM(Kırmızı, Yeşil, Mavi - RGB) renklerini ileten sinirlerin beyine ilettikleri uyarıların toplamı olarak ifade edilebilir. Burada her bir X, Y ve Z uyarımlarının tüm toplama oranları renkleri tanımlamaktadır. XYZ renk uzayı farklı farklı beyaz referansları baz alınarak Bu çalışma kapsamında D65 referans numaralı sRGB renk uzayı kullanılmıştır.

4.1.1.2 Kullanılan Bölütleme Yöntemleri

K-Ortalamalar (K-means): Bölütleme amacıyla kullanılan yöntemlerinden biridir. Bu yöntemde belirlenen sınıf sayısı kadar küme merkezine tüm örneklerin bu merkezlere olan mesafeleri hesaplanmaktadır. Hesaplanan uzaklıklara göre küme merkezleri güncellenerek örnekler kümelenmektedir [49]. Buna göre K-ortalamalarda gerçekleştirilen işlem adımları şu şekildedir:

- Küme merkezlerinin belirlenmesi
- Diğer örneklerin mesafelerine göre kümelenmesi
- Gerçekleştirilen sınıflandırmaya göre yeni merkezlerin hesaplanması

- Belirli bir iterasyon sayısı kadar veya küme merkezleri değişmeyene kadar 2. ve 3. adımların tekrarlanması

Bulanık C-Ortalamları(Fuzzy C-Means): Bulanık C-Ortalamlar (BCO) algoritması ilk defa 1973 yılında Dunn tarafından önerilmiş ve Bezdek tarafından geliştirilmiştir [50]. Bu algortmada kümelenecek veri seti için c adet küme sayısı belirlenir. Bu c sayısı 1'den büyük ve görüntüdeki N adet piksel sayısından küçük olmalıdır. Kümelemenin bulanıklığını belirleyen bulanıklık ağırlık üsteli m' ye 1'den büyük bir değ er atanır.

$$J(\mathbf{X}, \mathbf{U}, \mathbf{V}) = \sum_{j=1}^c J_j = \sum_{i=1}^N \sum_{j=1}^c (\mu_{ij})^m \|\mathbf{x}_i - \mathbf{v}_j\|_G^2 \quad (4.1)$$

Bu ifadede yer alan X görüntüde yer alan N adet pikselden oluşan matrisi göstermektedir. U, X'in bulanık üyelik matrisini ifade etmektedir. U'nun her bir elemanı kesin olarak 0 veya 1 yerine bu sayılar arasında değ erler alarak hangi kümeye hangi oranlarda dâhil olduklarını göstermektedir. Bu durum k-ortalamarından ayrılan en önemlik özelliktir. Bir piksele ait bulanık aitlik oranları toplamı 1'e eşittir. Hedef fonksiyonunda yer alan V ifadesi c adet küme merkezi değ erlerinin bulunduğu d uzunluklu vektörleri göstermektedir. Bulanık C-ortalamlar algoritmasının işlem adımları şu şekildedir:

- Başlangıçta U^0 üyelik matrisi rastgele değ erlerle oluşturulur, bulanık ağırlık üstelik m, durma şartı için eşik değ eri ϵ ve küme sayısı c belirlenir.
- Küme merkezleri belirlenir.
- Küme merkezleri ile her bir piksel arasındaki mesafe hesaplanır.
- Üyelik matrisi U^{t+1} güncellenir.
- Bir önceki adımdaki üyelik matrisi ile güncellenmiş üyelik matrisi arasındaki fark belirlenen eşik değ erinin üzerindeyse Adım 2'den aynı işlemler tekrarlanır.

Beklenti Ençoklama Algoritması (EM Algorithm): Beklenti ençoklama algoritması da K-ortalamlar gibi bölütleme amacıyla kullanılan temel algoritmalardan biridir. K-ortalamalardan ayrılan en belirgin özelliği bir küme merkezi ile karşılaştırılan örneklerin

kesin olarak bir kümeye atanamamasıdır. Bunun yerine örneklerin yine belirli oranlarda küme merkezlerine aidiyetleri hesaplanır. Son durumda örneklere ait kümeler belirlenir. EM algoritması iteratif bir algoritmadır. EM algoritması beklenti (Expectation) ve ençoklama (maximization) adımlarından oluşur. Bu adımlar belirli bir yakınsama kriteri sağlanana kadar ardışık olarak gerçekleştirilir [51].

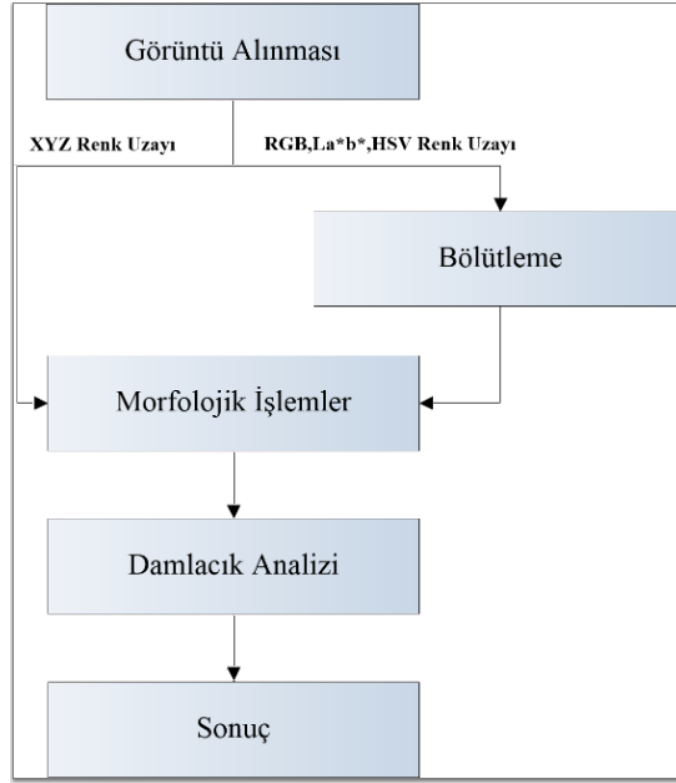
$$L(\theta, X, Z) = p(X, Z | \theta) \quad (4.2)$$

Denklem 4.2'te yer alan X değeri elde edilen veri kümesini, Z belirli olmayan verileri θ ise kestirim parametresini ifade etmektedir. Beklenti ençoklama yönteminin işlem adımları kabaca şu şekildedir:

- Öncelikle θ parametresi rastgele değerlerle ilklendirilir.
- En iyi Z değeri için verilen parametre değerleri kullanılarak hesaplama yapılır.
- Hesaplanan Z değeri daha iyi bir θ parametresi kestirimi için kullanılır.
- Belirli bir yakınsama değerine kadar 2. ve 3. adımları takip et.

4.1.2 Farklı Renk Uzaylarının Bölütlenmesi ile Elde Edilen Deneysel Sonuçlar

Uygulamada kullanılan veri kümesi Harvard Üniversitesi Beck Laboratuvarlarında uzman patoloğlar tarafından işaretlenmiş yüksek çözünürlüklü histopatolojik görüntüden oluşmaktadır. Çalışma kapsamında 64 adet hematoksilen ve Eosin(H&E) ile boyanmış böbrek renal kanseri görüntüleri kullanılmıştır.



Şekil 4.1 Çalışmada takip edilen işlem adımları

Şekil 4.1'de çalışmada takip edilen işlem adımları gösterilmiştir. Literatürde en çok kullanılan renk uzayları olan KYM, La*b*, HSV ve XYZ(D65) renk uzaylarının üç bandı öncelikle değerlendirilerek bölütleme başarımı hesaplanmıştır. Daha sonra her bir bandın başarımı da ayrıca hesaplanarak bir karşılaştırma yapılmıştır. Renk uzaylarının kendi içinde en başarılı sonucu veren bant veya bantların sonuçları sunulmuştur. Öncelikle KYM formatında verilen her bir görüntü farklı renk uzaylarına dönüştürülmüş daha sonra bu görüntüler K-ortalamlar, Bulanık C-Ortalamlar ve Beklenti-Ençoklama algoritmaları ile bölütlenmiştir. Bölütlenme sonucunda elde edilen görüntülerde hücrel yapılarıdaki aşınmanın önüne geçmek için 'kare' yapısal eleman kullanılarak genişleme işlemi gerçekleştirilmiştir. Son olarak bölütlenme sonucunda hücrel yapılarla aynı kümeye düşen fakat hücrel olmayan küçük parçacıklar damla analizi ile elimine edilmiştir. Eleme işlemi için ikili görüntüde yer alan en küçük hücrel yapının alanı eşik değeri seçilerek daha küçük yapıların silinmesi yeterli olacaktır.

Çizelge 4.1'de XYZ'nin sRGB renk uzayının doğruluk, F-ölçütü, kesinlik(K, Precision) ve hassasiyet (GÇ, Recall) değerleri verilmiştir. XYZ renk uzayının başarımlarının ayrı bir

tabloda verilmesinin sebebi dönüştürme işleminden sonra bölütleme gerektirmemesidir.

Çizelge 4. 1 XYZ Renk uzayındaki bütün görüntüler için ortalama doğruluk, F/F1-ölçümü, Kesinlik ve Hassasiyet değerleri

	Doğruluk	F-Ölçütü	Kesinlik	Hassasiyet
XYZ	0,94	0,60	0,57	0,74

Şekil 4.2' de veri kümesinden alınan yüksek çözünürlüklü bir görüntü ve bu görüntüye ait farklı renk uzaylarına dönüştürülmüş görüntüler yer almaktadır. Görüntünün bölütlenmesinden sonra hücrel yapıların ayrıştırılmasından sonraki ikili görüntü de yer almaktadır.

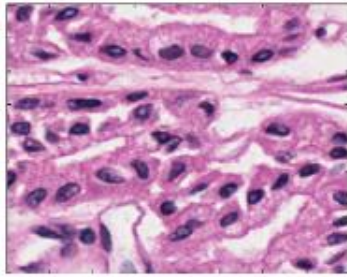
Çizelge 4. 2 KYM, La*b*, HSV renk uzaylarına ait başarımların değerleri

Doğruluk	K-Ortalamlar	Bulanık C-Ortalamlar	EM Algoritması
KYM	0,94	0,93	0,89
LA*B*	0,90	0,90	0,88
HSV	0,92	0,92	0,87

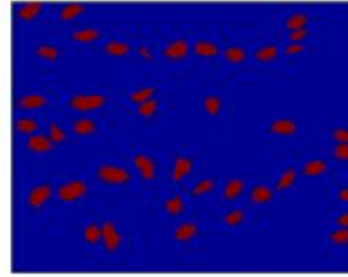
Çizelge 4.3' te KYM, La*b*, HSV gibi literatürde en çok kullanılan renk uzaylarındaki bölütleme başarımları karşılaştırılmıştır. Bölütleme algoritmaları olarak da yine literatürde en çok kullanılan yöntemlerin başında gelen K-ortalamları, Bulanık C-ortalamları ve Beklenti-Ençoklama algoritmaları kullanılmıştır. Her bir bölütleme algoritmaları için küme merkezleri sayısı üç olarak alınmıştır. Tabloda yer alan değerler veri kümesinde yer alan bütün görüntülerin ortalama değerlerini göstermektedir. Birçok görüntüde elde edilen sonuçlar yüksek iken bazı görüntülerde hücrel yapıların tam ayrışamamasından dolayı sonuçları çok düşük olan görüntülerin ortalamları düşürdüğü görülmüştür. Elde edilen sonuçlara göre KYM renk uzayında Kİ değerleri için en başarılı algoritma EM iken; K-ortalamlar ve Bulanık C-ortalamlar algoritmaları bütün renk uzayları için aynı sonucu vermiştir.

Çizelge 4. 3 KYM, La*b*, HSV renk uzaylarına ait kesinlik (K,precision), hassasiyet(H,recall) ve F-ölçütü (F,F-measure) başarımlar değerleri

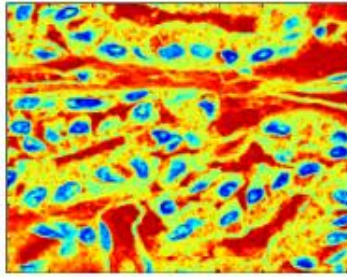
	K-Ortalamlar			Bulanık C-Ortalamlar			EM Algoritması		
	Kİ	H	F	Kİ	H	F	Kİ	H	F
KYM	0,77	0,64	0,68	0,76	0,64	0,68	0,85	0,49	0,59
LA*B*	0,80	0,54	0,61	0,80	0,54	0,61	0,73	0,47	0,51
HSV	0,68	0,70	0,67	0,68	0,70	0,78	0,78	0,53	0,59



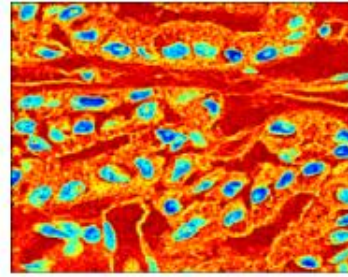
(a)



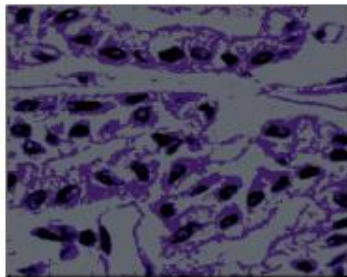
(b)



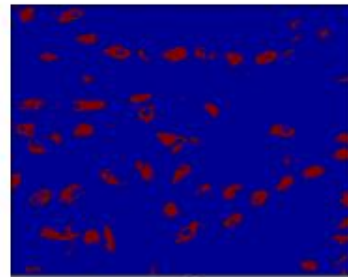
(c)



(d)



(e)



(f)

Şekil 4.2 Veri kümesinden alınan bir görüntüye ait (a) KYM renk uzayı (b) Kesin Referans (c) La*b* renk uzayının b* bandı (d) HSV renk uzayının v bandı (e) XYZ renk uzayı (f) Bölütme sonucu

Şekil 4.2'te veri kümesinden alınan yüksek çözünürlüklü bir görüntü ve bu görüntüye ait farklı renk uzaylarına dönüştürülmüş görüntüler yer almaktadır. Görüntünün bölütlenmesinden sonra hücrel yapıların ayrıştırılmasından sonraki ikili görüntü de yer almaktadır.

4.1.3 Sonuç ve Gelecek Çalışmalar

Histopatolojik görüntülerde hücrel yapıların yüksek başarımlı elde edilmesi kanser tespitinde çok büyük bir öneme sahiptir. Bu amaçla yukarıda tanıtılan veri kümesinde yer alan 64 yüksek çözünürlüklü histopatolojik görüntünün farklı renk uzaylarındaki bölütleme başarımlı hesaplanmış; hücrel yapıların en iyi ayrıştığı renk uzayı tespit edilmeye çalışılmıştır. Elde edilen sonuçlara göre K-ortalamlar ve Bulanık C-ortalamlar algoritmaları benzer sonuç verirken, Beklenti-Ençoklama algoritması KYM ve HSV renk uzaylarında kesinlik değerleri baz alındığında daha iyi sonuç vermiştir. Ayrıca KYM, La^*b^* , HSV renk uzayları arasında bir karşılaştırma yapılacak olunursa KYM ile HSV'nin bütün bölütleme algoritmalarında benzer sonuçlar verdiği görülmüştür. XYZ renk uzayına dönüştürülen görüntülerde üç farklı piksel grubu kendi içinde benzer piksel değerlerine sahip olduğundan dolayı bölütlemeye ihtiyaç duyulmamıştır. Bunun yerine basit morfolojik işlemler yardımıyla hücrel yapılar ayrıştırılmıştır.

4.2 Yüksek Çözünürlüklü Histopatolojik Görüntülerinin Bölütlenmesinde Süperpiksel Yaklaşımı

Histopatolojik görüntülerde hücrel yapıların yüksek başarı ile bölütlenmesi kanserli bölgelerin tespiti açısından çok önemlidir. Patologlar tarafından tanı koyarken dikkat edilen en önemli hususlardan biri de hücrel yapıların doğru bir şekilde bölütlenmesidir. Bu çalışmada TCGA (The Cancer Genome Atlas) veri kümesinden alınan böbrek hücre kanseri yüksek çözünürlüklü histopatolojik görüntülerinin bölütlenmesine süperpiksel yönteminin katkısı araştırılmıştır. Süperpiksel yöntemi sayısallaştırılmış görüntülerde yer alan piksellerin renk benzerliklerine ve uzamsal anlamda yakınlıklarına bağlı olarak kümeleme işlemini gerçekleştiren bir yöntemdir. Sonuçlar değerlendirildiğinde süperpiksel yönteminin hem bölütleme başarısına hem de çalışma zamanına olumlu katkıda bulunduğu gözlemlenmiştir.

4.2.1 SLIC (Simple Linear Iterative Clustering) Süperpiksel Yaklaşımı

Son yıllarda görüntüleme teknolojisinin de gelişmesi ile birlikte sayısal görüntü verilerinde artışa neden olmuştur. Verilerin nicelik olarak fazlalığı ve sürekli artış göstermesi bu verilerin etiketlenmesinin önündeki en büyük engeldir. Buna karşın görüntü işleme teknikleri ve makine öğrenme yöntemlerinin de gelişmesi ile bu verilerin etiketlenmeden işlenmesini mümkün hale getirmiştir. Son yıllarda güncel olan konulardan biri sayısal görüntülerin etiket bilgisi olmadan (eğiticişiz) tanımlanması işlemidir. Bu çalışmada sayısal histopatolojik görüntülerde etiket bilgisi olmadan hücrel yapıların bölütlenmesini gerçekleştiren süperpiksel yöntemi uygulanmıştır [52]. Çalışmada kullanılan süperpiksel yöntemi kümelenmesi amaçlanan görüntüde yer alan piksellerin renk benzerliklerine ve uzamsal anlamda yakınlıklarına bağlı olarak kümeleme işlemini gerçekleştiren bir yöntemdir. Buna göre her piksel değeri için spektral olarak alınan her bir parlaklık değeri ve sahne üzerindeki koordinat bilgisi özellik vektörü olarak hesaplanmaktadır. Denklem 4.3 çalışmada bölütleme yönteminin her bir piksel için gerçekleştirdiği işlemi ifade etmektedir.

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2} \quad (4.3)$$

Burada j değeri merkez kabul edilen pikseli, i değeri de kümelenmek istenen değeri ifade etmektedir. L, a, b değerleri La^*b^* renk uzayını temsil etmektedir. d değeri de ilgili pikselin ilgili merkeze ne kadar mesafede olduğunu gösteren değerdir. Çalışmada la^*b^* renk uzayının RGB renk uzayından daha düzgün dağılım gösterdiği ifade edildiğinden La^*b^* renk uzayı kullanıldığı belirtilmiştir. Denklem 4.4'de yine her bir piksel için sahnedeki koordinat bilgisinin nasıl hesaplandığını ifade etmektedir.

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (4.4)$$

Burada x_j, y_j değeri merkez kabul edilen pikselin yatay düşeydeki yerini, x_i, y_i değeri de kümelenmek istenen pikselin yatay ve düşeydeki bilgisini ifade etmektedir. Dikkat edilmesi gereken en önemli nokta bütün bu değerlerin aynı formül ile hesaplanmaması hassasiyeti değiştirdiğinden ayrı ayrı hesaplanarak değerlerinin normalizasyonu

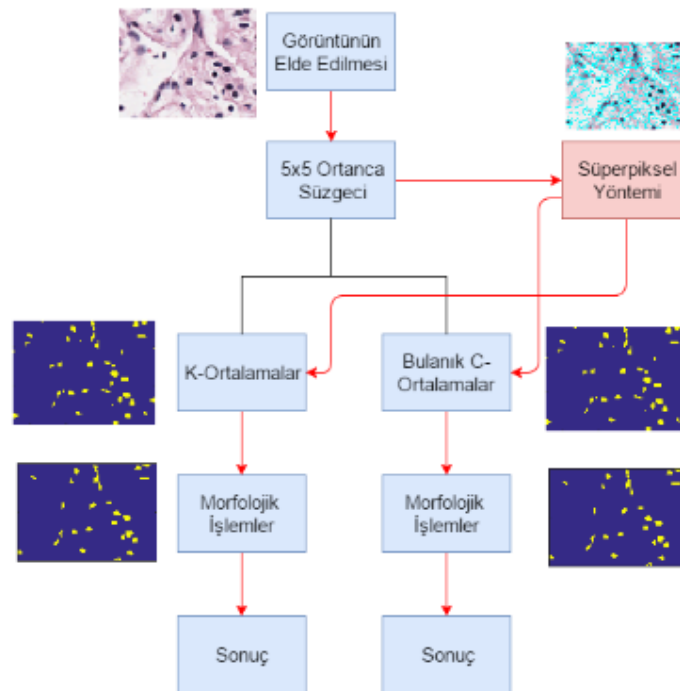
gerçekleştirildikten sonra ilgili pikselin bir merkeze atanması işleminin gerçekleştirilmesidir.

$$d = d_{lab} + x = \frac{m}{N} \cdot d_{xy} \quad (4.5)$$

Denklem 4.5 ise piksel parlaklık mesafeleri ve koordinat mesafelerinin ilklendirilen süperpiksel boyutları ve m sabiti ile normalize edilmesini ifade etmektedir. Bu şekilde mesafelerin ayrı ayrı hesaplanması koordinat bilgisinin parlaklık bilgisine baskınlığına dengelemek için yapılmaktadır.

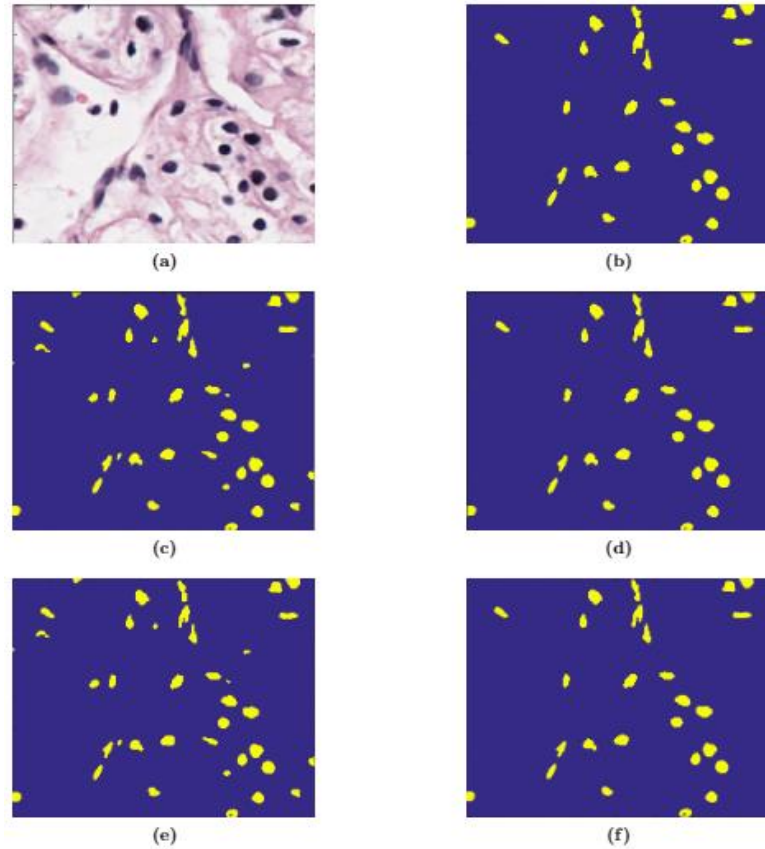
4.2.2 Süperpiksel Bölütleme Yaklaşımı ile Elde Edilen Deneysel Sonuçlar

Çalışmada kullanılan veri kümesi The Canser Genome Atlas (TCGA) veri kümesinden elde edilmiş böbrek renal hücre karsinomuna ait yüksek çözünürlüklü histopatolojik görüntülerden oluşmaktadır. Bu çalışma literatürde son dönemde oldukça popüler olan süperpiksel yönteminin bölütlemeye olan etkisi araştırılmaktadır. Bunun için çalışmanın ilk aşamasında histopatolojik görüntülerin bölütlenmesinde sıkça kullanılan yöntemlerden olan k-ortalamalar ve bulanık c-ortalamalar yöntemleri veri kümesine uygulanarak bölütlenme başarıları elde edilmiştir.



Şekil 4.3 Çalışmada önerilen yöntemin akış diyagramı

Şekil 4.3 veri kümesinden alınan bir görüntü 5x5 ortalama (medyan) süzgeçinden geçirilerek çeşitli artefaktların elimine edilmesi ve bölütleme öncesi görüntüde piksel geçişlerinin yumuşatılması sağlanmıştır. Görüntünün k-ortalamlar yöntemi ile bölütlenebilmesi için uygun matris formatına dönüştürülmektedir. Daha sonra 5 küme merkezi kullanılarak kümelenen matris yeniden görüntü boyutlarına dönüştürülmüş, hücresel yapıların ön planda kalması ve arka plan bilgisinin elimine edilmesi sağlanmıştır. Elde edilen bölütleme sonucunda hücresel yapılar ve bu yapılarla aynı piksel değerlerine sahip çeşitli bozukluklar bulunmaktadır. Son olarak elde edilen ikili görüntüye belli bir piksel alanından küçük gürültüleri giderilmesi için bir işlem, kalınlaştırma ve boşluk doldurma gerçekleştirilerek bu bozukluklar elimine edilip görüntüden en uygun bölütleme sağlanmıştır.



Şekil 4.4 (a) Veri kümesinden alınan histopatolojik görüntü ve (b) bu görüntüye ait kesin referans (ground truth), (c) görüntüye k-ortalamlar uygulanması sonucu, (d) k-ortalamlar sonucuna morfolojik işlem uygulanması, (e) bulanık c-ortalamlar sonucu, (f) bulanık c-ortalamlar sonucuna morfolojik işlem uygulanması

Bulanık c-ortalamlar yöntemi ile bölütlemeye, bir önceki paragrafta gerçekleştirilen işlemler benzer bir şekilde uygulanmıştır. Çizelge 4.4 k-ortalamlar ve bulanık c-

ortalamalar yöntemleri uygulandıktan sonra elde edilen bölütleme başarımını göstermektedir. Tabloda yer alan başarıım kıstasları orijinal çalışmada yer alan başarıım kıstasları referans alınarak elde edilmiştir [53]. Buna göre TPR (true positive ratio) sistemin hücre olarak işaretlediği piksellerin kesin referansında de hücre olarak yer aldığı bilgisini temsil etmektedir. Buna karşın TNR (true negative ratio) sistemin hücre olarak işaretlemediği ve kesin referansında hücre olarak işaretli olmayan piksellerin bilgisini temsil etmektedir.

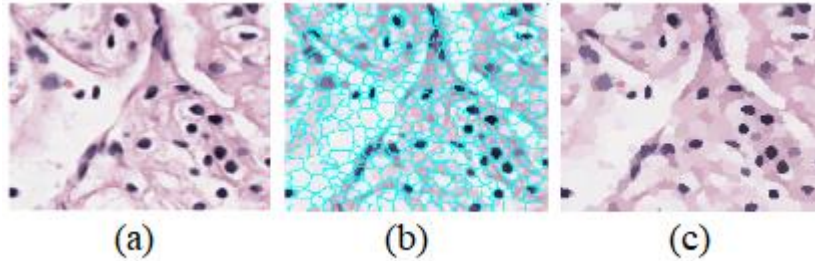
Çizelge 4. 4 KO ve BCO algoritmalarının veri kümesindeki görüntülere uygulandıktan sonra elde edilen başarımlar

Doğruluk	GPO (Gerçek Pozitif Oranı)	GNO (Gerçek Negatif Oranı)	Çalışma Zamanı
KO	0,6580	0,9713	86sn
BCO	0,6614	0,9708	853sn

Çizelge 4.4'e göre Irshad vd. gerçekleştirdikleri çalışmadaki sonuçlarla karşılaştırıldığında sonuçlar oldukça yakın çıkmaktadır. Buna göre k-ortalamalar algoritmasının tek başına kesin referans (ground-truth) hücrel bölge olarak işaretlenen piksel gruplarını hücrel bölge olarak tahmin etme başarımı 65.80% olmaktadır. Yine k-ortalamalar algoritması hücrel olmayan bölgeleri 97.13% başarı ile bölütleyebilmektedir. Buna karşın bulanık c-ortalamalar algoritması hücrel bölgeleri 66.14% başarı ile bölütleyebilmektedir. Bulanık c-ortalamalar hücrel olmayan bölgeleri 97.08% başarı ile bölütleyebilmektedir. Algoritmaların başarımları yaklaşık olarak aynı olmasına rağmen bulanık c-ortalamalar algoritması oldukça yavaş çalışmaktadır.

