

YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

79216

G-BANTLI KROMOZOMLARIN
BULANIK MANTIK SINIFLANDIRMA VE
ŞEKİL TANIMA YÖNTEMLERİ İLE
TANINMASINA YÖNELİK BİR SİSTEM

Bilgisayar Yük.Müh. Mine Elif KARSLIGİL

F.B.E. Bilgisayar Bilimleri Mühendisliği Anabilim Dalında
Hazırlanan

DOKTORA TEZİ

Tez Savunma Tarihi : 22 Nisan 1998
Tez Danışmanı : Prof. Mehmet Yahya KARSLIGİL (Y.T.Ü.)
Jüri Üyeleri : Prof. Dr. Asım CENANİ (İ.Ü.)
Doç. Dr. Lale AKARUN (B.Ü.)

79216

İSTANBUL, 1998



İÇİNDEKİLER

| | |
|---|-----|
| ŞEKİL LİSTESİ..... | iv |
| TABLO LİSTESİ..... | v |
| TEŞEKKÜR | vi |
| ÖZET | vii |
| ABSTRACT | ix |
| 1. GİRİŞ | 1 |
| 2. GENETİK İLE İLGİLİ TEMEL KAVRAMLAR..... | 6 |
| 2.1 Genetik Madde..... | 6 |
| 2.2 Kromozomun Tanımı..... | 6 |
| 2.3 Kromozomların Elde Edilmesi..... | 8 |
| 2.4 Kromozomların Sınıflandırılma Yöntemleri..... | 9 |
| 2.4.1 Kromozomların boy ve sentromer yerlerine göre sınıflandırılması..... | 10 |
| 2.4.2 Kromozomların bantlama tekniklerine göre sınıflandırılması..... | 12 |
| 2.4.2.1 Kromozomlarda sınırlar, bölgeler ve bantların tanımı..... | 13 |
| 2.4.2.2 Yüksek çözünürlüklü bantlama teknikleri..... | 14 |
| 2.5 Kromozomların Sıralanması (Karyotip)..... | 15 |
| 2.6 Kromozom Haritaları (İdiogram)..... | 16 |
| 3. G-BANTLI KROMOZOMLARIN SINIFLANDIRILARAK TANINMASI..... | 19 |
| 3.1 Kromozom Görüntüsünün Elde Edilmesi..... | 19 |
| 3.1.1 Görüntünün seçilmesinde dikkat edilmesi gereken kriterler..... | 19 |
| 3.2 Kromozom Sınıflandırma ve Tanıma İşlemini Olumsuz Etkileyen Faktörler ve Bunların Giderilmesi için Öneriler..... | 20 |
| 3.2.1 Metafaz üzerinde kromozom olmayan şekiller..... | 21 |
| 3.2.2 Kromozomların üst üste gelmesi..... | 22 |
| 3.2.3 Sentromer yerinin belirlenmesi..... | 23 |
| 3.2.4 Bant bilgilerinde görülen kayıplar..... | 23 |
| 3.3 Önerilen Kromozom Sınıflandırma ve Tanıma Sisteminin İşleyişi..... | 24 |
| 3.4 Kromozom Görüntülerinin Sınıflandırma ve Tanıma İşlemine Hazırlanması..... | 24 |
| 3.5 Kromozomların Sınıflandırılması ve Tanınması..... | 26 |
| 4. KROMOZOM SINIFLANDIRMA VE TANIMA SİSTEMİNDE KULLANILACAK ÖZELLİKLERİN ELDE EDİLMESİNDE YARARLANILAN YÖNTEMLER..... | 28 |
| 4.1 Kenar Belirleme..... | 28 |
| 4.1.1 Kenar belirleme işleminde eşik seviyesi kullanımı..... | 32 |
| 4.2 Sınır Belirleme..... | 34 |
| 4.3 Sınırları Belirlenmiş Bölgeye Ait Bilgilerin Elde Edilmesi..... | 37 |
| 4.3.1 Kapalı bir bölgenin içinin doldurulması..... | 37 |
| 4.3.2 Bir şeklin iskeletinin çıkarılması..... | 38 |
| 4.4 Kromozom Bilgilerinin Yönden ve Şekil Özelliklerinden Bağımsız Olarak Değerlendirilmesi..... | 40 |
| 4.4.1 Genişlik ve yoğunluk profilleri..... | 40 |
| 4.5 Kromozomların Sınıflandırılması ve Tanınmasında Kullanılan Kriterlerin Elde Edilmesi..... | 42 |
| 4.5.1 Kromozomun alanının ve boyunun bulunması..... | 42 |
| 4.5.2 Kromozomun sentromerinin bulunması..... | 43 |
| 4.5.3 Kromozomun p- ve q-kollarının belirlenmesi..... | 43 |
| 4.5.4 Kromozomun bant bilgilerinin bulunması ve değerlendirilmesi..... | 43 |

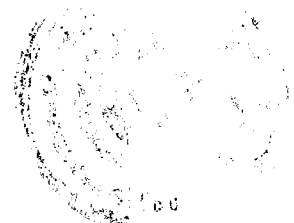


| | | |
|---------|---|-----|
| 5. | BULANIK MANTIK YÖNTEMLERİ İLE KROMOZOMLARIN SINIFLANDIRILMASI VE TANINMASI..... | 46 |
| 5.1 | Kromozomların Bulanık Mantık Kümeleri..... | 49 |
| 5.2 | Bulanık (Fuzzy) Üyelik Derecesi..... | 54 |
| 5.3 | Üyelik Derecesi Değerleri | 55 |
| 5.3.1 | Kromozomun alan ve boy üyelik derecelerinin elde edilmesi..... | 56 |
| 5.4 | Üyelik Derecesi Vektör ve Matrisleri..... | 59 |
| 5.4.1 | Alan ve boy üyelik derecesi vektörleri..... | 59 |
| 5.4.2 | Sentromer oranı üyelik derecesi vektörü..... | 60 |
| 5.4.3 | Bant bilgileri üyelik derecesi vektörü..... | 62 |
| 5.5 | Bulanık Mantık Kromozom Sınıflandırma Sisteminin Çalışması..... | 63 |
| 5.5.1 | Hazırlık aşaması | 63 |
| 5.5.2 | Eşdeğerlik bağıntısı benzerlik matrisinin oluşturulması..... | 64 |
| 5.5.2.1 | Klasik mantık eşdeğerlik ilişkisi (Denklik Bağı)..... | 64 |
| 5.5.2.2 | Bulanık mantık eşdeğer ilişkisi (Denklik Bağı)..... | 66 |
| 5.5.2.3 | Bulanık mantık ilişkisi kompozisyonu..... | 66 |
| 5.5.2.4 | Tolerans ilişkisinden eşdeğerlik ilişkisine geçiş..... | 67 |
| 5.5.3 | Kromozomların sınıflandırılması işlemi..... | 69 |
| 5.6 | Alfa (α) Kesmesi | 74 |
| 5.7 | Kromozom Sınıflandırma ve Tanıma İşlemi..... | 75 |
| 6. | UYGULAMA..... | 80 |
| 6.1 | Kromozomların Boy Bilgilerinin Değerlendirilmesi..... | 80 |
| 6.2 | Kromozomların Sentromer Bilgilerinin Değerlendirilmesi..... | 83 |
| 6.3 | Kromozomların Gruplara Ayrılması..... | 88 |
| 6.4 | Kromozomların Tanınması..... | 90 |
| 6.5 | Örnek Uygulama..... | 92 |
| 6.6 | Deneysel Sonuçlar ve Sistemin Başarısı..... | 93 |
| 7. | SONUÇ | 105 |
| | KAYNAKLAR | 108 |
| | ÖZGEÇMİŞ | 111 |



ŞEKİL LİSTESİ

| | | |
|------------|--|-----|
| Şekil 1.1 | Kromozomların mikroskop görüntüsü | 4 |
| Şekil 2.1 | Hücre sürecinin dört aşaması | 7 |
| Şekil 2.2 | Kromozomun sentromeri, p- ve q-kolları | 9 |
| Şekil 2.3 | Sentromer yerine göre kromozom şekilleri: (a) metasentrik, (b) submetasentrik, (c) akrosentrik | 10 |
| Şekil 2.4 | 20 numaralı kromozomun bantlarının numaralandırılması | 14 |
| Şekil 2.5 | İnsan kromozomlarına ait G-Bantlı karyotip örneği | 16 |
| Şekil 2.6 | 5 numaralı kromozoma ait 400, 550 ve 800 seviyeli idiogramlar | 17 |
| Şekil 2.7 | 550 bant seviyesine göre renklendirilmiş kromozomlarda bantların dağılımı | 18 |
| Şekil 3.1 | Kromozom Görüntüsünde Kromozom Olmayan Şekillere Örnek | 21 |
| Şekil 3.2 | Çakışan veya birbirine dokunan kromozomlar | 22 |
| Şekil 3.3 | Geliştirilen kromozom sınıflandırma ve tanıma sisteminin blok diyagramı | 24 |
| Şekil 4.1 | Kenar belirlemede birinci ve ikinci türevin durumu | 29 |
| Şekil 4.2 | Sobel operatörleri | 30 |
| Şekil 4.3 | İşlenmemiş metafaz örneği | 31 |
| Şekil 4.4 | Kenar belirleme işlemi sonucunda elde edilen metafaz örneği | 32 |
| Şekil 4.5 | Histogram ve Eşitlenmiş Histograma Örnek | 34 |
| Şekil 4.6 | Zincir koduna ait yön tanımları | 35 |
| Şekil 4.7 | Sınır üzerinde yer alan piksellerin birbirlerine göre konumları | 36 |
| Şekil 4.8 | İnceltme algoritmasında kullanılan komşuluk matrisi | 39 |
| Şekil 4.9 | İnceltme işleminin sonucu | 40 |
| Şekil 4.10 | Kromozom iskeletinden bant bilgilerinin elde edilmesi | 41 |
| Şekil 4.11 | Bant bilgileri elde edilen kromozomun şematik gösterimi | 45 |
| Şekil 5.1 | Sınıflandırma ve şekil tanıma sistemlerinin bağlantısı | 46 |
| Şekil 5.2 | Bir şekil tanıma sisteminin şematik yapısı | 47 |
| Şekil 5.3 | Boylarına göre sırasız A grubu | 48 |
| Şekil 5.4 | Boylarına göre sıralı A grubu | 48 |
| Şekil 5.5 | Farklı 1 numaralı kromozomlara ait bant bilgileri | 49 |
| Şekil 5.6 | Bir Üyelik Derecesi Fonksiyonu | 54 |
| Şekil 5.7 | Sıcaklık Üyelik Dereceleri | 56 |
| Şekil 5.8 | Alan Oranları İçin Üyelik Dereceleri | 58 |
| Şekil 5.9 | 46 kromozomun alan üyelik derecesi ilişkisi | 58 |
| Şekil 5.10 | Sentromer oranları üyelik dereceleri | 61 |
| Şekil 5.11 | Bant Bilgileri Üyelik Derecesi | 63 |
| Şekil 5.12 | Veri Akış Diyagramı | 76 |
| Şekil 6.1 | Boy benzerliklerine göre kromozomların görüldüğü gruplar | 80 |
| Şekil 6.2 | S_1, S_2, S_3, S_4 grupları arasında ISCN 1995'de tanımlı olan (A-G) gruplarına göre boy ve alan bilgisine bağlı olarak yapılan kesişimin sonucu | 86 |
| Şekil 6.3 | Örnek kromozom görüntüsü | 92 |
| Şekil 6.4 | Sınıflandırma ve tanıma işlemin sonunda örnek görüntüye ait karyotip | 101 |



TABLO LİSTESİ

| | | |
|------------|--|-----|
| Tablo 1.1 | Tıbbi genetik çalışmalarında kilometre taşı niteliğindeki önemli gelişmeler (Başaran 1996: 2) | 2 |
| Tablo 5.1 | Kromozomların İdiogram Bilgisine Göre Hesaplanan Oransal Bilgileri | 53 |
| Tablo 5.2 | Tolerans bulanık ilişki matrisi (00 değerleri 100 değerine denk düşmektedir) | 70 |
| Tablo 5.3 | Alan oranları için eşdeğer ilişki matrisi (00 değerleri 100 değerine denk düşmektedir) | 71 |
| Tablo 5.4 | İdiogram alan oranları için berrak ilişki matrisi | 72 |
| Tablo 6.1 | Oransal Boy Değerlerine Göre Gruplandırmaya 10 Örnek. Yanlış Grupta Olan Kromozomlar Koyu Renkte Gösterilmiştir. | 82 |
| Tablo 6.2 | Kromozomların Sentromer Özellikleri | 83 |
| Tablo 6.3 | 40 kromozom örneği için sentromer oranı aralığı değerleri | 84 |
| Tablo 6.4 | Sentromer Grupları | 85 |
| Tablo 6.5 | Sentromerlerin Değerlendirilmesi Sonucunda Gruplandırma Sonucu | 88 |
| Tablo 6.6 | Alan üyelik dereceleri matrisi (00 değerleri 100 değerine denk düşmektedir) | 94 |
| Tablo 6.7 | Eşdeğerlik ilişkisi matrisi (00 değerleri 100 değerine denk düşmektedir) | 95 |
| Tablo 6.8 | Berrak ilişki matrisi | 96 |
| Tablo 6.9 | Alan oranları benzerlik matrisi | 97 |
| Tablo 6.10 | Sentromer oranları eşdeğerlik ilişkisi matrisi (0 değerleri 100 değerine denk düşmektedir) | 98 |
| Tablo 6.11 | Sentromer oranları berrak ilişki matrisi | 98 |
| Tablo 6.12 | Örnek görüntü için kromozom tanıma sonuçları: Her kromozom içinde bulunduğu grupların şablon kromozomları ile korelasyon işlemine sokulmuş ve en çok benzediği şablon koyu renkte gösterilmiştir. | 99 |
| Tablo 6.13 | 40 kromozom görüntüsünün tanınması sonucunda elde edilen değerler. Her satır bir kromozom görüntüsündeki kromozomların tanıma sonucunda hangi numaralı kromozom olarak tanındığını göstermektedir. | 102 |
| Tablo 6.14 | 40 kromozom örneğine ait sınıflandırma ve tanıma işleminin sonucu | 103 |
| Tablo 6.15 | 40 kromozom görüntüsünün tanınması sonucunda elde edilen başarı | 104 |



TEŞEKKÜR

Doktora çalışmamda bilgisi, değerli görüşleri ve tecrübesi ile beni yönlendiren; kızı, öğrencisi ve asistanı olmaktan her zaman büyük gurur duyduğum ve kendime örnek aldığım, hocam Prof. M. Yahya Karşılıgil'e çok teşekkür ederim.

Bu çalışmanın temelini oluşturan Tıbbi Genetik konusunda verdikleri değerli destek için Prof. Dr. Asım Cenani yönetimindeki İstanbul Üniversitesi Cerrahpaşa Tıp Fakültesi Genetik ve Teratoloji Uygulama Araştırma Merkezi (GETAM)'ne; özellikle Tıbbi Genetik konusunda sağladığı değerli bilgiler, her türlü yardım ve önerileri için İstanbul Üniversitesi Cerrahpaşa Tıp Fakültesi Tıbbi Biyolojik Bilimler Genetik Anabilim Dalı öğretim üyesi Yrd. Doç. Dr. N. Aslı Silahtaroğlu'na; çalışmalarımda örnek olarak kullandığım kromozom görüntülerini sağlayan Vatan Hastanesi Prenetal Tanı Ünitesi'ne ve Genetik Uzmanı Dr. Jülide Caferler'e ve Kopenhag Üniversitesi Tıbbi Biyokimya ve Genetik Enstitüsü Tıbbi Genetik Anabilim Dalı'na ve Kim R. Rasmussen'e çok teşekkür ederim.

Tezimi inceleyerek değerli görüşleriyle beni yönlendiren jüri üyelerim Prof. Dr. Asım Cenani'ye ve Doç. Dr. Lale Akarun'a çok teşekkür ederim.

Doktora çalışmamın her aşamasında yaratıcı fikirleri, önerileri ve yardımlarıyla hep yanımda olan, özellikle tez kitabımın hazırlanması sırasında büyük yardımını ve desteğini gördüğüm sevgili arkadaşım Y.T.Ü. Bilgisayar Bilimleri ve Mühendisliği Bölümü Araştırma Görevlisi Ali Gökhan Yavuz'a; değerli önerileri için hocam Doç. Kirkor Harutunyan'a; yardımları için sevgili arkadaşlarım Y.T.Ü. Bilgisayar Bilimleri ve Mühendisliği Bölümü Araştırma Görevlileri Ahmet Tevfik İnan ve Banu Diri'ye, ve ablam Z. Gül Ünal'a çok teşekkür ederim.

En önemlisi, beni sevgi, özveri ve sabırla büyüten, sevgili annem Özgül Karşılıgil'e verdiği emekleri boşa çıkarmamış olmak ümidiyle çok teşekkür ederim.



ÖZET

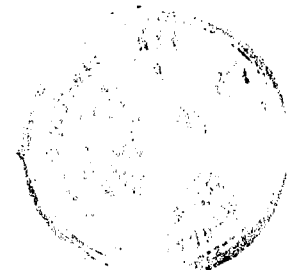
Bu çalışmada, genetik arařtırmaların esasını oluřturan kromozomların sınıflandırılmasını, Bulanık Mantık Görüntü İşleme ve Şekil Tanıma tekniklerini kullanarak gerçekleyen, bu konuda yapılan diđer çalışmalardan farklı olarak üç kademeli bir eleme yöntemine dayalı yeni bir sistem geliştirilmiř; sistem deęişik kromozom görüntülerine uygulanarak elde edilen sonuçlar deęerlendirilmiřtir.

Geliřtirilen bu sistemle sınıflandırılması yapılan G-Bantlama Teknięi ile boyanmıř kromozom görüntüleri, laboratuvar ortamında, ışık mikroskobuna baęlanan kamera aracılıęı ile veya kromozom fotoęraflarının uygun bir tarayıcı (scanner) üzerinden bilgisayara aktarılması ile elde edilmektedir.

Sınıflandırma yoluyla kromozom tanınması iřlemi üç ařamada gerçeşlenmektedir:

1. Kromozom görüntü kümesinden, görüntü işleme yöntemleri ile kromozomların ayrılması,
2. Elde edilen "iřlenmiř görüntüler" in renk, doku ve geometrik özelliklerinin elde edilmesi
3. Her kromozomun özelliklerinin bulanık sınıflandırma (Fuzzy Classification) kurallarına göre sınıflandırılıp tanınması

Görüntü işlemenin ilk ařamasında, kromozom görüntüleri uygun filtreler kullanılarak parazitlerden arındırılmaktadır. Kromozom kümesi halinde bir tek resimden oluřan temizlenmiř görüntü içerisinde her kromozomun yeri, kenar belirleme algoritmaları kullanılarak belirlenmektedir. Bu iřlem sonunda sınır bilgileri bulunan kromozomların ana görüntüden ayrılmaları saęlanmaktadır. Bu iřlemi izleyen ařamada her kromozomun uzunlamasına simetri eksenini (kromozomun iskeleti) ve buna baęlı olarak da genişlik ve yoğunluk profilleri elde edilmektedir.



Çalışmanın ikinci aşamasında genişlik ve yoğunluk profilleri kullanılarak kromozomların boyları, alanları, bant bilgileri ve renk (grilik seviyeleri) ve dokuları elde edilmektedir.

Son bölümde, kromozomlar boy ve alan bilgileri esas alınarak bulanık mantık benzerliklerine göre sınıflandırılmışlardır. Her kromozomun sentromer oranı, sadece içinde bulunduğu sınıftaki kromozomların sentromer oranlarına göre değerlendirilerek sentromer oranı benzerlik sınıfları matrisi elde edilmiş, boy, alan benzerlik sınıfları matrisi ile kesişim kümesi hesaplanarak kromozomların şekil özelliklerine göre gruplandırılması işlemi tamamlanmıştır. Kromozomların bant bilgilerini değerlendirmek için 22 otozom ve 2 cinsiyet kromozomu için 24 farklı kromozom şablonu oluşturulmuştur. Kromozomların bant bilgileri içinde bulunduğu gruba ait şablon kromozomlarla korelasyon yöntemi ile karşılaştırılarak her kromozomun kaç numaralı kromozom olduğu belirlenmiştir.



ABSTRACT

In this thesis a new approach to the classification problem of chromosomes has been proposed. The proposed approach is based on the image processing and pattern recognition and makes use of fuzzy classification methods using a three level decision system.

Images of chromosomes stained using G-Banding technique are transferred to the computer either via a camera attached to the light microscope, which is suitable for online laboratory research or from already taken photographs using a suitable scanner for later offline research and analysis.

The identification of chromosomes using classification techniques is accomplished in three steps:

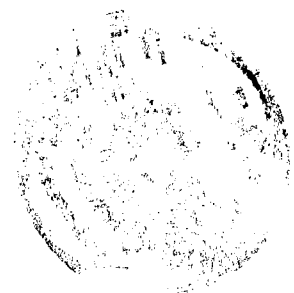
1. Separation of individual chromosomes from the whole metaphase image by means of image processing techniques.
2. Extraction of geometrical tissue and band characteristics of each chromosome from the processed images in step 1.
3. Classification and identification of chromosomes using Fuzzy Classification based on the characteristics obtained in step 2.

In the first step, an appropriate filter is applied to the whole metaphase image to reduce the noise caused by the environment and equipment. Resulting image is then undergone an edge detection process to distinguish the edges of chromosomes from the image background. Then each chromosome is separated from the image and stored individually.

Next step consists of the determination of vertical symmetry axis, length, band and tissue characteristics of each stored chromosome. These characteristics are then used to obtain distinctive features of the chromosome for further classification process.



In the last step fuzzy logic similarity relation matrices of chromosomes based on length, area and centromere ratios have been calculated. Each chromosomes centromere ratio is evaluated only among other chromosomes in its class and then the intersection of the aforementioned matrices have been calculated to obtain the final grouping of chromosomes. Templates for 22 autosomes and 2 genomes, which were prepared from 30 distinct chromosomes, are correlated with the chromosomes under test to find out each chromosomes number.

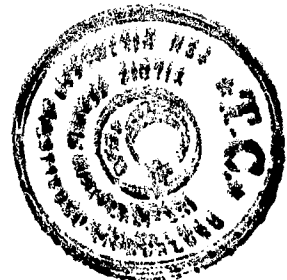


1. GİRİŞ

Uzay arařtırmalarının ilerlemesiyle birlikte bu arařtırmaların gerektirdiđi elektronik cihazlarda büyük geliřmeler gözlemlendi. Bu cihazların genel özellikleri, küçük boyutlu olmaları, az enerji tüketmeleri ve duyarlı olmalarıydı. Bu minyatür cihaz geliřimi tıbbın teřhis ve tedavide ihtiyaç duyduđu aletlerin yapımına yardımcı oldu. İnsan sađlığında gerekli olan iç organların görüntülenmesinde sadece röntgen cihazlarından yararlanılırken damar cidarlarının ekranda izlenmesine imkan veren aletler geliřtirildi. Yüksek frekanslı kristallerin esas eleman olarak kullanıldıđı ultrasonografi cihazlarıyla yumuřak dokulu iç organların sürekli (hareketli) görüntüleri izlenebilir hale geldi. Aynı řekilde Nükleer Manyetik Rezonans (MR) tekniđinin ve Lazer teknolojisinin tıpta kullanımıyla teřhis ve tedavide büyük ilerlemeler kaydedildi.

İnsan sađlığını tehdit eden unsurlar önceleri, enfeksiyonlar ve yetersiz beslenme olarak görülürken, özellikle geliřmiř ve geliřmekte olan ülkelerde sosyo-ekonomik kořullardaki iyileřme ve tıp bilimindeki bu hızlı geliřmelerle enfeksiyonlar kontrol altına alınırken **genetik** hastalıkların üzerine daha fazla ve dikkatli eğilinmesi geređi ortaya çıkmıřtır. Bugün için her yirmi kiřiden birisinin, genel hastanelere gelen her sekiz hastadan birisinin, çocuk hastanelerine gelen her beř hastadan ikisinin temel bir genetik bileřimi olan bir hastalık tařıdıđı bilinmektedir (Bařaran 1996:3).

Genetik dalında, özellikle 20.yüzyılın ortalarından beri kalıtsal hastalıkların önlenmesi ve iyi niteliklerin korunması konusunda önemli çalıřmalar yapılmaktadır. Tablo 1-1'de Tıbbi Genetik çalıřmalarındaki önemli geliřmeler görölmektedir (Bařaran 1996:2).



Tablo 1.1 Tıbbi genetik çalışmalarında kilometre taşı niteliğindeki önemli gelişmeler
(Başaran 1996: 2)

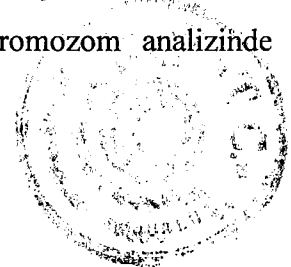
| | | |
|--|------|--------------------------------|
| Hücre teorisi | 1839 | Schleiden, Schwan |
| Evrimsel teori | 1859 | Darwin |
| Parçacıklı kalıtım | 1865 | Mendel |
| İlk kez kromozomun gözlenmesi | 1877 | Flemming |
| Kalıtsal hastalıklarda biyokimyasal temel | 1902 | Garrod |
| Genler kromozomlar tarafından taşınır | 1903 | Sutton, Boveri |
| İlk genetik kliniğin kurulması (ABD) | 1910 | Devenport |
| İlk kez bir genin yerinin saptanması | 1911 | Wilson |
| Kalıtımda DNA'nın rolünün ortaya konması | 1944 | Avery |
| DNA yapısının ortaya konması | 1953 | Watson, Crick |
| İnsanda 46 kromozom olduğunun bulunması | 1956 | Tijo, Levan |
| İnsanda ilk kromozomal anomalinin bulunması | 1959 | Lejeune |
| İlk prenatal cinsiyet tayini | 1960 | Riis, Fuchs |
| Kanda ilk kromozom analizi | 1960 | Moorhead |
| Moleküler hastalıklarda tarama testinin uygulanması | 1961 | Guthrie |
| Genetik kodun bulunması | 1961 | Nirenberg |
| İlk prenatal kromozom analizi | 1966 | Breg ve Steel |
| Rh immünizasyonunun önlenmesi | 1970 | Clarke |
| Kromozomların bantlanması | 1970 | Casperson |
| In vitro ilk gen sentezi | 1970 | Khorana |
| HLA ile hastalıkların ilişkilerinin ortaya konması | 1973 | Terasaki |
| İlk insan geninin klonlanması | 1977 | Shine |
| In vitro fertilizasyon | 1979 | Edwards, Steptoe |
| İnsülin'in genetik mühendislik yöntemleriyle üretilmesi | 1979 | Goeddel |
| Genetik mühendislik ürünlerinin satılmaya başlanması | 1982 | |
| DNA fingerprinting | 1985 | Jeffreys |
| Polimeraz Zincir Reaksiyonu (PCR) | 1985 | Saiki |
| İnsan kromozomlarında bağlantı çalışmalarının geliştirilmesi | 1987 | |
| İlk gen tedavisi | 1990 | Rosenberg, Anderson, Blease |
| p53 geninin bulunması | 1993 | |
| DNA tamir enzimi | 1994 | |



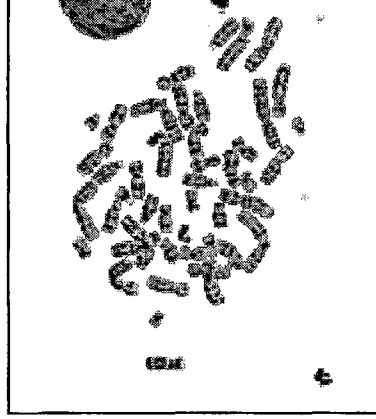
Özellikle, genetik materyalleri hücre seviyesinde inceleyen **sitogenetik** bilim dalındaki hızlı gelişmeler daha önceleri bilinmeyen bazı hastalıkların kromozomlara bağlı olduğunu göstermiştir. Kromozom elde etme ve kromozomları birbirinden ayırma tekniklerinde meydana gelen gelişmeler, monogenik ya da poligenik hastalık ya da sendromların, kromozomal hastalıklarla birlikte daha az etkili olması ya da sonraki kuşaklara daha az zarar vermesini sağlamak bakımından yapılan çalışmalara da zemin hazırlamış ve **genetik danışmanlık** artık ayrı bir çalışma alanı haline gelmiştir. Kalıtsal hastalıkların tedavisine yönelik çalışmalar da kromozomlardaki ayırma ve tanıma işlemlerine paralel olarak gelişmiştir. Yapılan bütün çalışmalar kromozomların sınıflandırılıp tanınması ve tanıma işleminden sonra da kalıtsal hastalıklara neden olabilecek bozukluğun bulunduğu kromozoma bağlı olan araştırmaların önemini arttırmıştır.

Gerçekleştirilen bu doktora çalışmasının konusu, “G-Bantlama Tekniği İle Boyanmış Kromozomların Bulanık Mantık Sınıflandırma ve Şekil Tanıma Yöntemleri ile Tanınmasına Yönelik Bir Sistem”dir. Geliştirilen bilgisayar destekli kromozom sınıflandırma sisteminin gerekliliğinin daha iyi anlaşılabilmesi için, laboratuvar ortamında kromozom analizinin nasıl yapıldığının incelenmesi yararlı olacaktır.

Kromozom analizi, metafaz aşamasındaki hücrenin mikroskop altında incelenip değerlendirilmesi yolu ile gerçekleştirilmektedir. Klasik yöntemlerde, mikroskoba bağlı bir fotoğraf makinasının yardımıyla kromozom görüntüsünün fotoğrafı çekilip, tab edilip kağıda basıldıktan sonra kromozomlar tek tek kesilir. Daha sonra, Bölüm 2.4’de anlatılan kromozom uzunluğu, bant bilgileri gibi ayırt edici özelliklerine göre sınıflandırılarak bir kart üzerine sıralı olarak yapıştırılır ve herhangi bir kromozomda bozukluk olup olmadığı gözle değerlendirilir. 46 tane objenin uzunluk, bant bilgileri gibi hassas ölçümler gerektiren bilgilerinin karşılaştırılmalı olarak değerlendirilmesinin zaman alıcı ve güç bir işlem olacağı açıktır. Ayrıca bir hasta için kromozom analizi tek bir hücre bilgisi üzerinden yapılmamaktadır. Normal olarak her hastadan 12-20 arası hücre alınarak incelenmekte ve bunların 3-5 tanesi analiz için seçilmektedir. Yani 12-20 hücrenin mikroskop altında fotoğrafı çekilmeden analizi yapılmakta, daha sonra bunların içinden seçilen 3-5 hücrenin fotoğrafları çekilip değerlendirilmektedir. Bu klasik yöntemde kromozom analizinde



uzman bir kiři bir günde en fazla iki hastanın bilgilerini deęerlendirebilmektedir. Őekil 1.1’de kromozomların mikroskop görüntüsü verilmiřtir.



Őekil 1.1 Kromozomların mikroskop görüntüsü

Kromozom analizi gerektiren durumlar incelenecek olursa, bunların ciddi kromozomal bozukluklar veya doęum öncesi tanı gibi sonucunun çok kısa zamanda elde edilmesi gereken vakalar oldukları görülür. Bu Őekilde klasik yöntemlerle yapılan analizlerin ise gereęinden çok zaman alıcı olacaęı açıktır. Bilinen en basit bilgisayar destekli kromozom analiz sistemleri bile incelendięinde bunların hız performansının klasik, elle yapılan sınıflandırmaya göre 40 kere daha hızlı olduęu belirlenmiřtir. Bundan bařka, hassas ölçümler gerektiren böyle bir konuda bilgisayarla yapılacak ölçüm sonuçlarının daha güvenilir oluřu da önemli bir özelliktir. Bütün bu noktalar deęerlendirildięinde, kromozom analizinde bilgisayardan yararlanmanın işlemleri hızlandırmak ve duyarlıęı arttırmak açısından yararı görülerek bu çalışmada bilgisayar yardımı ile kromozomları sınıflandırarak tanıyan bir sistem geliştirilmiřtir.

Bu tez çalışması, kromozomların, Őekilsel ve yapısal özelliklerinin deęerlendirilmesine dayalı bir çalışma olduęundan öncelikle bu çalışmada kullanılan, genetik ile ilgili temel kavramların bilinmesine ihtiyaç vardır. Bunun için, 2. bölümde genetik konusunda temel bilgiler verilmiř, insan kromozomlarının Őekilsel ve yapısal özellikleri tanıtılarak, laboratuvar ortamında kromozom sınıflandırma işleminin nasıl yapıldıęı anlatılmıřtır. 3.

bölümde, önerilen Kromozom Sınıflandırma Sistemi için gerekli kromozom bilgilerinin elde edilmesi ve sistemin işleyişi anlatılmıştır. 4. bölümde, kromozom görüntülerinin sınıflandırma işlemine hazır hale getirilmesi için yapılan görüntü işleme teknikleri anlatılmış ve bu tekniklerin uygulanmasıyla elde edilen kromozom bilgilerinden kromozomların sınıflandırılmasında kullanılacak kriterler olan boy, alan, sentromer yeri ve bant bilgilerinin elde edilişi anlatılmıştır. 5. bölümde, kromozomlardan elde edilen ayırt edici özellik bilgilerinin bulanık mantık (fuzzy logic) sınıflandırma ve şekil tanıma yöntemleri kullanılarak değerlendirilişi ve kromozomların nasıl tanındığı adım adım incelenmiştir. 6. bölümde, sistemin çalışmasının örnek kromozom görüntüleri üzerinde denenmesine yer verilmiş, 7. bölümde elde edilen sonuçlar değerlendirilmiştir.



