

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



AKCİĞER KANSERİNİN MAKİNE ÖĞRENME TEKNİKLERİ İLE ANALİZİ

Furkan ESMERAY

YÜKSEK LİSANS TEZİ

Fen Bilimleri Enstitüsü

Biyoteknoloji Ana Bilim Dalı

Danışman: Dr. Öğr. Üyesi Sevcan AYTAÇ KORKMAZ

ARALIK-2019

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

AKCİĞER KANSERİNİN MAKİNE ÖĞRENME TEKNİKLERİ İLE ANALİZİ

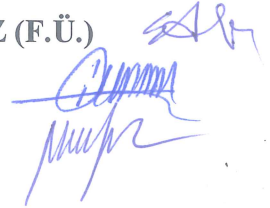
FURKAN ESMERAY

(172111102)

Tezin Verildiği Tarih: 18 Kasım 2019

Tezin Savunulduğu Tarih: 23 Aralık 2019

Tez Danışmanı: Dr. Öğr. Üyesi Sevcan AYTAÇ KORKMAZ (F.Ü.)
Diğer Jüri Üyeleri: Doç. Dr. İsmail TÜRKOĞLU (F.Ü.)
Dr. Öğr. Üyesi Abdulnasır YILDIZ (D.Ü.)



ARALIK-2019

ÖNSÖZ

Bu tez çalışmasında, konu seçiminde, yürütülmesinde ve gerçekleştirilmesinde isteklerimi göz önüne alarak bana her konuda desteğini, anlayışını ve en önemlisi sabrını esirgemeyen hocam Dr. Öğr. Üyesi Sevcan AYTAÇ KORKMAZ' a çok teşekkür ederim.

Ayrıca görüntü alma aşamalarında bana yardımcı olan Prof. Dr. İbrahim Hanifi ÖZERCAN'a yürekten teşekkür ederim.

FURKAN ESMERAY

ELAZIĞ-2019



İÇİNDEKİLER

ÖNSÖZ	II
İÇİNDEKİLER	III
ÖZET	V
ABSTRACT	VI
ŞEKİLLER LİSTESİ	VII
TABLolar LİSTESİ	VIII
KISALTMALAR LİSTESİ	IX
SEMBOLLER LİSTESİ	X
1. GİRİŞ	1
1.1. Çalışmanın Amacı ve Kapsamı	1
1.2. Literatür Araştırması	3
1.3. Çalışmanın Organizasyonu ve Bölümleri	6
2. GENEL BİLGİLER	8
2.1. Kanser	8
2.2. Akciğer Kanseri	9
2.2.1. Akciğer Kanseri Belirtileri	9
2.2.2. Akciğer Kanseri Risk Faktörleri	9
2.2.3. Akciğer Tümörü Tipleri	10
2.2.4. Akciğer Görüntüleme Yöntemleri	13
2.2.5. Akciğer Nodüllerini Gösteren Bilgisayarlı Tomografi Görüntüleri	17
2.2.6. Akciğer Tümörlerine Ait Histopatoloji Görüntüleri	18
3. MATERYAL VE METOT	22
3.1. Özellik Ekleme İçin Kullanılan Yöntemler	22
3.1.1. Maksimum Durağan Ekstrem Bölgeler (MDEB)	22
3.1.2. Hızlandırılmış Sağlam Özellikler (HSÖ)	24
3.1.3. Yönlendirilmiş Gradyanlar Histogramı (YGH)	26
3.2. Boyut İndirgeme İçin Kullanılan Yöntemler	27
3.2.1. Temel Bileşenler Analizi (TBA)	27