Uygulamanın ikinci aşamasında ise; süperpiksel uygulandıktan hemen sonra görüntülere ilk bölümde anlatılan KO ve BCO bölütleme algoritmaları uygulanarak performansları değerlendirilmiştir. Bu aşamada görüntüye ilk olarak süperpiksel yöntemi uygulanarak görüntünün belirli bir aralıkta piksel grupları ile temsil edilmesi sağlanmaktadır. Şekil 4.5 veri kümesinden alınan bir görüntü ve bu görüntüye

uygulanan süperpiksel yöntemi ile elde edilen piksel bölgelerini ve bu piksel bölgelerine denk düşen piksellerin değerlerinin güncellendikten sonraki çıktı görüntüsünü göstermektedir. Şekil 4.5 (c) aynı süperpiksele denk düşen piksellerin, KYM ortalaması alındıktan sonra değerlerinin güncellenmesi ile elde edilmiştir.



Şekil 4.5 Veri kümesinden alınan (a) örnek bir görüntü, (b) bu görüntüye uygulanan süperpiksel yöntemi sonucu elde edilen piksel bölgeleri ve (c) çıktı görüntüsü

Çizelge 4.5 görüntülere süperpiksel uygulandıktan sonra k-ortalamlar algoritmasının uygulanması sonucu elde edilen sonuçları göstermektedir. Tabloda her satırdaki SP değeri görüntünün kaç piksel grubuna bölüneceğini göstermektedir. İlk durumda 300 piksel grubuna bölünen görüntü direkt olarak k-ortalamlar uygulanmasından daha yüksek başarı ile bölütlenmektedir. Piksel grubu 400 değerine arttırıldığında k-ortalamlar en yüksek başarısına ulaşmaktadır. Süperpiksel uygulamanın çalışma zamanına da direkt olarak katkısının olduğu gözlemlenmektedir. İlk durumda yaklaşık 80 sn bütün görüntüler bölütlenirken süperpiksel uygulandıktan sonra yaklaşık 28sn'de bütün görüntüleri bölütlemektedir.

Çizelge 4. 5 K-Ortalamlar algoritmasının değişen süperpikselle göre kümeleme başarımı

	GPO (Gerçek Pozitif Oranı)	GNO (Gerçek Negatif Oranı)	Çalışma Zamanı
SP=300	0,694	0,9548	~28sn
SP=400	0,7267	0,9602	~28sn
SP=500	0,6942	0,9615	~28sn

4.2.3 Sonuç ve Çıkarımlar

Bu çalışma literatürde son dönemde oldukça popüler olan süperpiksel yönteminin The Canser Genome Atlas (TCGA) veri kümesinden elde edilmiş böbrek hücre kanserine ait yüksek çözünürlüklü histopatolojik görüntülerinin bölütlenmesine etkisi araştırılmıştır. Elde edilen sonuçlar süperpiksel yaklaşımının histopatolojik görüntülerin bölütlenmesinde kullanılabileceğini göstermiştir. Süperpiksel yöntemi bölütlemeye hem başarısız hem de çalışma zamanı olarak katkıda bulunmuştur. Çalışmada süperpikselin kullanılarak elde edilen sonuçların veri kümesinin paylaşıldığı çalışma ile karşılaştırıldığında süperpikselin başarımı iyileştirdiği gözlemlenmiştir. Gelecek çalışmalarda süperpikselin ön bölütlemeye kullanıldığı ve kanserli bölgelerin tanınmasına yönelik çalışma yapılması planlanmaktadır.

4.3 Superpixel Bölütleme ve Derin Öğrenme Algoritmaları ile Histopatolojik Görüntülerin Segmentasyonu

Histopatolojik görüntülerde hücresel yapıların yüksek başarı ile bölütlenmesi kanserli bölgelerin tespiti açısından çok önemlidir. Patologlar tarafından tanı koyarken dikkat edilen en önemli hususlardan biri de hücresel yapıların doğru bir şekilde bölütlenmesidir. Bu çalışmada TCGA veri kümesinden alınan böbrek hücre kanseri yüksek çözünürlüklü histopatolojik görüntülerinin bölütlenmesine süperpiksel yönteminin katkısı araştırılmıştır. Süperpiksel yöntemi sayısallaştırılmış görüntülerde yer alan piksellerin renk benzerliklerine ve uzamsal anlamda yakınlıklarına bağlı olarak kümeleme işlemini gerçekleştiren bir yöntemdir. Sonuçlar değerlendirildiğinde süperpiksel yönteminin hem bölütleme başarısına hem de çalışma zamanına olumlu katkıda bulunduğu gözlemlenmiştir.

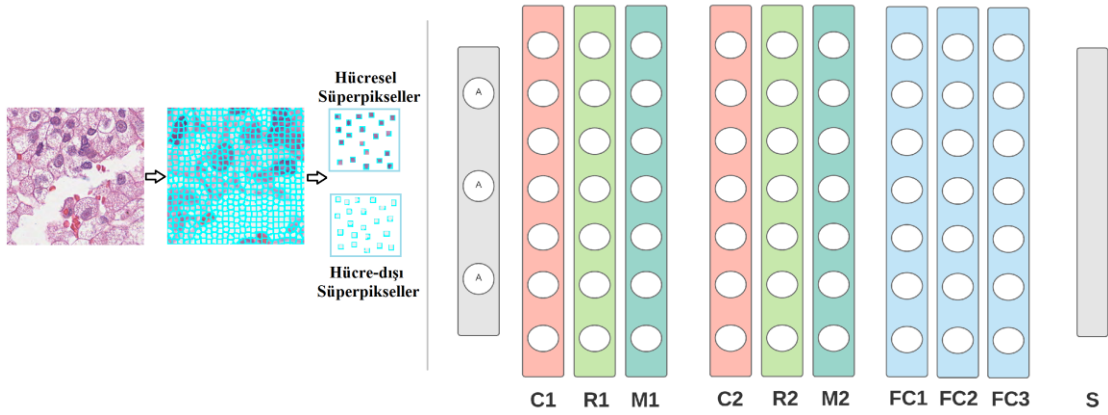
4.3.1 SLIC ve Evrimsel Sinir Ağları Yöntemlerinin Birleştirilmesi

Bu çalışmada yüksek çözünürlüklü histopatolojik görüntülerde elde edilen süperpikseller uzmanların işaretlediği kesin referans görüntüleri baz alınarak hücresel ve hücresel olmayan bölgelere denk gelmesine göre gruplandırılmıştır. Daha sonra iki gruba ayrılan bu süperpikseller önerilen evrimsel sinir ağları (ESA) modeli kullanılarak eğitilmiş ve yine bu ESA modeli ile test edilmiştir. Bu çalışmada histopatolojik

görüntülerde yer alan hücresel yapılar eğitici bir algoritma olan ESA yöntemi ile gerçekleştirilmeye çalışılmıştır. İşaretli veriler kullanılarak eğitilen ESA modeli sistemin hücresel yapılarını tespit etmek için kullanılmıştır. Görüntülerde elde edilen süperpiksellerin şekli m kompaktlık parametresine bağlıdır. m parametresinin düşük olması koordinat bilgisinin katsayısını düşürdüğünden parlaklık bilgisini baskın hale getirir. Bu durum her bir süper pikselin konveks veya konkav bir şekle sahip olmasını sağlayacağından sonraki adımda derin öğrenme yönteminde süperpiksellerin optimum bir girdi resmi olmamasına sebep olur. m parametresinin yüksek bir değere sahip olmasında ise her bir süperpiksel yaklaşık olarak karesel bir şekle sahip olur. Bölütleme işleminden sonra süperpiksellerin derin öğrenme yöntemi ile sınıflandırılması amaçlandığından yaklaşık olarak karesel bir yapıya yakın olması süperpikselin derin öğrenme yöntemi ile daha yüksek bir başarı oranı ile sınıflandırılmasını sağlar.

Evrişimsel sinir ağları (ESA), ilk olarak LeCun vd tarafından önerilen derin öğrenme algoritmalarından biridir [54]. İlk önerildiği zamanlarda küçük boyutlu verilerde uygulanan ESA zaman içinde donanım teknolojilerinin de gelişmesiyle birlikte büyük boyutlu verilerde de uygulanmaya çalışılmıştır [55-56]. ESA algoritmasında görüntü giriş katmanı tarafından elde edildikten ve analiz için hazırlandıktan sonra Evrişim (Convolutional) katmanına geçilir. Evrişim katmanında, M satır ve N sütun olmak üzere $M \times N$ boyutlu bir giriş görüntüsü olsun. Bu safhada görüntüye k adet $A \times A$ boyutlarında genelleştirilmiş süzgeç uygulanarak k adet özellik haritası elde edilir. Her bir özellik haritası giriş görüntüsüne uygulanan süzgecin boyutlarına bağlı olarak $M-A$ boyutlarına sahip olacaktır. Burada dikkat edilmesi gereken önemli bir nokta özellik haritalarının boyutlarının isteğe bağlı olarak azaltılamayabileceğidir. Küçük boyutlu görüntülerde daha fazla özellik elde edebilmek adına giriş görüntüsünün genişletilerek özellik haritalarının giriş görüntüsünün boyutlarına eşit boyutlara sahip olması sağlanabilir. Evrişimsel katman kabaca sahip olduğu süzgeç sayısı ve bu süzgece ait boyut olmak üzere 2 farklı parametreye sahiptir. Süzgeç sayısı çıktı olarak oluşacak özellik haritası boyutunu etkilerken, süzgeç boyutu ise özellik haritasında ayırteci özelliklerin oluşturulmasında katkıda bulunmaktadır. Aktivasyon katmanı evrişim katmanından sonra uygulanmaktadır. Bu katmanın amacı evrişim katmanı uygulanırken elde edilen verilerin doğrusal ortamdan doğrusal olmayan ortama aktarılıp analiz edilebilmesini

sağlamaktır. Havuzlama katmanı ise aktivasyon katmanından hemen sonra uygulanmaktadır. Bu katmanda amaç elde edilen içeriğin indirgenmesidir. Tam bağlı katman ise kendisinden önce yer alan bütün nöronlardaki değerlerin kendisinden sonra oluşturulan tüm nöronlarla eşleşmesini sağlayıp çıkarılabilecek en fazla sayıda ayırt edici özelliklerin elde edilmesini sağlamaktadır.



Şekil 4.6 Veri kümesinden alınan (a) örnek bir görüntü, (b) bu görüntüye uygulanan süperpiksel yöntemi sonucu elde edilen piksel bölgeleri ve (c) çıktı görüntüsü

4.3.2 Önerilen Yaklaşım ile Elde Edilen Sonuçların Değerlendirilmesi

Bu çalışma literatürde son dönemde oldukça popüler olan süperpiksel bölütleme yöntemi ve ESA algoritmalarının histopatolojik görüntülerde bölütlemeye olan etkileri araştırılmaktadır. Önerilen bu çalışma temel olarak superpiksellerin SLIC yöntemi ile elde edilmesi ve elde edilen bu süperpiksellerin ESA yöntemi ile sınıflandırılması olmak üzere iki aşamadan oluşmaktadır.

Çalışmanın ilk aşamasında veri kümesinde yer alan her bir girdi görüntüsü çeşitli ışık etkileri ve gürültülerden arındırılmak için 5x5 ortanca süzgeçinden geçirilmiştir. Gürültüden arındırılan görüntü daha sonra kendisini meydana getiren 3000 süperpiksele bölünerek her bir süperpikselin kesin referansta karşılığına bağlı olarak çekirdek veya çekirdek olmayan klasörlere kopyalanmıştır. İlgili süperiksel ground truth'ta çekirdek yapısına karşılık geliyorsa çekirdek klasörüne, arkaplan kabul edilen (yağ doku ve bağ doku) bilgisine karşılık geliyorsa çekirdek olmayan klasöre kaydedilmiştir. Burada iki önemli nokta ön plana çıkmaktadır. Süperpiksellerin sayısının neye göre belirleneceği ve elde edilen süperpiksellerin derin öğrenme algoritmalarına

nasıl uygun hale getirileceği konularıdır. Ortalama bir hücre yaklaşık olarak 30x30 mertebelerinde olduğu varsayıldığında veri kümesinde yer alan görüntüler için optimum süperiksel sayısı uzamsal benzerlikler de hesaba katıldığında yaklaşık olarak 250 ila 750 arasında değişebilir. Yani bir yağ doku parçacığı için ilgili süperpiksel hesaplanan değerden daha fazla çıkabiliyorken hücresel bir yapı için ilgili süperiksel hesaplanan değerden daha küçük oranlarda çıkabilir. Ayrıca elde edilen süperpiksel uzamsal benzerlikler dikkate alındığında konveks veya konkav bir şekle sahip olabilir. Bu durum ESA modeline girdi olacak bu görüntüde şekilsel bazda değişim gerektirdiğinden sınıflandırma performansını olumsuz etkileyebilir. Bu sebeplerden dolayı önerilen bu çalışmada histopatolojik görüntüler mümkün olduğunda yüksek sayıda süperpiksel sayısına bölünerek her bir süperpikselin atanacağı merkeze olan mesafelerine dolaylı yoldan kısıtlama getirilmesi düşünülmüştür. Bu çalışmanın ilk aşamasında her bir görüntü 3000 süperpiksele bölünmüştür. Bu sayı empirik olarak belirlenmiştir fakat ileride gerçekleştirilecek diğer çalışmalarda süperpiksel sayısının detaylı analizi gerçekleştirilerek bu sayının otomatize edilmesi mümkün hale gelebilir. Denklem 4.5' te m ile ifade edilen parametre süperpikselin kompaktlığını ifade eder. Bu parametrenin küçük sayıda verilmesi koordinat bilgisinin etkisini azaltmak anlamına gelmektedir. Derin öğrenme algoritmasına girdi olarak verilecek her bir süperpikselde karesel bir görüntü oluşturulabilirse bu durum görüntülerin sınıflandırılma başarısını arttırmaya katkı sağlar. m değeri piksel parlaklık değeri ve koordinat bilgisinin süperiksellerin hangi merkeze atanacağına hangi oranda katkıda bulunacağını gösterir. Başlangıçta kaç adet süperpiksel sayısı gerektiğini belirlemek için kullanılan

$\sqrt{\frac{\text{genişlik} \times \text{yükseklik}}{\text{süperpiksel sayısı}}}$ denklemine bağlı olarak m değerine yaklaşık olarak 50 verilmesi

karesel boyutlarda süperpikseller almamızı sağlayacaktır. Bu çalışmada m parametresi 50 olarak belirlenmiştir. Çalışmanın ilk aşamasında belirtilen parametrelere bağlı olarak çekirdek ve çekirdek olmayan süperpikseller belirlenip 2 sınıflı bir yapı kurulmuştur.

Çalışmanın ikinci aşamasında ise çekirdek ve çekirdek olmayan süperpikseller önerilen CNN algoritması yardımıyla sınıflandırılmışlardır. İlk aşamada elde edilen her bir süperpiksel yaklaşık olarak aynı en-boy (12x12) oranlarına sahiptir. Bu bağlamda görüntülerin CNN algoritmasına uygun hale getirmek için yeniden

boyutlandırıldıklarında çok fazla şekilsel değişime uğramayacaklardır. CNN algoritmasının girdi görüntülerinin boyutları 15x15 boyutlarındadır. CNN ağı için Bölüm 4.3.2'te belirtilen parametreler kullanıldığında elde edilen başarı sonuçları Çizelge 4.6'da gösterilmektedir.

Çizelge 4.6 süperpiksellerin önerilen ESA modeliyle sınıflandırma sonucunda elde edilen karışıklık matrisini göstermektedir. Tablodaki sütun bilgisi sırasıyla hücrel yapıyı temsil eden süperpikselleri, hücrel olmayan süperpikselleri ve algoritmanın çalışma zamanını ifade etmektedir. Satır bilgisi ise yine hücrel yapıyı oluşturan süperpikselleri ve hücrel olmayan süperpikselleri ifade etmektedir. Sonuçlara bakıldığında önerilen ESA yönteminin bütün hücrel yapıyı oluşturan süperpikselleri sınıflandırma başarıları %94.51 ve hücrel olmayan yapıları temsil eden süperpikselleri sınıflandırma başarıları %99.17 olarak elde edilmiştir. Sistemin başarıları %98.76 olarak gözlemlenmiştir. Çalışma zamanına bakıldığında önerilen ESA yönteminin bütün verileri sınıflandırması yaklaşık olarak 2.13 saat sürmüştür. ESA modelinin veriyi öğrenmesi ve yüksek başarı ile sınıflandırması için iterasyon sayısı oldukça önemlidir. Iterasyon sayısı arttıkça ESA modelinin veriyi öğrenmedeki başarıları artmaktadır. Bu durum aynı zamanda sonucun tesadüfi olmadığını göstermesi açısından önemlidir. Önerilen bu sistemde veriler her biri 100 epoch'lu olmak üzere 5 iterasyon ile eğitilmiştir. Sistemin başarıları 5-kat çapraz geçerlilik ile test edilmiştir.

Çizelge 4. 6 Bütün Veri kümesine ait K-ortalamlar yöntemine ait karışıklık matrisi(GPO ve GNO) ve Çalışma Zamanı. Satırlar gerçek verileri, sütunlar ise tahmin edilen verileri göstermektedir.

	Hücre SP	Hücre-dışı SP	Çalışma Zamanı
Hücre SP	0,9451	0,0549	2.13saat
Hücre-dışı SP	0,0083	0,9917	

Çizelge 4.7 ise önerilen ESA modelinin veri kümesinde yer alan bütün süperpiksellerin doğru ve yanlış sınıflandırma sayısını ifade etmektedir. Buna göre sistem 1038 hücrel yapıyı temsil eden süperpikselden sadece 57 tanesini yanlış sınıflandırmışken, hücrel

olmayan bölgeleri temsil eden 10702 süperpikselden sadece 89 tanesini yanlış sınıflandırmıştır.

Çizelge 4. 7 ESA modelinin veri kümesinde yer alan bütün süperpiksellerin doğru ve yanlış sınıflandırma sayısı. Satırlar gerçek verileri, sütunlar ise tahmin edilen verileri göstermektedir.

	Hücre SP	Hücre-dışı SP
Hücre SP	981	57
Hücre-dışı SP	89	10613

4.3.3 Sonuç ve Gelecek Çalışmalar

Bu çalışma literatürde son dönemde oldukça popüler olan süperpiksel yönteminin The Canser Genome Atlas (TCGA) veri kümesinden elde edilmiş böbrek hücre kanserine ait yüksek çözünürlüklü histopatolojik görüntülerinin bölütlenmesine etkisi araştırılmıştır. Elde edilen sonuçlar süperpiksel yaklaşımının histopatolojik görüntülerin bölütlenmesinde kullanılabileceğini göstermiştir. Süperpiksel yöntemi bölütlemeye hem başarısız hem de çalışma zamanı olarak katkıda bulunmuştur. Çalışmada süperpikselin kullanılarak elde edilen sonuçların veri kümesinin paylaşıldığı çalışma ile karşılaştırıldığında süperpikselin başarımı iyileştirdiği gözlemlenmiştir. Gelecek çalışmalarda süperpikselin ön bölütlemeye kullanıldığı ve kanserli bölgelerin tanınmasına yönelik çalışma yapılması planlanmaktadır.

4.4 Süperpiksel-Tabanlı İki Aşamalı Bölütleme Algoritması ile Histopatolojik Görüntülerde Otomatik Hücre Segmentasyonu

Yüksek çözünürlüklü sayısal histopatolojik görüntülerde hücresel yapıların morfolojik özelliklerinin analizi, kanserin tanı ve tedavisi için standart klinik uygulamadır. Halihazırda, patoloğların hücresel yapıları bu şekilde manuel olarak incelemesi oldukça yorucu bir süreçtir. Bu yorucu ve zaman alıcı süreci otomatikleştirmek, literatürdeki histopatolojik görüntü işleme çalışmalarında üzerinde çalışılan bir konudur. Bu yazıda böbrek karsinomunun yüksek boyutlu histopatolojik görüntülerinde hücresel yapıların

elde edilebilmesi için iki aşamalı bir segmentasyon yöntemi önerilmiştir. İlk olarak, veri kümesinden alınan her bir görüntü basit doğrusal yinelemeli kümeleme (BDYK) yöntemi ile süperpiksellere ayrılmıştır. Daha sonra elde edilen süperpikseller, hücre çekirdeğini oluşturan benzer süperpikselleri bulmak için literatürde en sık kullanılagelen kümeleme tabanlı bölütleme algoritmaları tarafından kümelenmiştir. Ayrıca, çalışmada kümeleme tabanlı bölütleme yöntemleri ve lokal uzaklık tabanlı süperpiksel bölütleme algoritmalarının performansı da karşılaştırılmıştır. Sonuçlar, süperpiksel bölütleme algoritmasının ön bölütleme yöntemi olarak kullanılmasının, basit kümeleme tabanlı bölütleme algoritmasına kıyasla hücre bölütlenmesi performansını geliştirdiğini göstermektedir. Gerçek pozitif oran (GPO), gerçek negatif oran (GNO), F ölçüsü, kesinlik ve örtüşme oranı (OR) bölütleme performansının değerlendirmesinde kullanılan ölçümlerdir. Algoritmaların hesaplama zamanları da karşılaştırmalı olarak değerlendirilmiş ve çalışmada sunulmuştur.

4.4.1 Süperpiksel Algoritmasının Ön-Bölütleme olarak Kullanılması

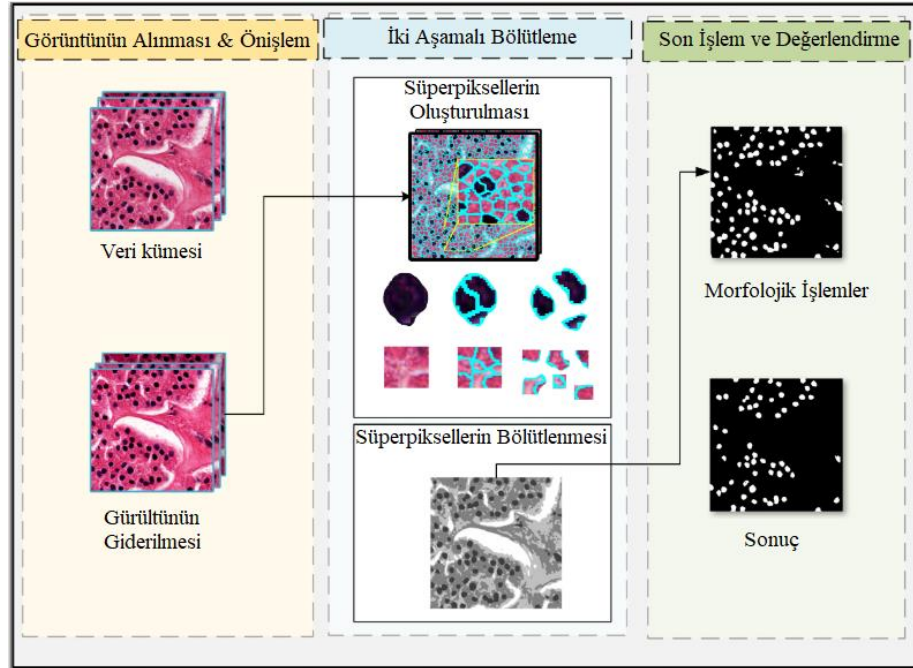
Histopatolojik görüntülerin bölütlenmesinde en çok kullanılan yöntemlerden biri kümeleme tabanlı bölütleme algoritmalarıdır. Bu algoritmalar temelde görüntüde yer alan konumundan bağımsız her bir pikselin parlaklık bilgilerini kullanarak belirli sayıda bir küme merkezine düşürmeyi amaçlamaktadır. Diğer bir deyişle, düzlemdeki (1,1) koordinatındaki piksel ile (100,100) koordinatındaki piksel aynı ağırlıklara (burada ağırlıktan kastedilen 'önemdir') sahip olarak hesaplanmaktadır. Bu da piksellerin birbirine olan konumlarından bağımsız olarak parlaklık bilgilerinin uzaklıklarını hesaplamaktadır. Süperpiksel yaklaşımı bu noktada kümeleme tabanlı bölütleme algoritmalarından ayrılmaktadır. Süperpiksel tabanlı bölütleme yöntemlerinde belirli bir lokal mesafeye kadar olan piksellerin hem koordinat uzaklığını hem de parlaklık uzaklığını hesaplamaya dayalı çalışan bir yöntemdir.

Bu çalışma ile literatüre aşağıdaki katkıların sağlanması amaçlanmaktadır:

- Global yaklaşımlı kümeleme tabanlı bölütleme algoritmaları ile lokal yaklaşımlı süperpiksel bölütleme algoritmalarının performanslarının karşılaştırılması,

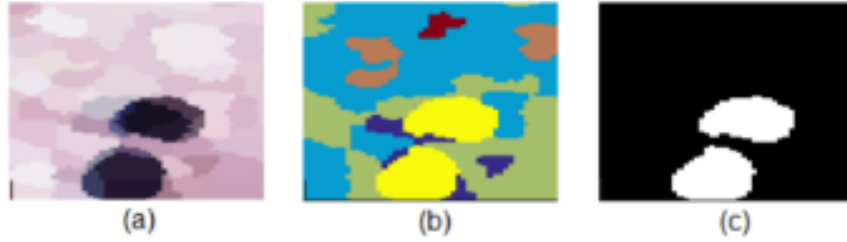
- Süperpiksel bölütleme algoritmalarının ön bölütleme algoritması olarak kullanıldığında bölütleme performansına olan etkisinin incelenmesi,
- Son dönemde sıkça kullanılan süperpiksel yöntemlerinin performanslarının karşılaştırılması,
- Bölütleme algoritmalarının çalışma zamanı performanslarının karşılaştırılması,
- SLIC süperpiksel bölütleme algoritmasının parametrelere bağlı bölütleme performansı şeklindedir.

Çalışmanın ilk aşamasında k-ortalamlar, bulanık c-ortalamar gibi literatürde en çok bilinen global yaklaşımlı bölütleme algoritmalarının performansı değerlendirilmiştir. Histopatolojik bir görüntüde 3 temel eleman (yağ doku, bağlayıcı doku ve çekirdek) bulunduğundan 3 ana küme merkezine bölütlenmesi başarılı sonuç verir diye düşünülebilir fakat gerek patolojik ön işlemlerden gerekse de verilerin taranıp sayısal ortama aktarılmasından kaynaklanan çeşitli gürültülü veriler oluşabilmektedir. Bu durumda 3 küme merkezi düzgün bölütler elde edebilmek için yeterli olamayabilmektedir. Bu çalışmada 3'ten 7'ye kadar bütün küme merkezi sayıları denenip en yüksek performansı veren küme merkezi sayısı kullanılmıştır.



Şekil 4.7 Önerilen bu çalışmada takip edilen işlem adımları.

Çalışmanın ikinci aşamasında, SLIC süperpiksel yöntemi ile elde edilen belirli sayıda süperpiksel kümeleme tabanlı bölütleme yöntemleri ile bölütlenmiştir. Şekil 4.7, bu çalışmada önerilen yöntemin işlem akışını ifade etmektedir. Bu aşamada ilk olarak görüntüler gürültülerden elimine edilmektedir. Bunun için 5x5 boyutlarında ortalama süzgeci uygulanmıştır. Daha sonra yumuşatılmış görüntü belirli sayıda süperpikselle bölütlenmektedir.

