

**T.C.
HALIÇ ÜNİVERSİTESİ
FEN BİLİMLER ENSTİTÜSÜ
ELEKTRONİK VE HABERLEŞME ANABİLİM DALI
ELEKTRONİK VE HABERLEŞMEMÜHENDİSLİĞİ PROGRAMI**

**AKUT LENFOSİT LÖSEMİNİN ÇEKİRDEK SAĞRI
REGRESYONU YÖNTEMİYLE TANINMASI**

YÜKSEK LİSANS TEZİ

**Hazırlayan
Merve Ayyüce KIZRAK**

**Danışmanı
Yard. Doç. Dr. Figen ÖZEN**

İstanbul – 2011

ÖNSÖZ

“Akut Lenfosit Lösemisinin Çekirdek Sağrı Regresyonu Yöntemiyle Tanınması” isimli araştırma Haliç Üniversitesi Fen Bilimler Enstitüsü Elektronik ve Haberleşme Anabilim Dalı Yüksek Lisans Programı’nda tez olarak hazırlanmıştır.

Günümüzde kanser çevre etkenleri ve genetik faktörlerin de etkisiyle sıklıkla karşılaşılan hastalıklarda istatistiksel olarak ön sıralarda yer almaktadır. Kanser türleri içinde lösemnin çocuk ve genç yetişkin nüfusta en çok rastlanan kanser türü olduđu saptanmıştır. Bu durum gelecekte lösemnin toplumlarda yaygın bir hal alacađı gerçeđini ortaya çıkarmıştır.

Bu çalışmada çocuklarda daha sıklıkla rastlanan, tedavi şansı yüksek ve tedavi edilmediđi takdirde ölümlle sonuçlanabilen akut lenfosit lösemi (ALL) hücrelerinin tanınması için verimli bir örüntü tanıma algoritması tasarlanmıştır.

Çalışmada ALL ve sağlıklı hücelere morfolojik bir takım ön işlemler uygulanmıştır. Sonrasında gri seviye eş oluşum matrisleri (GLCM) yöntemi kullanılarak öznelik çıkarımları gerçekleştirilmiştir. Elde edilen istatistiksel deđerler, çekirdek sağrı regresyon yöntemi kullanılarak sınıflandırılmıştır. Önerilen algoritma tüm detaylarıyla ve performans analizleriyle birlikte sunulmuştur.

Araştırmanın her aşamasında hiçbir konuda desteđini esirgemeyen danışmanım Haliç Üniversitesi öğretim üyesi Yard. Doç. Dr. Figen Özen’ e, araştırmalarımın çeşitli aşamalarında yorumlarıyla ışık tutan ve kaynak erişimimde destek sađlayan Yıldız Teknik Üniversitesi öğretim üyesi Yard. Doç. Dr. Bülent Bolat’a, araştırmalarım lösemi hücelerinin patolojileri hakkında sunduđu bilgilerle ilerlememi sađlayan hematopatolog Dr. Özden Özer’e, akut lenfosit lösemi ve sağlıklı hücelerden oluşan görüntü veri tabanını paylaştan Dr. Fabio Scotti’ye, araştırmalarım sırasında zaman zaman yollarımızın kesiştiđi Esra Parmaksız’a ve tüm bu çalışmalarımı sürdürürken desteklerini hep hissettiđim arkadaşım K. Sercan Bayram ve çok sevgili aileme şükranlarımı sunarım.

İstanbul, 2011

M. Ayyüce Kızrak

İÇİNDEKİLER

Sayfa No.

KISALTMALAR LİSTESİ.....	II
TABLolar LİSTESİ.....	III
ŞEKİLLER LİSTESİ.....	IV
ÖZET.....	VI
ABSTRACT.....	VII
1. GİRİŞ.....	1
1.1. Kanser.....	1
1.1.1. Kanser Türlerinin Sınıflandırılması.....	2
1.2. Lösemi.....	4
1.2.1. Akut Lösemi.....	5
1.2.1.1 Akut Lenfosit Lösemi (ALL).....	5
1.2.1.2 Akut Myeloid Lösemi (AML).....	5
1.2.2. Kronik Lösemi.....	5
1.2.2.1 Kronik Lenfosit Lösemi (KLL).....	6
1.2.2.2 Kronik Myeloid Lösemi (KML).....	6
1.2.3. Tüylü Hücreli Lösemi (HCL).....	6
1.2.4. T-Hücreli Prolenfositik Lösemi (T-PLL).....	6
1.2.5. Büyük Granüllü Lenfositik Lösemi.....	6
1.2.6. Yetişkin T-Hücreli Lösemi.....	7
2. LÖSEMİLİ HÜCRELERİN NORMAL HÜCRELERDEN AYIRT ETME YÖNTEMLERİ.....	8
2.1. Otsu Metodu (Otsu's Method).....	8
2.2. Renk Tonu, Doygunluk ve Yoğunluk (Hue, Saturation, Intensity).....	9
2.3. Yerel Kontrast Germe, Küresel Kontrast Germe, Kısmi Kontrast Germe, Parlaklık ve Karanlık Germe.....	10
2.3.1. Yerel ve Küresel Kontrast Germe.....	11
2.3.2. Kısmi Kontrast Germe.....	11
2.3.3. Karanlık Kontrast Germe.....	13
2.3.4. Parlaklık Kontrast Germe.....	14
2.4. Otomatik Morfolojik Analiz.....	15
2.4.1. FAB Metodu (French-American-British).....	15

2.5. Gradyan Yöntemi	16
2.6. Yüksek Verimli Eleme Algoritması	23
2.7. K-en Yakın Komşu Yöntemleri	25
2.8. Sınır Noktası (Watershed) Algoritması	26
2.9. Çekirdek Yöntemleri (Kernel Methods)	28
2.9.1. Çekirdek Sağrı Regresyon (Kernel Ridge Regression) Yöntemi.....	28
2.9.2. Destek Vektör Makineleri (Support Vektor Machine) Yöntemi.....	32
2.9.2.1 Vapnik-Chervonenkins Teoremi (VC).....	32
2.9.2.2 Destek Vektör Makinesi Algoritması.....	32
2.10. Zernike Momentleri.....	34
2.11. Dalgacık Dönüşümü.....	36
2.12. Genetik Algoritmalar.....	38
2.12.1. Genetik Algoritmaları Kullanarak Biyometrik Hiperspektral Veri Çalışmaları Bant Seçimi....	39
2.12.1.1 Kodlama Şeması	40
2.12.1.2 Uygunluk Fonksiyonu.....	40
2.12.1.3 Adaptif Operatörler.....	40
3. LÖSEMİLİ HÜCRELERİN TANIMLANMASI VE SİMÜLASYONLAR	43
3.1. Çekirdek Sağrı Regresyonu Sınıflandırma Yöntemi Kullanarak Örüntü Tanıma Algoritması	43
3.2. Gri Seviye Eş Oluşum Matrisleri (Gray Level Cooccurrence Matrices-GLCM) Öznitelik Çıkarımı..	57
3.2.1. Yamukluk (Skewness).....	58
3.2.2. Basıklık (Kurtosis).....	60
3.2.3. Korelasyon	62
3.2.4. Enerji	64
3.2.5. Grup Çıkıntısı (Cluster Prominence)	64
3.2.6. Yerel Homojenlik (Ters Fark Momenti-Inverse Difference Moment).....	64
4. PERFORMANS ANALİZİ	66
4.1. Topluluk Ortalaması (Ensemble Average).....	66
4.2. Ortalama Karesel Hataların Karekökü (Root Mean Squared Error).....	66
5. SONUÇLAR	67
6. KAYNAKLAR.....	69
7. EKLER.....	75
8. ÖZGEÇMİŞ	90

KISALTMALAR

ALL	Akut Lenfosit Lösemi
AML	Akut Myeloid Lösemi
KLL	Kronik Lenfosit Lösemi
KML	Kronik Myeloid Lösemi
HCL	Tüylü Hücreli Lösemi
T-PLL	T-Hücre Prolenfositik Lösemi
RGB	Kırmızı-Yeşil-Mavi Renk Uzayı
CYM	Cyan-Sarı-Magenta Renk Uzayı
HSI	Renk Tonu- Doygunluk-Yoğunluk
LCS	Yerel Kontrast Germe
FAB	French-American-British Morfolojik Sınıflandırması
KM	Çekirdek Metodu
PCA	Temel Bileşenler Analizi
LDA	Lineer Diskriminant Analizi
DVM	Destek Vektör Makineleri
KRR	Çekirdek Sağrı Regresyonu
GA	Genetik Algoritmalar
GLCM	Gri Seviye Eş Oluşum Matrisleri
OKHK	Ortalama Karesel Hataların Karekökü

TABLO LİSTESİ

Sayfa No.

Tablo 1.1. Yaşlara göre erkeklerde en sık görülen kanser tipleri.....	3
Tablo 1.2. Yaşlara göre kadınlarda en sık görülen kanser tipleri.....	3
Tablo 1.3. Ülkelere göre lösemiden kaynaklanan ölüm oranları.....	7
Tablo 2.1. Yönlü kenar belirleme için kullanılabilen süzgeçlerin dürtü yanıtları: (a) ve (b) düşey doğrultudaki kenarlar için; (c) ve (d) yatay doğrultudaki kenarlar için; (e) ve (f) ($\pm 45^\circ$) doğrultusundaki kenarlar için.....	21
Tablo 2.2. Gradyan maskeleri.....	22
Tablo 5.1. Laboratuvar sonuçları ile çekirdek sağrı regresyonu sınıflandırması sonuçlarının karşılaştırılması.....	67
Tablo 5.2. Algoritmanın 6 kez çalıştırılarak elde edilen performans.....	68

ŞEKİL LİSTESİ

Sayfa No.

Şekil 1.1. Kan hücre tipleri.....	4
Şekil 2.1. Otsu metodu uygulaması. (a) Kronik lösemi hücresi gri seviye görüntüsü. (b) Otsu dönüşüm sonucu.....	8
Şekil 2.2. Renk tonu, doygunluk ve yoğunluk (HSI) renk uzayı.....	9
Şekil 2.3. RGB ile HSI algoritmalarının sonuçlarının karşılaştırılması. (a) Akut lösemili hücre. (b) S eleman sonucu. (c) S eleman histogramı. (d) RGB renk uzayı sonucu. (e) HSI renk uzayı sonucu.....	10
Şekil 2.4. Kısmi kontrast germe metodu.....	13
Şekil 2.5. Karanlık kontrast germe metodu.....	14
Şekil 2.6. Parlaklık kontrast germe metodu.....	15
Şekil 2.7. FAB metodu ile lösemi hücrelerinin sınıflandırılması.....	16
Şekil 2.8. Türev operatörleri ile kenar belirleme: (a) Koyu arka plan üzerindeki beyaz bant görüntüsü; (b) Açık arka plan üzerindeki siyah bant görüntüsü.....	17
Şekil 2.9. 1-B kenar belirleme sisteminin blok diyagramı.....	18
Şekil 2.10. 1-B kenar belirleme işlemi.....	18
Şekil 2.11. 2-B kenar belirleme sisteminin blok diyagramı.....	19
Şekil 2.12. Kenar belirleme metodu. (a) Orijinal lösemi hücresi gri-ton görüntüsü. (b) İkili gradyan maskesi. (c) Genişletilmiş gradyan maskesi.....	22
Şekil 2.13. Yüksek verimli eleme algoritması uygulaması. (a) FMC44 tipi lösemi hücresi görüntüsü. (b) Kenar belirleme algoritması. (c) Özyineli segmentasyon. (d) Sınır çıkarımı.....	24
Şekil 2.14. Farklı lösemi hücre tiplerinde elde edilen algoritma sonuçları. (a) FMC44 tipi lösemi hücresi algoritma çıkışı. (b) 3E10 tipi lösemi hücresi algoritma çıkışı. (c) 17G10 tipi lösemi hücresi algoritma çıkışı.....	24
Şekil 2.15. Farklı lösemi hücreleri için topluluk istatistikleri.....	25
Şekil 2.16. K-en yakın komşu yöntemi diyagramı.....	26
Şekil 2.17. Watershed algoritmasının lösemi hücresinin tespit edilmesi. (a) Orijinal kronik lösemi hücre görüntüsü. (b) Gri seviyeli görüntü. (c) İkili yoğunluk dönüşümü. (d) Watershed haritası. (e) Watershed dönüşümü sonucu.....	27
Şekil 2.18. Çekirdek sağrı regresyonu.....	29
Şekil 2.19. Destek vektör makineleri ile sınıflandırma (Mercier ve Lennon, 2003).....	33
Şekil 2.20. Zernike momenti metodu kullanılarak görüntü tanıma. (a) Mikroskopik görüntü. (b) Sayısal hologram görüntüsü. (c) Zernike dönüşümü.....	36
Şekil 2.21. İki boyutlu dalgacık dönüşümü blok diyagramı ve matris gösterimi.....	37
Şekil 2.22. Dalgacık dönüşümünün lösemi hücresine uygulanması. (a) Orijinal lösemi hücre görüntüsü. (b) Dalgacık dönüşümü sonucu.....	38
Şekil 2.23. Bir GA algoritmasının kemik iliği patolojik görüntüsü üzerinde uygulanması. (a) Kemik iliği patolojik görüntüsü. (b) İkili gösterim. (c) İkili görüntünün tersi. (d) Ters alınmış ikili görüntüye sınır doldurma, morfolojik açma ve optimizasyon işlemi. (e) Ters alınmış ikili görüntüye sınır doldurma işlemi uygulanmadan, elde edilen morfolojik açma ve optimizasyon. (f) Önceki iki görüntü	

birbirinden çıkarılarak optimize edilen yeni görüntü. (g) Optimize edilmiş görüntünün hücre filtresi. (h) Görüntünün yer tespiti. (i) Orijinal görüntüdeki olgunlaşmamış öncül hücreyi gösteren görüntü.	42
Şekil 3.1. Çekirdek Sağrı Regresyonu Kullanılarak Tasarlanan ALL tanıma sistemi blok diyagramı.	44
Şekil 3.2. Dikdörtgen Top-Hat fonksiyonu.	45
Şekil 3.3. ALL Hücrelerinin teşhis edilmesi için tasarlanan algoritmada elde edilen çıktılar. (a) Gauss filtre çıkışı. (b) Top-Hat filtre çıkışı. (c) Kontrast ayarı çıkışı. (d) İkili eşikleme ve alan doldurma işlemi. (e) Özellikleri alınmış görüntüde ilgili bölgenin kesilmesi. (f) Üç renk bant bölgesinin birleştirilmesi ile elde edilen son görüntü.	47
Şekil 3.4. ALL Hücrelerinin teşhis edilmesi için tasarlanan algoritmada elde edilen çıktılar. (a) Gauss filtre çıkışı. (b) Top-Hat filtre çıkışı. (c) Kontrast ayarı çıkışı. (d) İkili eşikleme ve alan doldurma işlemi. (e) Özellikleri alınmış görüntüde ilgili bölgenin kesilmesi. (f) Üç renk bant bölgesinin birleştirilmesi ile elde edilen son görüntü.	49
Şekil 3.5. ALL Hücrelerinin teşhis edilmesi için tasarlanan algoritmada elde edilen çıktılar. (a) Gauss filtre çıkışı. (b) Top-Hat filtre çıkışı. (c) Kontrast ayarı çıkışı. (d) İkili eşikleme ve alan doldurma işlemi. (e) Özellikleri alınmış görüntüde ilgili bölgenin kesilmesi. (f) Üç renk bant bölgesinin birleştirilmesi ile elde edilen son görüntü.	51
Şekil 3.6. ALL Hücrelerinin teşhis edilmesi için tasarlanan algoritmada elde edilen çıktılar. (a) Gauss filtre çıkışı. (b) Top-Hat filtre çıkışı. (c) Kontrast ayarı çıkışı. (d) İkili eşikleme ve alan doldurma işlemi. (e) Özellikleri alınmış görüntüde ilgili bölgenin kesilmesi. (f) Üç renk bant bölgesinin birleştirilmesi ile elde edilen son görüntü.	53
Şekil 3.7. ALL Hücrelerinin teşhis edilmesi için tasarlanan algoritmada elde edilen çıktılar. (a) Gauss filtre çıkışı. (b) Top-Hat filtre çıkışı. (c) Kontrast ayarı çıkışı. (d) İkili eşikleme ve alan doldurma işlemi. (e) Özellikleri alınmış görüntüde ilgili bölgenin kesilmesi. (f) Üç renk bant bölgesinin birleştirilmesi ile elde edilen son görüntü.	55
Şekil 4.1. Yamukluk dağılım fonksiyonları.	60
Şekil 4.2. Basıklık dağılım fonksiyonu.	62
Şekil 4.3. Korelasyon fonksiyon dağılımı.	63

GENERAL KNOWLEDGE

Name and Surname : M. Ayyüce Kızrak
Field : Electronics and Communications Engineering
Program : Electronics and Communications Engineering
Supervisor : Assist. Prof. Dr. Figen ÖZEN
Degree Awarded and Date : Master – June 2011

DIAGNOSIS OF ACUTE LYMPHOCYTIC LEUKEMIA USING KERNEL RIDGE REGRESSION METHOD

ABSTRACT

Leukemia is the most common type of cancer throughout the world. It is seen both in children and adults. The rate of occurrence of leukemia varies between 28% to 33% according to age, gender and other specifications. Looking at the picture today, it can be concluded that there will be more leukemia cases in the future. Even though leukemia can be a fatal disease, early diagnosis and treatment can help prevent serious consequences.

In this thesis, an efficient pattern recognition algorithm to identify the acute lymphocytic leukemic (ALL) cells is presented. Children are more likely to develop acute lymphocytic leukemia. It can be treated when diagnosed in time. If not treated properly, it can be fatal.

In the study, morphological preprocessing techniques have been applied to both ALL and healthy cells. Later, using the gray-level co-occurrence matrix (GLCM) method, features are extracted. Statistical values obtained from GLCM have been classified using kernel ridge regression. Finally a decision as to whether the sample is acute lymphocytic leukemic or healthy is made.

The details of the algorithm and the performance analyses are included. Simulations have been done using the MATLAB program.

Keywords: Biomedical Image Processing, Pattern Recognition, Clustering Methods, Feature Extraction.

GENEL BİLGİLER

Adı ve Soyadı : M. Ayyüce Kızrak
Anabilim Dalı : Elektronik ve Haberleşme Anabilim Dalı
Programı : Elektronik ve Haberleşme Mühendisliği
Tez Danışmanı : Yard. Doç Dr. Figen Özen
Tez Türü ve Tarihi : Yüksek Lisans-Temmuz 2011

AKUT LENFOSİT LÖSEMİNİN ÇEKİRDEK SAĞRI REGRESYONU YÖNTEMİYLE TANINMASI

ÖZET

Günümüzde kanser türleri içinde lösemnin çocuk ve genç yetişkin nüfusta %28-%33 arasında deęişen oranla en çok rastlanan kanser türü olduęu saptanmıştır. Bu durum gelecekte lösemnin toplumlarda yaygın bir hal alacaęı gerçeęini ortaya çıkarmıştır. Lösemi ölümcül sonuçlar doğursa da, erken tanı ve tedavi ile önüne geçilmesi mümkündür.

Bu çalışmada çocuklarda daha sıklıkla rastlanan, tedavi şansı yüksek ve tedavi edilmedięi takdirde ölümlle sonuçlanabilen akut lenfosit lösemi (ALL) hücrelerinin tanınması için verimli bir örüntü tanıma algoritması tasarlanmıştır.

Çalışmada ALL ve sağlıklı hücrelere morfolojik bir takım ön işlemler uygulanmıştır. Sonrasında gri seviye eş oluşum matrisleri (GLCM) yöntemi kullanılarak öznitelik çıkarımları gerçekleştirilmiştir. Elde edilen istatistiksel deęerler, çekirdek sağrı regresyon yöntemi kullanılarak sınıflandırılmıştır. Algoritma tüm detaylarıyla ve performans analizleriyle birlikte sunulmuştur.

Simülasyon çalışmalarında MATLAB programı kullanılmıştır.

Anahtar Kelimeler: Biyomedikal Görüntü İşleme, Örüntü Tanıma, Grup Yöntemleri, Öznitelik Çıkarımı.

1. GİRİŞ

Vücudun tüm organlarını hücreler meydana getirir. Hücreler ancak mikroskopla görülebilecek kadar küçük yapıtaşlarıdır. Sağlıklı vücut hücreleri bölünebilme, yenilenebilme yeteneğine sahiptir. Ölen hücrelerin yenilenmesinde ve yaralanan dokuların onarılmasında önemli role sahiptirler. Fakat bu yetenekleri sınırlıdır. Sonsuz sayıda bölünemezler. Her hücrenin hayatı boyunca tekrar edebileceği bölünebilme sayısı vardır.

Buna karşın bazı hücreler bilincini kaybeder, kontrolsüz büyümeye ve bölünmeye başlar. Bu hücrelere kanser hücresi denir. Bu tip hücreler birikerek tümörleri oluşturur. Çevre dokulara yayılarak farklı dokularda çoğalmaya devam eder. Buna metastaz denir. Yerleştikleri dokuların ve organların işlevlerini bozmalarına ise invazyon denir (Vogstein, ve Kinzler, 2002).

Kanser, oluşmaya başladığı organ ve mikroskop altındaki görünüme göre sınıflandırılır. Farklı tipteki kanserler, farklı hızlarda büyür, farklı yayılma biçimleri gösterirler ve farklı tedaviler gerektirir (American Cancer Society, 2010). Tedavi yönteminin hekimlerce doğru olarak seçilebilmesi için hücrenin analizinin doğru şekilde yapılmış olması gerekmektedir. Bunun için gelişmiş elektronik tıp cihazları kullanılmaktadır. Ancak erken ve doğru tanı, seçilecek uygun tedavinin ilk adımını oluşturduğundan, her geçen gün görüntüleme tekniklerinde daha verimli sonuçlar elde etmek gerekmektedir.

1.1.Kanser

Kanser, hücrelerin normal sınırların ötesinde kontrolsüz büyüyerek ve bölünerek dokular üzerinde invazyon ve metastaz etkisine denir (Vogstein, ve Kinzler, 2002). Hücreler DNA replikasyonları esnasında meydana gelen bozulmalar nedeniyle yapı değiştirir (Carlo ve Croce, 2008). Normal vücut hücreleri, boyutlarını ve sayılarını normal koşullarda sabit tutabilirken kanser hücreleri bu duruma ters bir davranış benimser.

Kanser genellikle; % 90-95 oranında yaşam şekli, % 30-35 oranında yanlış beslenme (diyet ve obezite), % 25-30 oranında sigara vb., % 15-20 oranında enfeksiyon, radyasyon ve kanserojen faktörler, %10-15 kalımsal etkenler nedeniyle ortaya çıkmaktadır. Bu tip etkenler hücrelerin genetik yapısında anormalliklere neden olur (National Cancer Institute, 2010), (Anand ve diğ., 2008). Hücrelerde onkojenler ve anomalilerin tümör baskılayıcı genler tarafından düzenlenmesi ve yeniden yapılanması karmaşık işlemlerdir (Carlo ve Croce, 2008).

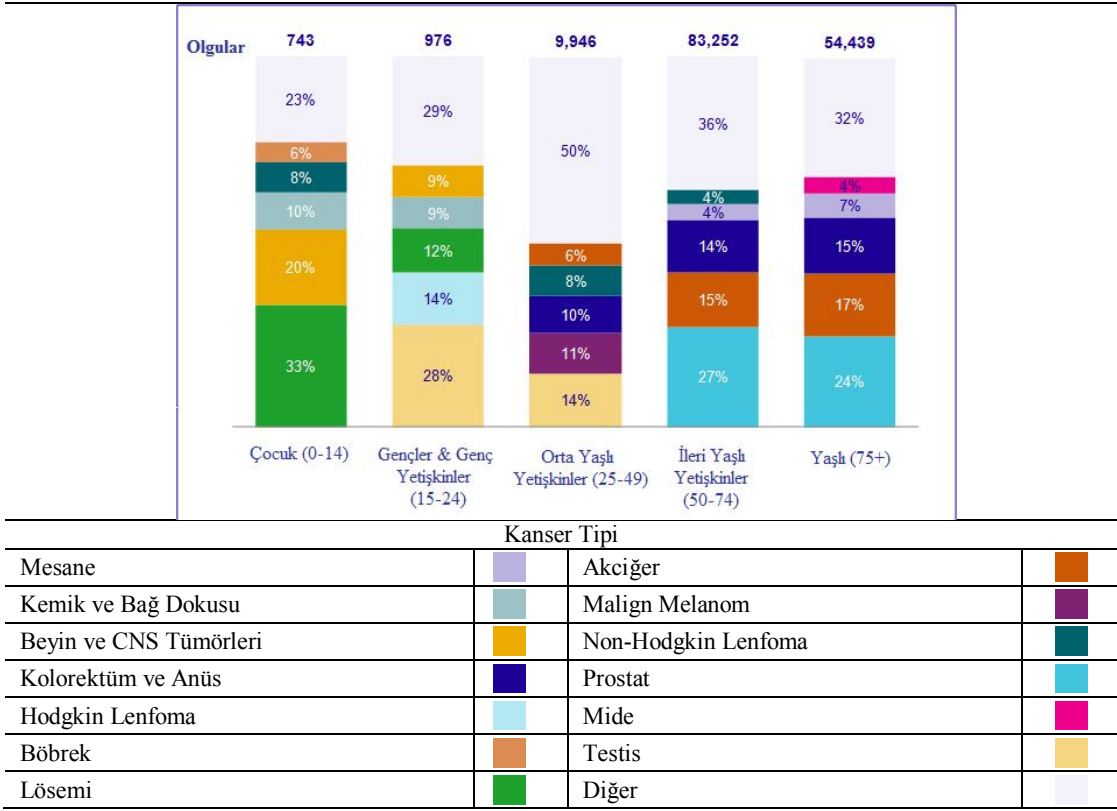
1.1.1. Kanser Türlerinin Sınıflandırılması

Kanserler hücre tipi ve tümör benzerliğine göre sınıflandırılır.

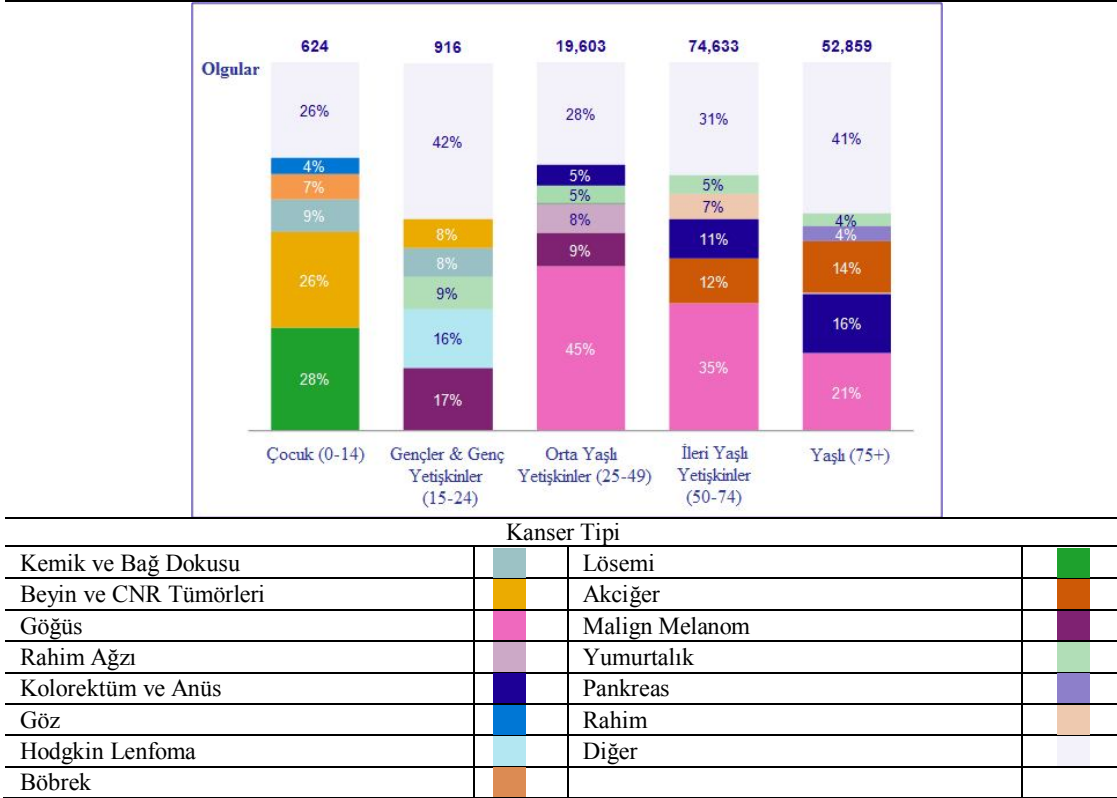
- Karsinom: Kanser, epitelyum hücre kökenlidir. Bu grubun en yaygın tipleri meme, prostat, akciğer ve kolon kanserleridir.
- Sarkom: Kanser, bağ dokusu veya mezokimal kökenlidir.
- Lenfoma ve lösemi: Kanser, hemotopietik (kan yapıcı) hücre kökenlidir.
- Germ hücreli tümör: Kanser, pluripotent hücre kökenlidir. Yetişkinlerde testis ve yumurtalıklarda bulunur, ancak bebek ve çocuklarda da yaygındır.
- Blastom: Kanser, olgunlaşmamış öncü hücrelerde ya da embriyak dokularda oluşur. Çoğunlukla çocuklarda rastlanan bir kanser tipidir (World Health Organization, 2010), (Anand ve diğ., 2010).

Kanser türlerinin görülme sıklıklarına yönelik olarak İngiltere'deki Cancer Research kurumu tarafından yapılan araştırmalarda elde edilen sonuçlar tablo 1.1 ve tablo 1.2.'de özetlenmiştir (Cancer Research UK, 2010).

Tablo 1. 1 Yaşlara göre erkeklerde en sık görülen kanser tipleri.



Tablo 1. 2 Yaşlara göre kadınlarda en sık görülen kanser tipleri.



Bu sonuçlarda çocuk ve gençlerde lösemi yüzdesinin diğer kanser türlerinden daha yoğun olduğu gözlenmektedir. Buradan lösemi hastalığı genetik geçişli bir hastalık olduğu için, gelecek nesillerde de görüleceği sonucu çıkar.

Bu konuda konacak olan tanıyı ve tanıya yönelik doğru tedavi yönteminin seçimini kolaylaştıracak çalışmalara ihtiyaç duyulmaktadır. Bu çalışmanın amacı lösemi kanser hücre yapısının sağlıklı hücreden ayırt edilmesi problemini çözmektir.

1.2.Lösemi

Lösemi, kan ve kemik iliğindeki beyaz kan hücrelerinin anormal şekilde artması karakteristik özelliğidir. Çeşitli türleri bulunan kanser tipidir. Erişkinlerde görülen kanser türlerinin içinde lösemi %90 oranındadır. Lösemi erkeklerde daha sık bir oranda görülmektedir. Ayrıca beyaz ırkta daha yaygındır. Yetişkinlerde lösemi tanısının konması sıklığı çocuklarınkinden %10 daha fazladır ve risk yaşla birlikte artar. Bu yüzden erken tanı çok önemlidir. Hastalık çocuklar arasında 4 yaş altında daha sık gözlenmektedir. Lösemi kısmen de olsa genetik aktarım nedeniyle açığa çıkar (The Leukemia & Lymphoma Society, 2010). Ayrıca, hastalık kan hücresi türünden de etkilenmektedir. Bu açıdan da ikiye ayrılmaktadır.

- Lenfosit

Lenfosit veya lenfoblastik olarak adlandırılırlar. Kanda veya bağışıklık sisteminin lenf bezi, dalak ve timus gibi doku organlarında bulunan enfeksiyonların olduğu bölgelerde bu enfeksiyonlarla savaşan hücredir.

- Myeloid

Beyaz kan hücrelerinin lenfotik olmayan gruplarıdır. Eritrosit, monosit, granülosit myeloid hücrelerine örnektir.



Şekil 1. 1 Kan hücre tipleri.

1.2.1. Akut Lösemi

Kemik iliğinde olgunlaşmamış kan hücreleri hızlı bir şekilde üretilir ve sonuçta sağlıklı (normal) kan hücrelerinden sayıca fazla hale gelebilir. Akut lösemi temel olarak olgunlaşmayan hücrelerin tipleri esas alınarak saptanır. Başlıca iki alt grubu vardır): akut lenfosit lösemi ve akut myeloid lösemi (The Leukemia & Lymphoma Society, 2010), (American Cancer Society, 2010).

1.2.1.1. Akut Lenfosit Lösemi (ALL)

Normalde lenfosit adı verilen olgun kan hücresi tipine dönüşmesi gereken lenfoblast isimli olgunlaşmamış kan hücrelerinin artması akut lenfosit lösemisinin karakteristik özelliğidir. Lenfoblastların sayıları çok miktarda artar ve genelde lenf düğümlerinde birikerek şişliklere neden olurlar. ALL, en sık gözlenen çocukluk çağı kanseridir ve 15 yaş altındaki çocuklarda gözlenen lösemilerin %80'ini oluşturur. Bazen yetişkinlerde de gözlenir, ancak 50 yaş üzerinde oldukça nadirdir.