3.2.2.	Genelleştirilmiş Diskriminant Analizi (GDA).....	30
3.2.3.	Doğrusal Diskriminant Analizi (DDA).....	30
3.2.4.	Yerel Doğrusal Gömme Analizi (YDGA)	31
3.2.5.	Çok Boyutlu Ölçekleme Analizi (ÇBÖA)	32
3.2.7.	Stokastik Yakınlık Gömme Analizi (SYGA).....	33
3.3.	Sınıflandırma İçin Kullanılan Yöntemler.....	35
3.3.1.	Karar Ağaç (KA)Sınıflandırıcısı	34
3.3.2.	Rastgele Orman (RO) Sınıflandırıcısı.....	36
3.3.3.	Naive Bayes (NB) Sınıflandırıcısı	37
4.	AKCİĞER KANSERİ TEŞHİSİ İÇİN GELİŞTİRİLEN YÖNTEMLERİN ANALİZ BULGULARI.....	39
4.1.	Histopatoloji Görüntüleri İçin HSÖ – Boyut İndirgeme - NB Yöntemi.....	40
4.2.	Histopatoloji Görüntüleri İçin HSÖ – Boyut İndirgeme - RO Yöntemi.....	41
4.3.	Histopatoloji Görüntüleri İçin HSÖ – Boyut İndirgeme - KA Yöntemi.....	42
4.4.	Histopatoloji Görüntüleri İçin MDEB – Boyut İndirgeme - NB Yöntemi	43
4.5.	Histopatoloji Görüntüleri İçin MDEB –Boyut İndirgeme - RO Yöntemi	44
4.6.	Histopatoloji Görüntüleri İçin MDEB –Boyut İndirgeme-KA Yöntemi	45
4.7.	Histopatoloji Görüntüleri İçin YGH – Boyut İndirgeme - NB Yöntemi	46
4.8.	Histopatoloji Görüntüleri İçin YGH – Boyut İndirgeme - RO Yöntemi	47
4.9.	Histopatoloji Görüntüleri İçin YGH – Boyut İndirgeme - KA Yöntemi.....	48
4.10.	Histopatoloji Görüntülerinin Doğruluk Teşhisi Oranlarının Kıyaslanması	49
5.	SONUÇLAR VE ÖNERİLER	51
	KAYNAKLAR	53
	ÖZGEÇMİŞ.....	60

ÖZET

Bu çalışmada, akciğer tomografisi, akciğer hücresi histopatolojisi ve ışık mikroskopundan alınan görüntüler ile daha iyi akciğer kanseri teşhisi hedeflenmiştir. Ayrıca elde edilen görüntüler iyi ve kötü huylu olmak üzere iki tip görüntü kullanılmıştır. Bu iki çeşit görüntülere ait optimum özellikler elde edebilmek için çeşitli boyut azaltma yöntemleri kullanılmıştır. Bu yöntemler Temel Bileşenler Analizi (TBA), Genelleştirilmiş Diskriminant Analizi (GDA), Doğrusal Diskriminant Analizi (DDA), Yerel Doğrusal Gömme (YDG), Klasik Çok Boyutlu Ölçekleme (KÇBÖ), Komşuluk Koruyan Gömme (KKG), Stokastik Yakınlık Gömme (SYG) yöntemleridir. Sınıflandırıcı olarak ise Naive Bayes (NB), Rastgele Orman (RO), Karar Ağaç (KA) yöntemleri kullanılmıştır. Ayrıca bu sınıflandırıcı yöntemlerinin her birine Maksimum Durağan Ekstrem Bölgeler (MDEB), Hızlandırılmış Sağlam Özellikler (HSÖ), Yönlendirilmiş Gradyanlar Histogramı (YGH) özellikleri eklenerek doğruluk değerinin durumu incelenmiştir.

Akciğer kanserinin erken ve doğru teşhisi için çeşitli yöntemler geliştirilmiştir. Bu çalışmada metodlar için toplam 240 adet görüntü kullanılmıştır. Bu görüntüler 60 adet hastadan, her biri için 4 adet kesit alınarak elde edilmiştir. Kullanılan görüntüler ile her bir sınıflandırıcının, her bir boyut azaltma yönteminin ve her bir özelliğin birbirine göre doğruluk oranının kıyaslanması yapılmıştır. Ayrıca en yüksek doğruluk oranına sahip yöntemler bulunmuştur. Kullanılan yöntemlerin bazıları ile %100 doğruluk oranına sahip sonuçlar bulunmuştur. Böylece en yüksek doğruluk oranını sağlayan yöntemler sayesinde iyi huylu ve kötü huylu tümörlerin ayrımı sağlanmıştır. Dünyada en fazla ölüm oranına sahip kanser türlerinden biri olan akciğer kanserinin erken teşhisi için doğruluk oranının yükseltilmesi de amaçlanmıştır. Böylece bilgisayar destekli tanı sistemlerine yüklenen akciğer görüntüleri bu sistemde çeşitli aşamalardan geçirilerek doğru bir sonucun elde edilmesi sağlanmıştır.