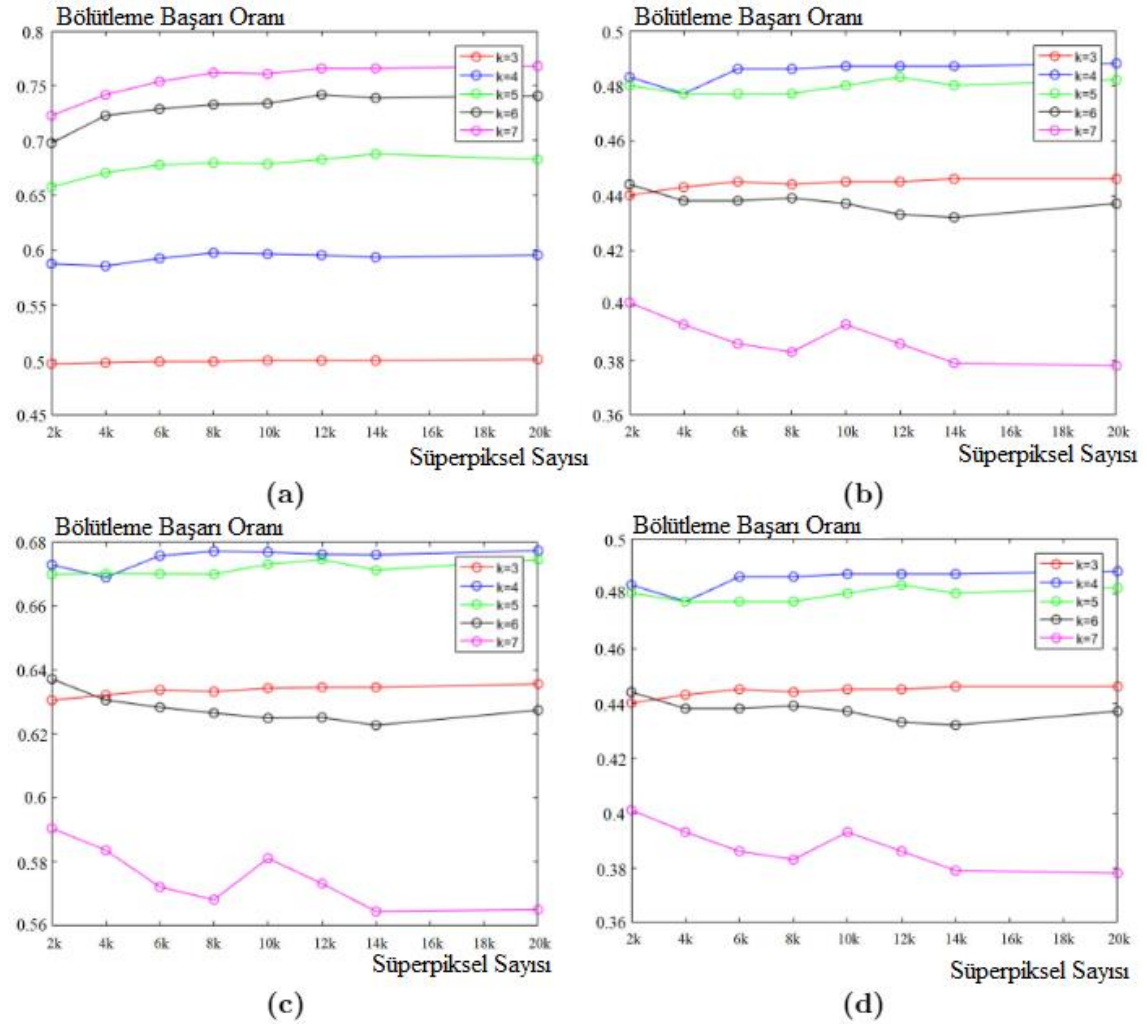


Şekil 4.8 Örnek bir görüntüye uygulanan (a) süperpiksel bölütleme çıktısı, (b) bu süperpikselle uygulanan k-ortalamar yönteminin çıktısı ve (c) elde edilen sonuç

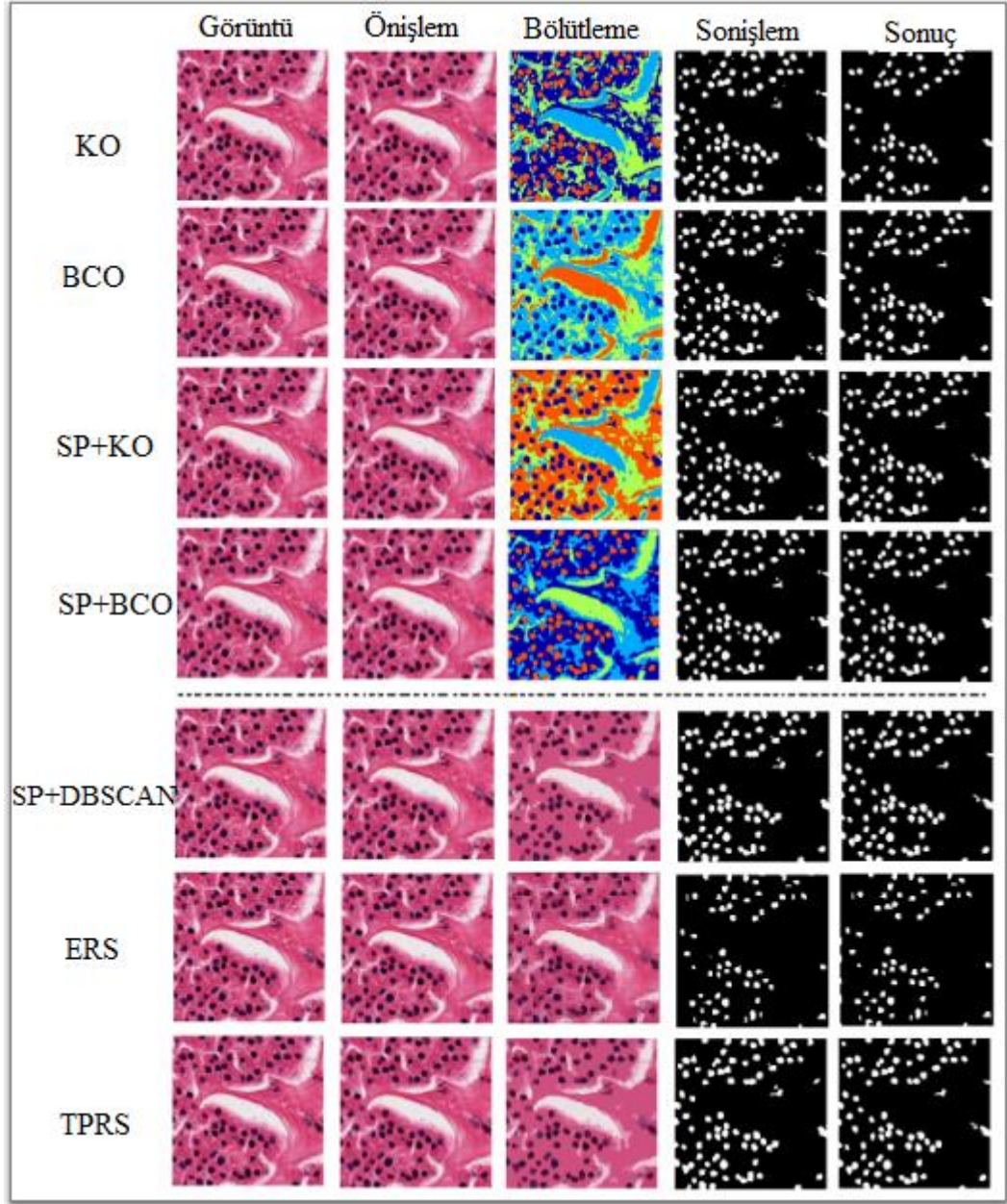
Çizelge 4. 8 K-ortalamları (KO), Bulanık C-Ortalamları (BCO), SLIC+KO, SLIC+BCO, SLIC+DBSCAN, ERS, TPRS ve veri kümesinin paylaşıldığı çalışmada yer alan yöntemle ait Gerçek pozitif Oranı (GPO), Kesin İsalet (Kİ), F-Ölçütü (F-Ö), Gerçek Negatif Oranı (GNO), Çakışma Oranı (ÇO) ve Çalışma Zamanı (Ç.Zamanı) oranlarının karşılaştırılması

	GPO	Kİ	F-Ö	GNO	ÇO	Ç. Zamanı
KO	0,60	0,69	0,61	0,97	0,45	60sn
BCO	0,60	0,68	0,61	0,973	0,447	546sn
SLIC+KO	0,69	0,67	0,63	0,96	0,472	27sn
SLIC+BCO	0,66	0,68	0,64	0,96	0,48	365sn
SLIC-DBSCAN	0,74	0,61	0,64	0,94	0,48	183sn
ERS	0,65	0,64	0,61	0,95	0,45	61sn
TPRS	0,65	0,63	0,61	0,96	0,45	487sn
İrshad vd.	0,76	0,62	0,65	0,96	0,49	-

Burada süperpiksel sayısının parametrik olması sistemin performansını direkt olarak etkilediğinden çok sayıda deney yapmayı gerektirmektedir. Bunun için her bir görüntü belirli sayı aralığında süperpiksellere bölütlendikten sonra kümeleme tabanlı algoritmalarla bölütlenip performans karşılaştırılması yapılmaktadır.



Şekil 4.9 Belirli süperpiksel sayıları ve küme merkezlerine karşılık önerilen yöntemin elde ettiği bölütleme başarısı. (a) Kesinlik, (b) Özgünlük, (c) F-ölçütü ve (çakışma) oranları sırasıyla gösterilmektedir.



Şekil 4.10 Veri kümesinden elde edilen bir görüntüye uygulanan algoritmalarından elde edilen sonuç.

4.4.2 Sonuç

Bu çalışma yüksek çözünürlüklü histopatolojik görüntülerde SLIC süperpiksel bölütleme yöntemiyle hücrelerin bölütlenmesini amaçlamaktadır. Çalışmada öncelikli olarak global yaklaşımlı kümeleme yöntemleri olan k-ortalamlar ve bulanık c-ortalamlar yöntemleri kullanılmıştır. Bu yöntemlerle karşılaştırma yapılması amacıyla süperiksel yöntemleri kullanılmıştır. Çalışmada önerilen yöntem SLIC süperiksel bölütleme yönteminin ön-bölütleme algoritması olarak kullanılmasıdır. Sonuçlar incelendiğinde

SLIC süperpiksel yönteminin ön bölütleme algoritması olarak kullanılması bölütleme performansını nispeten arttırmaktadır. Bunun yanında çalışma zamanına oldukça etki etmektedir.

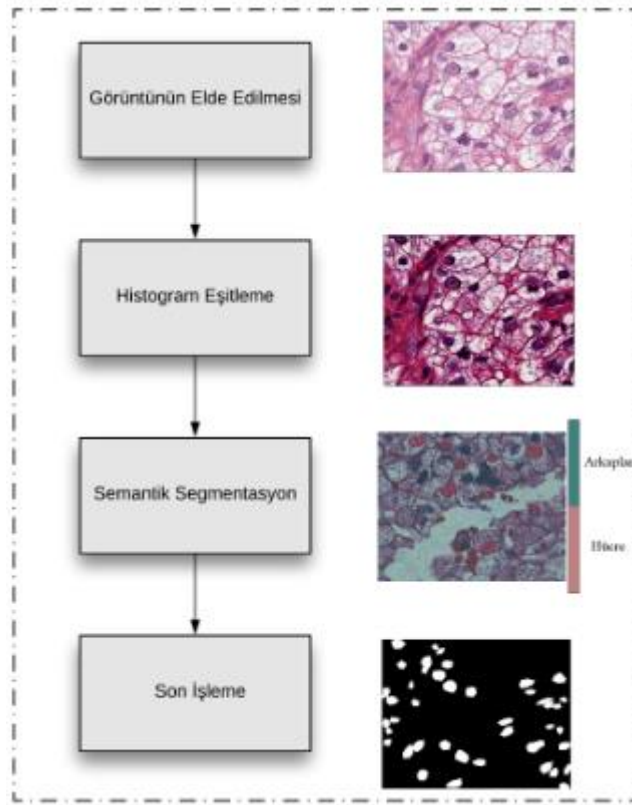
4.5 Histopatolojik Görüntülerde Kodlayıcı-Çözücü Tabanlı Derin Öğrenme Algoritması ile Hücresel Yapıların Bölütlenmesi

Önerilen bu çalışmada yüksek çözünürlüklü histopatolojik görüntülerde hücresel yapıların bölütlenmesi amacıyla derin öğrenme tabanlı SegNet yöntemi kullanılmıştır. SegNet, yol ve kapalı mekanlarda yer alan nesnelerin bölütlenmesi için önerilen derin evrişimsel kodlayıcı-çözücü mimariye sahip bir yöntemdir. Önerilen çalışmada SegNet algoritmasının hücresel yapıların bölütlenmesindeki başarısı elde edilmeye çalışılmıştır. Ayrıca SegNet'in literatürde iyi bilinen bölütleme algoritmaları ile performans karşılaştırılması da sunulmuştur. Elde edilen sonuçlara göre SegNet'in hücresel yapıların bölütlenmesinde sıkça kullanılan diğer yöntemlere göre oldukça başarılı olduğu gözlemlenmiştir.

4.5.1 Semantik Segmentasyon ile Bölütleme

Bu çalışmada kanser tespitinin gerçekleştirilmesinde dikkat edilen en önemli parametrelerden biri olan hücresel yapıların bölütlenmesi ile ilgili bir çalışma gerçekleştirilmiştir. Çalışma son yıllarda oldukça popüler olan 'semantik segmentation (SegNet)' yöntemi ile hücresel yapıları arka-plandan (yağ doku ve bağ doku) ayıran bir yöntem denenmiştir [57]. Uygulanan yöntem literatürde sıkça kullanılan ve geleneksel olarak ifade edilen çeşitli algoritmalarla karşılaştırılmıştır.

Çalışmada takip edilen işlem adımları Şekil 4.11'de ifade edilmektedir. Yüksek çözünürlüklü histopatolojik görüntülere ilk olarak histogram eşitleme yöntemi uygulanmaktadır. Histogram eşitleme, camların taranması sırasında oluşabilecek renk veya parlaklık farklılığını normalize ederek bütün görüntülerin belirli bir skala ile ifade edilmesini sağlamaktadır. Daha sonra, derin öğrenme tabanlı SegNet algoritması yardımıyla tüm görüntüler bölütlenmektedir.



Şekil 4.11 Çalışmada takip edilen işlem adımları özetle şu şekildedir: Elde edilen histopatolojik görüntü ilk olarak histogram eşitleme işleminden geçirilmektedir. Daha sonra SegNet yapısına girdi olarak verilen derin öğrenme tabanlı yöntem ile bölütlenmektedir. Son olarak elde edilen bölütleme çıktısına morfolojik işlem uygulanmaktadır.

Elde edilen bölütleme çıktılarında hücreler ve hücresel olmayan artefaktlar elde edilmektedir. Son olarak morfolojik işlemler (belirli bir piksel sayısından az olan piksel gruplarını elimine etmek, 1 boyutlarında disk operatörü ile genişletme ve nesnelerin içerisindeki boşlukları doldurma) yardımıyla giderilmektedir. Böylece hücresel yapılar elde edilerek algoritmanın performansı değerlendirilmektedir.

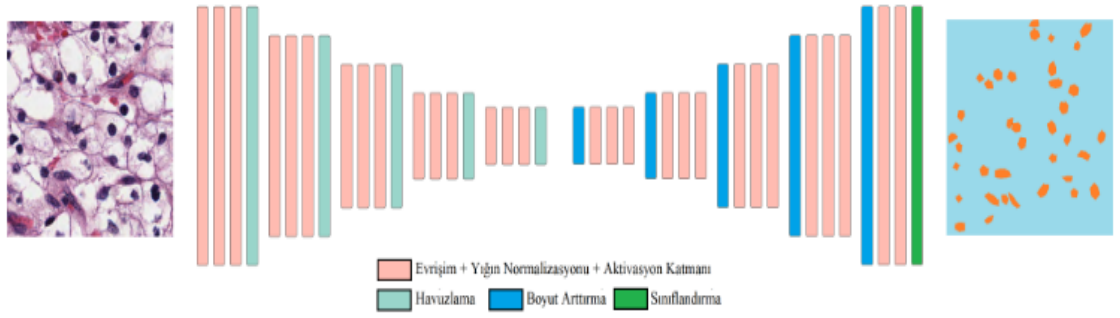
Semantik bölütleme algoritmaları temelde piksellerin etiketlenmesi yöntemini sağlayarak bölütleme işlemini gerçekleştirmektedir. Bu fikir ilk olarak rassal ormanlar, Boosting gibi algoritmaların tekil pikselleri sınıflandırılmasından ilham alarak ortaya çıkmıştır. Daha sonra sınıflandırma amacıyla önerilen derin öğrenme algoritmaları kullanılarak pikselleri etiketlenmesi ile semantik bölütleme yöntemine temel oluşturulmuştur [58-59].

Bu çalışmada klasik bölütleme algoritmalarından farklı olarak literatürde birçok alanda başarılı sonuçlar elde edilen derin öğrenme tabanlı semantik bölütleme algoritması önerilmektedir[60-64]. Algoritmanın kodlayıcı kısmında sınıflandırma amacıyla önerilen standart derin öğrenme yaklaşımlarından vgg16 ağı yer almaktadır. Vgg16 ağ yapısı Şekil 4.12'de ifade edildiği gibi evrişim (convolutional), yığın normalizasyonu (batch normalization), aktivasyon katmanı (Rectified Liner Unit-ReLU) ve havuzlama (pooling) katmanlarından oluşmaktadır. Dekoder kısmında ise kodlayıcıda gerçekleştirilen işlemlerin tam tersi sayılabilecek işlemler uygulanmaktadır. Dekoder kısmında kodlayıcıda boyut indirgeme işlemi için kullanılan havuzlama yönteminin tersi sayılan boyut artırma işlemi gerçekleştirilmektedir. Bu işlem her katmanda boyutu indirgenen görüntünün yeniden girdi görüntüsü boyutlarına dönüştürülmesi işlemidir. Böylece boyut indirgenirken kullanılan parametreler yardımıyla birbirleriyle ilişkili piksellerin boyut arttırılırken birlikte gruplanması sağlanmaktadır.

4.5.2 Semantik Segmentasyonun Başarımının Değerlendirilmesi

Çalışmanın bu bölümünde, derin öğrenme tabanlı SegNet semantik bölütleme algoritması ve literatürde sıkça kullanılan yöntemlerden k-ortalamar ve eşikleme tabanlı bölütleme algoritmalarının uzman patologlar tarafından işaretlenmiş yüksek çözünürlüklü böbrek karsinomum histopatolojik görüntüleri üzerindeki performansı değerlendirilmiş ve veri kümesinin paylaşıldığı çalışma ile karşılaştırılmıştır [49-65].

Bu çalışmada kullanılan veri seti, Harvard Üniversitesi'nde Beck Laboratuvarı'ndan alınmıştır. Veri kümesi, TCGA veri portalından seçilen böbrek karsinomunun yüksek çözünürlüklü histopatolojik görüntülerini içermektedir.



Şekil 4.12 SegNet algoritmasının temel yapısı. Algoritma enkoder (kodlayıcı) ve dekoder (çözücü) olmak üzere iki aşamadan oluşmaktadır. Kodlayıcıda sınıflandırma amacıyla kullanılan standart vgg16 algoritması yer almaktadır. Çözücüde ise sonucun girdi görüntüsüne benzetilmesi için boyut arttırımı işlemi de gerçekleştirilmektedir.

Çalışmanın ilk aşamasında, derin öğrenme tabanlı SegNet bölütleme algoritmasının başarımı elde edilmiştir. SegNet algoritmasının geliştirme ortamında çalışabilmesi için veri kümesinde yer alan kesin referans (ground truth) görüntülerinde yer alan her bir sınıf için bir renk ataması yapılmalıdır. Önerilen bu çalışma hücresel yapılar ve arkaplan (yağ doku ve bağ doku) olarak işaretlenen iki sınıflı bir problem olduğundan hücresel yapılar ve arkaplan iki farklı renk ile işaretlenmiştir. Veri kümesinde yer alan 64 adet görüntü SegNet algoritması kullanılarak 5-kat çapraz geçerlilik ile çalıştırılmış, elde edilen sonuçlar diğer yöntemlerle karşılaştırılmıştır. Uzman patoloğlar tarafından işaretlenen görüntüler baz alındığında hücresel yapıları temsil eden piksel sayısı arkaplan olarak ifade edilen ve hücresel olmayan piksel sayısından çok daha az sayıda olmaktadır. Normal şartlar altında örnek sayısından bağımsız olarak derin öğrenme tabanlı sınıflandırma algoritmaları yanlış sınıflandırılan her bir örnek için belirli bir hata değeri hesaplanmaktadır. Bu çalışmada kullanılan SegNet algoritması sınıfların dengesiz olduğu veri kümelerinde örnek sayısı az olan sınıflar için hata hesabını sayıya bağlı olarak ağırlıklandırabilmektedir. Yani sayısı az olan hücresel yapıları temsil eden pikseller yanlış sınıflandırıldığında hata değerini daha fazla olarak hesaplayabilmektedir. Bu durum SegNet algoritmasının hücresel yapıları (ya da sınıf sayısı az olan herhangi çok sınıflı örnekleri) diğer algoritmalarından daha iyi öğrenebilmesini sağlamaktadır.

Çalışmanın ikinci aşamasında literatürde sıkça kullanılan yöntemlerden kümeleme tabanlı bölütleme algoritması olan k-ortalamlar ve eşikleme tabanlı bölütleme

algoritmalarından Otsu algoritması karşılaştırma amaçlı kullanılmıştır. K-ortlamalar algoritması dört küme merkezli olacak şekilde kullanılmıştır.

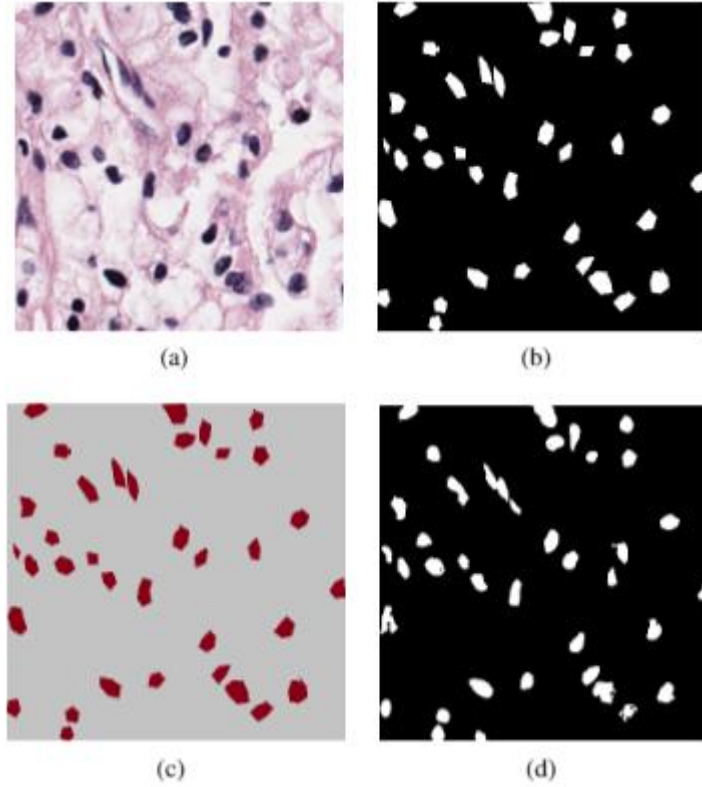
Çizelge 4.9'da böbrek karsinomu veri kümesinde dört algoritmanın bölütleme performansının karşılaştırılması sunulmuştur. Tablodaki sütun bilgileri sırasıyla: kesinlik (Kes.), f-ölçümü (F-Ö), gerçek negatif oranı (GNO) ve çakışma oranı (ÇO) değerlerini göstermektedir. Satır bilgisi ise; veri kümesinin paylaşıldığı çalışmada elde edilen performans, bu çalışmada önerilen SegNet yönteminde elde edilen performans ve karşılaştırma amacıyla literatürde sıkça kullanılan yöntemlerin kullanılmasıyla elde edilen performans sırasıyla verilmiştir.

Çizelge 4. 9 Böbrek karsinomu veri kümesinde dört algoritmanın bölütleme performansının karşılaştırılması.

	GPO	Kİ	F-Ö	GNO	ÇO	Çalışma Zamanı
Irshad vd.	0,76	0,62	0,65	0,95	0,49	-
Segnet	0,70	0,71	0,69	0,95	0,53	8sn (Test aşaması)
KO	0,60	0,69	0,61	0,97	0,45	60sn
Otsu	0,67	0,64	0,62	0,95	0,46	546sn

Sonuçlar değerlendirildiğinde f-ölçümü ve çakışma oranında SegNet algoritması en başarılı performans göstermektedir. Irshad vd. önerdiği algoritma ise kesinlik değerinde en iyi performans göstermektedir. GNO değerinde k-ortalamlar dışındaki diğer üç algoritma da neredeyse aynı sonucu verirken k-ortalamlar algoritması en başarılı sonucu vermiştir. Sonuçlara bakıldığında algoritmalar ortalama başarılı olarak ifade edilebilir. Burada uzman patolgoların veri kümesinde yer alan her bir hücrel yapıyı işaretlerken hücre sınırlarını tam anlamıyla işaretlememesinden kaynaklı olarak başarımlar düşmektedir. Literatürde bu anlamda yaygın bir veri kümesi olmadığından bu veri kümesi birçok çalışmada tercih edilmektedir. Şekil 4.13'te (a) veri kümesinden elde edilen histopatolojik görüntüyü, (b) bu görüntüye ait uzman patoloğ tarafından

işaretlenen kesin referans (ground truth) görüntüsünü (c) Segnet algoritmasının çalışması için kesin referans görüntülerinin dönüştürüldüğü görüntü tipini ve (d) SegNet algoritması sonucu elde edilen çıktı görüntüsünü ifade etmektedir.



Şekil 4. 13 Sırasıyla, (a) veri kümesinden elde edilen histopatolojik görüntüyü, (b) bu görüntüye ait uzman patolog tarafından işaretlenen kesin referans (ground truth) görüntüsünü (c) Segnet algoritmasının çalışması için kesin referans görüntülerinin dönüştürüldüğü görüntü tipini ve (d) SegNet algoritması sonucu elde edilen çıktı görüntüsünü ifade etmektedir.

Çalışma kapsamında gerçekleştirilen tüm deneyler 4.0 GHz Intel core i7-6700K işlemcili, NVIDIA GeForce TITAN X ekran kartı işlemcisi ve 48 GB RAM' e sahip bir iş istasyonunda gerçekleştirilmiştir. SegNet algoritması dışındaki diğer iki algoritma CPU ile gerçekleştirilmiştir. Derin öğrenme tabanlı SegNet algoritması ekran kartı işlemcisi ile gerçekleştirilmiştir. Algoritmalar farklı işlemciler ile gerçekleştirildiğinden zaman performansı karşılaştırılmamıştır.

4.5.3 Sonuç ve Çıkarımlar

Histopatolojik görüntülerde bilgisayar destekli teşhis sistemleri kullanılarak hücresel yapıların hücresel olmayan yapılardan yüksek başarı ile ayrılması kanserin erken

tespitinde oldukça büyük bir öneme sahiptir. Bu işlem kanserin erken tespitinde uzmana yardımcı ikincil karar destek sistemlerinin geliştirilmesi motivasyonunu arttıracak işlemlerden biri olacaktır. Bu amaçla yukarıda belirtilen veri kümesinde yer alan 64 yüksek çözünürlüklü histopatolojik görüntünün derin öğrenme tabanlı SegNet algoritması ile yüksek başarı ile bölütlenmeye çalışılmış, elde edilen sonuçlar literatürde sıkça kullanılan yöntemlerle karşılaştırılmıştır. Elde edilen sonuçlara göre SegNet algoritması Irshad vd, k-ortalamlar ve Otsu yöntemine göre göreceli olarak daha başarılı sonuç vermiştir. Buna karşın sayfa sayısı limitinden dolayı SegNet algoritmasının belirli parametrelere bağlı performans sonuçları değerlendirilememiştir. SegNet algoritması kodlayıcı-çözücü tabanlı bir derin öğrenme algoritması olduğundan gelecek çalışmalarda bu kodlayıcı-çözücü derinliğine bağlı bölütleme başarımları, sınıfların ağırlıklandırılması işlemine bağlı performans değerlendirmesi yapılması amaçlanmaktadır. Bunun yanında SegNet yönteminin U-Net gibi diğer kodlayıcı-çözücü tabanlı derin öğrenme yöntemleri ile karşılaştırılması da amaçlanmaktadır.

HİSTOPATOLOJİK GÖRÜNTÜLERİN SINIFLANDIRILMASI

Bu bölümde histopatolojik görüntülerde sınıflandırılma problemi üzerine yoğunlaşılacaktır. Histopatolojik görüntülerde hücresel bazda sınıflandırma yapılacağı gibi doku kesidi veya lezyon bazında da sınıflandırma işlemi yapmak mümkündür. Çalışma kapsamında bazı kanser türlerinin derecelendirilmesinde önemli bir yere sahip olan mitozlu hücrelerin tespitinde hücrelere yoğunlaşılmaktadır. Fakat lenf nodlarında tümörlü bölgelerin tespiti için teker teker hücrelerin morfometrik özelliklerinin analiz edimesindense bu hücrelerin dokuya etkisi incelenip bölgesel bazda görüntü kesitleri üzerinden sınıflandırma yapmak mümkündür. Bu bölümde her iki yaklaşıma ait çalışmalar anlatılacaktır.