2. GENETİK İLE İLGİLİ TEMEL KAVRAMLAR

Bu çalışmada geliştirilen kromozom sınıflandırma sisteminin değerlendirilebilmesi için, genetik ile ilgili temel kavramların bilinmesi gereklidir.

2.1 Genetik Madde

Mendel ve diğer arařtırmacıların deneyleri, canlıların tüm özelliklerinin kuşaktan kuşağa **gen** adı verilen maddelerle taşındığını göstermiştir. Gen denen parçacıklardan oluşan ve kuşaktan kuşağa aktarılan madde **genetik madde** adını almaktadır. Genetik maddeler üreme hücrelerinde **deoxyribonucleic acid (DNA)** ve kimi virüslerde de **ribonucleic acid (RNA)** olarak bulunur.

Genetik madde:

1. Kendisine tıpatıp benzeyen ya da kopyası olan maddeleri oluşturmak için kendi kendine çoğalır.
2. Enzimler ve hücre metabolizması için gerekli olan diğer makromoleküllerin sentezi ile bilgi aktarımı görevini yerine getirir.

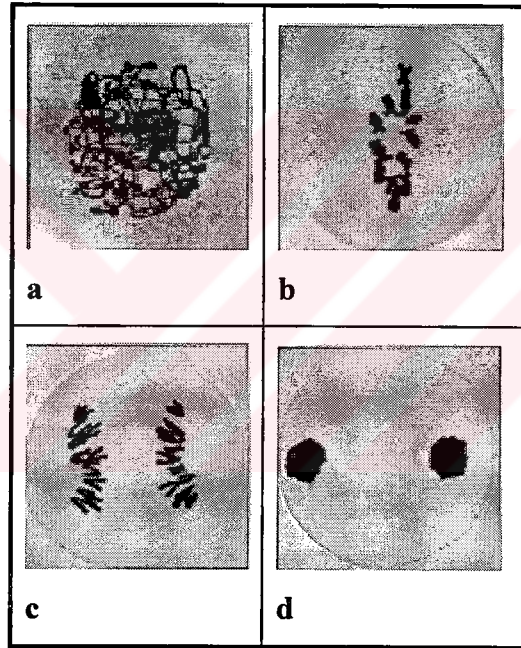
2.2 Kromozomun Tanımı

Yüksek yapılı canlılarda nükleik asitlerle histon ve/veya protamin gibi proteinler birleşerek hücre içinde genetik bilgilerin depolandığı **kromozomları** oluştururlar. Her türün vücut hücrelerinde kendine özgü ve tek bir kromozom yapısı ve sayısı vardır. İnsanlarda normal hücre çekirdeğinde **23 çift halinde 46 adet** kromozom bulunmaktadır. Her çiftin bir kromozomu **anne (maternal)** diğer kromozomu ise **baba (paternal)** kökenlidir. Erkek ve dişilerde 22 çift kromozom aynı yapıya sahip olup **otozom** olarak adlandırılmaktadırlar. 23.



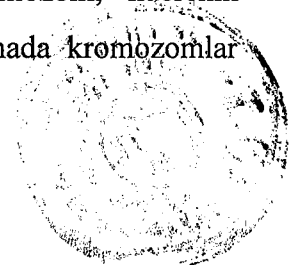
çift ise cinsiyet kromozomları olarak tanımlanmıştır ve dişilerde iki adet X, erkeklerde ise bir X ve bir Y kromozomundan oluşmaktadır.

Şekil 2.1’de ayrıntısı verilen hücre süreci esnasında, kromozomlar içinde yer alan DNA kendini kopyalar ve hücre de kendini orijinal hali ile özdeş iki eşit parçaya ayırır. DNA’nın kendini kopyalaması, **interfaz** olarak isimlendirilen aşamanın ortalarında gerçekleşir ve bu süreye S-faz’ı adı verilir. Kromozomlardaki değişik DNA sınıflarına sahip çeşitli bölgeler, S-faz’ı esnasında organize bir şekilde kopyalama işlemini gerçekleştirirler. Çekirdeğin ve sitoplazmanın parçalanması ise mitoz aşamasında gerçekleşir. Mitoz aşaması veya diğer adıyla M-faz’ı dört aşamadan oluşur:



Şekil 2.1 Hücre sürecinin dört aşaması

1. **Profaz:** Kromozomlar büzülürler, kalınlaşırlar ve daha yoğun şekilde renklenirler. Kromozom çiftleri, ayrık parçalar haline gelmeye başlarlar (Şekil 2-1.a).
2. **Metafaz:** Çekirdeğe ait membran kaybolur ve 46 adet kromozom, hücrenin çekirdeğinin merkezinde dairesel bir şekilde toplanırlar. Bu aşamada kromozomlar



oldukça kalınlaşmışlar ve birbiri ile özdeş iki kromatidden oluştukları iyice görülebilir hale gelmiştir. Kromatidler S-süreci sırasında üretilen genetik materyalin birebir kopyalarıdır. Kromozom bölünmesi sırasında kromatidlerin hareketini sağlayan iğ iplikleri de görülebilir hale gelmiştir (Şekil 2-1.b).

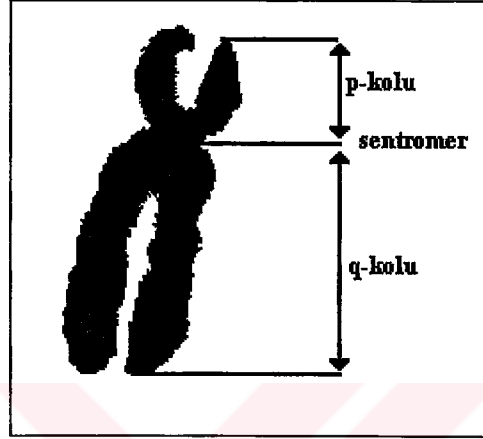
3. **Anafaz:** Bu aşamada kromatidler birbirlerinden ayrılırlar ve iğ iplik mekanizması her kromozoma ait kromatidleri hücrenin iki ucuna doğru harekete geçirir (Şekil 2-1.c).
4. **Telofaz:** Oluşan iki kromatid grubu birbirlerinden ayrılırlar ve iki adet yeni çekirdeği oluştururlar ve bunun hemen ardından sitoplazma da kendini ikiye böler (Şekil 2-1.d).

2.3 Kromozomların Elde Edilmesi

İnsanlara yönelik sitogenetik uygulamaların büyük çoğunluğu, bölünme özelliği gösteren, kemik iliği hücreleri, lenfositler gibi nükleuslu hücrelerden alınan örnekler üzerinde yapılmaktadır. Bu örnekler, **metafaz aşamasında** hücre bölünmesinin iğ ipliği hareketinin iç veya dış etkenler ile engellenerek durdurulması yoluyla elde edilir. Çekirdek membranı parçalanır, kromozomların yoğunlaşması devam eder fakat kromozomlar alışla geldiği üzere metafaz oluşumunu tamamlamazlar. Bu sayede kromozomlar sitoplazma içinde serbest şekilde hareket edebilirler. Böylece hücre içindeki kromozomlar bir yandan şekillerini korurlarken diğer yandan da işlenebilmek amacıyla daha rahat renklendirilebilirler. Metafaz aşamasında kromozomlar oldukça düzgün bir dağılıma sahip renklendirme teknikleri yardımıyla (**unbanded, solid-stained, block-stained**) işlenebilecekleri gibi, kromozom boyunca diferansiyel ton farklarının ortaya çıkmasını ve kromozomların daha kesin incelenmesini mümkün kılan tekniklerle (**banding**) de işlenebilirler.

Her normal metafaz kromozomu, '**primer darlık (primary constriction)**' adı verilen ve **sentromer** bölgesini tanımlayan noktada birbirine sıkı sıkıya bağlı iki kromatidden oluşur. Sentromer noktası aynı zamanda kromozomun mitoz sırasında iğ ipliği mekanizmasına bağlandığı noktadır. Bağlantının gerçekleştiği noktaya **kinetochore** adı verilir. Sentromer,

Şekil 2.2’de görüldüğü gibi, kromozomu iki ayrı kola bölen ve kromozomun eninin en dar olduğu bölgesidir. Bu kollardan kısa olanına ‘**petit**’ kelimesinden kısaltılarak **p-kolu**, uzun olanın ise ‘**grand**’ kelimesinden kısaltılarak **g-kolu** adı verilmiş, fakat sonraları p-kolu deyimine uyum sağlaması için g-kolu’na **q-kolu** denilmiştir. Tüm normal kromozomlarda, p-kolu kromozomun üst kısmındadır.

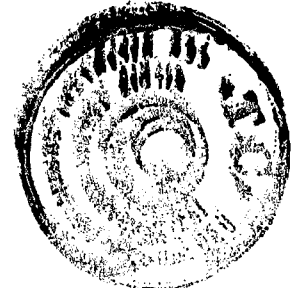


Şekil 2.2 Kromozomun sentromeri, p- ve q-kolları

Yukarıda kısaca bahsedilen diferansiyel renklendirme teknikleri bulunana kadar bir kromozomu diğer kromozomlardan ayırt etmek için p-kolu ve q-kolunun uzunlukları ile sentromerin kromozom üzerindeki yeri değerlendirilmiştir.

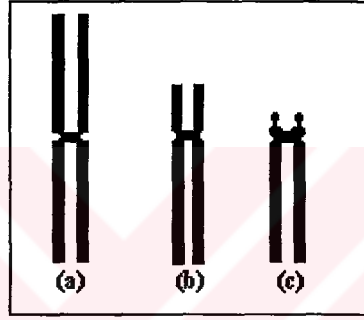
2.4 Kromozomların Sınıflandırılma Yöntemleri

Işık mikroskobu altında görülebilen kromozomlar gerekli inceleme ve değerlendirmelerin yapılabilmesi için önce belli kriterlere göre sınıflandırılmalıdırlar. Bu sınıflandırma işlemi ya kromozomun şekilsel özelliklerinden (sentromerin yeri, p-kolu, q-kolu uzunlukları) ya da kromozomun diferansiyel renklendirilmesi ile ortaya çıkan bant bilgilerinden yararlanılarak yapılır. Bu bölümde bu iki farklı sınıflandırma yöntemi ayrıntılı olarak incelenmiştir.



2.4.1 Kromozomların boy ve sentromer yerlerine göre sınıflandırılması

1960 yılında Denver’da temeli atılan ve daha sonra çeşitli toplantılarda (London 1963, Chicago 1966, Paris 1971, Paris 1975, Stockholm 1977, Paris 1980, Memphis 1994) yapılan çalışmalarla geliştirilen sisteme göre kromozomlar, boy ve sentromer özelliklerine göre değerlendirilirler. Şu anda, bu konudaki en geçerli olan çalışma **ISCN (An International System for Human Cytogenetic Nomenclature) 1995**’dir. Bu sisteme göre kromozomlar, Şekil 2.3’de görüldüğü gibi, sentromerlerinin kromozom üzerindeki konumuna göre üç farklı gruba ayrılırlar:



Şekil 2.3 Sentromer yerine göre kromozom şekilleri:

(a) metasentrik, (b) submetasentrik, (c) akrosentrik

1. **Metasentrik:** Sentromerin kromozomun yaklaşık olarak ortasında olduğu durumdur.
2. **Submetasentrik:** Sentromerin, p-kolunun en üstü ile ortası arasında olduğu kromozomdur.
3. **Akrosentrik:** Sentromerin, p-kolunun başına yakın olduğu kromozomdur. Bu kromozomlarda genelde p-kolunun ucuna bağcıklar ile bağlanmış, yuvarlak düğme biçiminde ve daha koyu renkli **uydu (satellite)** adı verilen yapılar görülür.



Ayrıca bazı kromozomlarda, sentromerden başka p-kolunda belirgin olarak bir darlık görülebilir. Özellikle 1-3-9 ve 16 numaralı kromozomlarda görülen bu daralma, **sekonder darlık (secondary constriction)** olarak adlandırılır.

ISCN 1995'e göre bant oluşturmeyen teknikler ile renklendirilmiş kromozomlar, boy ve sentromer pozisyonlarına göre büyükten küçüğe sıralı (A-G) diye isimlendirilen yedi gruba ayrılmışlardır:

1. **A Grubu (1-3 kromozom çiftleri):** Bu gruptaki kromozomlar diğer kromozomlara göre fark edilecek şekilde daha uzun ve metasentriktirler. 1 numaralı kromozom çifti, en uzun kromozom çiftidir. İki numaralı kromozomun sentromeri tam ortada olmadığı için 1 ve 3 numaralı kromozom çiftlerine göre daha fazla submedyandır.
2. **B Grubu (4-5 kromozom çiftleri):** Kısa kolun uzun kolun yaklaşık 1/4'ü kadar oluşu sebebiyle belirgin olarak submetasentrik olan bu kromozom çiftlerinden 5 numaralı kromozom 4 numaralı kromozomdan biraz daha küçüktür. Fakat birbirlerine çok benzer bir görünümde oldukları için bu iki kromozom çiftini birbirinden ayırmak oldukça güçtür.
3. **C Grubu (6-12 kromozom çiftleri ve X kromozomu):** Bu gruptaki kromozom çiftleri orta boy ve submetasentriktirler. Bu kromozomları sadece boy ve sentromer özelliklerine göre birbirinden ayırmak çok zordur.
4. **D Grubu (13-15 kromozom çiftleri):** Orta boylu, akrosentrik ve uydulu olan bu kromozomlar diğer kromozomlardan kolaylıkla ayrılacak bir görüntüye sahiptirler.
5. **E Grubu (16-18 kromozom çiftleri):** Boyları kısa sayılabilecek bu kromozom çiftlerinden 16 numaralı kromozom çifti metasentrik, 17-18 numaralı kromozom çiftleri submetasentrik özelliktedir. 16 numaralı kromozom çiftinin uzun kolunda sekonder darlık görülür.



6. **F Grubu (19-20 kromozom çiftleri):** Görünüşleri kelebeğe benzeyen bu kromozomlar, kısa ve metasentrik özelliğe sahiptirler.
7. **G Grubu (21-22 kromozom çiftleri ve Y kromozomu):** 21-22 numaralı kromozomlar kısa ve akrosentrikler ve uyduları vardır. Y kromozomu uzunluk olarak diğer iki kromozoma benzemekle birlikte, uydusu olmadığı için 21-22 numaralı kromozom çiftlerinden ayırt edilebilir.

2.4.2 Kromozomların bantlama tekniklerine göre sınıflandırılması

Denver Sistemine göre yapılan kromozom sınıflandırmaları incelendiğinde kromozomları tek tek tanımakta boy ve sentromer yeri özelliklerinin yetersiz kaldığı görülür. Kromozomları daha kesin sınıflandıracak özellikler bulmak için yapılan çalışmalar sonunda DNA segmentlerinin boya bağlama özelliğine dayanan ve kromozom bantlama olarak isimlendirilen farklı boyama teknikleri geliştirilmiştir. Bant kavramı ile anlatılmak istenen, kromozomun içinde yer alan ve üstündeki ve altındaki parçalardan, renginin açıklığı veya koyuluğu ile ayrılan bölgelerdir. Her kromozom farklı bant özelliğine sahip olduğu için, bantlama teknikleri alışlagelmiş diğer sınıflandırma yöntemlerine göre, daha kesin ve doğru sonuç vermektedirler. Bantlar kullanılan boyama tekniğine bağlı olarak bir yöntemde açık renk alabilirlerken bir diğer yöntem sonucunda koyu bir renge sahip olabilirler. Bantlama teknikleri iki temel gruba ayrılırlar:

1. Kromozomun tamamının açık koyulu bantlar halinde görüldüğü boyama teknikleri. **G, Q ve R** bantlama teknikleri bu gruba girer.
2. Kromozom üzerindeki belli bantların veya yapıların boyandığı teknikler. **C, T ve NOR** bantlama teknikleri bu gruba girer.

Bu çalışmada, günümüzde kromozom sınıflandırmada en çok tercih edilen bantlı boyama tekniklerinden biri olan G-bantlama (Giemsa banding) tekniği uygulanarak elde edilmiş kromozom görüntüleri kullanılmıştır.

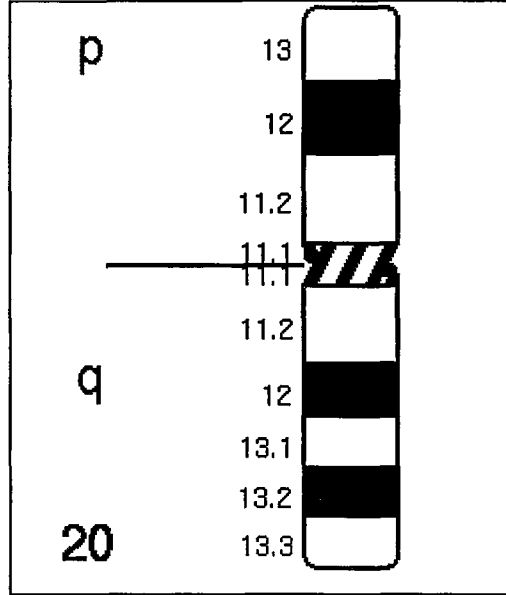


2.4.2.1 Kromozomlarda sınırlar, bölgeler ve bantların tanımı

Bölüm 2.4.2'de de tanımlandığı gibi, bant, kromozomda renklendirme sonucu açık veya koyu renkte görünen ve bu özelliği ile kendisine komşu diğer parçalardan ayrılan kısımlara verilen isimdir. Bantlar, kromozom kolları boyunca değişik bölgelerde yer almaktadırlar. Bu bölgeler birbirlerinden özel sınır işaretleri ile ayrılmışlardır. Bu sınır işaretleri, kromozomların belirlenmesinde ve sınıflandırılmasında rol oynayan tutarlı ve ayırt edici özelliğe sahip şekilsel özellikler olarak tanımlanmışlardır. En temel sınır işaretleri, kromozom kollarının uçları, sentromerin yeri ve belirgin özellikleri olan bantlar olarak sayılabilir. Bölge kavramı ise, yukarıda açıklanan sınır işaretlerinden herhangi ikisi tarafından çevrilmiş kapalı alan olarak düşünülebilir. Kromozomdaki bantlar ve buna bağlı olarak değişik bölgeler, sentromerden dışarı doğru, kromozomun numarası, p veya q kollarından hangisi üzerinde olduğu, bölge numarası ve bu bölgenin kaçınıcı bantı olduğunu gösterecek şekilde numaralandırılırlar.

Örneğin 1q21, 1 numaralı kromozomun q kolunun 2. bölgesinin 1. bandını göstermektedir. Koyu renk bir bant bölgesinde daha açık renkte veya açık renk bir bant bölgesinde daha koyu renkte gözüken ayrıntılı bantlar alt-bant olarak isimlendirilirler. Eğer bir kromozomda bu şekilde alt-bantlar varsa bu bantlar bir önceki gösterime '.' ile eklenerek gösterilirler. Örneğin 1q24.3, 1 numaralı kromozomun q kolunun 2. bölgesinin 4. bandının 3. alt bandını göstermektedir.





Şekil 2.4 20 numaralı kromozomun bantlarının numaralandırılması

Şekil 2.4’de 20 numaralı kromozomun bantlarının numaralandırılmış şematik gösterimi verilmiştir. En küçük bant numarası şematik gösterimde olmayan ve sentromerin yerini gösteren ‘10’ numaradır. Buna göre sentromerden dışarı doğru bantlamada bant çözünürlük seviyesine göre 11,11.1 veya 11.11 şeklinde numaralandırma yapılmaktadır.

2.4.2.2 Yüksek çözünürlüklü bantlama teknikleri

Çözünürlük seviyesi, 22 otozom ile X ve Y kromozomlarının tamamındaki toplam bant sayısı ile belirtilir. Normal bantlama teknikleri yaklaşık 250 bantlık bir çözünürlük ile bütün kromozomların yeterli hassasiyetle sınıflandırılmasını imkanı kılarken, pek çok kromozom bozukluğu bu çözünürlükte anlaşılabilir.

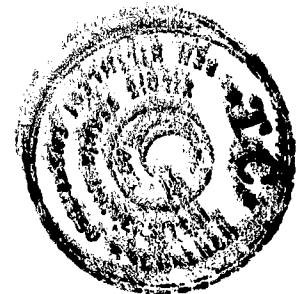
Yüksek çözünürlüklü bantlama teknikleri, mevcut bantlama tekniklerinin geliştirilmesi sonucunda elde edilmiştir. Bu teknikler hücre sürecinin değişik aşamalarındaki (profaz, prometafaz, interfaz) kromozomlara uygulanabilmektedir. Elde edilecek çözünürlük seviyesi normal bantlama tekniklerine göre kromozomların yoğunluklarından daha az

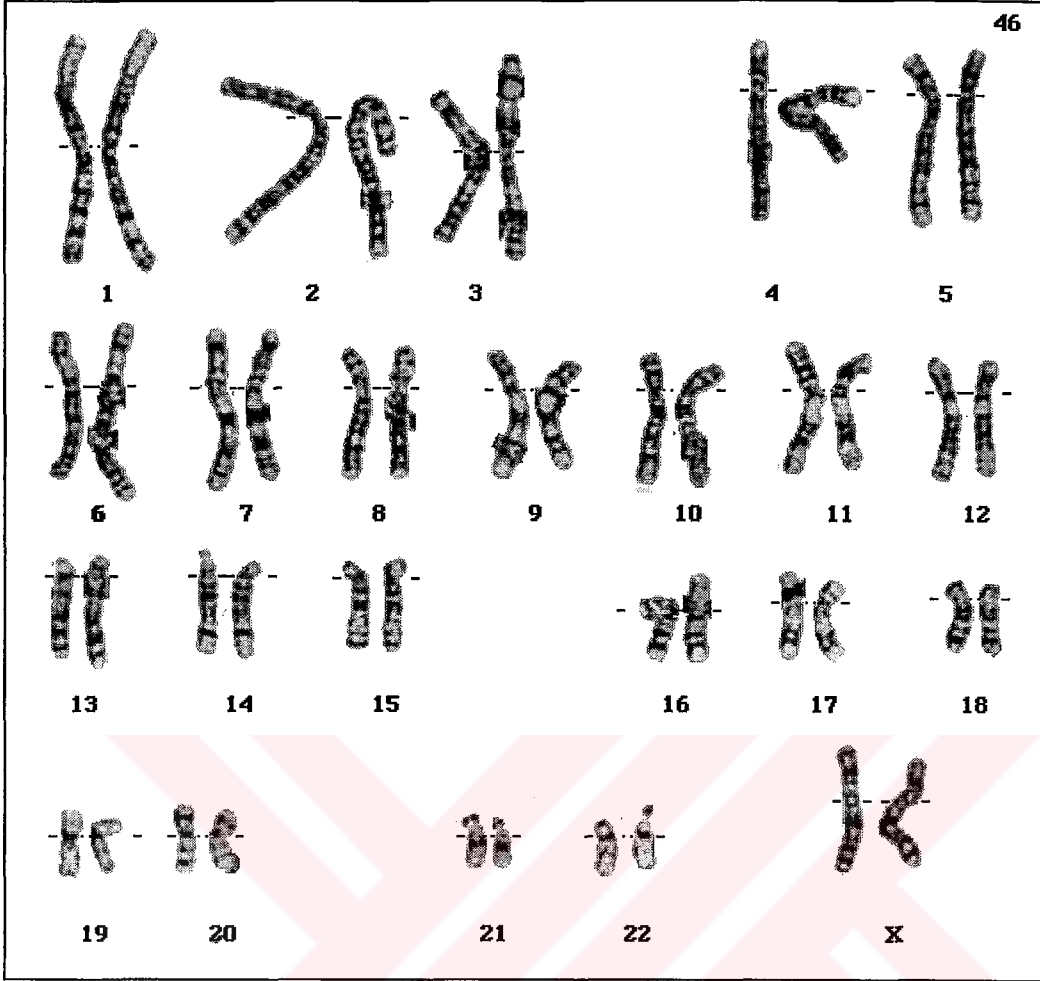
etkilenmektedir. Yüksek çözünürlük değerleri 400, 550 ve 850 bant olarak tanımlanmıştır. Örneğin 400 band seviyeli kromozom bilgileri, metafaz'ın hemen başlangıç aşamasında rahatlıkla elde edilebilir ve amniyosentez ve kan arařtırmalarında yüksek hassasiyet ile kullanılabilir. 550 band seviyesine ise profaz veya prometafaz aşamasındaki hücreler üzerinde yapılacak boyama teknikleri ile erişilebilir. 850 band seviyesi ise prometafaz aşamasındaki kromozomlarda trypsin-Giemsa bantlarının ölçülmesi ve bu işlem sırasında beş farklı renklendirme seviyesi kullanılması sonucunda elde edilebilir.

2.5 Kromozomların Sıralanması (Karyotip)

Kromozomların, Bölüm 2.4.1 ve 2.4.2'de anlatılan sınıflandırma teknikleri kullanılarak belli bir düzende sıralanmasına karyotip denir. Eğer değerlendirilecek kromozomlar bantsızsa, Bölüm 2.4.1'de belirtildiği gibi 7 grup halinde sınıflandırılıp, gruplar A' dan G' ye doğru sıralanırlar.

Bantlı kromozomların karyotipi de temelde bantsız kromozom karyotipi ile aynı olmakla birlikte, bant özelliklerinin sağladığı tek tek sınıflandırabilme üstünlüğü sayesinde, kromozomlar grup grup değil önce 1'den 22'ye kadar otozom kromozomlar, daha sonra sırasıyla X ve Y genozom kromozomları gelecek şekilde sıralanırlar. Şekil 2.5'de G-Bantlama tekniği ile sınıflandırılmış bir örnek görülmektedir.



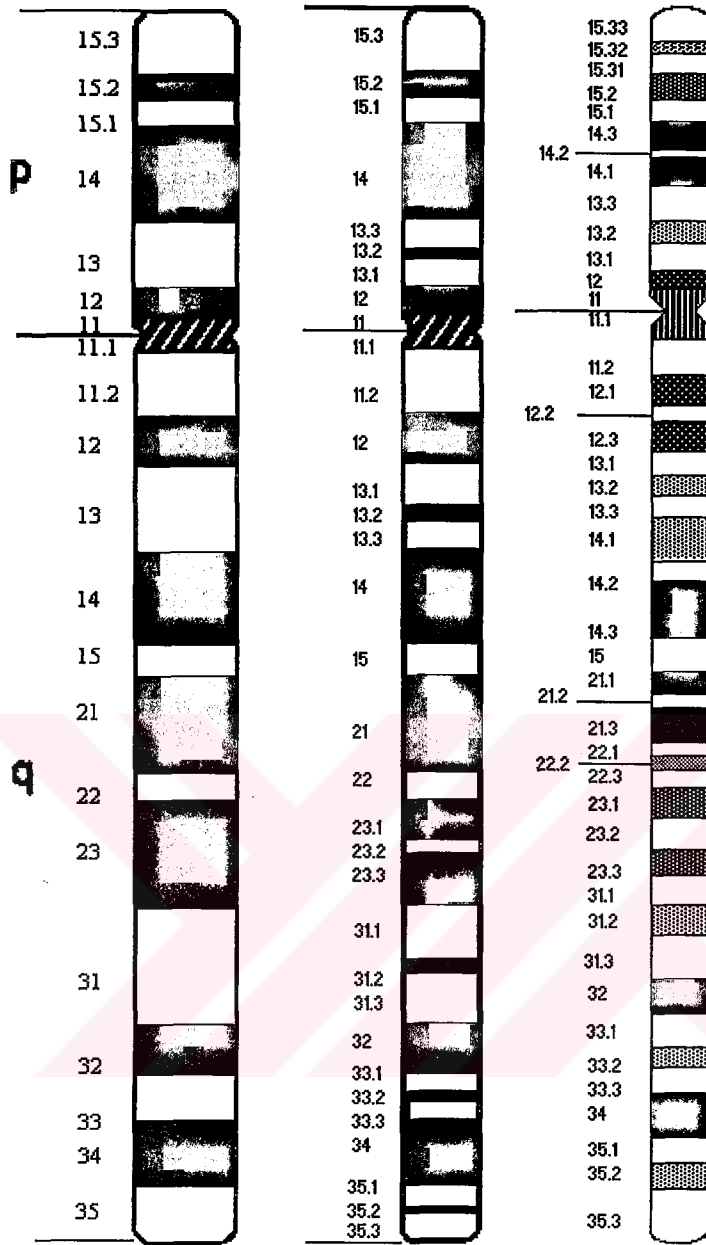


Şekil 2.5 İnsan kromozomlarına ait G-Bantlı karyotip örneği

2.6 Kromozom Haritaları (İdiogram)

Kromozom haritasının diyagram halinde gösterilmesine idiogram denir. 400, 550 ve 800 bant seviyeleri için farklı idiogramlar çıkarılmıştır.



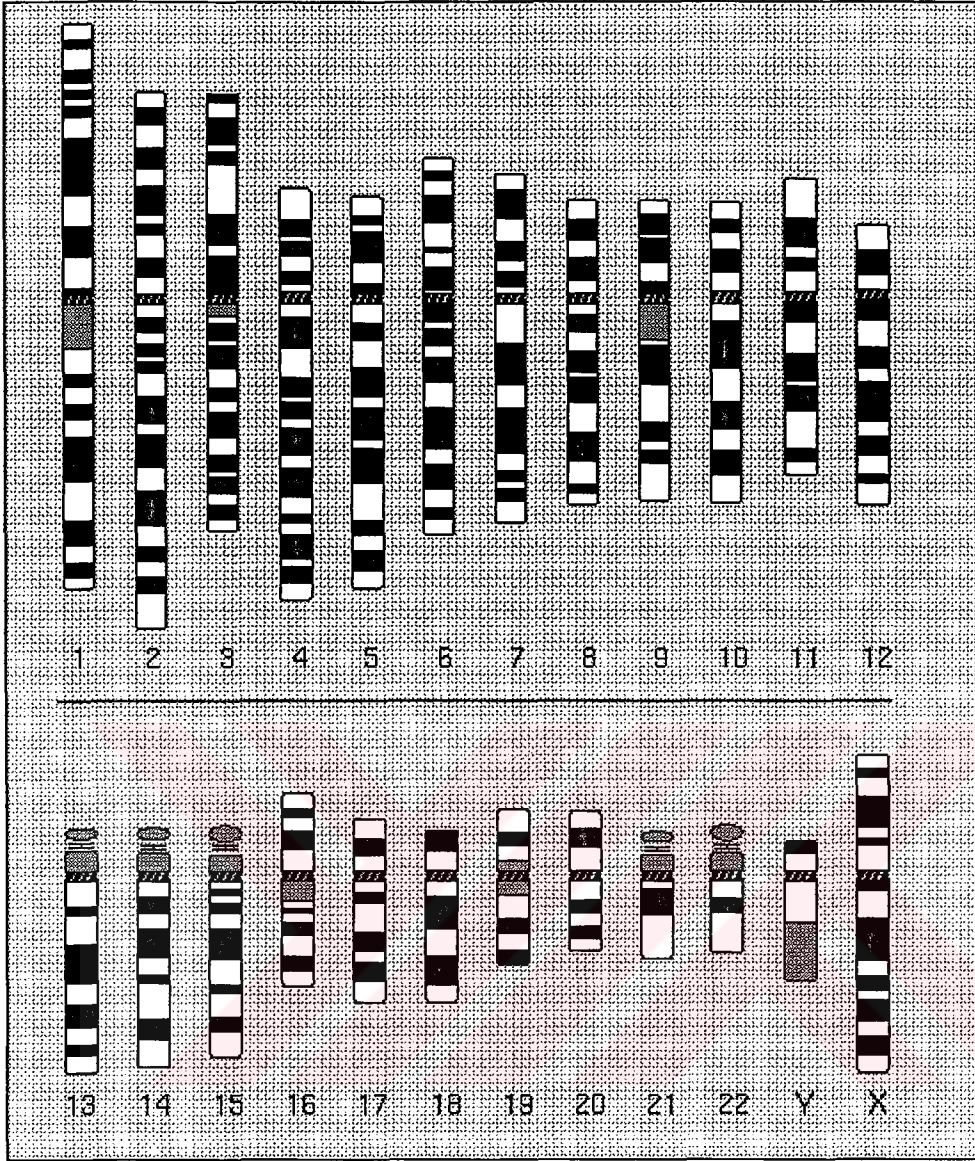


Şekil 2.6 5 numaralı kromozoma ait 400, 550 ve 800 seviyeli idiogramlar

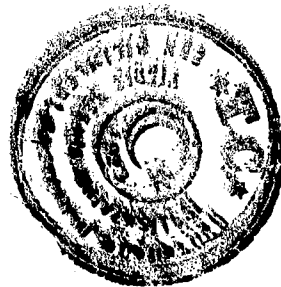
Şekil 2.6'da 5 numaralı kromozomun 400, 550 ve 850 bant seviyeleri için idiogramı görülmektedir.

Şekil 2.7'de idealde, 550 bant seviyeli kromozomlarda olması gereken bütün bant özelliklerini gösteren bantlı idiogram görülmektedir.