Anahtar Kelimeler: Akciğer Kanseri, Makine Öğrenme, Sınıflandırma, Görüntü İşleme, Histopatoloji, Akciğer Tomografisi, Bilgisayar Destekli Tanı, Bilgisayarlı Tomografi.

ABSTRACT

Analysis Of Lung Cancer With Machine Learning Techniques

In this study, we aimed to better diagnosis of lung cancer with lung tomography, lung cell histopathology and light microscopy images. In addition, two types of images, good and malignant, were used. Various size reduction methods have been used to obtain optimum properties of these two types of images. These methods include Principal Component Analysis (PCA), Generalized Discriminant Analysis (GDA), Linear Discriminant Analysis (LDA), Local Linear Embedding (LLE), Classical Multidimensional Scaling (MDS), Neighborhood Preserving Embedding (NPE), Stochastic Proximity Embedding (SPE)) methods. Naive Bayes (NB), Random Forest (RF), Decision Tree (DT) methods were used as the classifier. In addition, Maximum Stationary Extreme Regions (MSER), Speeded Up Robust Features (SURF), Histogram of Oriented Gradients (HOG) were added to each of these classifier methods and the status of the accuracy value was examined.

Various methods have been developed for the early and accurate diagnosis of lung cancer. In this study, a total of 240 images were used for the methods. These images were obtained from 60 patients by taking 4 sections for each. The accuracy of each classifier, each size reduction method and each feature was compared with the images used. In addition, the methods with the highest accuracy were found. Some of the methods used were found to have 100% accuracy. Thus, the distinction between benign and malignant tumors has been achieved thanks to the methods that provide the highest accuracy. It is also aimed to increase the accuracy rate for the early diagnosis of lung cancer, which is one of the cancer types with the highest mortality rate in the world. Thus, lung images loaded into computer assisted diagnostic systems were passed through various stages in this system to obtain an accurate result.

Key Words: Lung Cancer, Machine Learning, Classification, Image Processing, Histopathology, Lung Tomography, Computer Assisted Diagnosis, Computed Tomography.

ŞEKİLLER LİSTESİ

	Sayfa No
Şekil 2.1. Akciğer radyografisi örneği	14
Şekil 2.2. Bilgisayarlı tomografi cihazı ve üniteleri.....	15
Şekil 2.3. Hounsfield cetveli	16
Şekil 2.4. Bilgisayarlı tomografi sonucunda elde edilen görüntüler	16
Şekil 2.5. Örnek akciğer nodülleri.....	17
Şekil 2.6. Fırat Üniversitesi Tıp Fakültesi Hastanesi Patoloji Laboratuvarından alınmış boyalı olan preparat örneği	19
Şekil 2.7. Fırat Üniversitesi Tıp Fakültesi Hastanesi Patoloji Laboratuvarında kullanılan ışık mikroskobu.....	20
Şekil 2.8. Fırat Üniversitesi Tıp Fakültesi Hastanesi Patoloji Laboratuvarından elde edilen görüntülerin örnekleri	21
Şekil 4.1. Akciğer kanseri teşhisi için geliştirilen yöntemlerin akış şeması	39
Şekil 4.2. Akciğere ait tümörün histopatoloji görüntü örnekleri.....	40

TABLÖLAR LİSTESİ

Sayfa No

Tablo 2.1. Küçük hücreli olmayan akciğer kanserinin TNM sınıflandırılması	12
Tablo 2.2. Küçük hücreli olmayan akciğer kanserinin evreleri	13
Tablo 4.1. HSÖ-Boyut İndirgeme-NB yöntemi doğruluk teşhisi sonuçları	41
Tablo 4.2. HSÖ-Boyut İndirgeme-RO yöntemi doğruluk teşhisi sonuçları	42
Tablo 4.3. HSÖ-Boyut İndirgeme-KA yöntemi doğruluk teşhisi sonuçları	43
Tablo 4.4. MDEB-Boyut İndirgeme-NB yöntemi doğruluk teşhisi sonuçları.....	44
Tablo 4.5. MDEB-Boyut İndirgeme-RO yöntemi doğruluk teşhisi sonuçları.....	45
Tablo 4.6. MDEB-Boyut İndirgeme-KA yöntemi doğruluk teşhisi sonuçları.....	46
Tablo 4.7. YGH-Boyut İndirgeme-NB yöntemi doğruluk teşhisi sonuçları.....	47
Tablo 4.8. YGH-Boyut İndirgeme-RO yöntemi doğruluk teşhisi sonuçları.....	48
Tablo 4.9. YGH-Boyut İndirgeme-KA yöntemi doğruluk teşhisi sonuçları	49
Tablo 4.10. Histopatoloji görüntülerinin doğruluk teşhisi oranlarının kıyaslanması için yöntemlerin tek başlarına doğruluk teşhis oranları	50