1.2.1.2. Akut Myeloid Lösemi (AML)

Myeloblast adı verilen ve normal kan hücrelerine dönüşmesi gereken olgunlaşmamış kan hücrelerinin üretimi akut myeloid lösemisinin karakteristik özelliğidir. Olgunlaşmamış bu hücreler kemik iliğinde çok yüksek sayılara ulaşırlar ve normal kan hücrelerinin üretimini azaltırlar. Sonuçta anemi (kansızlık-kırmızı kan hücresi üretiminde azalma) ve sık enfeksiyona yakalanma (beyaz kan hücresi üretiminde azalma) durumu ortaya çıkabilir. Ergenlik çağında ve 20'li yaşlarda saptanan lösemilerin %50'sini, yetişkinlerdeki lösemilerin de %20'sini AML oluşturur.

1.2.2. Kronik Lösemi

Görünüşte olgun ancak normal kan hücrelerinin yaptıklarını yapamayan kan hücrelerinin aşırı üretimi kronik lösemisinin karakteristik özelliğidir. Kronik lösemi daha yavaş ilerler ve sonuçları daha az yıkıcıdır. Başlıca iki alt grubu vardır: kronik lenfosit lösemi ve kronik myeloid lösemi (The Leukemia & Lymphoma Society, 2010), (American Cancer Society, 2010).

1.2.2.1. Kronik Lenfosit Lösemi (KLL)

Olgun görünüşe sahip lenfositlerin kemik iliğinde aşırı üretimi ile kendini gösterir. Bu anormal hücreler tam olarak olgunlaşmamış normal lenfositler gibi görünür, ancak normal lenfositler gibi vücudu enfeksiyonlara karşı koruyamazlar. KLL’de kanser hücreleri kemik iliğinde, kanda ve lenf düğümlerinde bulunurlar ve lenf düğümlerinde şişmeler meydana gelir. KLL tüm lösemilerin %30’unu oluşturur. 30 yaşın altında nadiren görülür ancak görülme sıklığı yaşla birlikte artar. En çok 60-70 yaş arasında gözlenir.

1.2.2.2. Kronik Myeloid Lösemi (KML)

Olgun görünüşlü ancak fonksiyon kaybı bulunan myeloid hücrelerin (beyaz kan hücreleri gibi) aşırı üretimi ile kendini gösterir. Bu aşırı üretim hiç normal hücre kalmayana kadar devam eder. KML hastası olanlarda sıklıkla Philadelphia kromozomu denilen kromozom anomalisi ortaya çıkar. Bu kromozom anomalisinde bu hastalığa neden olan bir enzimin üretilmesine neden olan bir genin olduğu düşünülmektedir. KML yetişkinlerde gözlenen lösemilerin %20-30’unu meydana getirir. 25-60 yaşları arasında gözlenir. Bu hastalık kemik iliği nakli ile tedavi edilebilir.

1.2.3. Tüylü Hücreli Lösemi (HCL)

KLL’nin bir alt kümesi olarak anılsa da, aslında bu tip lösemnin kalıplarına uymamaktadır. HCL hastalarının %80’ini yetişkin erkekler oluşturur. Küçük çocuklarda da tespit edilmektedir. Bu hastalık çaresiz gibi görünse de erken teşhisinde tedavi edilebilir (Else ve diğ., 2005).

1.2.4. T-Hücre Prolenfositik Lösemi (T-PLL)

Genellikle yetişkinleri etkileyen çok nadir ve agresif bir lösemi türüdür. Kadınlardan çok erkeklerde görülmektedir. Tedavisi oldukça zordur (The Leukemia & Lymphoma Society, 2010).

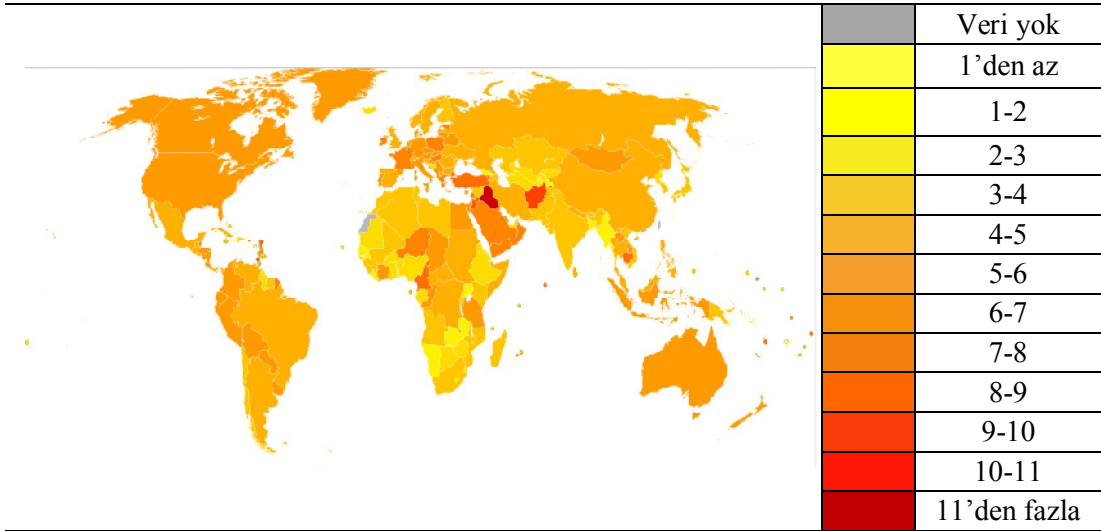
1.2.5. Büyük Granürlü Lenfositik Lösemi

Bu tip lösemi HCL gibi T veya NK hücreleri içerebilir. Nadir olarak görülür, agresif olmayan bir lösemi tipidir (The Leukemia & Lymphoma Society, 2010).

1.2.6. Yetişkin T-Hücreli Lösemi (HTLV)

Lösemi, tıpkı HIV gibi insan T-lenfotropik virüsünden kaynaklanır. HTLV, HIV gibi T hücreleri içinde çoğalır ancak HIV'in aksine onları yok etmez. HTLV enfekte olmuş T hücrelerini ölümsüzleştirir yani onlara normalin ötesinde çoğalma özelliği kazandırır (The Leukemia & Lymphoma Society, 2010).

Tablo 1. 3 Ülkelere göre lösemiden kaynaklanan ölüm oranları (per 100.000).



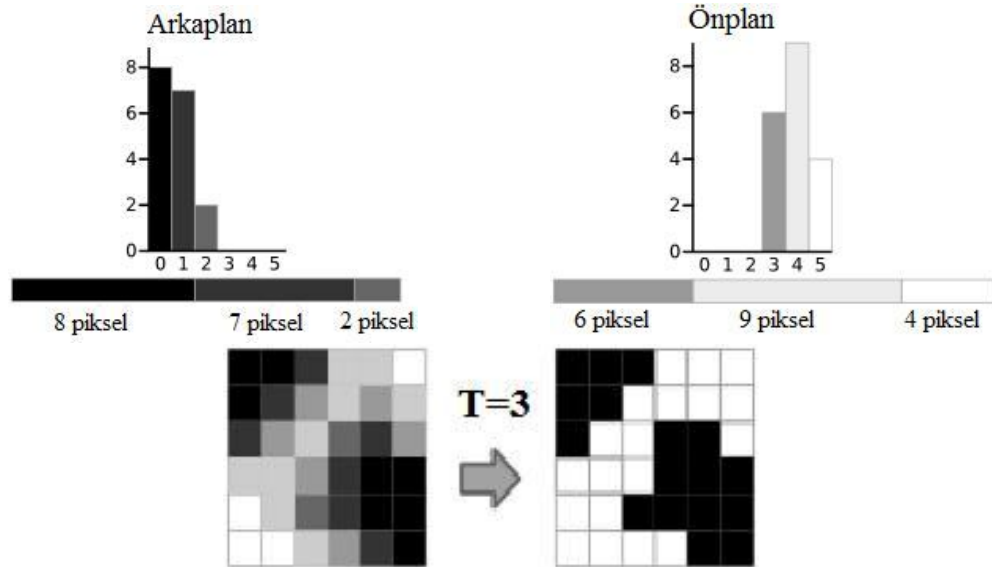
Literatürde lösemi teşhisinde kullanılan yöntemler incelendiğinde; lösemi hücrelerinin görüntüleme teknikleri kullanılarak tespit edilmesinde birbirinden farklı birçok yöntem geliştirilmiş ve kullanılmış olduğu görülür. Günümüzde bu yöntemler hakkında araştırma ve geliştirmeler sürmektedir. Bu çalışmaların amacı daha çok çocuklarda rastlanan ve teşhis edildiği durumda tedaviye açık olan, aksi halde ölümcül sonuçlar doğurabilen ALL tipi hücrelerde morfolojik bir örüntü tanıma algoritması ile erken ve daha doğru tanı sağlamaktır. Çalışmalar sırasında tesadüfen yeni kanser türleri de keşfedilmektedir. Bu yeni kanser türleri için yeni tedavi teknikleri geliştirilmektedir, ki bu da hastalıkla mücadeleyi olumlu yönde etkiler.

2. LÖSEMİLİ HÜCRELERİ NORMAL HÜCRELERDEN AYIRT ETME YÖNTEMLERİ

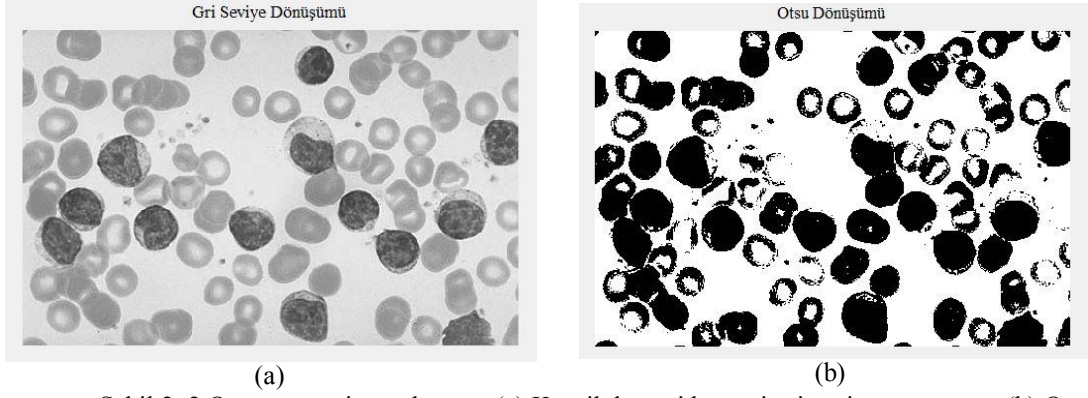
Aşağıda kan hücrelerinde oluşan bir kanser türü olan lösemi hücresinin görüntü işleme teknikleri kullanılarak tespit edilmesi ve tanınmasına yönelik olarak geliştirilen yöntemler ele alınmaktadır. Lösemi hücresi tanımda kullanılan bu metotların matematiksel anlamları ve görüntü üzerindeki etkileri açıklanmaktadır.

2.1. Otsu Yöntemi (Otsu's Method)

Görüntü işlemede otomatik histogram tabanlı eşikleme yapan bir algoritmadır. Gri seviyelerin azaltılması temeline dayanmaktadır. Görüntüde hedeflenen sonuca yönelik uygun eşik değeri seçilmektedir. Tek bir eşik değeri bulmak için ön ve arkaplan varyanslarının hesaplanması gerekir. Çoklu seviye eşikleme de gerçekleştirilir. (Otsu, 1979), (Liu ve Yu, 2009).



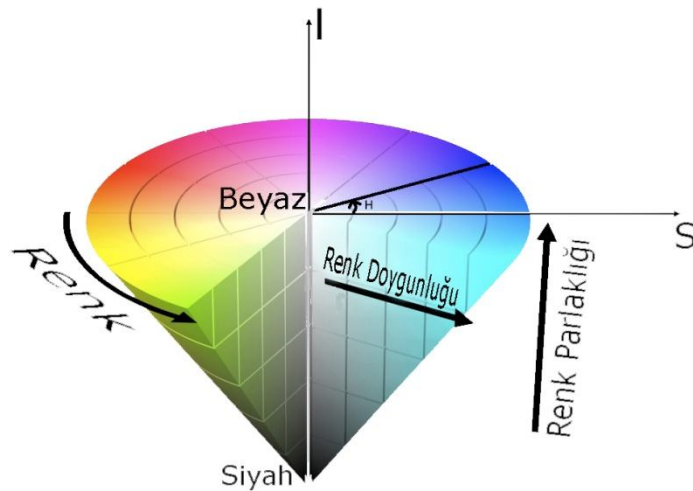
Şekil 2. 1 Otsu yöntemi uygulaması.



Şekil 2. 2 Otsu yöntemi uygulaması. (a) Kronik lösemi hücresi gri seviye görüntüsü. (b) Otsu dönüşüm sonucu.

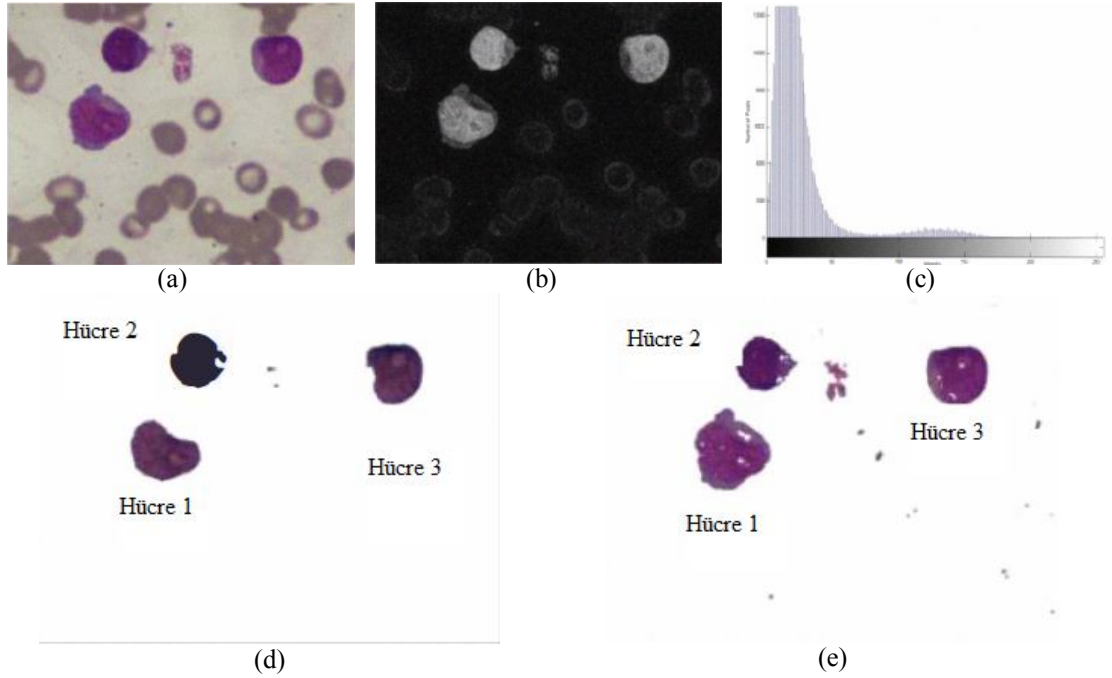
2.2. Renk Tonu, Doygunluk ve Yoğunluk (Hue, Saturation, Intensity) Yöntemi

Üç boyutlu bir renk uzayıdır. Bu özelliğinden dolayı RGB (red, green, blue) veya CYM'den (cyan, yellow, magenta) çok farklıdır. Koni şeklinin merkezi yoğunluğu simgelemektedir. Bu merkez eksenini gri değerleri, tabanı beyaz ve sivri ucu da siyahı temsil etmektedir. Görüntüyü modifiye etmek için RGB'den daha çok tercih edilen bir yöntemdir. Özellikle yoğunluk bilgisinden dolayı veri kaybı daha az ve istenen sonucu daha iyi verebilmektedir. Görüntü işleme alanındaki kullanımı genellikle, RGB görüntüden HSI (hue, saturation, intensity) uzayına geçiş yapılır, görüntü istenilen şekilde işlenir ve ardından tekrar RGB uzayına geçiş yapılır (Gonzales ve diğ., 2009).



Şekil 2. 3 Renk tonu, doyumluk ve yoğunluk (HSI) renk uzayı.

Bu bilgiler ışığında RGB ve HSI ile elde edilen sonuçları değerlendirmek gerekmektedir (Nor Hazlyna ve diğ., 2010).



Şekil 2. 4 RGB ile HSI algoritmalarının sonuçlarının karşılaştırılması. (a) Akut lösemili hücre. (b) S eleman sonucu. (c) S eleman histogramı. (d) RGB renk uzayı sonucu. (e) HSI renk uzayı sonucu.

2.3. Yerel Kontrast Germe, Küresel Kontrast Germe, Kısmi Kontrast Germe, Parlaklık ve Karanlık Kontrast Germe Yöntemleri

Görüntüyü iyileştirmek, geliştirmek ve/veya uygun forma dönüştürmek için görüntü insan ya da cihazlar tarafından analiz edilmektedir. Buna bağlı olarak yöntem, teknik ve süreçler geliştirilir. Burada görüntünün geliştirilmesi, görüntünün öznelilik belirleyici değerlerinin hakimiyetinin artırılması veya görüntünün farklı bölgelerdeki belirsizliklerin azaltılması olarak tanımlanmaktadır.

Görüntü iyileştirilme yöntemlerinden kontrast germe sıklıkla tıbbi görüntüler için tercih edilmektedir. Kontrast germe işlemleri genelde beş grupta incelenmektedir. Görüntü iyileştirme süreçlerinde uygulanan 4 adım bulunmaktadır.

- Akut lösemi kan hücresi görüntüsünün boyutunu 40 kez büyütme,
- .bmp uzantısı ile görüntüyü kaydetme,
- Bu görüntü içinden normal, aydınlık ve karanlık olmak üzere 3 görüntü daha ayrıştırılarak kaydedilmektedir.

Bu işlemler yapıldıktan sonra bahsedilen 5 tip kontrast germe işlemlerinden istenilen uygulanmaktadır (Mokhtar ve diğ., 2009).

2.3.1. Yerel ve Küresel Kontrast Germe

Görüntü üzerinde yerel görüntü elemanlarını ayarlamak, görüntünün birbirinden farklı koyu ve aydınlık bölgelerine ayrı ayrı uygulanarak gerçekleştirilen iyileştirme yöntemidir. LCS (Local Contrast Streching) olarak adlandırılmaktadır. LCS görüntü üzerinde pencereleme ve ayar merkez elemanı kullanılarak uygulanmaktadır. Aşağıdaki gibi formüle edilmektedir.

$$l_p(x, y) = 255[l_o(x, y) - \min]/(I_{max} - I_{min})$$

Burada

$l_p(x, y)$: kontrast germe işlemi sonucunda elde edilen (x,y) çıktı piksel için renk seviyesidir.

$l_o(x, y)$: veri girişi (x,y) için giriş görüntü renk seviyesi,

I_{max} : girdi görüntüsündeki renk seviyesi maksimum değeri,

I_{min} : girdi görüntüsündeki renk seviyesi minimum değeri.

Formülde (x, y) görüntüsünün maksimum ve minimum değerleri içinde çekirdek (kernel) değerini seçmektedir. Görüntü için RGB renk uzayında işlem tanımlanırsa her renk aralığındaki değerler kontrast germe süreci için kullanılmaktadır. Bu her renk aralığı için maksimum ve minimum değerleri verecektir. Buradan tüm renk yelpazesinin maksimum ve minimum değerlerini tek bir seferde belirlemek için küresel kontrast germe işlemi dikkate alınmaktadır.

RGB renk kombinasyonu, RGB renk için maksimum ve minimum olmak üzere yalnızca tek bir değer verecektir. Bu maksimum ve minimum değerler kontrast germe süreci için kullanılmaktadır (Mokhtar ve diğ., 2009).

2.3.2. Kısmi Kontrast Germe

Bir tür otomatik ölçeklendirme yöntemidir. Genellikle kontrast seviyesini ve görüntünün parlaklık düzeyini artırmak için kullanılan bir doğrusal haritalama fonksiyonudur. Bu tekniğin orijinal parlaklık, karanlık ayarlama ve kontrast seviyelerini esas almaktadır. Haritalama fonksiyonu aşağıdaki şekilde ifade edilmektedir.

$$P_k = \frac{(I_{max} - I_{min})}{(f_{max} - f_{min})} (q_k - f_{min}) + I_{min}$$

Burada

P_k : çıkış pikselin renk düzeyi,

q_k : giriş pikselin renk düzeyi,

f_{max} : giriş görüntüsünün maksimum renk düzeyi değeri,

f_{min} : giriş görüntüsünün minimum renk düzeyi değeri,

I_{max} : sırasıyla istenilen maksimum ve minimum renk seviyesi,

I_{min} : istenen minimum renk seviyesidir.

İşlemlere başlamadan önce her renk uzayı için giriş piksellerinin yoğunluğunun nerede yakınsandığı bulunmaktadır. Giriş görüntüleri RGB formatında olduğundan kırmızı, yeşil ve mavi yoğunluk aralıklarını bulmak gerekmektedir. Tüm bu işlemlerin ardından en üst ve en alt renk değerlerini her renk düzeyi için de bulabilmek için ortalama hesabı yapmak gerekmektedir. Bunun için aşağıdaki formül kullanılmaktadır.

$$max_{TH} = (max_{KIRMIZI} + max_{YEŞİL} + max_{MAVİ})/3$$

$$min_{TH} = (max_{KIRMIZI} + max_{YEŞİL} + max_{MAVİ})/3$$

max_{TH} ve min_{TH} renk uzayı için maksimum ve minimum renk seviyesi ortalamasıdır. Her üç renk içinde istenilen renk aralığı kullanılmaktadır. Üç renk düzeyinin amacı, aynı eşik seviyesinin geçerli renklerde yerleşmesini önlemektir. Bu işlemlerin ardından haritalama işlemi gerçekleştirmek gerekmektedir. Kısmi kontrast germe ifadesindeki doğrusal haritalama fonksiyonuna dayanarak piksel dönüşümü aşağıdaki gibi ifade edilmektedir.

$$çıkış(x, y) = \frac{giriş(x, y)}{min_{TH}} Nmin_{TH}, giriş(x, y) > min_{TH} \text{ için}$$

$$çıkış(x, y) = \left[\frac{(Nmax_{TH} - Nmin_{TH})}{max_{TH} - min_{TH}} (giriş(x, y) - f_{min}) \right] + I_{min},$$

$$min_{TH} < giriş(x, y) < max_{TH} \text{ için}$$

$$çıkış(x, y) = \frac{giriş(x, y)}{max_{TH}} Nmax_{TH}, giriş(x, y) < max_{TH} \text{ için}$$

Burada

$giriş(x, y)$: giriş piksel için renk seviyesi,

$çıkış(x, y)$: çıkış piksel için renk seviyesi,

min_{TH} : alt eşik seviyesi,

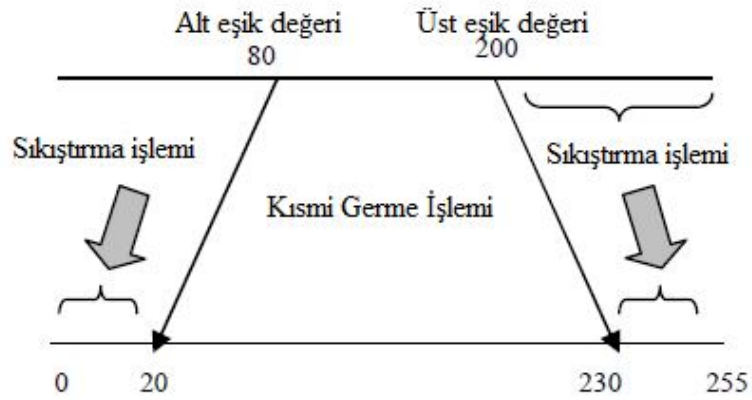
max_{TH} : üst eşik seviyesi,

$Nmin_{TH}$: yeni germe eşiği alt değeri,

$Nmax_{TH}$: yeni germe üst değeri ifade etmektedir.

min_{TH} ve max_{TH} yeni değerlerine yayılır ve diğer pikseller için sıkıştırma yapılır. Bu germe ve sıkıştırma işlemiyle görüntü üzerinde daha geniş ve parlak piksel değerlerine erişilebilir. Sonuçta görüntülerin kontrast ve parlaklık seviyeleri artmaktadır.

Şekil 2.4'te görüldüğü gibi orijinal görüntüde 80-200 piksel değerleri aralığında değişen değerler kısmi germe işlemi sonucunda 20-230 aralığına yayılmaktadır. 80 ve altındaki renk seviyeleri 0-20, 200 ve üzerindeki renk seviyeleri de 230-255 aralığına yerleşmektedir (Mokhtar ve diğ., 2009).



Şekil 2. 5 Kısmi kontrast germe metodu.

2.3.3. Karanlık Kontrast Germe

Bir tür kısmi kontrast germe işlemidir. Germe işleminin koyu renkli piksel değerlerine uygulanarak elde edilmektedir. Karanlık kontrast germe işlemi parlaklık kontrast germe işleminin ters simetriği olan bir süreçtir. Renk seviyelerinin seçilmesi aşağıdaki denkleme dayanmaktadır.

$$çıkış(x, y) = \frac{giriş(x, y) - TH}{255 - TH} Yeni_{TH}, giriş(x, y) < TH$$

$$çıkış(x, y) = \left[\frac{(giriş(x, y) - TH)}{255 - TH} (255 - Yeni_{TH}) \right] + I_{min}, giriş(x, y) > TH$$

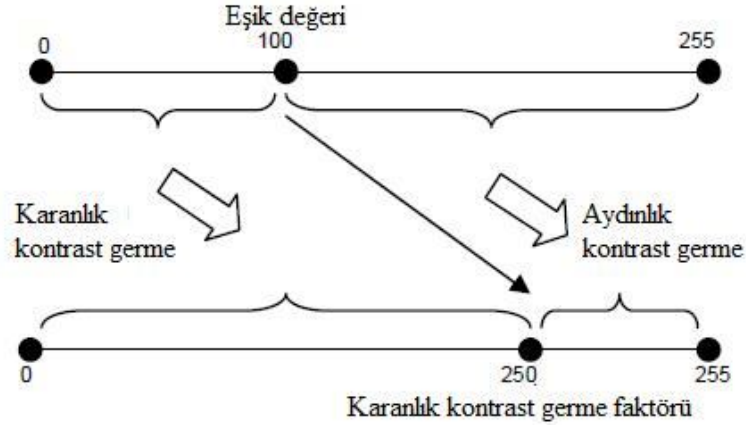
Burada

$giriş(x, y)$: giriş görüntüsü için yerel piksel seviyelerini,

TH : eşik değerlerini,

$Yeni_{TH}$: karanlık değer germe faktörünü ifade etmektedir.

Karanlık kontrast germe işlemi uygulandığında, 0-100 aralığındaki piksel değerleri 0-200 aralığına gerilmekte ve böylece karanlık bölgeler gerilmekte aydınlık bölgeler sıkıştırılmaktadır.



Şekil 2. 6 Karanlık kontrast germe metodu.

2.3.4. Parlaklık Kontrast Germe

Görüntünün parlaklık ve kontrast seviyesini artırmak için otomatik bir ölçeklendirme yöntemidir. Bu yöntem kısmi kontrast germe temeline dayanmakta ve karanlık kontrast germe işleminin ters simetrik bir sürecidir. Aşağıdaki denklem ile ifade edilmektedir.

$$çıkış(x, y) = \frac{giriş(x, y)}{TH} Yeni_{TH}, \quad giriş(x, y) < TH$$

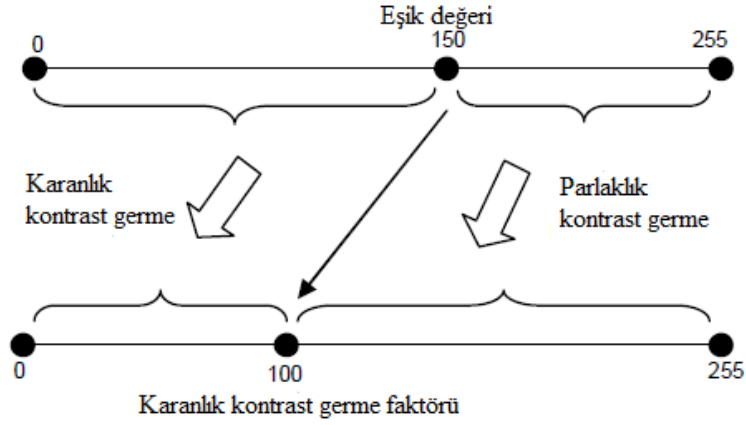
$$çıkış(x, y) = \left[\frac{(giriş(x, y) - TH)}{255 - TH} (255 - Yeni_{TH}) \right] + I_{min}, \quad giriş(x, y) > TH$$

Burada

TH : eşik değerleri,

$Yeni_{TH}$: aydınlık değer germe faktörü olarak ifade edilmektedir.

$Yeni_{TH}$, kırmızı, yeşil ve mavi eşik değerleri için parlaklık kontrast germe işleminin yeni boyutudur. $giriş(x, y)$ giriş görüntüsünün renk seviyesi bir piksel için belirlenen değeri ifade etmektedir. Şekil 2.6'da aydınlık bölgeler için germe, karanlık bölgeler için sıkıştırma tekniği gösterilmektedir.



Şekil 2. 7 Parlaklık kontrast germe metodu.

2.4. Otomatik Morfolojik Analiz

Görüntülerdeki morfolojik farklılıkları ayırt edebilmek için her ne görüntüsü ise bunun çözünürlük ve okunabilirlik açısından nitel ve kaliteli olması koşullarını sağlaması gerekir. Ancak bu şekilde görüntü üzerinde ayıklama yapılabilir (Scotti, 2005).

Patolojik incelemelerde immünofenotipleme ve stogenetik yaygın analiz yöntemleridir. Lökositlerin yüzeyinde hücreye özel antijenik yapıya sahip olan moleküllerin saptanmasıyla neoplastik hücrelerin incelenmesinde immünojenetik yöntemler de kullanılmaya başlanmıştır. İmmunohistokimyasal boyama yöntemleri kullanarak lösemili hücrelerin fenotiplerinin tespit edilmesi işlemi de immünofenotipleme olarak adlandırılır.

2.4.1. FAB Metodu (French-American-British)

FAB, lösemi hücrelerinin tespitinde günümüzde hala geçerliliği olan görüntü tabanlı morfolojik tanımlama yöntemidir. Mikroskoptan alınan kötü huylu hücre görüntülerini inceleyerek sitogenetik kullanılarak kromozom anomalilerini tanımlamak için kullanılır. FAB sınıfları üç bölümde toplanabilir (Scotti, 2005), (Yenerel, 2009).

FAB-1:

- Patlamalar küçük ve homojendir.
- Çekirdeklerinde yuvarlak ve düzenli küçük yarıklar ve göze çarpmayan nükleoller vardır.
- Stoplazma genellikle yetersiz ve vakuollerden yoksundur.

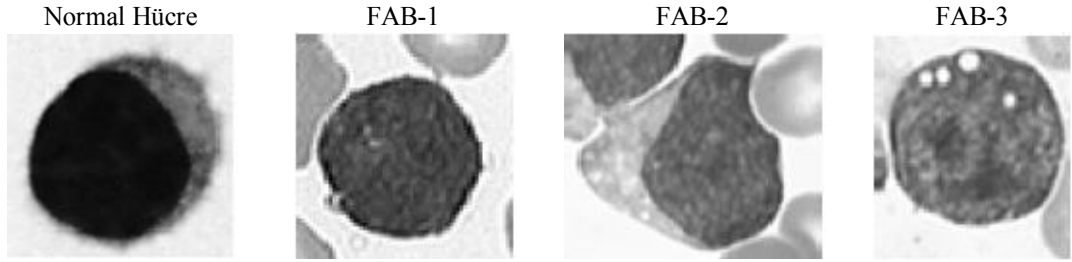
FAB-2:

- Patlamalar büyük ve heterojendir.
- Çekirdekler düzensiz ve çokça yarıklar bulunur.
- Bir veya daha fazla büyük nükleol mevcuttur.
- Stoplazmanın hacmi değişkendir ama sık sık ve bol vakuoller içerebilir.

FAB-3:

- Patlamalar ortaboylu ve homojendir.
- Çekirdeklerin şekli oval ve düzenlidir.
- Bir veya daha fazla belirgin nükleol mevcuttur.
- Stoplazmanın hacmi orta ve belirgin vakuoller içermektedir.

Normal hücre ile özellikler verilen üç sınıf aşağıdaki şekilde gösterilmektedir (Scotti, 2005).



Şekil 2. 8 FAB metodu ile lösemi hücrelerinin sınıflandırılması.