5.1 Histopatolojik Görüntülerde Mitozlu Hücrelerin Tamamlanmış Yerel İkili Örüntüler Kullanılarak Belirlenmesi

Bu çalışmada 2014 yılında ICPR tarafından gerçekleştirilen bir yarışma için oluşturulan yüksek çözünürlüklü histopatolojik görüntülerdeki mitozlu hücrelerin belirlenmesi ve normal hücrelerden ayrıştırılıp ayrıştırılamadığı incelenmiştir. Normal şartlar altında patoloji uzmanları tarafından mikroskopla incelenip tespit edilmesi zor olan mitozlu hücreler bilgisayar destekli teşhis (BDT- Computer Aided Diagnosis) sistemleri ile otomatik elde edilmeye çalışılmış, böylelikle patoloğlara yardımcı olunması amaçlanmıştır. Bir önceki bölümde vurgulanan ve olası tüm hücrelerin bölütlenmesini kapsayan işlemde sonra hücrelerin özniteliklerinin çıkarılarak sınıflandırılması işlemi gerçekleştirilmektedir. Bunun için literatürde dokusal özniteliklerin çıkarılması amacıyla

önerilen tamamlanmış yerel ikili örüntüler (TYİL, completed local binary pattern) yöntemi uygulanmıştır.

Önerilen bu çalışmada ilk olarak k-ortalamlar algoritması kullanılarak hücresel yapıların hücresel olmayan yapılardan ayırt edilmesi sağlanmıştır. Tespit edilen bu hücresel yapıların tamamlanmış yerel ikili örüntü (CLBP) yöntemi ile mitozlu olup olmadıklarının belirlenmesi için özellikleri çıkarılmıştır. Öznitelikleri çıkarılan hücreler rassal ormanlar algoritması ile sınıflandırılmış ve buna göre modeller oluşturulmuştur. Elde edilen sonuçlara bakıldığında, eşit sayıda veya yakın oranlara sahip mitozlu ve mitozsuz hücreye sahip veri kümelerinde kesinlik, hassasiyet ve F-ölçütü değerleri oldukça başarılı sonuç verirken, verinin düzgün dağılım göstermediği, mitotik olmayan hücrelerin lehine dengesiz oluşturulan kümelerde bu başarının düştüğü gözlemlenmektedir.

5.1.1 Yerel İkili Örüntü (YİÖ, Local Binary Pattern)

Yerel ikili örüntüler (local binary pattern), piksellerin komşu pikselleri ile olan ilişkilerini inceleyerek ürettiği LBP kodları ile örüntü bulmaya çalışır. Bir pikseli LBP kodu, komşuları ile kıyaslanarak aşağıdaki gibi hesaplanmaktadır:

$$LBP_{P,R} = \sum_{p=0}^{p-1} s(g_p - g_c) 2^p \quad (5.1)$$

$$s(x) \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Burada g_c merkezde yer alan pikselin gri seviyedeki değeri, g_p ise komşu piksellerin gri seviyedeki değerlerini göstermektedir. P toplam komşu sayısını, R ise komşuların merkez piksele olan uzaklığını göstermektedir. Diğer bir deyişle, R uzaklığı kaç adet komşunun dikkate alınacağını, P sayısını, belirlemektedir. Bu çalışmada $R=1$ olarak alınmıştır. Bu durumda merkez pikselin $P=8$ adet komşusu bulunmaktadır. s fonksiyonu ise eşik fonksiyonudur. Komşu pikselin, merkez pikselden küçük olup olmadığını inceler.

Elde edilen LBP kodlarından, veriye ait bir histogram elde edilir. Verilen MxN boyutlarına sahip bir resim için histogram aşağıdaki gibi formülize edilebilir (K, LBP kodunun alabileceği en büyük değeri göstermek üzere $k \in [0, K]$):

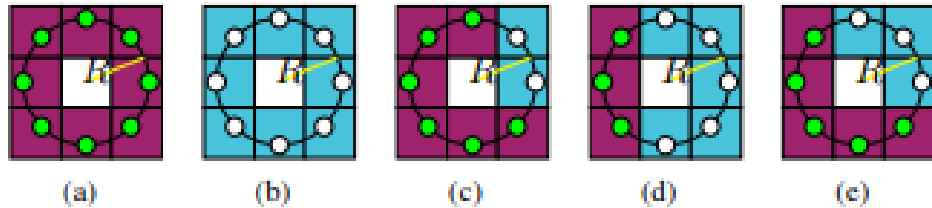
$$H(k) = \sum_{i=1}^M \sum_{j=1}^N f(LBP_{p,R}(i,j), k) \quad f(x, y) \begin{cases} 1 & x = y \\ 0 & \text{diğer durumlar} \end{cases} \quad (5.2)$$

LBP kodlarında 0-1 veya 1-0 geçişlerinin sayısı U değeri olarak aşağıdaki gibi tanımlanmaktadır:

$$U(LBP_{p,R}) = |s(g_{p-1} - g_c) - s(g_0 - g_c)| + \sum_{j=1}^N |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (5.3)$$

U değeri hesaplanırken, LBP kodunun dairesel olduğu düşünülerek hesaplama yapılır. Düzgün dağılımlı LBP örüntülerinde iki geçiş ya da hiç geçiş olmayan ikili gösterimler ($U \leq 2$) kullanılır. Bu $LBP_{p,R}^{U2}$ notasyonu ile gösterilir. Şekil 5.1'de bu notasyona göre örnek geçişlerin ne anlama geldikleri gösterilmiştir.

Bu durumda 2^p değer üzerinden değil, $P \times (P-1) + 3$ değer üzerinden histogram hesaplanır. Bu çalışmada $P=8$ olduğundan 59 değer mevcut bulunmaktadır. $U > 2$ olan tüm değerler 59. değer olarak sayılmaktadır. Haliyle $U \leq 2$ şartını sağlayan 58 değer bulunmaktadır.



Şekil 5.1 Merkez pikselden büyük ve eşit değerde olan komşular; içi dolu ve koyu zemin üzerinde, küçük değerde olan komşular ise içi boş ve açık zemin üzerinde gösterilmiştir. 0 ve 2 renk geçişi olduğunda elde edilen bu örnek durumlar şu şekilde yorumlanabilir:

(a) Nokta (b) Nokta/Düz (c) Doğru (d) Kenar (e) Köşe

LBP yönteminde, merkez piksel eşik değeri kabul edilerek LBP kodları üretilmekteydi. Tamamlanmış LBP (CLBP) yönteminde ise, LBP yönteminde elde edilenlere ek olarak komşu piksellerin merkez piksele olan uzaklığı da değerlendirilir [66]. Bu durumda

merkez pikselin komşulara benzerliği de değerlendirilmiş olmaktadır. Şekil 5.2'de örnek bir piksel için CLBP yönteminden elde edilen değerler gösterilmiştir.

74	95	45
60	60	52
67	10	16
(a)		
1	1	0
1	g_c	0
1	0	0
(b)		
14	35	15
0	g_c	8
7	50	44
(c)		

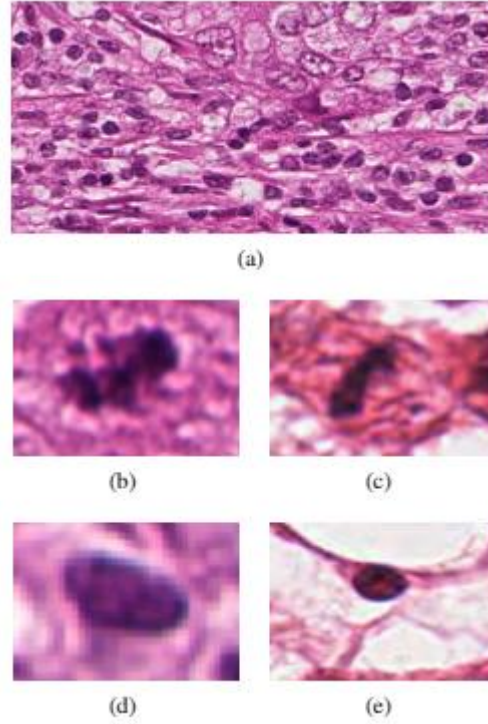
Şekil 5. 2 Örnek 3x3'lük blok (b) LBP kodu (c) Merkeze olan uzaklıkların (fark) mutlak değerleri

Elde edilen uzaklık değerleri belirlenen bir eşik değerine göre, LBP'de olduğu gibi kodlar üretilir (eşik değeri, resimdeki tüm farkların ortalama değeri olarak belirlenmiştir).

Bu çalışmada $P=8$, $R=1$ alındığında LBPu2P,R yöntemine göre 59 adet özellik elde edilmiştir. Ayrıca uzaklık (fark) bilgileri için oluşturulan değerler için $U \leq 2$ yöntemi ile özellik elde edilirse buradan da 59 özellik elde edilmektedir. Bu durumda bir görüntüden 118 adet özellik elde edilmiş olur.

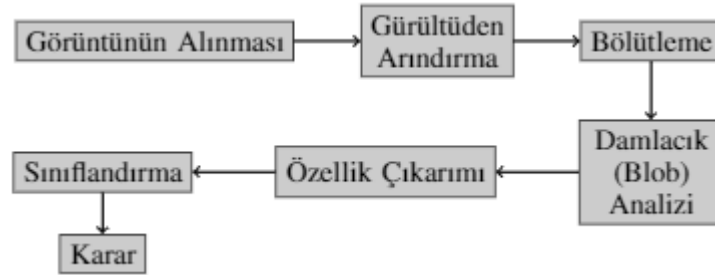
5.1.2 Yerel İkili Örüntülerin Sınıflandırma Başarısının Değerlendirilmesi

Bu çalışmada, 2014 yılında International Conference on Pattern Recognition (ICPR) tarafından gerçekleştirilen yarışma kapsamında paylaşılan göğüs kanseri histopatolojik görüntüleri incelenmiştir. Paylaşılan veri kümesinde toplamda 1136 görüntü ve 749 mitotik hücre yer almaktadır. Şekil 5.3'te veri kümesinden alınmış bir görüntü kesiti, iki adet mitozlu hücre örneği Şekil 5.3 (b-c) ve iki adet mitozsuz hücre örneği Şekil 5.3 (d-e) yer almaktadır.



Şekil 5.3 (a) Veri kümesinden alınmış taranmış histopatolojik görüntü örneği (b)-(c) mitozlu hücre örneği (d)-(e) mitozsuz hücre örneği

Şekil 5.4 modelin işlem adımlarını göstermektedir. Buna göre ilk olarak, önışlemlerden geçirilen görüntü bölütlendikten sonra, özellik çıkarımı gerçekleştirilmektedir. Elde edilen özelliklerden modeller oluşturularak sınıflandırma gerçekleştirilir.



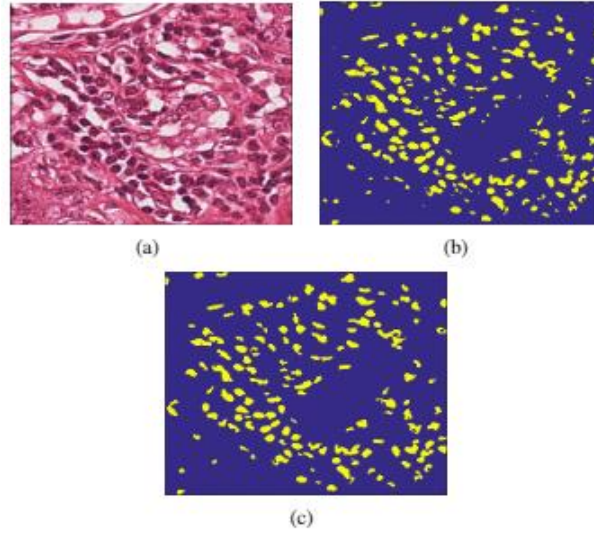
Şekil 5.4 Çalışmada takip edilen işlem adımları

Önişleme Adımları: Yüksek çözünürlüklü histopatolojik görüntülerde mitozlu hücrelerin tespiti için hücresel ve hücresel olmayan bölgelerin ayrıştırılması büyük önem arz etmektedir. Bu ayrıştırma için temelde kullanılan yöntem eğiticişiz bölütleme işlemidir. Ancak salt bölütleme işleminde hücresel olmayan ve genel olarak gürültü kabul edilen küçük alanların bulunması söz konusu olabilmektedir.

Çalışmada ilk olarak gürültü kabul edilen alanların minimize edilmesini sağlamak amacıyla görüntülere bölütleme öncesinde 32x32'lik görüntülere 9x9 ortanca filtresi uygulanmıştır. Bu sayede hücresele olmayan küçük alanların yumuşatılması sağlanmıştır.

Bölütleme amacıyla görüntülere K-ortalamlar algoritması uygulanmış ve görüntüler 6, 7, 8, 9 küme sayısı değerleri denenerek hücresele bölgelerin en iyi derecede ayrıştırılmasını sağlayan küme sayısı bulunmuştur. Denemeler sonucunda hücresele yapıların en iyi 7 bölütte ayrıştırıldığı sonucu çıkarılmıştır. Bölüt sayısının 7'den az olduğu durumlarda hücre çevresinin hücrenin bir parçası olduğu, 7'den fazla bölüt sayısında ise hücresele yapıların sınırlarının daraldığı ve bu durumun hücreyi olduğundan daha küçük göstermekte olduğu görülmüştür. Bölütleme sonucu ortaya çıkan hücresele olmayan belirli bir piksel sayısından küçük yapılar da şekilbilimsel (1 boyutunda 'disk' operatörü ve genişletme işlemi) işlemlerle giderildikten sonra hücresele yapıların özelliklerinin çıkarılması safhasına geçilmiştir.

Özellik Çıkarımı Yöntemi: Mitotik ve mitotik olmayan hücresele yapıların özellikleri CLBP algoritması kullanılarak çıkarılmıştır. Şekil 5.5'de örnek olarak gösterilen görüntüdeki gibi bütün görüntülerdeki tüm hücresele yapılar çevrelenerek bu yapıların kapladığı alanların dokusal özellikleri çıkarılmıştır. Burada dikkat edilmesi gereken en önemli nokta mitoz geçiren bir hücrede gerçekleşen morfolojik değişikliklerdir. Literatürde mitoz geçiren bir hücrede gerçekleşen çekirdeğin erimesi, kromozonların merkeze hareketi belirli ayırtedici değişikliklere sebep olurken mitoz geçirmeyen hücreler genel itibarıyla oval yapıda ve oval çekirdeğe sahip olması nedeniyle bu hücrelerin ayırtedici olduğu düşünülmektedir. Bu noktada hücresele bölgelerde gerçekleşen mitozun görüntüdeki bu alanda parlaklık değerlerini değiştirmekte ve bu bölgelerin dokusal özelliklerinin incelenmesini gerektirmektedir.



Şekil 5.5 (a) Orijinal görüntü (b) Bölütlenmiş görüntü (c) gürültülü alanların temizlenmesi

Sınıflandırma: Mitotik ve mitotik-olmayan tüm hücresel yapılara ait özellik vektörleri çıkarıldıktan sonra sınıflandırma aşamasına geçilmiştir. Sınıflandırma algoritması olarak topluluk tabanlı rassal ormanlar algoritması kullanılmıştır. Rassal ormanlar algoritması için kullanılan karar ağaçları sayısı 250 olarak belirlenmiştir. Bunun yanında çapraz geçerlilik olarak 5-kat çapraz geçerlilik uygulanmıştır. Veri kümesinde yer alan test görüntülerine ait kesin referansında paylaşılmadığı için eğitim veri kümesi ikiye bölünerek yarısı eğitim diğer yarısı test verisi olarak kullanılmıştır.

Çizelge 5. 1 Algoritmaya ait sınıflandırma başarısı

Mitozlu/Mitozsuz	Kesinlik	Hassasiyet	F-Ölçütü
500/500	0,874	0,879	0,876
500/1000	0,766	0,882	0,819
500/1500	0,676	0,891	0,768
500/2000	0,602	0,895	0,719
500/2500	0,55	0,889	0,678
500/5000	0,38	0,931	0,539

Çizelge 5.1’de veri kümesinden kesin referansı belirli olan 11 klasörden 500 mitotik hücreye karşılık 1, 2, 3, 4, 5 ve 10 katını içeren mitotik-olmayan hücreler rastgele olarak seçilmiştir. Buna göre 500 mitozluya karşılık 500 mitozsuz hücre ele alınınca kesinlik değeri %87,4 iken, hassasiyet değeri %87.9 olmaktadır. Kesinlik ve hassasiyet değerlerinin harmonik ortalaması olan F-ölçütü değeri de %87.6 olmaktadır. Mitozlu hücre sayısı sabit 500 olurken mitozsuz hücre sayısının periyodik olarak arttırılması mitozlu hücrelerin tespitini zorlaştırırken mitozsuz hücrelerin tespitini kolaylaştırmaktadır. Bu durum verinin dağılımının değişmesinden kaynaklanmaktadır. Bu gibi verilerde sınıflandırma başarısını arttırmanın en iyi yolu düzgün dağılım göstermeyen verilere duyarlı algoritmaların geliştirilmesidir.

5.1.3 Sonuç

Bu çalışmada, ICPR tarafından 2014 yılında paylaşılan yüksek çözünürlüklü histopatolojik görüntülerdeki mitozlu hücreler tespit edilmeye çalışılmıştır. Mitoz geçiren hücrelerde çekirdeğin erimesi, kromatid ipliklerin merkezde belirmesi gibi morfolojik değişiklikler gerçekleştiğinden dokusal tanımlayıcılarla bu hücrelerin daha kolay tespit edildiği düşünülmektedir. Bu bağlamda mitozlu hücrelerin özellik vektörleri dokusal öznitelik tanımlayıcı yöntemlerden biri olan CLBP kullanılarak çıkarılmıştır. Çıkarılan özellikler rassal ormanlar (random forests) algoritması ile sınıflandırılıp sistemin başarısı gözlemlenmiştir.

Elde edilen sonuçlara bakıldığında sistemin, eşit sayıda veya yakın oranlara sahip mitozlu ve mitozsuz hücrelerin kesinlik, hassasiyet ve F-ölçütü değerleri oldukça başarılı sonuç verirken, verinin düzgün dağılım göstermediği çeşitli oranlarda bu başarının düştüğü gözlemlenmektedir. Örneğin 500 mitozlu hücreye karşılık 2500 mitozsuz hücre alındığında kesinlik %87’den %55’e düşerken, F-ölçütü değeri %87’den %67’e düşmektedir. Bu durum verinin dağılımının değişmesinden kaynaklanmaktadır. Bu noktada gelecek çalışmalar kapsamında düzgün dağılım göstermeyen verilerde sınıflandırma başarısının yükseltilmesi konusunda çalışmalar yapılması planlanmaktadır.

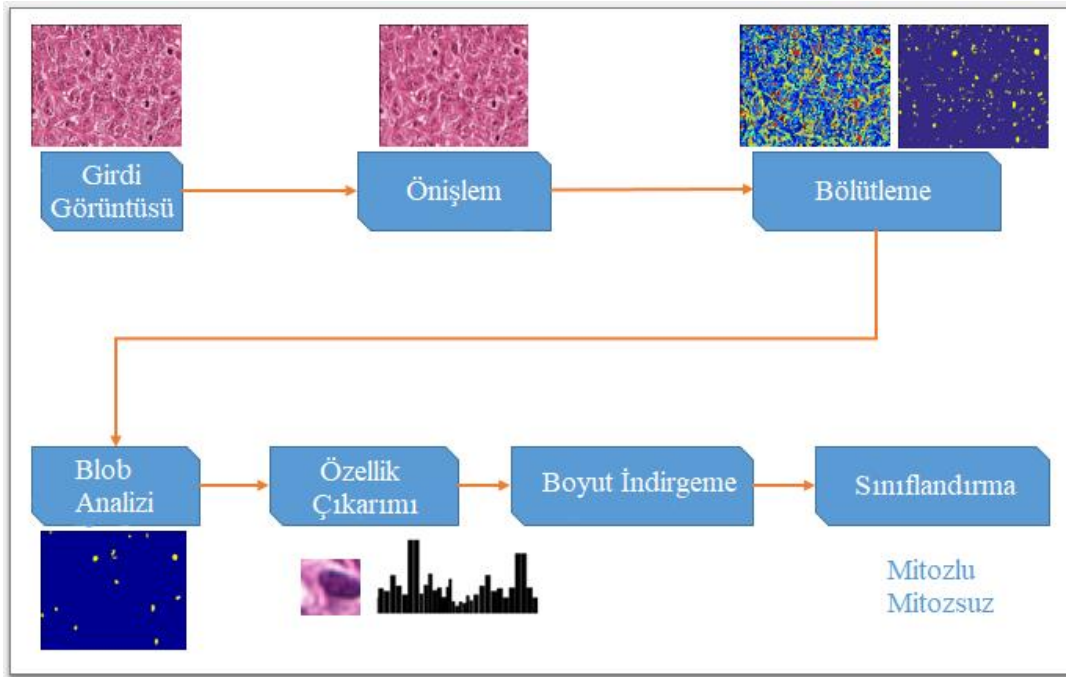
5.2 Mitozlu Hücrelerin Derin Öğrenme Tabanlı Öznelikler Kullanılarak Belirlenmesi

Dünya Sağlık Örgütü'ne (DSÖ) göre meme kanseri kadınlarda ikinci kanser ölüm nedenidir. Bilgisayar destekli diagnostik (CAD) sistemlerinin geliştirilmesi, doğru tanı ve tedavi süreci için ikincil okuyucu sistemleri olarak büyük önem taşımaktadır. Bu çalışmada histopatolojik görüntülerde mitozlu hücrelerin tespitinin gerçekleştirilmesi için evrışimsel sinir ağıları algoritması kullanılmıştır. Önerilen bu çalışma International Conference on Pattern Recognition (ICPR) konferansında düzenlenen bir yarışma için hazırlanan veri kümesi üzerinde çalıştırılmıştır. Veri kümesinde yer alan her bir görüntü yumuşatma işleminden sonra bölütlenmiş ve öznelikleri çıkarılmak üzere çerçeveye alınmıştır. Görüntülerde bölütlenen her bir hücresel yapı kesilerek evrışimsel sinir ağıları algoritmasına girdi olarak verilmiştir. Burada ayırtedici olduğu düşünülen öznelikler çıkarıldıktan sonra temel bileşenler analizi(TBA) ve doğrusal ayrıştırma analizi (DAA) yöntemleri ile boyutları indirgenmiştir. Son olarak Destek Vektör Makinaları (DVM) yöntemiyle sınıflandırılmıştır. Elde edilen sonuçlar incelendiğinde benzer sayılarda (dengeli veri kümesinde) mitozlu hücre ve mitozsuz hücrelerde algoritmanın başarılı olduğu söylenebilir.

5.2.1 Evrışimsel Sinir Ağları Tabanlı Mitozlu Hücre Tespiti Yaklaşımı

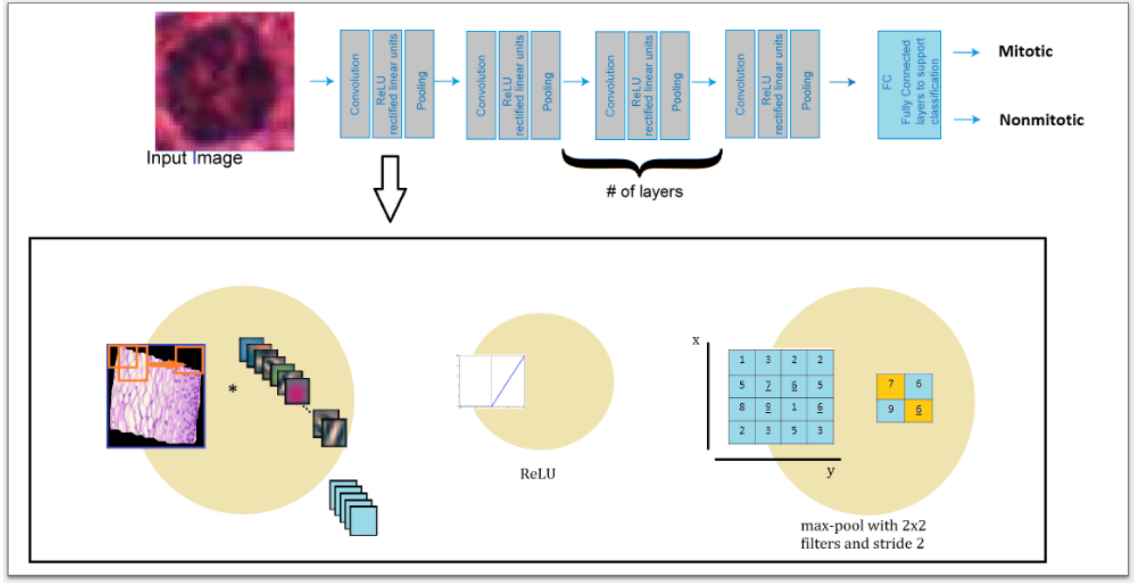
Önerilen bu çalışma Şekil 5.6'da belirtilen işlem adımları takip edilerek gerçekleştirilmiştir. Yüksek çözünürlüklü histopatolojik görüntülere ilk olarak 32x32 boyutlarındaki girdi görüntüsüne gürültülerin giderilmesi amacıyla 7x7 ortanca filtresi uygulanmıştır. Ortanca filtresinin uygulanmasının sebebi yüksek frekanslar temzilenirken hücresel yapıların kenar geçişlerinin korunmasını sağlamaktır. Bölütleme amacıyla görüntülere K-ortalamar algoritması uygulanmış ve görüntüler 6, 7, 8 küme sayısı denenerek hücresel bölgelerin en iyi derecede ayrıştırılmasını sağlayan küme sayısı bulunmuştur. Hücresel yapıların en iyi 7 küme merkezinde ayrıştırıldığı gözlemlenmiştir. Küme merkezinin sayısının 7'den az olduğunda hücre çevresinin hücre ile bütünleştiği, 7'den fazla olması durumunda ise hücresel yapıların parçalandığı görülmüştür. Bölütleme işleminden sonra ortaya çıkan küçük artefaktlar şekilbilimsel(1 boyutunda 'disk' operatörü ve genişletme işlemi) işlemlerle giderildikten sonra hücresel

yapıların özelliklerinin çıkarılması safhasına geçilmiştir. Bu aşamada hücrelere evrimsel sinir ağı yöntemi uygulanmıştır.



Şekil 5.6 Çalışma kapsamında takip edilen işlem adımları

Şekil 5.7 bir ESA modelinin tüm katmanlarını ifade etmektedir. ESA algoritması birçok gizli katman barındıran nöral ağılar tabanlı bir algoritma olarak da ifade edilebilir. Standart bir ESA ağında evrişim katmanı, aktivasyon katmanı, havuzlama katmanı, tam bağlı katman ve sınıflandırma katmanı olmak üzere beş ana katman bulunmaktadır. Evrişim katmanında girdi olarak elde edilen görüntüye çeşitli sayıda, belirli bir boyuta sahip genelleştirilmiş filter uygulanmaktadır. Bu işlemden sonra uygulanan filter sayısı kadar aktivasyon görüntüsü (activation maps) oluşmaktadır. Bu aşamadan sonra evrişim katmanından çıkan görüntüler aktivasyon katmanına girdi olarak iletilmektedir. Daha sonra havuzlama katmanında ayırteci olduğu düşünülen bölgeler bir sonraki katmanda bulunan ikinci bir evrişim katmanına iletilmektedir. Böylece her bir katmanda boyut azala azala tek boyuta kadar düşmektedir. Bundan sonra tam bağlı katmanda bütün nöronlar birbirlerine bağlanarak özellik vektörü oluşturulmaktadır. Son katman olan sınıflandırma (softmax) katmanında ise oluşan öznitelik vektörüne bağlı olarak görüntünün hangi sınıfa ait olduğu saptanmaktadır.



Şekil 5.7 ESA modelinin iç yapısı

Uygulamada önerilen ESA modelinin ilk evrişim katmanında girdi görüntüsüne 96 adet 11x11 boyutlarında genelleştirilmiş filter uygulanmaktadır. Daha sonra aktivasyon katmanından geçirilen her bir evrişim sonucu elde edilen görüntüler 3x3 boyutlarına sahip maksimum havuzlama katmanından geçirilmektedir. Boyutları 3'te 1 oranında azalan girdi görüntüsü ikinci bir evrişim katmanından geçirilmektedir. Bu aşamada 384 adet 3x3 boyutlarında genelleştirilmiş filter uygulandıktan sonra aktivasyon katmanından geçirilip yine 3x3 boyutlardaki maksimum havuzlama katmanından geçirilmektedir. Filtre sayısının fazlaca uygulanmasının sebebi küçük boyutlara sahip hücresel yapıların olabildiğince zengin filter çeidinden geçirilip maksimum ayırtedici özneliliklerin yakalanmasının amaçlanmasıdır. 3 ve 4. Katman grupları (evrişim+relu+havuzlama) yine ikinci katman grubunda uygulanan parametrelerle aynı şekilde uygulanmıştır. 5. katmanda 3x3 boyutlarında 128 filtre ile konvolve edilen görüntü aktivasyon katmanından sonra 3x3 boyutlarında havuzlama katmanından geçirilip tam bağlı katmana iletilmektedir. Tam bağlı katmanda 4096 adet nöron yer almaktadır. Bu aşamada 4096 adet nörondaki bilgiler alınıp vektörel olarak dışarıya aktarıldıktan sonra TBA ve DAA yöntemleri ile boyutları sırasıyla 1024 ve 256'a indirgenmektedir. Son olarak DVM yöntemi ile sınıflandırma işlemi gerçekleştirilmektedir.