KISALTMALAR LİSTESİ

BT	: Bilgisayarlı Tomografi
YSA	: Yapay Sinir Ağı
CAD	: Bilgisayar Destekli Algılama
FCM	: Fuzzy C Means
MRG	: Manyetik Rezonans Görüntüleme
PET	: Pozitron Emisyon Tomografi
FDG	: Florodeoksiglikoz
MDEB	: Maksimum Durağan Ekstrem Bölgeler
KDF	: Kümülatif Dağılım Fonksiyonu
HSÖ	: Hızlandırılmış Sağlam Özellikler
YGH	: Yönlendirilmiş Gradyanlar Histogramı
TBA	: Temel Bileşenler Analizi
GDA	: Genelleştirilmiş Diskriminant Analizi
DDA	: Doğrusal Diskriminant Analizi
YDGA	: Yerel Doğrusal Gömme Analizi
ÇBÖA	: Çok Boyutlu Ölçekleme Analizi
KKGA	: Komşuluk Koruyan Gömme Analizi
SYGA	: Stokastik Yakınlık Gömme Analizi
KA	: Karar Ağaç
RO	: Rastgele Orman
NB	: Naive Bayes

SEMBOLLER LİSTESİ

\pm	: Hem artan hem azalan eşik değeri
Ω	: Etiket kümesi
$ \cdot $: Bölge alanı
I	: Görüntü konumu
$I_{\Sigma}(a)$: Görüntü tamamlayıcı özellik
σ	: Görüntü Ölçeği
d_a	: Yatay yönde Haar dalgacığı
d_b	: Düşey yönde Haar dalgacığı
G	: Gradyan
Θ	: Gradyan yönelimi
S_y	: Yatay Sobel filtresi
S_d	: Düşey Sobel filtresi
\check{A}	: Ortalaması çıkartılmış veri matrisi
C	: Kovaryans matrisi
W	: İz düşüm matrisi
S_M	: Saçılım matrisi
S_F	: Sınıf içi saçılım matrisi
ε	: Komşuluk değeri
δ	: Öğrenme oranı azaltma katsayısı
$h(x, T_K)$: Sınıflandırıcı
T	: Eğitim veri seti
T_K	: Yer değiştirmeli eğitim veri seti
$f_b(F)$: Bayesian sınıflandırıcı
M_i	: Bir sınıfa ait ihtimal
λ	: Öğrenme oranı
d_{ij}	: Öklid mesafesi
r_c	: Kesme mesafesi
\widehat{d}_{ij}	: Bireyler arası veri uzaklığı
\in	: Eleman
\forall	: Her

- I_x : Yatay düzlemde imge kenarı
 I_y : Düşey düzlemde imge kenarı
 w : Hessian determinanı sabiti
 D_{aa} : a yönünde yönlendirilmiş ağırlıklandırılmış kutu filtre yaklaşımı
 D_{ab} : ab yönünde yönlendirilmiş ağırlıklandırılmış kutu filtre yaklaşımı
 D_{bb} : b yönünde yönlendirilmiş ağırlıklandırılmış kutu filtre yaklaşımı
 H : Hessian matrisi
 P : Dönüşüm matrisi
 A_i^* : Uç bölge
 $d_{bhr}(b)$: Hesaplanan eşik değeri
 μ : Örnek küme ortalaması
 E_b : Etiket görüntüsü
 E_{b+1} : Yeni etiket görüntüsü