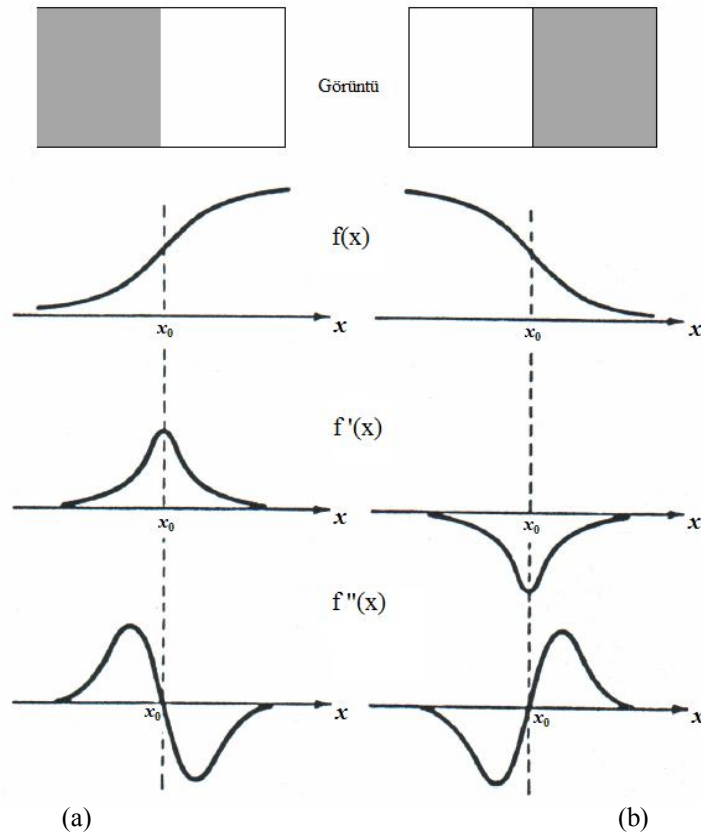
2.5. Gradyan Yöntemi

Bir görüntü içerisindeki kenarları belirlemek için uygulanabilecek en verimli yöntemlerden birisi, ani renk seviye değişimlerini tespit etmektir. Bu amaç için birçok kenar belirleme yönteminin kullandığı temel yaklaşım, bölgesel türev hesabına dayanır. Görüntü bilgisinde alınan türevler yerel maksimum, yerel minimum ve sıfır geçiş değerlerini verdiğinden görüntünün kenarlarının belirlenmesinde kullanılmaktadır.

Görüntünün 1. türevi gradyan yöntemi, 2. türevi de Laplace operatörü yöntemi olarak adlandırılmaktadır. Aşağıda gösterildiği gibi tipik bir-boyutlu (1-B) kenarı temsil eden bir analog $f(x)$ fonksiyon incelenmektedir. Tipik problemlerde, şekildeki x_0 'ın değerini bir kenar noktası olarak incelemek mantıklıdır. x_0 'ı tanımlamanın bir yolu, $f(x)$ fonksiyonunun 1. türevini ($f'(x)$) veya 2. türevini ($f''(x)$) hesaplamaktır. Şekilde x_0 değeri, $f'(x)$ 'in bölgesel ekstremumu

(maksimum veya minimumu) veya $f''(x)$ 'in sıfır geçişi ($f''(x)$ 'in işaretini değiştirdiği nokta) aranarak tanımlanabilir. İkinci türevin yani $f''(x)$ 'in sıfır geçişi, Laplace operatörü tabanlı yöntemler grubunda incelenmektedir.

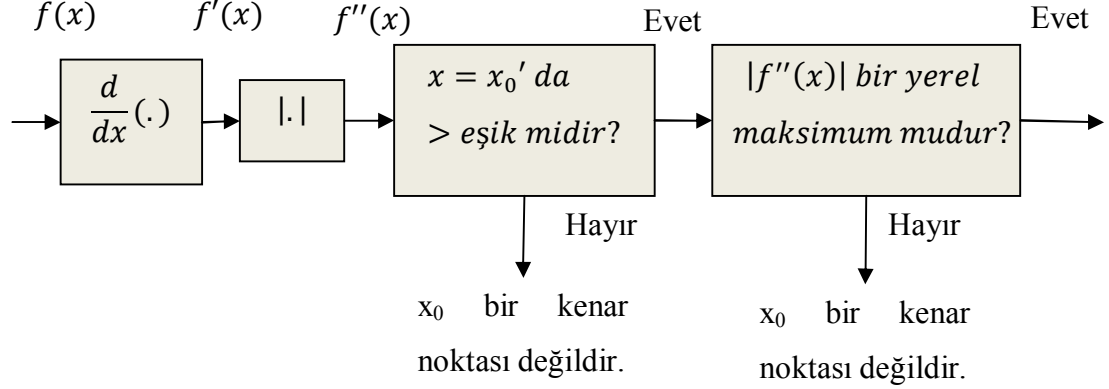
Şekil 2.8'den de görüleceği gibi 1-B çizimlerde kenar noktaları, ani değişimden ziyade biraz daha yumuşak bir geçiş biçiminde modellenmektedir. Bu tip modelleme, gerçekliğe yakın bir gösterim için tercih edilmektedir. Şöyle ki, örneklemenin bir sonucu olarak sayısal görüntülerdeki kenarlar genel olarak hafif bulanıklaşmaktadır. Şekilde gösterilen muhtemel kenar noktası x_0 'ı tanımlamaya ek olarak, $f'(x)$ aynı zamanda kenarın yönünü ve büyüklüğünü kestirmede kullanılabilir. Eğer $|f'(x)|$ çok büyük ise, $f(x)$ çok hızlı değişir ve bu durum parlaklıkta hızlı bir değişime karşı düşer. Eğer $f'(x)$ pozitif ise, $f(x)$ artan bir fonksiyondur. Bu fikirlere dayalı olarak kenar belirleme için bir yaklaşım aşağıdaki sistemi kullanmaktadır.



Şekil 2.9 Türev operatörleri ile kenar belirleme: (a) Koyu arka plan üzerindeki beyaz bant görüntüsü; (b) Açık arka plan üzerindeki siyah bant görüntüsü.

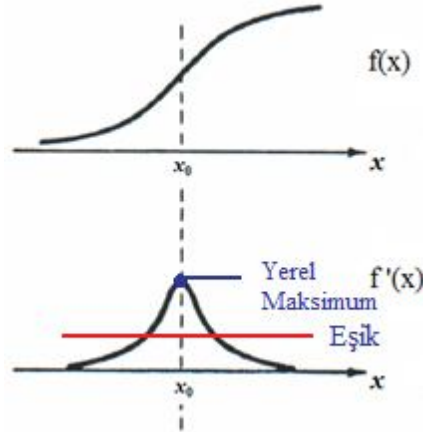
Şekil 2.9'da verilen sistemde ilk olarak $f(x)$ 'den $|f'(x)|$ hesaplanmaktadır. Burada $|f'(x)|$, belli bir eşik değerinden büyükse bu görüntü pikseli, bir kenar

adayı olacaktır. İlgilenilen kenar noktasında eğer bu şart birden fazla x değerleri için sağlanırsa bu durumda bir kenar, noktadan ziyade çizgi olarak görünmekte ve kalın kenarların oluşmasına neden olmaktadır.



Şekil 2. 10 1-B kenar belirleme sisteminin blok diyagramı.

Bu sorunun üstesinden gelmek için en iyi yaklaşım, sadece $|f'(x)|$ değerlerinde yerel maksimuma sahip olan noktaları bulmak olacaktır ve bu noktalar kenar noktaları olarak belirlenmektedir. Bu durum aşağıdaki şekil 2.10'da gösterilmektedir.



Şekil 2. 11 1-B kenar belirleme işlemi.

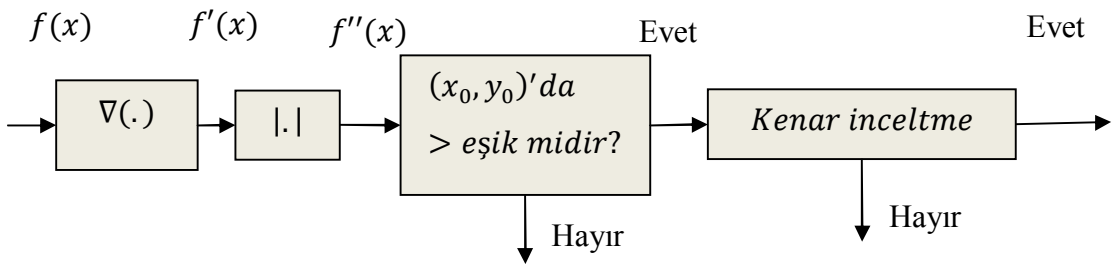
Kenar belirleme işleminde kullanılan eşik değerinin seçimi, uygulamaya bağlı olarak değişiklik arz eder ve kenar görüntüsünün performansı, seçilen eşik değeri ile yakından ilgilidir. Eşik değeri, $|f'(x)|$ değerlerinin büyük olanlarını küçüklerinden ayıracak şekilde seçilmelidir. Uygun eşik değeri, $|f'(x)|$ değerlerinin histogramına bakılarak belirlenmektedir. Ayrıca uyarlamalı olarak da eşik değerini belirlemek

mümkündür. Genellikle eşik değeri, en büyük gradyan %5 ile %10'u olarak seçilmektedir.

2-B bir $f(x, y)$ fonksiyonu için $f'(x)$ 'in genellemesi, $f(x, y)$ 'nin gradyanı olarak tanımlanır ve aşağıdaki denklem ile ifade edilmektedir.

$$\nabla f(x, y) = \frac{\partial f(x, y)}{\partial x} \vec{i}_x + \frac{\partial f(x, y)}{\partial y} \vec{i}_y$$

Burada \vec{i}_x , x-yönündeki \vec{i}_y ise y-yönündeki birim vektörlere karşı düşmektedir. $\nabla f(x, y)$ 'ye dayalı olarak, 1 boyutlu kenar belirleme sistemi şekil 2.11.'deki gibi geliştirilmektedir.



(x_0, y_0) bir kenar noktası değildir.

Şekil 2. 12 2-B kenar belirleme sisteminin blok diyagramı.

Şekil 2.11'deki sisteme göre, ilk olarak $f(x, y)$ 'nin gradyanı olan $\nabla f(x, y)$ fonksiyonunun genliği hesaplanmakta ve bu genlik değeri daha sonra aday kenar noktalarını tanımlamak için bir eşik değeri ile karşılaştırılmaktadır. Eğer $\nabla f(x, y)$ fonksiyonunun genliğinin belirli bir eşik değerinden büyük olduğu bütün (x, y) noktaları kenar noktası olarak tanımlandığından, oluşan kenar görüntüsü ince çizgilerden ziyade kalın şeritler halinde görünmektedir. Kalın şeritlerden ince çizgiler elde etmek için, kenar görüntü inceltme (thinning) işlemine tabi tutulmaktadır. Basit bir kenar inceltme algoritmasında, en azından bir yönde $|\nabla f(x, y)|$ büyüklüğünün yerel maksimum olup olmadığını kontrol edilerek kenar noktaları seçilmektedir. Birçok durumda, sadece yatay ve düşey yöndeki yerel maksimumu kontrol etmek yeterli olmaktadır. Potansiyel bir kenar noktasında, belirlenen yönlerden herhangi biri boyunca eğer $|\nabla f(x, y)|$ yerel maksimum ise bu potansiyel kenar noktası bir kenar noktası olarak değerlendirilmektedir. Diğer taraftan, bu basit kenar inceltme algoritması güçlü kenar çizgileri etrafında çok

sayıda hatalı kenar çizgileri oluşturmaktadır. Bu hatalı kenar çizgilerinin çoğunu yok etmek için basit bir yöntem, aşağıdaki ek sınırlamalarıyla sağlamaktadır.

- Eğer yatay yönde (x_0, y_0) noktasında $|\nabla f(x, y)|$ yerel maksimuma sahip ise, yani

$$\left| \frac{\partial f(x, y)}{\partial x} \right|_{(x, y) - (x_0, y_0)} > k \left| \frac{\partial f(x, y)}{\partial y} \right|_{(x, y) - (x_0, y_0)}$$

olduğunda (x_0, y_0) kenar noktası olmaktadır.

- Eğer düşey yönde (x_0, y_0) noktasında $|\nabla f(x, y)|$ yerel maksimuma sahip ise, yani

$$\left| \frac{\partial f(x, y)}{\partial y} \right|_{(x, y) - (x_0, y_0)} > k \left| \frac{\partial f(x, y)}{\partial x} \right|_{(x, y) - (x_0, y_0)}$$

olduğunda (x_0, y_0) bir kenar noktası olmaktadır.

Yukarıdaki iki eşitlikte k tipik olarak 2 civarında seçilmektedir. Şartlar sağlandığında yatay yön boyunca parlaklık değişim oranı düşey yönden önemli derecede büyük olmakta ve düşey yön boyunca parlaklık değişim oranı yatay yönden daha büyük olmaktadır. Gradyan tabanlı kenar belirleme sistemleri iki değişik şekilde uygulanabilmektedir. Bunlar, yönlü ve yönsüz kenar sezicileri olarak adlandırılmaktadır. $|\nabla f(x, y)|$ fonksiyonunu kullanan kenar belirleme sistemlerine yönsüz kenar sezicisi adı verilir ki bu sistem herhangi bir doğrultu için ayarlı değildir ve her yön için eşit ağırlıklı sonuçlar üretmektedir. Eğer kenar belirleme sistemi herhangi bir yön için ayarlanmışsa bu sistemlere yönlü kenar sezicisi adı verilmektedir. Örneğin, 2 boyutlu kenar belirleme sisteminde $|\nabla f(x, y)|$ fonksiyonu yerine sadece $|\partial f(x, y) / \partial x|$ fonksiyonu kullanılırsa, bu durumda sistem sadece düşey yöndeki kenarları tanımakta ve yatay yöndeki kenarlara ise cevap vermemektedir. 2-B sayısal fonksiyonlarda türev işlemleri, fark denklemleri ile ifade edilebilmektedir. Bu anlamda $\partial f(x, y) / \partial x$ ve $\partial f(x, y) / \partial y$ ile ifade edilen kısmi türevler fark denklemi ifadeleri ile tanımlanabilmektedir. Örneğin, $\partial f(x, y) / \partial x$ kısmi türevi aşağıdaki fark denklemleri ile yer değiştirmektedir:

$$\frac{\partial f(x, y)}{\partial x} \cong \frac{1}{T} [f(n_1, n_2) - f(n_1 - 1, n_2)]$$

$$\frac{\partial f(x, y)}{\partial x} \cong \frac{1}{T} [f(n_1 + 1, n_2) - f(n_1, n_2)]$$

$$\frac{\partial f(x, y)}{\partial x} \cong \frac{1}{T} [f(n_1 + 1, n_2) - f(n_1 - 1, n_2)]$$

Kenar belirleme sisteminde, hesaplanan türev değerleri bir eşik değeri ile karşılaştırıldığından ve bu eşik değeri ayarlanabilir olduğundan eşitlikteki $1/T$ ve $1/2T$ ölçeklendirme katsayıları ihmal edilebilmektedir. Tanımlanan ifadelerin daha iyi yaklaşım için aşağıdaki gösterimler kullanılabilir:

$$\begin{aligned}\frac{\partial f(x, y)}{\partial x} &\cong [f(n_1 + 1, n_2 + 1) - f(n_1 - 1, n_2 + 1)] \\ &\quad + [f(n_1 + 1, n_2) - f(n_1 - 1, n_2)] \\ &\quad + [f(n_1 + 1, n_2 - 1) - f(n_1 - 1, n_2 - 1)]\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x, y)}{\partial x} &\cong [f(n_1 + 1, n_2 + 1) - f(n_1 - 1, n_2 + 1)] \\ &\quad + 2[f(n_1 + 1, n_2) - f(n_1 - 1, n_2)] \\ &\quad + [f(n_1 + 1, n_2 - 1) - f(n_1 - 1, n_2 - 1)]\end{aligned}$$

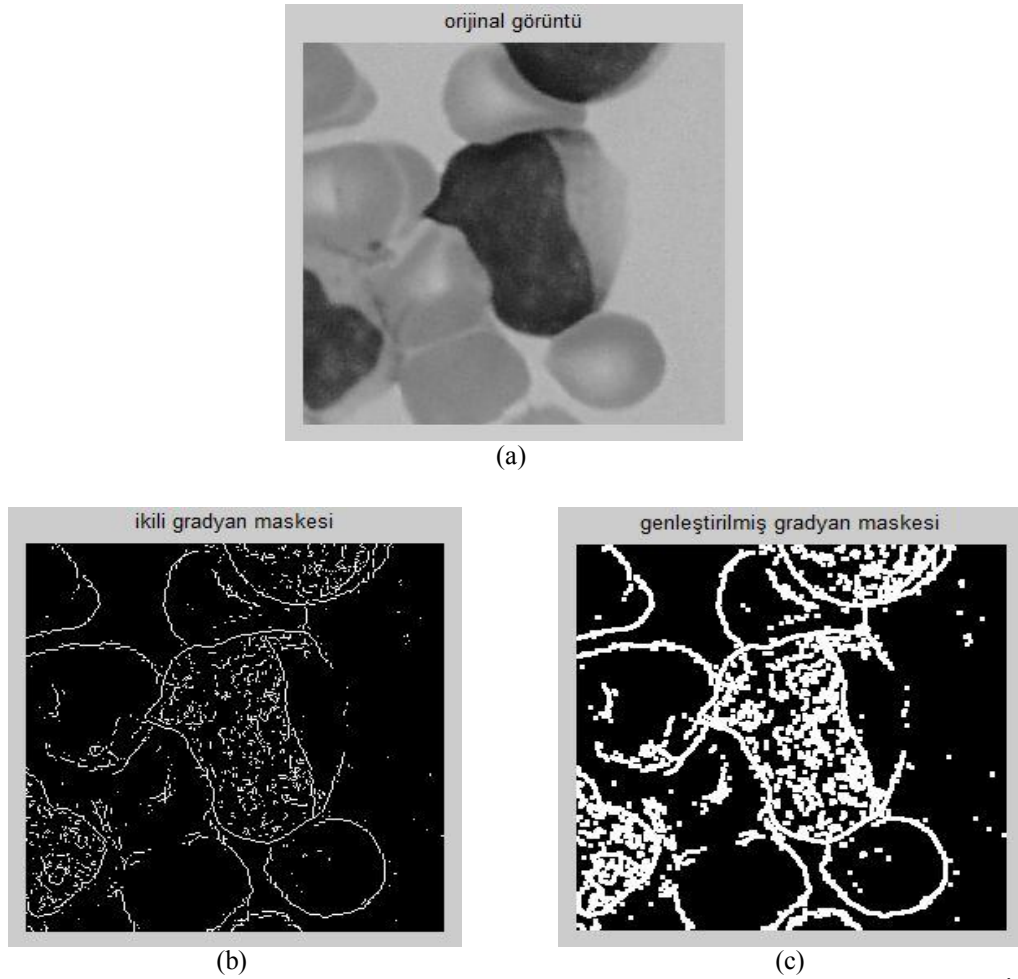
Yukarıda ifade edilen nedenden dolayı, ölçekleme katsayıları bu eşitliklerde ihmal edilmiştir. Eşitliklerdeki fark alma işlemleri, $f(n_1, n_2)$ ile dürtü yanıtı $h(n_1, n_2)$ olan bir süzgecin evrişimi olarak yorumlanabilmektedir. Bu anlamda eşitlikleri sağlayan ve yönlü kenar sezicilerine karşı düşen süzgeç dürtü yanıtları Tablo 2.1'de (a) ve (b)'de gösterilmektedir. Ayrıca aynı mantıkla yatay yöndeki ve $\pm 45^\circ$ doğrultusundaki kenarları bulmak için kullanılacak süzgeç dürtü yanıtları Tablo 2.1'de (c)-(f)'de verilmektedir.

Tablo 2. 1 Yönlü kenar belirleme için kullanılabilen süzgeçlerin dürtü yanıtları: (a) ve (b) düşey doğrultudaki kenarlar için; (c) ve (d) yatay doğrultudaki kenarlar için; (e) ve (f) $\pm 45^\circ$ doğrultusundaki kenarlar için.

$(n_1 - 1, n_2 + 1)$			$(n_1, n_2 + 1)$			$(n_1 + 1, n_2 + 1)$											
$(n_1 - 1, n_2)$			(n_1, n_2)			$(n_1 + 1, n_2)$											
$(n_1 - 1, n_2 - 1)$			$(n_1, n_2 - 1)$			$(n_1 + 1, n_2 - 1)$											
-1	0	1	-1	0	1	1	1	1	1	2	1	0	1	1	1	1	0
-1	0	1	-2	0	2	0	0	0	0	0	0	-1	0	1	1	0	-1
-1	0	1	-1	0	1	-1	-1	-1	-1	-2	-1	-1	-1	0	0	-1	-1
(a)			(b)			(c)			(d)			(e)			(f)		

Tablo 2.1'deki yönlü kenar maskelerini kullanarak Şekil 2.12 (a)'daki orijinal görüntüye ilişkin elde edilmiş kenar görüntüleri gösterilmektedir. Görüldüğü gibi maskeler sadece ilgili doğrultudaki kenarları tanımakta, diğer doğrultulardaki kenarları ise tanıyamamaktadırlar. Her bir maske ile elde edilen gradyan görüntüsünden kenar görüntüsüne ulaşmak için kullanılan eşik değerleri, her bir gradyan görüntüsünün en büyük genlik değerinin % 5'i alınarak belirlenmektedir. Bu

eşik değerinden büyük gri seviyeler 255, küçük olanlar ise 0 alınarak kenar görüntülerine ulaşmaktadır.



Şekil 2. 13 Kenar belirleme yöntemi. (a) Orijinal lösemi hücresi gri-ton görüntüsü. (b) İkili gradyan maskesi. (c) Genişletilmiş gradyan maskesi.

Yönsüz kenar sezicileri 3 grupta toplanmaktadır. Bunlar, Sobel, Prewitt ve Robert olarak adlandırılmaktadır. Bu tip yönsüz kenar sezici süzgeçlerin maskeleri Tablo 2.2'deki gibidir (Gonzales ve diğ., 2009).

Tablo 2. 2. Gradyan maskeleri.

-1	0	1	-1	2	1	-1	0	1	1	1	1	0	0	1	0	1	0
-2	0	2	0	0	0	-1	0	1	0	0	0	0	-1	0	0	0	-1
-1	0	1	-1	-2	1	-1	0	1	-1	-1	-1	0	0	0	0	0	1
Sobel			Prewitt			Robert											

Ayrıca gradyan yöntemi çeşitli tıbbi görüntülerde ön işleme veya ara işleme olarak da tercih edilmektedir. Bu alanda watershed algoritması içinde veya destek

vektör makineleri ile birlikte kullanımı söz konusu olmaktadır (Morkiewwicz ve diğ., 2005), (Gavlasova ve diğ., 2008).

2.6. Yüksek Verimli Eleme Algoritması

Görüntü işleme teknikleri kullanılarak biyolojik hücre görüntülerinin de analizi yapılmaktadır. Bu algoritmalar dört farklı sınıfla ifade edilmektedir.

- Karakteristik eşikleme özelliği veya kümeleme özelliğine dayalı,
- Kenar algılama özelliklerine dayalı,
- Bölge çıkarma türlerine dayalı,
- Bazı algoritmaların kombinasyonlarından oluşturulan yeni algoritmalar olarak tanımlanmaktadır.

Genel olarak sıralanan bu algoritmalarda arkaplanın sabit istenmesi sadece birkaç izole hücre görüntülerinde başarımlı oranını kabul edilebilir seviyede sağlamaktadır. Diğer yandan yarı otomatik sistemler kullanılarak kullanıcı tarafından belirlenen mekansal bölgelerde piksel saymak için geliştirilmektedirler.

Sjostrom ve diğ., önerdikleri yapay sinir ağlarına (YSA) dayalı bir hücre sayma sistemidir. Reddedilen atıkları ve diğer nesnelere hücre olmadıklarını sınıflandırması için ilgili giriş görüntüsünün taramada akıllı bir filtre kullanılmalıdır. Ancak doğruluk ve yüksek hesaplama yükü sınırlı olmaktadır.

Loukas ve diğ., büyük ölçekli histolojik görüntülerin hücre çekirdeğinin tespit edilmesinde çok aşamalı bir algoritma önermektedirler. Önerilen algoritma lösemi hücrelerinin temeli, sayısı, boyutu ve konumunu tanımlamada kullanılması için tasarlanmıştır.

Şekil 2.13.(a)'da FMC44 tipi bir lösemi hücresinin tipik mikroskopik görüntüsü verilmektedir. Önerilen algoritma kenar belirleme için 3×3'lük Sobel filtre kullanmaktadır. Şekil 2.13.(b)'de gösterilen sonuç üretilmektedir. Daha sonra görüntü yatay olarak bölümlenmiştir ve histogram piksel yoğunlukları kullanıcı tanımlı eşik değeri aşağıdaki eşitlik ile ifade edilmektedir.

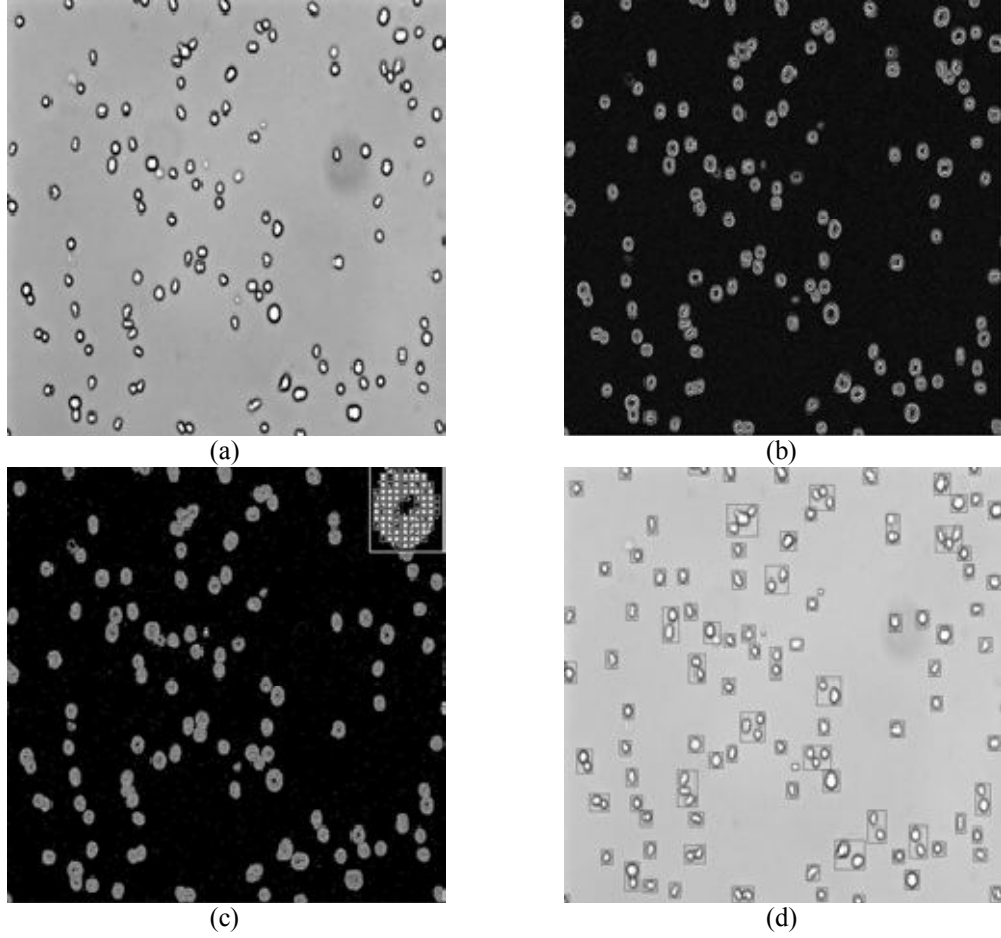
$$\sum_{X_{N_1}}^{X_{N_2}} I \geq \text{Eşik değeri}$$

Burada

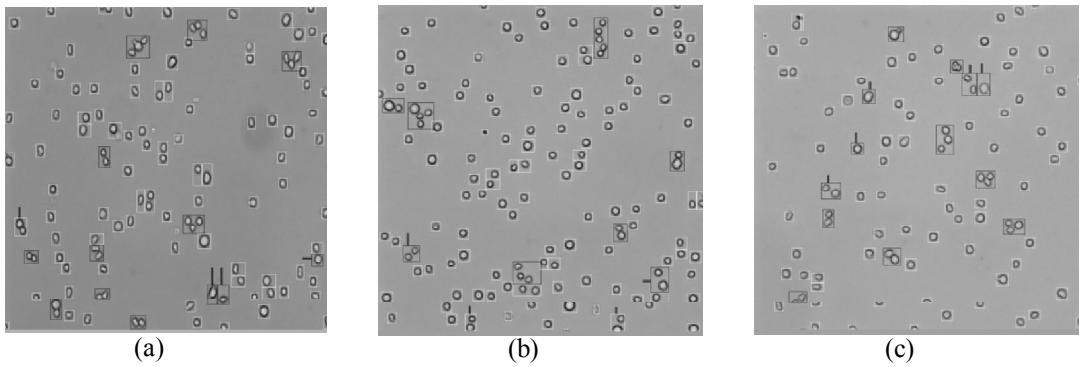
I : histogram değerindeki piksel yoğunlukları,

$[X_{N_1}, X_{N_2}]$: bir bölümün başlangıç ve bitiş noktaları olarak ifade edilmektedir.

Şekil 2.13.(c)'de yedi parçaya bölütlenmiş olan görüntü gösterilmektedir. Bölütleme işlemi hücrelerin arkaplandan ayrışmasını sağlamaktadır. Şekil 2.13.(d)'de hücrelerin sınırlarıyla kutular içinde gösterilmektedir. Tüm hücreler belirlenmekte ancak bazen bir kutu içinde birden çok hücre kümelenmektedir.



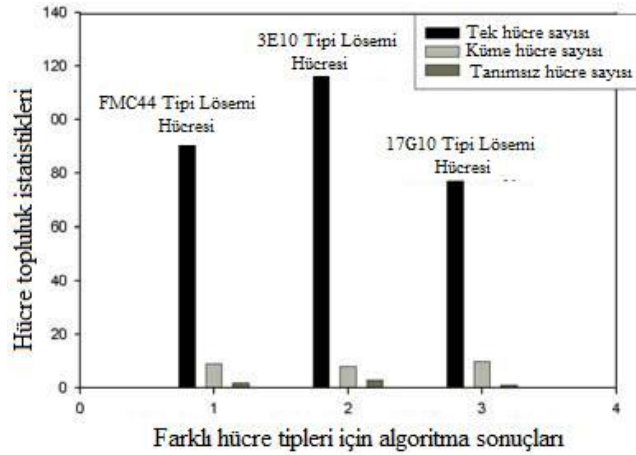
Şekil 2. 14 Yüksek verimli eleme algoritması uygulaması. (a) FMC44 tipi lösemi hücresi görüntüsü. (b) Kenar belirleme algoritması. (c) Özyineli segmentasyon. (d) Sınır çıkarımı.



Şekil 2. 15 Farklı lösemi hücre tiplerinde elde edilen algoritma sonuçları. (a) FMC44 tipi lösemi hücresi algoritma çıkışı. (b) 3E10 tipi lösemi hücresi algoritma çıkışı. (c) 17G10 tipi lösemi hücresi algoritma çıkışı.

Şekil 2.15'te FMC44, 3E10 ve 17G10 tipi lösemi hücreleri için algoritma sonucunda hücrelerin tanımlanmasına dair görüntüler gösterilmektedir. Şekilde her lösemi hücresi için 40 kez büyütülmüş ve 1000×1000 çözünürlüğüne sahip parlak bir görüntü elde edilmektedir.

Önerilen algoritma hücre görüntüsünden otomatik olarak hücre sayısını, boyutunu ve görece hücre konumunu belirlemektedir. Bunlar topluluk içinde istatistiksel eleme işlemleri sonucunda elde edilmektedir. Buna ek olarak kümeler içinden tek hücreleri sınıflandırmaktadır. Bu hücreleri beyaz olarak belirlenmiş sınır kutuları ile göstermektedir. Yanlış sınıflandırılan hücreler ise siyah olarak belirlenmiş sınır kutuları ile gösterilmektedir. Tek hücrelerin ve hücre kümelerinin çapları ve özlerinin önerilen algoritma sonuçları grafikte verilmektedir.



Şekil 2. 16 Farklı lösemi hücreleri için topluluk istatistikleri.

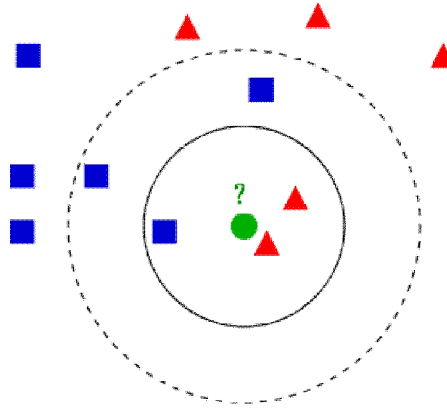
Ek olarak sınıflandırma adımları ile verilen görüntü içindeki hücreleri yüksek doğrulukla saymakta ve belirlemektedir. Önerilen algoritma yüksek hassasiyet ve hız ile hücre görüntülerini analiz etmektedir (Prasad ve diğ., 2006).

2.7. K-En Yakın Komşu Yöntemleri

K-en yakın komşu algoritması, uzayda en yakın örneklere dayalı nesnelere sınıflandırmada kullanılan bir görüntü tanıma yöntemidir. K-en yakın komşu tabanlı öğrenmenin temel işlevi yerel olarak yaklaşık nerede olduğu sınıflanana kadar vektör testi oluşturur. Yakınlık minimum mesafe olarak alınır (Asadi ve diğ., 2006). K-en yakın komşu algoritması ve tüm basit makine öğrenme algoritmaları arasında bir nesnenin nesne sınıfının en yakın komşuları olan k arasında ortak atanmış olmak,

komşularının oyu ile sınıflandırılır (k pozitif bir tam sayı, genellikle küçüktür). Eğer $k=1$ ise, nesne basit onun en yakın komşu sınıfına atanır. Aynı yöntem basitçe nesnenin özellik değerini atayarak komşuların en yakın olan k değerlerini ortalama regresyon değeri olarak kullanabilir. Bu komşu değerlerinin yakın olanları uzak olanlara göre komşuların ağırlığına daha fazla katkıda bulunabilir. Her komşu için ortak ağırlık ($1/d$), komşu mesafesi de d olarak genellenmektedir (Bremmer ve diğ., 2005).