5.2.2 Önerilen ESA Yönteminin Başarısının Değerlendirilmesi

Önerilen bu çalışma International Conference on Pattern Recognition (ICPR) konferansında düzenlenen bir yarışma için hazırlanan veri kümesi üzerinde çalıştırılmıştır. Veri kümesinde toplamda 181000 adet mitozsuz ve 749 adet mitozlu hücre yer almaktadır. Veri kümesinde yer alan tüm örnekler dengesiz olduğundan sınıflandırma başarısını oldukça düşürecektir. Bu noktada mitozlu hücrelerin sayıları baz alınarak bir veri kümesi hazırlanmıştır. Hazırlanan veri kümesinde bölümlenebilen 725 mitozlu hücrelerin %80'i olan 580 adet mitozlu ve buna eşit 580 adet mitozsuz hücre eğitim verisi olarak kullanılmıştır. Daha sonra eşit sayıda mitozlu ve mitozsuz hücre test olarak kullanılmıştır. Çizelge 5.2'de boyut indirgmeden önce önerilen yöntemin sınıflandırma başarısını göstermektedir.

Çizelge 5. 2 Boyut indirgeme öncesinde algoritmaya ait sınıflandırma başarısı

	Mitozlu (Tahmin)	Mitozsuz (Tahmin)	Başarım
Mitozlu (Tahmin)	73	72	0,50
Mitozsuz (Tahmin)	4	141	0,97
Sınıflandırma Başarısı			0,73

Çizelge 5.3 ise TBA ve DAA uygulandıktan sonra, boyut 1x256 iken elde edilen sınıflandırma başarısını göstermektedir. Sonuçlar incelendiğinde küçük sayıda örnekler olmasına rağmen boyut indirgeme öncesinde hücrelerin sınıflandırma başarısı nispeten düşükken, boyut indirgeme sonrasında başarımın ciddi oranda arttığı gözlemlenmektedir. Veri kümesindeki örnek sayısı periyodik olarak arttığında başarımın periyodik olarak azalması beklenir ancak veri kümesinde yer alan örnekler kısıtlı olduğundan bu çalışmada deneme imkanı olmamıştır.

Çizelge 5. 3 Boyut indirgeme sonrasında algoritmaya ait sınıflandırma başarısı. Satırlar gerçek verileri, sütunlar ise tahmin edilen verileri göstermektedir.

	Mitozlu (Tahmin)	Mitozsuz (Tahmin)	Başarım
Mitozlu (Tahmin)	136	9	0,937
Mitozsuz (Tahmin)	0	145	1
Sınıflandırma Başarısı			0,968

5.2.3 Sonuçların Değerlendirilmesi

Bu çalışmada derim öğrenme tabanlı evrimsel sinir ağı algoritması ile yüksek çözünürlüklü histopatolojik görüntülerdeki mitozlu hücrelerin tespitine yönelik bir yöntem geliştirilmiştir. Veri kümesinde toplamda 180.000 mitozsuz ve 749 mitozlu hücre yer almaktadır. Verilerin dengesi olması uygulanacak algoritmaların sınıflandırma başarısını direkt olarak etkilemektedir. Bu çalışmada az sayıda bulunan mitozlu hücrelerin sayısı baz alınarak eğitim verisi hazırlanmış ve buna göre de test verileri oluşturulmuştur. Sonraki çalışmalarda dengesiz verilere karşı gürbüz sonuçlar veren Boosting tabanlı sınıflandırma algoritmaları üzerinde durulacaktır. Bunun yanında ESA modeli ile geleneksel yöntemlerin karşılaştırılması da mitozlu hücrelerin tespitini gerçekleştiren çalışmalara katkıda bulunacaktır.

5.3 Mitozlu Hücrelerin Tespitinde Derin Öğrenme ve Geleneksel Özellik Çıkarımı Algoritmalarının Karşılaştırılması

Dünya Sağlık Örgütü istatistiklerine göre meme kanseri kadınlar arasında en sık görülen ve en çok ölüme neden olan kanser türlerinden biridir. Buna karşın bu kanser türünün erken teşhis edilmesi hastalarının hayatının kurtarılması oranını arttırmaktadır. Meme

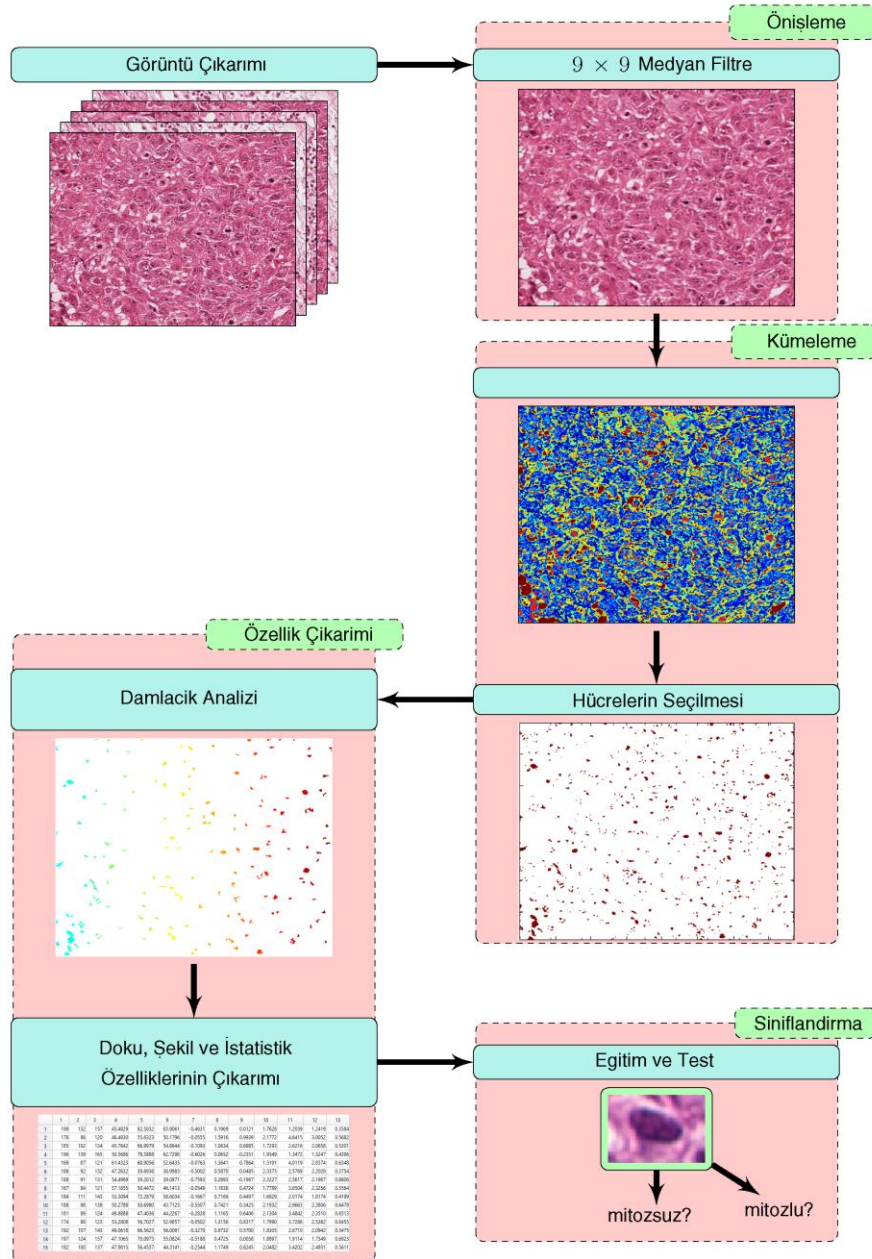
kanseri tanısında en önemli süreçlerden biri mitozlu hücrelerin saptanmasıdır. Belirli bir alanda yer alan mitozlu hücre sayısı meme kanserinin düzeyini belirleme noktasında oldukça önemlidir. Bu çalışmada literatürde gerçekleştirilen geleneksel yaklaşımlardan farklı olarak histopatolojik görüntülerde yer alan mitozlu hücrelerin tespitini amaçlayan evrimsel sinir ağları tabanlı öznelik çıkarma yöntemi kullanılmıştır. Bu yöntemin kullanılması ile elde edilen başarımlar değerleri geleneksel yaklaşımlarla karşılaştırılmıştır. Veri kümesinde yer alan 180.000 mitozsuz ve 749 mitozlu hücre, dengesiz bir veriye sahip olduğundan bu problemi çözmek için boosting tabanlı random under-sampling boosting (RUSBoost) algoritması sınıflandırma algoritması olarak kullanılmıştır. Önerilen yöntem %79.42 hassasiyet, %96,78 kesinlik ve %86.97 F-ölçütü değerleri ile bu veri kümesinde literatürdeki en iyi sonuçlarından birini elde etmiştir.

5.3.1 Önerilen ESA Yöntemi

Bölüm 5.2’te ifade edilen evrimsel sinir ağları yöntemi 2012 yılında Alex vd tarafından önerilen yöntemdir. O çalışmada Alex vd’den farklı olarak tam bağlı katmanlar aşamasında algoritma tarafından oluşturulan öznelik vektörleri alınarak boyut indirgenmiş ve softmax’ten farklı bir sınıflandırma algoritmasına verilerek kısıtlı sayıda mitozlu ve mitozsuz hücre örnekleri sınıflandırılmaya çalışılmıştır.

Bu çalışmada ise yukarıda vurgulanan ESA modelinden farklı olarak çalışma kapsamında küçük boyutlara sahip görüntülere uygulanabilecek bir ESA modeli geliştirilmiştir. Mitozlu veya mitozsuz herhangi bir hücre ortalama 32x32 boyutlarına sahiptir. AlexNet veya GoogleNet gibi ESA modelleri 227x227 veya 224x224 gibi boyutlara sahip görüntülere uygulanmak için geliştirilmiş ESA modelleridir. Küçük boyutlu görüntülerin bu modellerde çalışabilmesi için yeniden boyutlandırılması gerekmektedir. Yeniden boyutlandırma işleminde görüntünün boyutunun büyütülmesi görüntü kalitesini düşüreceğinden ESA modelinin veriyi iyi öğrenememesi anlamına gelebilir. Bunun yerine daha küçük boyutlara sahip görüntülere uygun modellerin geliştirilmesi daha uygun olacaktır. Bölüm 5.2’te ifade edilen önlemler tekrar edilerek hücresel yapıların bölütlenmesi işlemi gerçekleştirilmiştir. Daha sonra elde edilen hücresel yapıların tamamı önerilen bu algorithma kullanılarak mitozlu hücrelerin tespiti işlemi gerçekleştirilmeye çalışılmıştır. Şekil 5.8, bu çalışma kapsamında gerçekleştirilen işlem

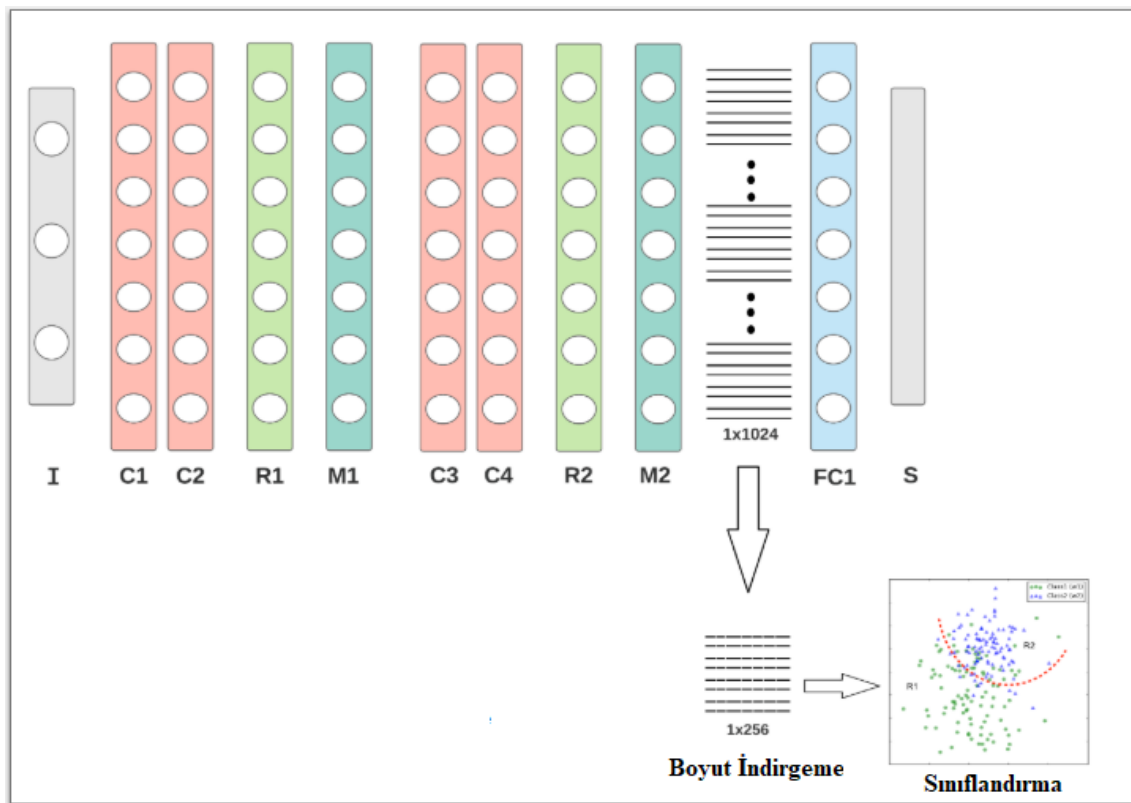
adımlarını ifade etmektedir. Buna göre elde edilen görüntü öncelikle 9x9 ortanca süzgecinden geçirilmiştir. Gürültüden arındırılan görüntü daha sonra 7 küme merkezine ayrılarak hücresel olduğu düşünülen yapılar elde edilmiştir. Son olarak damlacık analizi ile hücresel yapı olamayacak küçüklükteki nesneler elimine edilmiştir. Böylelikle hücresel yapı olduğu düşünülen görüntü kesitleri ESA modeline girdi olarak verilmiştir.



Şekil 5.8 Çalışmada takip edilen işlem adımları

Bu çalışma kapsamında önerilen ESA modelinin katmanları Şekil 5.9'da gösterilmektedir. I, C,R, M, FC ve S sırasıyla girdi görüntüsü, evrişim katmanı,

aktivasyon katmanı, havuzlama katmanı, tam-bağı katmanı ve sınıflandırma katmanlarını ifade etmektedir. Elde girdi görüntüsü klasik ESA modellerinden farklı olarak iki ayrı evrişim katmanından geçirilmiştir. Bunun temel sebebi havuzlama katmanından hemen sonra boyutun en yarıya düşmesinden dolayı yeterince ayırt edici özniteliklerin çıkarılamama ihtimalidir. Bu çalışmayı diğerlerinden ayıran en büyük özelliği girdi görüntü kesitinin boyutunun küçük olmasından dolayı arka arkaya evrişim katmanları kullanılarak ayırt edici olduğu düşünülen özniteliklerin maksimum sayıda çıkarılmaya çalışılmasıdır. Daha sonra havuzlama ile boyutları yarıya düşürülen görüntü kesiti arka arkaya iki evrişim katmanından daha geçirilerek öznitelikler boyutları indirgenmek üzere PCA algoritmasına girdi olarak verilmiştir. Şekil 5.9 önerilen çalışmada kullanılan ESA modelini ve daha sonra veriye uygulanan boyut indirgeme yöntemini ifade etmektedir.



Şekil 5.9 Çalışma kapsamında uygulanan ESA modelinin iç yapısı

Toplamda 180.000 mitozsuz ve 749 mitozlu hücre yer aldığından standart sınıflandırma algoritmaları veriyi öğrenmekte zorlanmaktadır. Genel itibariyle mitozsuz hücreler yüksek başarı oranına yakın bir sınıflandırma başarısı ile sınıflandırılmakta bu durum mitozlu hücrelerin tamamına yakınına mitozsuz denmesine yol açabilmektedir. Her ne

kadar ortalama başarısı yüksek olsa da mitozlu hücrelerin tamamına yakınına mitozsuz demek sistemin hassasiyetinin düşük olduğu anlamına gelecektir. Bu çalışmanın sınıflandırma aşamasında RusBoost algoritması, kullanılarak mitozsuz hücrelerin fazlaca olan sayısının belirli bir benzerlik gösteren örnekleri azaltılarak verinin modellenmesini sağlaması bakımından standart algoritmalarından ayrılmaktadır.

5.3.2 Önerilen Yöntemin Konvansiyonel Yöntemlerle Karşılaştırılması

Önerilen bu çalışmada ESA modeli ile literatürde en çok kullanılan şekil, doku ve istatistiksel yöntemlerin mitozlu hücrelerdeki başarımları karşılaştırılmıştır. Çizelge 5.4'te algoritmaların sınıflandırma performansları gösterilmektedir. Sonuçlar incelendiğinde önerilen yöntemin diğer geleneksel yöntemlerden oldukça başarılı sonuçlar verdiği gözlemlenebilir.

Çizelge 5. 4 Önerilen yöntemin diğer geleneksel özellik çıkarma algoritmaları ile karşılaştırılarak elde edilen başarı oranları

	Kesinlik	Hassasiyet	F-Ölçütü
CLBP	0,857	0,746	0,798
HOG	0,843	0,726	0,780
Haralick	0,807	0,720	0,761
LPQ	0,818	0,695	0,751
SFTA	0,898	0,760	0,824
İstatistik	0,875	0,806	0,839
Önerilen	0,967	0,794	0,869

Mitoz geçiren hücrelerde şekilbilimsel değişimlerin olduğu gözlemlenmektedir. Bu durum hücresel yapılarda gerek dokusal, gerek şekilsel gerekse de istatistiksel olarak değişimlere sebebiyet vermektedir. Bu çalışma kapsamında meme kanseri

histopatolojik görüntülerinde mitozlu hücrelerin tespitini gerçekleştiren ESA algoritmasının kullanıldığı bir yöntem önerilmiştir. Önerilen yöntemin literatürdeki en iyi bilinen algoritmalarından bazıları ile karşılaştırılması yapılarak başarımların performansı değerlendirilmiştir. Elde edilen sonuçlara bakıldığında ESA modelinin mitozlu hücrelerin tespitinde konvansiyonel yöntemlerden daha iyi sonuç verdiği söylenebilir. Ayrıca önerilen yöntem ile diğer yöntemlerin dengesiz verilerde tek başına yüksek başarımlar sağlaması oldukça güçtür. Bu problemi aşmak için dengesiz verilere karşı gürbüz yöntemlerden biri olan RusBoost yöntemi kullanılmıştır.

5.4 Büyük Boyutlu Medikal Görüntülerin Evrimsel Sinir Ağları ile Sınıflandırılması

Sayısal görüntü işleme, görüntüleme teknolojisine sahip cihazlardan elde edilen yüksek boyutlu görüntülerin bilgisayar ortamına aktarıldıktan sonra çeşitli işlem ve analizlerin gerçekleştirilmesi olarak ifade edilmektedir. Bu işlem ve analizler düşük seviye olarak ifade edilebilen, görüntülerin taranması ile ilgili çeşitli problemlerin giderilmesi olabildiği gibi, orta ve yüksek seviyeli işlemler olan hedef bir objenin veya bölgenin tespiti ve takibini amaçlayan işlemler de olabilmektedir [67].

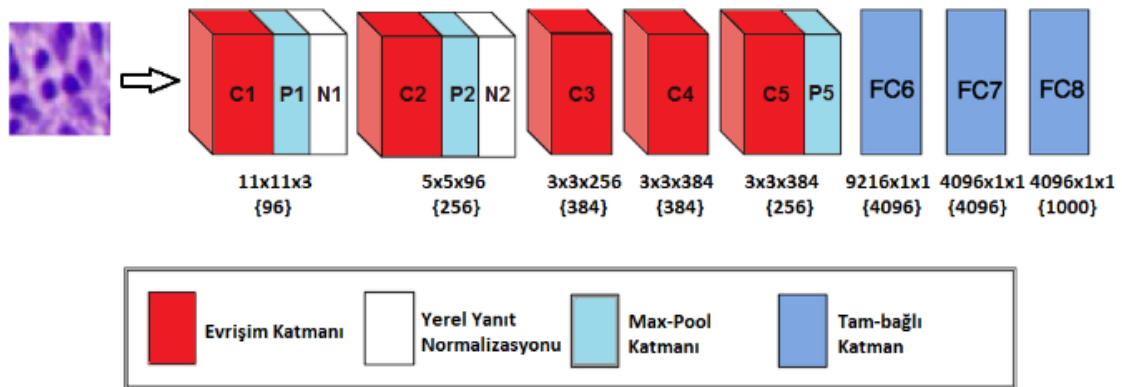
Büyük veri, geleneksel yöntemlerin kolaylıkla analiz edemediği bilgi kümesine verilen isimdir [68]. Teknolojik gelişimlere bağlı olarak birçok kaynaktan elde edilen bilginin işlenmesi ve bu bilgiden anlamlı veri elde edilmesi klasik veri işleme mimarisinde zor ve zaman alıcı olabilmektedir. Veri boyutunun orantısız büyümesi buna karşın klasik veri işleme yöntemlerinin limitli olması bu verilerin analizini zorlaştırmaktadır. Örnek olarak; yüksek çözünürlüklü histopatolojik görüntülerin yüksek tarama özelliğine sahip yeni cihazlarla tarandıktan sonra bilgisayar ortamına aktarılıp işlenmesi klasik veri işleme yöntemlerinde günler hatta haftalar alabilmektedir. Buna karşın son yıllarda yaygın hale gelen çok-düğümlü veya çok çekirdekli paralel işlemciler kullanılarak elde edilen verilerin paralel olarak işlenmesi ve anlamlı verilerin daha hızlı çıkarılması mümkün hale gelmiştir. Literatürde yüksek boyutlu ve yüksek çözünürlüklü histopatolojik görüntülerinin işlenmesi ile ilgili çok sayıda çalışma gerçekleştirilmiştir [69].

Yüksek çözünürlüklü histopatolojik görüntülerde kanserli bölgelerin yüksek başarımla sınıflandırılması hastalığın önlenmesi açısından büyük bir öneme sahiptir. Görüntüleme

teknolojisinin gelişmesi ile birlikte hastalığa ait bilgilerin elde edilmesi kolay hale gelmiştir. Özellikle görüntülerin yüksek boyutlu olması ve sayıca fazlalığı düşünüldüğünde verilerin analizi patologlar için iş yükü oluşturmaktadır. Bunun için patoloji uzmanlarına yardımcı büyük boyuttaki verilerin analizine katkı sağlayan BDT sistemleri gerçekleştirilmektedir. Bu çalışmada ESA yöntemi kullanılarak sentinel lenf düğümünde yer alan metastaz bölgelerinin otomatik tespitini gerçekleştiren bir sistem tasarlanmıştır. 5-kat çapraz geçerlilik ile gerçekleştirilen sınıflandırmada ESA modelinin metastaz bölgelerini tanıma başarımı %98.42 olarak elde edilmektedir.

5.4.1 Görüntü Kesitlerinin ESA Yöntemi ile Sınıflandırılması

Bu çalışmada ESA algoritması, yüksek boyutlu sentinel lenf düğümlerindeki metastaz bölgelerinin tespiti amacıyla kullanılmıştır. Şekil 5.10 çalışmada kullanılan ağ yapısını göstermektedir [70].



Şekil 5.10 Önerilen çalışma kapsamında uygulanan ESA modeli

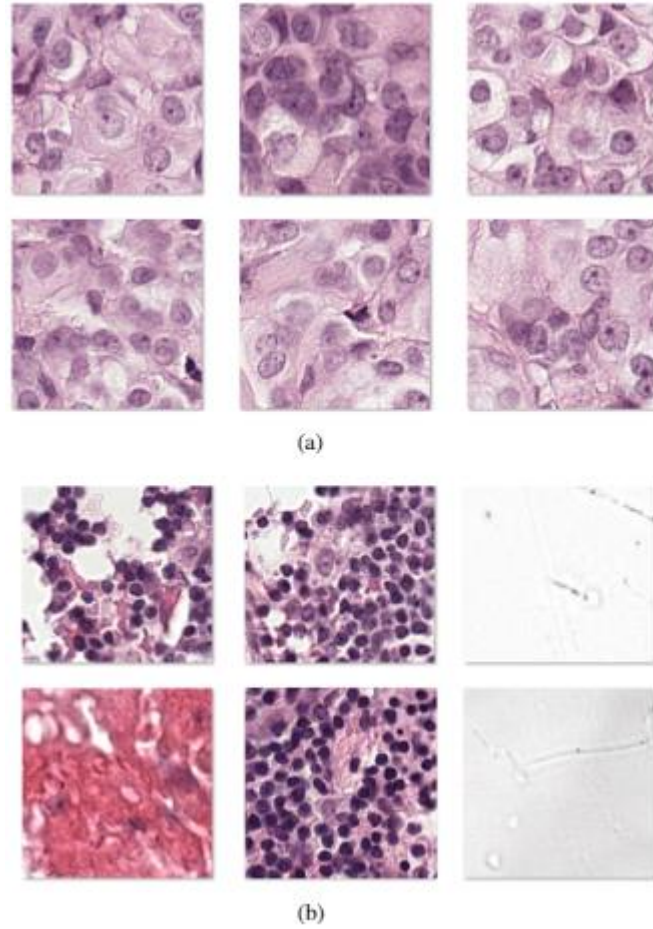
Bu çalışmada önerilen ağın C ile ifade edilen evrişim katmanında girdi görüntüye 11x11 pencere boyutlarında 96 adet genelleştirilmiş filtre uygulanmakta böylece özellik haritaları oluşturulmaktadır. Bu özellik haritaları aktivasyon fonksiyonundan geçirilerek veriyi temsil eden nöronlar aktifleştirilerek bir sonraki katmana girdi olarak sunulmaktadır. P1 katmanı olarak gösterilen havuzlama katmanında ise bu özellik haritalarına ait aktifleştirilmiş bütün değerler tasarıma bağlı olarak ortalama veya maksimumu alınarak geriye kalan değerler sönümlenir. Bir sonraki katmanda bir önceki üç katmana benzer bir yapı söz konusu olduğundan farklı çekirdek yapıları ve farklı özellik haritaları yine aktivasyon fonksiyonlarından geçirilerek üst seviye öznetelik vektörleri elde edilmektedir. Derin öğrenmeyi diğer geleneksel yöntemlerden ayıran en

önemli özelliği veriyi adaptif olarak öğrenebilmesidir. Yalnızca dokusal özniteliklerin ayırt edilmesinde kullanılan yöntemler şekil gibi diğer özellik türlerinde başarılı olamazken, derin öğrenme algoritmaları verinin türüne göre genelleştirilmiş filtreler üreterek ayırtedici öznitelikler elde etmeye çalışmaktadır. Derin öğrenme algoritmaları ilk katmanda dokusal özniteliklere odaklanırken daha üst katmanlarda şekle ve objeye ait özniteliklere odaklanarak en ayırt edici öznitelikleri çıkarmaya çalışılmaktadır. Veri kümesinden 256x256 boyutlarında onbinlerce görüntü kesiti alınarak metastaz bölgesi ve arka plan şeklinde işaretlendikten sonra veriyi modellemesi için evrimsel sinir ağlarına girdi görüntüsü olarak verilmektedir. Daha sonra veri kümesinden alınan çeşitli hastalara ait büyük boyutlu görüntülerden elde edilen 256x256 boyutlarına sahip kesitler ilgili kesitin ön bilgi veya arka plan bilgisi olup olmadığı saptanmaktadır.

5.4.2 Elde Edilen Sonuçların Değerlendirilmesi

Uygulamada kullanılan veri kümesi Radboud University Medical Center (Nijmegen, the Netherlands) ve University Medical Center Utrecht (Utrecht, the Netherlands) tarafından sentinel lenf düğümlerindeki metastaz bölgelerinin tespitinin yapılması amacıyla hazırlanmıştır. Veri kümesinde yer alan bütün görüntüler, patoloji uzmanları tarafından işaretlenmiş yüksek çözünürlüklü ve büyük boyutlu histopatolojik görüntülerden oluşmaktadır. Veri kümesinde yer alan görüntüler hematoxylin ve eosin (H&E) kimyasal boya ile boyanmış ortalama 100.000x200.000 genişlik ve yüksekliğe sahiptir.