1.GİRİŞ

Her organ gibi akciğere ait olan hastalıklarda bulunmaktadır. Bu hastalıkların başında ise akciğer kanseri olarak isimlendirilen bir kanser türü vardır. Ülkemizde olduğu gibi diğer ülkelerde de kanser ölümlerinin başlıca nedenlerinden biri olarak kabul edilen akciğer kanseri gelmektedir [1, 2]. Önceleri akciğer kanseri ender görülen bir hastalık iken şu an tüm dünyada en fazla ölüme yol açan kanser türüdür. Akciğer kanseri bütün dünyada tüm kanser türlerinin %12,8'ini ve kanserden dolayı ölümlerin %17,8'ini oluşturmaktadır. Ayrıca tüm dünyada vaka sayısı her yıl artmaktadır [3]. 2,09 milyon kişi akciğer kanseri vakasıyla karşımıza çıkmaktadır. Dünya Sağlık Örgütü verilerine göre dünyada, 2018 yılında tahmini 9,6 milyon kişi kanser sebebiyle ölmüştür. Ölümlü kanser vakalarının arasında ise akciğer kanseri, 1,76 milyon ölümlü vaka sayısına sahiptir. 1,76 milyon sayısı, ölümlü kanser vakaları içerisinde akciğer kanserinin birinci sırada yer almasına sebep olmuştur [4]. Türkiye'de kanser sıklığı diğer dünya ülkeleriyle benzerdir. 2009 yılında Türkiye'de yaşa bağlı olarak kanser hızı erkeklerde her 100.000 kişide 269,7, kadınlarda ise her 100.000 kişide 173,3'tür. Akciğer kanseri ülkemizde de en sık erkeklerde görülen kanser türüdür. Akciğer kanseri yıllık insidansı erkeklerde her 100.000 kişide 66, kadınlarda ise her 100.000 kişide 8,1 olarak belirlenmiştir. Bu oran gelecekte akciğer kanserinin önemli bir sorun olacağını göstermektedir [5].

Akciğer dokularındaki hücreler kontrolsüz bir şekilde çoğalarak akciğer kanserine sebep olmaktadır. Akciğer dokularında kontrolsüz çoğalan bu hücrelerin başka bir yakın olan hücreyi etkilemesi ve başka dokulara temas etmesi halinde hastalıklı hücre sayısında artış meydana gelir. Akciğer kanserinin belirtileri arasında öksürük, kilo kaybı ve nefes darlığı gibi şikayetler bulunmaktadır. Ancak bu belirtiler diğer hastalıkların belirtileriyle de ortak olduğu için hastalığın tanısı zor olabilir ve uzun zaman alabilir. Akciğer kanserine ait bir tanı yapabilmek için bilgisayarlı tomografiden faydalanılmaktadır. Kesin tanıyı koyabilmek için ise biyopsi kullanılmaktadır. Bu işlem bilgisayarlı tomografi yardımıyla tespit edilen nodüle yapılan biyopsi ile yapılmaktadır. Bilgisayarlı tomografi ile akciğerde kansere sebep olan tümörlerin belirlenmesi oldukça önemlidir [6]. Bilgisayarlı tomografi, gelişmiş görüntüleme sistemleri ile hastalığın erken teşhis edilebilmesi için kullanılabilir. Özellikle bilgisayarlı tomografi sayesinde küçük boyutlu tümörlerin akciğer görüntülerinin izlenmesi erken teşhiste önemli bir rol oynamaktadır. Fakat bunun yanında kullanılan sistemlerden dolayı bir karmaşıklıkta ortaya çıkmıştır. Her bir hasta için

çok fazla sayıda kesite sahip olan bilgisayarlı tomografi görüntülerinin radyologlar tarafından incelenmesi uzun zaman almakta ve yorucu bir sürece işaret etmektedir. Öte yandan bir radyologun yaptığı tanı ile bir başka radyologun yaptığı tanı arasında fark bulunmaktadır. Çünkü tanı bulacak kişiye göre hem bilgi birikimi hem de deneyimi arasında farklar vardır. Bu sebepten dolayı insanlar arasındaki farklı tanıları ortadan kaldırmak ve insan hatasını azaltmak için bilgisayar destekli sistemlerin kullanımına ihtiyaç duyulmuştur. Böylece bilgisayar destekli tespit/teşhis sistemleri, hastalığın tanı/teşhisine ait karar verme sürecinde önemli bir konuma sahip olmuştur [7]. Teknolojinin gelişmesi ile beraber kanser tanılarında çeşitli sistemlerin kullanılarak bu tespiti yapacak kişiye yardımcı olacak yeni sistemlere gereksinim duyulmaktadır. Akciğer kanserinin tanısında da yeni sistemlerin kullanılmasına ihtiyaç duyulmaktadır.