Eğitim örnekleri çok boyutlu bir örnek uzayda bir sınıf etiketiyle tanımlanır. K-en yakın komşu yöntemi, düzgün bir çekirdek ile çekirdek yoğunluğu ‘balon’ kestirimi değişken bant genişliği özel durumudur (Terrell ve Scott, 1992). K-en yakın komşu yöntemi şekil 2.16’da gösterilen örnekle açıklanabilir: Bu yeşil yuvarlak test örneği olarak tanımlanmaktadır. Birinci sınıf olarak mavi kareler, ikinci sınıf olarak kırmızı üçgenler tanımlanmalıdır. Eğer $k=3$ ise, yeşil yuvarlak ikinci sınıf olarak belirlenir. Çünkü sadece iki üçgen ve bir kare seçilen dairenin içinde bulunmaktadır. Eğer $k=5$ ise, yeşil yuvarlak birinci sınıf olarak belirlenir. Çünkü ikinci seçilen daire içinde 3 kare ve iki üçgen bulunmaktadır.



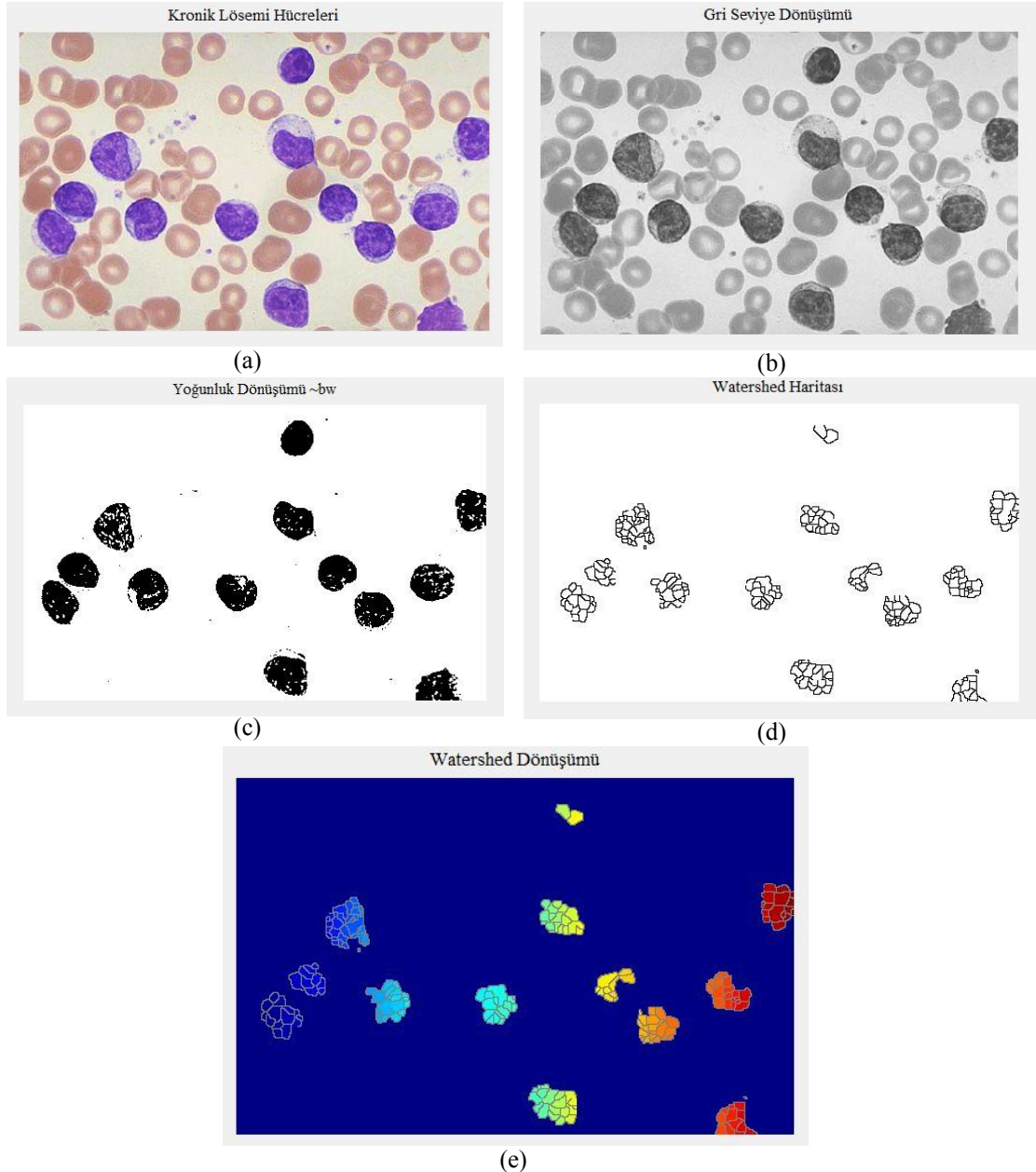
Şekil 2. 16 K-en yakın komşu yöntemi diyagramı.

2.8. Sınır Noktası (Watershed) Algoritması

Görüntü segmentasyonunda morfolojik temelli bir yöntemdir. Bir görüntünün gradyan genliği watershed dönüşümü için topolojik bir yüzey olarak kabul edilir (Burges, 1998), (Morkiewwicz ve diğ., 2005). Watershed haritası farklı yollarla bulunabilir ve görüntü işlemede farklı alanlarda kullanılmaktadır. Watershed çizgileri düğümler, kenarlar veya hem düğüm hem kenarlar üzerinde tanımlı olabilirler.

Watershed dönüşümü çoğunlukla görüntünün bölümlenmesi sayesinde görüntü üzerindeki geçişlerin tahmin edilmesine dayalıdır. Bu tip bir segmentasyonun amacı görüntü üzerindeki geçiş uzunluğu ve yüksekliğinin bilgi olarak yorumlanmasıdır (Morkiewwicz ve diğ., 2005).

Lösemi hücreleri üzerinde watershed algoritmasının temel kullanımını incelendiğinde elde edilen sonuçlar Şekil 2.17'deki gibi gösterilmektedir.



Şekil 2. 17 Watershed algoritmasının lösemi hücrelerinin tespit edilmesi. (a) Orijinal kronik lösemi hücre görüntüsü. (b) Gri seviyeli görüntü. (c) İkili yoğunluk dönüşümü. (d) Watershed haritası. (e) Watershed dönüşümü sonucu.

2.9. Çekirdek Yöntemi (Kernel Methods)

Çekirdek yöntemleri bir örüntü tanıma algoritmaları sınıfıdır. Bu yaklaşım 1900'li yılların ortalarından beri geliştirilmektedir. Dışbükey optimizasyonu veya özproblemler çoğunlukla çekirdek algoritmalar ile istatistiksel olarak verimli sonuçlar vermektedir. Bu yüzden tercih edilen yöntemlerin başında gelmektedir. Örüntü analizinin genel görevi kümeler, sıralamalar, temel bileşenler analizi (PCA-principle components analysis), korelasyon, sınıflandırma gibi genel türdeki veriler (dizileri, metinleri, set noktaları, vektörler ve görüntüler) içinden çalışma sınıfındaki ilişkileri bulmakta ve analiz etmektedir (Welling, 2009).

Çekirdek yöntemi yaklaşımı bir Öklid uzayında çeşitli veri ilişkilerini bulmak için kullanılabilir. Bunlar veri dönüştürme problemi, yüksek boyutlu özellik uzayında verileri haritalama problemleri olabilir.

Bu yöntemle tüm çekirdek çiftleri arasında iç hesaplamalar ile öznelikler hesaplanmaktadır. Çekirdek fonksiyonlarının veri dizisi, grafik, metin, görüntü gibi vektörler için geliştirilmiştir.

Algoritmaları çekirdekler ile çalışma yeteneğine sahip yöntemler destek vektör makineleri (DVM), Gauss süreçleri, Fisher'in doğrusal diskriminant analizi (LDA), temel bileşenler analizi (PCA), kanonik korelasyon analizi, çekirdek sağrı regresyon (KRR), spektral kümeleme, doğrusal adaptif filtreler olarak sıralanabilir (Sewell, 2009), (Theodoridis ve Koutroumbas, 2009: 214).

2.9.1. Çekirdek Sağrı Regresyon (Kernel Ridge Regression) Yöntemi

Andrey Tychonoff tarafından geliştirilen Tikhonov Regülasyonu iyi tanımlanmamış (ill-posed) problemlerin çözümünde en sık kullanılan yöntemdir. İstatistikte bu yöntem sağrı regresyonu olarak bilinir. Sınıflandırma için verinin en küçük karelere (least squares) ayrıştırılarak incelendiği yöntem denir. Eğer bu doğrusal olarak gerçekleştiriliyorsa, buna Fisher'in diskriminant sınıflandırması (LDA) da denir (Sewell, 2009).

Çekirdek sağrı regresyonda farklı nesnelerin karar sınırlarına yakın bölgelerine vurgu yapılmaz. Çözüm ise verilerden elde edilen eğitim örneklerine dayanır. Bununla birlikte, bu yöntem sadece birden fazla ve örneklere sahip olan veri kümelerine uygundur. Doğrusal yöntem için çekirdek hilesi (kernel trick) kullanımı özelliklerin sayısı fazla ve örneklerin sayısı az olmalıdır. Aksi durumda çekirdek hilesi kullanılmamalıdır. Böylece veri kümesinde az özellik ile çok sayıda örneğin

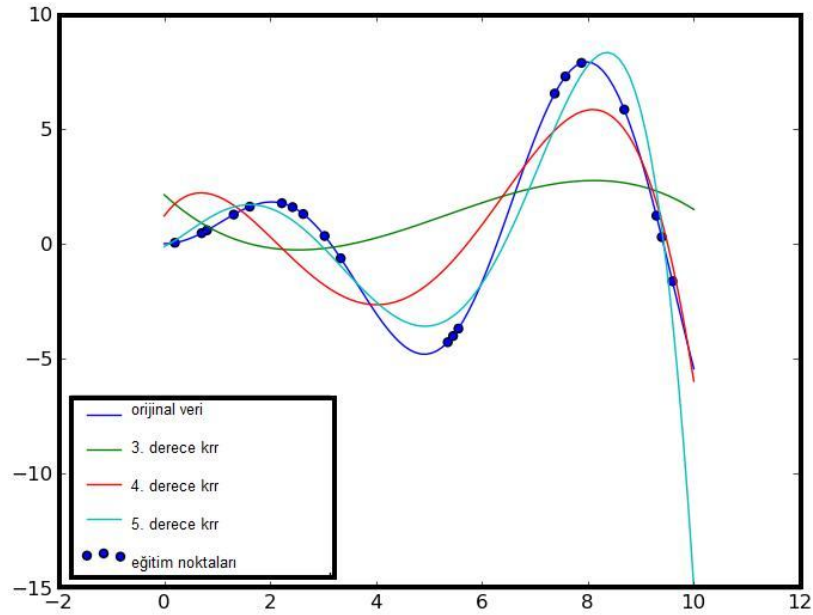
kullanılması için doğrusal sınıflandırmalı sağrı regresyon kullanımı olanağı sağlar. En küçük kareler yöntemi ω 'yı bulmak için en popüler yaklaşımdır. Kare farkları toplamını minimize eden ω için:

$$L_{SQ}(\omega) = \sum_{i=1}^T (\omega'x_i - y_i)^2$$

Yukarıdaki ifadeyi minimize etmek, n adet v_1, v_2, \dots, v_n vektör tarafından üretilen \mathbb{R}^T 'nin alt uzayında $Y = (y_1, y_2, \dots, y_T)$ vektörünün izdüşümünü almaya eşdeğerdir.

Alt uzaydaki izdüşüm her zaman tanımlı ve tektir. Öte yandan v_1, v_2, \dots, v_n lineer bağımsızsa, eğer örnek küçükse ve $T < n$ ise bu her zaman geçerlidir. Bu tek izdüşüm bunların lineer birleşimi olarak çoklu temsile sahiptir. En küçük kareler metodu tek bir regresörü zorunlu olarak tanımlamaz. Daha genel bir problem düşünülürse, $a > 0$ için:

$$L_{RR}(\omega) = a\|\omega\|^2 + \sum_{i=1}^T (\omega'x_i - y_i)^2$$



Şekil 2.18 Çekirdek sağrı regresyonu.

Bu A vektörünü minimize eden ω sağrı regresyonu çözümdür. En küçük kareler yaklaşımının $a > 0$ olduğu özel durum da sağrı regresyonu verir. Parametre olarak $a > 0$ kullanılmasının 3 ana nedeni bulunmaktadır:

- **Birincil Form:** Sağrı regresyonunun $a > 0$ için çözümü incelenirse, $T \times n$ matrisinin satırları x_i sütunları v_i olmak üzere $X = (x_1, x_2, \dots, x_n)'$ şeklinde ifade edilen matris kullanılarak L_{RR} 'nin ω 'ya göre türe alınıp sonuç sıfıra eşitlenirse

$$L_{RR}(\omega) = a\|\omega\|^2 + \|X\omega - Y\|^2$$

$$2a\omega - 2X'Y + X'X\omega = 0$$

$$\omega = (aI + X'X)^{-1}X'Y$$

I : birim matris olarak tanımlandığında ve $\lim \omega \rightarrow \infty$ iken $a\|\omega\|^2$ olduğundan dolayı $\lim L_{RR} \rightarrow \infty$ gerçekleşir.

$X'X$, I ve $aI + X'X$ matrislerinin boyutları $n \times n$ şeklindedir. $X'X$ matrisi yarı-pozitif tanımlıdır ve aI ile toplanarak matrisi pozitif tanımlı hale getirilir. Her pozitif tanımlı matrisin aynı zamanda tekil olmayan matristir ve $aI + X'X$ 'in tersi vardır. Eğer $a > 0$ ise sağrı regresyon için her zaman tek bir çözüm vardır.

Eğer $a = 0$ ise ve X 'in derecesi n 'den küçükse, matris tekil hale gelir. $\lim a \rightarrow 0$ iken matris tekilliğe yaklaşmaktadır ve ω değerini bulmak için kullanılan değerler daha kararsız hale gelirler. Çok büyük veya çok küçük sonuçlar ve yuvarlama hataları ile karşılaşılır. Regresyonun $a > 0$ seçilmesi hesaplamayı kararlı hale getirmektedir.

Eğer a çok büyük seçilirse aI baskın olur $X'X$ 'i tamamen etkisizleştirir ve tahmin performansını kötüleştirir. Eğer a çok küçük seçilirse nümerik sorunlarla karşılaşılabilir. Optimal değer ne çok büyük, ne çok küçük olmalıdır. $X'X$ elemanları ile uyumlu olmalıdır ve a seçimi veri kümesine özgü olmalıdır.

- **İkili Form:** $A[(aI + A'A)^{-1}]' = [(aI + A'A)^{-1}]'A$, $x \in \mathbb{R}^n$ ve $\omega'x$ özelliği kullanılarak sağrı regresyon çözümü şu şekilde ifade edilir.

$$\omega'x = ((aI + X'X)^{-1}X'Y)'x$$

$$= Y'X(aI + X'X)^{-1}x$$

$$= Y'(aI + XX')Xx$$

Tekil olmanın benzer özellikleri $aI + X'X$ üzerinde de kullanılır. $T \times T$ boyutunda olan matris ilk formula kıyaslandığında dezavantajlı olarak yorumlanabilir ve T örnek boyutu çok büyükse ancak bu formül önemli geneller yapabilme imkanı tanımaktadır.

$$\omega = Y'(aI + XX')^{-1}X$$

$$\omega'x = Y'(aI + K)^{-1}k$$

Burada K matrisi karşılıklı skaler yapıdadır.

$$K = \begin{pmatrix} (x_1, x_1) & (x_1, x_2) & \cdots & (x_1, x_T) \\ (x_2, x_1) & (x_2, x_2) & \cdots & (x_2, x_T) \\ \vdots & \vdots & \ddots & \vdots \\ (x_T, x_1) & (x_T, x_2) & \cdots & (x_T, x_T) \end{pmatrix}_{T \times T}$$

ve $k = k(x)$ vektör formundadır.

$$k = \begin{pmatrix} (x_1, x) \\ (x_2, x) \\ \vdots \\ (x_T, x) \end{pmatrix}$$

Tüm x_i ve x 'ler karşılıklı skaler durumdadır.

• **Lineer Olmayan Regresyon:** $n = 1$ ve tüm x 'lerin birer sayı olduğu 3. dereceden polinom yaklaşımları ile ilgilidir. Bu polinomların türevi alınıp ikili formda olduğu gibi basit bir argüman oluşturularak $L(\omega)$ şeklinde yazılabilir.

Rastgele seçilmiş bir x sinyali için $\Phi: X \rightarrow S$ haritalaması için S skaler bir vektör uzayı özellik uzayı olarak da tanımlanmaktadır. Bu özellik matrisi sağrı regresyon uzayında kullanılmaktadır.

$$\gamma_{RR} = Y'(aI + K)^{-1}k$$

Burada

$$K = \begin{pmatrix} \kappa(x_1, x_1) & \kappa(x_1, x_2) & \cdots & \kappa(x_1, x_T) \\ \kappa(x_2, x_1) & \kappa(x_2, x_2) & \cdots & \kappa(x_2, x_T) \\ \vdots & \vdots & \ddots & \vdots \\ \kappa(x_T, x_1) & \kappa(x_T, x_2) & \cdots & \kappa(x_T, x_T) \end{pmatrix}$$

ve

$$k = \begin{pmatrix} \kappa(x_1, x) \\ \kappa(x_2, x) \\ \vdots \\ \kappa(x_T, x) \end{pmatrix}$$

S uzayı sonlu boyutlu olmak zorunda değildir ve bununla birlikte tüm skaler vektörler Hilbert uzayı ile ifade edilebilir.

$\kappa: x^2 \rightarrow \mathbb{R}$ olduğundan $\kappa(x_1, x_2) = (\Phi(x_1), \Phi(x_2))$ ve κ bilindiğinde regresyon işlemi gerçekleştirilmektedir (Kernel Methods Article, 2011).

2.9.2. Destek Vektör Makinesi (Support Vector Machine) Yöntemi

Bir çeşit doğrusal veri sınıflandırma yöntemidir. Kullanılan örnek sayısı önemli değildir. Farklı sınıflardan durumlarda bulunduğu verilerde bir hedef noktası belirler. Bu hedef noktası üzerinden verileri sınıflandırır. Son zamanlarda çeşitli görüntü sınıflandırma algoritmalarında elde edilen başarımların yüksek olduğundan tercih edilen bir yöntemdir. Destek vektör makinelerinin anlaşılabilmesi için Vapnik Chervonenkins teoremini ele almak gerekmektedir (Burges, 1998).

2.9.2.1. Vapnik-Chervonenkins Teoremi (VC)

VC teoreminin temel fikri, genelleme yeteneğinin kontrol edilebilmesi için bir algoritma oluşturulmasına dayanır. İkili örüntü tanıma problemini incelemekte ve elde ettiği sonuçları m’li durum için genellemektedir (Costes ve Vapnik, 1995).

2.9.2.2. Destek Vektör Makinesi Algoritması

İki farklı sınıfın belirlenmesi problemine çözüm olarak üretilen yöntemdir. D adında n noktalı bir veri seti verilmiş olsun

$$D = \{(x_i, y_i) | x_i \in \mathbb{R}^p, y_i \in \{-1, 1\}\}_{i=1}^n$$

N eğitim örneği bulunmaktadır, $\{(y_i, x_i), i = 1, 2, \dots, N\}$ küme çiftleri ile temsil edilmektedir. Burada x_i ve y_i ’nin ait olduğu sınıflar belirtilmiştir. Boyutu p olan bir vektör x_i ile ifade edilir. Bu -1 ve 1 değerlerini alır. y_i ’de bu kümenin bir elemanını temsil etmektedir; y ve α sınıflandırıcı parametreleri $f(x, \alpha)$ fonksiyonu tarafından temsil edilmektedir. DVM yöntemi hiperdüzlemi optimum olarak sınıflara ayırmaktadır. Böylece

- Hiperdüzlemin her iki tarafında bulunan $y = \pm 1$ etiketleri ile örneklendirilmektedir.
- En yakın vektörlerin mesafesi normalin her iki tarafında bulunan maksimum hiperdüzlemdir.

Buna şekil 2.19’da gösterildiği gibi optimum paylı mesafe ve destek vektörleri denir. Bu hiperdüzlem aşağıdaki gibi tanımlanmaktadır.

$$\omega x + b = 0$$

Burada ω ve b hiperdüzlemin parametreleridir. Bu vektörler hiperdüzlem üzerinde değilse $\omega x + b \leq 0$ ’a gider. Bunların sınıflandırıcısı $f(x, \alpha) = \text{sgn}(\omega x + b)$ olarak

tanımlanır. Destek vektör makineleri paralel iki hiperdüzlem boyunca uzanmaktadır. Optimal hiperdüzleme ilişkin eşitlik aşağıdaki gibi ifade edilmektedir.

$$\omega x + b = \pm 1$$

İki destek vektör hiperdüzlemi denklemi ile payın maksimizasyonu sınırlı (kısıtlı) optimizasyon sorununa yol açmaktadır.

$$\min \left\{ \frac{1}{2} \|\omega\|^2 \right\}$$

Problemin kısıtları:

$$y_i(\omega x + b) \geq 1, \quad i = 1, \dots, N$$

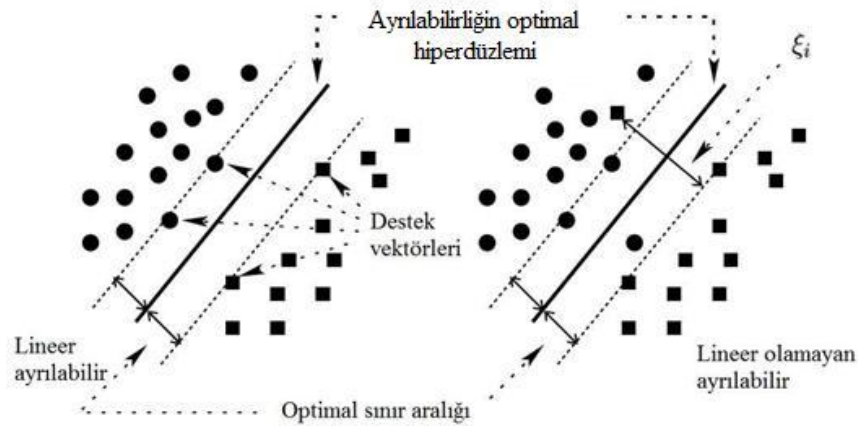
Eğer eğitim örnekleri doğrusal olarak ayrılabilir değilse Şekil 2.19'da gösterildiği gibi sınıflandırılmamış vektörlerin ağırlığını azaltmak için düzenli hale getirme parametresi C ve hata değişkenleri ξ_i olarak sınıflandırılır. Bu optimizasyon problemi Lagrange çarpanları kullanılarak çözülebilir.

$$\begin{cases} \min \left\{ \sum_{i=1}^N \lambda_i - \frac{1}{2} \sum_{i,j=1}^N \lambda_i \lambda_j y_i y_j x_i \cdot x_j \right\} \\ 0 \leq \lambda_i \leq C, & \forall i = 1, 2, \dots, N \\ \sum_{i=1}^N \lambda_i y_i = 0, & \forall i = 1, 2, \dots, N \end{cases}$$

Burada

λ_i : Lagrange çarpanlarıdır ve sadece destek vektörleri için sıfırdan farklıdır.

Hiperdüzlem parametreleri (ω, b) ve sınıflandırıcı fonksiyonu $f(x; \omega, b)$ olarak tanımlanmaktadır. Bunlar optimizasyon süreciyle hesaplanabilir (Burges, 1998), (Cortes ve Vapnik, 1995), (Mercier ve Lennon, 2003).



Şekil 2. 19 Destek vektör makineleri ile sınıflandırma (Mercier ve Lennon, 2003).

2.10. Zernike Momentleri Yöntemi

Zernike momentleri, birim çember içinde karmaşık tam dikgen küme çok terimliler tanımlanmaktadır. Bunlar Kartezyen koordinatlarda $V_{nm}(x, y)$ olarak gösterilmektedir. Zernike çokterimlileri kutupsal koordinatlarda aşağıdaki gibi ifade edilmektedir.

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta)$$

Burada

n : negatif olmayan bir tam sayıyı,

m : $n - |m|$ her zaman çift ve $|m| \leq n$ olmak üzere pozitif veya negatif bir tamsayıyı,

ρ : (x, y) noktasının orijine olan Öklid uzaklığını,

θ : (x, y) vektörünün x eksenine ile saat yönünde yaptığı açığı ifade etmektedir.

Zernike momentlerinin hesabı için önce görüntü belirli bir β oranıyla yeni bir boyuta normalize edilir (Çapar ve diğ., 2002), (Khotanzad ve Hong).

Zernike momentleri karmaşık polinomların bir setidir. Saat yönünün tersine x eksenine ile r vektörü arasındaki $R_{nm}(r)$ radyal polinomlar aşağıdaki gibi tanımlanmaktadır:

$$R_{nm}(r) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)!}{s! \left[\frac{n+|m|}{2} - s \right]! \left[\frac{n-|m|}{2} - s \right]!} r^{n-2s}$$

$f(x, y)$ için tekrar eden m ve n dereceli 2 boyutlu Zernike momenti aşağıdaki gibi tanımlanmaktadır.

$$Z_{nm} = \frac{n+1}{\pi} \iint_{\text{birim çember}} f(x, y) V_{nm}^*(x, y) dx dy$$

Burada

$$V_{nm}^*(x, y) = V_{n, n-m}(x, y)$$

Bir sayısal görüntünün Zernike momentini hesaplamak için sadece integrali toplam ile değiştirilmesi gerekmektedir.

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}^*(x, y)$$

Burada

$$x^2 + y^2 \leq 1$$

Zernike momentlerinin tanımlanmamış özelliklerinin değişmeyen rotasyonları vardır. Ölçek ve rotasyon değişmezliği sağlamaktadır. Görüntüleri ilk

olarak düzgün Zernike momentleri kullanarak normalize etmek gerekmektedir. Orijinal görüntünün rotasyon değişmezliği $f(x, y)$ için $f(x + \bar{x}, y + \bar{y})$ olmaktadır. Burada, $x = m_{10}/m_{00}$ ve $\bar{y} = m_{01}/m_{00}$ olarak ifade edilmektedir. Diğer bir deyişle orijinal görüntünün merkezi Zernike momentleri hesaplamadan önce ağırlık merkezine taşınmaktadır (Qader ve diğ., 2007).

Ölçek değişmezliği, büyütme veya her şekli azaltarak görüntünün sıfırinci düzenli moment denkleminde önceden belirlenmiş bir β değeridir. İkili görüntü için m_{00} görüntüdeki şekil piksel sayısına eşittir. Görüntü ölçeği için $f(\alpha x, \alpha y)$,

$$m'_{pq} = \alpha^{p+q+2} m_{pq} m_{pq}$$

$f(x, y)$ için düzenli momentler $\alpha\sqrt{\beta/m_{00}}$ ile $m'_{00} = \beta$, $m'_{00} = \alpha^2$, $m_{00} = \beta$ sağlanabilmektedir.

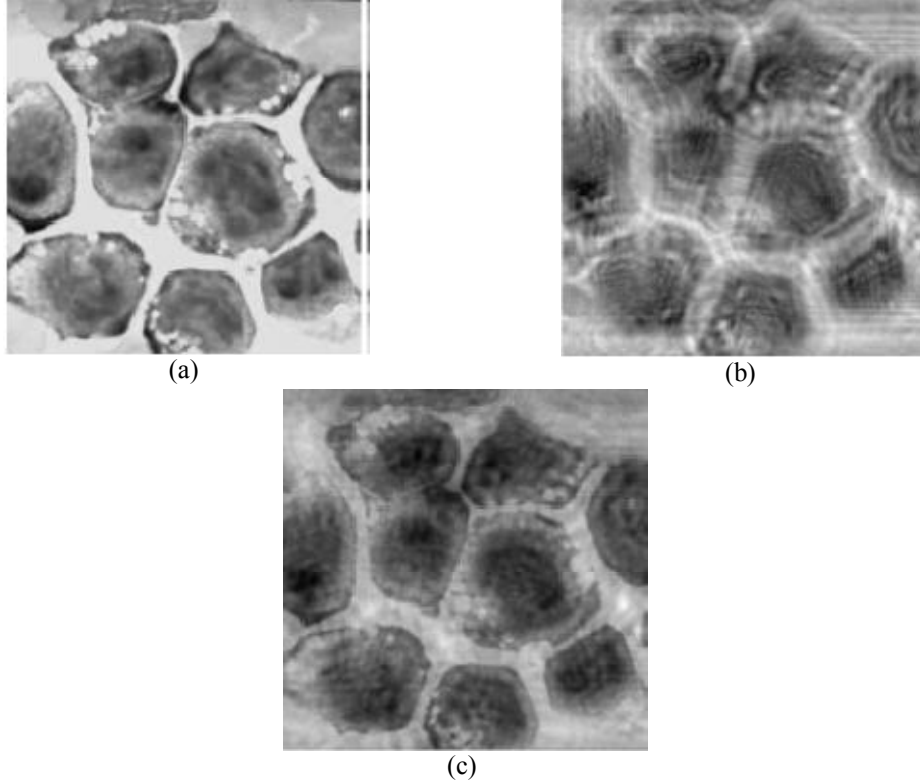
Belirtildiği gibi Zernike momentlerinin temel özelliği rotasyonel değişmezliktir. Eğer bir α açısıyla $f(x, y)$ döndürülürse bu durumda Zernike momentleriyle (Z_{nm}) döndürülmüş görüntü aşağıdaki gibi ifade edilmektedir (Quitti ve diğ., 2007).

$$Z'_{nm} = Z_{nm} \exp(-jm\alpha)$$

Böylece Zernike momentlerinin büyüklükleri rotasyon değişimsizdir, görüntü özellikleri olarak kullanılabilir. Zernike momentleri çeşitli yöntemlerle birlikte farklı alanlarda kullanılabilirdiği gibi tıbbi görüntü işleme tekniklerinde de yer bulmaktadır. Sayısal hologramdan gelen lösemi hücrelerinin yeniden yapılandırılması üzerine sınıflandırma işleminin performansı incelenmektedir. Zernike momentlerinin doğruluğunu arttırmak ve görüntü rotasyonunu ortadan kaldırmak için geometrik momentler elde edilmiştir (Chun-Lin, 2010).

Hologram görüntülerinde Zernike momentleri ile lösemi hücrelerinin tanınmasıyla ilgili çalışmalar incelendiğinde şöyle sonuçlar elde edilmektedir. Hücrelerin tanımlanması işlemi odak uzaklıklı iki lens ile her noktayı $+z$ ve $-z$ üzerinden dönüşüm işlemi ile çok benzerlik göstermektedir. Burada z hologram yüzeyindeki nesne noktaları arasındaki fark olarak ifade edilmektedir. Öyle ki, nesne yüzeyinden iki görüntü elde edilmektedir. Hologram yüzeyin diğer tarafında ve nesnenin bulunduğu yer olan bölgede sanal bir görüntüden gerçek görüntü elde edilmektedir. Biyolojik hücreler gibi düzlemsel olmayan nesnelere için, farklı nesne alanları ve bu nedenle farklı z mesafeleri, hücrelerin farklı şekillerde 3 boyutlu yorumlanmasını gerektirmektedir. Bir tıbbi mikroskop ile alınmış lösemili hücre

görüntüsü, bu görüntünün sayısal hologramı ve bu hücrenin Zernike momentleri kullanılarak tanımlanmış görüntüsü şekil 2.20'deki gibi elde edilmektedir (Asadi ve diğ., 2006).



Şekil 2. 20 Zernike momenti yöntemi kullanılarak görüntü tanıma. (a) Mikroskopik görüntü. (b) Sayısal hologram görüntüsü. (c) Zernike dönüşümü.

2.11. Dalgacık Dönüşümü

Dalgacık dönüşümü yöntemi, Bulgar matematikçi Alfréd Haar ve Belçikalı matematikçi Ingrid Daubechies tarafından geliştirilmiştir. Son 25 yıldır kullanılan en popüler yöntemlerden biridir. Görüntü sıkıştırma yöntemlerinde sıklıkla tercih edilmektedir. Bunun dışında yine görüntüde cismin bulunması, tanınması gibi çeşitli görüntü işleme uygulamalarında da etkin olarak kullanılmaktadır. Bir fonksiyonun dalgacık dönüşümünün alınması o fonksiyonun değişik çözünürlüklere ayırarak dalgacık katsayılarının belirlenmesi ile elde edilmektedir. Bu işlem için ana dalgacık adı verilen bir fonksiyon farklı zaman ve genişliklerde dönüşümü yapılmak istenen fonksiyon ile korelasyona tabi tutularak karşılık gelen dalgacık katsayıları elde edilmektedir.