Şekil 2.7 550 bant seviyesine göre renklendirilmiş kromozomlarda bantların dağılımı



3. G-BANTLI KROMOZOMLARIN SINIFLANDIRILARAK TANINMASI

Bu tez çalışmasında kromozomların Bulanık Mantık Sınıflandırma ve Şekil Tanıma yöntemleri kullanılarak, tanınmasına yönelik yeni bir sistem geliştirilmiştir.

Geliştirilen sistem iki ana bölümden oluşmaktadır:

1. Kromozom görüntülerinin görüntü işleme yöntemleri ile sınıflandırma işlemine hazırlanması
2. Bulanık Mantık Sınıflandırma ve Tanıma Yöntemleri kullanılarak kromozomların sınıflandırılması ve tanınması

3.1 Kromozom Görüntüsünün Elde Edilmesi

Bu çalışmada geliştirilen sistem, laboratuvar ortamlarında ışık mikroskobundan elde edilen ve kamera aracılığıyla bilgisayara aktarılan görüntülerdeki kromozomların sınıflandırılmasında kullanılabileceği gibi, var olan kromozom fotoğraflarının uygun bir tarayıcı (scanner) aracılığıyla, araştırma yapmak üzere bilgisayara aktarılmasıyla elde edilen görüntülerdeki kromozomların sınıflandırılmasına da uygundur. Önerilen sistemin performansı, kromozom görüntülerinin elde edilmesi sırasında gösterilecek özene ve bu görüntünün bilgisayar ortamına aktarılmasında kullanılan cihazların duyarlılığına doğrudan bağlıdır.

3.1.1 Görüntünün seçilmesinde dikkat edilmesi gereken kriterler

Kromozomlar, metafaz aşamasında ışık mikroskobunda görülebilir hale gelirler. Giriş bölümünde belirtildiği gibi, bir hastadan alınan 15-20 hücre, mikroskop altında incelendikten sonra bunlardan ancak seçilen 3-5 tanesi değerlendirilmektedir.



Sınıflandırma işleminin başarısı, sınıflandırma işlemi için seçilecek görüntünün özellikleri ile yakından ilgilidir.

Sınıflandırma için örnek seçerken şu noktalara dikkat edilmelidir:

1. Kromozomların metafazın hangi evresinde oldukları önem taşır. Erken metafaz evresinde kromozomlar ip şeklinde uzun, geç metafaz evresinde ise yapıları bozulmaya başladığından, kolları açık halde olduğundan bu dönemlerden elde edilecek görüntüler sınıflandırma için uygun değildir.
2. Kromozomların çok geniş, ince veya kısa görünmemelerine dikkat edilmelidir.
3. Seçilecek görüntüdeki bant kalitesi de önemli bir faktördür. Kromozomların bantları göz ile ne kadar iyi ayırt edilebiliyorsa, bilgisayarla yapılacak sınıflandırmada da o kadar iyi sonuç alınacaktır.
4. Kromozomlarda üst üste çakışmaların mümkün olduğunca az olduğu görüntüler seçilmelidir.

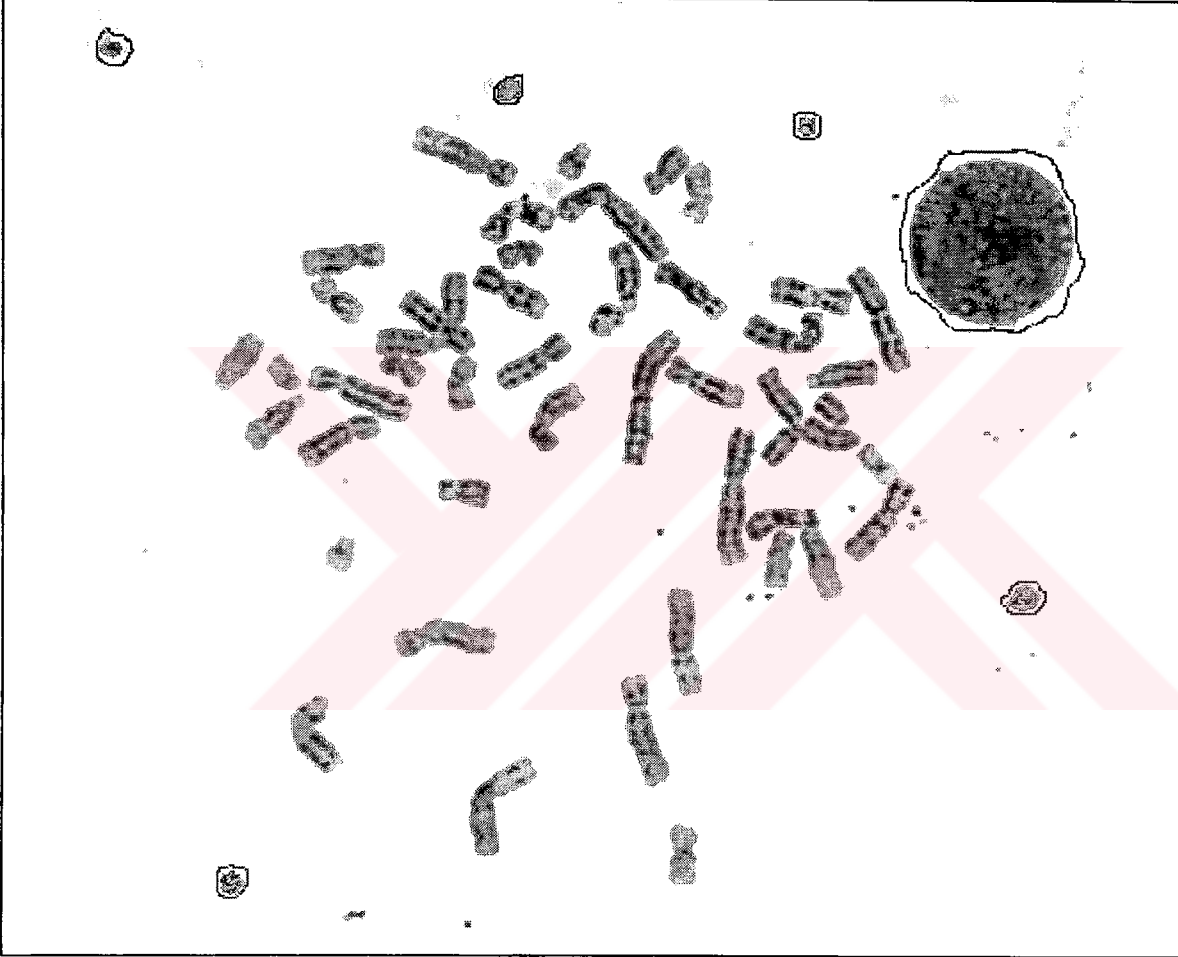
3.2 Kromozom Sınıflandırma ve Tanıma İşlemini Olumsuz Etkileyen Faktörler ve Bunların Giderilmesi için Öneriler

Farklı kromozom metafazları üzerinde yapılan sınıflandırma çalışmaları sırasında, çözümü olmayan bazı faktörlerin sınıflandırmanın başarısını olumsuz yönde etkilediği görülmüştür. Bu problemlere çözüm olarak, sınıflandırmanın belli aşamalarında kullanıcının onayının alınması ve kullanıcının gerekli düzeltmeleri yapmasına izin verilmesi uygundur.



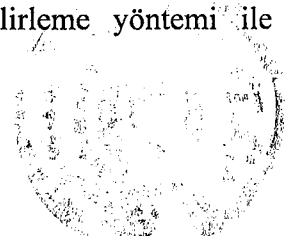
3.2.1 Metafaz üzerinde kromozom olmayan şekiller

İncelenen metafazlarda kromozom olmayan şekillerin de olduğu görülmüştür. Şekil 3.1’de çerçeve içine alınmış şekiller, interfaz aşamasında başka bir hücre, boya parçacıkları gibi kromozom olmayan bilgilerdir.



Şekil 3.1 Kromozom Görüntüsünde Kromozom Olmayan Şekillere Örnek

Ortalama bir kromozomun boyutundan büyük olan şekillerin kromozom olmadığını değerlendirmek mümkün olmakla birlikte görüntü üzerindeki boya parçacıklarını kromozomdan ayırt etmek her zaman için mümkün değildir. Bunun için görüntüdeki bütün kromozomların yerleri Bölüm 4.3’de anlatıldığı gibi sınır belirleme yöntemi ile



belirlendikten sonra sonuç kullanıcıya gösterilerek eğer kromozom olmayan şekiller de değerlendirilmişse bunların kullanıcı tarafından elimine edilmesi sağlanmalıdır.

3.2.2 Kromozomların üst üste gelmesi

Bir hastaya ait metafazlardan en iyisinin seçildiği görüntülerde bile en az iki veya üç kromozomun üst üste geldiği gözlenmiştir. Şekil 3.2’de Bölüm 4.2’de Sınır Belirleme Algoritması sonucunda belirlenen sınır değerlere göre birbirine dokunan veya çakışan kromozomlar çerçeve içine alınmıştır. Burada da görülebileceği gibi, çakışmanın çok düzgün olduğu özel durumlar dışında, genel olarak iki kromozom üst üste geldiğinde hangi parçanın hangi kromozoma ait olduğunu otomatik olarak bulmak mümkün değildir.



Şekil 3.2 Çakışan veya birbirine dokunan kromozomlar

Bu aşamada kullanıcının müdahalesine izin verilerek çakışan kromozomları doğru bir şekilde ayırması sağlanmalıdır. Kromozomların çakışması ile ortaya çıkan bir problem de çakışmanın olduğu noktadaki bant bilgilerinin kaybolarak çakışan kromozomların bant



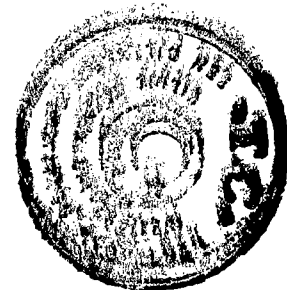
bilgilerinin değerlendirilmesini güçleştirmesidir ki çakışma noktasındaki bilgilerin değerlendirilebilir hale gelmesini sağlamak mümkün değildir.

3.2.3 Sentromer yerinin belirlenmesi

Üzerinde çalışılan pek çok kromozom görüntüsü incelendiğinde, gerek metafazda kromozomun boya tutma kalitesine bağlı olarak gerekse bilgilerin mikroskoptan bilgisayara aktarılana kadar oluşabilecek kayıplardan ötürü, bazı kromozomlarda sentromerden farklı herhangi bir yerin sentromerden daha dar olduğu görülmektedir. Sentromerin yanlış belirlenmesine bağlı olarak, p- ve q-kolları oranlarının belirlenmesi, kromozomun metasentrik, submetasentrik veya akrosentrik kromozom gruplarından hangisine dahil olduğu, p- ve q- kollarındaki bant bilgilerinin belirlenmesi aşamalarında da yanlış değerlendirilmeler yapılmış olacaktır. Tamamen görüntünün özelliklerine bağlı olan ve otomatik olarak düzeltilmesi mümkün olmayan böyle bir problemle karşılaşılması için sentromer yerleri belirlendikten sonra sonuçlar kullanıcıya gösterilerek hatalı olarak belirlenmiş sentromer yerleri varsa bir sonraki adıma geçmeden bunların düzeltilmesi sağlanmalıdır.

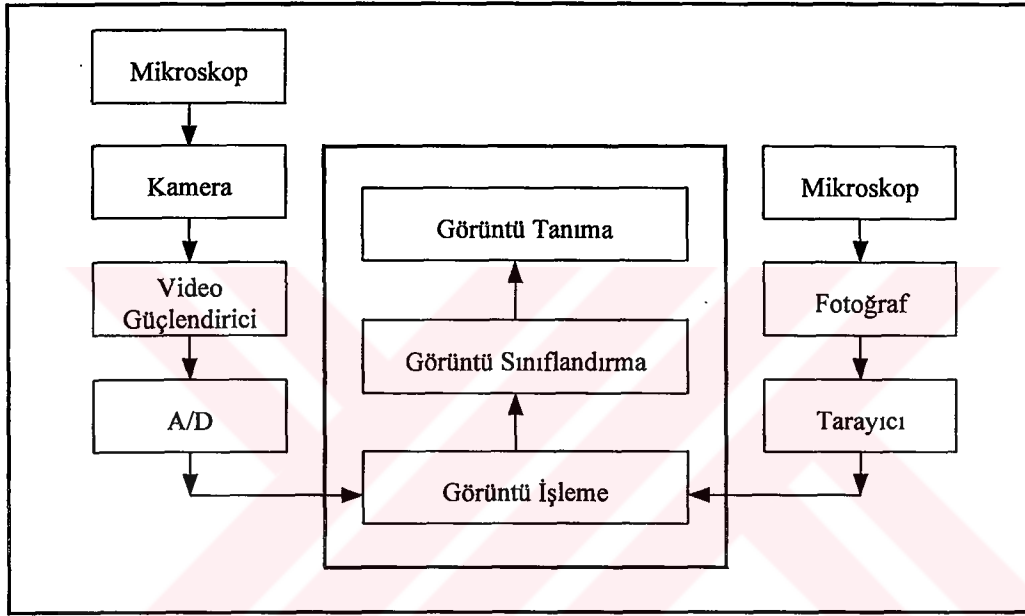
3.2.4 Bant bilgilerinde görülen kayıplar

Bantlı kromozom karyotipleri incelendiğinde, her karyotipte her bandın görülmediği görülür. Kromozomlarda bütün bantların her zaman görünmemesinin sebebi kısaca şöyle açıklanabilir: Hücre süreci sırasında DNA iplikçığı kromozom haline paketlenip sonra tekrar çözülmeye başlar. Kromozomlardaki bazı bölgeler, bu açılıp kapanmalar sırasında formlarını korurken, bazı bölgeler uzayıp kısalırlar ve bu da boyama özelliklerini etkiler. Bütün bant bilgilerinin idealde olması gerektiği gibi görünmemesi sınıflandırma performansını olumsuz yönde etkileyen ve algoritmik yöntemlerle çözümlenemeyecek bir problemdir.



3.3 Önerilen Kromozom Sınıflandırma ve Tanıma Sisteminin İşleyişi

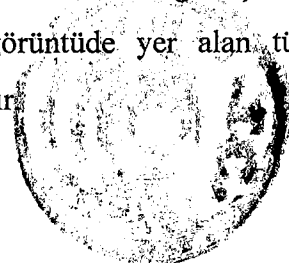
Bu çalışmada geliştirilen Kromozom Sınıflandırma ve Tanıma Sistemi, kromozom görüntülerinin sınıflandırma işlemine hazırlanmasını sağlayan sayısal görüntü işleme aşamalarından ve bu işlemlerin sonucunda elde edilen sonuç görüntüye uygulanan sınıflandırma ve tanıma yöntemlerinden oluşmaktadır. Sistemin işleyişi Şekil 3.3'deki Blok diyagramı ile gösterilmiştir.



Şekil 3.3 Geliştirilen kromozom sınıflandırma ve tanıma sisteminin blok diyagramı

3.4 Kromozom Görüntülerinin Sınıflandırma ve Tanıma İşlemine Hazırlanması

Bu çalışmada geliştirilen sistem, 8-bit (256) gri seviyesine sahip, TIFF (Tag Image File Format) yöntemine göre kodlanmış görüntüler ile çalışmak üzere tasarlanmıştır. İlk adım olarak, bilgisayara aktarılmış, metafaz aşamasındaki kromozomlardan oluşan görüntüye kenar belirleme algoritması uygulanarak, 0 (siyah) ile 255 (beyaz) arasında değerlere sahip gri seviyelerinden oluşan görüntüde kenarların beyaz ile gösterilerek belirginleşmesi, diğer kısımların ise siyah yapılması sağlanmıştır. Böylece görüntüde yer alan tüm kromozomların kenar hatları görüntünün geri kalanından ayrılmıştır.

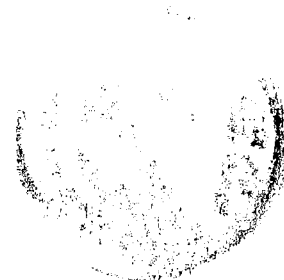


Bu şekilde işlenmiş görüntü, lineer olarak satır satır taranarak beyaz renkteki ilk nokta bulunmuştur. Bu noktanın görüntü içindeki kromozomlardan birinin kenarı üzerinde olduğu bilindiği için, bu noktadan başlayarak görüntüye sınır belirleme algoritması uygulanmıştır. Böylece bu kromozoma ait sınır bilgileri x-y eksen sisteminde her nokta için x-y çiftleri şeklinde elde edilmiştir. x-y çiftleri içinde x ve y değerlerinin minimum ve maksimumları bulunarak bu kromozomun görüntü içinde kapladığı alan belirlenmiştir. Bu alanın tamamını içine alacak şekilde $m = x_{\max} - x_{\min}$, $n = y_{\max} - y_{\min}$ boyutlarında $S(m,n)$ matrisi oluşturulmuştur. İlk değeri '0' olan bu matriste sınıra karşılık gelen gözlemlere '1' değeri atanmıştır. Böylece $S(m,n)$ matrisinde kromozomun sınır değerleri saklanmıştır.

Kromozomun iskeletini çıkarmak ve bant bilgilerini elde etmek için bir önceki adımda kullanılan $S(m,n)$ matrisi ile aynı boyutlarda $I(m,n)$ iskelet, $B(m,n)$ bant matrisleri oluşturulmuştur. $S(m,n)$ matrisindeki sınır bilgisinden yararlanarak, $I(m,n)$ matrisi sınır içinde kalan matris gözlemleri '1' değeri ile $B(m,n)$ matrisi ise orijinal görüntüde kromozomun iç bölgesini oluşturan gri seviyesi değerleri ile doldurulmuştur. $I(m,n)$ matrisinde, sınır içinde kalan alanın '1' değeri ile dolu olması iskelet çıkartma işlemi sonucunda uniform ve tek piksel kalınlığında iskelet elde edilmesini sağlamaktadır.

Bu aşamada, kromozoma ait, ilerideki işlemler için gerekli bilgiler elde edildikleri ve $S(m,n)$, $I(m,n)$, $B(m,n)$ matrislerinde saklandıkları için bu kromozomun orijinal görüntüde kapladığı alan '0' değerleri ile doldurularak silinmektedir. Böylece değerlendirilmiş kromozomun, orijinal görüntüde diğer kromozomların belirlenme işlemini etkilememesi sağlanmıştır.

Görüntüde işlenecek kromozom kalmayınca kadar yukarıda belirtilen işlemler tekrar edilerek tüm kromozomlar kendilerine ait iskelet ve bant matrislerine ($I(m,n)$, $B(m,n)$) yerleştirilmiştir.



İskelet matrisleri, Bölüm 4.3.2’de anlatılan inceltme yöntemleri kullanılarak işlenmiş ve her kromozoma ait ‘1’ piksel kalınlığında iskeletler elde edilmiştir. Bu iskeletler, kromozomların dik eksenine göre simetri eksenini oluşturduklarından kromozomların görüntü içindeki yön ve şeklinden bağımsız olarak işlenebilmelerini sağlamaktadırlar.

İskelet matrisi $I(m,n)$, içinde ‘0’ değerinden farklı ilk piksel bulunarak, iskelet üzerinde zincir kodları yardımıyla ilerlenmiştir. Bu işlemde önemli olan iskeletin iki uç noktasını bularak iskeletin matris içindeki doğrultusunun belirlenebilmesidir. Bunun için şöyle bir yöntem geliştirilmiştir; iskelet üzerinde zincir kodu ile ilerlenirken p_n noktasının bir adım sonraki komşusu olan p_{n+1} , p_{n-1} adımıdaki nokta ise, iskelet üzerinde geri dönmeye başlanmış demektir. Böylece p_n noktasının kromozomun uçlarından biri olduğu bulunmuştur. Zincir kodlamaya iskelet üzerinde diğer yönde de aynı durum oluşuncaya kadar devam edilerek kromozomun diğer ucu da belirlenmiştir.

İki uç noktası belirlenmiş iskeletlerden yararlanarak bant matrisleri, $B(m,n)$, üzerinde ilerlenerek, kromozomların bant bilgileri, kromozomlar düz ve y-eksenine paralel konumdaymış gibi düzenlenerek bir $X(k,l)$ matrisine yerleştirilmiştir. Bu matriste k , kromozomun eninin en geniş olduğu yeri, l ise kromozom iskeletinin uzunluğunu göstermektedir.

3.5 Kromozomların Sınıflandırılması ve Tanınması

Kromozom Sınıflandırma ve Tanıma sisteminde, kromozomları gruplara ayırmak için boy, alan ve sentromer bilgileri, gruplara ayrılmış kromozomları tek tek tanımak için ise bant bilgileri değerlendirilmiştir.

Bu bilgileri elde etmek için, $X(k,l)$ matrisi kullanılarak kromozomun genişlik ve gri seviyesi yoğunluğu profilleri Bölüm 4.4.1’de anlatıldığı gibi hesaplanarak, genişlik profili bir $GP(l)$ vektörüne, yoğunluk profili ise bir $YP(l)$ vektörüne yerleştirilmiştir.



Her kromozoma ait $X(k,l)$ matrisinden alınan l bilgilerinin 46 kromozomun toplam boy değerine oranlanmasıyla elde edilen değerlerin $L(46)$ vektöründe toplanmasıyla tüm kromozomlara ait boy bilgileri elde edilmiştir. $GP(l)$ vektörlerinin yatay toplamlarından elde edilen alan bilgileri ise, 46 kromozomun toplam alan değerine oranlanarak $A(46)$ vektöründe saklanmıştır. Ayrıca Bölüm 4.5.2'de anlatılan yöntemin $GP(l)$ vektörlerine uygulanması ile sentromer yerlerini içeren $S(46)$ vektörü elde edilmiştir. $YP(l)$ vektörlerinin Bölüm 4.5.4'de açıklanan yöntemle işlenmesi sonucunda her kromozoma ait bant bilgileri hesaplanmış ve $B(l)$ vektörlerinde saklanmıştır.

Boy ve alan bilgileri Bölüm 5'de anlatılan Bulanık Mantık Sınıflandırma yöntemleri ile değerlendirilerek kromozomlar boy ve alan benzerliklerine göre sınıflandırılmışlardır. Her sınıftaki kromozomların sentromer özellikleri de değerlendirilerek kromozomların (A-G) arası yedi gruba yerleştirilmeleri sağlanmıştır. Daha sonra her gruptaki kromozomların bant bilgileri o gruba ait şablon kromozomların bant bilgileri ile karşılaştırılarak hangi kromozoma en çok benzediğine, yani kaç numaralı kromozom olduğuna karar verilmiştir.



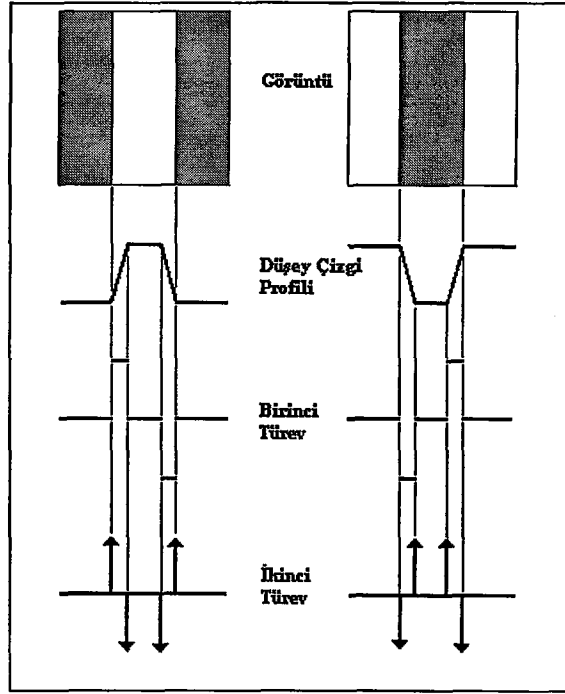
4. KROMOZOM SINIFLARDIRMA VE TANIMA SİSTEMİNDE KULLANILACAK ÖZELLİKLERİN ELDE EDİLMESİNDE YARARLANILAN YÖNTEMLER

4.1 Kenar Belirleme

Belirgin şekilde gri seviyesi farkı olan iki bölge arasında kalan sınır çizgisi kenar (edge) olarak isimlendirilir. Normal kenar belirleme yöntemlerinin kullanılabilmesi için, yukarıda belirtilen bölgelerdeki grilik seviyesi dağılımının yeterince homojen olması gereklidir. Aksi takdirde kenar belirlemek için segmentasyon (Gonzalez,1987,335) yöntemlerinin kullanılması daha uygun olacaktır.

Temel anlamda, kenar belirleme işlemi, sayısal görüntü üzerinde gerçekleştirilen kısmi bir türev alma işlemi olarak düşünülebilir. Sayısal görüntüdeki kenarlar, sayısal görüntünün örnekleme yöntemi ile elde edildiği de göz önüne alınacak olursa, gri seviyesindeki ani değişikliklerden çok, rampa (basamak) halinde yumuşak geçiş olarak modellenir. Anlatıldığı şekilde modellenmiş bir kenara ait ilk türev, gri seviyesi değişmediği sürece '0' değerini, gri seviyesinin değiştiği bölgede ki, burası kenarın olduğu bölgedir, sabit bir n değerini alır. İkinci türev ise, gri seviyesinin başladığı ve bittiği noktalar hariç '0' değerine sahiptir. Şekil 4.1'de solda koyu zemin üzerine açık renk görüntü, sağda açık zemin üzerine koyu renk görüntü için, birinci ve ikinci türev değerlendirmesi gösterilmektedir. Bu bilgiler ışığında, birinci türevin genliği inceleme yapılan sayısal görüntüde bir kenarın var olup olmadığını anlamak, ikinci türevin pozitif veya negatif oluşu, geçişin koyu gri seviyelerinden açık gri seviyelerinin olduğu bir bölge yönünde mi yoksa tam tersi yönde mi olduğunu belirlemek amacıyla kullanılabilir.





Şekil 4.1 Kenar belirlemede birinci ve ikinci türevin durumu

Sayısal bir görüntüde, herhangi bir noktaya ait ilk türev değeri o noktaya gradient operatörünün uygulanması ile belirlenir. Gradient operatörünün tanımı Eşitlik 4.1'de verilmiştir.

$$G[f(x,y)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (4.1)$$

Kenar belirlemek amacıyla gradient operatörünün genliği Eşitlik 4.2'de belirtildiği gibi hesaplanmalıdır.

$$G[f(x,y)] = [G_x^2 + G_y^2]^{1/2} \quad (4.2)$$



Gradient operatörünün kullanımında dikkat edilmesi gereken önemli bir ayrıntı, işlemin uygulandığı piksele ait kaç adet komşu noktadan yararlanıldığıdır. Küçük bir komşuluk alanı seçilecek olursa, sayısal görüntü içindeki gürültüden daha fazla etkilenme durumu ortaya çıkar. Bu nedenle 3x3'lük bir alan kullanılarak, 2x2'lik bir alan kullanımına göre daha düzgün ve net bir kenar belirlemenin yanı sıra gürültüye karşı olan hassasiyet de dengelenmiş olur. Bundan başka, Gradient operatörünün kullanımının getirdiği matematiksel karmaşıklık da dikkate alındığında, 3x3'lük komşuluk alanı kullanımı daha çok tercih edilmektedir. Gradient operatöründen yararlanılarak çıkarılan Roberts, Smoothed, Isotropic, Sobel operatörleri gibi operatörler vardır. Bu çalışmada kenar belirlemek için Sobel operatöründen yararlanılmıştır. Sobel operatörünün yatay ve dikey kenarları belirlemek için iki ayrı maskesi vardır. Bu maskelerin değerleri Şekil 4.2'de verilmiştir.

| | | |
|----|----|----|
| -1 | -2 | -1 |
| 0 | 0 | 0 |
| 1 | 2 | 1 |

(a)

| | | |
|----|---|---|
| -1 | 0 | 1 |
| -2 | 0 | 2 |
| -1 | 0 | 1 |

(b)

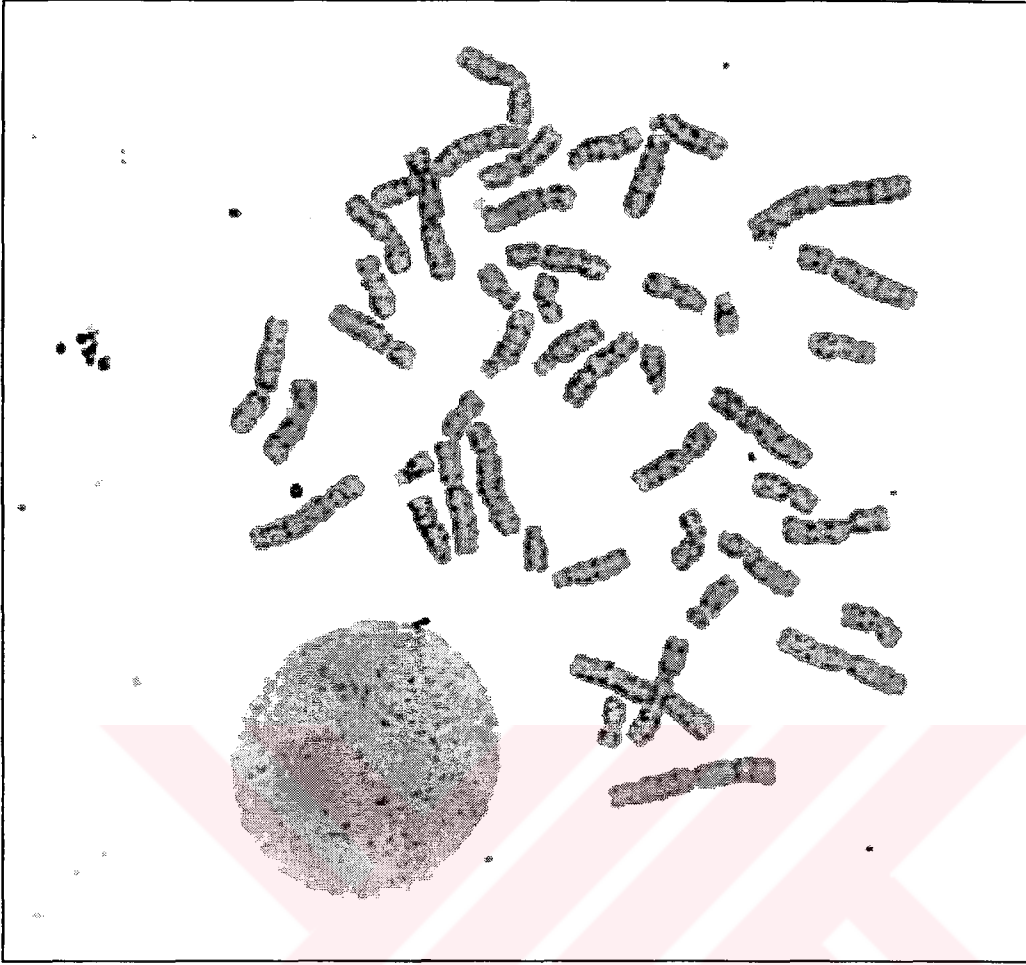
Şekil 4.2 Sobel operatörleri,

(a) 3x3'lük bir alanda G_x değerinin hesaplanması için kullanılan Sobel maskesi,

(b) 3x3'lük bir alanda G_y değerinin hesaplanması için kullanılan Sobel maskesi

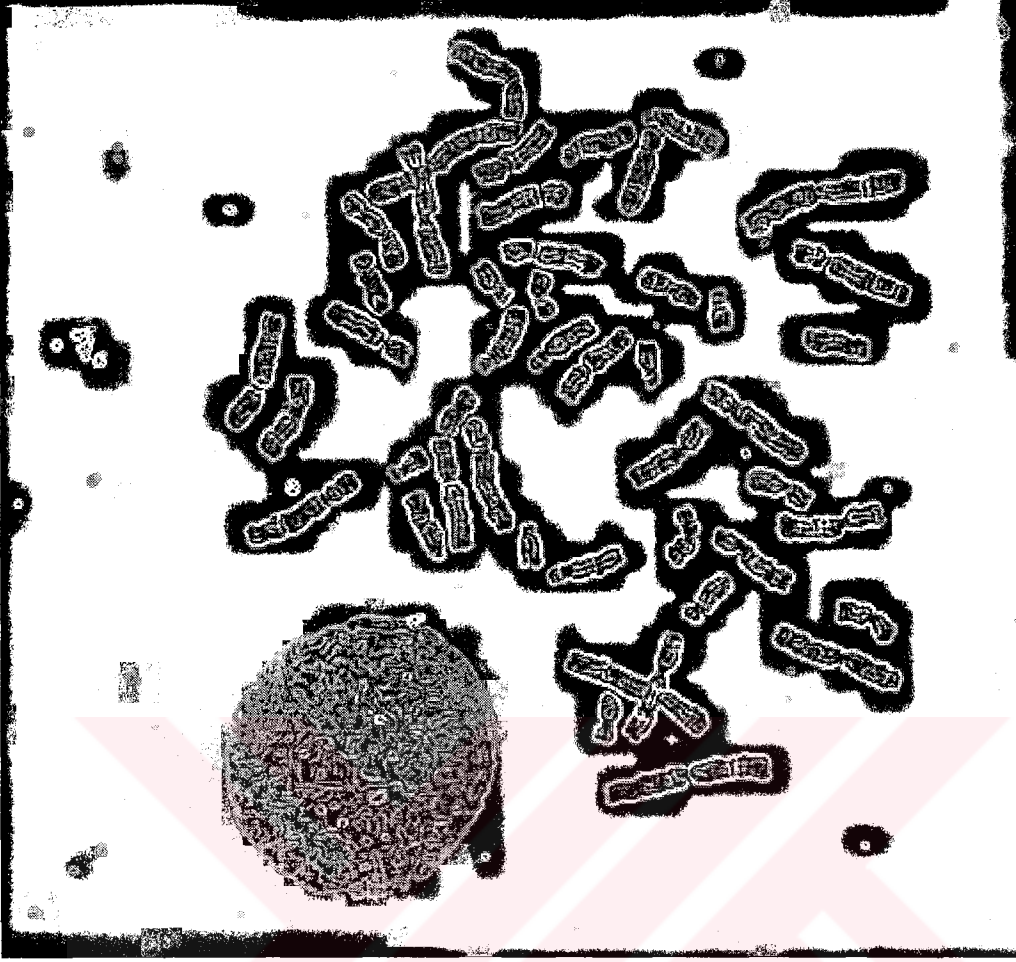
Bir noktanın komşularına her iki maske birden uygulanıp bunlardan elde edilen toplamın sonuçları Eşitlik 4.2'deki gibi birlikte değerlendirilerek noktanın yeni değeri belirlenir. Böylece, görüntü üzerindeki yatay ve dikey bütün kenarlar, yönleri dikkate alınmaksızın değerlendirilmiş olur. Şekil 4.3'de işlenmemiş bir kromozom görüntüsü, Şekil 4.4'de ise Kenar Belirleme işleminden sonra elde edilen çıkış görüntüsü görülmektedir.