Çalışma kapsamında gerçekleştirilen basit bir kullanıcı arayüzü ile veri kümesinde yer alan bütün görüntülerden yaklaşık olarak 48 000 ön bilgi ve arka plan bilgisi içeren 256x256 boyutlarında kesitler alınmıştır. Şekil 5.11 veri kümesine ait çeşitli ön bilgi ve arka plan bilgisini temsil eden görüntüler içermektedir.



Şekil 5.11 Veri kümesinden alınan görüntülerden elde edilen (a) metastaz bölgelerine ait görüntü örnekleri (b) Lam, yağ doku ve sağlıklı dokuları temsil eden görüntü örnekleri

Elde edilen bu görüntüler evrimsel sinir ağı yöntemi ile 5 kat çapraz geçerlilik yöntemi ile sınıflandırılması gerçekleştirilmiştir. Modelin karışım matrisi Çizelge 5.5'de ifade edilmektedir.

Çizelge 5. 5 Çalışma kapsamında önerilen ESA modelinin 5-kat çapraz geçerliliğe sahip karışım matrisi

İlgi Bölgesi / Arkaplan	Doğru	Yanlış
Doğru	4720	80
Yanlış	72	4728

Çizelge 5.6'a göre, 48000 adet görüntü yer alan veri kümesinin 4/5 ile eğitildikten sonra geriye kalan 1/5 ile test edilmesi sonucu elde edilen karışım matrisi ifade edilmektedir. Buna göre, veri kümesinden alınan metastaz görüntülerinin her bir test aşamasında kullanılan 4800 görüntüden 4720 adedi doğru olarak sınıflandırılmış, sadece 80 adet görüntü yanlış olarak sınıflandırılmıştır. Buna karşın metastaz bölgesinin dışındaki bütün normal doku, cam ve kimyasal boya ile işaretli çeşitli dokuların 4800 görüntüsünden 4728 tanesi doğru bir şekilde sınıflandırılırken 72 tanesi metastaz olarak sınıflandırılmıştır.

Çizelge 5. 6 ESA modelinin doğruluk, hassasiyet, F-Ölçütü ve başarı değerleri

İlgi Bölgesi / Arkaplan	Doğruluk	Hassasiyet	F-Ölçütü	Başarı
ESA Modeli	0,9849	0,9833	0,9849	0,9842

Çizelge 5.6 bu çalışma kapsamında uygulanan ESA modelinin doğruluk, hassasiyet, F-ölçütü ve başarı değerlerini göstermektedir. Tabloya göre sistemin doğruluk ve hassasiyet değerlerine bakıldığında sistemin veriyi çok iyi öğrendiği söylenebilir. F-ölçütü sistemin genel başarısını tespit etmede oldukça büyük bir öneme sahiptir. Diğer alanlardan farklı olarak biyomedikal görüntülerin sınıflandırma başarısı değerlendirilirken sadece doğruluk veya sadece başarı değeri referans alınmamalıdır. Tümör içermeyen bir görüntüde sistemin görüntüde tümör bulmaması sistemin sınıflandırma başarısını vurgulamak için tek başına yeterli bir parametre değildir. Buna ek olarak tümör içeren bir görüntüde tümörün varlığının da sistem tarafından tespit edilmesi gerekmektedir. Bu bakımdan hassasiyet (recall) değeri çok önemlidir.

Uygulama kapsamında gerçekleştirilen ESA modeline ait zaman performansı şu şekildedir: 48000 görüntünün 4/5'inin eğitim 1/5'inin de test olarak kullanıldığı durumda zaman performansı yaklaşık olarak 48 saat olmaktadır. ESA modelinin 50 'epoch' ve 100 iterasyon ile gerçekleştirilmesine rağmen 48 saat çok kısa bir süre olarak değerlendirilebilir. Bu noktada yüksek çözünürlüklü ve büyük boyutlu verilerin işlenmesini kolaylaştıran grafik işlemcilerinin katkısı oldukça fazladır. Çalışma 64-bit Windows 10 işletim sistemli, Intel(R) Core(TM) i7-6700K CPU 4.00 GHz, 48 GB RAM ve

NVIDIA Geforce GTX Titan X(12 GB RAM and 3072 Cuda Cores) ekran kartı işlemcisine sahip bir iş istasyonunda gerçekleştirilmiştir. ESA modeline ait bütün deneysel çalışmalar Matlab 2016a ortamında gerçekleştirilmiştir.

Bu çalışmada Radboud University Medical Center (Nijmegen, the Netherlands) ve University Medical Center Utrecht (Utrecht, the Netherlands) tarafından hazırlanan yüksek çözünürlüklü ve büyük boyutlu sentinel lenf düğümleri veri kümesindeki metastaz bölgelerinin tespiti amaçlayan bir derin öğrenme modeli gerçekleştirilmiştir. Çalışma kapsamında 256x256 boyutlarına sahip toplamda 48000 adet görüntü yüksek başarı ile sınıflandırılmaya çalışılmıştır. ESA modeli metastaz görüntülerini 5-kat çapraz geçerlilikle %98.42 başarı oranı ile sınıflandırabilmektedir.

5.5 Histopatolojik Görüntülerde Tümörlü Bölgelerin Evrimsel Sinir Ağlarıyla Sınıflandırılması

Görüntü işleme tekniklerinin hastalık şüphesi bulunan doku veya organların tanısında yardımcı yöntem olarak kullanılması amaçlanan çalışmalarda, ilk olarak ilgi bölgesinin belirlenmesi gerekmektedir. Günümüzde kullanılan mevcut sistemlerde ilgi bölgesinin belirlenmesi patolog tarafından çoğu zaman manuel olarak gerçekleştirilmektedir. Bunun için ilgi bölgesinin sınırları belirlendikten sonra arkaplan bilgisi temizlenmektedir. Daha sonra şüpheli bölgelerde yer alan hücre ve dokular analiz edilerek hastalığa ait teşhisin konması amaçlanmaktadır. Görüntülerin yüksek boyutlu olması ve sayıca fazlalığı düşünüldüğünde bu işlem patologlar için iş yükü oluşturmaktadır. Bu işlem bilgisayar destekli sistemler tarafından otomatik olarak gerçekleştirildiğinde uzmana büyük kolaylık sağlamaktadır [71-72].

Son yıllarda kanser ve kanserin neden olduğu hastalıklarda yaşanan ölümler diğer hastalıkların sebep olduğu ölümlerden daha fazla ön plana çıkmaktadır. Kanser ve ona bağlı hastalıkların erken teşhisi hastalığın tedavi edilebilmesi açısından oldukça önem arz etmektedir. Görüntüleme cihazlarının gelişmesi ile birlikte hastalığın görüntülenmesi, takibi ve Bilgisayar Destekli Teşhis (BDT) sistemlerinin de yardımıyla tedavi edilebilmesi mümkün hale gelmiştir. Özellikle yüksek çözünürlüğe sahip tarayıcılar sayesinde doku ve organlarda yer alan tümörlü bölgelerin BDT tabanlı sistemler tarafından otomatik olarak tespiti mümkündür. Bütün lam görüntüleme

(whole slide imaging) olarak isimlendirilen bu yöntem son dönemde sıkça kullanılan tekniklerden biridir. Bütün slayt görüntüleme mantığı ile taranan her bir görüntünün ortalama boyutu 4-5 gigabayt'lar mertebesinde. Görüntüler Matlab gibi ortamlarda işlenmesi için yüklendiğinde 50 gigabayt'larda bellek gereksinimi duymaktadır. Bu boyuttaki görüntülerin kişisel bilgisayarlarda işlenebilmesi mümkün olamamaktadır. Bu sebeple görüntüler yüksek bellek kapasiteli (48GB) bir sunucu bilgisayarında işlenmiştir. Yüksek kapasiteli bilgisayar kullanımına alternatif olarak görüntülerin tamamının okunması yerine kesit kesit okunabilmesi de mümkündür.

Bu çalışmada lenf nodunda yer alan tümörlü bölgelerin derin öğrenme algırtmaları yardımıyla tespitini sağlayan bir uygulama geliştirilmiştir. Çalışmada Radboud University Medical Center ve University Medical Center Utrecht tarafından hazırlanan sentinel lenf düğümlerindeki metastaz bölgelerinin tespitinin yapılması amacıyla hazırlanan veri kümesine derin öğrenme yöntemlerinden biri olan ESA uygulanmıştır. Sonuçlara bakıldığında uygulanan ESA modelinin 5-kat çapraz geçerlilikle tümörlü bölgeleri tespit etmedeki ortalama başarısı %99.15 olduğu gözlemlenmiştir. Ayrıca ESA modelinin eğitim veri kümesinde yer almayan 2000 adet görüntüdeki başarısı ise %95 mertebesinde. ESA'da eğitim aşamasında veriyi öğrenme süreci oldukça yavaş olurken test aşamasında her bir görüntünün sınıflandırılması oldukça hızlı olabilmektedir.

5.5.1 Lenf Nodlarındaki Tümörün Tespiti için Önerilen ESA Modeli

Bu çalışmada lenf düğümlerinde tümörlü bölgelerin tespiti işlemi için evrimsel sinir ağları (ESA, Convolutional Neural Network) yönteminden faydalanılmıştır. ESA yöntemi literatürde obje tanıma, metin sınıflandırma, biyomedikal görüntü işleme ve içerik tabanlı görüntü tarama olmak üzere birçok alanda başarılı sonuçlar veren bir derin öğrenme algoritmasıdır [73-75]. Çalışmamızda öncelikli olarak elde edilen veri kümesinde yer alan görüntülerden rastgele olarak eşit sayıda elde edilen tümörlü ve tümörsüz bölgelerden kesitler çıkarılmıştır. Daha sonra elde edilen bu görüntüler çeşitli oranlarda eğitim ve test verisine ayrılmıştır. Son olarak algoritmanın işlem adımları Bölüm 5.2.1'de anlatılan ESA yönteminin tümörlü bölgelerin tespitindeki başarısı saptanmaya çalışılmıştır. Bu çalışmada kullanılan girdi görüntüleri ise 256x256

boyutlarındadır. Çalışmada ESA'nın evrişim katmanlarındaki pencere boyutlarının sınıflandırma başarısına etkisi ayrıca incelenmiştir.

5.5.2 Önerilen Yöntemin Başarımının Değerlendirilmesi

Çalışmada kullanılan veri kümesi Bölüm 3.3'te anlatılan Camelyon veri kümesidir. Çalışmada yüksek çözünürlüklü histopatolojik görüntülerin metastaz içeren doku bölgeleri ve normal doku bölgelerinden rastgele olarak görüntü kesitleri seçilmektedir. Elde edilen metastaz ve normal dokuya ait görüntülerin sınıflandırılması, metastazların işaretli olduğu siyah-beyaz görüntüler referans alınarak gerçekleştirilmiştir. Önerilen model, veri kümesinden elde edilen belirli sayıda ilgi bölgesi kesitleri ve arkaplan kesitlerinden oluşturulan eğitim verisi ile eğitilmiştir. Model daha sonra bir başka test veri kümesi ile test edilerek modele ait başarımların performansı değerlendirilmiştir.

Veri kümesinde yer alan bütün görüntülerin en ve boy uzunluklarının 256 ile tam bölünebildiğinden ve 256x256 boyutları tümörlü bölgelerden optimum büyüklüğe sahip olduğundan her bir kesitin 256x256 boyutlarında alınması tercih edilmiştir. Daha büyük boyutlarda görüntü kesitleri sınıflandırma başarısını çok etkileyemeyebilir ancak tümörlü bölgelerin sınırlarında hata payını arttıracaktır. Veri kümesinde yer alan görüntülerden 5000 adet normal doku bölgelerinden ve 5000 adet tümörlü bölgelerden alınmış olmak üzere eğitim aşamasında toplam 10000 adet görüntü kesiti kullanılmıştır. Görüntüler 5-kat çapraz geçerlilik ile eğitilmiştir. Eğitim aşamasında ESA yönteminin evrişim katmanının farklı pencere boyutlarında veriyi öğrenme başarısı Çizelge 5.7'de ifade edilmektedir. Tabloda satır bilgileri sırasıyla ESA modelinde yer alan evrişim katmanlarının boyutlarını ifade ederken; sütun bilgileri ise sırasıyla ESA modelinin veriyi öğrenmede ortalama başarımlar, hata oranı ve veriyi öğrenme süresini ifade etmektedir. ESA modeli için evrişim katmanlarında öğrenme katsayısı 0.01, maliyet fonksiyonunu minimize etmek amacıyla SGD (stochastic gradient descent) kullanılmıştır. Havuzlama katmanlarında 2x2 boyutlu bir pencere kullanılarak veri indirilmesi yapılmıştır.

Çizelge 5. 7 ESA modelinin evrişim katmanında yer alan filtrelerin farklı pencere boyutlarına sahip olması durumunda ESA'nın sınıflandırma performansı

	Doğruluk	Hata Oranı	Çalışma Zamanı(sn)
3x3	0,993	0,023	~2000sn
5x5	0,991	0,019	~2260sn
7x7	0,952	0,259	~2660sn

Çizelge 5.7'de görüleceği üzere evrişim katmanında pencere boyutlarının artması başarımlarını azaltırken, hata oranı ve zaman karmaşıklığını arttırmaktadır. Fakat veriyi öğrenme başarısı her üç farklı pencere boyutlarında da yüksek orandadır.

Çizelge 5. 8 Eğitim aşamasından sonra kaydedilen ESA modelinin farklı 1000 tümörlü ve 1000 normal doku olmak üzere toplamda 2000 adet görüntü kesitindeki sınıflandırma başarısı

	Gerçek Pozitif	Gerçek Negatif	Çalışma Zamanı(sn)
3x3	0,99	0,950	~34,97sn
5x5	1	0,933	~39,14sn
7x7	0,924	0,90	~48,28sn

Çalışmanın ikinci aşamasında ise; ilk bölümde anlatılan eğitim safhasında kaydedilen model ile eğitim safhasında kullanılan verilerden farklı olarak elde edilen 1000 adet normal, 1000 adet tümörlü olmak üzere toplamda 2000 adet görüntü kesitleri kullanılmıştır. Çizelge 5.8 evrişim katmanında farklı pencere boyutlarına sahip ESA ile eğitilen modellerin, eğitim aşamasında kullanılmayan 2000 adet görüntü kesitlerindeki başarımlarını ifade etmektedir. Buna göre evrişim katmanlarında 3x3 pencere boyutları kullanılan ESA modeli, 999 adet tümörlü görüntüyü doğru tahmin ederken, 1 adet görüntüyü yanlış sınıflandırmaktadır. Buna karşılık 950 adet normal görüntüyü doğru sınıflandırırken 50 adet normal görüntüyü yanlış sınıflandırmaktadır. Çizelge 5.8'deki

sütun bilgileri sırasıyla Gerçek Pozitif (GP, True Positive-TP) değerinin tüm pozitiflere oranını (TPR) ve gerçek negatif (GN, True Negative-TN) tüm negatiflere oranını (TNR) göstermektedir. Sistemin her bir görüntü kesitini sınıflandırma süresi ortalama olarak 0.04'tür.

5.5.3 Sonuç

Bu çalışma ESA yönteminin lenf düğümlerindeki metastaz bölgelerinin tespitindeki başarısı incelenmiştir. Çalışma iki aşamadan oluşmaktadır. İlk aşamada yüksek boyutlu histopatolojik görüntülerden 256x256 boyutlarında 5000 adet normal ve 5000 adet tümörlü doku bölgelerinden elde edilen kesitlerin ESA modeli ile eğitilmesi gerçekleştirilmiştir. Eğitilen ESA modelinin evrişim katmanında farklı pencere boyutlarının öğrenmeye etkisi ayrıca incelenmiştir. ESA modeli 5-kat çapraz geçerlilik ile eğitildikten sonra her bir ağ yapısına ait model oluşturulmuştur. Evrişim katmanında 3x3, 5x5 ve 7x7 pencere boyutları kullanıldığında ESA'nın veriyi öğrenme başarısı sırasıyla 0.993, 0.991, ve 0.952 olarak hesaplanmıştır. Çalışmada elde edilen en belirgin sonuç pencere boyutları azaldığında ayırt edici genelleştirilmiş filtreler üretilirken, pencere boyutu arttığında başarı düşmektedir. Bunun dışında daha düşük pencere boyutlarında ESA modeli veriyi daha hızlı öğrenebilmektedir. İkinci aşamada ise ilk aşamada elde edilen modeller kullanılarak eğitim aşamasında hiç kullanılmayan görüntü kesitleri ile modellerin sınıflandırma başarısı sunulmuştur. ESA'da eğitim aşaması oldukça yavaş olurken test aşamasında her bir görüntünün sınıflandırılması oldukça hızlı olabilmektedir. Gelecek çalışmalarda ESA modellerinin öğrenme katsayısı, katman sayısı, filtre sayısı, filtre boyutu farklı aktivasyon fonksiyonlarının sistemin başarımına etkisinin incelenmesi planlanmaktadır.

5.6 Rahim Ağzı (Serviks) Kanseri Öncüsü Lezyonlarının Derecelendirilmesinde

Kullanılmak üzere Oluşturulan Veri Kümesi ve Önerilen Yöntem

Rahim ağzı (serviks) kanseri Human Papilloma Virus (HPV) etkisi ile kanser öncüsü lezyonlarda gelişen ve kadınlar arasında en çok görülen kanser türlerinden biridir. Bu lezyonlara yönelik gerçekleştirilen tarama ile erken teşhis ve tedavi edilebilir bir hastalıktır. Kanseri öncüsü lezyonlar, servikal intraepitelyal neoplazi (CIN) ve skuamöz

intraepitelyal lezyon (SIL) olmak üzere iki farklı sistemde derecelendirilmektedir. Hastalığın tanısında gözlemciler arası uyumsuzluk yüksek olup bu durum aşılmaya çalışılmaktadır.

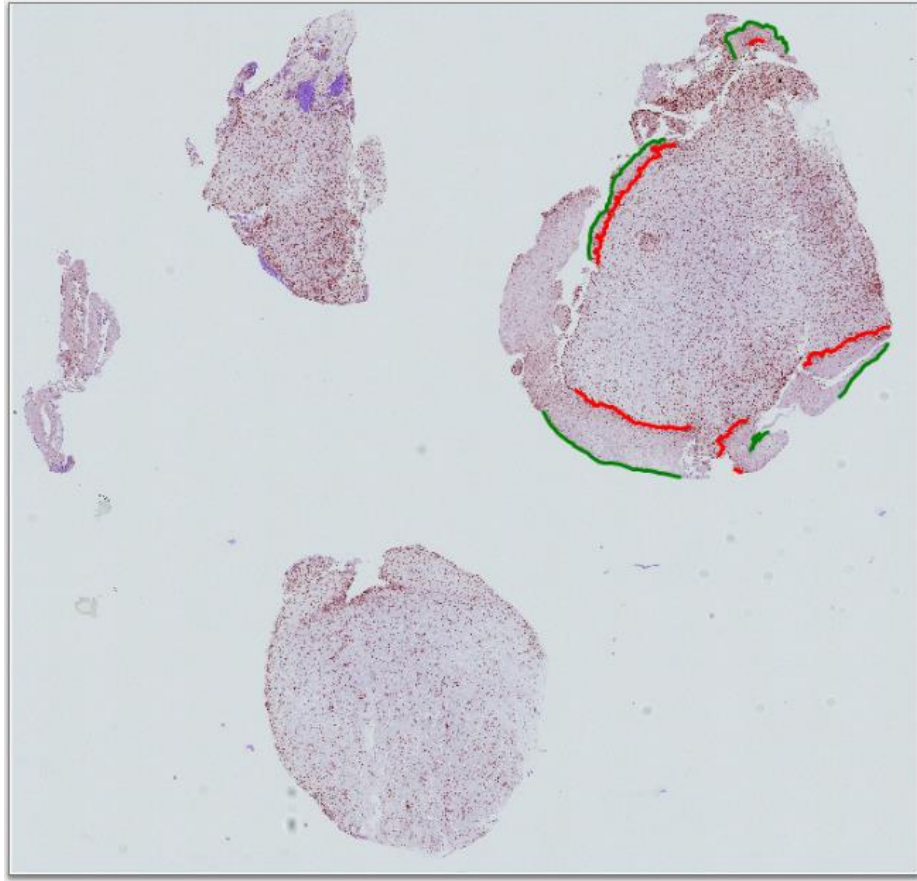
Human Papilloma Virüsünün (HPV) neden olduğu kanser öncüsü lezyonlarda gelişen serviks kanseri, periyodik taramanın yardımıyla önlenabilir kanserlerden biridir. Patologlar arasında yaygın olarak kabul edilen servikal intraepitelyal neoplazi (CIN) ve skuamöz intraepitelyal lezyon (SIL) olmak üzere iki derecelendirme tipi söz konusudur. Bunun yanında kesin tanının konması için gözlemciler (patologlar) arası uyum ve uyumsuzluk oldukça önemlidir. Bu çalışmada son yıllarda oldukça popüler olan bütün slayt görüntüleme tekniği ile elde edilmiş rahim ağzı kanser öncüsü lezyonlarının derecelendirilmesi amaçlanmaktadır. Kanser öncüsü lezyonlarda yer alan hücresel yapıların morfolojik özellikleri ve bazale yakınlıklarına göre derecelendirmeye çalışılmıştır. Elde edilen sonuçlar veri kümesini işaretleyen iki ayrı patolog ile karşılaştırılmıştır. Ayrıca patologlar arası uyum ve uyumsuzluk da bu çalışmanın bir diğer zenginliğidir. Oluşturulan veri kümesi daha sonra bu alanda çalışan bilim insanları ile paylaşılacaktır.

5.6.1 Materyal ve Metod

Bu başlık altında rahim ağzı kanser öncüsü lezyonunda derecelendirme için önerilen morfolojik özelliklerin baz alınarak sınıflandırılması işlemini gerçekleştiren BDT sistemi detaylandırılacaktır.

5.6.1.1 Verinin Elde Edilmesi ve Saklanması

Çalışma kapsamında İstanbul Medipol Üniversitesi Patoloji Laboratuvarında toplamda 54 hastaya ait 127 bütün slayt yüksek çözünürlüklü tarayıcı yardımıyla taranmıştır. HE boyalı preparatlara ait görüntülerin sayısal ortama aktarılmasında tarayıcı olarak Sakura VisionTek Live Digital Microscope kullanılmıştır. Tarayıcı, 64-bit işlem gücüne sahip Windows 7 ortamında, 24 inç çözünürlüklü bir monitör ile uyumlu çalışabilmektedir. Tarayıcı optik olarak 20x, dijital olarak 40x çözünürlüğe kadar tarama özelliğine sahiptir.



Şekil 5.12 İlgi bölgesinin işaretlenmesi (a) Veri kümesinden elde edilen bir görüntü ve, (b) daha detaylı analiz edilmek üzere ilgi bölgesinin işaretlenmesi. Kırmızı ile ifade edilen çizgi bazal membran, yeşil ile ifade edilen çizgi ise üst membranı göstermektedir. Epitel doku bu iki çizgi arasında yer almaktadır.

Hematoksilen-eosin, ki67 ve p16 immünhistokimyasal boyaları ile muamele edilen tüm slaytlar sayısal ortama aktarılmıştır. Bu boyalar uzmanın tanı koymak amacıyla dokulardaki hücresel yapı ile diğer yapıları ayırdetmek için kullanılan kimyasal boyalardır. Çalışma kapsamında geliştirilen bir program yardımıyla sayısal ortama aktarılan bütün slaytlarda epitel (ilgi bölgesi) bölgesi uzman patolog tarafından işaretlenmiştir. Şekil 5.12 veri kümesinden alınan yüksek çözünürlüklü bir adet histopatolojik görüntüyü ve bu görüntüye ait epitel dokunun işaretlenmiş versiyonunu ifade etmektedir.

Alt ve üst sınırları belirlenen epitel daha sonra yaklaşık olarak sabit aralıklarla küçük epitel parçalarına bölünmüştür. Tek başına bütün epitel aynı anda birden fazla tanı içerebileceğinden küçük epitellere bölünmesi yoluna gidilmiştir. toplamda 957 adet

küçük epitel parçaları oluşturulmuştur. Bu epitel parçaları daha sonra iki uzman tarafından işaretlenerek sayısal ortamda saklanmıştır.

Yüksek çözünürlüklü görüntülerin, hematoksilen-eosin (HE) ve immünohistokimyasal (İHK) incelemeleri bir arada yorumlanarak kabul edilen kesin tanıların KEP'lere göre dağılımı üçlü sistemde; 471 adet (%49,2) normal, 240 adet (%25,1) CIN1, 107 adet (%11,2) CIN2, 139 adet (%14,5) CIN3 ve BEP'lere göre dağılımı 150 adet (%46,9) normal, 79 adet (%24,7) CIN1, 34 adet (%10,6) CIN2, 57 adet (%17,8) CIN3'dür (Çizelge 5.9)) İkili sisteme ait tanı dağılımları ise Çizelge 5.10'de gösterilmiştir.

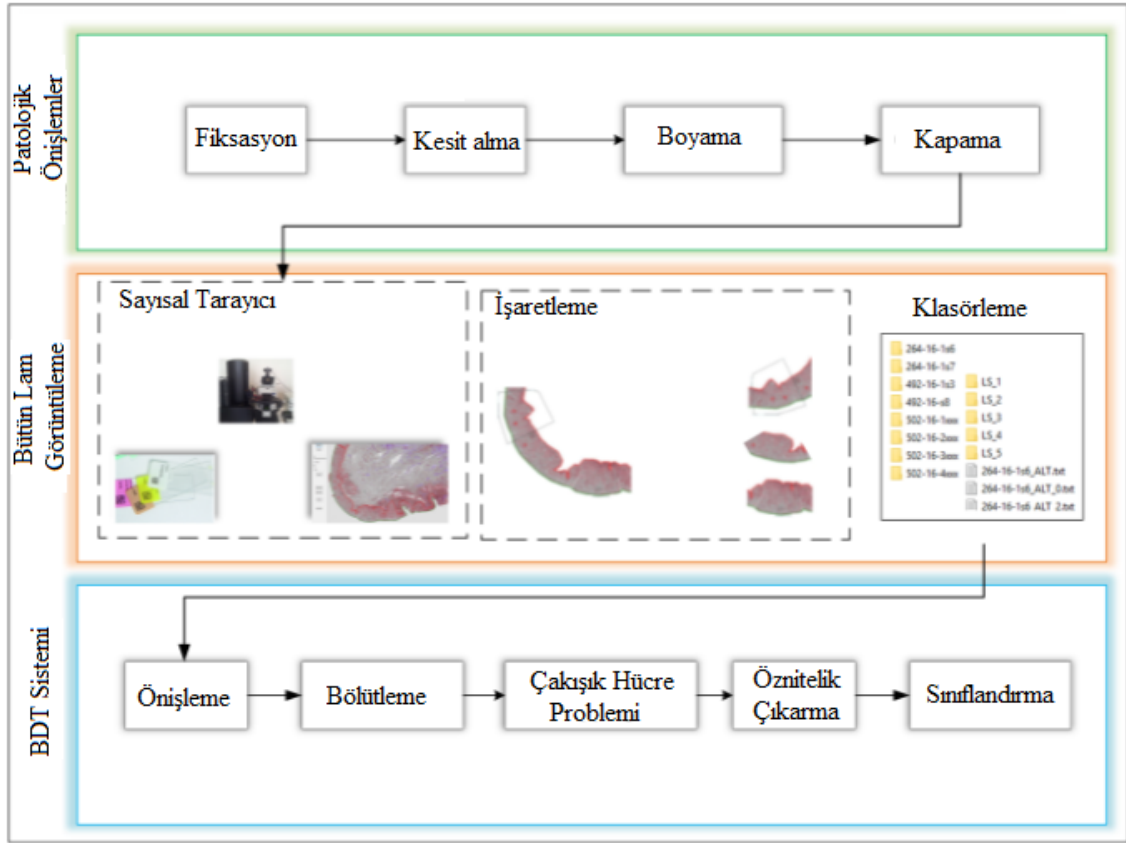
Çizelge 5. 9 CIN tabanlı derecelendirmede büyük epitel parçası (BEP) ve küçük epitel parçası (KEP) sayıları

		Normal	CIN1	CIN2	CIN3	Toplam
CIN	KEP	471	240	107	139	957
Tabanlı Der.	BEP	150	79	34	57	320

Aynı yaklaşım ile kesin tanıların ikili sisteme göre dağılımı 471 adet (%49,2) normal, 240 adet (%25,1) LSIL ve 246 adet (%25,7) HSIL; BEP'lere göre dağılımı 150 adet (%46,9) normal, 79 adet (%24,7) LSIL, 91 adet (%28,4) HSIL'dir. Şekil 5.13 önerilen BST sisteminde takip edilen işlem adımlarını ifade etmektedir. Patolojik önışlemlerden geçirilen doku parçası yüksek çözünürlüklü tarayıcılar yardımıyla sayısal ortama aktarılmaktadır. Daha sonra sayısal ortama aktarılan bu görüntüler görüntü işleme teknikleri ve makine öğrenmesi yöntemleri ile analiz edilmektedir.