İki boyutlu işaretlerde dalgacık dönüşümü görüntünün tekrarlı bir şekilde bir dizi alçak ve yüksek geçiren filtrelerden geçirilmesi ile gerçekleştirilmektedir. Her

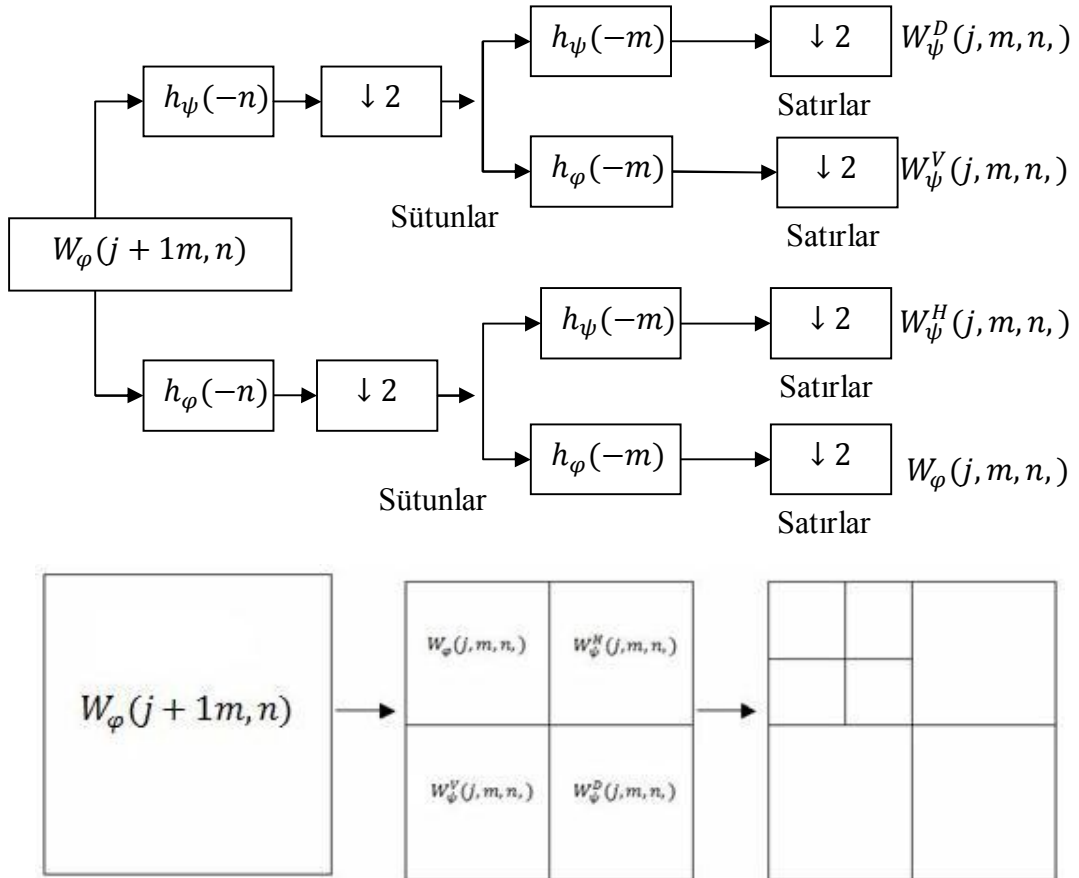
filtreleme görüntüye ait farklı bir çözünürlükte özet ve detay katsayılarının elde edilmesini sağlamaktadır. Bu işlem görüntü tek piksele indirgenene kadar devam edebilmektedir. Matematiksel olarak iki boyutlu dalgacık dönüşümü aşağıdaki gibi ifade edilmektedir.

$$W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \varphi_{j_0, m, n}(x, y)$$

$$W_\psi^i(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \varphi_{j, m, n}^i(x, y), i = \{H, D, V\}$$

$$f(x, y) = \frac{1}{\sqrt{MN}} \sum_m \sum_n W_\varphi(j_0, m, n) \varphi_{j_0, m, n}(x, y) + \frac{1}{\sqrt{MN}} \sum_{i=H,V,D} \sum_{j=j_0}^{\infty} \sum_m \sum_n W_\varphi^i(j, m, n) \varphi_{j, m, n}^i(x, y)$$

Herhangi bir çözünürlükteki dalgacık dönüşümü işleminin blok diyagramı aşağıdaki gibi ifade edilmektedir.



Şekil 2. 17 İki boyutlu dalgacık dönüşümü blok diyagramı ve matris gösterimi.

Şekil 2.21'de $W_{\varphi}, W_{\psi}^H, W_{\psi}^D, W_{\psi}^V$ değerleri görüntünün j çözünürlüğündeki özet, yatay ve diyagonal ve dikey dalgacık katsayılarına karşılık gelmektedir. Alçak ve yüksek geçiren filtrelerin katsayıları kullanılan ana dalgacığa göre belirlenmektedir.

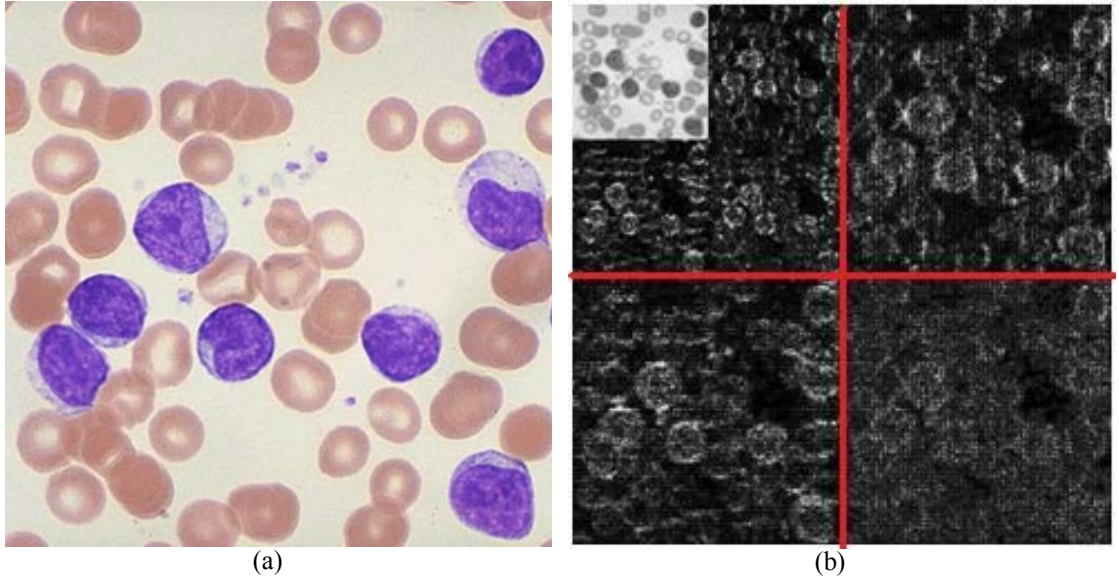
Burada

HL: x eksenini boyunca olan değişiklikleri,

LH: y eksenini boyunca olan değişiklikleri ifade etmektedir.

LL bandında güç daha küçüktür. Kodlama açısından düşük frekans bandı ve yüksek frekans bandının en düşük biti veya bunların sıfır setleri kullanılmaktadır. İki ekseninde görüntü ayırtırmaya ek olarak görüntünün doğal kenarları boyunca ayırtırma yöntemi tanımlanmıştır (Chun-Lin, 2010), (Petrosian, 2002).

Dalgacık dönüşü birçok görüntüde kullanıldığı gibi tıbbi görüntülerin işlenmesinde de çokça tercih edilen yöntemlerdendir. Kronik lösemi hücreleri bulunan mikroskop görüntüsü üzerindeki sonucu şekil 2.22'deki gibidir.



Şekil 2. 22 Dalgacık dönüşümünün lösemi hücrelerine uygulanması. (a) Orijinal lösemi hücre görüntüsü. (b) Dalgacık dönüşümü sonucu.

2.12. Genetik Algoritmalar (GA)

GA genellikle doğrusal olmayan çoklu fonksiyon optimizasyonu için global optimum çözümü belirler. Doğada gözlenen evrimsel sürece benzer şekilde çalışan arama ve optimizasyon yöntemidir (Holland, 1992). Problemlere tek bir çözüm üretmek yerine bir çözüm kümesi önerirler. Böylelikle, arama uzayında aynı anda birçok nokta değerlendirilmekte ve sonuçta bütünsel çözüme ulaşma olasılığı

yükselmektedir. Çözüm kümesindeki çözümler birbirinden tamamen bağımsızdır. Her biri çok boyutlu uzay üzerinde bir vektördür. Genetik algoritmalar problemlerin çözümünde evrimsel süreci bilgisayar ortamında taklit ederler. Genetik algoritmaların başarısına karar vermedeki en önemli faktör, problemin çözümünü temsil eden bireylerin gösterimidir. Nüfus içindeki her bireyin problem için çözüm olup olmayacağına karar veren bir uygunluk fonksiyonu vardır. Bunun yanı sıra çoklu mod fonksiyonunda yerel çözümün belirlenmesi için nişler azaltılır. Probleme ait en iyi çözümün bulunabilmesi için;

- Bireylerin temsili doğru bir şekilde yapılmalı,
- Uygunluk fonksiyonu etkin bir şekilde oluşturulmalı,
- Doğru genetik işlemciler seçilmelidir.

Bu durumda çözüm kümesi problem için bir noktada birleşecektir. Genetik algoritmalar, diğer optimizasyon yöntemlerini kullanılırken büyük zorluklarla karşılaşılan, oldukça büyük arama uzayına sahip problemlerin çözümünde başarı göstermektedir. Bir problemin bütünsel en iyi çözümünü bulmak için garanti vermezler. Ancak problemlere makul bir süre içinde, kabul edilebilir, iyi çözümler bulurlar. Genetik algoritmaların asıl amacı, hiçbir çözüm tekniği bulunmayan problemlere çözüm aramaktır. Kendilerine has çözüm teknikleri olan özel problemlerin çözümü için mutlak sonucun hızı ve kesinliği açısından genetik algoritmalar kullanılmazlar. Genetik algoritmalar ancak;

- Arama uzayının büyük ve karmaşık olduğu,
- Mevcut bilgiyle sınırlı arama uzayında çözümün zor olduğu,
- Problemin belirli bir matematiksel modelle ifade edilemediği,

geleneksel optimizasyon yöntemlerinden istenen sonucun alınmadığı alanlarda etkili ve kullanışlıdır (Beasley ve diğ., 1993 a, b).

2.12.1. Genetik Algoritmaları Kullanarak Biyomedikal Hiperspektral Veri Çalışmaları Bant Seçimi

Amaç hiperspektral doku çalışmaları için genetik algoritmalar kullanılarak niceliksel analiz gerçekleştirmek için en iyi bant kombinasyonu çözmektir.

2.12.1.1. Kodlama Şeması

Bant sayısı pozitif tamsayı ile temsil edilir. İkili kodlama stratejisi uygulanmaktadır. Hiperspektral veri küpünde maksimum bant sayısı 240 olarak seçilmektedir. 8 bitlik ikili dizi şeklinde ifade edilmektedir. Bu nedenle bir kromozom kodlamasının 3 bant sayısının dizi uzunluğu 24'tür. 8 bitlik alt dizeler içinde 0~239 değerlerine ihtiyaç varken, 240'tan 255'e kadar olan dizeler gereksiz ve etkisizdir. Örneğin; bir bireyin kodlama dizesi "10001011, 00101110, 01001101" ise kod çözme sonucunda 139, 46, 77 seçilir (Chunni ve diğ., 2009).

2.12.1.2. Uygunluk Fonksiyonu

Her bir popülasyon (nüfusa) bir başarı kriteri atamak eldeki problemin kromozom çözümlerinin ne kadar iyi olduğuna bağlıdır. Hiperspektral veri için bant seçimi performansını değerlendirmek amacıyla kullanılan OIF (Optimum Index Factor) etkili bir endekstir.

$$OIF = \frac{\sum_{i=1}^n \sigma_i}{\sum_{i=1}^n \sum_{j=i+1}^n |R_{i,j}|}$$

Burada

σ_i : standart sapma,

$R_{i,j}$: iki grup arasındaki korelasyon katsayısı,

Diğer seçenekler bu durumda yüksek standart sapma ve düşük korelasyonlar olacak şekilde belirlenir (Chunni ve diğ., 2009).

2.12.1.3. Adaptif Operatörler

Genetik algoritmalar, genel olarak bazı koşullarda kolayca etkili global optimum çözümler bulmaya yarar (Holland, 1992). Bu nedenle optimal arama sürecine GA için bazı adaptif stratejiler kullanılmaktadır. Çaprazlama ve mutasyon işlemlerinin ele alındığı tür bir genetik algoritma ele alınmaktadır (Chunni ve diğ., 2009). Pozitif bir tamsayı olan N_f nesiller için değişmediyse en iyi çözüm, çaprazlama ve mutasyon olasılıkları arasında aşağıdaki gibi bir ilişki vardır:

$$P_c = P_{c0} + \frac{N - N_f}{N} (\alpha - P_{c0})$$
$$P_m = P_{m0} + \frac{N - N_f}{N} (\beta - P_{m0})$$

Burada

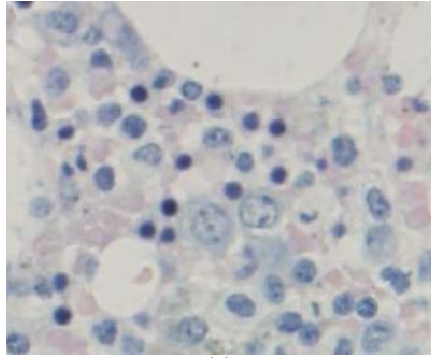
α, β : reel sabitler,

N : en iyi çözüm için nesil sayısı,

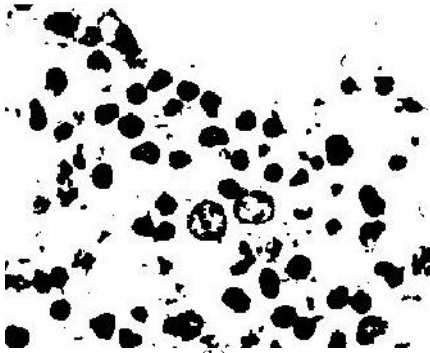
P_{c0}, P_{m0} : çaprazlama ve mutasyon olasılıklarıdır.

Aslında, α ve β , P_c ve P_m olarak belirtilenler maksimum değerlerdir. Eğer $N < N_f$ ise, ya da en iyi çözümdeki N sıfırlanıyorsa, P_{c0} ile P_{m0} birbirine eşit olur.

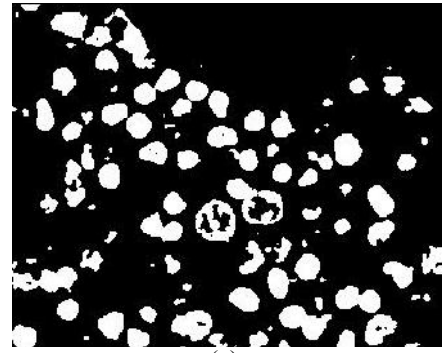
Verilen problemlerin çözümü için tasarlanacak algoritmalara genetik algoritma özelliği kazandırılarak daha iyi sonuçların elde edilmesi hedeflenebilmektedir. Aşağıda oluşturulan genetik algoritma ile kemik iliğindeki öncül hücrelerin tespiti gerçekleştirilmiştir. Bu algoritma incelenirse, kemik iliği patolojik görüntüsü ikili hale indirgenmekte ve görüntüye filtreleme ile optimizasyon uygulanmaktadır. Elde edilen görüntünün tersi (negatifi) alınmakta ardından sınırları doldurma işlemi uygulanmaktadır. Görüntüye morfolojik açma işlemi de uygulandıktan sonra sınırın doldurulması işlemi çıkışındaki görüntü ile birbirinden çıkarılmakta ve optimize edilmektedir. Hücre belirleme filtresi uygulandıktan sonra hücre yer tespiti ve segmentasyon işlemleri gerçekleştirilmektedir. Sonuçta elde edilen test sonucu görüntüsüdür (Li ve diğ., 2010).



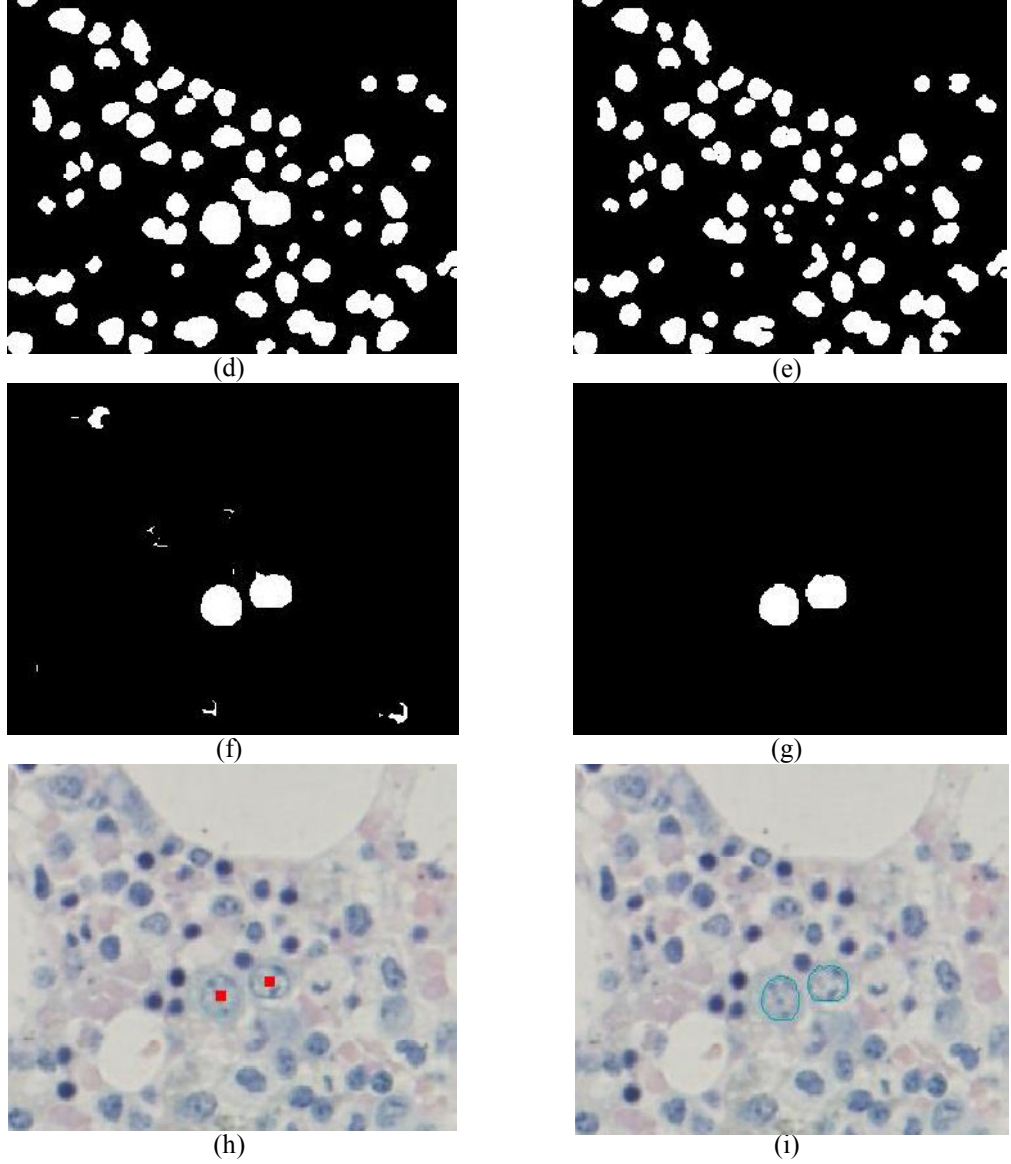
(a)



(b)



(c)



Şekil 2. 23 Bir GA algoritmasının kemik iliği patolojik görüntüsü üzerinde uygulanması. (a) Kemik iliği patolojik görüntüsü. (b) İkili gösterim. (c) İkili görüntünün tersi. (d) Ters alınmış ikili görüntüye sınır doldurma, morfolojik açma ve optimizasyon işlemi. (e) Ters alınmış ikili görüntüye sınır doldurma işlemi uygulanmadan, elde edilen morfolojik açma ve optimizasyon. (f) Önceki iki görüntü birbirinden çıkarılarak optimize edilen yeni görüntü. (g) Optimize edilmiş görüntünün hücre filtresi. (h) Görüntünün yer tespiti. (i) Orijinal görüntüdeki olgunlaşmamış öncül hücreyi gösteren görüntü.

GA, genel olarak bazı koşullarda kolayca etkili global optimum çözüm bulmaya yaradığından, verilen problemlerin çözümü için tasarlanacak algoritmalara genetik algoritma özelliği kazandırılarak daha iyi sonuçların elde edilmesi mümkündür. Benzer şekilde amaca yönelik farklı biçimlerde GA tasarımlarla gerçekleştirmek mümkündür.

3. LÖSEMİLİ HÜCRELERİN TANIMLANMASI VE SİMÜLASYONLAR

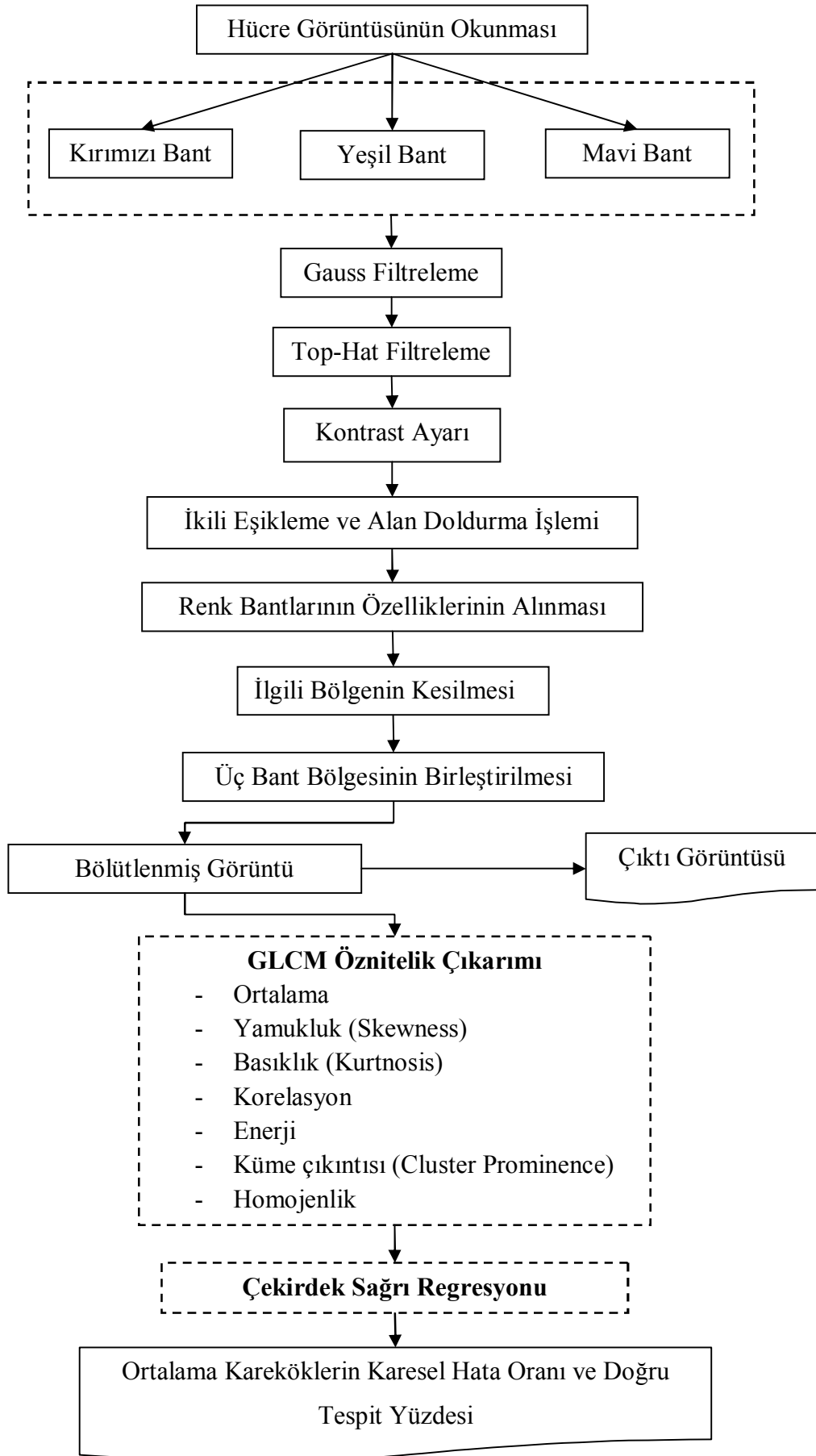
Lösemi, özellikle beyaz kan hücreleri veya lökositlerde oluşan anomaliler ile kan hücrelerinin normal üretimini olumsuz olarak etkileyerek ortaya çıkan bir kanser türüdür.

Hastalığın teşhis edilmesi ve doğru tedavi yönteminin belirlenmesi için güvenilir sonuç verecek algoritmaların tasarlanması ve uygulanması gereklidir. Bu çalışmada daha çok çocuklarda rastlanan ve teşhis edildiği durumda tedaviye açık olan aksi halde ölümcül sonuçlar doğurabilen ALL tipi hücrelerde morfolojik ön işlemeli ve çekirdek tabanlı bir örüntü tanıma algoritması ile erken ve daha doğru tanı sağlamaktır. ALL hücreleri normal hücrelerden ayrılıp tanınması için önerilen algoritma matematiksel ifadeleri ve elde edilen görüntü çıktıları ile açıklanmaktadır.

Lösemili hücrelerin sağlıklı hücrelerden ayrılması için tasarlanan örüntü tanıma algoritmaları içeriğinde çeşitli morfolojik özelliklerden faydalanarak ön eliminasyonlar gerçekleştirilmektedir. Ardından gri seviye eş oluşum matrisleri kullanılarak öznitelik çıkarım işlemleri istatistiksel temellere dayandırılarak gerçekleştirilmiştir. ALL hücreleri ile sağlıklı hücrelerin ayırt edilmesinde çekirdek sağrı regresyon yöntemi kullanılmıştır. Bu sınıflandırmanın ardından tasarlanan algoritmanın performans analiz sonuçları tablolar halinde sunulmuştur.

3.1. Çekirdek Sağrı Regresyonu Sınıflandırma Yöntemi Kullanarak Örüntü Tanıma Algoritması

Önerilen algoritmada tanıma işlemi, lösemi hücrelerinin boyutları, sayıları, çekirdek yuvarlaklık oranları ve olgunlukları gibi parametreler dikkate alınarak tasarlanmaktadır. Bu çalışmada ALL hücrelerinin tanınması ile ilgili bir örüntü tanıma algoritması tasarlanmıştır. Şekil 3.1'de algoritmanın akış diyagramı gösterilmektedir.



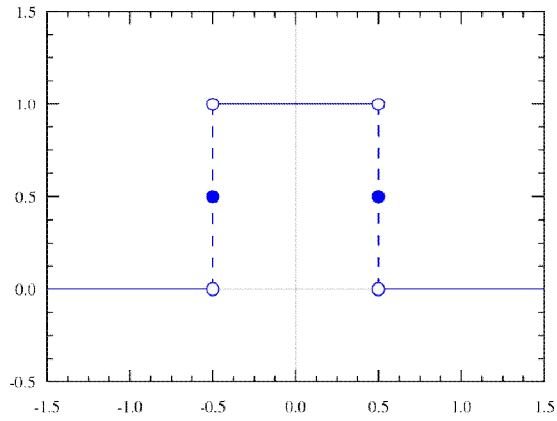
Şekil 3. 1 Çekirdek Sağrı Regresyonu Kullanılarak Tasarlanan ALL tanıma sistemi blok diyagramı.

Lösemi hastalığına sahip hücrelerin sağlıklı hücrelerden ayırt edilmesi için tasarlanan algorithmada 260 hücre görüntüsünden oluşan veri kümesine sırayla uygulanmaktadır.

Sayısal ortama aktarılan hücre görüntüleri önce RGB (Kırmızı-Yeşil-Mavi) renk bantlarına ayrıştırılmaktadırlar. Elde edilen üç farklı banttaki görüntü çıktılarına algoritma işlemleri adım adım uygulanmaktadır.

Gauss filtresi: Görüntü üzerindeki gri seviye ve arkaplan gürültülerinin bastırılmasında etkin sonuç vermesinden dolayı kullanılmaktadır.

Top-Hat filtresi: Gerçek uzay veya Fourier uzayında filtreleme işlemi uygulamaya olanak sağlayan dikdörtgen fonksiyona sahip bir filtre türüdür.



Şekil 3. 2 Dikdörtgen Top-Hat fonksiyonu.

Görüntü özellikle arkaplanda meydana gelen aydınlık ve karanlık seviyelerin ayrıştırılmasında kullanılır. Böylece görüntüde tepe ve çukur (pick point) bölgeler belirlenir. Açma işlemi uygulanan görüntünün orijinal görüntüden çıkarılmasıyla aydınlık bölgeler elde edilir. Kapama işlemi uygulanan görüntünün orijinal görüntüden çıkarılmasıyla ile de karanlık bölgeler elde edilir.

Kontrast ayarı: Hücre görüntüleri üzerinde filtreleme işlemleri gerçekleştirildikten sonra özellikle yüksek frekanslı bölgelerin belirlenmesi için kullanılmaktadır. Böylece hücre çerçeveleri belirlenmeye başlamaktadır.

İkili eşikleme ve alan doldurma işlemi: Hücre görüntüsü hücrenin hesaplanması için gerekli olan bölgesi ve arkaplanı ayıracak şekilde eşiklenmiştir. Hücre alanına dahil olup da arkaplan gibi değerlendirilen oyuklara alan doldurma işlemi uygulanarak bu probleminde üstesinden gelinmektedir.

Renk bantlarının özelliklerinin alınması: Kırmızı, yeşil ve mavi renk bantlarına bölütlenmiş gri seviyeli işlenmiş hücre görüntülerinin her üçünden ayrı

ayrı özellik çıkarımı gerçekleştirilir. Hücrelerin bulunduğu alanın satır ve sütun bilgileri birer dizi halinde tutulur.

Görüntü üzerinden ilgili bölgenin kesilmesi: Bu işlemde her renk bandında mavi hücrelerin belirlenerek etiketlenmesi ve hücre bilgilerinin dizilerde saklanması tasarlanmıştır. Bu etiketlenen görüntüler üzerinden ölçümler gerçekleştirilir. Hücrelerin boyutlarına göre görüntünün diğer kalanının kesilerek sadece ilgili bölgeye ulaşılmıştır.

Üç renk bant bölgelerinin birleştirilmesi: Görüntünün geri kalanından çıkarılan kırmızı, yeşil ve mavi renk bantlarındaki hücreler birleştirilerek görüntü yeniden boyutlandırılır. Böylece çekirdek sağrı regresyonu uygulanacak ve istatistiksel çıkarımlar yapılacak olan son görüntü elde edilmektedir.

Görüntüden özellik çıkarılması: Görüntü üzerinden ortalama, yamukluk (skewness), basıklık (kurtosis), enerji, grup öncelliği (cluster prominence) ve homojenlik özelliklerine göre görüntüyü analiz etmektedir. Böylece görüntünün belirleyici özelliklerinden faydalanarak eliminasyon gerçekleştirilir ve veri kümesindeki hasta ve sağlıklı hücreler tespit edilir.

Öğrenme ve test sınıflandırıcı: Sonuçta hasta olmayıp doğru tespit edilen, hasta olmayıp yanlış tespit edilen, hasta olup doğru tespit edilen, hasta olup yanlış tespit edilen olasılıklar üzerinden hücre görüntülerini veri kümesi üzerinden hesaplayarak başarımlar analizleri elde edilir.

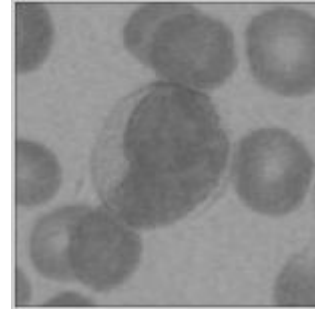
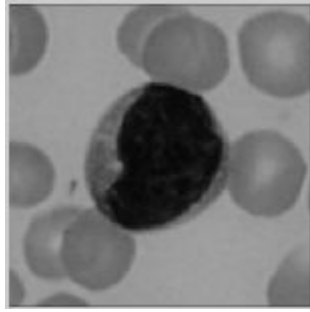
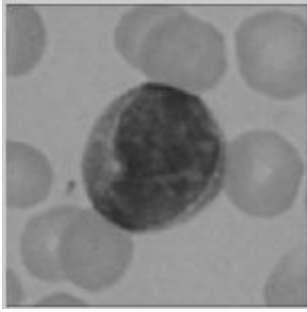
Başarım analizi (Performans kriteri): Çıktı görüntüsü, kareköklerin ortalama karesel hata oranı (root mean square error) ve doğru teşhis yüzdesi performans kriteri olarak hesaplanmaktadır (Papoulis ve Pillai, 2002: 590).