Şekil 4.3 İşlenmemiş metafaz örneği

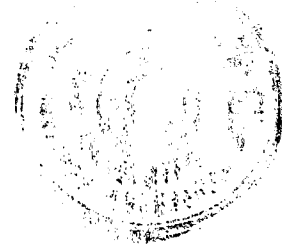




Şekil 4.4 Kenar belirleme işlemi sonucunda elde edilen metafaz örneği

4.1.1 Kenar belirleme işleminde eşik seviyesi kullanımı

Eşitlik 4.2’de hesaplanan genlik değeri doğrudan pikselin yeni değeri olarak kullanıldığında, elde edilen çıkış görüntüsünde kenarlar beyaza yakın bir gri tonu ile belirginleşirken, diğer değerler siyaha yakın ve farklı tonlarda gri değerlerde görülürler. Daha sonraki işlemlerde kullanılmak üzere, görüntü üzerindeki şekillerin yerlerinin kesin olarak belirlenebilmesi için kenar değerlerin beyaz, diğer bilgilerin siyah ile gösterildiği bir çıkış görüntüsü elde edilmelidir. Bunun için, elde edilen genliğin değeri bir T eşik seviyesi ile karşılaştırılarak bir p_n noktasının yeni değeri Eşitlik 4.3’deki gibi belirlenmelidir.



$$p_n = \begin{cases} 255 & p_n \geq T \\ 0 & p_n < T \end{cases} \quad (4.3)$$

Eşik seviyesinin belirlenebilmesi için, görüntünün gri seviyesi dağılımı bilinmelidir. Bir görüntünün gri seviyesinin dağılımını, histogramını hesaplayarak elde etmek mümkündür. Histogram, görüntü üzerindeki farklı gri seviyelerinin bağıl kullanım sıklığını gösterir. Histogram içinde yer alan herhangi bir gri seviyesinin bulunma olasılığı Eşitlik 4.4'deki gibi gösterilir (Gonzalez,1987,149).

$$p_r(r_k) = \frac{n_k}{n} \begin{cases} 0 \leq r_k \leq 1 \\ k = 0,1,\dots,L-1, \end{cases} \quad (4.4)$$

- L : gri görüntüdeki gri seviyesi sayısı
 $p_r(r_k)$: k. gri seviyenin bulunma olasılığı
 n_k : görüntüdeki k. seviyedeki nokta sayısı
 n : görüntüdeki toplam nokta sayısı

n adet görüntü için $p_r(r_k)$ değerleri, görüntünün histogramını verir. Histogram Eşitleme (Histogram Equalization) yöntemi ile, görüntünün uniform histogramı Eşitlik 4.5'de olduğu gibi elde edilir.

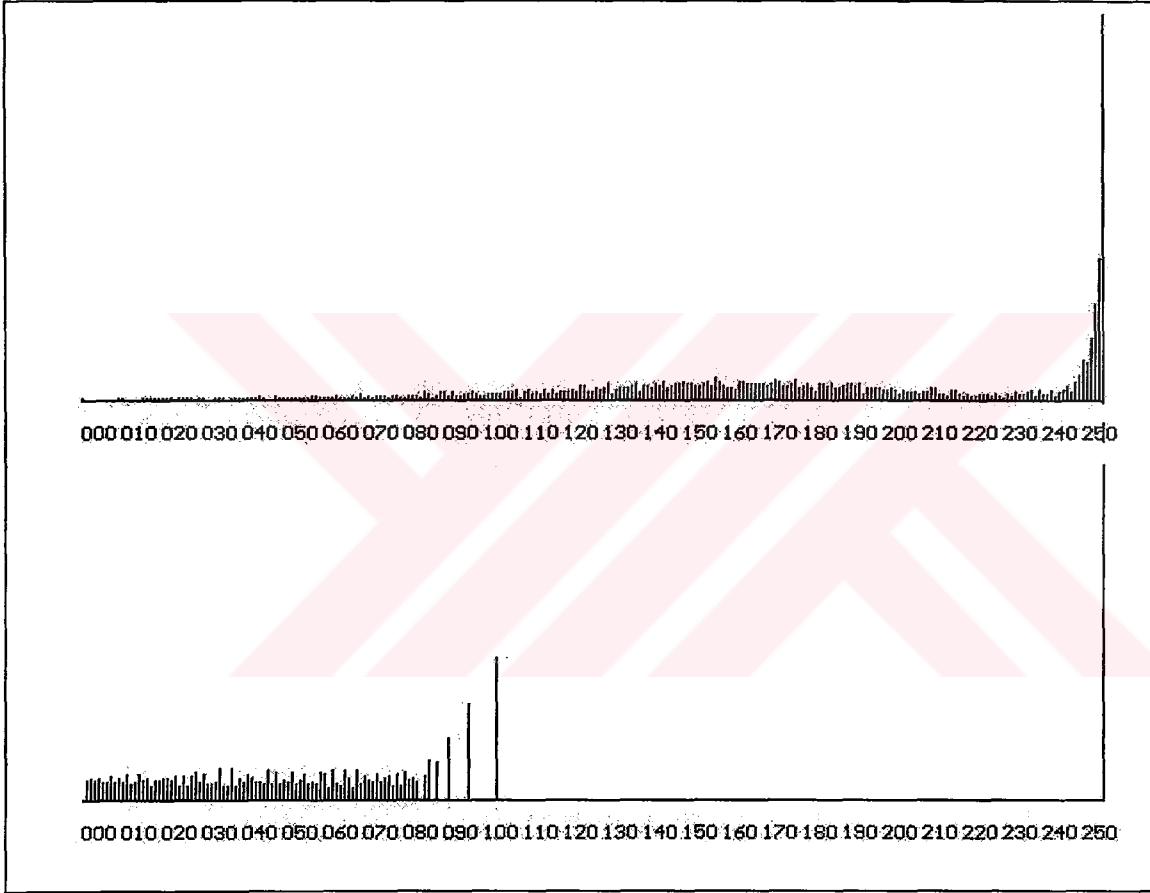
$$s_k = T(r_k) = \sum_{j=0}^k \frac{n_j}{n} \quad (4.5)$$

$$s_k = \sum_{j=0}^k p_r(r_j) \begin{cases} 0 \leq r_k \leq 1 \\ k = 0,1,\dots,L-1, \end{cases}$$

Hesaplanan s_k değerleri 0-1 arasındadır. Bu değerler, 255 gri seviyesi için 255 eşit aralık olması gerektiğinden kendilerine en yakın $\frac{x}{255}$ değerine dönüştürülürler. Bu şekilde kaç farklı s_k değeri elde ediliyorsa, her s_k değerine karşılık düşen yeni toplam piksel



değerleri hesaplanır. Örneğin s_3 'e karşılık düşen piksel değerleri $p_1 + p_2 + p_3$ toplamı ile elde edilir. Bu şekilde, Histogram Eşitleme işlemi sonucunda 255 gri seviyesi için elde edilen s_k değerlerinin 0'dan başlayıp artarak giderken maksimum olduğu gri seviyesi, sınır belirleme işleminde eşik seviyesi olarak kullanılmak üzere seçilir. Şekil 4.5'de üst kısımda 0-255 gri seviyesinde bir görüntünün histogramı, alt kısımda ise aynı görüntünün histogram eşitleme işlemi sonucunda elde edilen grafiği görülmektedir.



Şekil 4.5 Histogram ve Eşitlenmiş Histograma Örnek

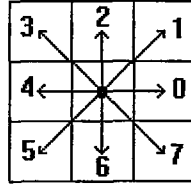
4.2 Sınır Belirleme

Bir şeklin çevresini takip eden ve birim büyüklükteki piksellerden oluşan piksel sırası, o şeklin ana hattını (outline) oluşturur. Her piksel konumuna göre, sağ, sol, yukarı, aşağı veya 4 diyagonalden biri aracılığıyla bir sonraki piksele bağlanır. Bir şekle ait bu özelliğe



sahip bütün pikseller belirlendiği zaman, o şeklin sınırı belirlenmiş olur. Sayısal görüntülerde şekli meydana getiren piksellerin, görüntü üzerindeki yerleri x,y değer çiftleri ile gösterilir. Şeklin sınırlarını gösteren kenar piksellerine ait bilgilerin daha sonra şeklin yapısını belirlemek için kullanılacağı düşünülecek olursa, bu piksellere ait x,y değerlerinin kullanılmasının hesaplama teknikleri ve algoritmalar açısından yeterli olmayacağı görülür. Bunun için, sınır değerlerine ait x-y değerlerinden başka, ardarda gelen sınır değerlerinin yön bilgilerini gösteren bilgiler de zincir kodlama yöntemi (chain coding) kullanılarak belirlenmiştir.

Zincir kodlama yönteminde amaç, sınır üzerindeki bir kenar pikseli bilinirken bir sonraki kenar pikselini bulabilmektir. Ana hat tanımından yola çıkarsak, bilinen kenar pikseline komşu sekiz pikseli inceleyerek bir sonraki kenar pikselinin yerini belirleyebiliriz. Kenar piksellerine, buldukları konumu ifade edecek şekilde, 0-7 arası değerler verilerek zincir kodu ifade edilir. Şekil 4.6'daki zincir koduna ait yön tanımlarında görüldüğü gibi bir sonraki kenar pikseli o anki kenar pikselinin sağında ise "0", sağ üst köşesinde ise "1", üstünde ise "2" v.b. değerleri ile ifade edilir. Kenar piksellerine ait yön sayısı belirleme işlemi bütün ana hat boyunca, sınır tanımlamasına başlanan piksele gelene kadar tekrarlandığı zaman şeklin sınır bilgilerine ait zincir kodu elde edilmiş olur.



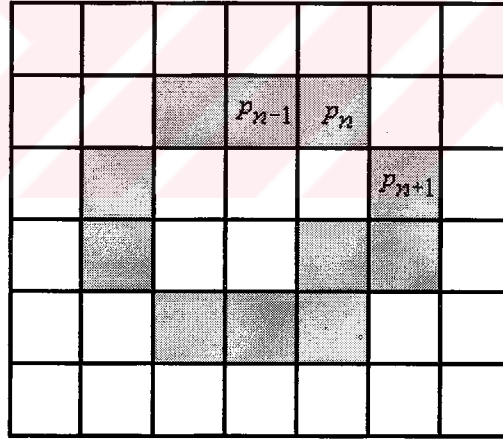
Şekil 4.6 Zincir koduna ait yön tanımları

Zincir kodu belirlenirken dikkat edilmesi gereken iki önemli bilgi vardır:

1. Bir önceki pikselin o anki piksele göre yönü;
2. Sınır belirlenmesine başlanan piksele ait x,y değerleri.



İlk bilgi sayesinde o anki kenar pikselinden bir önceki kenar pikseline yanlışlıkla dönülmesi engellenmiş olur. Şöyle ki, Şekil 4.7’de p_n noktası o anki, p_{n-1} noktası bir adım önceki kenar pikselleri olsunlar. Eğer p_{n+1} kenar pikselinin aranmasına ‘0’ yönünden başlanacak olursa yanlışlık sonucu p_{n+1} yerine p_{n-1} bir sonraki kenar pikseli olarak belirlenecektir. Halbuki, p_{n-1} yönünden p_n ’e doğru gelindiği için bir sonraki kenar pikselinin aranmasına ‘5’ yönünden itibaren başlanmalı ve “7” yönünde örnek şekil için doğru piksel olan p_{n+1} bulunmalıdır. Buna göre her adımda elde edilen yön bilgisini d_n ile ifade edecek olursak bir adım sonrası için arama işlemine $d_{n+1} = (d_n + 1) \bmod 8$ bağıntısından elde edilecek yönden başlanmalı ve saatin ters yönünde ilerlenerek arama işlemine devam edilmelidir. Burada saatin ters yönünde ilerlenmesi özellikle önem taşır. Zira saat yönünde ilerlenecek olursa şeklin iç bölgesindeki piksellere doğru gidilerek kontrol yapıldığından sonuçta elde edilen kenar bir piksel kalınlığında olmayacak ve yanıltıcı sonuçlar elde edilecektir.



Şekil 4.7 Sınır üzerinde yer alan piksellerin birbirlerine göre konumları

İkinci bilgi, bir adım sonra bulunan kenar pikselin, sınır belirlenmesine başlanılan piksele eşit olup olmadığının kontrolünü ve böylece kenar belirleme işleminin ne zaman sonuçlandırılacağını tanımlar. Bu yöntemle elde edilen zincir kodu, yönler için daima sabit bir orijine göre 0 ile 7 arası değerler verdiği için “mutlak zincir kodu” olarak adlandırılır.



4.3 Sınırları Belirlenmiş Bölgeye Ait Bilgilerin Elde Edilmesi

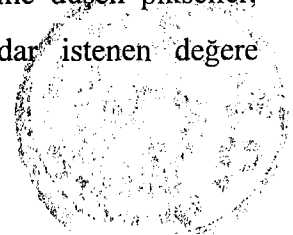
Bölüm 4.2’de kromozomların görüntü üzerindeki yerlerini belirleyen sınır bilgileri, x-y koordinat değerleri ve zincir kodları şeklinde elde edilmiştir. Kromozomun şekil bilgilerinin değerlendirilebilmesi için, sınır değerlerinden başka bu sınırlar içinde kalan değerlerin de bilinmesine ihtiyaç vardır. Bunun için, Bölüm 3.2’de anlatıldığı gibi, x-y koordinat çiftleri içinde x ve y değerlerinin minimum ve maksimumları bulunarak bu kromozomun görüntü içinde kapladığı alan belirlenmiştir. Bu alanın tamamını içine alacak şekilde $a = x_{\max} - x_{\min}$, $b = y_{\max} - y_{\min}$ boyutlarında $S(a,b)$ matrisi oluşturulmuştur. İlk değeri ‘0’ olan bu matriste sınıra karşılık gelen gözlere ‘1’değeri atanmıştır. Böylece $S(a,b)$ matrisinde kromozomun sınır değerleri saklanmıştır.

4.3.1 Kapalı bir bölgenin içinin doldurulması

Kromozomların şekil, yapı ve renk özelliklerinin incelenebilmesi için, sınır bilgilerinden başka, bu sınırlar içinde kalan değerlerin de bilinmesi gereklidir. Bir kromozomun iç bölgesindeki değerlere erişebilmek için, kapalı bir bölgenin içini dolduran bir algoritma kullanılmıştır.

Sınır bölgesinin içinde kalan alanı doldurmakta kullanılacak değerler, Bölüm 3.2’de anlatıldığı gibi, inceltme işlemi sonunda elde edilecek şekile ait iskelet değerlerinin saklanacağı $I(m,n)$ matrisi için ‘1’, kromozomun iç bölgesini oluşturan gri seviyesi değerlerin saklanacağı $B(m,n)$ matrisi için ise orijinal görüntüde kromozomun iç bölgesini oluşturan gri seviyesi değerleridir.

Sınır çizgisi üzerindeki piksel bilgilerinin aynı değerlere sahip olduğu, fakat sınırlar içindeki piksellerden farklı olduğu bir bölgenin içini doldurmak için, önce sınır bölgesi içinde kalan bir pikselin yeri belirlenmiştir. Bu pikselin yerini gösteren (x,y) değerlerinden sağdaki ve soldaki sınırlara çizili olduğu varsayılan çizgilerin üzerine düşen pikseller, sağdaki ve soldaki sınırları belirleyen piksellere erişilinceye kadar istenen değere



eşitlenerek ilerlenmiştir. Bu işlem sonucunda sağda ve solda erişilen en son piksellere ait x değerleri saklanmıştır. Bu değerler sağ taraf için x_r , sol taraf için x_l olarak gösterilecek olursa x_l 'den x_r 'ye giden hat boyunca yukarıda anlatılan yöntem rekürsif olarak altta ve üstte kalan pikseller için de tekrarlandığı zaman sınırları belli bölgenin iç kısmındaki piksellerin tamamı istenen değere eşitlenmiş olur.

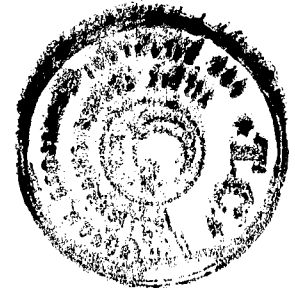
4.3.2 Bir şeklin iskeletinin çıkarılması

Belirli biçimdeki bir bölgeyi ifade etmek için en uygun yöntemlerden biri o bölgeye ait grafi belirlemektir. Bu işlem çoğunlukla o bölgenin inceltme (thinning) algoritması ile inceltirilerek iskeletinin çıkartılması sayesinde gerçekleştirilir. Bir bölgeye ait iskelet, MAT (medial axis transformation) yardımıyla tanımlanabilir. Buna göre, B sınırlarına sahip R bölgesinin MAT'ı R bölgesindeki her p noktasına ait B sınırındaki en yakın noktanın hesaplanması ile ifade edilir. (Blum, 1967) Burada “en yakın” kavramı, MAT uygulamasını gerçekleyen kişinin tanımına kalmıştır. MAT, prensip itibariyle bir bölgeye ait en ideal iskeleti vermesine rağmen, “en yakın” kavramının tanımından ötürü bilgisayara uygulanabilirlik açısından zordur. Bu sebeple bilgisayara uygunluk açısından kullanılan algoritmalar,

- uç noktaların silinmemesi
- var olan bağlantıların kopartılmaması
- bölge içinde aşırı erimeye sebebiyet verilmemesi

prensiplerine dayalı olmaktadır.

Bu çalışmada Zhung ve Suen (1984) tarafından geliştirilen algoritma kullanılmıştır. Bu algoritma iki adımdan oluşur.



| | | |
|-------|-------|-------|
| p_9 | p_2 | p_3 |
| p_8 | p_1 | p_4 |
| p_7 | p_6 | p_5 |

Şekil 4.8 İnceltme algoritmasında kullanılan komşuluk matrisi

Şekil 4.8’de komşuları ile birlikte verilen bir p_1 noktasının inceltme işleminin birinci adımının sonucu silinmeye aday olması,

- a) $2 \leq N(p_1) \leq 6$
- b) $S(p_1) = 1$
- c) $p_2 * p_4 * p_6 = 0$
- d) $p_4 * p_6 * p_8 = 0$

koşullarına bağlıdır. Burada $N(p_1)$, p_1 ’in 0’dan farklı komşularının sayısı, yani $N(p_1) = p_2 + p_3 + \dots + p_8 + p_9$ şeklinde, $S(p_1)$ ise $p_2, p_3, \dots, p_8, p_9$ sıralı dizisindeki 0-1 geçişlerinin sayısı olarak tanımlanmıştır.

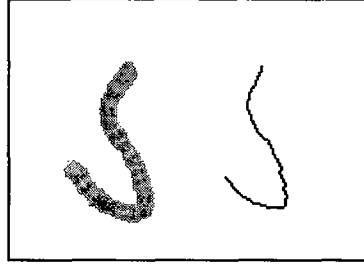
İkinci adımda birinci adıma ait (a) ve (b) şartları aynen kalmasına rağmen (c) ve (d) şartları,

- (c) $p_2 * p_4 * p_8 = 0$
- (d) $p_2 * p_6 * p_8 = 0$

şeklinde değişirler. Birinci adım iskeleti çıkartılacak bölgedeki her kenar piksele uygulanır. Eğer her dört şart da sağlanıyorsa o piksel silinmek için işaretlenir. Tüm kenar pikselleri böylece işlendikten sonra, silinmek için işaretlenenler silinir ve geriye kalan piksellere ikinci adım uygulanır. İkinci adımda da aynen birinci adımda olduğu gibi işlem yapılır. Birinci ve ikinci adımlar dönüşümlü olarak silinmeye aday piksel çıkmayınca kadar tekrarlanır.



Bu çalışma açısından inceltme ve iskelet çıkartma algoritmaları, kromozomların metafaz halinde iken buldukları açı, eğim, diklik, eğiklik gibi unsurlardan bağımsız olarak değerlendirilebilmelerini sağladığı için önem taşımaktadırlar. Kromozomların simetri ekseninden yola çıkılarak yapılacak işlemlerin açı, eğim, diklik, eğiklik gibi unsurlardan etkilenmeyeceği göz önünde tutularak, iskelet çıkartma algoritması ile kromozomlara ait iskelet, yani kromozomun ekseni elde edilmiştir. Şekil 4.9'da sol taraftaki kromozomun inceltme işlemi sonucunda elde edilen iskeleti, sağ taraftaki gibi elde edilmiştir.



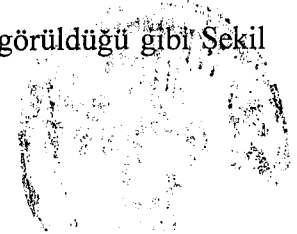
Şekil 4.9 İnceltme işleminin sonucu

4.4 Kromozom Bilgilerinin Yönden ve Şekil Özelliklerinden Bağımsız Olarak Değerlendirilmesi

Bu çalışmada, kromozom bilgilerinin yönden ve şekil özelliklerinden bağımsız olarak değerlendirilebilmelerini sağlamak için kromozomların genişlik ve yoğunluk profilleri kullanılmıştır.

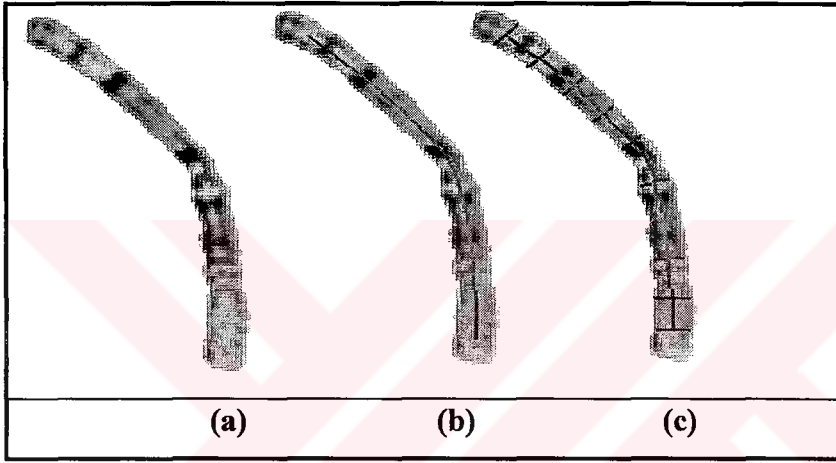
4.4.1 Genişlik ve yoğunluk profilleri

Bu çalışmanın kapsamında profil terimi, kromozoma ait herhangi bir özelliğin (genişliği, renk dağılımı vb.) kromozomun simetri eksenini boyunca gösterdiği dağılımı tanımlamaktadır. Bu amaçla, iskeleti belirlenmiş kromozomdan sınıflandırma işlemi esnasında ihtiyaç duyulan, sentromerin yerinin ve bant özelliklerinin elde edilmesini kolaylaştırmak amacıyla, kromozoma ait genişlik ve yoğunluk profillerinin çıkarılması uygun görülmüştür. Profil elde etme işlemi esnasında, Şekil 4.10'da görüldüğü gibi Şekil



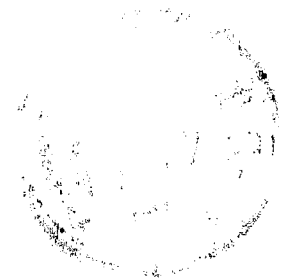
4.10.b'deki iskelete dik olarak Şekil 4.10.c'deki gibi sağa ve sola çizildiği varsayılan çizgiler boyunca ilerlenerek elde edilen piksel değerleri satır bilgisi olacak şekilde, matris formunda saklanmıştır (Şekil 4.10.c'deki kromozom eksenine dik çizilen çizgiler gerçekte eksen üzerinde var olan bütün noktalara uygulanmakla birlikte, şeklin anlaşılır olması için aralıklı olarak gösterilmiştir).

Bu sayede kromozoma ait bilgilerin kromozom düz ve dik olduğunda erişilebileceği düzene getirilmesi sağlanmıştır.



Şekil 4.10 Kromozom iskeletinden bant bilgilerinin elde edilmesi

Profil çıkarma sonucunda elde edilen bilgiler sentromer yerinin tespitinde kullanılmak üzere gerekli olan **genişlik profilini** ve bant özelliklerinin belirlenmesinde kullanılmak üzere **yoğunluk profilini** elde etmek amacıyla kullanılırlar. **Genişlik profili**, incelenen kromozoma ait bilgilerin bulunduğu matristeki her satıra ait eleman sayılarının, toplam satır adedi uzunluğunda tek boyutlu bir vektöre yerleştirilmesi ile elde edilir. **Yoğunluk profili** ise her satıra ait grilik yoğunluğu değerlerinin aritmetik ortalamasının alınması ve bu ortalamaların toplam satır adedi uzunluğunda ikinci bir tek boyutlu vektöre yerleştirilmesi ile elde edilir.



4.5 Kromozomların Sınıflandırılması ve Tanınmasında Kullanılan Kriterlerin Elde Edilmesi

Kromozomların sınıflandırılma ve tanınmasında kullanılacak özellikler olan boy-alan, sentromer özelliği ve bant bilgileri genişlik ve yoğunluk profilleri değerlendirilerek bulunur.

4.5.1 Kromozomun alanının ve boyunun bulunması

Kromozomlar, görüntünün elde edildiği cihazın (kamera, tarayıcı, vb.) x-y eksenine rasgele açılarla ve değişik formlarda (düz, kıvrık, vb.) yer alırlar. Kromozomları geometrik olarak düz ve x-y eksenin dik hale getirmek için yapılan işlemlerin hem çok uzun ve zahmetli hem de bilgi kaybına sebep olabileceği düşünülerek, bu çalışmada kromozomların boyu, iskelet bilgisi üzerinden piksel sayma yöntemi ile belirlenmiştir.

Değerlendirilen farklı kromozom görüntüleri incelendiğinde, görüntünün alınışına bağlı olarak, kromozom boylarının görüntüden görüntüye değiştiği görülmüştür. Bu sebeple, kromozomun boy bilgisini (iskeletini oluşturan nokta sayısı) doğrudan kullanmak yerine, her kromozomun boyu, aynı görüntüdeki bütün kromozomların boylarının toplamına bölünerek normalizasyon yapılmış ve bu sayede oransal kromozom boyu elde edilmiştir.

Kromozom görüntülerinin değerlendirmesi sırasında kromozomun şeklindeki aşırı bozuklukların (başka kromozomlarla üst üste gelmesi, çok kıvrık durması vb.) iskeleti etkilediği ve bunun sonucu olarak boy bilgisinin her zaman yeterince hassas bir kriter olmadığı tespit edilmiştir. Bu durumda yapılacak sınıflandırmanın doğruluğunu arttırmak ve bu cins etkenlere karşı daha az duyarlılıkla daha kesin sonuçlar elde etmek amacıyla boy bilgisinin yanı sıra kromozomlara ait alan bilgisinin de kullanılması uygun görülmüştür. Bir kromozoma ait alan bilgisi, kapladığı bölge içindeki toplam piksel sayısı ile ifade edilebilir. Burada da dikkat edilmesi gereken nokta, alan bilgisinin de aynen boy bilgisinde olduğu gibi görüntüdeki tüm kromozomlara ait alan toplamına oranlanmak sureti ile normalize edilerek oransal hale getirilmesi gerektiğidir.



4.5.2 Kromozomun sentromerinin bulunması

Bir kromozomun sentromerinin, kromozomun genişliğinin en az olduğu yer olduğu bilindiğine göre, kromozoma ait genişlik profilinin satır satır taranması sonucunda, en az sayıda elemana sahip olan satır, sentromerin yerini verecektir. Sentromer bilgisi, kromozomun metasentrik, akrosentrik veya submetasentrik kromozom gruplarından hangisine ait olduğunun belirlenmesini sağlar. Kromozomun toplam boyu, sahip olduğu iskeletin piksel sayısı cinsinden ifade edilecek olursa sentromerin kromozom boyundan bağımsız, sınıflandırmaya uygun bağıl yeri, sentromerin kromozomun p-kolunun başlangıcına olan uzaklığının, p+q uzunluğuna bölünmesi ile elde edilir.

4.5.3 Kromozomun p- ve q-kollarının belirlenmesi

Bölüm 2.3'de anlatıldığı gibi sentromer, kromozomu kısa olanı p-kolu ve uzun olanı q-kolu olarak isimlendirilen ve p-kolunun daima üst kısımda bulunduğu iki ayrı kola böler. Sentromerin kromozomun iki uç noktasına olan uzaklığı hesaplanarak bulunan değerlerden küçük olanı p-kolunu, büyük olanı ise q-kolunu gösterir. p-kolunun tesbit edilmesi, kromozomun başlangıcının görüntü içindeki yerinden, yönünden ve konumundan bağımsız olarak belirlenmesini sağlar. Kromozomun başlangıcının nerede olduğunun bilinmesi, bantların doğru değerlendirilebilmesi için gereklidir.

4.5.4 Kromozomun bant bilgilerinin bulunması ve değerlendirilmesi

G-Bantlama tekniği ile boyanmış kromozomların, boyama işlemi sonucunda üstündeki ve altındaki parçalardan renginin koyuluğu veya açıklığı ile ayrılan bölgeleri bant olarak isimlendirilir.

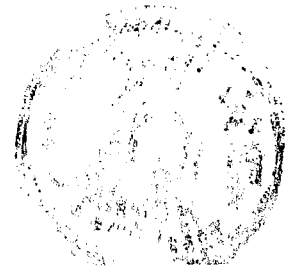
Kromozomlardaki bantların grilik seviyeleri görüntünün alındığı şartlara bağlı olarak değiştiğinden, hatta aynı görüntüde bazı kromozomlar daha açık renklerde bazıları daha koyu renklerde görülebildiğinden farklı kromozom görüntüleri ve aynı görüntüdeki farklı

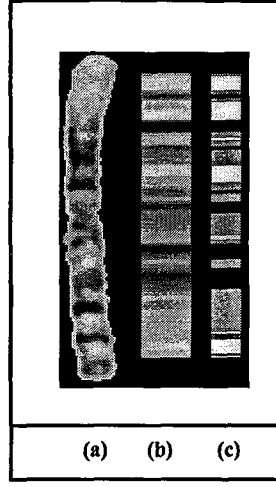


kromozomlar için bant geçişlerini gösteren sabit gri seviyesi değerleri belirlemek mümkün değildir. Bunun yerine, her görüntü için bant geçişlerini adaptif olarak belirlemek gereklidir. Kromozomların gri seviyesi bilgilerini içinde buldukları görüntüden bağımsız olarak değerlendirilerek grilik seviyesi geçişlerini belirlemek amacıyla aşağıdaki yöntem kullanılmıştır:

1. Farklı kromozom görüntüleri incelendiğinde, bir kromozomun bant geçişlerini ayırt edebilmek için 0-255 arasında değişen gri seviyesi değerlerinin hepsinin ayrı ayrı değerlendirilmeyip, bantların **siyah**, **gri**, **beyaz** olmak üzere üç temel grupta incelenmesine karar verilmiştir.
2. Bölüm 4.1.1'de kullanılan *Eşitlenmiş Histogram* yöntemi, her kromozomun yoğunluk profilindeki değerlere uygulanarak, 255 gri seviyesi için beyaz-gri ve gri-siyah geçişlerini gösterecek 2 ayrı eşik seviyesi (t_1 ve t_2) belirlenir.
3. Yoğunluk profilindeki tüm değerler, t_1 ve t_2 değerleri ile karşılaştırılarak, t_1 'den küçük değerler **siyah** bant bölgelerini, t_1 'den büyük fakat t_2 'den küçük değerler **gri** bant bölgelerini, t_2 'den büyük değerler ise **beyaz** bant bölgelerini gösterecek şekilde kromozomların bantlarının yerleri belirlenmiş olur. Kromozomları tek tek tanımak için kullanılacak bu bant bilgileri siyah 0, gri 1, beyaz 2 ile gösterilerek saklanır.