Tarama teknolojilerindeki ilerlemeler sayesinde akciğer tümörü teşhis oranı artmaktadır. Akciğere ait tümörün tespiti sonrasında bilgisayar destekli tanı sistemleri kullanılarak radyologlara bu tanı sisteminin yardımcı olması sağlanmıştır. Bilgisayar destekli tanı sistemlerinde iki önemli nokta bulunmaktadır. Bunlar akciğerde bulunan tümörlerin belirlenmesi ve sınıflandırılmasıdır. Böylece hem daha hızlı hem de daha doğru tanı konulması sağlanmıştır [8].

Tomografi yöntemleriyle belirlenen tümörün boyut ve çeşidine göre iyi huylu veya kötü huylu olduğunu tespit etme noktasında tomografi yöntemlerinde bazı yetersiz durumlar söz konusudur. Bu durumlarda histopatoloji gibi bağlantılara da ihtiyaç duyulmaktadır. Çünkü histopatoloji görüntüleri piksel olarak daha geniştir. Ayrıca histopatoloji görüntüleri farklı renklerde olabilir. Ancak radyoloji görüntüleri sadece gri renkle sınırlıdır. Bunun yanında histopatoloji görüntülerinde daha düşük büyütme oranı alınır. Bu sayede doku sınırlarının belirlenmesi daha kolay olur [9].

Bilgisayarlı tomografi ile tespit edilen tümörden biyopsi yöntemiyle bir parça alınmaktadır. Alınan parça üzerindeki dokuya ait hücreler patoloji doktoru tarafından incelenir. Histopatoloji görüntüleri ışık mikroskopuyla incelenebilmektedir. Işık mikroskopu ile incelenen dokuya ait kesit görüntüleri bir kamera tarafından bilgisayar ortamında görüntülenebilir. Histopatoloji görüntüleri sayesinde akciğer kanserinin erken teşhisi ve tanısı yapılabilmektedir.

1.1. Çalışmanın Amacı ve Kapsamı

Akciğer kanseri bilinen ölümcül kanser türleri arasında en ilk sıralarda yer almaktadır. Fakat erken teşhis ölüm oranını azaltır. Daha erken kanserler tespit edilebilirse daha etkin bir tedavi sağlanabilir. Ancak erken teşhis, aynı zamanda iyi ve kötü huylu tümörleri ayırt edebilmek için doğru ve güvenilir bir tanı gerektirir. Akciğerdeki kitlelerin iyi huylu, kötü huylu ayırımını yapabilmek için duyarlı yöntemlere gereksinim vardır. Bu yöntemler akciğer kanserinin tespiti ve erken tanısı için oldukça önemlidir. Akciğer kanserinin tespiti için akciğer kanseri şüphesi olan bir doku biyopsi yöntemiyle alınır. Daha sonra bu doku içerisindeki hücrelere ışık mikroskobu ile bakılarak tanı koyulmaktadır. Işık mikroskobu ile bakılan dokunun görüntüsü bir kamera vasıtasıyla fotoğrafı çekilmektedir. Bu sayede histopatoloji görüntüleri elde edilmektedir. Histopatoloji görüntüleri elde edildikten sonra artık bu görüntülere bakılarak kanser teşhisi rahatlıkla yapılabilir.

Bu nedenle akciğer kanserinin erken teşhisi ile akciğer kanserinin sebep olduğu ölüm oranlarını azaltabilmek için akciğer kanseri tanı sistemleri yardımıyla doktorlara bir ön fikir verebilme amaçlanmıştır. Akciğer kanseri tanı sistemini oluşturabilmek için yüksek tespit doğruluk oranı ve en kısa süre harcaması hedeflenmiştir. Bu sistemin oluşturulması için görüntü işleme teknikleri ve yapay zeka yöntemleri kullanılmıştır.