Çalışma için kullanılan veri kümesinden rastgele 5 ALL hücresini özellik çıkarımı için gerekli olan son görüntü elde edilmektedir. Bunun için algoritmanın her adımdan üç renk bandı için ekran çıktıları incelenmektedir. Sonuçlar şekil 3.3-3.7'de gösterilmiştir.

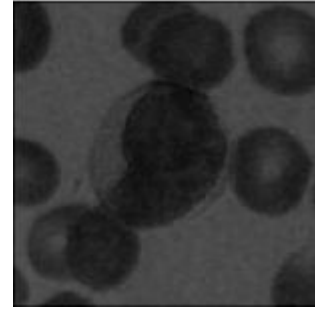
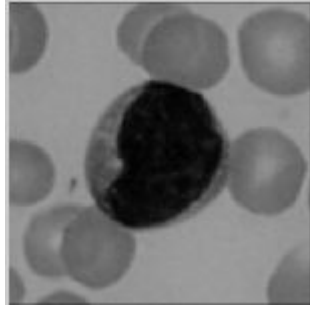
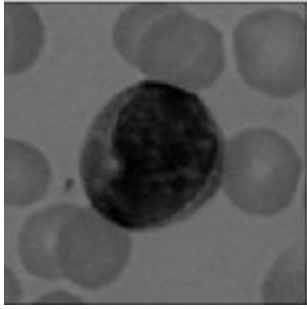
Kırmızı Bant

Yeşil Bant

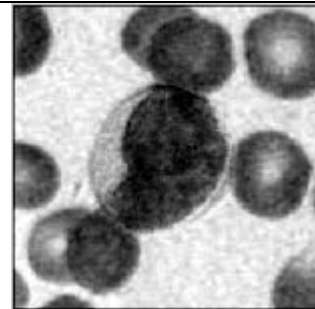
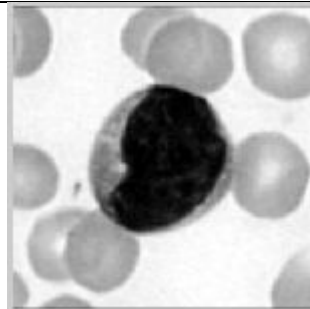
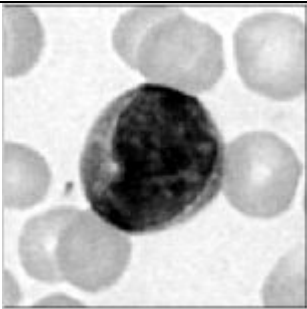
Mavi Bant



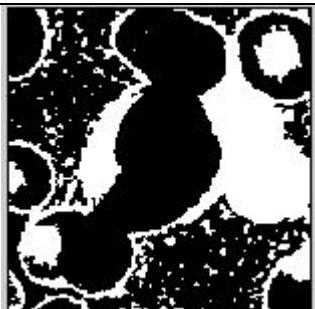
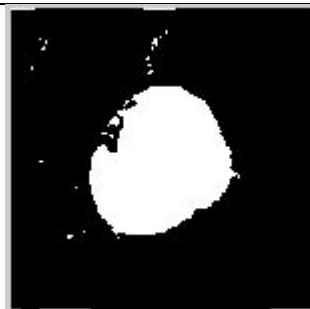
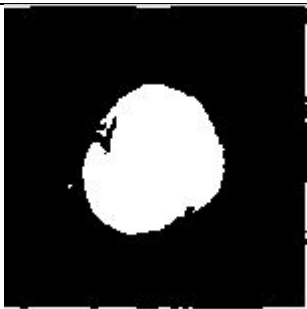
(a)



(b)



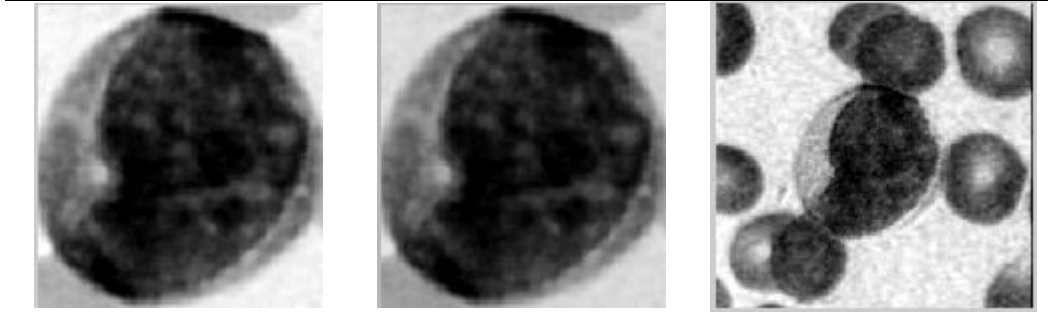
(c)



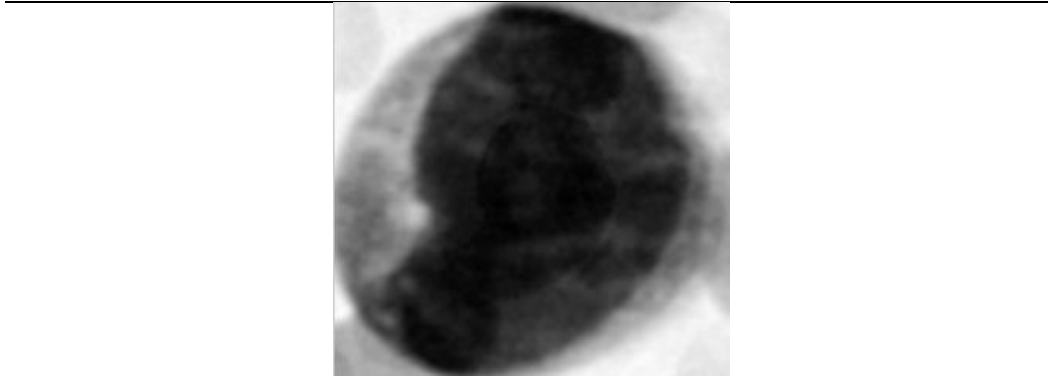
(c)



(d)

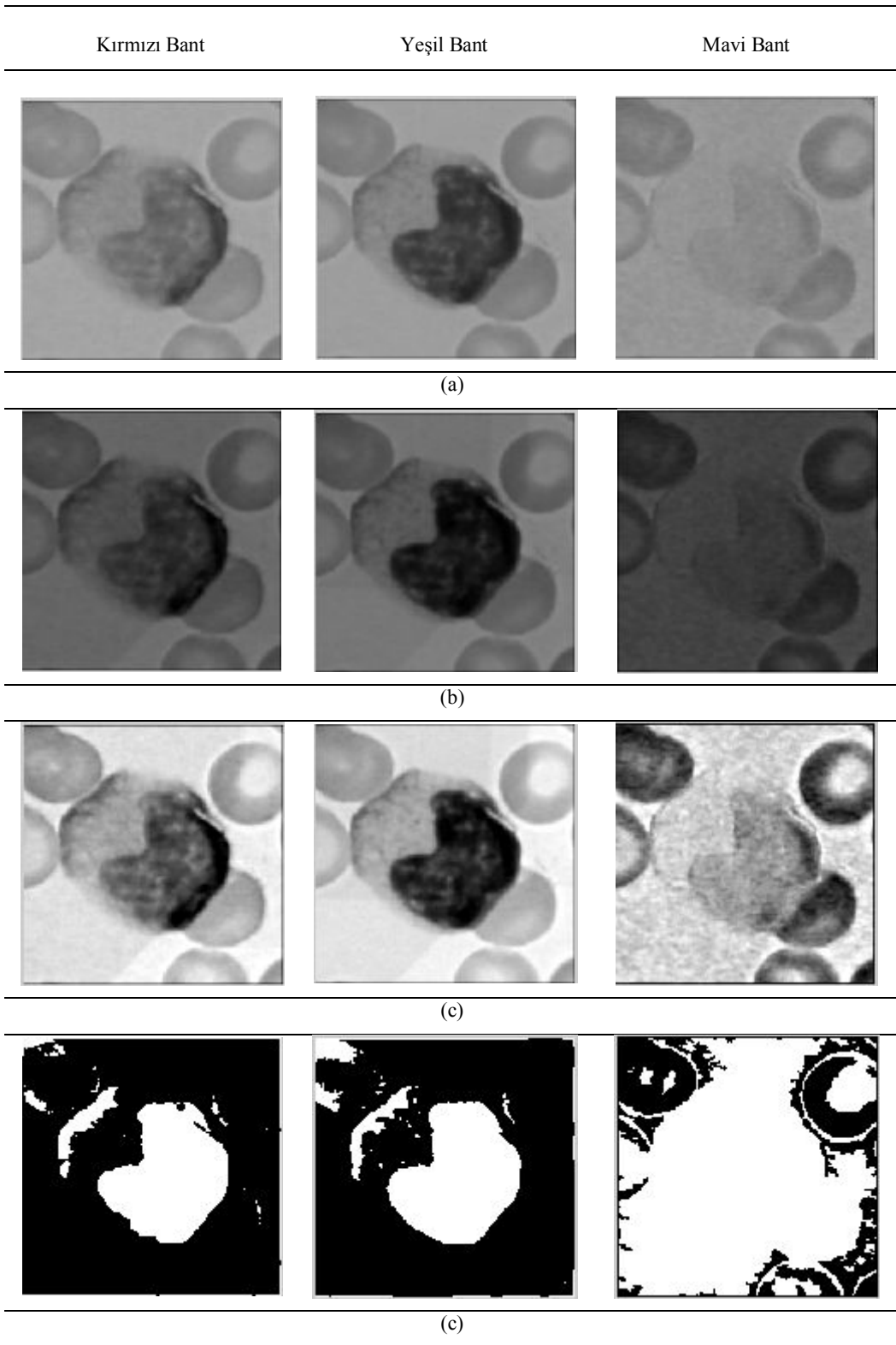


(e)



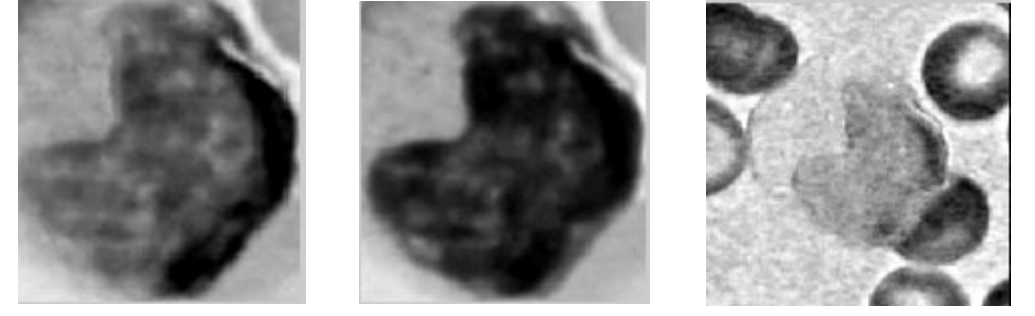
(f)

Şekil 3. 3 ALL Hücrelerinin teşhis edilmesi için tasarlanan algorithmada elde edilen çıktılar. (a) Gauss filtre çıkışı. (b) Top-Hat filtre çıkışı. (c) Kontrast ayarı çıkışı. (d) İkili eşikleme ve alan doldurma işlemi. (e) Özellikleri alınmış görüntüde ilgili bölgenin kesilmesi. (f) Üç renk bant bölgesinin birleştirilmesi ile elde edilen son görüntü.

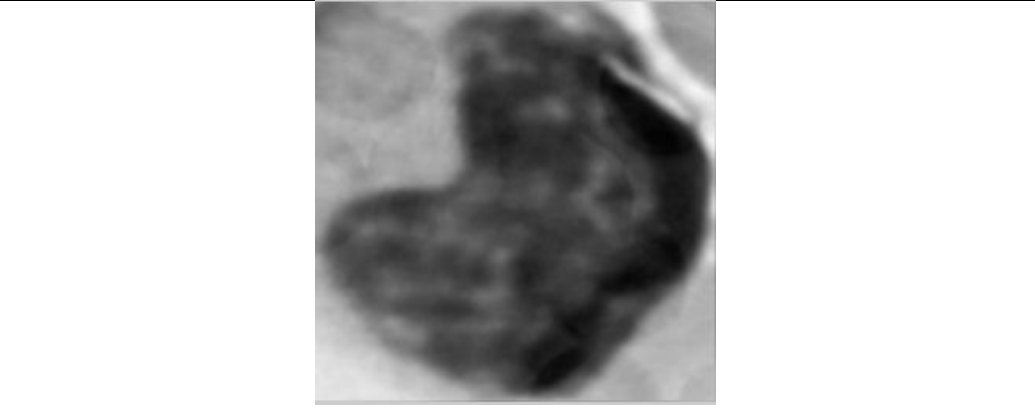




(d)



(e)



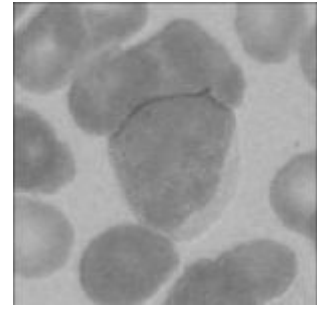
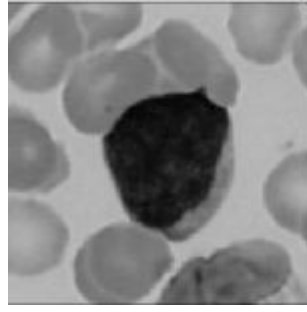
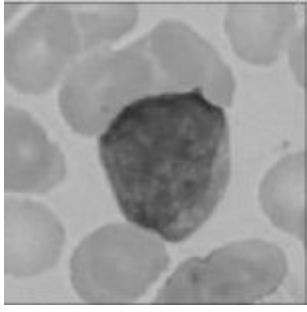
(f)

Şekil 3. 4 ALL Hücrelerinin teşhis edilmesi için tasarlanan algorithmada elde edilen çıktılar. (a) Gauss filtre çıkışı. (b) Top-Hat filtre çıkışı. (c) Kontrast ayarı çıkışı. (d) İkili eşikleme ve alan doldurma işlemi. (e) Özellikleri alınmış görüntüde ilgili bölgenin kesilmesi. (f) Üç renk bant bölgesinin birleştirilmesi ile elde edilen son görüntü.

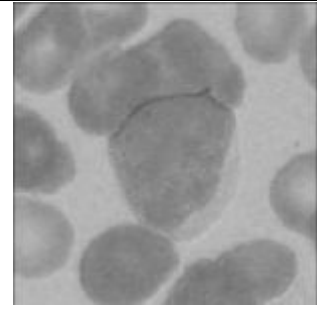
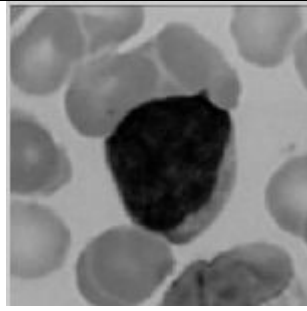
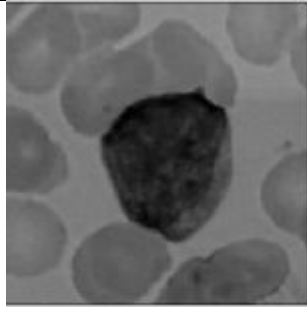
Kırmızı Bant

Yeşil Bant

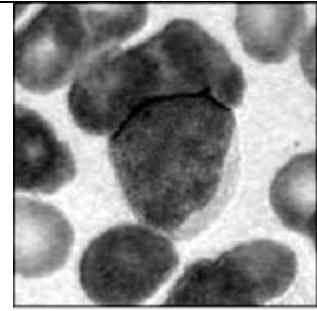
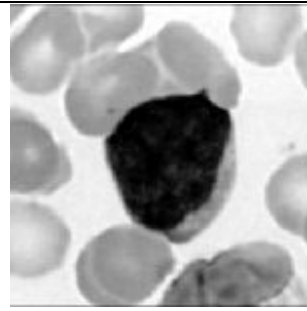
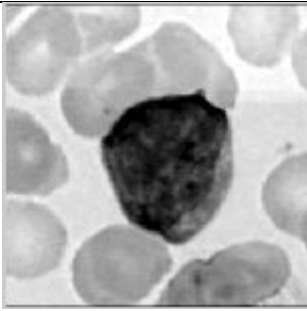
Mavi Bant



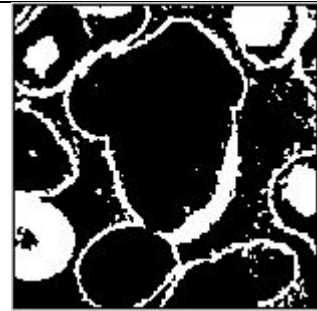
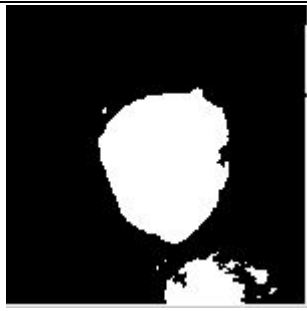
(a)



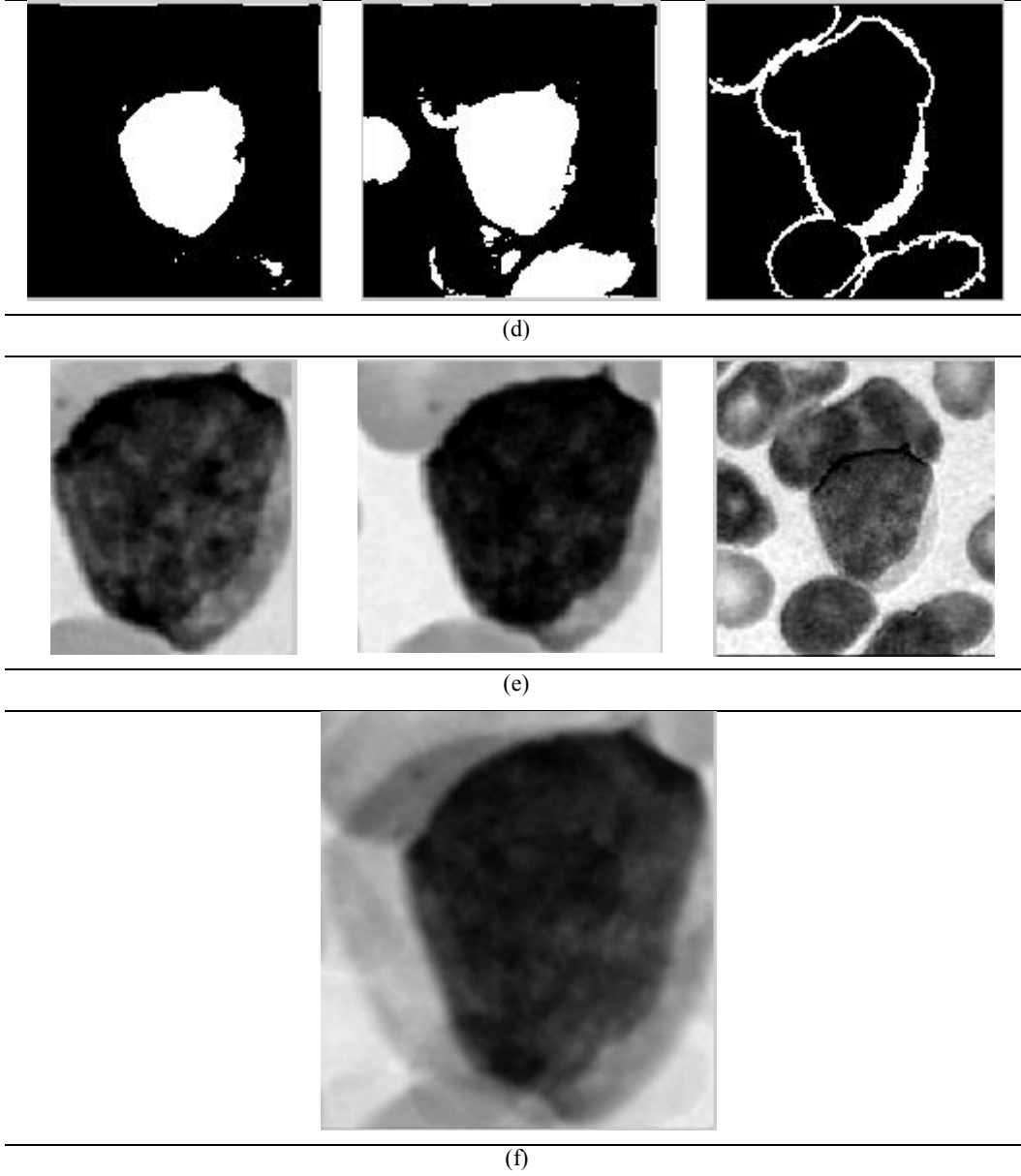
(b)



(c)



(c)

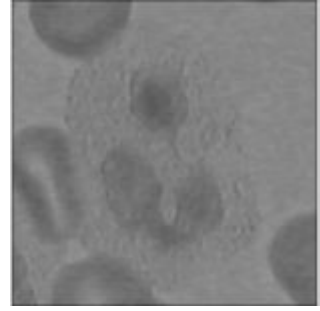
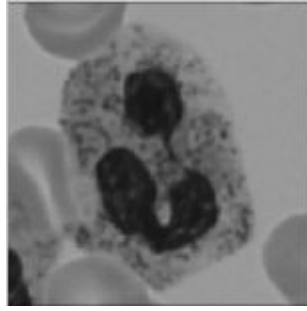
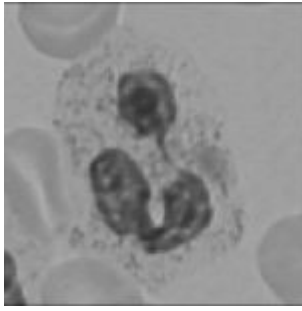


Şekil 3. 5 ALL Hücrelerinin teşhis edilmesi için tasarlanan alırtmada elde edilen çıktılar. (a) Gauss filtre çıkışı. (b) Top-Hat filtre çıkışı. (c) Kontrast ayarı çıkışı. (d) İkili eşikleme ve alan doldurma işlemi. (e) Özellikleri alınmış görüntüde ilgili bölgenin kesilmesi. (f) Üç renk bant bölgesinin birleştirilmesi ile elde edilen son görüntü.

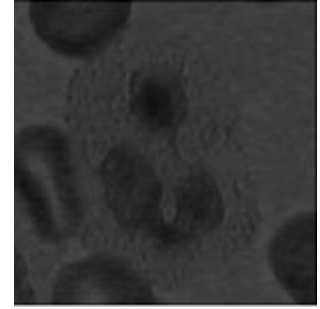
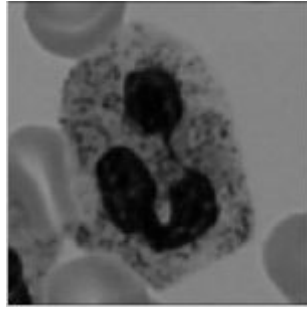
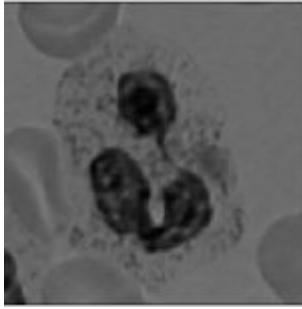
Kırmızı Bant

Yeşil Bant

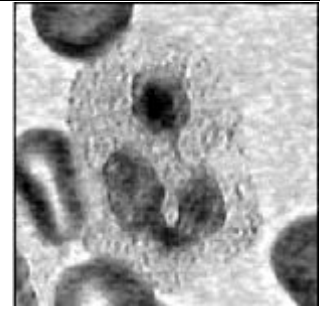
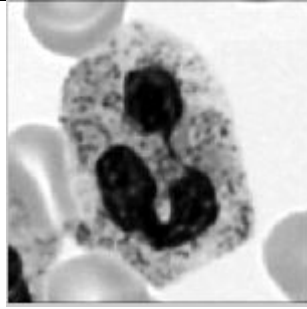
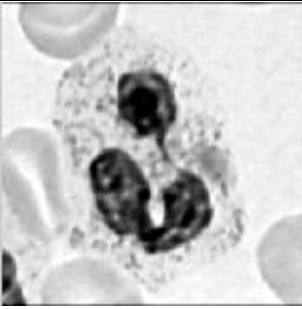
Mavi Bant



(a)



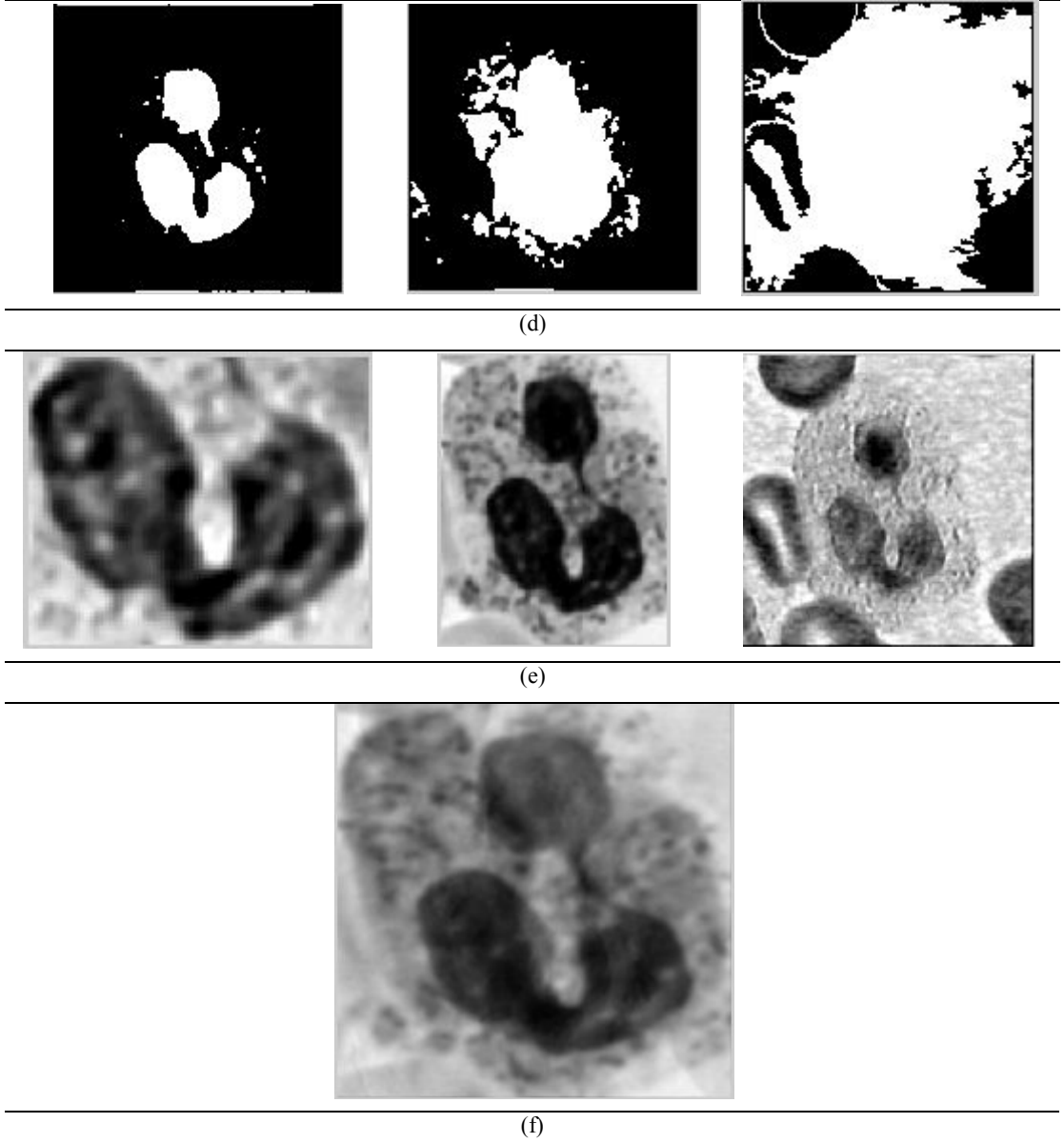
(b)



(c)



(c)

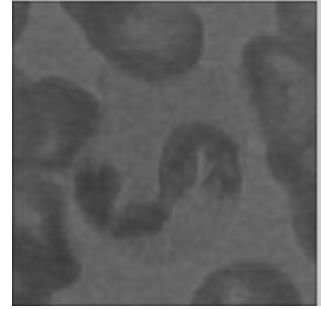
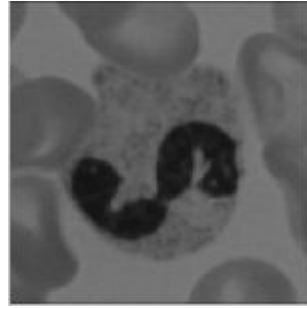
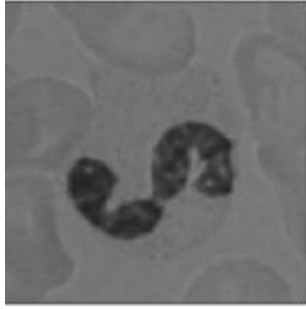


Şekil 3. 6 ALL Hücrelerinin teşhis edilmesi için tasarlanan algoritmada elde edilen çıktılar. (a) Gauss filtre çıkışı. (b) Top-Hat filtre çıkışı. (c) Kontrast ayarı çıkışı. (d) İkili eşikleme ve alan doldurma işlemi. (e) Özellikleri alınmış görüntüde ilgili bölgenin kesilmesi. (f) Üç renk bant bölgesinin birleştirilmesi ile elde edilen son görüntü.

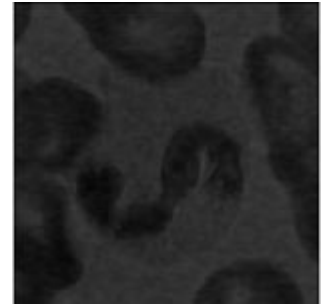
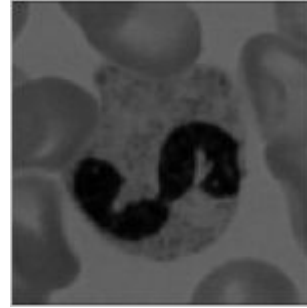
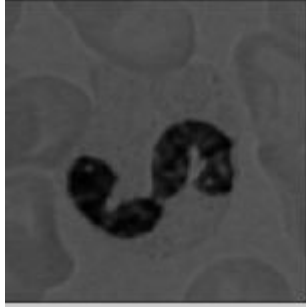
Kırmızı Bant

Yeşil Bant

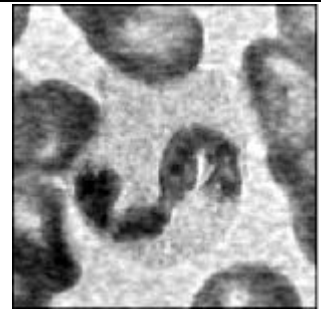
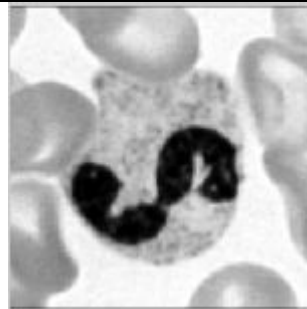
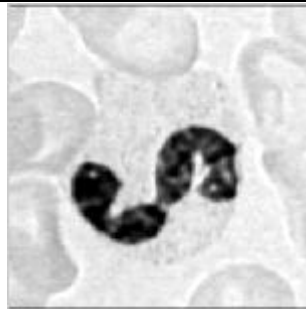
Mavi Bant



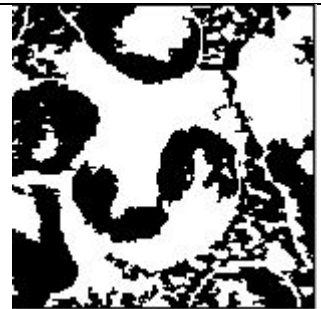
(a)



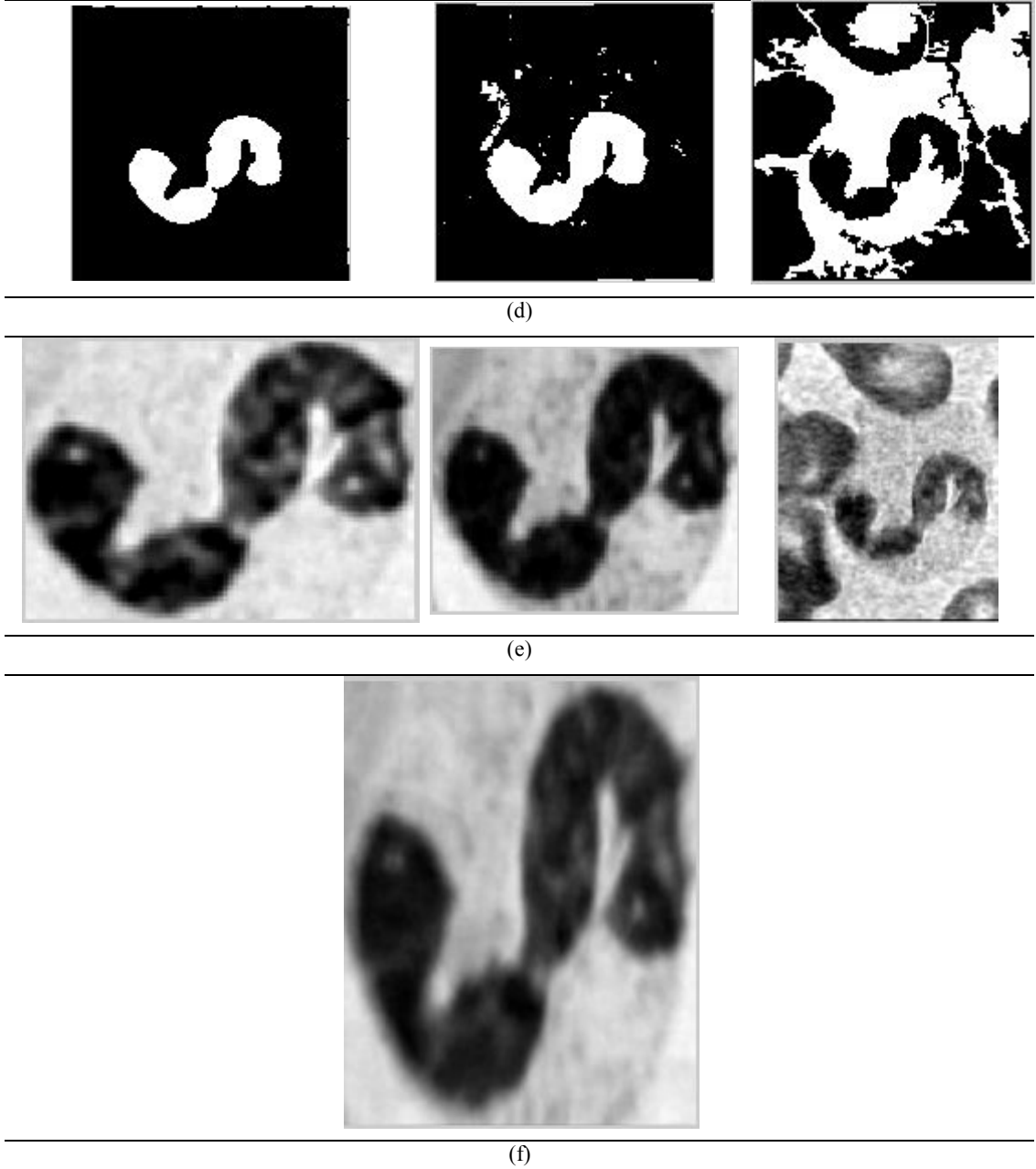
(b)



(c)



(c)



Şekil 3. 7 ALL Hücrelerinin teşhis edilmesi için tasarlanan algorithmada elde edilen çıktılar. (a) Gauss filtre çıkışı. (b) Top-Hat filtre çıkışı. (c) Kontrast ayarı çıkışı. (d) İkili eşikleme ve alan doldurma işlemi. (e) Özellikleri alınmış görüntüde ilgili bölgenin kesilmesi. (f) Üç renk bant bölgesinin birleştirilmesi ile elde edilen son görüntü.