Şekil 4.11.a'daki kromozomun, yoğunluk profili Şekil 4.11.b'de görüldüğü gibi elde edilmiştir. Bant belirleme işlemi sonucunda bant bilgilerinin gri seviyelerinin Histogram Eşleme işleminden geçirilmesi sonucunda gri seviyesi değerlerinin siyah, beyaz ve gri değerlere dönüştürülmesi sonucunda elde edilen yeni bant bilgilerinin şematik gösterimi ise Şekil 4.11.c'deki gibidir.



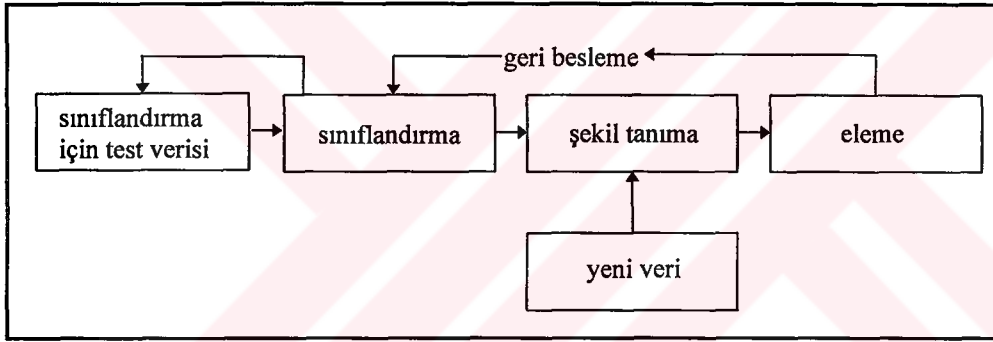


Şekil 4.11 Bant bilgileri elde edilen kromozomun şematik gösterimi



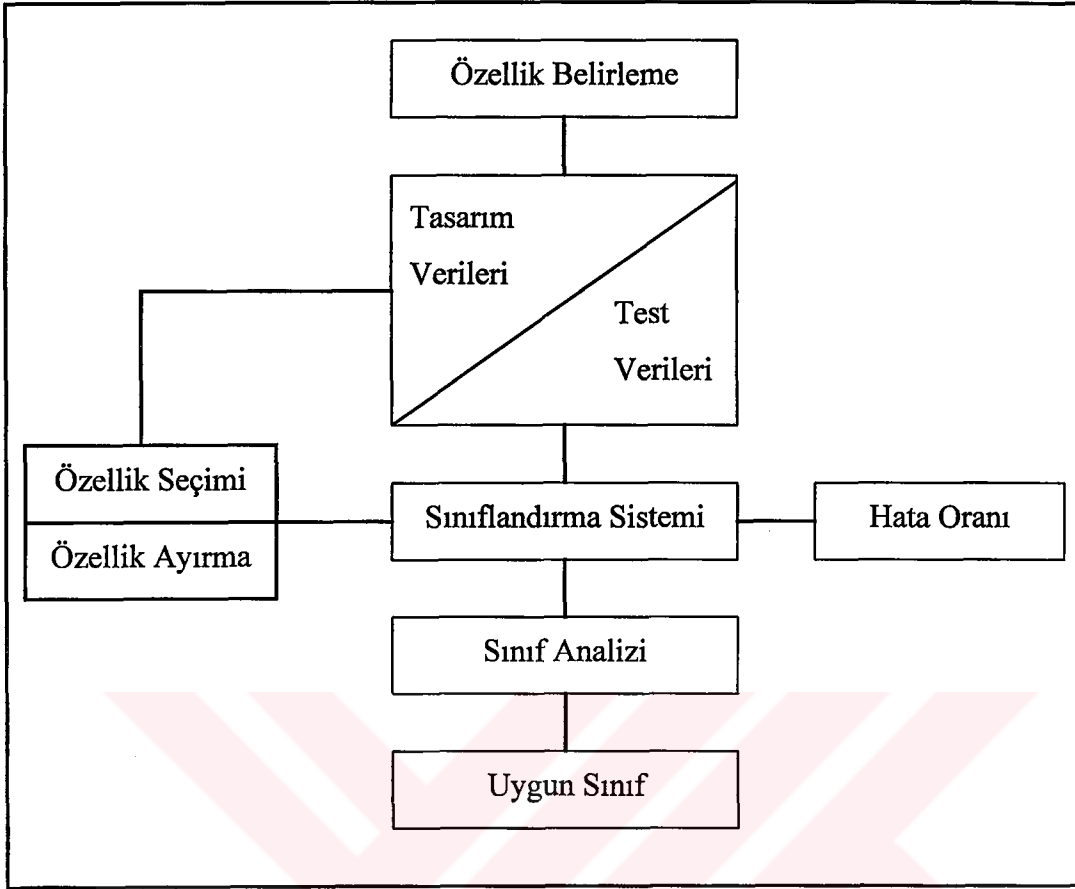
5. BULANIK MANTIK YÖNTEMLERİ İLE KROMOZOMLARIN SINIFLANDIRILMASI VE TANINMASI

Bir **şekil tanıma** işleminin ilk adımı, tanınması istenen bir örneği, özelliklerinin uygun olduğu sınıfa yerleştirmektir. Bu örneğin karşılaştırılacağı sınıfların yapısı ve özellikleri ise **sınıflandırma** işlemi ile belirlenir. Sınıflandırma ve şekil tanıma sistemlerinin bağlantısını Şekil 5.1'deki gibi göstermek mümkündür. Buna göre sınıflandırma işlemi, gelen veriyi inceleyerek veriyi oluşturan yapıyı belirlerken, şekil tanıma işlemi de yeni gelen bir verinin sınıflandırma işlemi sonucunda ortaya çıkan sınıflardan hangisine ait olduğunu belirler. Sistemdeki geri besleme, sınıflandırma işleminin daha duyarlı olması istendiğinde veya şekil tanıma işleminde başarılı bir sonuç elde edilemediğinde kullanılır.



Şekil 5.1 Sınıflandırma ve şekil tanıma sistemlerinin bağlantısı

Bir şekil tanıma sistemi şematik olarak Şekil 5.2'deki gibi gösterilebilir. Bu sistemin başarısı, sınıflandırma sisteminde kullanılan özelliklerin seçimi ile yakından ilgili olduğundan sınıflandırma kriteri olarak, örneklerin belirgin şekilde birbirinden ayırt edilebilmelerini sağlayacak özellikler seçilmelidir.



Şekil 5.2 Bir şekil tanıma sisteminin şematik yapısı

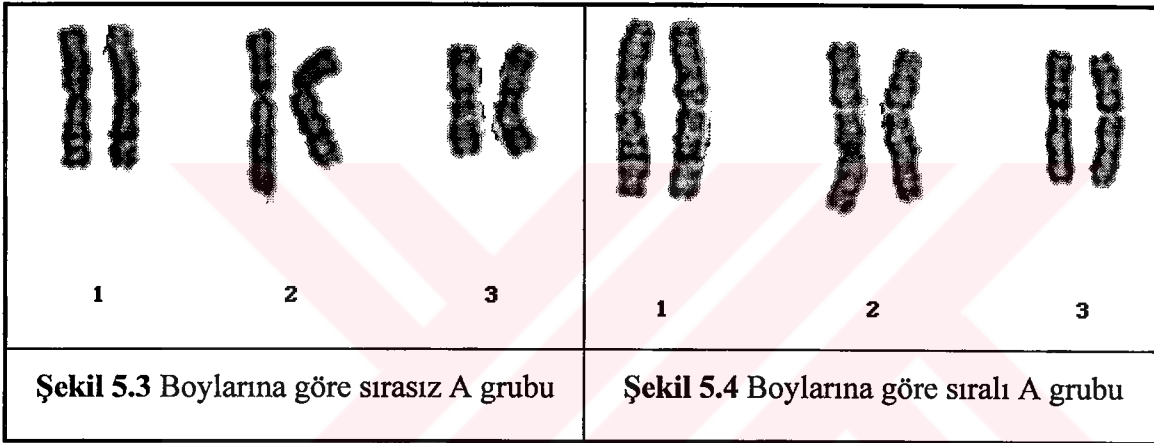
Bu tez çalışmasında geliştirilen Kromozom Sınıflandırma ve Tanıma Sisteminde, daha önceki bölümlerde de değinildiği gibi, kromozomları sınıflandırmada kullanılacak özellikler **boy, alan, sentromer özelliği ve bant bilgileri** olarak seçilmiştir. Bölüm 2’de kromozomların yapısal özelliklerinin tanıtımında da ayrıntılı olarak anlatıldığı gibi boy, alan ve sentromer yeri bilgileri kromozomları gruplara ayırmakta kullanılan özelliklerdir. Bu nedenle, bu çalışmada kromozomlar önce boy, alan ve sentromer özelliği bilgilerine göre gruplara ayrılmış, daha sonra 24 farklı kromozom için hazırlanan bant bilgisi şablonları her grup için ayrı ayrı kullanılarak kromozomların tek tek tanınması sağlanmıştır.

Kromozomlar şekil özelliklerine göre yedi gruba (A-G) ayrılırlar. Kromozomlar bu gruplar içinde boylarına göre büyükten küçüğe sıralı olarak bulunurlar. Kromozomlar



sınıflandırma özelliklerinin belirlenmesi için çok sayıda kromozom görüntüsü üzerinde yapılan incelemelerde kromozomların oransal boy değerlerinin sabit olmadığı, görüntüden görüntüye değişebildiği, daha kısa boylu olması gereken bir kromozomun daha uzun olabildiği görülmüştür.

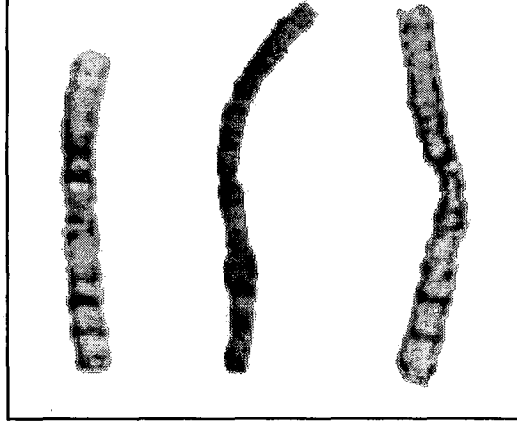
Şekil 5.3'deki A grubu kromozomlarının boyları incelendiğinde 2 numaralı kromozomlardan ilkinin boyunun 1 numaralı kromozomlardan daha uzun olduğu, Şekil 5.4'de ise kromozomların boylarının 1 numaralı kromozomdan 3 numaralı kromozoma doğru azaldığı görülmektedir.



Örnek kromozom görüntülerinin sentromer oranları değerlendirildiğinde, metasentrik, submetasentrik ve akrosentrik kromozomlar için belirlenen ideal oranlardan farklılıklar gösterebildiği belirlenmiştir.

Bant bilgileri incelendiğinde, Bölüm 3.2.4'deki nedenlere bağlı olarak bir kromozomun bant bilgilerinin her zaman aynı sayıda ve şekilde elde edilemediği görülmektedir. Şekil 5.5'de G-Bantlama tekniği ile boyanarak elde edilmiş farklı kromozom görüntülerinden seçilmiş 1 numaralı kromozomlar görülmektedir. Görüntüler incelendiğinde her birinde kromozomların farklı yerlerinde ve farklı sayıda bant olduğu görülür.





Şekil 5.5 Farklı 1 numaralı kromozomlara ait bant bilgileri

Kromozomların ayırt edici özellikleri olan boy, sentromer yeri ve bant özelliklerinde görülen bu belirsizlikler göz önüne alındığında, kromozomların sınıflandırılmasında sınıf yapılarının kesin özelliklere sahip olduğu klasik sınıflandırma yöntemlerinin kullanılmasındansa kesin olmayan, yaklaşık bilgilerin kabul edilir olduğu “bulanık (fuzzy) sınıflandırma” sisteminin geliştirilmesinin daha uygun olacağı görülmüştür.

5.1 Kromozomların Bulanık Mantık Kümeleri

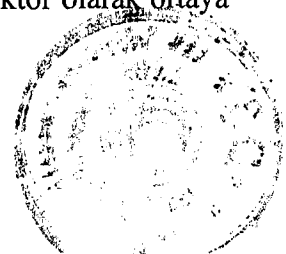
Kromozomların sınıflandırma yöntemiyle ayrılmalarında bulanık sınıflandırma metotlarından yararlanmak üzere 22 çift ile x ve y 'den oluşan kromozom kümesini :

$$X = \{x_1, x_2, x_3, \dots, x_n\}$$

şeklinde ifade edebiliriz. Burada bir x_i elemanı m değişik özelliğe sahip ise;

$$x_i = \{x_1, x_2, x_3, \dots, x_m\}$$

olarak gösterilir. Bu ifadelerde x_i , (x) uzayında m boyutlu, n elemanlı vektör olarak ortaya çıkar.



Kromozomlarda ayırt edici özellik olan alan, boy, sentromer ve bant rengi bilgileri

$$F = \{\text{alan, boy, sentromer, bant}\}$$

esas alındığında,

$$x_i = \{x_{i1}, x_{i2}, x_{i3}, x_{i4}\}$$

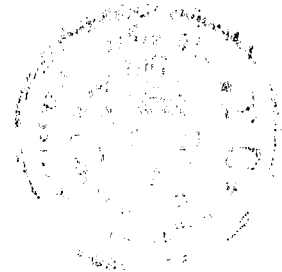
olarak dört boyutlu özellikler uzayında 46 elemandan oluşan bir kümenin incelenmesi gerekir. Böylece sınıflandırma yoluyla ayırımı yapılacak kromozom kümesi elemanlarına, sahip oldukları özellikler için, benzeme oranlarına eşdeğer olan **bulanık üyelik derecesi** verilir.

Bir kromozomun görüntü olarak x-y eksen sistemindeki konumuna, yani simetri ekseninin yatay, düşey veya eğik olarak bulunmasına bakılmaksızın en az hatayla ölçülebilen özelliği alanıdır. Bir kromozom kümesindeki her elemanın alanı, görüntünün alındığı duruma bağlı olarak değişir. Bu değişimden gelecek hatayı önlemek için ölçüm ile elde edilen değerlerin normalize edilmesi gerekir. Bu çalışmada, her kromozomun alanını tek tek normalize etmek yerine, her kromozomun alan değerini, 46 kromozomun toplam alanına oranlayarak oransal alan değerleri hesaplanmıştır.

Aynı şekilde, kromozomların ayırt edici özelliklerinden bir diğeri olan boyları için de, her kromozomun boyunun 46 kromozomun toplam boy değerine oranları hesaplanarak oransal boy değerleri kullanılmıştır.

Kromozomların üçüncü belirleyici özellikleri olan sentromer yerlerinin değeri ise kromozomun p-kolu ve q-kolu uzunlukları bulunduktan sonra, p-kolu/(p-kolu+q-kolu) oranı kullanılarak hesaplanmıştır.

Bu üç ana özelliği Eşitlik 5.1'deki notasyonla gösterebiliriz.



$$Alan = \{A_1, A_2, \dots, A_n\}$$

$$Boy = \{L_1, L_2, \dots, L_n\}$$

$$Sen_p = \{p_1, p_2, \dots, p_n\}$$

$$Sen_q = \{q_1, q_2, \dots, q_n\}$$

(5.1)

A_i : i. kromozomun alanı

L_i : i. kromozomun boyu

p_i : i. kromozomun p-kolunun uzunluğu

q_i : i. kromozomun q-kolunun uzunluğu

Buna göre;

$$A_T = \sum_{i=1}^n A_i \quad \% a_i = 100 * \frac{A_i}{A_T} \quad \text{oransal alan,}$$

$$L_T = \sum_{i=1}^n L_i \quad \% l_i = 100 * \frac{L_i}{L_T} \quad \text{oransal boy,}$$

$$\% s_i = 100 * \frac{p_i}{p_i + q_i} \quad \text{sentromer oranı}$$

bağıntıları çıkarılır.

Böylece, $X = \{x_1, x_2, x_3, \dots, x_n\}$ ve $x_i = \{x_{i1}, x_{i2}, x_{i3}, \dots, x_m\}$ genel bağıntılarında $m=3$ olup;

$x_i = \{x_{i1}, x_{i2}, x_{i3}\}$ şeklini alır.

Test Kümesi olarak isimlendireceğimiz kromozom kümesinin elemanlarının boy, alan ve sentromer özelliklerine göre sınıflandırılması, her kromozomun bu 3 özelliğinin diğer kromozomlara göre değerlendirilmesi ile yapılmıştır. Sınıflandırma işlemi sonunda kromozomlar 7 gruba (A-G) yerleştirilmiştir. Kromozomları ayrı ayrı tanımak, yani her kromozomun kaç numaralı kromozom olduğunu bulmak için, her kromozomun bant bilgileri 30 kromozom örneğinin bant bilgilerinden elde edilen şablonlarla karşılaştırılarak değerlendirilmiştir.



Bu çalışmada ayırt edici özelliklerde örnek kromozom görüntülerinden hesaplanan değerler kullanılmıştır. Bu değerlerin ideal değerlere göre karşılaştırılmasının yapılabilmesini kolaylaştırmak amacıyla Tablo 5.1’de idiogramlardan elde edilen boy, alan, sentromer ve bant bilgileri verilmiştir.



Tablo 5.1 Kromozomların İdiogram Bilgisine Göre Hesaplanan Oransal Bilgileri

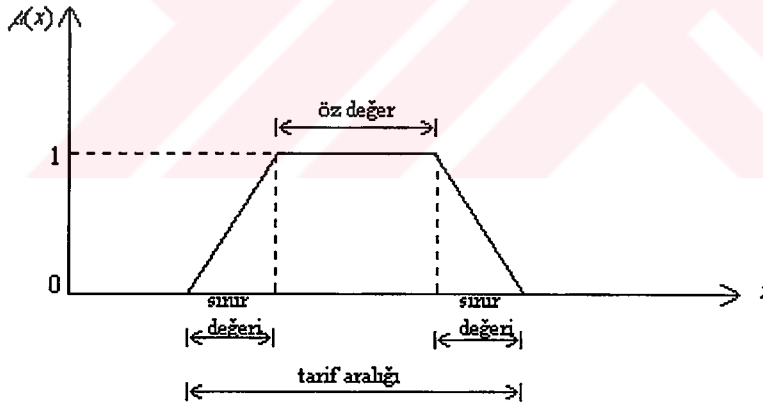
| Kromozom Grupları | A | A | A | B | B | C | C | C | C | C | C | D | D | D | E | E | E | F | F | G | G | X | Y | |
|------------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Kromozom Numaraları | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | X | Y |
| Alan Oranları (%) | 8.41 | 7.86 | 6.54 | 6.15 | 5.88 | 5.64 | 5.10 | 4.56 | 4.47 | 4.47 | 4.45 | 4.20 | 3.32 | 3.20 | 3.02 | 2.89 | 2.75 | 2.56 | 2.35 | 2.09 | 1.54 | 1.50 | 4.00 | 2.09 |
| Boy Oranları (%) | 8.79 | 8.37 | 6.75 | 6.32 | 6.03 | 5.77 | 5.28 | 4.59 | 4.50 | 4.50 | 4.48 | 4.19 | 3.24 | 3.10 | 2.91 | 2.76 | 2.60 | 2.40 | 2.17 | 1.90 | 1.30 | 1.25 | 4.80 | 1.90 |
| Sentromer Oranları (%) | 48 | 38 | 46 | 24 | 24 | 36 | 35 | 28 | 31 | 30 | 39 | 24 | 6 | 4 | 5 | 42 | 27 | 21 | 40 | 46 | 10 | 11 | 37 | 18 |
| Bant Sayıları | 11 | 6 | 8 | 3 | 4 | 5 | 4 | 4 | 3 | 3 | 5 | 3 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 5 | 1 |
| Üst | | | | | | | | | | | | | | | | | | | | | | | | |
| Alt | 8 | 12 | 8 | 10 | 11 | 10 | 7 | 7 | 6 | 9 | 9 | 8 | 7 | 6 | 8 | 3 | 4 | 4 | 2 | 2 | 2 | 3 | 7 | 3 |
| Toplam | 19 | 18 | 16 | 13 | 15 | 15 | 11 | 11 | 9 | 12 | 14 | 11 | 8 | 7 | 9 | 5 | 5 | 5 | 3 | 3 | 3 | 4 | 12 | 4 |



5.2 Bulanık (Fuzzy) Üyelik Derecesi

Bulanık Kümeler Teorisinde, belli bir özelliğe göre oluşmuş bir sınıfın bir elemanı (bir üyesi) esas alındığında, diğer üyelerin bu elemana olan benzerliği (μ) **üyelik derecesi** ile gösterilir. İncelenen küme bir (A) kümesi ise, bu kümedeki bir x elemanın üyelik derecesi $\mu_A(x)$ olarak gösterilir. Tam benzemede $\mu = 1$, hiç benzememe halinde $\mu = 0$ olarak alınır. Bu durumda $0 \leq \mu \leq 1$ olur. (μ) üyelik derecesi $\mu(x) = f(x)$ şeklinde bir sürekli fonksiyonla ifade edilebildiği gibi (x)'in belli değerlerine karşı kesikli olarak da gösterilebilir.

$\mu(x) = f(x)$ sürekli fonksiyonunda (x) bağımsız değişkeninin belli değerler (sınır değerleri) aralığında tarifli olması gerekir. Şekil 5.6'da bir $\mu(x)$ fonksiyonunun sınır değerleri, plato şeklindeki ($\mu(x) = 1$) öz değer ve sınır değerleri görülmektedir.



Şekil 5.6 Bir Üyelik Derecesi Fonksiyonu



5.3 Üyelik Derecesi Değerleri

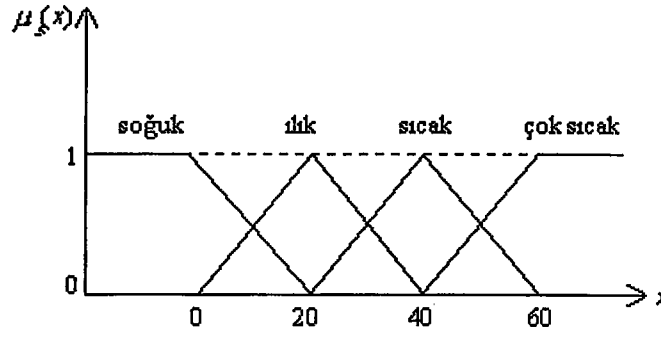
Bulanık Mantık değişkenlerine üyelik derecesi atamaları çok değişik şekillerde yapılır. Bu değer atamaları sezgisel yoldan olabildiği gibi bazı algoritmik ve mantıksal işlemlerle de gerçekleştirilebilir. Bulanık Mantık değişkenlerine üyelik derecesi değeri atama için kullanılan bazı yöntemler şöyle sıralanabilir:

- Sezgisel Yaklaşım
- Çıkarım
- Tümevarım
- Genetik Algoritmalar
- Bulanık Mantık İstatistiği
- Yapay Sinir Ağları
- Derecelendirme

Bu tez çalışmasında üyelik derecelerinin belirlenmesi için sezgisel yaklaşım, tümevarım ve bulanık mantık istatistiği yöntemleri kullanılmıştır.

Sezgisel yaklaşımda, uygulanacak alanın durumuna göre (x) değişkeninin fiziksel yapısına bağlı olarak değişim gösterecek bir $\mu(x)$ fonksiyonu tarif edilir. Şekil 5.7'deki örnekte "sıcaklık" kavramıyla ilgili olarak soğuk, ılık, sıcak ve çok sıcak sınıflarına ait $\mu_s(x)$ üyelik derecesi fonksiyonlarının sıcaklık derecesi (C°) değerlerine bağlı olarak değişimleri görülmektedir.





Şekil 5.7 Sıcaklık Üyelik Dereceleri

5.3.1 Kromozomun alan ve boy üyelik derecelerinin elde edilmesi

Kromozomların alan ve boy özellikleri için çıkarılacak üyelik derecesi fonksiyonlarında göz önünde bulundurulması gereken üç önemli nokta şunlardır:

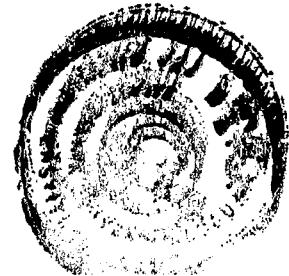
1. Her kromozom kendine benzer: $\mu_{ij} = 1$
2. (i) ve (j) gibi iki kromozomun karşılıklı benzerlik oranları eşittir: $\mu_{ij} = \mu_{ji}$
3. Üyelik derecesi fonksiyonları, bağımsız değişkenin (min-max) değerleri arasında tariflidir.

Bu durumda oransal alanı (a_i) olan bir (i) kromozomunun, (a_j) alanlı bir (j) kromozomuna benzerliği, (a_i) esas alındığında

$$(a_i > a_j) \text{ için } \mu_{ij}(a) = \frac{a_j}{a_i} \text{ olur.}$$

Eğer seçilen (j) kromozomunun oransal alanı ($a_j > a_i$) şeklindeyse

$$\mu_{ij}(a) = \frac{a_i}{a_j} \text{ olarak yazılır.}$$



O zaman oransal alanlar için üyelik dereceleri genel olarak

$$\mu_{ij}(a) = \begin{cases} \frac{a_j}{a_i} & \rightarrow a_i > a_j \\ \frac{a_i}{a_j} & \rightarrow a_i < a_j \end{cases} \quad (5.2)$$

yazılır. Aynı işlem oransal kromozom boyları için de yapılırsa

$$\mu_{ij}(l) = \begin{cases} \frac{l_j}{l_i} & \rightarrow l_i > l_j \\ \frac{l_i}{l_j} & \rightarrow l_i < l_j \end{cases} \quad (5.3)$$

olur.

Alan üyelik derecesi fonksiyonunu (a)'yı bağımsız değişken olarak inceleyelim. 46 elemanlık alanlar kümesinin üyelik derecelerinin değişimini bir (a_i) 'yi esas alarak bulalım.

Bu fonksiyonda $a = a_i$ için $\mu(a) = 1$ olacaktır. $j=1,2,3..46$ için a_i , a_j benzerliklerine bakarsak,

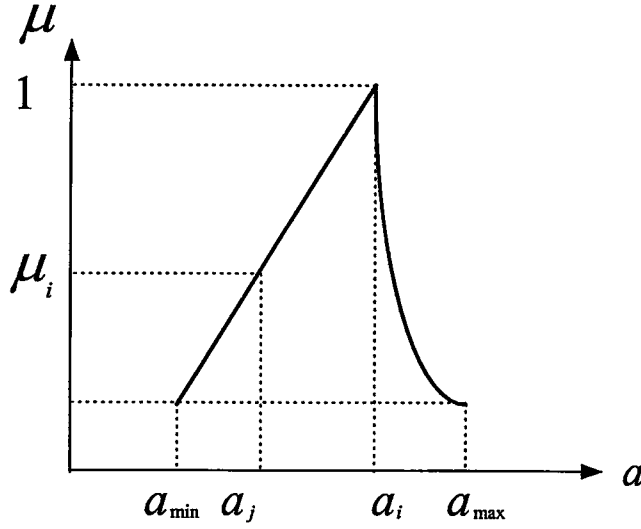
1. $a_{\min} < a_j < a_i$ aralığında $\mu_{ij}(a) = \frac{a_j}{a_i}$ olup, bu aralıkta (a_i) sabit olup bu değişim

$f(x) = m * x + n$ tarzında olup **lineerdir**.

2. $a_i < a_j < a_{\max}$ aralığında $\mu_{ij}(a) = \frac{a_i}{a_j}$ olup, buradaki değişim $f(x) = \frac{m}{x}$

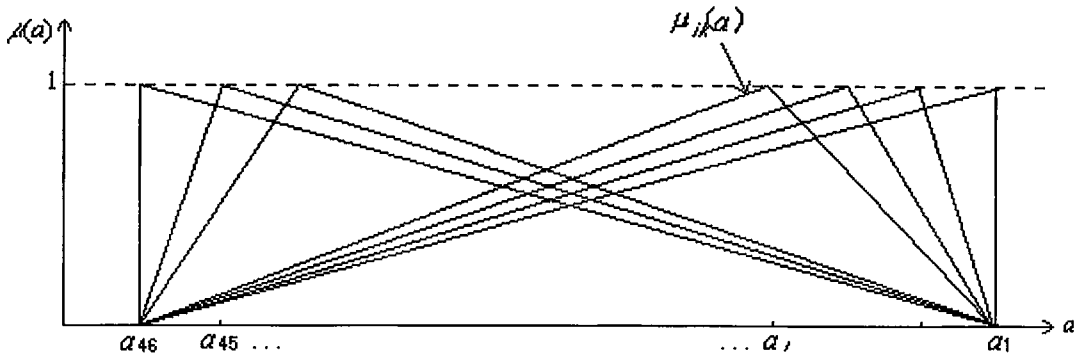
tarzındadır. Bu fonksiyon bir hiperbolün değişimini gösterir. Bu iki aralıktaki değişim Şekil 5.8'deki gibidir.





Şekil 5.8 Alan Oranları İçin Üyelik Dereceleri

Elde edilen bu bağıntıda (i) değeri, bütün kromozomları tarayacak şekilde değiştirilirse, $j \rightarrow k$ ve $k=1,2,\dots,46$ alınır, genel olarak $\mu_{ik}(a)$ üyelik derecesi fonksiyonlar ailesi elde edilir. Bu $\mu_{ik}(a)$ fonksiyonunun tarif aralığı $(a_{\min} - a_{\max})$ olarak görülmektedir. Burada $k=1,2,\dots,46$ 'dır. $i=1,2,\dots,46$ değerleri için aynı eksen sisteminde sırasıyla $\mu_{1k}(a_1), \mu_{2k}(a_2) \dots \mu_{46k}(a_{46})$ şeklinde 46 fonksiyonun değişimi Şekil 5.9'daki gibi görülecektir.



Şekil 5.9 46 kromozomun alan üyelik derecesi ilişkisi



5.4 Üyelik Derecesi Vektör ve Matrisleri

Kromozomların alan ve boy oranlarına göre çıkartılan üyelik derecesi bağıntıları yardımıyla önce bu özellikler için üyelik derecesi vektörleri, sonra da üyelik derecesi matrisleri elde edilir.

5.4.1 Alan ve boy üyelik derecesi vektörleri

Kromozom görüntülerinden elde edilen ve her kromozomun alan oranlarını gösteren “Alan-Oran” vektörünü ele alalım. Bu vektörün 1 numaralı elemanı esas alınır, 46 eleman için,

$$\mu_{ij}(a) = \begin{cases} \frac{a_j}{a_i} & \rightarrow a_i > a_j \\ \frac{a_i}{a_j} & \rightarrow a_i < a_j \end{cases}$$

ifadesinde $j=1,2,\dots,46$ için 46 elemanlı bir vektör oluşur. Bu vektör 1 numaralı kromozomun üyelik derecesi vektörü olup sembolik olarak

$$\text{Üye_Alan_Oran_1} = \{\mu_{1,1}, \mu_{1,2}, \dots, \mu_{1,46}\}$$

şeklinde yazılabilir. Aynı işlem kromozom kümesinin 2-46 numaralı elemanları için yapıldığında

$$\mu(A) = \begin{bmatrix} \mu_{1,1} \mu_{1,2} \dots \mu_{1,46} \\ \mu_{2,1} \mu_{2,2} \dots \mu_{2,46} \\ \vdots \\ \mu_{46,1} \mu_{46,2} \dots \mu_{46,46} \end{bmatrix}$$

(5.4)



şeklindeki **kromozom alan oranları üyelik derecesi** matrisi elde edilmiş olur.

Aynı işlem boy oranı değerlerine uygulanarak $\mu(l)$ matrisi elde edilir.

5.4.2 Sentromer oranı üyelik derecesi vektörü

Çok sayıda kromozom örneğinin sentromer oranları incelendiğinde metasentrik, submetasentrik ve akrosentrik kromozomları gösteren grupların ayırımında bir tek sayının yeterli olmadığı bu oranların, alt ve üst limitleri belirleyen bir a-b sayı aralığında gösterilmesi gerektiği görülmüştür. Örnek olarak 1 numaralı kromozomun sentromer oranı %45-%50 arasında değişen değerler almaktadır.

Bu bakımdan sentromerler için bulanık mantık üyelik dereceleri belirlenirken, $\mu(s)$ fonksiyonunun elde edilmesinde bağımsız (s) değişkeninin $S_{as} - S_{us}$ tarif aralığı içinde değişik platoları dikkate alınmıştır.

Bu çalışmada 46 kromozomun sentromer oranları, Bölüm 6.2'de ayrıntılı olarak incelenen değerlendirme sonuçlarına bağlı olarak S_1, S_2, S_3, S_4 gibi 4 grup altında incelenmiştir.