Çizelge 5. 10 SIL tabanlı derecelendirmede büyük epitel parçası (BEP) ve küçük epitel parçası (KEP) sayıları

		Normal	LSIL	HSIL	Toplam
SIL	KEP	471	240	246	957
Tabanlı Der.	BEP	150	79	91	320



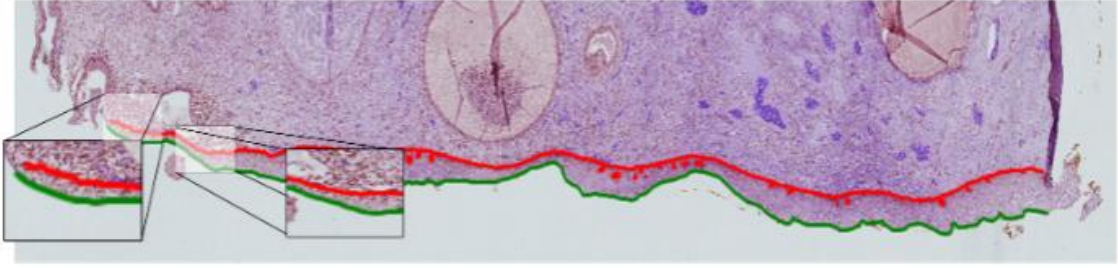
Şekil 5.13 Önerilen bu çalışmada takip edilen işlem adımları. İlk satırda belirtilen işlemler patoloji laboratuvarında gerçekleştirilmiştir. İkinci satırda belirtilen işlemler tıbbi araştırmacılar ve bilgisayar bilimcileri tarafından ortaklaşa yapılmıştır. Üçüncü satırda belirtilen işlemler ise BDT sistemlerinin geliştirilmesi aşamasını göstermektedir.

5.6.1.2 Gözlemciler arası uyum/uyumsuzluk

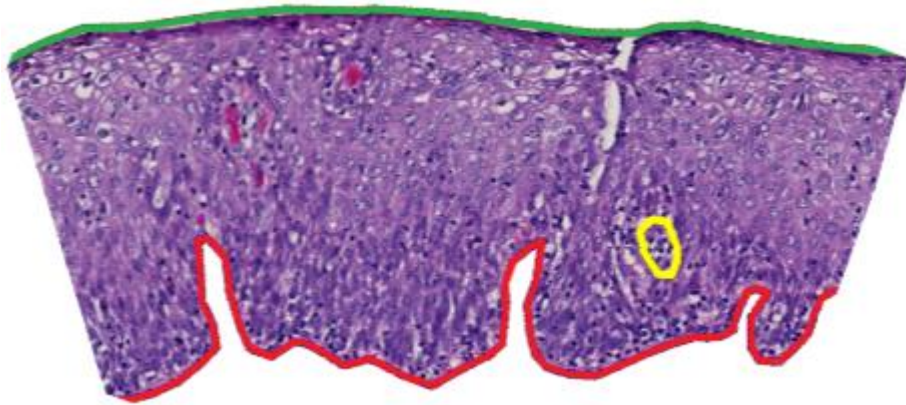
Displaziyi temsil eden şekilbilimsel değişikliklerin yorumları farklı uzman patoloğlar için farklılık gösterebilmektedir. Bunun yanında bazen aynı patoloğ için farklı zamanlarda bakıldığı zaman farklı tanı konabilmektedir. Altın standart oluşturulmaya çalışılırken ki67 ve p16 gibi immunhistokimyasal boyalar yardımıyla kesin tanıya ulaşmak durumunda kalınabilmektedir. Önerilen bu çalışmada servikal displazide tanıya yardımcı bir BDT geliştirilmesi amaçlanmıştır. Bunun için literatürdeki çalışmalara kıyasla oldukça geniş bir veri seti (320 BEP, 957 KEP) iki patoloğ tarafından değerlendirilmiştir. Uyumsuz vakalar tekrar HE, P16 ve Ki67 ile değerlendirilerek kesin tanının doğrulaması gerçekleştirildi. Böylelikle BDT geliştirilirken kullanılan referanslardaki gözlemciler arası uyum ve farklılığı sorunu aşılmıştır. Bu açıdan, mevcut çalışma literatürdeki benzer çalışmalarda görülen gözlemciler arası uyum farklılığına dair dezavantajı ortadan kaldırmaktadır.

5.6.1.3 Morfometrik Öznitelikler Kullanılarak Rahim Ağzı Kanseri Öncüsü Lezyonlarının Sınıflandırılması

Veri kümesinden alınan her bir görüntü Şekil 5.14'te gösterildiği gibi küçük epitellere bölünmektedir. Daha sonra bu epitel parçaları üzerinden tanı işlemi gerçekleştirilmektedir.



Şekil 5.14 Çalışma kapsamında veri kümesinden alınan ve belirli büyüklükte kesilen epitel parçalarının gösterimi.

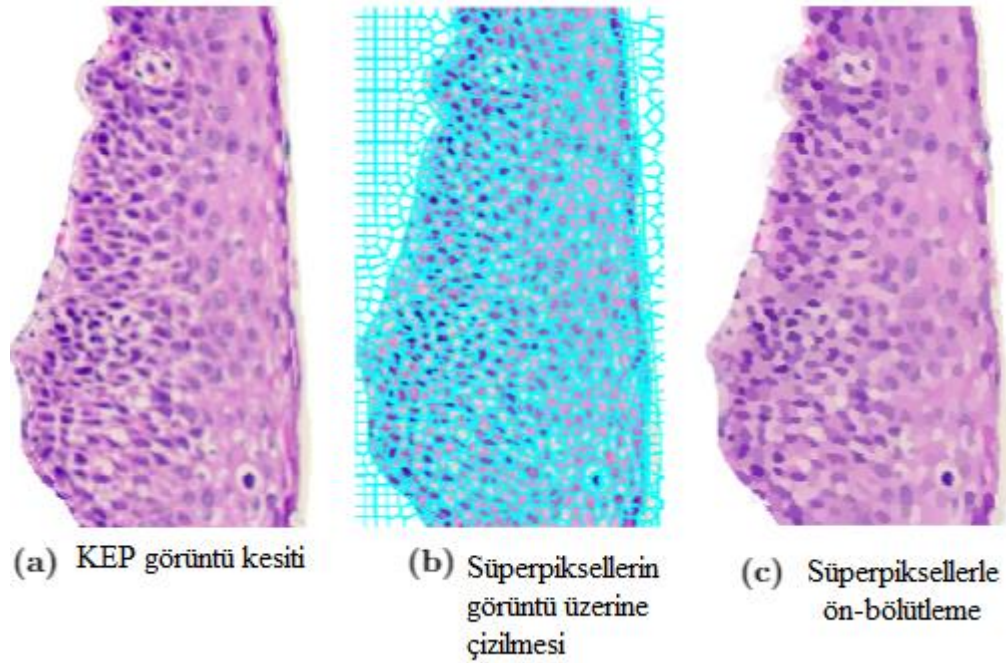


Şekil 5.15 Çalışma kapsamında veri kümesinden alınan ve belirli büyüklükte kesilen küçük epitel parçası. Kırmızı bazal membranı, yeşil çizgi üst membranı ve sarı çizgi ise papilla bölgelerini ifade etmektedir.

Şekil 5.15 sınıflandırılmak üzere elde edilen küçük epitel parçasını ifade etmektedir. Kırmızı ile ifade edilen çizgi bazal membranı, yeşil ile ifade edilen çizgi ise üst membranı ifade etmektedir. Sarı çizgi ile ifade edilen kısım ise tanjansiyel kesitten kaynaklanan bir çıkıntı olan papillayı nitelemektedir.

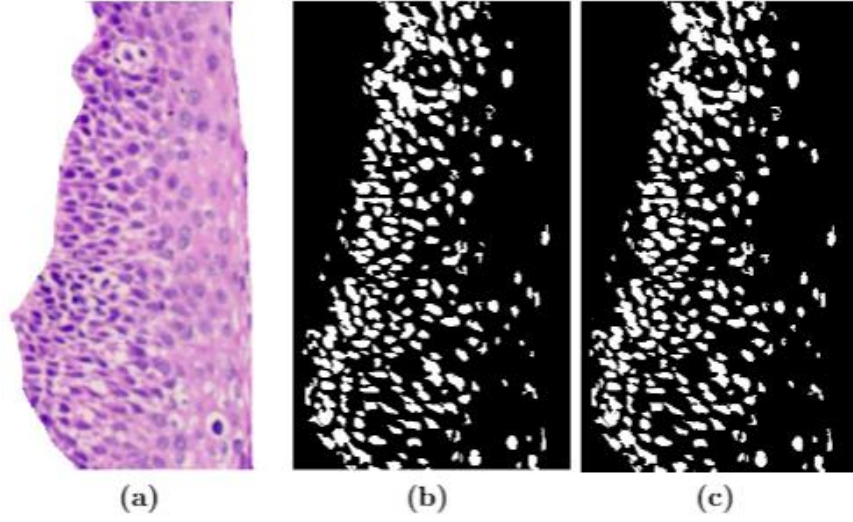
Sınıflandırılmak istenen görüntü ilk olarak 9x9 boyutlarında ortanca süzgecinden geçirilmiştir. Ortanca süzgecinin, ortalama veya gauss gibi diğer süzgeçlerden farkı görüntüdeki gürültüleri giderirken kenar geçişlerine, obje kenarlarına neredeyse hiç

etki etmemesidir. Bu amaçla görüntüler ortanca süzgeci ile yumuşatılmıştır. Ortanca süzgecinden geçirilen her görüntü daha sonra 3000 adet süperpiksele bölünmüştür. Daha sonra görüntüye 4 küme merkezli k-ortalamlar algoritması uygulanmaktadır.



Şekil 5.16 Veri kümesinden alınan örnek bir epitele SLIC süperpiksel bölütme algoritması uygulanması (a) Örnek KEP, (b) 3000 süperpikselin görüntü üzerine yansıtılması, (c) her bir süperpikselin ortalaması hesaplandıktan sonra elde edilen sonuç görüntüsü.

Süperpiksel ve k-ortalamlar algoritmaları uygulandıktan sonra çeşitli küçük artefaktlar elimine edilmiştir. Bu artefaktlar toplamda belirli bir sayının altında (50 bağlı piksel) olan ve hücresel olmayan piksel grubu olarak ifade edilebilir. Bu piksel grupları elimine edildikten sonra çakışık hücre problemi giderilmeye çalışılmıştır. Özellikle bazal membranın çevresinde yer alan çakışık hücrelerin sayıca fazlalığı önerilen BDT sisteminin başarımını direkt olarak etkilemektedir. Hücrelerin yüksek başarı ile ayrıştırılması sistemin performansını direkt olarak arttırmaktadır.

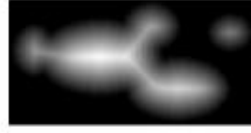


Şekil 5.17 Önerilen yöntemin veri kümesinden elde edilen (a). Örnek girdi görüntüsü (b) bölütleme sonucu ve (c) şekilbilimsel işlemlerden sonra oluşan çıktı görüntüsü.

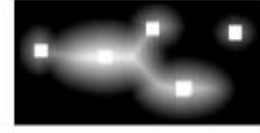
Çakışan hücreler, KEP görüntülerinde yoğun olarak bulunur. Hücre çekirdeklerinin morfolometrik özelliklerinin elde edilebilmesi için hücrelerin çakışması sorunu giderilmelidir. Çakışık hücre probleminin giderilmesi için Şekil 5.18 (a)'da gösterilen ikili formattaki görüntüde her bir beyaz piksel kendisine en yakın siyah piksele olan mesafesi hesaplanmıştır. Böylece Şekil 5.18' (b) gösterilen görüntü elde edilir. Bu görüntüde daha parlak beyaz renk ile ifade edilen pikseller olası çekirdek merkezleri olarak ifade edilebilir. Son olarak Şekil 5.18 (c)'de ifade edilen görüntüde parlaklık değeri en yüksek olan piksel grupları 11x11 boyutlarında bir pencere gezdirilerek lokal maksimumlar tespit edilmeye çalışılır. Bulunan maksimum değer olası merkez olarak işaretlenir. Bu şekilde görüntüde lokal maksimumlar belirlendikten sonra Gaussian Karışım Modeli (GKM, Gaussian Mixture Model) yöntemi ile her bir merkeze elips uydurulmaya çalışılır. Elips formu, hücre şekillerini matematiksel olarak iyi bir şekilde modelleyebilir. GKM, hücre yığınları üzerinde elips uydurmak için uygulanabilecek en iyi yöntemlerden biridir. GKM, istatistiksel veri modellemesi için en yaygın kullanılagelen algoritmalarından biridir [76]. GKM algoritmasının temel amacı, dağılımları, ağırlıklı Gaussian karışımlarının toplamı olarak ifade etmektir.



(a) İkili Görüntü

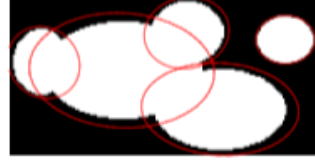


(b) Dönüşüm Transformü



(c) Tahmini Merkezler

Şekil 5.18 Çakışık olarak elde edilen hücresel yapılar. (a) çakışık hücreleri temsil eden ikili görüntü (b) bu ikili görüntüye ait mesafe dönüşümü (Distance Transform) (c) Lokal maksimumların elde edilmesi.



(a) Düzgün Dağılım



(b) Normal Dağılım

Şekil 5.19 Merkezi konumların kestiriminden sonra düzgün ve normal dağılım temelli elipslerin üretilmesi.

5.6.2 Deneysel Sonuçlar

Çizelge 5.11, veri kümesinde yer alan tüm görüntüleri işaretleyen patolojiler arasındaki uyum ve uyumsuzlukları ifade etmektedir. Patolojiler arasındaki uyumsuz görüntü kesitleri daha sonra ki67 ve p16 immunhistokimyasal boyaları ile muamele edilmiş görüntülere bakılarak kesin tanı konmaktadır.

Çizelge 5. 11 Patolog1 ve patolog2'nin birbirlerine göre uyum ve uyumsuzlukları

	Patolog2					
		Normal	CIN1	CIN2	CIN3	Toplam
Patolog1	Normal	354	37	0	0	391
	CIN1	88	156	8	0	252
	CIN2	13	52	71	15	151
	CIN3	4	9	22	128	163
	Toplam	459	254	101	143	957

Çizelge 5.12; patoloğ1, patoloğ2 ve önerilen yöntemin kesin tanı ile uyum ve uyumsuzluğunu CIN-tabanlı derecelendirme sisteminde göstermektedir. Sonuçlar incelendiğinde patoloğ2'nin kesin tanı ile uyumu patoloğ1'e göre daha fazladır. Buna karşın, CIN1 örneklerinde patoloğ2 patoloğ1'e göre daha uyumlu sonuç vermektedir. Kesin tanı ile uyum göz önünde bulundurulduğunda en kötü sonuç önerilen yöntemle ait olduğu gözlemlenmektedir.

Çizelge 5. 12 Kesin Tanının Patoloğ1, Patoloğ2 ve Önerilen Yöntem ile uyumve uyumsuzlukları

CIN Tabanlı Derecelendirme		P1				P2				Önerilen Yöntem			
		N	C1	C2	C3	N	C1	C2	C3	N	C1	C2	C3
Kesin Tanı	N	377	74	18	2	424	44	3	0	386	61	17	7
	C1	13	176	37	14	26	193	16	5	116	91	19	14
	C2	0	2	87	18	8	17	79	3	13	30	47	17
	C3	1	0	9	129	1	0	3	135	4	15	18	102
	Toplam	391	252	151	163	459	254	101	143	519	197	101	140

Çizelge 5.13 patoloğ1 ve patoloğ2'nin kesin tanı ile uyum ve uyumsuzluğunu SIL-tabanlı derecelendirmeye göre göstermektedir. SIL tabanlı sonuçlar incelendiğinde yine patoloğ1'in kesin tanı ile daha uyumlu sonuç verdiği gözlemlenebilir. Çizelge 5.12 ve Çizelge 5.13 karşılaştırıldığında patoloğların SIL tabanlı derecelendirme sisteminde CIN tabanlı derecelendirme sistemine göre kesin tanıyla daha uyumlu sonuçlar verdiği gözlemlenebilir.

Çizelge 5. 13 Kesin tanının patoloj1 (P1) ve patoloj2 (P2)'ye göre uyum ve uyumsuzluk durumu

SIL Tabanlı Derecelendirme		P1			P2		
		N	LSIL	HSIL	N	LSIL	HSIL
Kesin Tanı	N	377	74	20	424	44	3
	LSIL	13	176	51	26	193	21
	HSIL	1	2	243	9	17	220
	Toplam	391	252	314	459	254	244

Çizelge 5. 14 Önerilen yöntemin literatürde diğer çalışmalarda kullanılan Delaunay Triangulation yöntemi ile SIL tabanlı derecelendirmede karşılaştırılması sonucu elde edilen sonuçlar

SIL Tabanlı Derecelendirme		Önerilen			DT		
		N	LSIL	HSIL	N	LSIL	HSIL
Kesin Tanı	N	377	74	20	471	0	0
	LSIL	13	176	51	240	0	0
	HSIL	1	2	243	246	0	0
	Toplam	391	252	314	957	0	0

Çizelge 5.14 önerilen yöntemin SIL tabanlı derecelendirme sistemine göre kesin tanı ile uyum ve uyumsuzluğunu göstermektedir. Önerilen yöntem normal dokuları tespit etmede patoloj1'den daha başarılı sonuç vermesine rağmen diğer örneklerin kestiriminde daha başarısız olduğu söylenebilir.

5.6.3 Sonuç

Önerilen bu çalışmada rahim ağzı kanser öncüsü lezyonlarının derecelendirilmesinde kullanılmak üzere BDT yöntemi önerilmiştir. Önerilen bu yöntem İstanbul Medipol Üniversitesi Hastanesi Patoloji Laboratuvarında taranan 54 hastaya ait yüksek çözünürlüklü görüntülere uygulanmıştır. Sistemin başarısı iki farklı patolog ile karşılaştırılmış, uyum ve uyumsuzluklar tablolarla ifade edilmiştir. Oluşturulan veri kümesi bu alanda çalışan bilim insanları ve araştırma grupları ile paylaşarak problemin çözümüne katkıda bulunulması amaçlanmaktadır. Ayrıca gözlemciler arası uyum ve uyumsuzluklar da analiz edildiğinden literatürdeki birçok çalışmadan o yönüyle ayrılmaktadır. Gelecek çalışmalarda hücrel yapıların daha yüksek başarı ile bölütlenmesi amacıyla farklı bölütleme yaklaşımları uygulanarak sınıflandırma başarısının yükseltilmesi hedeflenmektedir. Bunun yanında hücrel yapılardan ziyade lokal bölgelerin istatistiki özellikleri çıkarılarak mevcut çalışma ile karşılaştırılacaktır.

SONUÇ VE ÖNERİLER

Hastalıkların tanı ve tedavisinde önemli bir yeri olan patolojinin görüntüleme teknolojileri ve yapay zekânın gelişmesi ile birlikte yakın bir gelecekte tanıda büyük bir ilerleme kaydetmesi yadsınamaz bir gerçektir. Patoloji uzmanlarının özellikle yüksek çözünürlüklü tarayıcılar yardımıyla doku ve organlardaki değişimi daha detaylı gözlemleyebiliyor olması karar verme sürecini hem daha hızlı hem de daha doğru yapmasını sağlamaktadır. Bunun yanında taranan biyopsilerin uzaktan erişilebiliyor olması da birden fazla uzmanın görüş alışverişi ile karara yardımcı olmasını sağlamaktadır. Bu durum sayısal patolojinin önemini daha da arttırmaktadır.

Görüntüleme teknolojilerinin yanında bilgisayarla görü, makine öğrenmesi ve görüntü işleme tekniklerinin de gelişmesi ile birlikte uzmana yardımcı ikinci karar destek sistemlerinin gelişmesi de hastalığın tanı ve tedavisine katkı sunan bir diğer parametredir. Hesaplamalı patoloji (computational pathology) sayesinde uzmanın sezgisel olarak değerlendirdiği birçok vakada tespit edilmesi istenen hedef nesne ile ilgili bilgisayar tabanlı sistemler kullanılarak numerik değerlerin elde edilmesi daha nesnel bir sonuç elde edilmesi bakımından oldukça önemlidir. Böylece uzman daha nesnel kıstaslarla sonuca ulaşabilecektir. Özellikle son dönemde geliştirilen görüntü işleme teknikleri sayesinde bazı hastalıkların tespiti otomatik olarak gerçekleştirilmekte, uzmana düşen iş yükü azaltılmaya çalışılmaktadır.

Bu çalışmada geliştirilen makine öğrenmesi ve görüntü işleme teknikleri yardımıyla çeşitli kanser hatalıklarında kullanılmak üzere uzmana yardımcı ikincil karar destek sistemlerinin ön adımı sayılabilecek çeşitli algoritmalar geliştirilmiştir. Önerilen

yöntemler meme kanseri, böbrek karsinomu ve rahim ağzı kanserinde önemli bir yere sahip olan hücresel yapıların bölütlenmesi, sınıflandırılması ve analizinde kullanılmıştır.

Çizelge 5. 15 Çalışma kapsamında önerilen bölütleme yöntemleri, kullanılan süperpiksel yöntemlerinin literatürdeki bilinen algoritmaların performans analizi

	GPO	Kİ	F-Ö	GNO	ÇO	Ç. Zamanı
Irshad vd.	0,76	0,62	0,65	0,96	0,49	-
KO	0,60	0,69	0,61	0,97	0,45	60sn
BCO	0,60	0,68	0,61	0,973	0,447	546sn
Otsu	0,67	0,64	0,62	0,95	0,46	10sn
SLIC+KO	0,69	0,67	0,63	0,96	0,472	27sn
SLIC+BCO	0,66	0,68	0,64	0,96	0,48	365sn
SLIC-DBSCAN	0,74	0,61	0,64	0,94	0,48	183sn
ERS	0,65	0,64	0,61	0,95	0,45	61sn
TPRS	0,65	0,63	0,61	0,96	0,45	487sn
SegNet	0,70	0,71	0,69	0,95	0,53	8sn (Test aşaması)

Çalışmanın bölütleme başlığı altında süperpiksel yaklaşımı ve derin öğrenme tabanlı semantik bölütleme algoritmaları hücresel yapıların bölütlenmesinde kullanılmıştır. Özellikle son yıllarda bilgisayarla görü alanında sıkça kullanılan süperpiksel bölütleme yöntemlerinden SLIC, SLIC-DBSCAN, ERS ve TPRS algoritmalarının yüksek çözünürlüklü histopatolojik görüntülerde hücresel yapıların bölütlenmesindeki başarımlarını elde edilmeye çalışılmıştır. Çizelge 5.15 incelendiğinde Hipotez bölümünde vurgulanan ERS ve TPRS süperpiksel bölütleme algoritmalarının bölütleme performansı literatürdeki yöntemlere nispeten benzer sonuç verdiği gözlemlenmiştir.

SLIC-DBSCAN algoritmasının nispeten gerçek pozitif oranında daha iyi sonuç verdiği gözlemlenmiştir. Ayrıca SLIC süperpiksel bölütleme algoritmasının ön bölütleme olarak kullanıldığı durumda bölütleme başarımını arttırdığı gözlemlenmiştir. Çizelge 5.15 derin öğrenme tabanlı SegNet algoritmasının hücresel yapıların bölütlenmesinde diğer geleneksel yöntemler ve süperpiksel yöntemlerine göre daha başarılı sonuç verdiği gözlemlenmiştir. Bu durum hipotezde vurgulanan durumu onaylaması bakımından oldukça önemlidir.

Hipotez bölümünde irdelenmesi amaçlanan süperpiksel bölütleme algoritmalarının literatürde kullanılan bölütleme algoritmalarına göre nispeten daha iyi sonuç verdiği gözlemlenmiştir. Çalışmada ayrıca SLIC süperpiksel yöntemi ile kümeleme tabanlı bölütleme algoritmaları birleştirilerek farklı bir yaklaşım önerilmiştir. Bununla hücresel yapıların tek başına kümeleme tabanlı algoritmalar yardımıyla bölütlenmesinin hem bölütleme performansı hem de zaman performansı artırılmıştır. Süperpiksel yaklaşımları hiper-parametrik yöntemler olduğundan (süperpiksel sayısı, kompaktlık parametresi kullanıcı tarafından belirlenir) elde edilen görüntünün kaç süperpiksele bölüneceği ve ne kadar kompakt olacağı gibi durumların otomatize edilmesi de sağlanmaya çalışılmıştır. Çalışmada SLIC süperpiksel bölütleme algoritması ayrıca ESA modeli ile birleştirilerek hücresel yapıların bölütlenmesi gerçekleştirilmiştir. Bu yaklaşımda kompaktlık parametresinin değeri süperpiksellerin yaklaşık olarak karesel görüntü alınması sağlandıktan sonra ESA ağına girdi olarak verilmiştir. Böylece görüntüyü oluşturan tüm süperpikseller ESA modeli ile hücresel veya hücresel olmayan süperpikseller olarak iki farklı sınıfa ayrıştırılmaya çalışılmıştır.

Çalışmanın “Histopatolojik Görüntülerin Sınıflandırılması” bölümünde mitozlu hücrelerin çeşitli öznelik çıkarma yöntemleri ve son dönemde oldukça popüler olan evrimsel sinir ağları yöntemi ile tespiti gerçekleştirilmiştir. Mitozlu hücrelerin tespitinde klasik öznelik çıkarma yöntemleri ile ESA yöntemi karşılaştırılarak performans analizi gerçekleştirilmiştir. Mitozlu hücrelerin mitozsuz hücrelere göre sayılarının oldukça az olması (dengesiz ver-imbalanced data) sınıflandırma algoritmalarının veriyi modelleyememesine sebep olmaktadır. Bölüm 5.1, Bölüm 5.2 her iki sınıftan belirli sayıda mitozlu ve mitozsuz hücre elde edilerek oluşturulan küçük bir veri kümesi üzerinde uygulanan yöntemleri tanıtmaktadır. Bölüm 5.3’te ise çalışma

kapsamında kullanılan RusBoost yöntemi ile dengesiz verilerin yüksek başarı ile sınıflandırılması anlatılmıştır.

Çizelge 5. 16 MITOS-ATYPIA-14 veri kümesinde test verisi üzerindeki sınıflandırma başarıları

	K	H	F-Ö
STRASBOURG	-	-	0,024
YILDIZ	-	-	0,167
MINES-CURIE-INSERM	-	-	0,235
CUHK	0,448	0,300	0,356
CasNN [77]	0,411	0,478	0,442
DeepMitosis [78]	0,431	0,443	0,437

Çizelge 5.16 MITOS-ATYPIA-14 veri kümesinde yer alan mitozlu hücrelerin tespiti için gerçekleştirilen çalışmaları göstermektedir. Çizelgedeki ilk üç takım yarışmaya katılan 127 araştırma grubu içinde sonuç gönderen 17 takımdan ilk üçe giren araştırmacıları ifade etmektedir. Sınıflandırma başarısının düşük olmasında test verisi üzerinden anlık olarak değerlendirme yapılması ve dengesiz veri ihtiva etmesinin önemi büyüktür. Bu tezin bir bölümünü oluşturmak amacıyla yarışmaya YILDIZ adı ile sonuç gönderilmiştir. Yarışmada sıralama f-ölçütü üzerinden değerlendirildiğinden K_i ve H değerleri katılımcılarla paylaşılmamıştır. Çizelgedeki son üç satır ise daha sonra gerçekleştirilen çalışmaların sınıflandırma başarısını göstermektedir. Önerilen bu çalışmalar da ilgili veri kümesinin test verileri kullanılarak performans analizi yapmaktadırlar. Çizelge’de dikkat çeken bir diğer ayrıntı da yöntemlerin herhangi bir çapraz geçerlilik olmadan direkt olarak test verisi üzerinde uygulanmasıdır. Başarımın bu kadar farkediyor olmasında ilk üç satırdaki sonuçlar yarışma sürecinde gerçekleşen ve o anda koşturulan programların çıktıları olurken, son üç satırdaki sonuçlar algoritmaların parametrelerinin daha yüksek başarı elde etmek için ayarlanması ile açıklanabilir. Benzer veri kümesinde 5-kat çapraz geçerlilik yöntemi ile elde edilen başarımların sonuçları da Çizelge 5.17’de ifade

edilmektedir. İlk üç satır literatürde gerçekleştirilen çeşitli yaklaşımları göstermektedir. Son 7 yöntem ise bu tez çalışması kapsamında uygulanan yöntemlerin sonuçlarını göstermektedir. Sonuçlar incelendiğinde en iyi kesinlik (precision) değerinin bu tez çalışması kapsamında önerilen ESA modeline ait olduğu görülmektedir. Buna karşın en iyi hassasiyet ve f-ölçütü değerlerinin de DeepCNN algoritması ile elde edildiği gözlemlenmektedir. Bu iki çalışma da ESA'nın önemini vurgulaması açısından oldukça önemlidir.