Elde edilen son görüntüler üzerinden örüntü tanıma algoritması çekirdek sağrı regresyonu için GLCM yöntemi ile özellik çıkarımları yapılır. Ortalama, yamukluk (skewness), basıklık (kurtosis), korelasyon, enerji, grup çıkıntısı (cluster prominence) ve homojenlik özellikleriyle elde edilen sayısal sonuçlar karşılaştırılarak başarımları analizleri elde edilir.

3.2. Gri Seviye Eş Oluşum Matrisleri (Gray Level Coocurrence Matrices GLCM) ile Öznitelik Çıkarımı

İçerik tabanlı görüntü erişim sistemleri, görüntüyü tanımlamak için görsel tanımayıcılar kullanır. Bu tanımlayıcılar renk, doku ve şekil olmak üzere üç temel grupta toplanabilir. Çalışmanın bu aşamasında dokunun istatistiksel anlamda tanımlanmasıyla ilgilenilmektedir.

Doku, sezgisel bir kavram olması sebebiyle yıllardır tanımlanması zor bir kavram olmuştur. İstatistiksel yaklaşımlar, dokuyu bir alandaki yoğunluk dizilişinin nicel bir ölçüsü olarak tanımlar. Her doku, çok boyutlu öznitelik alanında bir noktayı gösteren özelliklerin öznitelik vektörü ile betimlenir. İstatistiksel yaklaşımlar, konumsal alan, sinyal işleme ve model tabanlı yöntemlerini kapsar. Konumsal alan yöntemlerinde öznitelikler, özellikle ikinci sıra istatistiklerden elde edilir, çünkü insanlar ikinci sıradaki istatistiklere karşı daha duyarlıdır. Bu tür istatistiklerin örnekleri, gri seviye eş oluşum matrisi ve gri seviye varyans matrisidir. Gri seviye dizi uzunluğu, Fourier güç spektrumu ve otokorelasyon işlevi gibi daha yüksek düzeydeki istatistikler de ölçülebilir. Bu çalışmada kapsamında, istatistiksel uzaysal yaklaşım yöntemleri ile çalışılmaktadır.

Uzaysal alan yöntemlerinde öznitelikler, bazı istatistikler sonucunda elde edilir. İnsan görme sistemi de bu istatistiksel bilgilere duyarlıdır. Bu tür istatistiklerin örnekleri, gri seviye eş oluşum matrisi ve gri seviye varyans matrisi, gri seviye dizi uzunluğu matrisleridir. İmgenin geneli üzerinde olasılıksal bir hesaplama sonucu bilgiler sunan uzaysal alan yöntemleri frekans tabanlı yöntemlerinden ayırdırlar ve görsel sistemle ve alt bant resimlerle herhangi bir işlem yürütmezler.

GLCM bir matris olarak tanımlanmaktadır. Bu matristeki satır ve sütun sayıları G görüntüsü içindeki gri seviyelerin sayısına eşittir. Matris elemanı $P(i, j | \Delta_x, \Delta_y)$ bağıl frekansı ile iki piksel arasındaki mesafe (x, y) ile ayrılmıştır. Verilen komşulukta i ve j yoğunlukları ifade etmektedir. Matris elemanı $P(i, j | d, \theta)$ iken belirli gri seviyeler olan i ve j arasındaki değişiklikleri istatistiki olasılıkları ikinci derecedendir. Burada d deplasman mesafesi, θ belirli bir açı değeridir. Verilen $M \times N$ komşuluğu giriş görüntüsünün gri seviyelerini içermektedir. Buna $f(m, n)$ denirse m örnek yoğunluğu, n komşuluk çizgisidir.

$$P(i, j | \Delta_x, \Delta_y) = WQ(i, j | \Delta_x, \Delta_y)$$

Burada

$$W = \frac{1}{(M - \Delta_x)(N - \Delta_y)}$$

$$Q(i, j | \Delta_x, \Delta_y) = \sum_{n=1}^{N-\Delta_y} \sum_{m=1}^{M-\Delta_x} A$$

$$A = \begin{cases} 1, & \text{eğer } f(m, n = i \text{ ve } f(m + \Delta_x, n + \Delta_y) = j) \\ 0, & \text{diğer} \end{cases}$$

GLCM ile görüntünün gri seviyeleri üzerinden doku özellikleri elde edilmektedir. Bu bilgiler ışığında aşağıdaki tanımlamalar yapılabilmektedir. μ , P 'nin ortalama değeridir.

μ_x, μ_y : P_x ve P_y 'nin ortalama değeri,

σ_x, σ_y : P_x ve P_y 'nin standart sapması olarak tanımlanır.

$P_x(i)$, i . dereceden marjinal olasılık matrisinin $P(i, j)$ 'nin satırlarının toplanmasıyla elde edilmektedir.

$$P_x(i) = \sum_{j=0}^{G-1} P(i, j)$$

$$P_y(i) = \sum_{i=0}^{G-1} P(i, j)$$

$$\mu_x = \sum_{i=0}^{G-1} i \sum_{j=0}^{G-1} P(i, j) = \sum_{i=0}^{G-1} i P(i, j)$$

$$\mu_y = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} j P(i, j) = \sum_{j=0}^{G-1} j P(i, j)$$

$$\sigma_x^2 = \sum_{i=0}^{G-1} (i - \mu_x)^2 \sum_{j=0}^{G-1} P(i, j) = \sum_{i=0}^{G-1} (P_x(i) - \mu_x(i))^2$$

$$\sigma_y^2 = \sum_{j=0}^{G-1} (j - \mu_y)^2 \sum_{i=0}^{G-1} P(i, j) = \sum_{j=0}^{G-1} (P_y(j) - \mu_y(j))^2$$

Buradan

$$P_{x+y}(k) = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i, j)$$

burada, $i + j = k$ ve $k = 0, 1, \dots, 2(G - 1)$

$$P_{x-y}(k) = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i, j)$$

burada, $|i + j| = k$ ve $k = 0, 1, \dots, G - 1$.

GLCM yöntemi görüntü üzerinden farklı özelliklerin çıkarılması için farklı istatistiksel yaklaşımlar gerektirmektedir (Albregtsen, 2008).

3.2.1. Yamukluk (Skewness)

Olasılık ve istatistik teorisinde yamukluk bir asimetri ölçüsü olarak reel değerli rastlantısal değişkendir. Yamukluk değeri pozitif, negatif veya tanımlanmamış olabilir. Niteliksel olarak, negatif yamukluk olasılık yoğunluk fonksiyonunun sol tarafındaki kuyruğu pozitif yamukluk da sağ tarafındaki kuyruğu temsil etmektedir.

Olasılık yoğunluk dağılımı pozitif ve negatif düzlemde simetrik dağılıma sahip olduğunda sıfır değerlidir denir. Ancak her sıfır değerli fonksiyonun dağılımı simetrik olmayabilir (Haralick ve Shanmugam, 1973).

$$g_1 = \frac{m_3}{m_2^{3/2}} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{3/2}}$$

Burada

g_1 : örnek yamukluk fonksiyonu,

n : yamukluk için örnek sayısı,

\bar{x} : ortalama değer,

m_2 : örnek varyans,

m_3 : 3. merkezi moment.

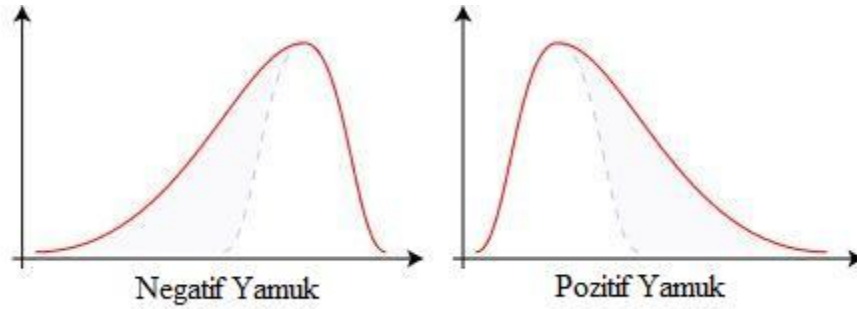
Yamukluk olağan popülasyon (nüfus) dağılımı kestirimi aşağıdaki gibi ifade edilebilir.

$$G_1 = \frac{k_3}{k_2^{3/2}} = \frac{\sqrt{n(n-1)}}{n-2} g_1$$

Burada

k_3 : 3. kümülant tek yansız simetrik kestirim,

k_2 : 2. kümülant tek yansız simetrik kestirim.



Şekil 4. 1 Yamukluk dağılım fonksiyonları.

Örneğin merkezleri -99, 0.5 ve 2'de bulunan ve ağırlıkları 0.01, 0.66 ve 0.33 olan dar Gauss dağılımlarının karması olan dağılımın yamukluğu -9.77 civarındadır. Ancak üç örnek alındığında G_1 'in beklenen değeri 0.32'dir. Her üç örnekte dağılımın diğer yöne yamuk olan pozitif değerli kısmındadır.

3.2.2. Basıklık (Kurtosis)

Olasılık ve istatistik teorisine göre basıklık reel değerli rastlantısal olasılık dağılımına sahip bir "sivrilik" ölçüsüdür. And not peakedness, is what is really being measured by kurtosis (Haralick ve Shanmugam, 1973).

Yüksek basıklık varyansın sonucu frekansı daha yüksek düşük ölçekli sapmalara karşı düşük frekanslı sapmalar anlamına gelir. Basit bir basıklık fonksiyonu tanımlanırsa:

$$g_2 = \frac{m_4}{m_2^2} - 3 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} - 3$$

Burada

m_4 : 4. örnek merkezi momenti,

m_2 : örnek varyansı,

\bar{x} : örnek ortalaması,

n : örnek boyutu.

$$D = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

$$E = \frac{1}{nD^2} \sum_{i=1}^n (x_i - \bar{x})^4 - 3$$

Burada

D : varyans,

\bar{x} : aritmetik ortalama. Basıklık fonksiyonunun popülasyon kesitirimi aşağıdaki gibi ifade edilebilir:

$$\begin{aligned} G_2 &= \frac{k_4}{k_2^2} = \frac{n^2((n+1)m_4 - 3(n-1)m_2^2)(n-1)^2}{(n-1)(n-2)(n-3) n^2 m_2^2} \\ &= \frac{n-1}{(n-2)(n-3)} \left((n+1) \frac{m_4}{m_2^2} - 3(n-1) \right) \\ &= \frac{n-1}{(n-2)(n-3)} ((n+1)g_2 + 6) \\ &= \frac{(n+1)n(n-1)}{(n-2)(n-3)} \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{(\sum_{i=1}^n (x_i - \bar{x})^2)^2} - 3 \frac{(n-1)^2}{(n-2)(n-3)} \\ &= \frac{(n+1)n}{(n-1)(n-1)(n-3)} \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{k_2^2} - 3 \frac{(n-1)}{(n-2)(n-3)} \end{aligned}$$

Burada

k_4 : 4. kümülant tek yansız simetrik kestirim,

k_2 : Popülasyon varyansının yansız kestirimi,

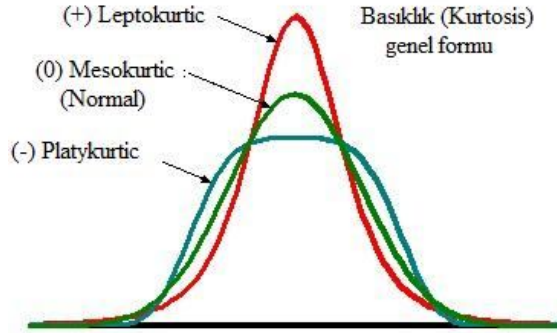
m_4 : 4. örnek ortalama momenti,

m_2 : örnek varyansı,

\bar{x} : örnek ortalaması,

n : örnek boyutu.

Ancak G_2 genel olarak yanlıdır. Normal dağılımlarda ise yansızdır.



Şekil 4. 2 Basıklık dağılım fonksiyonu.

3.2.3. Korelasyon

Korelasyon, olasılık kuramı ve istatistikte iki rastlantı değişkeni arasındaki doğrusal ilişkinin yönünü ve gücünü belirtir. Genel istatistiksel kullanımda korelasyon, bağımsızlık durumundan ne kadar uzaklaşıldığını gösterir. Farklı durumlar için farklı korelasyon katsayıları geliştirilmiştir. Bunlardan en iyi bilineni Pearson çarpım-moment korelasyon katsayısıdır. İki değişkenin kovaryansının, yine bu değişkenlerin standart sapmalarının çarpımına bölünmesiyle elde edilir.

Korelasyon katsayısı, bağımsız değişkenler arasındaki ilişkinin yönü ve büyüklüğünü belirten katsayıdır. Bu katsayı, (-1) ile (+1) arasında bir değer alır. Pozitif değerler direkt yönlü doğrusal ilişkiyi; negatif değerler ise ters yönlü bir doğrusal ilişkiyi belirtir. Korelasyon katsayısı 0 ise söz konusu değişkenler arasında doğrusal bir ilişki yoktur.

Matematikte beklenen değerler μ_X ve μ_Y , standart sapmaları σ_x ve σ_Y olan iki bağımsız değişken X ve Y arasındaki Pearson'un çarpım-moment korelasyon katsayısı ($\rho_{X,Y}$), şu şekilde tanımlanır:

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_x \sigma_Y} = \frac{E((X - \mu_X)(Y - \mu_Y))}{\sigma_x \sigma_Y}$$

E değişkenin matematiksel beklenen değerini, cov ise kovaryansı ifade eder,

$\mu_X = E(X)$ olduğunda, $\sigma_x^2 = E(X^2) - E^2(X)$ ve

Y , için de aynısı geçerli olduğundan, şu ifadeyi yazabiliriz:

$$\rho_{X,Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}}$$

Korelasyon, yalnızca standart hataların ikisi de sonlu ve sıfırdan farklı ise, tanımlıdır. Korelasyon katsayısının mutlak değeri olarak 1'i geçemeyeceği ise Cauchy-Schwarz eşitliğinin doğal bir sonucudur.

Tam bir artan doğrusal ilişkinin varlığı halinde korelasyon katsayısı 1 değerini alır, tam bir azalan ilişkinin varlığı halinde ise korelasyon katsayısı -1 değerini alır. Katsayının alabileceği diğer tüm değerler ise ilişkinin doğrusallığına bağlı olarak bu iki değer arasında olacaktır. Katsayı +1'e veya -1'e ne kadar yakınsa ilişkinin doğrusallığı o kadar güçlüdür (Haralick ve Shanmugam, 1973).

Değişkenler istatistiksel olarak bağımsız ise korelasyon 0'dır fakat bunun tersi doğru değildir, çünkü korelasyon katsayısı yalnızca doğrusal olan ilişkiyi belirler.

KORELASYON DAĞILIMI



Şekil 4.3 Korelasyon fonksiyon dağılımı.

Sonuç olarak korelasyon eşitliği aşağıdaki gibi tanımlanmaktadır.

$$Corr = \frac{\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i \times j\} \times P(i, j) - \mu_i \times \mu_j}{\sigma_x \times \sigma_y}$$

Standart sapma ve ortalama değerleri denklemde yerine konulduğunda;

$$Corr = \frac{W}{\sigma_i \sigma_j} \sum_{n=1}^{N-\Delta_y} \sum_{m=1}^{M-\Delta_x} (f(m, n) - \mu_i)(f(m + \Delta_x, n + \Delta_y) - \mu_j)$$

Bu arada

$$\mu_i = W \sum_{n=1}^{N-\Delta_y} \sum_{m=1}^{M-\Delta_x} f(m, n)$$

$$\mu_j = W \sum_{n=1}^{N-\Delta_y} \sum_{m=1}^{M-\Delta_x} f(m + \Delta_x, n + \Delta_y)$$

$$\sigma_i^2 = R \left[T \sum_{n=1}^{N-\Delta_y} \sum_{m=1}^{M-\Delta_x} f(m, n)^2 - \left(\sum_{n=1}^{N-\Delta_y} \sum_{m=1}^{M-\Delta_x} f(m, n) \right)^2 \right]$$

$$\sigma_j^2 = R \left[T \sum_{n=1}^{N-\Delta_y} \sum_{m=1}^{M-\Delta_x} f(m + \Delta_x, n + \Delta_y)^2 - \left(\sum_{n=1}^{N-\Delta_y} \sum_{m=1}^{M-\Delta_x} f(m + \Delta_x, n + \Delta_y) \right)^2 \right]$$

$$T = (M - \Delta_x)(N - \Delta_y)$$

$$R = W \frac{1}{T(T - 1)}$$

3.2.4. Enerji

Enerji ya da açısıl ikinci moment benzerliği olarak da literatürde ifade edilmektedir. Bir çeşit görüntü homojenliği ölçüleri arasındadır. Görüntü ne kadar homojen olursa değer o ölçüde büyümektedir. G normalize edilmiş GLCM matrisi olarak tanımlanırsa enerji aşağıdaki gibi ifade edilmektedir (Albregtsen, 2008).

$$Eng = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} G^2(i, j)$$

3.2.5. Grup Çıkıntısı (Cluster Prominence)

Grup çıkıntısı görüntü üzerindeki bulunan gruplar için istatistiksel bir çıkarımdır. Matematiksel olarak aşağıdaki gibi ifade edilmektedir. Orijinal grup çıkıntısı eşitliğinden yola çıkılmaktadır (Albregtsen, 2008).

$$Cprom = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i + j - \mu_i - \mu_j\}^4 \times P(i, j | d)$$

Burada

$$\mu_i = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i, j | d) \quad \mu_j = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} j \times P(i, j | d)$$

Denklemin g görüntüsü için yeniden düzenlemek gerekirse

$$Cprom = W \sum_{n=1}^{N-\Delta_y} \sum_{m=1}^{M-\Delta_x} (g(m, n) - \mu(g)_{i+j})^4$$

Burada

$$\mu(g)_{i+j} = W \sum_{n=1}^{N-\Delta_y} \sum_{m=1}^{M-\Delta_x} g(m, n)$$
$$g(m, n) = f(m, n) + f(m + \Delta_x, n + \Delta_y)$$

3.2.6. Yerel Homojenlik (Ters Fark Momenti-Inverse Difference Moment)

Görüntünün homojenlik bilgisini istatistiksel olarak elde etmek için kullanılır. Ağırlık faktörü $(1 + (i - j)^2)^{-1}$ katkıları homojen olmayan bölgelerden küçük olacaktır ($i \neq j$). Sonuç olarak homojen olmayan görüntüler için düşük yerel homojenlik değeri ve homojen görüntüler için kıyasla daha yüksek değerlidir. Yerel homojenliğin istatistiksel ifadesi incelenirse:

$$LHom = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1 + (i - j)^2} P(i, j | \Delta_x, \Delta_y)$$
$$LHom = W \sum_{n=1}^{N-\Delta_y} \sum_{m=1}^{M-\Delta_x} g(m, n)$$

Burada

$$g(m, n) = \frac{W}{1 + (f(m, n) - f(m + \Delta_x, n + \Delta_y))^2}$$

Tüm bu bilgiler ışığında çalışmada tasarlanan algoritma sonucunda elde edilen istatistiksel değerler, çekirdek sağrı regresyon sınıflandırıcısıyla sınıflandırılmıştır. Veri kümesi içinden hastalıklı ve sağlıklı hücrelerin doğru tanınması için kullanılmıştır. 260 hücre görüntüsünden oluşan veri kümesi için çeşitli performans analizleri yapılmıştır (Albregtsen, 2008).

4. PERFORMANS ANALİZİ

Bu çalışmada oluşturulan morfolojik ön işlemlerin ardından GLCM yöntemi ile öznelik çıkarımı gerçekleştirilmektedir. Elde edilen istatistiksel değerler üzerinden çekirdek sağrı regresyonu ile ALL ve sağlıklı hücrelerin tanınması işlemi gerçekleştirilir. Elde edilen sonuçları değerlendirebilmek için algoritmanın performansını analiz etmek gerekmektedir. Bunun için veri kümesinden elde edilen rastgele sonuçlar aşağıda verilmektedir. Ayrıca performans kriteri olarak topluluk ortalaması ve ortalama karesel hata karekökü (root mean squared error) kullanılmaktadır.

4.1. Topluluk Ortalaması (Ensemble Average)

Topluluk ortalaması, toplam hata oranının bulunmasına yarayan bir niceliktir. Ortalama karesel hata azaldıkça, iki veri arasındaki fark azalır. Dolayısıyla hata sıfıra yaklaştıkça hücre görüntüleri daha güvenilir olarak tanınmaktadır. Topluluk ortalaması şöyle ifade edilir:

$$x(t_1) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N x^{(i)}(t_1)^2$$

4.2. Ortalama Karesel Hataların Karekökü (Root Mean Squared Error)

Ortalama karesel hata karekökü, ölçüm değerleri ile model tahminleri arasındaki hata oranını belirlemek amacıyla kullanılmakta ve bu kriter değerinin sıfıra yaklaşması modelin tahmin yeteneğinin artması anlamına gelmektedir. Ortalama karesel hataların karekökü aşağıdaki gibi hesaplanmaktadır.

$$OKHK = \sqrt{\frac{\sum_{i=0}^n (y_{orj,i}, y_{son,i})^2}{n}}$$

Burada

$(y_{orj,i}, y_{son,i})$ model tahminlerini, n gözlem sayısını ifade etmektedir.

5. SONUÇLAR

Akut lenfosit lösemili hücrelerin bir takım morfolojik ön işlemler ardından üç gri seviye eş oluşum matrisleriyle (GLCM) öznitelikleri çıkarılmıştır. Ardından çekirdek sağrı regresyonu ile örüntü tanıma yapıları kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Algoritmanın tanı sonuçları ve laboratuvar elde edilen sonuçlar Tablo 5.1 ve Tablo 5.2’de sunulmaktadır.

Tablo 5. 1. Laboratuvar sonuçları ile çekirdek sağrı regresyonu sınıflandırması sonuçlarının karşılaştırılması.

Örnek Hücre Sayısı	Laboratuvar Deney Sonuçları	Çekirdek Sağrı Regresyonu Tanısı
1	ALL	ALL
2		ALL
3		ALL
4		ALL
5		ALL
6		Sağlıklı
7		ALL
8		ALL
9		ALL
10		ALL
11		Sağlıklı
12		ALL
13		ALL
14		ALL
15		ALL
16		ALL
17		ALL
18		ALL
19		ALL
20		ALL

Program tablolarında sunulmak için otomatik olarak rastgele 10 tane örnek seçerek ve 6 kez çalıştırılmıştır. Tüm hücrelerin analiz edilmesiyle elde edilen değerlerde topluluk ortalaması ve ortalama karesel hata karekökü performans kriteri olarak hesaplanmıştır.

Tablo 5. 2. Algoritmanın 6 kez çalıştırılarak elde edilen performans.

Run #1	Başarım (%)	Run #2	Başarım (%)	Run #3	Başarım (%)
1	93.17	1	93.84	1	97.43
2	95.89	2	93.84	2	99.28
3	98.32	3	95.89	3	98.97
4	96.95	4	97.78	4	97.66
5	94.76	5	92.86	5	95.97
6	95.89	6	93.17	6	97.43
7	95.89	7	96.95	7	93.84
8	92.86	8	97.43	8	95.97
9	97.43	9	93.17	9	98.43
10	97.66	10	92.82	10	92.84
Ortalama	95.88	Ortalama	94.74	Ortalama	96.78

Run #4	Başarım (%)	Run #5	Başarım (%)	Run #6	Başarım (%)
1	93.84	1	94.80	1	97.78
2	95.78	2	99.28	2	92.82
3	92.82	3	98.97	3	95.97
4	98.76	4	95.89	4	98.32
5	96.95	5	93.17	5	94.85
6	93.84	6	97.43	6	96.28
7	95.97	7	97.66	7	96.95
8	94.76	8	95.89	8	97.66
9	98.79	9	92.86	9	94.89
10	98.32	10	94.85	10	97.43
Ortalama	95.98	Ortalama	96.08	Ortalama	96.20

Algoritmanın rastgele veriyle çalıştırılmasının ardından elde edilen ortalama başarımlar %94.75 ile %96.78 arasında değişmektedir. Bu işlemler algoritma içinde 100 kez rastgele olarak gerçekleştirildikten sonra topluluk ortalaması %96.43 olarak hesaplanmıştır.

ALL hastalığı için erken tanının tedavide olumlu sonuç erdiği bilinmektedir. Elde edilen sonuç, bu tezde önerilen yöntemin, yani morfolojik ön işlemlerin ardından GLCM ile öznelik çıkarımının ve bu özneliklerin çekirdek sağrı regresyonu kullanılarak akut lenfosit lösemili hücrelere tanı konması problemini başarıyla çözen bir yöntem olduğunu göstermektedir.

6. KAYNAKLAR

Albregtsen, F. (2008). Statistical Texture Measures Computed from Gray Level Cooccurrence Matrices. *University of Oslo, Department of Informatics, Image Processing Laboratory*. Retrieved: 10 May 2011, http://www.uio.no/studier/emner/matnat/ifi/INF4300/h08/undervisningsmateriale/glc_m.pdf

American Cancer Society. (2010). Retrieved: 7 Kasım 2010, <http://www.cancer.org/Cancer/LeukemiaAcuteLymphocyticALLinAdults/DetailedGuide/leukemia-acute-lymphocytic-classified>

Anand, P., Kunnumakara, A. B., Sundaram, C., Herikumar, K. B., Tharakan, S. T., Lai, O. S., Sung, B. and Aggarwal, B. B. (2008). Cancer is a Preventable Disease That Requires Major Lifestyle Changes, *Pharmaceutical Research*, No: 9, Vol: 25. 2097-2116. Retrieved: November 18, 2010, Springer PubMed Central Journal List.

Asadi, M. R., Vahedi, A. and Amindavar, H., (2006). Leukemia Cell Recognition with Zernike Moments of Holographic Images, *Proceedings of the Seventh Nordic Signal Processing Symposium*, 214–217. Retrieved: December 4, 2010, IEEE Xplore Digital Library.

Beasley, D., Bull, D. R. and Martin R. R. (1993 a). Part 1, Fundamentals. *An Overview of Genetic Algorithms* 15 (2), 58-69.

Beasley, D., Bull, D. R. and Martin R. R. (1993 b). Part 2, Research Topics, *An Overview of Genetic Algorithms* 15 (4), 170-181.

Bremner, D., Demaine, E., Erickson, J., Iacono, J., Langerman, S., Morin, P. and Toussaint, G. (2005). Output-Sensitive Algorithms for Computing Nearest-Neighbor Decision Boundaries, *Discrete and Computational Geometry*, Vol: 33 (4): 593-604. Retrieved: December 4, 2010, ACM Digital Library.

Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition, *Bell Laboratories, Lucent Technologies Data Mining and Knowledge Discovery*, 121-167. Retrieved: February 7, 2011.

Cancer Research UK. (2010). Retrieved: January 9, 2010. <http://www.cancerresearchuk/>

Carlo, M. and Croce, M. D. (2008). Oncogenes and Cancer. *The New England Journal of Medicine*, Vol: 358, 502-511. Retrieved: January 9, 2010. <http://www.nejm.org/doi/full/10.1056/NEJMra072367>

Chun-Lin, L. (2010). A Tutorial of the Wavelet Transforms. *National Taiwan University, Department of Electrical Engineering, National Taiwan University*, Retrieved: February 10, 2011. <http://disp.ee.ntu.edu.tw/tutorial/WaveletTutorial.pdf>

Chunni, D., Qingli, L. and Jingao, L. (2009). Band Selection for Biomedical Hyperspectral Data Studies Using Genetic Algorithms, *3rd International Conference on Bioinformatics and Biomedical Engineering (ICBBE)*, 1-4. Retrieved: November 11, 2010, IEEE Xplore Digital Library.

Cortes, C. and Vapnik, V. (1995). Support Vector Networks, *SpringerLink, Machine Learning*, No: 3, Vol: 20. 273-297. Retrieved: February 3, 2011, <http://www.springerlink.com/content/w08253ul7m3780v8/fulltext.pdf>

Çapar, A., Taşdemir, K., Kılıç, Ö. ve Gökmen, M. (2002). Türkçe Elyazısı Tanıma Sistemlerinde Öznitelik Çıkarma ve Sınıflandırma Yöntemlerinin Karşılaştırması, *Conference on Union Radio Science International (URSI)*. Erişim tarihi: 30 Ekim 2010, <http://www.ursi.org.tr/2002-1.Ulusal%20Kongre/ursicd1/C14.pdf>

Else, M., Ruchlemer, R., Osuji, N., Del Giudice, I., Matutes, E., Woodman, A., Wotherspoon, A., Swansbury, J., Deorden, C. and Catovsky D. (2005). Long Remission in Hairy Cell Leukemia with Purine Analogs. *American Cancer Society, Cancer* Vol: 102, Issue 11, pp. 2442-2448. Retrieved: November 7, 2010, <http://onlinelibrary.wiley.com/doi/10.1002/cncr.21447/full>

Gavlasova, A., Prochazka, A., Pozivil, J. and Vysata, O. (2008). Functional Transforms in MR Image Segmentation, *3rd International Symposium on Communications Control and Signal Processing*, 1357-1360. Retrieved: November 13, 2010, IEEE Xplore Digital Library.

Gonzales, R. C., Woods, R. E. and Eddins, A. L. (2009). *Digital Image Processing Using MATLAB*. (2nd Ed). USA, New Jersey: Gatesmark Publishing.

Haralick, R., M., ve Shanmugam, K., (1973). Computer Classification of Reservoir Sandstones. *IEEE Transaction on Geo. Eng., Vol. GE-11*, pp. 171-177. Retrieved: 10 May 2011, IEEE Xplore Digital Library.

Holland, J. H. (1992). *Adaption in Natural and Artificiation Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. USA, Massachusetts: MIT Press.

<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.117.3731>

Kernel Methods Artcle. (2011). Eriřim tarihi: 14 Mayıs 2011, <http://onlineprediction.net/?n=Main.KernelMethods#toc3>

Khotanzad, A. and Hong, Y. H. (1990). Invariant Image Recognition by Zernike Moments, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, No. 5., Vol: 12. 489-497. Retrieved: February 15, IEEE Xplore Digital Library.

Li, L., Cao, G., Shi, J. and Zhang, X. (2010). Detecting Immature Precursor Cells in Pathological Images of Bone Marrow Based on Morphology, *Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, 2190-2194. Retrieved: December 14, 2010, IEEE Xplore Digital Library.

Liu, D. and Yu, J. (2009). Otsu Method and K-means, *Ninth International Conference on Hybrid Intelligent Systems (HIS)*, Vol:1.344-349. Retrieved: December 4, 2010, IEEE Xplore Digital Library.