(S_1) grubundaki bir kromozomun sentromer oranı %50-%45 aralığındadır. Bu gruptaki bir kromozomun grup üyelik derecesi $\mu(S_1)$, bu aralık içinde (1) olacak, sağ ve soldaki iki sınır değerinde azalarak değişecektir. Bu durumda (S_1) sentromer grubu için üyelik derecesi(yani herhangi bir kromozomun s sentromer oranı değerine göre (S_1) grubunda olma ihtimali) şöyle gösterilebilir:



$$S_1 \text{ sentromer_üyelik: } \mu_{ij}(S_1) = \begin{cases} 1 & \rightarrow S_{1as} < S_j < S_{1üs} \\ \frac{S_j}{S_{1as}} & \rightarrow S_j < S_{1as} \\ \frac{S_{1üs}}{S_j} & \rightarrow S_j > S_{1üs} \end{cases}$$

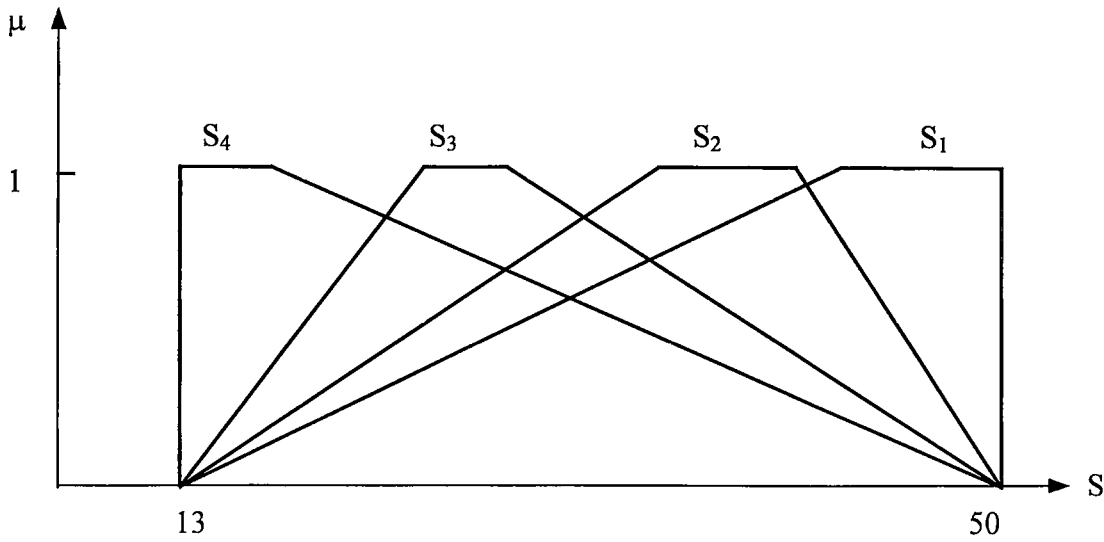
Genel olarak, $k=1,2,3,4$ olmak üzere (k .) sentromer grubu için üyelik derecesi için

$$S_k \text{ sentromer_üyelik: } \mu_{ij}(S_k) = \begin{cases} 1 & \rightarrow S_{kas} < S_j < S_{küs} \\ \frac{S_j}{S_{kas}} & \rightarrow S_j < S_{kas} \\ \frac{S_{küs}}{S_j} & \rightarrow S_j > S_{küs} \end{cases} \quad (5.5)$$

bağıntısı elde edilir. Bu bağıntılarda

S_{kas} = (k .) grubun sentromer oranı alt sınırı,

$S_{küs}$ = (k .) grubun sentromer oranı üst sınırı değerlerini göstermektedir.



Şekil 5.10 Sentromer oranları üyelik dereceleri

Şekil 5.10'da sentromer oranları üyelik dereceleri için elde edilen değerlerin şematik gösterimi verilmiştir.

5.4.3 Bant bilgileri üyelik derecesi vektörü

Bant özelliklerine göre kromozom tanımada korelasyon yöntemi kullanılmıştır. Korelasyon yönteminde kullanmak amacıyla, önce her kromozom için karakteristik özellikleri gösterecek kadar çok sayıda örnek kromozomdan bir şablon oluşturulmalıdır. Daha sonra boy, alan ve sentromer özelliklerine göre gruplara ayrılmış her test kromozomu, bulunduğu gruba ait şablon kromozomlar ile karşılaştırılarak hangi kromozoma en çok benzediği belirlenir.

Şablon kromozomun beklenen değeri

$$E(x)=\eta_x = \frac{x(\xi_1)+\dots+x(\xi_n)}{n} \quad (5.6)$$

iken,

$$E\left\{ (x-\eta_x)(y-\eta_y) \right\} \quad (5.7)$$

bağıntısı, şablon kromozom değeri ile test kromozom arasındaki kovaryansı gösterir. Buna göre şablon kromozom ile test kromozomu arasındaki korelasyon

$$r(x) = \frac{E\left\{ (x-\eta_x)(y-\eta_y) \right\}}{\sqrt{E\left\{ (x-\eta_x)^2 \right\} * E\left\{ (y-\eta_y)^2 \right\}}} \quad (5.8)$$

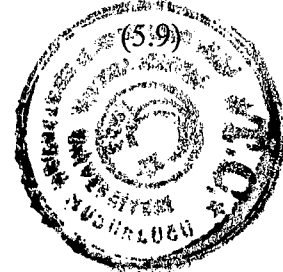
şeklinde yazılabilir. Burada $-1 \leq r(x) \leq 1$ 'dir.

$$r(x) = -1 \text{ için } \mu(x) = 0$$

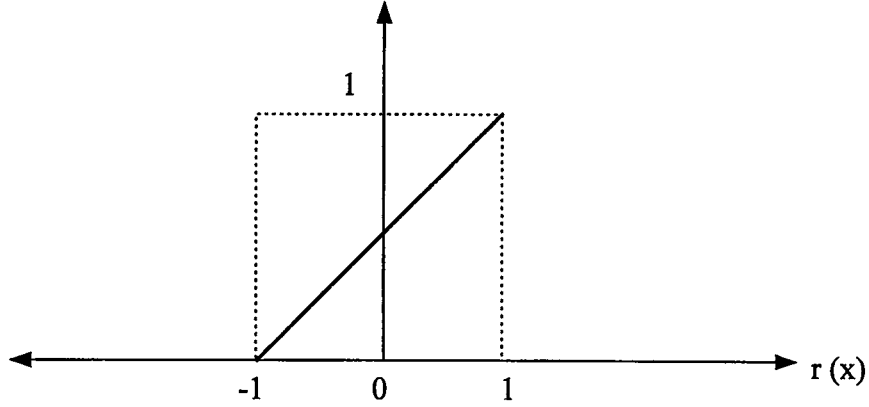
$$r(x) = 1 \text{ için } \mu(x) = 1$$

olarak değişen bir bulanık mantık üyelik derecesi fonksiyonu,

$$\mu(x) = \frac{1}{2} * r(x) + \frac{1}{2}$$



olarak elde edilir (Şekil 5.11).



Şekil 5.11 Bant Bilgileri Üyelik Derecesi

5.5 Bulanık Mantık Kromozom Sınıflandırma Sisteminin Çalışması

Kromozom görüntülerinin işlenmesi sonucu elde edilen bilgiler yardımıyla, bu test kromozomları kümesinin elemanlarının tanıma işlemine hazırlanması adım adım şöyle gerçekleşir:

5.5.1 Hazırlık aşaması

Test kromozomlarının toplam alan (A_T), toplam boy (L_T), sentromer oranları ve bant bilgileri yardımıyla alan oranları (a), boy oranları (l) ve sentromer oranları (s) ve bant bilgileri (b) vektörleri oluşturulur. Burada alan, boy ve sentromer oranları vektörlerinin büyükten küçüğe sıralı oluşturulması gerekir.



5.5.2 Eşdeğerlik bağıntısı benzerlik matrisinin oluşturulması

Kromozomların alan ve boy oranları için elde edilen üyelik derecesi matrislerinden yararlanarak kromozom gruplarının belirlenmesi (sınıflandırma) için bu matrislerin **benzerlik sınıfları** matrislerine dönüştürülmesi gerekir. Bu işlemi yapmak üzere, önce bulanık mantıkta eşdeğerlik ilişkisini inceleyelim.

5.5.2.1 Klasik mantık eşdeğerlik ilişkisi (Denklik Bağı)

$$X = \{x_1, x_2, \dots, x_n\}$$

$$Y = \{y_1, y_2, \dots, y_n\}$$

kümelerinde $x \in X$ $y \in Y$ olmak üzere (X) ve (Y) kümeleri arasındaki bağıntı

$$R : x \rightarrow y \quad x R y$$

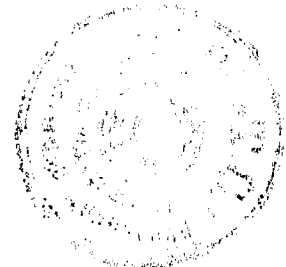
olarak gösterilir.

Çoğu zaman (R) bağıntısı bir (X) kümesinin kendi elemanları arasında olup:

$$R : x \rightarrow x, \quad x R x \text{ olur.}$$

Bir (R) bağıntısı aşağıdaki üç özelliğe sahipse (R) bir **Eşdeğerlik İlişkisi** (Denklik Bağı) demektir:

- Yansıma: $(x_i, x_i) \in R$
- Simetri: $(x_i, x_j) \in R \rightarrow (x_j, x_i) \in R$ (5.10)
- Geçiş (Transitivity): $\left. \begin{array}{l} (x_i, x_j) \in R \\ (x_j, x_k) \in R \end{array} \right\} \rightarrow (x_i, x_k) \in R$



Bu üç özellikten geçiş bağıntısını, R 'nin bir benzerlik ilişkisi olması halinde ele alırsak x_i, x_j 'ye ve x_j, x_k 'ya benziyorken x_i, x_k 'ya benziyorsa (R) geçiş özelliğine sahip demektir.

Bu benzerlik bağıntısını (R) ikilileri yerine $N \times N$ elemanlı bir (M_R) ilişki matrisi üzerinde gösterirsek geçiş özelliği için

$$\left. \begin{array}{l} m_{ij} = 1 \\ m_{ij} = 1 \end{array} \right\} \text{ iken } m_{ik} = 1 \quad \text{olmalıdır.}$$

Klasik mantıkta bir (R) bağıntısının bazı elemanları arasında geçiş yoksa, (M_R) matrisine geçişi sağlayacak bir eleman eklenir. Yani $(m_{ij} = 1 \text{ ve } m_{jk} = 1)$ iken (R) 'ye geçiş özelliği kazandırmak için $m_{ik} = 1$ yapılır. Ekleme işlemi için (M_R) matrisinin $M_{R^2} = M_R \circ M_R$ şeklinde lojik çarpım yapılır. Eklenen yeni elemanların varlığı ile geçiş özelliği yeniden bozulan (R) bağıntısına (birinci grup yeni elemanların gerektirdiği ikinci grup) yeni elemanların eklenmesi gerekebilir. Bu durumda (M_R) matrisi sabit oluncaya kadar matris çarpımı (lojik çarpım) $M_{R^2} = M_R \circ M_R, M_{R^3} = M_R \circ M_{R^2} \dots$ işlemi tekrarlanır. Böylece (R) bağıntısının *Geçiş Zarfı* (*Transitive Closure*) elde edilir.

(M_R) matrisinin sabit hale gelmesi (N^3) karmaşıklığındaki lojik çarpma işleminin en fazla $(N-1)$ kere tekrarını gerektirir. O zaman *Geçiş Zarfı* işleminin karmaşıklığı (N^4) olur. Klasik lojikte *Geçiş Zarfı* için karmaşıklığı (N^3) olan Warshall Algoritması kullanılır.



5.5.2.2 Bulanık mantık eşdeğer ilişkisi (Denklik Bağı)

Bir $X = \{x_1, x_2, x_3, \dots, x_n\}$ uzayı üzerinden (R) bağı yine (X) 'den (X) 'e bir bağıntıdır. Burada da (R) bağıntısı aşağıdaki üç özelliğe sahipse bu bir *Bulanık Mantık Eşdeğerlik İlişkisi (Denklik Bağı)* demektir. Bu üç özellik şu ifadelerle gösterilir:

$$\begin{aligned}
 \text{Yansıma:} \quad & \mu_R(x_i, x_i) = 1 \\
 \text{Simetri:} \quad & \mu_R(x_i, x_j) = \mu_R(x_j, x_i) \\
 \text{Geçiş:} \quad & \left. \begin{aligned} \mu_R(x_i, x_j) &= \lambda_1 \\ \mu_R(x_i, x_k) &= \lambda_2 \end{aligned} \right\} \mu_R(x_i, x_k) = \lambda
 \end{aligned} \tag{5.11}$$

Burada $\lambda \geq \min(\lambda_1, \lambda_2)$ olmalıdır.

5.5.2.3 Bulanık mantık ilişkisi kompozisyonu

$$\begin{aligned}
 X &= \{x_1, x_2, x_3, \dots, x_n\} \\
 Y &= \{y_1, y_2, y_3, \dots, y_n\} \\
 Z &= \{z_1, z_2, z_3, \dots, z_n\}
 \end{aligned} \quad \text{iken}$$

$P(X, Y)$ ve $Q(Y, Z)$, (Y) kümesini müşterek kullanan iki tane, ikili klasik ilişki olsun. Bir (P) ve (Q) kompozisyonu $[P \circ Q]$ olarak gösterilir. $P \circ Q : X \rightarrow Z$ olarak gösterilen bu bağıntıda $(x, z) \in P \circ Q$ şartı

$$y \in Y \text{ ve } \left\{ \begin{aligned} (x, y) &\in P \\ (y, z) &\in Q \end{aligned} \right\} \rightarrow \text{için } (x, z) \in P \circ Q \tag{5.12}$$

olmalıdır.



Bir $[P \circ Q]$ kompozisyonu iki şekilde gerçekleşir:

a) Maksimum-Minimum Kompozisyonu:

$P(X,Y)$ ve $Q(Y,Z)$ iki bulanık mantık ilişkisi ise,
 $P \circ Q : X \rightarrow Z$ $(x,z) \in X \times Z$ 'dir.

Üyelik derecesi üzerinden maksimum-minimum kompozisyonu

$$\mu_{P \circ Q}(x,z) = \max[\min(\mu_P(x,y), \mu_Q(y,z))] \quad (5.13)$$

şeklinde ifade edilir.

b) Maksimum-Çarpım Kompozisyonu:

Aynı kompozisyon işlemi uygulamanın gerektirdiği bazı durumlarda, maksimum çarpım olarak,

$$\mu_{P \circ Q}(x,z) = \max_{y \in Y} [\mu_P(x,y) \circ \mu_Q(y,z)] \quad (5.14)$$

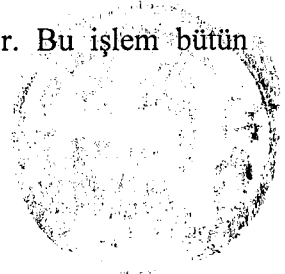
şeklinde elde edilir.

5.5.2.4 Tolerans ilişkisinden eşdeğerlik ilişkisine geçiş

Gerek klasik, gerekse bulanık mantıkta bir (R) ilişkisi sadece simetri ve yansıma özelliğine sahipse bu ilişki *Tolerans İlişkisi* (*Tolerance Relation*) olarak isimlendirilir. Sadece

tolerans özelliği olan bir (R) ilişkisinde, (M_R) matrisi üzerinde $\left. \begin{array}{l} m_{ij} = 1 \\ m_{jk} = 1 \end{array} \right\} \text{ iken } m_{ik} = 0$

ise ardışık $(M_R \circ M_R)$ kompozisyonları alınarak $m_{ik} = 1$ yapılır. Bu işlem bütün



($i=1,2,\dots,n$), ($j=1,2,\dots,n$), ($k=1,2,\dots,n$) değerleri için yapılırsa (R) “tolerans ilişkisi” “eşdeğerlik ilişkisine” dönüşür.

Sonuç olarak, bir (R) ilişkisinde tolerans ilişkiden eşdeğerlik ilişkisine geçmek için (R)’ye ait (M_R) matrisi üzerinde $M_{R^2} = M_R \circ M_R$, $M_{R^3} = M_R \circ M_{R^2} \dots$ ardışık kompozisyon alma işlemi yapılır. (M_R) matrisinin değişimi duruncaya kadar işleme devam edilir. (M_R)’nin sabit hale gelmesi en çok ($N-1$) adımlık işlem gerektirir. O zaman maksimum sistem karmaşıklığı $\approx N^4$ olur.

Bu doktora çalışmasında, yansıma ve simetri özelliği gösteren benzerlik bağıntılarından (*tolerans ilişkisi*), yansıma, simetri ve geçiş özelliği gösteren benzerlik bağıntılarına (*eşdeğerlik ilişkisi*) geçmek üzere benzerlik üyelik dereceleri matrisleri max-min kompozisyon işlemine sokulmuştur.

Uygulamada, kromozomlar için $N=46$ olduğundan her kompozisyon, $\mathcal{G} = 46 * 46 * 46 = 97.336$ işlem gerektirmiştir. Bu durumda, üyelik derecesi matrisinin sabit hale gelmesi için adım sayısı $46-1=45$ iken $\mathcal{G} = 46^3 * 45 = 4.380.120$ işlem gerekirken ilk defa bu çalışmada ortaya atılan *Ardışık Kuvvet Alma Algoritması* ile:

$$T = M_{R^{n-1}} = \prod_{i=1}^{n-1} M_R \quad (5.15)$$

işlemi yerine

$$T_1 = M_R * M_R, T_2 = T_1 \circ T_1, T_3 = T_2 \circ T_2, \dots, T_r = T_{r-1} \circ T_{r-1}$$

$$T = ((((M_R \circ M_R)^2)^2) \dots)^r \quad (5.16)$$

sonucuna varılmış, böylece ($N-1$) adımlık çarpma işlemi $\lceil \log_2(N-1) \rceil$ adıma indirgenmiştir.



$N=46$ için $\lceil \log_2(45) \rceil = 6$ olup $T_1 = M_{R^2}, T_2 = M_{R^4} \dots T_6 = M_{R^{64}}$ değeri hesaplanmıştır. Böylece işlem sayısı $45 * 46^3 = 4.380.120$ yerine $6 * 46^3 = 584.016$ 'ya düşmüştür. Böylece kompozisyonlar için performans $(P = \frac{N}{\log_2 N})$ kadar arttırılmıştır.

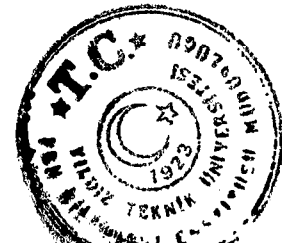
5.5.3 Kromozomların sınıflandırılması işlemi

Kromozomların alan veya boy oranları üyelik derecesi matrisi bir *benzerlik ilişkisi* matrisidir. Bu matris, yansıma ve simetri özelliği olan Tolerans Bulanık İlişki (Tolerance Fuzzy Relation) matrisidir. Tablo 5.2'de ISCN 1995'den alınan 450 bantlı ideal kromozom bilgilerinin oransal boy değerleri için elde edilen Tolerans Bulanık İlişki matrisi verilmiştir. Bu matrisler üzerinde benzerlik sınıfları oluşturmak üzere, *Ardışık Kuvvet Alma Algoritması* kullanılarak yapılan 6 adımlık kompozisyon işlemi sonunda *Eşdeğer İlişki (Equivalence Relation)* matrisi elde edilmiştir (Tablo 5.3).

Eşdeğerlik İlişkisi matrisinin (0-1) aralığındaki ondalık kesirli sayılardan oluşan her elemanı bir $(0 < \alpha < 1)$ gibi bir (α) sayısı ile karşılaştırılarak α -Kesmeleri(α -cut) denen bir işlemle

$$\mu_{ij}(a) = \begin{cases} 1 & \rightarrow \mu \geq \alpha_a \\ 0 & \rightarrow \mu < \alpha_a \end{cases} \quad (5.17)$$

şeklinde, elemanları (0) ve (1)'lerden oluşan yeni matris elde edilir. $\alpha_a = 95$ için elde edilen, alan oranları için benzerlik sınıfları matrisi Tablo 5.4'de görülmektedir. Bu yeni ilişki matrisi, *berrak ilişki (defuzzified relation)* matrisidir. Bu matrisin iki temel özelliği ana diyagonalin elemanlarının (1) olması(yansıma) ve diğer elemanların bu diyagonale göre simetrik olmasıdır.



Berrak İlişki matrisinin 1. satırındaki ilk eleman (1)'dir ($m_{11} = 1$). Bu satırdaki diğer (1)'lerin bulunduğu sütun numaraları, 1 numaralı kromozomun da içinde bulunduğu ve (α) kesmesine göre oluşan birinci kromozom grubunun elemanlarını gösterir. Bu gruba girmeyen, aynı satırda $\mu = 1$ olan ikinci ve sırayla diğer kromozom gruplarının elemanları bulunur.

(α) sayısının $\alpha = 1$ ve $\alpha = 0$ limit değerlerinin kullanılması halinde oluşacak gruplara bakalım:

a) $\alpha = 1$ için oluşan $\mu_{\alpha=1}$ matrisinde yalnız ana diagonal (1)'lerden oluşur. Bu durum benzerlik sınıf sayısının (46) olduğunu gösterir. Yani (46) kromozomun herbiri kendi sınıfını oluşturmuş demektir.

b) $\alpha = 0$ alındığında $\mu_{\alpha=0}$ matrisinin bütün elemanları (1) olur ki bu durum da bütün kromozomların bir tek sınıfta toplandığını gösterir.

Oluşacak benzerlik gruplarının sayısı (G) ise $G = f(\alpha)$ olduğu görülür. Uygulamada, beklenen sınıf sayısına yakın bir (G) değeri elde etmek üzere, iterasyonlarla, uygun (α) sayısı seçilmiş olur.

Böylece (α_a)'ya bağlı olarak bulunan alan benzerlik grupları G(a) elde edilmiş olur.

Aynı işlem sırasıyla, bu defa boy oranları üyelik derecesi matrisinden (α_l) kesmeleri berrak ilişki matrisi ve dolayısıyla boy benzerlik grupları G(l) elde edilmiş olur.

Üçüncü olarak hazırlanan sentromer benzerlik ilişkisi matrisinin ardışık kompozisyonları alınarak *sentromer eşdeğerlik ilişkisi matrisi* elde edilir.

4 satır, 46 sütundan oluşan bu matrisin elemanları da $0 < \alpha_s < 1$ için

$$\mu_{\bar{y}}(s) = \begin{cases} 1 & \rightarrow \mu \geq \alpha_s \\ 0 & \rightarrow \mu < \alpha_s \end{cases} \quad (5.18)$$

işlemine göre (0) ve (1)'e dönüşür. Böylece S_1, S_2, S_3, S_4 sentromer grupları için *Berrak İlişki Matrisi* elde edilir. Burada da iterasyonlarla uygun (α_s) seçilip kromozomların sentromer oranları bakımından $(G(s))$ grupları elde edilir.

Sonuç olarak kromozomların üç ana özelliği için her biri $(G = \{\{g_1\}, \{g_2\}, \dots, \{g_k\}\})$ şeklinde alt kümelerden oluşan $G(a), G(l), G(s)$ grupları elde edilir.

5.6 Alfa (α) Kesmesi

Bir (A) bulanıklık kümesi üzerinden çıkarılan, (A) elemanlarının bir 'test' elemanına benzerliğini simgeleyen $\mu_A(x)$ üyelik dereceleri bir küme oluşturur. $0 \leq \mu_A(x) \leq 1$ aralığında, benzerliğin en yüksek olduğu eleman için $\mu_A(x) = 1$ olurken, hiç benzememe hali için $\mu_A(x) = 0$ olur.

Bir küme elemanının, benzerlik üyesi derecesine göre tanınması işleminde, üyelik derecesi belli bir (α) eşik değerinin altındaysa, $\mu_A(x) < \alpha$ ise $\mu_A(x) = 0$ yapılır. Bu işlem 'alpha kesmesi' olarak tarif edilir ve şu şekilde yazılır:

$$A_\alpha = \{x | \mu_A(x) \geq \alpha\}$$

Böylece tanıma olayı için yüksek üyelik dereceli bir benzerlik gerekiyorsa α büyük seçilerek A_α kümesinin eleman sayısı azaltılır. Bu ise tanımda seçiciliği artırır.

Kromozomların sınıflandırılarak tanınmasında test elemanlarından çıkarılan üyelik dereceleri (benzerlik ilişkisi) matrisleri, kompozisyon işlemleriyle, benzerlik sınıflarını oluşturmak üzere *eşdeğerlik ilişkisi* matrislerine dönüştürülmüştür. $(0,1)$ aralığındaki



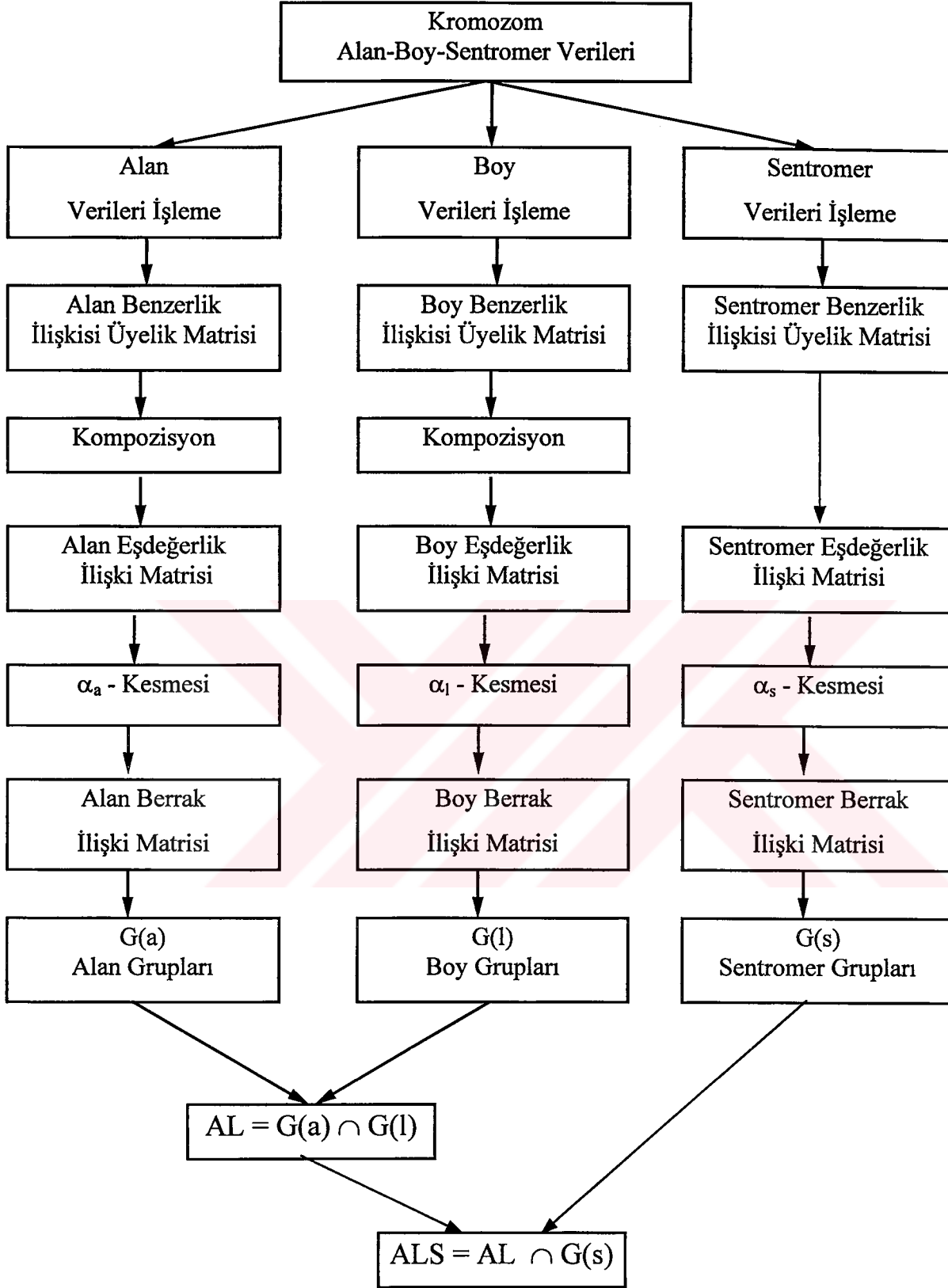
ondalık kesirli sayılardan oluşan bu matrisler de alan, boy ve sentromer oranları için seçilen uygun ($\alpha_a, \alpha_l, \alpha_s$) değerleriyle α -kesmeleri işlemine sokulup her üç özellik için elemanları (0) ve (1)'lerden oluşan *berrak ilişki* matrisleri elde edilmiştir. Buradaki α -kesmesi uygulamalarında

$$\begin{aligned} \mu_A(x) < \alpha &\rightarrow \mu_A(x) = 0 \\ \mu_A(x) \geq \alpha &\rightarrow \mu_A(x) = 1 \end{aligned} \quad \text{yapılmıştır.}$$

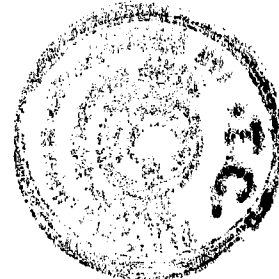
5.7 Kromozom Sınıflandırma ve Tanıma İşlemi

Kromozom görüntülerinin işlenmesiyle elde edilen, alan, boy ve sentromerlere ait veriler, benzerlik ilişkisi matrisinin oluşumuna kadar değişik işlemlerden geçerler. Bu işlemlerin sırasını gösteren veri-akış diyagramı Şekil 5.12'dedir. (0) ve (1) elemanlarından oluşan berrak ilişki matrisinde kümelenen (1)'ler kromozom sınıflarının elemanlarını veren alt kümelere dönüşürler.





Şekil 5.12 Veri Akış Diyagramı



İdeal bir sınıflandırmada boy-alan için yedişer alt sınıf, sentromerler için de dört sınıf ortaya çıkar. Küme olarak isimlendireceğimiz bu sınıflar şöyle yazılabilir.

$$\text{Alan Kümeleri: } G(a) = \{\{g_1(a)\}, \{g_2(a)\}, \dots, g_k(a)\}$$

$$\text{Boy Kümeleri: } G(l) = \{\{g_1(l)\}, \{g_2(l)\}, \dots, g_k(l)\}$$

$$\text{Sentromer Kümeleri: } G(s) = \{\{g_1(s_1)\}, \{g_2(s_2)\}, \{g_3(s_3)\}, \{g_4(s_4)\}\}$$

Bu kümelerden alan ve boy kümelerinin $g(a)$ ve $g(l)$ alt kümelerinin kesişmeleriyle kromozom sınıfları belirlenir.

$$AL = \cap(g(a), G(l)) = \cap_{i=1,k} [\{g_i(a)\}, \{g_i(l)\}]$$

işlemi sonunda;

$$AL(1) = [g_1(a)] \cap [g_1(l)]$$

$$AL(2) = [g_2(a)] \cap [g_2(l)]$$

...

$$AL(k) = [g_k(a)] \cap [g_k(l)]$$

kromozom sınıflarının elemanları belirlenir. Yeni oluşan bu kesişim kümeleri, verilerin giriş sıra numarasına göre kromozom numaralarını içerir.

Kromozomlar boy-alan ve sentromer bilgilerine göre (A,B,C,D,E,F,G) olarak yedi grupta toplanırlar. Elde edilen (AL) kesişim kümesinin alt kümelerinin bu (A-G) gruplarına bire bir denk gelmesi istenen sonuçtur. Fakat uygulamada:

1. Kromozom gruplarından bazılarının elemanları yan grupların içinde yer alırsa (AL)'nin alt küme sayısı yediden küçük olur.



2. Ayrı grupların elemanları yeni bir grup oluşturursa (AL)'nin alt küme sayısı yediden büyük olur.

Bu çalışmada $\alpha_a = 98\%$ ile $\alpha_a = 93\%$ arasındaki α değerleri için G(a)'nın alt küme sayısı için $k=25$ ile $k=4$ arasında sonuçlar elde edilmiştir.

(AL) kümesindeki sınıfların elemanlarının daha duyarlı bir şekilde belirlenmesi için (AL) alan-boy kesişimleri kümesinin alt birimleri G(s) sentromer kümesiyle kesiştirilirler.

$$ALS = \bigcap_{\substack{1 \leq i \leq 7 \\ 1 \leq j \leq 4}} \left\{ [AL(i)], [G(S_j)] \right\}$$

şeklinde gösterilen bu işlemde (AL)'yi oluşturan gruplar (A,B,C,D,E,F,G) iken G(S)'i oluşturan gruplar (S_1, S_2, S_3, S_4) şeklindedir. Fakat (A-G) gruplarının elemanlarının sentromerlerine bakıldığında:

$$A \rightarrow (S_1, S_2), B \rightarrow (S_3), C \rightarrow (S_2, S_3), D \rightarrow (S_4), E \rightarrow (S_1, S_2), F \rightarrow (S_1), G \rightarrow (S_2)$$

olduğu görülür. O halde (ALS) kesişiminde kesişim kümeleri sembolik olarak aşağıdaki gibi olacaktır.

$$ALS(1) = AL(1) \cap g(S_1) \cap g(S_2)$$

$$ALS(2) = AL(2) \cap g(S_3)$$

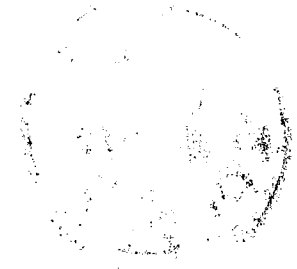
$$ALS(3) = AL(3) \cap g(S_2) \cap g(S_3)$$

$$ALS(4) = AL(4) \cap g(S_4)$$

$$ALS(5) = AL(5) \cap g(S_1) \cap g(S_2)$$

$$ALS(6) = AL(6) \cap g(S_1)$$

$$ALS(7) = AL(7) \cap g(S_2)$$



Boy, alan ve sentromer bilgilerinin deęerlendirilmesi sonucunda yedi gruba yerleřtirilen kromozomların bant bilgileri deęerlendirilerek her kromozomun kaç numaralı kromozom olduęu belirlenir. Bu alıřmada, test kromozomlarının bant bilgileri ile 30 kromozom grntsndeki 30*22 otozom ve 30*2 cinsiyet kromozomunun bant deęerlerinden elde edilen řablon deęerler arasında korelasyon iřlemi yapılmıř ve elde edilen sonu deęerlere gre her kromozomun kaç numaralı kromozom olduęuna karar verilmiřtir. Test kromozomların 24 kromozom řablonunun herbiri ile deęil, sadece iinde buldukları grubun kromozom řablonları ile korelasyon iřlemine girmesi, sistemin bařarı oranını arttıran nemli bir faktrdr.