Çizelge 5. 17 MITOS-ATYPIA-14 veri kümesinde 5-kat çapraz geçerlilik ile elde edilen sınıflandırma başarıları

Yöntemler	K	H	F-Ö
Support vector machine [79]	-	-	0,837
Random forest tree [80]	0,86	0,72	0,78
Deep CNN with handcrafted feature [81]	0,92	0,88	0,90
CLBP	0,857	0,746	0,798
HOG	0,843	0,726	0,780
Haralick	0,807	0,720	0,761
LPQ	0,818	0,695	0,751
SFTA	0,898	0,760	0,824
İstatistik(ort.,std,var,...)	0,875	0,806	0,839
Önerilen ESA modeli	0,967	0,794	0,869

Bu tez çalışmasında önerilen ve uygulanan öz nitelik çıkarım yöntemlerinin başarılarının yüksek çıkmasındaki en büyük neden dengesiz verilere karşı gürbüz olan yöntemlerden biri olan RusBoost sınıflandırma algoritmasının kullanılmasıdır. RusBoost algoritması

kabaca fazla olan verinin azaltılarak veriyi dengeli hale getirip böylece sınıflandırma işlemi yapmaktadır. Bunun yersi ise over-sampling olarak ifade edilen ve az olan sayının kopyalanması ile verilerin dengelenmesini sağlayan algoritmalarıdır. Her iki yaklaşımın birbirlerine göre avantaj ve dezavantajları vardır. Özellik vektörlerinin silinmesi önemli olabilecek özelliklerin kaybolmasına sebep olurken az sayıda bulunan özellik vektörlerinin arttırılması ise sınıflandırma algoritmasının veriyi ezberlemesine sebep olacaktır. RusBoost algoritması kullanılarak elde edilen başarı değerleri Çizelge 5.17’de gösterilmiştir.

Lenf nodlarında yer alan tümörlü bölgelerin tespiti amacıyla oluşturulan veri kümesinde görüntülerin 80.000x100.000 mertebelerinde olduğu düşünüldüğünde hücresel yapıların teker teker analiz edilmesi sayısal sistemler açısından oldukça zor ve zaman alıcı bir işlem olacaktır. Bunun yerine düşük düzeyde hücrelerde gerçekleşen şekilbilimsel değişimin dokuya nasıl etki ettiği analiz edilerek daha kolay çözümler bulunabilir. Bu amaçla çalışma kapsamında önerilen ESA modeli ile tümörlü ve tümörsüz bölgelerden alınan görüntü kesitleri kullanılarak tümörlü bölgelerin yüksek başarı ile sınıflandırılması sağlanmıştır.

Bu tez çalışması kapsamında oluşturulan rahim ağzı kanser öncüsü lezyonların derecelendirilmesini içeren veri kümesi yüksek çözünürlüklü bütün slayt görüntüleme (whole slide imaging) tekniği ile elde edilmiştir. Veri kümesi son dönemde oldukça revaçta olan bütün slaytın taranarak çevrimiçi ortama aktarıldıktan sonra uzman tarafından bilgisayar ortamında incelenmesi bakımından mikroskopla incelemekten daha avantajlıdır. Bu durum uzaktan daha fazla kişi ile konsültasyonu kolaylaştırdığından hastalıkla ilgili daha doğru tespit yapmayı kolaylaştıracaktır. Ayrıca oluşturulan veri kümesinin bilgisayar ortamında bulunması çeşitli görüntü işleme teknikleri uygulanarak verinin anlamlandırılması açısından büyük öneme sahiptir. Oluşturulan veri incelenecek olursa, literatürde bilim insanları tarafından kullanılabilen rahim ağzı kanseri verisi pek bulunamamaktadır. Mevcut çalışmalarda kullanılan veri kümeleri çeşitli ülkelerin lokal hastanelerinden temin edilen klinik veriler olduğundan ve genel olarak paylaşılmadığından yaygın olarak kullanılacak ortak bir veri kümesi bulmak neredeyse imkansızdır. Bu çalışma kapsamında oluşturulan veri kümesi bu anlamda dünyanın her yerinde bu alanda çalışan bilim insanları ile paylaşılabilmesi için

ortak veri kümesi olacaktır. Son olarak, veri kümesindeki her bir görüntü iki uzman patolog tarafından işaretlenmiştir. Özellikle rahim ağzı kanseri öncü lezyonlarının derecelendirilmesinde patologlar arası uyum/uyumsuzluk çok büyük bir öneme sahiptir. Bu veri kümesi ile patologlar ve önerilecek sistemler arası uyum/uyumsuzlukların incelenmesi mümkün olacaktır.

Elde edilen sonuçlar değerlendirildiğinde önerilen yöntemlerin gerek hücresel yapıların, gerek mitozlu hücrelerin tespiti gerekse de uzamsal özellikler kullanılarak tümörlü bölgelerin tespitinde başarılı sonuçlar verdiği söylenebilir. Gelecekte bu alanda çalışmak isteyen bilim insanlarının özellikle dengesiz verilerin sınıflandırılması, çakışık hücre problemlerinin giderilmesi ve büyük boyutlu medikal verilerin analizinin gerçekleştirilmesi noktalarında çalışma yapmaları bu alana oldukça büyük katkı sunacaktır.

KAYNAKLAR

- [1] Gültekin, M., ve Boztaş, G. (2014). "Türkiye kanser istatistikleri.", Sağlık Bakanlığı, Türkiye Halk Sağlığı Kurumu, 43.
- [2] Siegel, R., Ma, J., Zou, Z., ve Jemal, A. (2014). "Cancer statistics, 2014." CA: a cancer journal for clinicians, 64(1): 9-29.
- [3] Huzarski, T., Górecka-Szyld, B., Huzarska, J., Psut-Muszyńska, G., Wilk, G., Sibilski, R., vd. (2017). "Screening with magnetic resonance imaging, mammography and ultrasound in women at average and intermediate risk of breast cancer." Hereditary cancer in clinical practice, 15(1)
- [4] Monticciolo, D. L., Newell, M. S., Moy, L., Niell, B., Monsees, B., ve Sickles, E. A. (2018). "Breast Cancer Screening in Women at Higher-Than-Average Risk: Recommendations from the ACR." Journal of the American College of Radiology, 15(3): 408-414.
- [5] Azizi, S., Bayat, S., Yan, P., Tahmasebi, A., Kwak, J. T., Xu, S., vd. (2018). "Deep Recurrent Neural Networks for Prostate Cancer Detection: Analysis of Temporal Enhanced Ultrasound." IEEE transactions on medical imaging.
- [6] Babaie, M., Kalra, S., Sriram, A., Mitcheltree, C., Zhu, S., Khatami, A., vd. (2017). "Classification and retrieval of digital pathology scans: a new dataset.", 22 May, Hawaii.
- [7] Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., vd. (2017). "QuPath: Open source software for digital pathology image analysis." Scientific reports, 7(1): 16878.
- [8] Mills, A. M., Gradecki, S. E., Horton, B. J., Blackwell, R., Moskaluk, C. A., Mandell, J. W., vd. (2018). "Diagnostic Efficiency in Digital Pathology.", The American journal of surgical pathology, 42(1): 53-59.
- [9] Griffin, J., ve Treanor, D. (2017). "Digital pathology in clinical use: where are we now and what is holding us back?.", Histopathology, 70(1): 134-145.
- [10] Wang, P., Hu, X., Li, Y., Liu, Q., ve Zhu, X. (2016). "Automatic cell nuclei segmentation and classification of breast cancer histopathology images.", Signal Processing, 122, 1-13.
- [11] Tey, W. K., Kuang, Y. C., Ooi, M. P. L., ve Khoo, J. J. (2018). "Automated quantification of renal interstitial fibrosis for computer-aided diagnosis: A

- comprehensive tissue structure segmentation method.”, *Computer methods and programs in biomedicine*, 155: 109-120.
- [12] Su, H., Xing, F., Kong, X., Xie, Y., Zhang, S., ve Yang, L. (2017). “Robust Cell Detection and Segmentation in Histopathological Images Using Sparse Reconstruction and Stacked Denoising Autoencoders.”, In *Deep Learning and Convolutional Neural Networks for Medical Image Computing*, Springer, Cham.
 - [13] Wahab, N., Khan, A., ve Lee, Y. S. (2017). “Two-phase deep convolutional neural network for reducing class skewness in histopathological images based breast cancer detection.”, *Computers in biology and medicine*, 85, 86-97.
 - [14] AlZubaidi, A. K., Sideseq, F. B., Faeq, A., ve Basil, M. (2017). “Computer aided diagnosis in digital pathology application: Review and perspective approach in lung cancer classification.”, In *New Trends in Information & Communications Technology Applications (NTICT)*, 7-9 March, Baghdad.
 - [15] Greer, M. D., Lay, N., Shih, J. H., Barrett, T., Bittencourt, L. K., Borofsky, S., vd. (2018). “Computer-aided diagnosis prior to conventional interpretation of prostate mpMRI: an international multi-reader study.”, *European radiology*, 1-11.
 - [16] Gonzalez, R. C., ve Woods, R. E. 1. (2008). “Digital image processing (3rd ed.).”, Prentice Hall, Upper Saddle River, N.J.
 - [17] Sigirci, I. O., Albayrak, A., ve Bilgin, G. (2015). “Detection of mitotic cells using completed local binary pattern in histopathological images.” In *Signal Processing and Communications Applications Conference (SIU)*, 16-19 May, Malatya.
 - [18] Karabulut, E. M., ve Turgay, I., (2016). “Texture analysis of Melanoma Images for Computer-aided Diagnosis.”, In *Annual Int'l Conference on Intelligent Computing, Computer Science & Information Systems (ICCSIS16)*, Pattaya.
 - [19] Jaworek-Korjakowska, J. (2016). “Computer-aided diagnosis of micro-malignant melanoma lesions applying support vector machines.”, *BioMed research international*, 2016.
 - [20] Albayrak, A., ve Bilgin, G. (2013). “Breast cancer mitosis detection in histopathological images with spatial feature extraction.”, In *Sixth International Conference on Machine Vision (ICMV 2013)*, 16-17 November, London.
 - [21] Mercan, C., Aksoy, S., Mercan, E., Shapiro, L. G., Weaver, D. L., ve Elmore, J. G. (2018). “Multi-Instance Multi-Label Learning for Multi-Class Classification of Whole Slide Breast Histopathology Images.”, *IEEE transactions on medical imaging*, 37(1): 316-325.
 - [22] Su, H., Xing, F., Kong, X., Xie, Y., Zhang, S., ve Yang, L. (2015). “Robust cell detection and segmentation in histopathological images using sparse reconstruction and stacked denoising autoencoders.”, In *International*

Conference on Medical Image Computing and Computer-Assisted Intervention, 5-9 October, Munich.

- [23] Krishnan, M. M. R., Chakraborty, C., Paul, R. R., ve Ray, A. K. (2012). "Hybrid segmentation, characterization and classification of basal cell nuclei from histopathological images of normal oral mucosa and oral submucous fibrosis.", *Expert Systems with Applications*, 39(1): 1062-1077.
- [24] Zhang, X., Su, H., Yang, L., ve Zhang, S. (2015). "Fine-grained histopathological image analysis via robust segmentation and large-scale retrieval.", In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 7-12 June, Boston.
- [25] Kulikova, M., Veillard, A., Roux, L., ve Racoceanu, D. (2012). "Nuclei extraction from histopathological images using a marked point process approach.", In *Medical Imaging 2012: Image Processing* (Vol. 8314, p. 831428). International Society for Optics and Photonics, February, California.
- [26] Xing, F., Xie, Y., ve Yang, L. (2016). "An automatic learning-based framework for robust nucleus segmentation.", *IEEE transactions on medical imaging*, 35(2): 550-566.
- [27] Liu, F., Xing, F., ve Yang, L. (2014). "Robust muscle cell segmentation using region selection with dynamic programming.", *IEEE 11th International Symposium on In Biomedical Imaging (ISBI)*, 29 Apr-2 May, Beijing.
- [28] Zhang, L., Kong, H., Chin, C. T., Liu, S., Chen, Z., Wang, T., ve Chen, S. (2014). "Segmentation of cytoplasm and nuclei of abnormal cells in cervical cytology using global and local graph cuts.", *Computerized Medical Imaging and Graphics*, 38(5): 369-380.
- [29] Shen, P., Qin, W., Yang, J., Hu, W., Chen, S., Li, L., vd. (2015). "Segmenting multiple overlapping Nuclei in H&E stained breast cancer histopathology images based on an improved watershed.", In *IET International Conference on Biomedical Image and Signal Processing (ICBISP 2015)*, 19 November, Beijing.
- [30] Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., ve Madabhushi, A. (2016). "Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images.", *IEEE transactions on medical imaging*, 35(1): 119-130.
- [31] Lu, C., Mahmood, M., Jha, N., ve Mandal, M. (2012). "A robust automatic nuclei segmentation technique for quantitative histopathological image analysis.", *Analytical and Quantitative Cytology and Histology*, 34: 296-308.
- [32] Xu, H., Lu, C., ve Mandal, M. (2014). "An efficient technique for nuclei segmentation based on ellipse descriptor analysis and improved seed detection algorithm.", *IEEE journal of biomedical and health informatics*, 18(5): 1729-1741.
- [33] Al-Lahham, H. Z., Alomari, R. S., Hiary, H., ve Chaudhary, V. (2012). "Automating proliferation rate estimation from Ki-67 histology images.", In *Medical Imaging 2012: Computer-Aided Diagnosis*, 23 February, California.

- [34] Beucher, S., ve Meyer, F. (1992). "The morphological approach to segmentation: the watershed transformation.", Optical Engineering-New York-Marcel Dekker Incorporated-, 34: 433-433.
- [35] Schüffler, P. J., Schapiro, D., Giesen, C., Wang, H. A., Bodenmiller, B., ve Buhmann, J. M. (2015). "Automatic single cell segmentation on highly multiplexed tissue images.", Cytometry Part A, 87(10): 936-942.
- [36] Mouelhi, A., Sayadi, M., Fnaiech, F., Mrad, K., ve Romdhane, K. B. (2013). "Automatic image segmentation of nuclear stained breast tissue sections using color active contour model and an improved watershed method.", Biomedical Signal Processing and Control, 8(5): 421-436.
- [37] Zhang, L., Kong, H., Liu, S., Wang, T., Chen, S., ve Sonka, M. (2017). "Graph-based segmentation of abnormal nuclei in cervical cytology.", Computerized Medical Imaging and Graphics, 56: 38-48.
- [38] Malon, C. D., & Cosatto, E. (2013). "Classification of mitotic figures with convolutional neural networks and seeded blob features.", Journal of pathology informatics, 4.
- [39] Xia, Z., Zhu, Y., Sun, X., Qin, Z., ve Ren, K. (2018). "Towards privacy-preserving content-based image retrieval in cloud computing.", IEEE Transactions on Cloud Computing, 6(1): 276-286.
- [40] Reddy, K. S., Anandan, R., Kalaivani, K., ve Swaminathan, P. (2018). "A comprehensive survey on content based image retrieval system and its application in medical domain.", International Journal of Engineering & Technology, 7(2.31): 181-185.
- [41] Bressan, R. S., Alves, D. H., Valerio, L. M., Bugatti, P. H., ve Saito, P. T. (2018). "DOCToR: The Role of Deep Features in Content-Based Mammographic Image Retrieval.", In 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), 18-21 June, Karlstad.
- [42] Saha, M., Chakraborty, C., & Racocanu, D. (2018). Efficient deep learning model for mitosis detection using breast histopathology images. Computerized Medical Imaging and Graphics, 64: 29-40..
- [43] Hebbar, H., Mushigeri, S., ve Niranjana, U. C. (2014). "Medical image retrieval–performance comparison using texture features.", International Journal of Engineering Research and Development, 9(9): 30-34.
- [44] Tibbi Patoloji, <https://patoloji.gen.tr/>, 31.10.2018
- [45] Irshad, H., Montaser-Kouhsari, L., Waltz, G., Bucur, O., Nowak, J. A., Dong, F., vd. (2014). "Crowdsourcing image annotation for nucleus detection and segmentation in computational pathology: evaluating experts, automated methods, and the crowd.", In Pacific Symposium on Biocomputing, 3-7 January, Hawaii.
- [46] International Conference on Pattern Recognition, <https://mitos-atypia-14.grand-challenge.org/>, 31.10.2018

- [47] Bejnordi, B. E., ve van der Laak, J. (2017). "Camelyon16: grand challenge on cancer metastasis detection in lymph nodes 2016."
- [48] Cervical Cancer Precursor Lesions Data Set, <http://simplab.yildiz.edu.tr/?q=sources>, 31.10.2018.
- [49] MacQueen, J. (1967). "Some methods for classification and analysis of multivariate observations.", In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability (Vol. 1, No. 14, pp. 281-297).
- [50] Bezdek, J. C., Ehrlich, R., ve Full, W. (1984). "FCM: The fuzzy c-means clustering algorithm.", Computers & Geosciences, 10(2-3): 191-203.
- [51] Dempster, A. P., Laird, N. M., ve Rubin, D. B. (1977). "Maximum likelihood from incomplete data via the EM algorithm.", Journal of the royal statistical society. Series B (methodological), 1-38.
- [52] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., ve Süsstrunk, S. (2012). "SLIC superpixels compared to state-of-the-art superpixel methods.", IEEE transactions on pattern analysis and machine intelligence, 34(11), 2274-2282.
- [53] Irshad, H., Montaser-Kouhsari, L., Waltz, G., Bucur, O., Nowak, J. A., Dong, F., vd. (2014). "Crowdsourcing image annotation for nucleus detection and segmentation in computational pathology: evaluating experts, automated methods, and the crowd.", In Pacific Symposium on Biocomputing, 3-7 January, Hawaii.
- [54] LeCun, Y., Bottou, L., Bengio, Y., ve Haffner, P. (1998). "Gradient-based learning applied to document recognition.", Proceedings of the IEEE, 86(11): 2278-2324.
- [55] Lv, Y., Duan, Y., Kang, W., Li, Z., ve Wang, F. Y. (2015). "Traffic flow prediction with big data: A deep learning approach.", IEEE Trans. Intelligent Transportation Systems, 16(2): 865-873.
- [56] Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., ve Muharemagic, E. (2015). "Deep learning applications and challenges in big data analytics.", Journal of Big Data, 2(1): 1.
- [57] Badrinarayanan, V., Kendall, A., ve Cipolla, R. (2015). "Segnet: A deep convolutional encoder-decoder architecture for image segmentation.", arXiv preprint arXiv:1511.00561.
- [58] Shotton, J., Johnson, M., ve Cipolla, R. (2008). "Semantic texton forests for image categorization and segmentation.", In Computer vision and pattern recognition, 2008. CVPR 2008. 23-28 June, Anchorage.
- [59] Sturges, P., Alahari, K., Ladicky, L., ve Torr, P. H. (2009). "Combining appearance and structure from motion features for road scene understanding.", In BMVC-British Machine Vision Conference, 7-10 September, London.

- [60] Isobe, S., ve Arai, S. (2017). "Deep convolutional encoder-decoder network with model uncertainty for semantic segmentation.", In INnovations in Intelligent SysTems and Applications (INISTA), 3-5 July, Gdynia.
- [61] Wang, Z., Li, T., Pan, L., ve Kang, Z. (2017). "Scene Semantic Segmentation From Indoor Rgb-D Images Using Encode-Decoder Fully Convolutional Networks.", International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 18-22 September, Wuhan.
- [62] Naseer, M., Khan, S. H., ve Porikli, F. (2018). "Indoor Scene Understanding in 2.5/3D: A Survey.", arXiv preprint arXiv:1803.03352.
- [63] Chen, X., Kundu, K., Zhang, Z., Ma, H., Fidler, S., ve Urtasun, R. (2016). "Monocular 3d object detection for autonomous driving. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition", 26 June-1st July, Las Vegas.
- [64] Kampffmeyer, M., Salberg, A. B., ve Jenssen, R. (2016). "Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks.", In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 26 June-1st July, Las Vegas.
- [65] Otsu, N. (1979). "A threshold selection method from gray-level histograms.", IEEE transactions on systems, man, and cybernetics, 9(1), 62-66.
- [66] Guo, Z., Zhang, L., ve Zhang, D. (2010). "A completed modeling of local binary pattern operator for texture classification.", IEEE Transactions on Image Processing, 19(6): 1657-1663.
- [67] Gonzalez, R. C., Woods, R. E., ve Eddins, S. L. (2004). "Digital image processing using MATLAB", (Vol. 624), New Jersey: Pearson-Prentice-Hall, Upper Saddle River.
- [68] John Walker, S. (2014). "Big data: A revolution that will transform how we live, work, and think.", 181-183.
- [69] Spanhol, F. A., Oliveira, L. S., Petitjean, C., ve Heutte, L. (2016). "A dataset for breast cancer histopathological image classification.", IEEE Transactions on Biomedical Engineering, 63(7): 1455-1462.
- [70] Krizhevsky, A., Sutskever, I., ve Hinton, G. E. (2012). "Imagenet classification with deep convolutional neural networks.", In Advances in neural information processing systems, Nevada.
- [71] Albayrak, A., ve Bilgin, G. (2013). "Breast cancer mitosis detection in histopathological images with spatial feature extraction." In Sixth International Conference on Machine Vision (ICMV 2013)", 16-17 November, London.
- [72] Mercan, C., Aksoy, S., Mercan, E., Shapiro, L. G., Weaver, D. L., ve Elmore, J. G. (2018). "Multi-Instance Multi-Label Learning for Multi-Class Classification of Whole Slide Breast Histopathology Images.", IEEE transactions on medical imaging, 37(1): 316-325.

- [73] Zhao, R., Ouyang, W., Li, H., ve Wang, X. (2015). "Saliency detection by multi-context deep learning.", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 7-12 June, Boston.
- [74] Girshick, R., Donahue, J., Darrell, T., ve Malik, J. (2016). "Region-based convolutional networks for accurate object detection and segmentation.", IEEE transactions on pattern analysis and machine intelligence, 38(1): 142-158.
- [75] Bušta, M., Neumann, L., ve Matas, J. (2017). "Deep textspotter: An end-to-end trainable scene text localization and recognition framework.", In IEEE International Conference on Computer Vision (ICCV), 22-29 October, Venice.
- [76] Najar, F., Choura, S., El-Borgi, S., Abdel-Rahman, E. M., ve Nayfeh, A. H. (2004). "Modeling and design of variable-geometry electrostatic microactuators.", Journal of micromechanics and microengineering, 15(3): 419.
- [77] Chen, H., Dou, Q., Wang, X., Qin, J., & Heng, P. A. (2016). "Mitosis Detection in Breast Cancer Histology Images via Deep Cascaded Networks.", In AAAI, 12-17 February, Phoenix.
- [78] Li, C., Wang, X., Liu, W., & Latecki, L. J. (2018). "DeepMitosis: Mitosis detection via deep detection, verification and segmentation networks.", Medical image analysis, 45: 121-133.
- [79] Nateghi, R., Danyali, H., & Helfroush, M. S. (2017). "Maximized Inter-Class Weighted Mean for Fast and Accurate Mitosis Cells Detection in Breast Cancer Histopathology Images.", Journal of medical systems, 41(9): 146.
- [80] Paul, A., & Mukherjee, D. P. (2014). "Enhanced random forest for mitosis detection.", In Proceedings of the 2014 Indian conference on computer vision graphics and image processing, 14-18 December, Bangalore.
- [81] Saha, M., Chakraborty, C., & Racocceanu, D. (2018). "Efficient deep learning model for mitosis detection using breast histopathology images.", Computerized Medical Imaging and Graphics, 64: 29-40.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Abdülkadir ALBAYRAK
Doğum Tarihi ve Yeri : 01.03.1987
Yabancı Dili : İngilizce
E-posta : drkadiralbayrak@gmail.com

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Y. Lisans	Bilgisayar Mühendisliği	Yıldız Teknik Üniversitesi	2013
Lisans	Bilgisayar Mühendisliği	Fatih Üniversitesi	2010
Lise	Fen Bilimleri	Kızıltepe Anadolu Lisesi	2004

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2012-2018	Yıldız Teknik Üniversitesi Bilgisayar Mühendisliği	Araştırma Görevlisi
2011-2012	Dicle Üniversitesi Bilgisayar Mühendisliği	Araştırma Görevlisi

YAYINLARI

Makale

1. Albayrak, A., & Bilgin, G. (2018). "Automatic cell segmentation in histopathological images via two-staged superpixel-based algorithms.", Medical & Biological Engineering & Computing, 1-13.
2. Turkmen, H. I., Albayrak, A., Karşligil, M. E., & Kocak, I. (2017). "Superpixel-based segmentation of glottal area from videolaryngoscopy images.", Journal of Electronic Imaging, 26(6): 061608.
3. Sigirci, I. O., Albayrak, A., & Bilgin, G. (2018)., " Detection of mitotic cells in histopathological images using deep versus handcrafted features", Medical & Biological Engineering & Computing, (Değerlendirme aşamasında)
4. Albayrak, A., Unlu, A., Calik N., Capar, A., Bilgin, G., Toreyin, B.U., Muezzinoglu, B., Turkmen, I., Durak-Ata, L., " A Whole Slide Image Grading Benchmark and Tissue Classification for Cervical Cancer Precursor Lesions with Inter-Observer Variability", Scientific reports (Değerlendirme aşamasında)

Bildiri

1. Albayrak, A., & Bilgin, G. (2013). "Breast cancer mitosis detection in histopathological images with spatial feature extraction. In Sixth International Conference on Machine Vision (ICMV 2013) (Vol. 9067, p. 90670L), London.
2. Sigirci, I. O., Albayrak, A., & Bilgin, G. (2015). "Detection of mitotic cells using completed local binary pattern in histopathological images. In Signal Processing and Communications Applications Conference (SIU), 16-19 May, Malatya.
3. Albayrak, A., & Bilgin, G. (2016). "Effects of color spaces to the cell segmentation in histopathological images.", In Signal Processing and Communication Application Conference (SIU), 16-19 May, Zonguldak.

- 4 Albayrak, A., & Bilgin, G. (2016). "Mitosis detection using convolutional neural network based features.", In Computational Intelligence and Informatics (CINTI), 17-19 November, Budapest.
- 5 Albayrak, A., Ünlü, A., Çalık, N., vd. (2017). "Segmentation of precursor lesions in cervical cancer using convolutional neural networks.", In Signal Processing and Communications Applications Conference (SIU), 15-18 May, Antalya.
- 6 Albayrak, A., & Bilgin, G. (2017). "Superpixel approach in high resolution histopathological image segmentation.", In Signal Processing and Communications Applications Conference (SIU), 15-18 May, Antalya.
- 7 Gündüz, K., Albayrak, A., Bilgin, G., & Karşlıgil, M. E. (2018). "Classification of tumor regions in histopathological images using convolutional neural networks. In 2018 26th Signal Processing and Communications Applications Conference (SIU), İzmir.
- 8 Albayrak, A., & Bilgin, G. (2018). "A Hybrid Method of Superpixel Segmentation Algorithm and Deep Learning Method in Histopathological Image Segmentation. In 2018 Innovations in Intelligent Systems and Applications (INISTA), Thessaloniki.

Proje

1. "Yüksek Boyutlu İşaret ve Görüntülerin Grafik İşlem Birimleri ile Paralel Analiz Sistemi", Bilimsel Araştırma Projesi, , Araştırmacı, 2016

ÖDÜLLERİ

1. Albayrak A, Sığircı İ.O., Bilgin G., "MITOS & ATYPIA 14 Contest (3rd Ranking), International Conference on Pattern Recognition (ICPR), Ağustos 2014