Mercier, G. and Lennon, M. (2003). Support Vector Machines for Hyperspectral Image Classification with Spectral-Based Kernels. *Proceedings of the IEEE*

International Geoscience and Remote Sensing Symposium, Vol: 1, 288-290.
Retrieved: February 21, 2011, IEEE Xplore Digital Library.

Mokhtar, N. R., Harun, N. H., Mashor, M. Y., Roseline ,H., Mustafa, N., Abdollah ,
R., Adilah, H. and Mohd Nasir, N. H., (2009). Image Enhancement Techniques
Using Local, Global, Bright, Dark and Partial Contrast Stretching For Acute
Leukemia Images, *Proceedings of the World Congress on Engineering*, Vol I.
Retrieved: November 4, 2010,
http://www.iaeng.org/publication/WCE2009/WCE2009_pp807-812.pdf

Morkiewicz, T., Osawski, S., Marinska, B. and Moszcynski, L. (2005). Automatic
Recognition of the Blood Cells of Myelogenous Leukemia Using SVM, *Proceedings
of the International Joint Conference on Neural Networks*, Vol: 4. 2496-2501.
Retrieved: December 4, 2010, IEEE Xplore Digital Library.

National Cancer Institute. (2010). Retrieved: November 7, 2010,
<http://seer.cancer.gov/statfacts/html/clyl.html>

Nor Hazlyna H., Mashor, M. Y., Mokhtor, N. R., Aimi Salihah, A. A. N., Hassan, R.,
Raof, R. A. A. and Osman, M. K. (2010). Comparison of Acute Leukemia Image
Segmentation Using HSI and RGB Color Space, *10th International Conference on
Information Science, Signal Processing and Their Applications (ISSPA)*, 749-752.
Retrieved: December 11, 2010, IEEE Xplore Digital Library.

Otsu, N. (1979). A Threshold Selection Method From Gray-Level Histogram, *IEEE
Transaction on Systems, Man and Cybernetics*, Vol: 9 (1), 62. Retrieved: December
6, 2010, IEEE Xplore Digital Library.

Papoulis, A. ve Pillai, U., S., (2002). *Probability, Random Variables and Stochastic
Processes*. USA, McGraw Hill Publishing.

Petrosian, A. (2002). New Classes of Hybrid Hadamard-Wavelet Transforms for
Signal-Image Processing. *Proceedings of Second Joint Engineering in Medicine and
Biology, 24th Annual Conference and the Annual Fall Meeting of the Biomedical*

Engineering Society EMBS/BMES Conference, 153-154, Retrieved: February 10, 2011, IEEE Xplore Digital Library.

Prasad, B., Choi, J-S. I. and Badawy, W. (2006). A High Throughput Screening Algorithm for Leukemia Cells. *Canadian Conference on Electrical and Computer Engineering*, 2094–2097. Retrieved: February 17, 2011, IEEE Xplore Digital Library.

Qader, H. A., Ramli, A. R. and Al-Haddad, S. (2007). Fingerprint Recognition Using Zernike Moments, *The International Arab Journal of Information Technology*, No. 4., Vol: 4. 372-376. Retrieved: February 5, 2011, <http://www.ccis2k.org/iajit/PDF/vol.4.no.4/12-Qader.pdf>

Quiti, T. C., Vivanco, P. A. and Agustin, G. F. (2007). Classification of Motion-Blurred Images Using Zernike and Wavelet-Fourier Moments, *Image and Signal Processing for Remote Sensing XIII, Proceedings of the SPIE*, Vol: 6748, Retrieved: February 5, 2011, SPIE Digital Library.

Scotti, F. (2005). Automatic Morphological Analysis for Acute Leukemia Identification in Peripheral Blood Microscope Image, *IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA)*, 96-101. Retrieved: November 5, 2010, IEEE Xplore Digital Library.

Sewell, M. (2009). *Kernel Methods*. University College London, Department of Computer Science. Erişim tarihi: 2 Mayıs 2010, <http://www.svms.org/kernels/kernel-methods.pdf>

Terrell, D. G. and Scott, D. W., (1992). Variable Kernel Density Estimation, *Project Euclid, Annals of Statistics*, Vol: 20. 1236-1265. Retrieved: October 30, 2010, http://projecteuclid.org/DPubS/Repository/1.0/Disseminate?view=body&id=pdf_1&handle=euclid.aos/1176348768

The Leukemia & Lymphoma Society. (2010). Retrieved: November 7, 2010, http://www.leukemia-lymphoma.org/all_page?item_id=7026

Theodoridis, S. Ve Koutroumbas, K. (2009). *Pattern Recognition Fourth Edition*. USA: Elsevier Inc.

Vogelstein, B. and Kinzler, W. K. (2002). *The Genetic Basis of Human Cancer*. USA: McGraw Hill Professional.

Welling, M. (2009). *Kernel Ridge Regression*. University of Toronto, Department of Computer Science. Eriřim tarihi: 2 Mayıs 2011, http://www.ics.uci.edu/~welling/classnotes/papers_class/Kernel-Ridge.pdf

World Health Organization. (2010), Retrieved: November 7, 2010, <http://www.who.int/publications/en/>

Yenerel, M. N. (2009). *Tanısal Hematoloji: Çok Fazla Test Yapmayın*, İstanbul: İstanbul Üniversitesi, İstanbul Tıp Fakültesi, İç Hastalıkları Anabilim Dalı, Hematoloji Bilim Dalı. Eriřim tarihi: 6 Kasım 2010, http://www.thd.org.tr/doc/kurs_pdf/5_IBK_03.pdf

7. EKLER

EK 1: Algoritmanın tüm adımlarının çağırıldığı genel program.

```
tic; % Programın zaman sayacı
clc;
clear all;
close all;

disp('Program çalıştırılıyor...'); % komut penceresine mesaj
gönderir

ProjectInit;

% Görüntüyü okuma
for j = 1:NumImages
% for j = 1:7
    originalImage = im{j};
%     size(originalImage)
    RedImage = originalImage(:, :, 1);
    GreenImage = originalImage(:, :, 2);
    BlueImage = originalImage(:, :, 3);

%% Filtreler: Gauss ve Top-Hat filtresi ekleme

    % Görüntülere Gauss filtresi uygula

    RedImage = imfilter(RedImage, H); % kırmızı banta filtre uygula
    GreenImage = imfilter(GreenImage, H); % yeşil banta filtre
    uygula
    BlueImage = imfilter(BlueImage, H); % mavi banta filtre uygula

    % Görüntülere Top-Hat filtresi uygula

    tophatFilteredRed = imtophat(RedImage, seRed);
    tophatFilteredGreen = imtophat(GreenImage, seGreen);
    tophatFilteredBlue = imtophat(BlueImage, seBlue);

    % Kontrast ayarlarının yapılması
    contrastAdjustedRed = imadjust(tophatFilteredRed);
    contrastAdjustedGreen = imadjust(tophatFilteredGreen);
    contrastAdjustedBlue = imadjust(tophatFilteredBlue);

    RedImage = contrastAdjustedRed;
    GreenImage = contrastAdjustedGreen;
    BlueImage = contrastAdjustedBlue;
```

```

%% İkili görüntü için eşikleme

    % Kırmızı bant için eşikleme

    RedBinaryImage = RedImage < RedThresholdValue; % Koyu kısımlar
izole edilir
    % Arkaplan leke piksellerinden kurtulmak için "hole fill" ile
    % doldurma işlemi
    RedBinaryImage = imfill(RedBinaryImage, 'holes');
    RedLabeledImage = bwlabel(RedBinaryImage, 8);

    % Yeşil bant için eşikleme

    GreenBinaryImage = GreenImage < GreenThresholdValue; % Koyu
kısımlar izole edilir
    % Arkaplan leke piksellerinden kurtulmak için "hole fill" ile
    % doldurma işlemi
    GreenBinaryImage = imfill(GreenBinaryImage, 'holes');
    GreenLabeledImage = bwlabel(GreenBinaryImage, 8);

    % Mavi bant için eşikleme

    BlueBinaryImage = (BlueImage < BlueThresholdValue) & (BlueImage
> 120); % İlgili kısım izole ediliyor
    % Arkaplan leke piksellerinden kurtulmak için "hole fill" ile
    % doldurma işlemi
    BlueBinaryImage = imfill(BlueBinaryImage, 'holes');
    BlueLabeledImage = bwlabel(BlueBinaryImage, 8);

%% Kırmızı bandın hücre özelliklerini al
    RedBlobMeasurements = regionprops(RedLabeledImage, 'all');
    RedNumberOfBlobs = size(RedBlobMeasurements, 1);

    % bwboundaries() her hücre / sütun görüntü bir nesnenin
koordinatları
    % satır içeren bir hücre dizisi döndürür.
    % bwboundaries tarafından döndürülen koordinatları kullanarak
    % orijinal gri tonlarında bir fotoğraf tüm lekeler
sınırları.Arsa
    % bwboundaries tarafından döndürülen koordinatları kullanarak
orijinal
    % gri tonlarında bir fotoğraf tüm lekeler sınırları çizdirir.
    RedBoundaries = bwboundaries(RedImage);
    RedNumberOfBoundaries = size(RedBoundaries);

    RedBlobECD = zeros(1, RedNumberOfBlobs);
%% Yeşil bandın hücre özelliklerini al

    GreenBlobMeasurements = regionprops(GreenLabeledImage, 'all');
    GreenNumberOfBlobs = size(GreenBlobMeasurements, 1);

    % bwboundaries() her hücre / sütun görüntü bir nesnenin
koordinatları
    % satır içeren bir hücre dizisi döndürür.

```

```

    % bwboundaries tarafından döndürülen koordinatları kullanarak
    % orijinal gri tonlarında bir fotoğraf tüm lekeler
sınırları.Arsa
    % bwboundaries tarafından döndürülen koordinatları kullanarak
orijinal
    % gri tonlarında bir fotoğraf tüm lekeler sınırları çizdirir.
GreenBoundaries = bwboundaries(GreenImage);
GreenNumberOfBoundaries = size(GreenBoundaries);

GreenBlobECD = zeros(1, GreenNumberOfBlobs);
%% Mavi bandın hücre özelliklerini al

BlueBlobMeasurements = regionprops(BlueLabeledImage, 'all');
BlueNumberOfBlobs = size(BlueBlobMeasurements, 1);

    % bwboundaries() her hücre / sütun görüntü bir nesnenin
koordinatları
    % satır içeren bir hücre dizisi döndürür.
    % bwboundaries tarafından döndürülen koordinatları kullanarak
    % orijinal gri tonlarında bir fotoğraf tüm lekeler
sınırları.Arsa
    % bwboundaries tarafından döndürülen koordinatları kullanarak
orijinal
    % gri tonlarında bir fotoğraf tüm lekeler sınırları çizdirir.
BlueBoundaries = bwboundaries(BlueImage);
BlueNumberOfBoundaries = size(BlueBoundaries);

BlueBlobECD = zeros(1, BlueNumberOfBlobs);

%% İlgili bölgenin (mavi hücre) Kırmızı banttı seçilmesi

RedAllBlobIntensities = [RedBlobMeasurements.Solidity];
RedAllBlobAreas = [RedBlobMeasurements.Area];

RedAllowableIntensityIndexes = (RedAllBlobIntensities > 0.5);
RedAllowableAreaIndexes = RedAllBlobAreas > 0;
RedKeeperIndexes = find(RedAllowableIntensityIndexes &
RedAllowableAreaIndexes);

    % Aslında sadece bu kriter ile hücreler belirlenir
RedKeeperBlobsImage = ismember(RedLabeledImage,
RedKeeperIndexes);
    % Leke olarak belirlenen hücreler tekrar etiketlenir.
RedLabeledDimeImage = bwlabel(RedKeeperBlobsImage, 8);      % Her
leke etiketlenir
    % İlgili bölgenin (mavi hücre) Yeşil banttı seçilmesi

GreenAllBlobIntensities = [GreenBlobMeasurements.Solidity];
%[Will be replaced with intensity]
GreenAllBlobAreas = [GreenBlobMeasurements.Area];

GreenAllowableIntensityIndexes = (GreenAllBlobIntensities >
0.6);
GreenAllowableAreaIndexes = GreenAllBlobAreas > 0;
GreenKeeperIndexes = find(GreenAllowableIntensityIndexes &
GreenAllowableAreaIndexes);

```

```

GreenKeeperBlobsImage = ismember(GreenLabeledImage,
GreenKeeperIndexes);

GreenLabeledDimeImage = bwlabel(GreenKeeperBlobsImage, 8); %
Hücreler etiketlenir ve ölçüm yapılabilir
%% İlgili bölgenin (mavi hücre) Mavi banttan seçilmesi

BlueAllBlobSolidity = [BlueBlobMeasurements.Solidity];
BlueAllBlobCentroid = [BlueBlobMeasurements.Area]; %
Değiştirilmesi için

BlueAllowableSolidityIndexes = (BlueAllBlobSolidity > 0);
BlueAllowableCentroidIndexes = (BlueAllBlobCentroid > 5000);
BlueKeeperIndexes = find(BlueAllowableSolidityIndexes &
BlueAllowableCentroidIndexes);

BlueKeeperBlobsImage = ismember(BlueLabeledImage,
BlueKeeperIndexes);

BlueLabeledDimeImage = bwlabel(BlueKeeperBlobsImage, 8);

%% Kırmızı bant için ilgili bölgenin kesilmesi

Red_k = find(RedAllBlobAreas == max(RedAllBlobAreas));
RedBlobsBoundingBox = RedBlobMeasurements(Red_k).BoundingBox; %
Geçerli hücre pikselini alır.
RedFinalImage = imcrop(RedImage, RedBlobsBoundingBox);
%% Yeşil bant için ilgili bölgenin kesilmesi

Green_k = find(GreenAllBlobAreas == max(GreenAllBlobAreas));
GreenBlobsBoundingBox =
GreenBlobMeasurements(Green_k).BoundingBox; % Geçerli hücre
pikselini alır.
GreenFinalImage = imcrop(GreenImage, GreenBlobsBoundingBox);
%% Mavi bant için ilgili bölgenin kesilmesi

Blue_k = find(BlueAllBlobCentroid == max(BlueAllBlobCentroid));
BlueBlobsBoundingBox = BlueBlobMeasurements(Blue_k).BoundingBox;
% Geçerli hücre pikselini alır.
BlueFinalImage = imcrop(BlueImage, BlueBlobsBoundingBox);

%% Üç bant bilgisinin birleştirilmesi

[RedX RedY] = size(RedFinalImage);
[GreenX GreenY] = size(GreenFinalImage);
[BlueX BlueY] = size(BlueFinalImage);

% Maksimum satır boyutu ve maksimum sütun boyutunu al
RowArray = [RedX, GreenX, BlueX];
ColArray = [RedY, GreenY, BlueY];
RowSize = max(RowArray);
ColSize = max(ColArray);

```



```

% Aynı boyuta tüm görüntüleri yeniden boyutlandırma
RedIm = imresize(RedFinalImage, [RowSize ColSize]);
GreenIm = imresize(GreenFinalImage, [RowSize ColSize]);
BlueIm = imresize(BlueFinalImage, [RowSize ColSize]);

FinalImage = RedIm;
FinalImage(:, :, 2) = GreenIm;
FinalImage(:, :, 3) = BlueIm;

ROI_Image{j,1} = FinalImage; %Görüntünün ilgili bölgesi (SON
GÖRÜNTÜ)
ROI_Image{j,2} = Tag{j};

ROI_Image{j,1} = rgb2gray(ROI_Image{j,1});
ROI_Image{j,1} = im2double(ROI_Image{j,1});

    % Ekran çıktıları
if toc>2
    fprintf('Bölütlenen Görüntü%d\n', j);
    tic;
end
end

%% Görüntüden Özellik Çıkarımı

offsets = [0 1;-1 1;-1 0;-1 -1];

for j = 1:NumImages

    % Ekran çıktıları
if toc>1
    fprintf('Görüntünün öznitelikleri çıkarılıyor%d\n', j);
    tic;
end

    % graycomatrix Fonksiyonunun kullanılması
GLCMS = graycomatrix(ROI_Image{j,1}, 'offset', offsets);
stats_01 = graycoprops(GLCMS);

    % GLCM_features1 Fonksiyonunun kullanılması
stats_02 = GLCM_features1(GLCMS,0);

Feat_Mean = mean(mean(ROI_Image{j, 1}));
Feat_Skewness = mean(skewness(ROI_Image{j,1}));
Feat_kurtosis = mean(kurtosis(ROI_Image{j,1}));

Feat_contrast = mean(stats_01.Contrast);
Feat_Correlation = mean(stats_01.Correlation);
Feat_energy = mean(stats_01.Energy);

Feat_cprom = mean(stats_02.cprom); %Küme önemleri
Feat_hom = mean(stats_02.homom);

FeatureMatrix(j, 1) = Feat_Mean;
FeatureMatrix(j, 2) = Feat_Skewness;
FeatureMatrix(j, 3) = Feat_kurtosis;
FeatureMatrix(j, 4) = Feat_contrast;

```

```

    FeatureMatrix(j, 5) = Feat_Correlation;
    FeatureMatrix(j, 6) = Feat_energy;
    FeatureMatrix(j, 7) = Feat_cprom;
    FeatureMatrix(j, 8) = Feat_hom;

    TagArray(j) = str2num(ROI_Image{j,2});
end

%% Öğrenme ve test sınıflandırıcısı (KRR-Çekirdek Sağrı Regresyonu)

TrainSet = FeatureMatrix(1:180,:);
Trainset = normalise(TrainSet);
TrainLabels = TagArray(1:180)';

TestSet = FeatureMatrix(181:260,:);
TestSet = normalise(TestSet);
TestLabels = TagArray(181:260)';

% TrainSet = FeatureMatrix(1:5,:);
% Trainset = normalise(TrainSet);
% TrainLabels = TagArray(1:5)';
%
% TestSet = FeatureMatrix(6:7,:);
% TestSet = normalise(TestSet);
% TestLabels = TagArray(6:7)';

[E,PredictedTestLabels] = optimisekrr(TrainSet,TrainLabels,TestSet)

```

EK 2: GLCM ile öznitelik çıkarımı.

```
function [out] = GLCM_Features1(glcm, pairs)

% Features computed
% Autocorrelation: [2] (out.autoc)
% Contrast: matlab/[1,2] (out.contr)
% Correlation: matlab (out.corr)
% Correlation: [1,2] (out.corrp)
% Cluster Prominence: [2] (out.cprom)
% Cluster Shade: [2] (out.cshad)
% Dissimilarity: [2] (out.dissi)
% Energy: matlab / [1,2] (out.energ)
% Entropy: [2] (out.entro)
% Homogeneity: matlab (out.homom)
% Homogeneity: [2] (out.homop)
% Maximum probability: [2] (out.maxpr)
% Sum of squares: Variance [1] (out.sosvh)
% Sum average [1] (out.savgh)
% Sum variance [1] (out.svarh)
% Sum entropy [1] (out.senth)
% Difference variance [1] (out.dvarh)
% Difference entropy [1] (out.denth)
% Information measure of correlation1 [1] (out.inflh)
% Information measure of correlation2 [1] (out.inf2h)
% Inverse difference (INV) is homom [3] (out.homom)
% Inverse difference normalized (INN) [3] (out.indnc)
% Inverse difference moment normalized [3] (out.idmnc)
%
if ((nargin > 2) || (nargin == 0))
    error('Çok fazla veya çok az sayıda girdi argüman. GLCM ve çift girin.');
```

```
elseif (nargin == 2)
    if ((size(glcm,1) <= 1) || (size(glcm,2) <= 1))
        error('İki ya da Üç Boyutlu Matris Olmalı.');
```

```
    elseif (size(glcm,1) ~= size(glcm,2))
        error('Each GLCM should be square with NumLevels rows and NumLevels cols');
```

```
    end
elseif (nargin == 1)
    pairs = 0;
    if ((size(glcm,1) <= 1) || (size(glcm,2) <= 1))
        error('İki ya da Üç Boyutlu Matris Olmalı.');
```

```
    elseif (size(glcm,1) ~= size(glcm,2))
        error('Each GLCM should be square with NumLevels rows and NumLevels cols');
```

```
    end
end

format long e
if (pairs == 1)
    newn = 1
```

```

    for nglcm = 1:2:size(glcm,3)
        glcm(:,:,newn) = glcm(:,:,nglcm) + glcm(:,:,nglcm+1);

        newn = newn + 1;
    end
elseif (pairs == 0)
    glcm = glcm;
end

size_glcm_1 = size(glcm,1);
size_glcm_2 = size(glcm,2);
size_glcm_3 = size(glcm,3);

out.autoc = zeros(1,size_glcm_3); % Autocorrelation: [2]
out.contr = zeros(1,size_glcm_3); % Contrast: matlab/[1,2]
out.corr = zeros(1,size_glcm_3); % Correlation: matlab
out.corrp = zeros(1,size_glcm_3); % Correlation: [1,2]
out.cprom = zeros(1,size_glcm_3); % Cluster Prominence: [2]
out.cshad = zeros(1,size_glcm_3); % Cluster Shade: [2]
out.dissi = zeros(1,size_glcm_3); % Dissimilarity: [2]
out.energ = zeros(1,size_glcm_3); % Energy: matlab / [1,2]
out.entrop = zeros(1,size_glcm_3); % Entropy: [2]
out.homom = zeros(1,size_glcm_3); % Homogeneity: matlab
out.homop = zeros(1,size_glcm_3); % Homogeneity: [2]
out.maxpr = zeros(1,size_glcm_3); % Maximum probability: [2]

out.sosvh = zeros(1,size_glcm_3); % Sum of squares: Variance [1]
out.savgh = zeros(1,size_glcm_3); % Sum average [1]
out.svarh = zeros(1,size_glcm_3); % Sum variance [1]
out.senth = zeros(1,size_glcm_3); % Sum entropy [1]
out.dvarh = zeros(1,size_glcm_3); % Difference variance [4]
%out.dvarh2 = zeros(1,size_glcm_3); % Difference variance [1]
out.denth = zeros(1,size_glcm_3); % Difference entropy [1]
out.inflh = zeros(1,size_glcm_3); % Information measure of
correlation1 [1]
out.inf2h = zeros(1,size_glcm_3);
out.indnc = zeros(1,size_glcm_3);
out.idmnc = zeros(1,size_glcm_3);

glcm_sum = zeros(size_glcm_3,1);
glcm_mean = zeros(size_glcm_3,1);
glcm_var = zeros(size_glcm_3,1);

u_x = zeros(size_glcm_3,1);
u_y = zeros(size_glcm_3,1);
s_x = zeros(size_glcm_3,1);
s_y = zeros(size_glcm_3,1);

p_x = zeros(size_glcm_1,size_glcm_3);
p_y = zeros(size_glcm_2,size_glcm_3); p_xplusy =
zeros((size_glcm_1*2 - 1),size_glcm_3);
p_xminusy = zeros((size_glcm_1),size_glcm_3);
hxy = zeros(size_glcm_3,1);
hxy1 = zeros(size_glcm_3,1);
hx = zeros(size_glcm_3,1);
hy = zeros(size_glcm_3,1);
hxy2 = zeros(size_glcm_3,1);

```

```

for k = 1:size_glcm_3
    glcm_sum(k) = sum(sum(glcm(:,:,k)));
    glcm(:,:,k) = glcm(:,:,k)./glcm_sum(k);
    glcm_mean(k) = mean2(glcm(:,:,k));
    glcm_var(k) = (std2(glcm(:,:,k)))^2;

    for i = 1:size_glcm_1

        for j = 1:size_glcm_2

            out.contr(k) = out.contr(k) + (abs(i -
j))^2.*glcm(i,j,k);
            out.dissi(k) = out.dissi(k) + (abs(i - j)*glcm(i,j,k));
            out.energ(k) = out.energ(k) + (glcm(i,j,k).^2);
            out.entro(k) = out.entro(k) -
(glcm(i,j,k)*log(glcm(i,j,k) + eps));
            out.homom(k) = out.homom(k) + (glcm(i,j,k)/(1 + abs(i-
j) ));
            out.homop(k) = out.homop(k) + (glcm(i,j,k)/(1 + (i -
j)^2));
            out.sosvh(k) = out.sosvh(k) + glcm(i,j,k)*((i -
glcm_mean(k))^2);
            out.indnc(k) = out.indnc(k) + (glcm(i,j,k)/(1 + (abs(i-
j)/size_glcm_1) ));
            out.idmnc(k) = out.idmnc(k) + (glcm(i,j,k)/(1 + ((i -
j)/size_glcm_1)^2));
            u_x(k) = u_x(k) + (i)*glcm(i,j,k);
            u_y(k) = u_y(k) + (j)*glcm(i,j,k);
        end

    end

    out.maxpr(k) = max(max(glcm(:,:,k)));
end

for k = 1:size_glcm_3

    for i = 1:size_glcm_1

        for j = 1:size_glcm_2
            p_x(i,k) = p_x(i,k) + glcm(i,j,k);
            p_y(i,k) = p_y(i,k) + glcm(j,i,k);
            if
(ismember((i + j), [2:2*size_glcm_1]))
                p_xplusy((i+j)-1,k) = p_xplusy((i+j)-1,k) +
glcm(i,j,k);
            end
            if (ismember(abs(i-j), [0:(size_glcm_1-1)]))
                p_xminusy((abs(i-j))+1,k) = p_xminusy((abs(i-
j))+1,k) +...
                glcm(i,j,k);
            end
        end
    end
end

end

for k = 1:(size_glcm_3)

```

```

        for i = 1:(2*(size_glcm_1)-1)
            out.savgh(k) = out.savgh(k) + (i+1)*p_xplusy(i,k);
            out.senth(k) = out.senth(k) -
                (p_xplusy(i,k)*log(p_xplusy(i,k) + eps));
        end

    end
    for k = 1:(size_glcm_3)

        for i = 1:(2*(size_glcm_1)-1)
            out.svarh(k) = out.svarh(k) + (((i+1) -
                out.senth(k))^2)*p_xplusy(i,k);
        end

    end
    for k = 1:size_glcm_3
        for i = 0:(size_glcm_1-1)
            out.denth(k) = out.denth(k) -
                (p_xminusy(i+1,k)*log(p_xminusy(i+1,k) + eps));
            out.dvarh(k) = out.dvarh(k) + (i^2)*p_xminusy(i+1,k);
        end
    end

    for k = 1:size_glcm_3
        hxy(k) = out.entro(k);
        for i = 1:size_glcm_1

            for j = 1:size_glcm_2
                hxy1(k) = hxy1(k) - (glcm(i,j,k)*log(p_x(i,k)*p_y(j,k) +
                    eps));
                hxy2(k) = hxy2(k) -
                    (p_x(i,k)*p_y(j,k)*log(p_x(i,k)*p_y(j,k) + eps));
            end
            hx(k) = hx(k) - (p_x(i,k)*log(p_x(i,k) + eps));
            hy(k) = hy(k) - (p_y(i,k)*log(p_y(i,k) + eps));
        end
        out.inf1h(k) = ( hxy(k) - hxy1(k) ) / ( max([hx(k),hy(k)] ) );
        out.inf2h(k) = ( 1 - exp( -2*( hxy2(k) - hxy(k) ) ) ) ^0.5;
    end
    corm = zeros(size_glcm_3,1);
    corp = zeros(size_glcm_3,1);
    for k = 1:size_glcm_3
        for i = 1:size_glcm_1
            for j = 1:size_glcm_2
                s_x(k) = s_x(k) + (((i) - u_x(k))^2)*glcm(i,j,k);
                s_y(k) = s_y(k) + (((j) - u_y(k))^2)*glcm(i,j,k);
                corp(k) = corp(k) + ((i)*(j))*glcm(i,j,k);
                corm(k) = corm(k) + (((i) - u_x(k))*((j) -
                    u_y(k))*glcm(i,j,k));
                out.cprom(k) = out.cprom(k) + (((i + j - u_x(k) -
                    u_y(k))^4)*...
                    glcm(i,j,k));
                out.cshad(k) = out.cshad(k) + (((i + j - u_x(k) -
                    u_y(k))^3)*...
                    glcm(i,j,k));
            end
        end
        s_x(k) = s_x(k) ^ 0.5;
        s_y(k) = s_y(k) ^ 0.5;
        out.autoc(k) = corp(k);
    end
end

```

```
out.corrp(k) = (corp(k) - u_x(k)*u_y(k)) / (s_x(k)*s_y(k));  
out.corm(k) = corm(k) / (s_x(k)*s_y(k));
```

EK 3: Çekirdek sağrı regresyonu.

```
function [yy,rss,prss] = krr(X,t,Q,s,g)

p = length(t);
K = exp(-dist2(X,X)/(2*s));

Ki =pinv( K'*K + g*eye(size(K)))* K';

a = Ki*t;

Kq = exp(-dist2(Q,X)/(2*s));

yy = 2*(Kq * a > 0) - 1;

r = t-K*a;
rss = sum(r).^2/p;
prss = sum((r./diag(1-K*Ki)).^2)/p;

function n2 = dist2(X,C)
[ndata, dimx] = size(X);
[ncentres,dimc] = size(C);
n2 = (ones(ncentres, 1)*sum((X.^2)', 1))' + ...
ones(ndata, 1)*sum((C.^2)', 1) - ...
2.*(X*(C'));

function [ypred] = elmreg(X,tn,Q,nh)

[p, d] = size(X);
Xa = [ones(p,1),X];
Wh = rand(d+1, nh);

H = 1./(1+exp(-Xa*Wh));

w = pinv(H) * tn;

n = size(Q,1);
Qa = [ones(n,1),Q];
Ht = 1./(1+exp(-Qa*Wh));

ypred = 2*(Ht*w>0) - 1;

function X = normalise(X)

[p,d]=size(X);

X = X - ones(p,1)*mean(X);
X = X*diag(1./(s+(s==0)));
```



```

function [E,yy] = optimisekrr(X,t,Q)
sspan = linspace(1.5, 11.313, 100);
gspan = logspace(-12, -2, 100);

E = [];
t0 = clock;
for s = sspan
    prssmin = inf;
    for g = gspan
        [yy, rss, prss] = krr(X,t,Q,s,g);
        if prss < prssmin
            prssmin = prss;
        else
            gbest = g;
            prssmin;
            break
        end
    end
    E = [E, [s; g; prssmin]];
end
ElapsedTime = etime(clock,t0)

[prssmin, idx] = min(E(3,:));
msePESS = prssmin
sbest = E(1, idx)
gbest = E(2, idx)

```

EK 4. Kullanıcı Arayüzü.

```
function varargout = User_Interface(varargin)
gui_Singleton = 1;
gui_State = struct('gui_Name',       mfilename, ...
                  'gui_Singleton',  gui_Singleton, ...
                  'gui_OpeningFcn', @User_Interface_OpeningFcn, ...
                  'gui_OutputFcn',  @User_Interface_OutputFcn, ...
                  'gui_LayoutFcn',  [] , ...
                  'gui_Callback',   []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end

if nargout
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end

function User_Interface_OpeningFcn(hObject, eventdata, handles,
varargin)
handles.val = 3;

handles.output = hObject;

guidata(hObject, handles);

function varargout = User_Interface_OutputFcn(hObject, eventdata,
handles)
varargout{1} = handles.output;

function radiobutton2_Callback(hObject, eventdata, handles)

    if(get(hObject, 'Value')== 1)
        handles.val = 1;

        guidata(hObject,handles)
    end

function radiobutton3_Callback(hObject, eventdata, handles)
if(get(hObject, 'Value')== 1)
    handles.val = 2;
    guidata(hObject,handles)
end
```

```

function radiobutton4_Callback(hObject, eventdata, handles)

if(get(hObject,'Value')== 1)
    handles.val = 3;
    guidata(hObject,handles)
end

function pushbutton2_Callback(hObject, eventdata, handles)

h = hObject;
col = get(h, 'backg');
set(h, 'str', 'ÇALIŞIYOR...', 'backg', [1 .6 .6])

pause(.01)

switch(handles.val)
    case 1
        InputOneImage
        set(h, 'str', 'ÇALIŞIYOR...', 'backg', [1 .6 .6])
        if toc>5
            set(h, 'str', 'Sınıflandırıcı Çalışıyor...', 'backg', [1 .6
.6])
            tic;
        end
    case 2
        InputMultipleImages
        set(h, 'str', 'ÇALIŞIYOR...', 'backg', [1 .6 .6])
        if toc>2
            set(h, 'str', 'Sınıflandırıcı Çalışıyor...', 'backg', [1 .6
.6])
            tic;
        end
    case 3
        JustRunClassifier
        set(h, 'str', 'ÇALIŞIYOR...', 'backg', [1 .6 .6])
        if toc>2
            set(h, 'str', 'Sınıflandırıcı Çalışıyor...', 'backg', [1 .6
.6])
            tic;
        end
end

set(h, 'str', 'ÇALIŞTIR', 'backg', col)

```

ÖZGEÇMİŞ

1987 yılında İstanbul'da doğdu. İlk, orta ve lise öğrenimlerini İstanbul'da tamamladı. 2005 yılında Haliç Üniversitesi, Elektronik ve Haberleşme Mühendisliği Bölümünde lisans eğitimine başladı. 2009 yılında yine Haliç Üniversitesi Fen Bilimleri Enstitüsü, Elektronik ve Haberleşme Anabilim Dalı, Elektronik ve Haberleşme Mühendisliği programında yüksek lisans programına başladı ve aynı süre zarfında üniversite bünyesinde çalıştı.

İstanbul, 2011

M. Ayyüce Kızrak