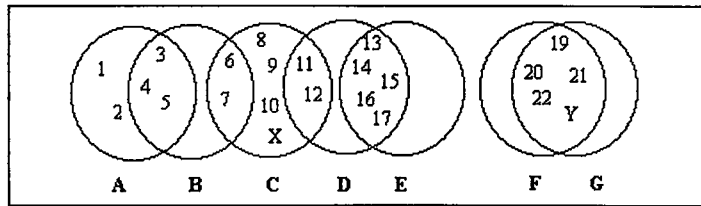


6. UYGULAMA

Bu tez çalışmasında geliştirilen Kromozom Sınıflandırma ve Tanıma Sistemi, 40 kromozom görüntüsündeki 40*22 otozom ve 40*2 cinsiyet kromozomu üzerinde denenerek sistemin başarısı değerlendirilmiştir.

6.1 Kromozomların Boy Bilgilerinin Değerlendirilmesi

ISCN 1995'e göre kromozomlar, şekil özellikleri olan boy ve sentromer yeri bilgilerine göre (A-G) olarak isimlendirilen yedi gruba ayrılırlar. Bu gruplar incelendiğinde, kromozomların boylarına göre büyükten küçüğe sıralı olduğu görülür. Buna göre, teorik olarak A grubundaki bir kromozomun B grubundaki bir kromozomdan daha uzun olması gerekir. Kromozomları gruplara ayırmak için, boy ve alan bilgilerine göre sıralamanın pratikte ne kadar doğru sonuç verdiğini görmek amacıyla her kromozomun kaç numaralı kromozom olduğunun bilindiği test kromozom görüntülerinin boy-alan bilgileri her görüntü için ayrı ayrı değerlendirilmiş ve her görüntünün kromozomları boy-alan bilgisine göre büyükten küçüğe sıralanarak gruplandırmalar yapılmıştır. Gruplandırmalar incelendiğinde özellikle aralarında belirgin boy farkı olmayan kromozomlarda, daha kısa boylu olması beklenen bir kromozomun daha uzun olabildiği ve buna bağlı olarak yanlış gruplandırmalar yapılabildiği görülmüştür. Şekil 6.1'de kromozomların boy ve alan bilgileri değerlendirilerek yapılan gruplandırma işlemi sonucunda her kromozomun hangi gruplarda görülebildiği gösterilmiştir.



Şekil 6.1 Boy benzerliklerine göre kromozomların görüldüğü gruplar



Kromozomları boy-alan bilgisine göre büyükten küçüğe sıralayarak yapılan gruplandırmalara 10 örnek Tablo 6.1'de verilmiştir. Yanlış grupta olan kromozomlar koyu renkte gösterilmiştir.



Tablo 6.1 Oransal Boy Değerlerine Göre Gruplandırmaya 10 Örnek. Yanlış Grupta Olan Kromozomlar Koyu Renkte Gösterilmiştir.

| Grup | Kromozom Numaraları | Grup Başarı (%) | Grup | Kromozom Numaraları | Grup Başarı (%) |
|------|----------------------------------|-----------------|------|----------------------------------|-----------------|
| A | 1 2 1 2 3 3 | 100 | A | 1 1 2 2 3 4 | 83 |
| B | 5 4 4 5 | 100 | B | 3 5 5 4 | 75 |
| C | 6 6 X 7 7 9 8 11 8 11 10 9 10 12 | 100 | C | 6 6 7 9 11 11 7 8 10 8 X 9 10 12 | 100 |
| D | 13 15 14 15 13 14 | 100 | D | 13 15 15 13 14 14 | 100 |
| E | 17 16 17 16 18 18 | 100 | E | 16 17 17 18 18 16 | 100 |
| F | 20 19 20 19 | 100 | F | 19 20 22 20 | 75 |
| G | 22 22 21 21 Y | 100 | G | 19 22 21 21 Y | 80 |
| A | 1 2 1 2 3 4 | 83 | A | 1 2 1 2 3 3 | 100 |
| B | 3 4 5 5 | 75 | B | 4 4 5 5 | 100 |
| C | 8 8 7 6 6 9 11 7 X 12 11 10 10 9 | 100 | C | 6 6 7 X X 7 9 8 8 10 9 10 11 12 | 100 |
| D | 13 13 15 14 15 16 | 83 | D | 13 15 13 14 14 15 | 100 |
| E | 14 16 17 17 19 18 | 66 | E | 16 17 16 18 17 18 | 100 |
| F | 18 20 19 22 | 50 | F | 19 19 20 20 | 100 |
| G | 22 19 21 21 | 75 | G | 22 21 21 22 | 100 |
| A | 1 1 2 2 3 3 | 100 | A | 1 2 1 2 4 3 | 75 |
| B | 5 4 5 4 | 100 | B | 3 5 6 6 | 25 |
| C | 6 6 7 7 X 8 X 11 12 10 10 11 8 | 100 | C | 4 5 7 7 8 8 9 X 11 9 10 10 12 | 86 |
| D | 13 14 13 15 15 14 | 100 | D | 13 14 15 13 14 15 | 100 |
| E | 16 17 16 17 18 18 | 100 | E | 16 17 17 16 18 18 | 100 |
| F | 19 20 19 21 | 75 | F | 19 22 20 20 | 50 |
| G | 20 21 22 22 | 75 | G | 19 21 Y 22 21 | 80 |
| A | 2 1 2 1 3 4 | 75 | A | 1 1 2 2 3 3 | 100 |
| B | 3 5 5 4 | 75 | B | 4 5 4 5 | 100 |
| C | 6 6 10 7 7 10 X X 8 8 9 9 11 12 | 93 | C | 6 6 7 7 8 X 9 8 9 10 12 10 11 11 | 100 |
| D | 12 13 14 15 14 15 | 83 | D | 13 13 14 15 14 15 | 100 |
| E | 16 16 17 17 18 18 | 100 | E | 16 16 17 17 18 18 | 100 |
| F | 19 20 19 20 | 100 | F | 20 19 19 20 | 100 |
| G | 21 22 22 21 | 100 | G | 22 21 21 22 Y | 100 |
| A | 1 1 2 2 3 3 | 100 | A | 2 1 1 2 3 5 | 83 |
| B | 4 5 4 6 | 75 | B | 4 5 4 6 | 75 |
| C | 5 6 7 X 7 8 X 8 10 11 10 11 9 9 | 93 | C | 3 6 8 10 12 7 8 X X 9 10 9 7 8 | 87 |
| D | 13 14 13 14 15 15 | 100 | D | 11 15 15 16 14 16 | 50 |
| E | 16 16 18 17 17 18 | 100 | E | 13 17 14 17 18 18 | 66 |
| F | 20 19 19 20 | 100 | F | 20 20 19 22 | 75 |
| G | 21 22 22 21 | 100 | G | 22 21 19 21 | 75 |



Tablo 6-1'deki gruplandırma başarı oranları incelendiğinde bu tarz sınıflandırmanın başarısını arttırmak ve gruplar arasındaki istenmeyen geçişleri belirleyip engellemek için boy-alan bilgisine göre sıralanmış kromozomların sentromer yeri bilgilerinin de değerlendirilerek gruplandırma işleminin tamamlanması gerektiği görülmüştür.

6.2 Kromozomların Sentromer Bilgilerinin Değerlendirilmesi

ISCN 1995'e göre kromozomlar sentromer yerlerine göre metasentrik, submetasentrik ve akrosentrik olarak üç temel gruba ayrılırlar (Tablo 6.2).

Tablo 6.2 Kromozomların Sentromer Özellikleri

| Grup | Kromozom Numaraları | Sentromer |
|------|---------------------|----------------|
| A | 1 3 | Metasentrik |
| | 2 | Submetasentrik |
| B | 4 5 | Submetasentrik |
| C | 6 7 8 9 10 11 12 X | Submetasentrik |
| D | 13 14 15 | Akrosentrik |
| E | 16 | Metasentrik |
| | 17 18 | Submetasentrik |
| F | 19 20 | Metasentrik |
| G | 21 22 Y | Akrosentrik |

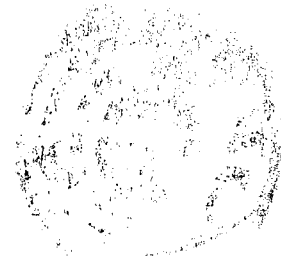
40 adet kromozom görüntüsü üzerinde yapılan deneysel çalışmalar sonucunda her kromozomun sentromer oranı değerleri Tablo 6.3'deki şekilde elde edilmiştir. Elde edilen bu değerlerin ideal sentromer oranı değerleri ile karşılaştırılabilmesi için, tabloda ISCN 1995'deki 450 bantlı ideal kromozom değerleri için sentromer oranları da gösterilmiştir.



Tablo 6.3 40 kromozom örneđi için sentromer oranı aralıđı deđerleri

| Kromozom No | Sentromer Oranı Aralıđı(%) | İdiogram Sentromer Oranı(%) | Kromozom No | Sentromer Oranı Aralıđı(%) | İdiogram Sentromer Oranı(%) |
|-------------|----------------------------|-----------------------------|-------------|----------------------------|-----------------------------|
| 1 | 45 - 50 | 48 | 13 | 13 - 22 | 6 |
| 2 | 35 - 42 | 38 | 14 | 13 - 22 | 6 |
| 3 | 44 - 50 | 46 | 15 | 13 - 22 | 6 |
| 4 | 24 - 30 | 24 | 16 | 41 - 45 | 42 |
| 5 | 24 - 30 | 24 | 17 | 28 - 37 | 27 |
| 6 | 34 - 42 | 36 | 18 | 23 - 33 | 21 |
| 7 | 34 - 42 | 35 | 19 | 42 - 50 | 40 |
| 8 | 33 - 38 | 28 | 20 | 41 - 50 | 46 |
| 9 | 32 - 40 | 31 | 21 | 22- 30 | 10 |
| 10 | 30 - 37 | 31 | 22 | 22- 30 | 11 |
| 11 | 35 - 45 | 39 | X | 36-41 | 37 |
| 12 | 24 - 30 | 24 | Y | 28-34 | 18 |

Tablo 6.3 incelendiđinde bazı submetasentrik ve akrosentrik kromozomların sentromer oranlarının içinde buldukları sentromer özelliđi grubundaki diđer kromozomların sentromer oranlarından belirgin olarak farklı olduđu görölmüştür. Örneđin genel olarak submetasentrik kromozomların sentromer oranı aralıđı %34-%42 iken, yine submetasentrik olan 4 ve 5 numaralı kromozomların sentromer oranları %23-%30 aralıđında deđerler almaktadır. Akrosentrik kromozomlar olan 13, 14 ve 15 numaralı kromozomların sentromer oranları %13-%22 arasında deđişirken 21 ve 22 numaralı kromozomların uyduları çok belirgin olduđundan kromozomun devamı gibi deđerlendirilmekte ve sentromer oranları %22-%30 gibi büyük deđerler çıkmaktadır. Bu nedenle bu çalışmada, kromozomların sentromer özellikleri idealdeki metasentrik, submetasentrik ve akrosentrik kavramlarına göre deđerlendirilmemiş, ölçümler sonunda elde edilen deđerler kullanılarak, sentromer oranlarının benzerliđine göre Tablo 6.4'de gösterildiđi gibi dört grup altında incelenmişlerdir.



Tablo 6.4 Sentromer Grupları

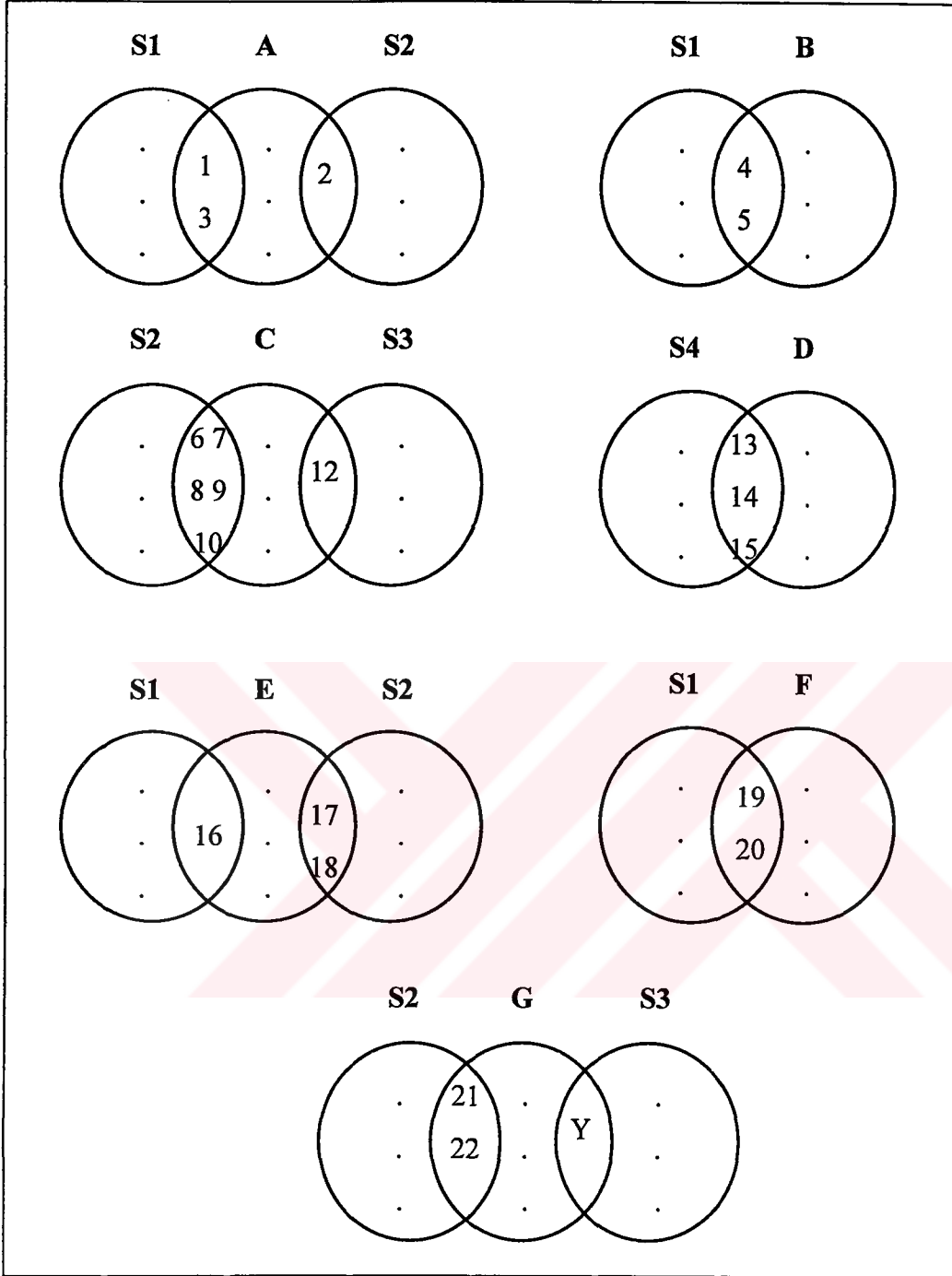
| Sentromer Grubu | Sentromer Oranı(%) | Kromozom Numaraları |
|-----------------|--------------------|-------------------------------|
| S1 | 50-42 | 1 3 16 19 20 |
| S2 | 41-30 | 2 6 7 8 9 10 11 X 17 18 21 22 |
| S3 | 34-28 | 4 5 12 Y |
| S4 | 22-13 | 13 14 15 |

ISCN 1995'de tanımlı olan (A-G) gruplarının komşu olanlarının S_1, S_2, S_3, S_4 sentromer grupları ile kesişim kümeleri aşağıdaki gibidir.

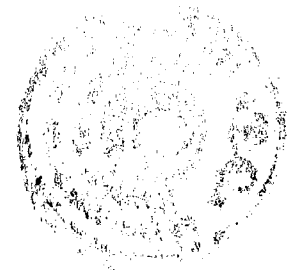
$$A \cap B = \emptyset \quad B \cap C = S_2 \quad C \cap D = \emptyset \quad D \cap E = \emptyset \quad E \cap F = S_1 \quad F \cap G = \emptyset$$

Her gruptaki kromozomların S_1, S_2, S_3, S_4 sentromer grupları ile kesişim kümeleri de Şekil 6.2'deki gibidir.





Şekil 6.2 S_1, S_2, S_3, S_4 grupları arasında ISCN 1995’de tanımlı olan (A-G) gruplarına göre boy ve alan bilgisine bağlı olarak yapılan kesişimin sonucu



Kromozomlar, boy ve alanlarına göre gruplandıklarında en fazla karışıklık komşu gruplar arasında olduğu için sentromer özellikleri aynı olan kromozomlar boyları benzemediği sürece doğru gruplanabilmektedirler.

Bu durumu örnekleyecek olursak; 4 (B), 5(B) ve 12 (C) numaralı kromozomlar S_2 grubunda, 16 (E), 19(F) ve 20 (F) numaralı kromozomlar ise S_1 grubundadırlar. B ve C ile E ve F komşu gruplar olmalarına rağmen 4, 5 numaralı kromozomların boyu 12 numaralı kromozoma, 16 numaralı kromozomun boyu da 19 ve 20 numaralı kromozomlara benzemediği için gruplandırma sırasında bu kromozomlar birbirine karıştırılmamalıdır. Komşu olmayan ama sentromer özellikleri aynı olan diğer kromozom gruplarının da boy benzerliği olmadığı sürece birbirlerine karıştırılmamaları gerekir.

Bu nedenle bu çalışmada kromozomlar önce bulanık mantık yöntemleri ile boy ve alan bilgilerine göre benzerliklerine göre sınıflandırılmışlar; daha sonra her kromozomun sentromer bilgisi sadece benzediği kromozom grupları içinde değerlendirilerek A-G arası yedi gruba yerleştirilmişlerdir. Tablo 6.1'de boy bilgilerine göre gruplandırma sonuçları görülen 10 örneğin sentromer bilgilerinin değerlendirilmesi sonucunda elde edilen gruplandırma sonuçları Tablo 6.5'de verilmiştir. Tablo 6.1'de yanlış grupta olan kromozomların sentromer özelliklerinin değerlendirilmesinden sonra doğru gruplara yerleştiğini vurgulamak için bu kromozomlar Tablo 6.5'de koyu renkte gösterilmiştir.



Tablo 6.5 Sentromerlerin Değerlendirilmesi Sonucunda Gruplandırma Sonucu

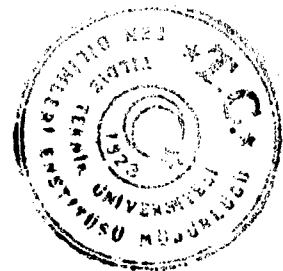
| Grup | Kromozom Numaraları | Grup Başarı (%) | Grup | Kromozom Numaraları | Grup Başarı (%) |
|------|-------------------------------------|-----------------|------|-------------------------------------|-----------------|
| A | 1 2 1 2 3 3 | 100 | A | 1 1 2 2 3 3 | 100 |
| B | 5 4 4 5 | 100 | B | 4 5 5 4 | 100 |
| C | 6 6 X 7 7 9 8 11 8 11 10 9 10 12 12 | 100 | C | 6 6 7 9 11 11 7 8 10 8 X 9 10 12 12 | 100 |
| D | 13 15 14 15 13 14 | 100 | D | 13 15 15 13 14 14 | 100 |
| E | 17 16 17 16 18 18 | 100 | E | 16 17 17 18 18 16 | 100 |
| F | 20 19 20 19 | 100 | F | 19 20 19 20 | 100 |
| G | 22 22 21 21 Y | 100 | G | 22 22 21 21 Y | 100 |
| A | 1 2 1 2 3 3 | 100 | A | 1 1 2 2 3 3 | 100 |
| B | 4 4 5 5 | 100 | B | 4 4 5 5 | 100 |
| C | 8 8 7 6 6 9 11 7 X 11 10 10 9 X 12 | 100 | C | 6 6 7 X X 7 9 8 8 10 9 10 11 12 11 | 100 |
| D | 13 13 15 14 15 14 | 100 | D | 13 15 13 14 14 15 | 100 |
| E | 16 16 17 17 18 18 | 100 | E | 16 17 16 18 17 18 | 100 |
| F | 19 20 19 20 | 100 | F | 19 19 20 20 | 100 |
| G | 22 22 21 21 | 100 | G | 22 21 21 22 | 100 |
| A | 1 1 2 2 3 3 | 100 | A | 1 1 2 2 3 3 | 100 |
| B | 5 4 5 4 | 100 | B | 4 5 4 5 | 100 |
| C | 6 6 7 7 X 8 X 11 12 10 10 11 8 9 | 100 | C | 6 6 7 7 8 8 9 X 11 9 10 10 12 11 | 100 |
| D | 13 14 13 15 15 14 | 100 | D | 13 14 15 13 14 15 | 100 |
| E | 16 17 16 17 18 18 | 100 | E | 16 17 17 16 18 18 | 100 |
| F | 19 20 19 21 | 75 | F | 19 19 20 20 | 100 |
| G | 20 21 22 22 | 75 | G | 22 21 Y 22 21 | 100 |
| A | 1 2 2 1 3 3 | 100 | A | 1 1 2 2 3 3 | 100 |
| B | 4 5 5 4 | 100 | B | 4 5 4 5 | 100 |
| C | 6 6 10 7 7 10 X X 8 8 9 9 11 12 11 | 100 | C | 6 6 7 7 8 X 9 8 9 10 12 10 11 11 12 | 100 |
| D | 13 13 14 15 14 15 | 100 | D | 13 13 14 15 14 15 | 100 |
| E | 16 16 17 17 18 18 | 100 | E | 16 16 17 17 18 18 | 100 |
| F | 19 20 19 20 | 100 | F | 20 19 19 20 | 100 |
| G | 21 22 22 21 | 100 | G | 22 21 21 22 Y | 100 |
| A | 1 1 2 2 3 3 | 100 | A | 1 1 2 2 3 3 | 100 |
| B | 4 5 4 5 | 100 | B | 4 5 4 5 | 100 |
| C | 6 6 7 X 7 8 X 8 10 11 10 11 9 9 12 | 100 | C | 6 6 8 10 12 7 8 X X 9 10 9 7 8 12 | 100 |
| D | 13 14 13 14 15 15 | 100 | D | 13 15 15 14 13 14 | 100 |
| E | 16 16 18 17 17 18 | 100 | E | 16 16 17 17 18 18 | 100 |
| F | 20 19 19 20 | 100 | F | 20 20 19 19 | 100 |
| G | 21 22 22 21 | 100 | G | 22 21 22 21 | 100 |

6.3 Kromozomların Gruplara Ayrılması

Bu çalışmada kromozomların gruplara ayrılması işlemi iki adımda yapılmaktadır. İlk adımda, kromozomlar boy ve alan bilgileri kullanılarak bulanık mantık yöntemleri ile benzerliklerine göre sınıflandırılmakta, ikinci aşamada ise her sınıftaki kromozomların sentromer özellikleri değerlendirilerek (A-G) arası yedi grup oluşturulmaktadır.



Kromozomların alan değerleri için elde edilen büyükten küçüğe sıralı verilerden bulanık mantık alan üyelik derecesi fonksiyonu yardımıyla benzerlik ilişkisi matrisi oluşturulur. Bu matris üzerinde yapılan kompozisyonlarla eşdeğerlik ilişkisi matrisi elde edilir. Matris elemanlarına α -kesmesi uygulanarak berrak ilişki matrisi oluşturulur. 1 ve 0'lardan oluşan bu matris üzerinde 1'lerin kümelendiği bölgeler kromozomun alan gruplarını oluşturur. Aynı işlem sırasıyla boy grupları elde edilir. Alan ve boy kümelerinin kesişimi ile kromozomlar alan ve boy benzerliklerine göre sınıflandırılmış olur. Bu işlem sonucunda kaç sınıf elde edileceği, seçilen α -kesmesinin değerine bağlıdır. α -kesmesinin değeri büyüdükçe sınıflandırmanın hassasiyeti artar. Boy ve alan benzerliklerine göre elde edilen bu sınıflardaki kromozomların sentromer özellikleri değerlendirilerek kromozomların (A-G) arası yedi gruba ayrılması sağlanır. Bunun için kromozomların sentromer oranları değerlendirilerek bulanık mantık sentromer benzerlik ilişkisi matrisleri elde edilmiştir. Boy, alan benzerlik ilişkisi matrisleri ile sentromer benzerlik ilişkisi matrislerinin kesişim kümesinin hesaplanmasıyla kromozomların yedi gruba yerleştirilmesi işlemi tamamlanmıştır. Bu işlemi şöyle bir örnekle açıklayabiliriz. Boy ve alan bilgisi olarak en başta bulunan kromozomun (1) numaralı kromozom olması gerekir. Eğer bu kromozomun sentromer özelliği de S_1 ise bu kromozomun (1) numaralı kromozom olabileceği ve A grubuna ait olduğu söylenebilir. Eğer sentromer özelliği farklı ise, (1) numaralı kromozomu bulmak için aynı sınıftaki sentromer özelliği S_1 olan en uzun kromozom aranır. Eğer bulunamazsa, sentromer özelliği S_1 'e en çok benzeyen en uzun kromozomun (1) numaralı kromozom olduğu kabul edilir. Bu şekilde yapılan gruplandırma işleminde (3) numaralı kromozomun da sentromer özelliği S_1 olduğu için boy-alan bilgisine göre daha önde yer almışsa (3) numaralı kromozom (1) numaralı kromozomla karıştırılabilir. Ama bu aşamada yapılan işlem sadece doğru grupları oluşturmaktır. Bir sonraki aşamada her grup içindeki kromozomlar bant bilgilerine göre değerlendirilerek tanıma işlemi yapılacaktır.



6.4 Kromozomların Tanınması

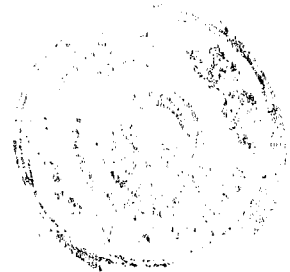
Boy-alan ve sentromer bilgilerinden yararlanılarak A-G gruplarına ayrılmış olan kromozomların sınıflandırılması ve hangi kromozom olduğunun belirlenmesi amacıyla bant bilgilerinden yararlanılmıştır.

ISCN 1995’de tarif edildiği gibi, her kromozomun diğer kromozomlardan farklı bir bant dağılımı olmasına rağmen bu bilgi ideal durumu göstermektedir. Bir kromozoma ait bant bilgileri, kullanılan boyama yönteminin hassasiyetine, boyamanın gerçekleştirildiği ortamın şartlarına bağlı olarak ISCN 1995’de belirtilen ideal durumdan oldukça fazla sapma gösterebilmektedir. Bu durumu göz önüne alarak her kromozoma ait tüm bant bilgilerini değerlendirmek yerine karakteristik özellik gösteren bantlar ile ilgilenmek ve bu bantları temel alan sınıflandırma mekanizmaları kullanmak yerinde olacaktır. Bu konuda çalışma yapan uzmanların da bantları tek tek sayarak ISCN 1995’de verilenlerle karşılaştırmak yerine karakteristik bantların varlığına ve yerine göre karar vermeleri de bu konuda uygulanacak yöntem için ayrı bir çıkış noktası olmuştur. Karakteristik bant bilgisinin en az duyarlıklı boyama tekniklerinin kullanılması durumunda bile görünür hale gelmesinden yola çıkılarak kromozomlara Histogram Eşleme yönteminin uygulanması ile bu karakteristik bantlar daha belirgin ve işlenebilir hale getirilmiştir. Farklı kromozom görüntülerinde aynı numaralı kromozomların gri tonlarının birbirine göre koyulu açıklı olması, karakteristik bantların birbirlerine göre uzaklıklarının, yerlerinin ve genişliklerinin farklılıklar göstermesi, görüntü işleme teknikleri ile bu kromozomların boyları eşitlene bile basit matematiksel bir karşılaştırma yapılmasını imkansız kılmaktadır. Böyle basit bir karşılaştırma yerine her kromozomun fonksiyonunun ne olduğu bilinmeyen rasgele değişkenler (random variable) ile ifade edilmesi ve bu rasgele değişkenlerin birbirlerine benzerliğinin korelasyon yöntemi ile belirlenerek kromozom benzerliğinin tespit edilmesi yoluna gidilmiştir. Histogram Eşleme sonunda bant bilgileri siyah, gri ve beyaz bantlar için 1,2,3 olmak üzere üç ayrık değere getirildiği ve korelasyon hesabında karşılaştırılan rasgele değişkenlerin hem ortalama değerleri hem de varyans ve kovaryansları işleme girdiği için kromozomların gri tonları açısından gösterdikleri farklılıklar elde edilen sonucu en az etkiler hale getirilmiştir.



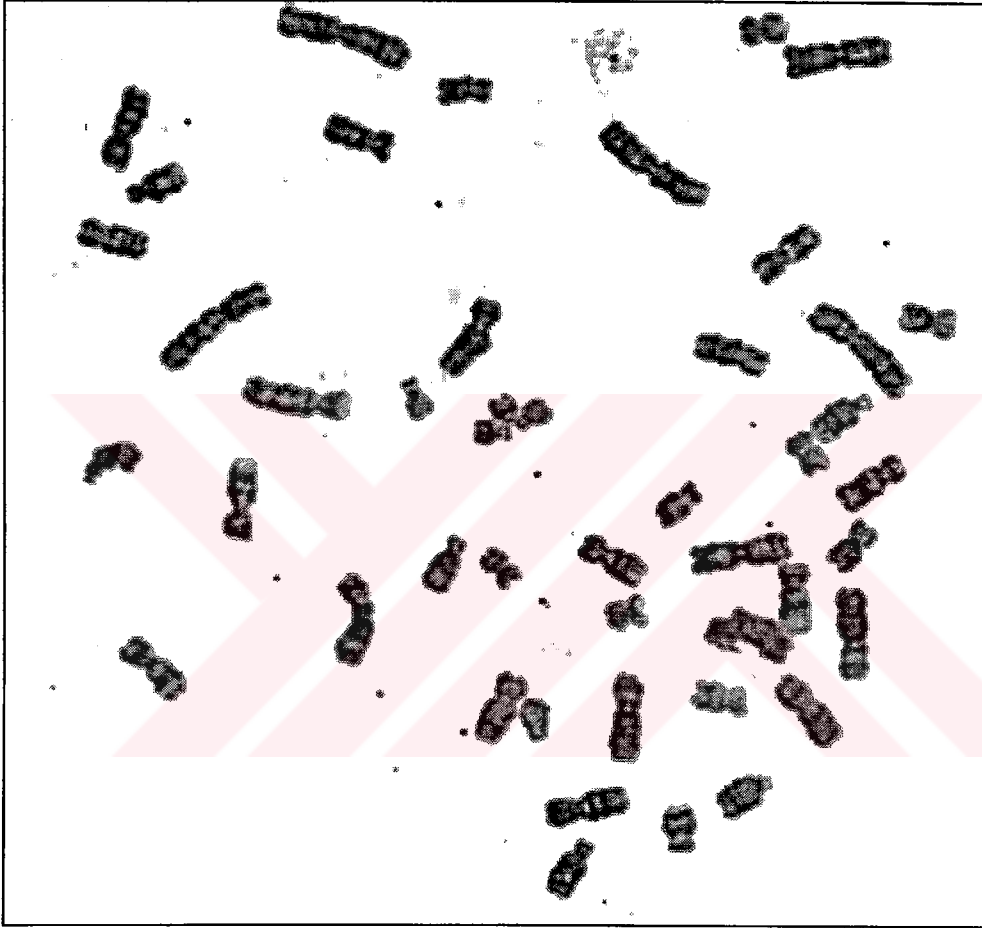
Korelasyonda kullanılmak amacıyla öncelikle her grup (A-G) için 30 kromozom görüntüsündeki 60*22 otozom ve 30*2 cinsiyet kromozomunun kullanılmasıyla elde edilen şablonlar (template) hazırlanmıştır. Şablon hazırlama işleminde her kromozoma ait örnekler öncelikle görüntü işleme teknikleri ile aynı boya getirilmiş ve satır satır incelenerek her satırda ağırlıklı olarak yer alan piksel değeri şablonda o satıra ait piksel değeri olarak kabul edilmiştir.

Hangi kromozom olduğu belirlenecek kromozomun ait olduğu grup daha önceki adımlarda boy, alan ve sentromer bilgilerinden yararlanılarak belirlendiği için, bu kromozom şablon kromozomlar içinde yalnızca içinde bulunduğu gruba ait kromozomlar ile korelasyon işlemine girer. Hesaplanan korelasyon katsayıları -1 ile +1 arasında değerler alır. Burada -1 hiç benzeşmemeyi, +1 ise tam benzeşmeyi göstermektedir. Tüm katsayıların negatif çıkması durumunda ise test edilen kromozomlar arasında ters benzerlik söz konusu demektir. Elde edilen korelasyon katsayıları arasından maksimum olan, test edilen kromozomun o grup içinde en çok hangi kromozoma benzediğini belirler.



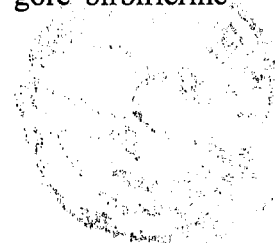
6.5 Örnek Uygulama

Geliştirilen Kromozom Sınıflandırma ve Tanıma Sisteminin çalışmasını Şekil 6.3'deki örnek görüntü üzerinde adım adım inceleyelim:



Şekil 6.3 Örnek kromozom görüntüsü

Örnek görüntünün oransal alan değerleri için Tablo 6.6'daki alan üyelik dereceleri matrisi elde edilmiştir. Kompozisyon sonunda Eşdeğerlik İlişkisi (Benzerlik Kümeleri) Matrisi Tablo 6.7'deki gibi elde edilmiştir. $\alpha = 93$ için Berrak İlişki Matrisi Tablo 6.8'de verilmiştir. Alan oranları benzerlik matrisi ise Tablo 6.9'da verilmiştir. Bu tabloda (1-4), (5-25), (26-43), (44-46) aralıklarındaki kromozomların alan bilgisine göre birbirlerine



benzedikleri görülmektedir. Aynı işlemler boy oranları bilgileri için de tekrarlanarak bu iki matrisin kesişiminden alan ve boy oranları için benzerlik matrisi elde edilmiş olur.

Sentromer oranları için de önce Eşdeğerlik İlişkisi matrisi (Tablo 6.10) daha sonra da bu matristen yararlanarak Berrak İlişki matrisi (Tablo 6.11) hesaplanmıştır. Sentromer oranları Berrak İlişki matrisi ile boy-alan benzerlik matrisinin kesişim kümesi ile (A-G) grupları elde edilmiştir.

Gruplandırma işlemi tamamlandıktan sonra her kromozomun bant bilgileri içinde bulunduğu kromozom grubundaki şablon kromozomların bant bilgileri ile korelasyon işlemine sokulduğunda Tablo 6.12'de görülen sonuçlar elde edilmiştir. Buna göre örneğin (1) numaralı test kromozomu (A) grubunda yer aldığı için 1,2,3 numaralı kromozomların şablonları ile karşılaştırılmıştır. Elde edilen sonuçlara göre en büyük değer Şablon-1 ile korelasyona girmesi sonucunda elde edilen 0.886 değeri olduğu için, (1) numaralı test kromozomunun gerçekte (1) numaralı kromozom olduğuna karar verilmiştir. Bu tabloda **sonuç değer**, korelasyon işlemine göre sistemin verdiği kararı, **gerçek değer** ise gerçekte bu kromozomun kaç numaralı kromozom olduğunu göstermektedir. Sonuçlar incelendiğinde, bu örnek için 7,9,11,12,22 ve Y kromozomlarının yanlış tanındığı görülür. Bu sonuçlara göre elde edilen karyotip Şekil 6.4'de verilmiştir. Şekil 6.4'de yanlış tanınan kromozomların hangi kromozomlar olduğu altı çizili olarak gösterilmiştir.

6.6 Deneysel Sonuçlar ve Sistemin Başarısı

Sistemin 40 kromozom görüntüsü üzerinde uygulanması sonucunda elde edilen başarı Tablo 6.13'de verilmiştir. Tanıma işleminin başarısını değerlendirebilmek için her kromozomun kaç defa doğru tanındığı ve hangi kromozomlarla karıştırıldığı hesaplanarak Tablo 6.14'de gösterilmiştir. Her satıra ait başarı oranı ise Tablo 6.15'de verilmiştir. Kromozomları tanıma işleminin örnek görüntüdeki bütün kromozomlara göre değil sadece içinde buldukları gruba bağlı olarak yapılmasının sisteminin başarısını arttırdığı ve bant bilgileri benzese de şekil olarak hiç benzemediği kromozomlarla karıştırılmasının önlenildiği görülmüştür.

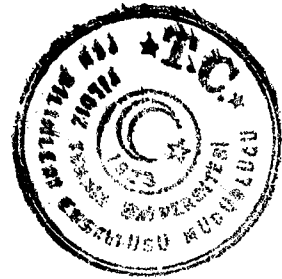


Tablo 6.10 Sentromer oranları eşdeğerlik ilişkisi matrisi (0 değerleri 100 değerine denk düşmektedir)

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | | |
| 1 | 0 | 0 | 90 | 0 | 0 | 61 | 59 | 61 | 54 | 92 | 90 | 90 | 92 | 90 | 80 | 88 | 83 | 80 | 73 | 92 | 97 | 52 | 64 | 47 | 42 | 47 | 47 | 40 | 45 | 0 | 95 | 71 | 85 | 88 | 78 | 0 | 0 | 97 | 59 | 59 | 92 | 85 | 95 | 95 | | | |
| 2 | 82 | 83 | 0 | 0 | 89 | 85 | 86 | 83 | 86 | 76 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 73 | 90 | 66 | 60 | 66 | 66 | 66 | 56 | 63 | 95 | 0 | 0 | 0 | 0 | 0 | 83 | 91 | 97 | 0 | 83 | 83 | 0 | 0 | 0 | | |
| 3 | 62 | 63 | 81 | 81 | 67 | 64 | 0 | 0 | 95 | 79 | 81 | 81 | 79 | 81 | 91 | 83 | 88 | 91 | 0 | 79 | 75 | 91 | 0 | 83 | 75 | 83 | 83 | 70 | 79 | 72 | 77 | 0 | 86 | 83 | 93 | 63 | 68 | 73 | 75 | 0 | 0 | 79 | 86 | 77 | 77 | | |
| 4 | 44 | 44 | 57 | 57 | 47 | 45 | 84 | 88 | 84 | 95 | 56 | 57 | 56 | 57 | 64 | 59 | 62 | 64 | 70 | 56 | 53 | 0 | 81 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 51 | 55 | 73 | 61 | 59 | 66 | 44 | 48 | 52 | 53 | 88 | 88 | 56 | 61 | 55 | 55 |

Tablo 6.11 Sentromer oranları berrak ilişki matrisi

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | | | | |
| 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | | |
| 2 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | |
| 3 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

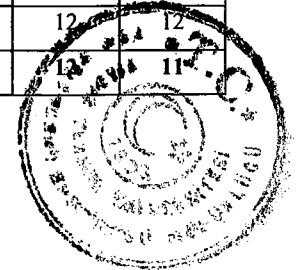


Tablo 6.12 Örnek görüntü için kromozom tanıma sonuçları: Her kromozom içinde bulunduğu grupların şablon kromozomları ile korelasyon işlemine sokulmuş ve en çok benzediği şablon koyu renkte gösterilmiştir.

| A Grubu(1-3) | | | | | |
|--------------|--------------|--------------|--------------|-------------|--------------|
| No | Şablon-1 | Şablon-2 | Şablon-3 | Sonuç Değer | Gerçek Değer |
| 1 | 0.886 | 0.312 | 0.178 | 1 | 1 |
| 2 | 0.805 | 0.221 | 0.141 | 1 | 1 |
| 3 | 0.335 | 0.759 | -0.277 | 2 | 2 |
| 4 | 0.206 | 0.590 | -0.134 | 2 | 2 |
| 5 | 0.086 | -0.376 | 0.922 | 3 | 3 |
| 6 | 0.201 | -0.136 | 0.636 | 3 | 3 |

| B Grubu(4-5) | | | | |
|--------------|--------------|--------------|-------------|--------------|
| No | Şablon-4 | Şablon-5 | Sonuç Değer | Gerçek Değer |
| 7 | 0.691 | -0.269 | 4 | 4 |
| 8 | 0.107 | 0.872 | 5 | 5 |
| 9 | 0.563 | 0.195 | 4 | 4 |
| 10 | 0.268 | 0.706 | 5 | 5 |

| C Grubu(6-12,X) | | | | | | | | | | |
|-----------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------------|--------------|
| No | Şablon 6 | Şablon 7 | Şablon 8 | Şablon 9 | Şablon 10 | Şablon 11 | Şablon 12 | Şablon X | Sonuç Değer | Gerçek Değer |
| 11 | 0.656 | 0.126 | 0.15 | -0.241 | 0.217 | 0.198 | -0.316 | 0.437 | 6 | 6 |
| 12 | 0.890 | -0.358 | -0.086 | -0.167 | 0.501 | -0.231 | -0.234 | -0.362 | 6 | 6 |
| 13 | -0.103 | 0.879 | -0.178 | 0.623 | -0.248 | 0.100 | 0.009 | 0.007 | 7 | 7 |
| 14 | -0.202 | 0.593 | 0.226 | 0.761 | -0.159 | 0.001 | -0.069 | -0.247 | 7 | 9 |
| 15 | -0.333 | 0.245 | 0.453 | 0.894 | -0.405 | -0.294 | 0.074 | -0.049 | 9 | 9 |
| 16 | -0.551 | -0.328 | 0.236 | 0.007 | -0.316 | 0.388 | 0.076 | 0.751 | X | X |
| 17 | -0.042 | -0.068 | -0.005 | -0.396 | 0.388 | 0.754 | 0.586 | -0.583 | 11 | 12 |
| 18 | 0.040 | -0.004 | 0.855 | -0.009 | -0.158 | -0.369 | -0.140 | -0.101 | 8 | 8 |
| 19 | -0.062 | 0.330 | -0.226 | -0.287 | 0.004 | 0.698 | -0.148 | -0.502 | 11 | 11 |
| 20 | 0.307 | -0.052 | -0.453 | -0.342 | 0.772 | -0.155 | -0.382 | -0.108 | 10 | 10 |
| 21 | -0.352 | 0.153 | 0.623 | 0.016 | -0.292 | 0.155 | -0.006 | -0.177 | 8 | 8 |
| 22 | -0.192 | -0.189 | 0.071 | -0.067 | 0.548 | 0.135 | 0.064 | -0.445 | 10 | 10 |
| 23 | -0.417 | 0.442 | 0.000 | 0.691 | -0.085 | 0.000 | 0.148 | -0.269 | 9 | 7 |
| 24 | 0.040 | -0.198 | -0.151 | 0.462 | -0.158 | -0.300 | 0.878 | 0.266 | 12 | 12 |
| 25 | -0.02 | 0.055 | -0.244 | 0.203 | -0.329 | 0.524 | 0.853 | 0.152 | 12 | 11 |



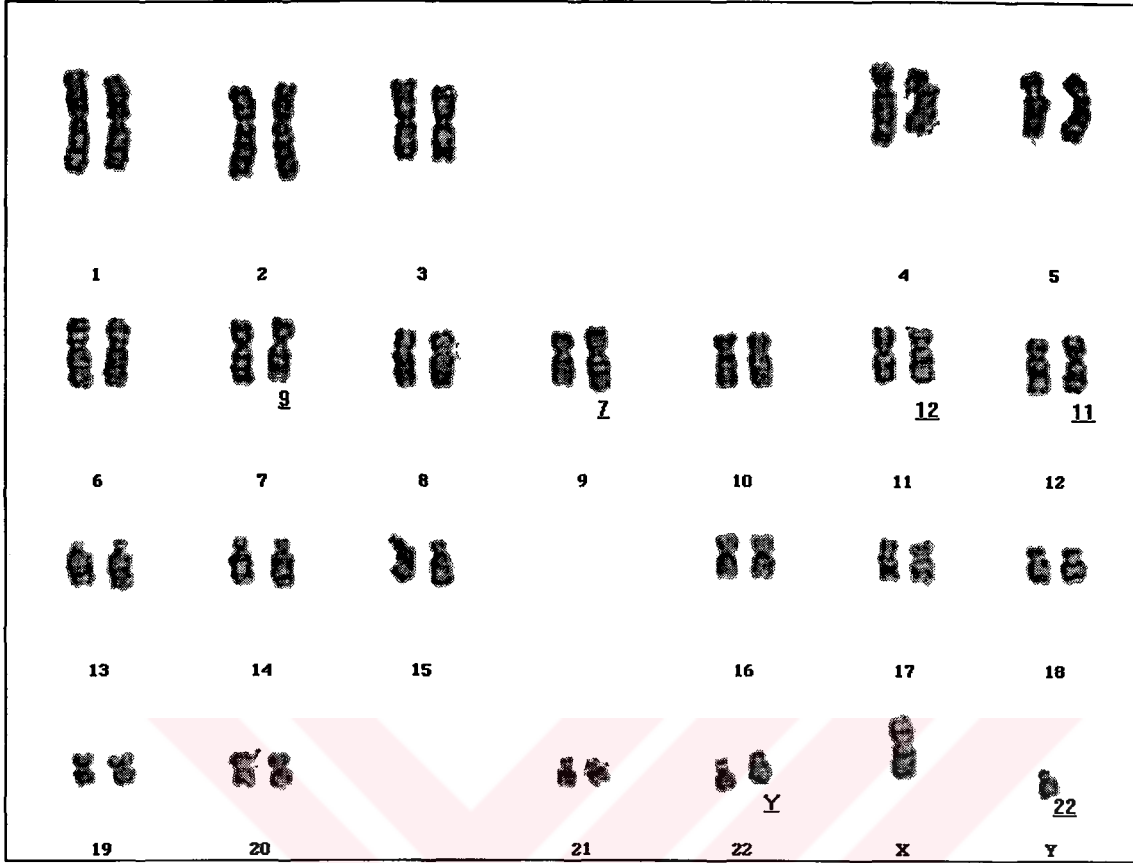
| D Grubu(13-15) | | | | | |
|-----------------------|--------------|--------------|--------------|----------------|-----------------|
| No | Şablon 13 | Şablon 14 | Şablon 15 | Sonuç Değer | Gerçek Değer |
| 26 | 0.913 | -0.402 | -0.376 | 13 | 13 |
| 27 | -0.530 | 0.472 | 0.416 | 14 | 14 |
| 28 | -0.140 | 0.000 | 0.837 | 15 | 15 |
| 29 | 0.669 | -0.116 | 0.255 | 13 | 13 |
| 30 | -0.294 | 0.427 | 0.010 | 14 | 14 |
| 31 | -0.060 | 0.242 | 0.762 | 15 | 15 |

| E Grubu(16-18) | | | | | |
|-----------------------|--------------|--------------|--------------|----------------|-----------------|
| No | Şablon 16 | Şablon 17 | Şablon 18 | Sonuç Değer | Gerçek Değer |
| 32 | 0.741 | -0.146 | 0.756 | 16 | 16 |
| 33 | 0.796 | 0.161 | 0.436 | 16 | 16 |
| 34 | -0.016 | 0.546 | -0.342 | 17 | 17 |
| 35 | 0.362 | 0.000 | 0.945 | 18 | 18 |
| 36 | 0.465 | 0.409 | 0.072 | 17 | 17 |
| 37 | 0.440 | 0.000 | 0.496 | 18 | 18 |

| F Grubu(19-20) | | | | |
|-----------------------|--------------|--------------|----------------|-----------------|
| No | Şablon 19 | Şablon 20 | Sonuç Değer | Gerçek Değer |
| 38 | 0.428 | -0.017 | 19 | 19 |
| 39 | 0.624 | 0.017 | 19 | 19 |
| 40 | 0.224 | 0.498 | 20 | 20 |
| 41 | 0.513 | 0.695 | 20 | 20 |

| G Grubu(21-22,Y) | | | | | |
|-------------------------|--------------|--------------|-------------|----------------|-----------------|
| No | Şablon 21 | Şablon 22 | Şablon Y | Sonuç Değer | Gerçek Değer |
| 42 | -0.353 | 0.683 | 0.267 | 22 | 22 |
| 43 | -0.152 | 0.792 | 0.562 | 22 | Y |
| 44 | -0.708 | 0.302 | 0.456 | Y | 22 |
| 45 | 1.000 | 0.475 | 0.354 | 21 | 21 |
| 46 | 1.000 | -0.292 | 0.126 | 21 | 21 |





Şekil 6.4 Sınıflandırma ve tanıma işlemin sonunda örnek görüntüye ait karyotip



Tablo 6.14 40 kromozom örneğine ait sınıflandırma ve tanıma işleminin sonucu

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | X | Y | T | % | | |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|----|---|----|----|----|
| 1 | 73 | 3 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 73 | 91 | |
| 2 | 5 | 73 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 73 | 91 |
| 3 | 2 | 4 | 74 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 74 | 92 |
| 4 | 0 | 0 | 0 | 75 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 75 | 93 |
| 5 | 0 | 0 | 0 | 5 | 75 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 75 | 93 |
| 6 | 0 | 0 | 0 | 0 | 0 | 69 | 1 | 2 | 3 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 69 | 86 |
| 7 | 0 | 0 | 0 | 0 | 0 | 3 | 67 | 1 | 3 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 67 | 83 |
| 8 | 0 | 0 | 0 | 0 | 0 | 1 | 3 | 72 | 3 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 72 | 90 |
| 9 | 0 | 0 | 0 | 0 | 0 | 1 | 4 | 2 | 66 | 5 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 66 | 82 |
| 10 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 5 | 64 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 64 | 80 |
| 11 | 0 | 0 | 0 | 0 | 0 | 3 | 2 | 1 | 1 | 4 | 65 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 65 | 81 |
| 12 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 3 | 0 | 1 | 7 | 73 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 60 | 75 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 74 | 3 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 74 | 92 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 73 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 73 | 91 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 3 | 72 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 72 | 90 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 73 | 4 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 73 | 91 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 72 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 72 | 90 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 5 | 72 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 72 | 90 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 73 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 73 | 91 |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 73 | 0 | 0 | 0 | 0 | 0 | 0 | 73 | 91 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 74 | 4 | 0 | 2 | 0 | 0 | 74 | 92 |
| 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 74 | 0 | 3 | 0 | 0 | 74 | 92 |
| X | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 1 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 45 | 0 | 0 | 45 | 77 |
| Y | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 3 | 3 | 16 | 69 | |



Tablo 6.15 40 kromozom görüntüsünün tanınması sonucunda elde edilen başarı

| Test No | Başarı Oranı (%) | Test No | Başarı Oranı (%) |
|---------|------------------|---------|------------------|
| 1 | 91 | 21 | 85 |
| 2 | 93 | 22 | 89 |
| 3 | 95 | 23 | 89 |
| 4 | 83 | 24 | 89 |
| 5 | 93 | 25 | 89 |
| 6 | 93 | 26 | 85 |
| 7 | 89 | 27 | 91 |
| 8 | 83 | 28 | 85 |
| 9 | 85 | 29 | 89 |
| 10 | 87 | 30 | 95 |
| 11 | 83 | 31 | 91 |
| 12 | 83 | 32 | 89 |
| 13 | 91 | 33 | 93 |
| 14 | 83 | 34 | 80 |
| 15 | 87 | 35 | 91 |
| 16 | 89 | 36 | 87 |
| 17 | 95 | 37 | 87 |
| 18 | 91 | 38 | 80 |
| 19 | 85 | 39 | 89 |
| 20 | 89 | 40 | 85 |



7. SONUÇ

Hücre içinde genetik bilgilerin depolandığı kromozomların sınıflandırılması, genetik araştırmaların esasını oluşturan temel unsurlardan birisidir. Yapılan bu doktora çalışmasında, laboratuvar ortamında ışık mikroskobundan elde edilen ve kamera aracılığıyla bilgisayara aktarılan kromozom görüntülerinin boy, alan, sentromer yeri ve bant bilgilerinin bulanık mantık yöntemleri ile sınıflandırılarak tanınmasına yönelik bir sistem geliştirilmiştir.

Bu tez çalışmasının özelliği, kromozomların konuyla ilgili bilimsel yayınlarda yer alan tüm özelliklerinin birebir kontrol edilmesi yerine genetik uzmanlarının kişisel sezgilerini de kullanarak her kromozomu belirleyen temel özelliklere göre gerçekleştirdikleri sınıflandırma ve tanıma işleminin bulanık mantık yöntemlerinin uygulanması ile bilgisayar ortamında başarı ile gerçekleştirilmiş olmasıdır. Bu tez çalışması kapsamında değişik aşamalarda benzerlik sınıflarının oluşturulması için, kromozomlar arasında benzerliği gösteren üyelik derecesi fonksiyonları tarif edilmiş, buradan hareketle benzerlik matrisleri oluşturulmuştur. Sadece yansıma ve simetri özellikleri olan bu benzerlik ilişkisi matrisleri kompozisyon işlemlerine sokularak eşdeğerlik ilişkisi matrisi elde edilmiştir. Bu işlemde, bu doktora çalışması sırasında geliştirilen ardışık kare alma metodu kullanılarak işlem performansı ($n/\log_2 n$) oranında artırılmıştır. Daha sonra adaptif α -kesmeleri kullanılarak kromozom sınıflarının benzerlik sınıfı (berrak ilişki) matrisleri elde edilmiştir.

Boy ve sentromer özellikleri değerlendirilerek ideal durumda boylarına göre büyükten küçüğe sıralı olacak şekilde (A-G) harfleri ile gösterilen yedi gruba ayrılan kromozomlara ait görüntüler incelendiğinde, boy sıralamasının özellikle boyları birbirine yakın olan kromozomlarda karışabildiği görülmüştür. Kromozomların sadece boy özelliklerinin değerlendirilmesi sonucunda oluşabilecek hataları ortadan kaldırmak amacıyla boy bilgisinin yanı sıra alan bilgisinin de değerlendirilmesi yoluna gidilmiş ve kromozomların boy ve alan değerleri için bulanık mantık benzerlik sınıfları elde edilmiştir.



Boy, alan ve sentromer özellikleri, kromozomları (A-G) gruplarına yerleştirmek için yeterli olmalarına rağmen bu bilgilere bağlı olarak (A-G) grupları oluşturulurken dikkat edilmesi gereken en önemli problem boyları benzeyen ve sentromer özelliği de aynı olan fakat farklı gruplarda olması gereken kromozomların gruplanmasının doğru yapılabilmesidir. Özellikle birbirine komşu grupların kromozomlarının boyları yakın olduğu için, sentromer özellikleri de aynı olduğunda gruplar arası karışmaları engellemek güçleşecektir. Kromozomların sentromer yerleri değerlendirildiğinde, submetasentrik ve akrosentrik kromozomların sentromer oranı aralığının çok geniş olduğu ve buna bağlı olarak sentromer yerindeki farklılık gözle ayırt edilebilecek kadar belirgin olan bazı kromozomların aynı sentromer grubunda yer aldığı görülmüş ve gruplandırma başarısını arttırabilmek için, sentromer özellikleri idealde belirlenen metasentrik, submetasentrik ve akrosentrik özelliklere göre değil, sentromer yeri benzerliklerine göre dört grupta (S_1 , S_2 , S_3 , S_4) incelenmişlerdir. Kromozomlara ait sentromer oranlarının S_1 , S_2 , S_3 ve S_4 gruplarına göre bulanık mantık sentromer benzerlik ilişkisi matrisleri elde edilmiştir. Boy, alan benzerlik ilişkisi matrisleri ile sentromer benzerlik ilişkisi matrislerinin kesişim kümesinin hesaplanmasıyla kromozomların yedi gruba ayırma işlemi tamamlanmıştır.

Sentromer özelliklerinin dört grupta incelenmesi, kromozomları yedi gruba yerleştirme işleminin başarısını arttırması bakımından önemli bir özelliktir. Boy, alan ve sentromer bilgilerine göre elde edilen bu yedi grup sayesinde şekilleri hiç benzemeyen kromozomların bant özellikleri benzediği için birbirlerine karıştırılmaları önlenmiştir.

Bant bilgilerinin değerlendirilmesinde, her karyotipte idealde tanımlı bütün bantların olmadığı ve elle karyotip hazırlamasında genetik uzmanlarının da kromozomun bantlarını tek tek saymayıp, karakteristik bantların varlığına ve yerine göre karar vermeleri dikkate alınarak, kromozomların tanınması işleminde bütün bant bilgilerini değerlendirmek yerine, karakteristik bant özelliklerinin değerlendirilmesi öngörülmüştür. Karakteristik bant özelliklerini belirlemek amacıyla, 30 örnek kromozom görüntüsündeki 660 otozom ve 60 cinsiyet kromozomunun bant bilgilerinden kromozom şablonları oluşturulmuştur. Kromozom görüntüsünün alındığı ortama bağlı olarak gri tonlarında oluşan farklılıkları dengelemek için Histogram Eşleme yöntemi ile bant değerleri beyaz, gri ve siyah karşı



düŖen 1,2,3 sayıları ile gösterilmiŖlerdir. Aynı boya getirildikten sonra ađırlıklı olarak bulunan piksellere bađlı olarak karakteristik bant bilgileri kaybedilmeyecek Ŗekilde Ŗablonlar hazırlanmıŖtır. Her kromozomun kaç numaralı kromozom olduđu, bulunduđu gruba ait Ŗablon kromozomlarla korelasyon yöntemi ile karŖılaŖtırılarak belirlenmiŖtir.

Sonuç olarak bu tez çalıŖmasında, kromozomlar ayırt edici özelliklerine göre bu konuda daha önce yapılan çalıŖmalarda örneđi görülmeyen ve bu tez çalıŖmasında geliŖtirilen, aŖađıda ayrıntısı verilen üç kademeli bir eleme yöntemi ile sınıflandırılmıŖ ve tanınmıŖlardır.

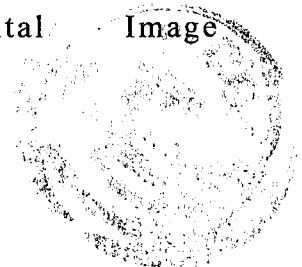
- Birinci adımda kromozomlar boy ve alan bilgilerinin benzerliklerine göre sınıflara ayrılmıŖlardır.
- İkinci aŖamada, boy, alan benzerliklerine göre sınıflara ayrılan her kromozomun sentromer oranı yalnızca içinde bulunduđu sınıftaki kromozomlara göre deđerlendirilerek boyları benzeyen ama sentromer özellikleri farklı olan kromozomların ayırt edilebilmeleri sađlanmıŖtır.
- Üçüncü aŖamada, her kromozomun bant bilgisi içinde bulunduđu gruba ait Ŗablon bant bilgileri ile deđerlendirilerek tanıma iŖlemi tamamlanmıŖ, her kromozomun kaç numaralı kromozom olduđu belirlenmiŖtir.

Son aŖamada kullanılan korelasyon yönteminin temel aldıđı Ŗablon yapısı, sistemin deđerŖen kromozom yapılarına daha rahat uyum sađlamasını ve Ŗablona eklenen her dođru tanınmıŖ yeni kromozom bilgisi ile sistem baŖarısının artmasını sađlamıŖtır. Hem bulanık mantık yöntemlerinin hem de adaptif Ŗablon yapısının kullanılması geliŖtirilen sistemin belirleyici ve özgün özellikleridir.

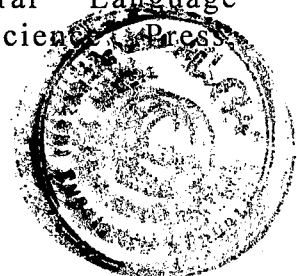


KAYNAKLAR

- Başaran, N., (1996), Tıbbi Genetik, Bilim Teknik Yayınevi, İstanbul
- Bezdek, J., (1981), Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, NY
- Carothers, A. ve Piper J., (1994), 'Computer-Aided Classification of Human Chromosomes: A Review', Statistics and Computing, vol.4, no.3, 161-171
- Cohen, P. ve Feigenbaum E., (1982), The Handbook of Artificial Intelligence, Addison-Wesley, Massachusetts
- Dyer, C., (1979), 'Thinning Algorithms for Grayscale Pictures', IEEE Transactiond on Pattern Analysis and Machine Intelligence, vol.PAMI-5, no.6, pp:621-623
- Fu, K.S., (1982), Syntactic Pattern Recognition and Applications, Prentice-Hall, N.J.
- Fukunaga, K., (1990), Introduction to Statistical Pattern Recognition, Academic Press
- Guthrie, C., Gregor, J. ve Thomson, M., (1993), 'Constrained Markov Networks for Automated Analysis of G-banded Chromosomes', Comput. Biol. Med., vol 2, 105-114
- Gonzalez, R.C., ve Wintz, P., (1987), Digital Image Processing, Addison-Wesley Publishing Company, Massachusetts.
- Graham, J. ve Piper, J., (1994), 'Automatic Karyotype Analysis', Methods in Molecular Biology, vol.29, 141-183
- Granum, E. ve Thomason, MG., (1990), 'Automatically Inferred Markov Network Models for Classifacation of Chromosomal Band Pattern Structures', Cytometry, vol. 11.1, pp:26-39
- Gregor, J. ve Granum, E., (1991), 'Finding Chromosome Centromeres using Band Pattern Information', Comput. Biol. Med., vol. 21, 55-67
- Ibach P., (1987), 'Flaechenfüll-Algorithmus', MC Die Mikrocomputer-Zeitschrift, pp:63-65
- Jain,A.K., (1989), Fundamentals of Digital Image Processing,Prentice-Hall, New York



- Jain, R., (1976), 'Decision-Making in the Presence of Fuzziness and Uncertainty', Proceedings IEEE Conference Decision Control, New Orleans, pp:698-703
- Ji,L., (1994), 'Fully Automatic Chromosome Segmentation', Cytometry 17:196-208
- Johnston, D., Tang, K. ve Zimmerman S., (1993), 'Band Features as Classification Measures for G-banded Chromosome Analysis', Comput. Biol.Med, vol.23, 115-129
- Klir, G., ve Floger, T., (1988), Fuzzy Sets, Uncertainty and Information, Prentice Hall, NJ
- Lam, L., Lee, S.W. ve Suen C.Y., (1992), 'Thinning Methodologies-A Comprehensive Survey', IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.14, no.9, 869-885
- Mitelman, F., (1995), An International System for Human Cytogenetic Nomenclature (1995), Karger, Switzerland
- Papoulis, A., (1965), Probability, Random Variables, and Stochastic Processes, Mc Graw Hill, New York
- Piper, J. ve Granum,E., (1989), 'On Fully Automatic Feature Measurement for Banded Chromosome Classification', Cytometry 10:242-255
- Pratt W.K., (1978), Digital Image Processing, Wiley, New York
- Richardson, A., (1991), ACT Cytogenetics Laboratory Manual, Raven Press, New York
- Rooney, D.E. ve Czepulkowski, B.H., (1992), Human Cytogenetics, Volume I, Oxford University Press, Oxford
- Ross, T.J., (1995). Fuzzy Logic with Engineering Applications, Mc Graw Hill, Inc., New York
- Stanley, R., Keller, J., Caldwell, C. ve Gader, P., (1996), 'A Centromere Attribute Integration Approach to Centromere Identification', Biomed. Sci. Instrum. vol.32, 23-29
- Schumecker, K.J., (1982), Fuzzy Sets, Natural Language Computations, and Risk Analysis, Computer Science Press, Washington D.C.



- Tan, H., (1989), 'A Comparative Cost Function Approach to Edge Detection', IEEE Transactions on Systems, Man and Cybrenetics, vol.19, pp:1337-1346
- Toussaint, G., (1982), 'Computational Geometric Problems in Pattern Recognition', Pattern Recognition Theory and Applications, Reidel, New York, pp:73-91
- Tou, J. ve Gonzalez R., (1974), Pattern Recognition Principles, Addison Wesley Publishing Company, Massachusettes
- Wang, L., (1997), A Course in Fuzzy Systems and Control, Prentice Hall, NJ
- Weska, J., (1978), 'A Survey of Threshold Selection Techniques', Computer Graphics Image Processing, vol.7, pp:259-265
- Widman, E.,Loparo, K. ve Nielsen N., (1989), Artificial Intelligence, Simulation and Modelling, Wiley, New York
- Whaite, P. ve Ferrie, F., (1990), 'From Uncertainty to Visual Exploration', IEEE Transactions on Pattern Analysis and Machine Intelligence, pp:690-697
- Zhang, T.Y. ve Suen, C.Y., (1984), 'A Fast Parallel Algorithm for Thinning Digital Patterns', Comm.ACM, vol.27, no.3, pp:236-239



ÖZGEÇMİŞ

Soyadı KARSLIGİL
Adı Mine Elif
Doğum Tarihi 25 Nisan 1968
Doğum Yeri İstanbul

Eğitim

1978-1984 Kadıköy Kız Lisesi, İstanbul

1984-1988 Yıldız Üniversitesi

Mühendislik Fakültesi
Bilgisayar Bilimleri ve Mühendisliği Bölümü
(Bilgisayar Bilimleri Mühendisi)

1988-1991 Yıldız Üniversitesi

Fen Bilimleri Enstitüsü, Yüksek Lisans
Bilgisayar Bilimleri Mühendisliği
(Bilgisayar Bilimleri Yüksek Mühendisi)

1991-1998 Yıldız Teknik Üniversitesi

Fen Bilimleri Enstitüsü, Doktora
Bilgisayar Bilimleri Mühendisliği

Yabancı Dil

İngilizce, Almanca



İş Deneyimleri

1988-1998 Araştırma Görevlisi

Yıldız Teknik Üniversitesi

Elektrik Elektronik Fakültesi

Bilgisayar Bilimleri ve Mühendisliği Bölümü

Yıldız - İstanbul