Bu çalışmanın amacı görüntü işleme yöntemlerini ve yapay zeka yöntemlerini kullanarak akciğer kanserinin tespit ve tanısı bir bilgisayar destekli tanı sistemi oluşturmaktır. Bu çalışmanın sonucunda akciğer kanseri görüntüleri iki sınıfa ayrılacaktır. Bu iki sınıf iyi huylu ve kötü huylu tümör görüntüsüdür. Akciğer kanseri görüntülerini iki sınıfa ayırabilmek için kullanılan her bir yöntemin hangisi veya hangilerinin en yüksek doğruluk oranına sahip olduğunu göstermede hedeflenmiştir. Ayrıca kullanılan yöntemlerin kendi aralarında kıyaslanmaları da amaçlanmaktadır. Böylece hangi yöntemin en uygun yöntem olduğu tespit edilmiş olacaktır. Gerekli çalışma yapıldıktan sonra akciğer kanseri bilgisayar destekli tanı sistemi oluşturabilmek için başka yöntem önerilerinin yapılması da amaçlanmıştır.

1.2. Literatür Araştırması

Giger ve arkadaşları (1994), çeşitli bilgisayar görüntüleme tekniklerini ve polmoner nodüllerin morfolojik özellikleri hakkında bilgileri kullanarak, akciğer sınırları içindeki özelliklerin gri seviye eşiklemine tabii tutmuşlardır. Bu çalışma sonucunda da BT

taramalarında pulmoner nodüllerin saptanması için otomatik bir sistem geliştirilmesinin radyologlara yardımcı olacağı sonucuna varmışlardır [10].

Xu ve arkadaşları (1997), ileri yönlü yapay sinir ağları kullanımıyla nodüllerin tespitini gerçekleştirmişlerdir. Bu çalışma kullanılan yöntemden dolayı en ilk çalışmalardan birisi olma özelliğine sahiptir [11].

Jianzhong ve arkadaşları (2002), radyologların nodül tespitinde verdikleri kararların doğruluğunu arttıran bir çalışma gerçekleştirmişlerdir. Çalışma sonucunda radyologların sadece kendi yorumu ile tespit edebildiği nodül sayısından daha fazla sayıda nodül tespit edilebildiğini savunmuşlardır [12].

Edis ve arkadaşları (2007), akciğer kanserinin Türkiye'ye maliyetinin hesaplanmış olan bir çalışmadan yola çıkmışlardır. Akciğer kanserine sebep olan nodüllerin erken ve doğru tespit edilmesinin yanında, bilgisayar destekli teşhis (CAD) sistemlerinin kullanılmasıyla da akciğer kanseri hastalığının tedavisinin maliyetinde azalacağını savunmuşlardır [13].

Dehmeshki ve arkadaşları (2007), nodüllerin tespit edilebilmesi için şekil tabanlı bir algoritma şablon eşleştirme yöntemi kullanmışlardır. Bu çalışmanın sonucunda %90'lık bir tespit oranına ulaşmışlardır [14].

Hirose ve arkadaşları (2008), yaptıkları çalışmada çoklu dedektör ile alınan akciğer görüntülerinden nodül tespit edilebilmesi için bilgisayar destekli teşhis (CAD) sisteminin kullanılmasıyla nodül belirleme hassasiyetinin arttığını gözlemlemişlerdir [15].

Aubry ve arkadaşları (2006), yaptıkları çalışmada radyologların nodül tespit performansları üzerinde bilgisayar destekli teşhis (CAD) sisteminin durumunu incelemişlerdir. Radyologlar tarafından kullanılan sistemde nodül belirleme duyarlılığının arttığı tespit edilmiştir [16].

Lee ve arkadaşları (2009), yaptıkları çalışmada bayesian sınıflandırıcı, saf bayesian sınıflandırıcısı, destek vektör makineleri sınıflandırıcısı yöntemlerini kullanarak kıyaslama yapmışlardır. En iyi sonucu destek vektör makineleri sınıflandırıcısının verdiğini tespit etmişlerdir [17].

Penedo ve arkadaşları (1998),iki seviyeli bir yapay sinir ağı (YSA) yöntemi ile sağlanan ve yerel görüntü eğriliklerini de kullanarak bilgisayar destekli bir tanı sistemi geliştirmişlerdir. Gerçek olarak 90 sayıda olmak üzere, 288 simüle edilmiş nodül YSA ile sınıflandırılmıştır ve duyarlılık % 89- % 96 oranında, görüntü başına yanlış pozitif sayısı ise 5-7 olarak elde edilmiştir [18].

Retico ve arkadaşları (2008), düşük dozlu ve ince dilimli BT taramalarında küçük pulmoner nodüllerin tanımlanması için bilgisayar destekli bir algılama sistemi geliştirilmişlerdir. Bilgisayar destekli algılama sistemi ile BT taramaları veritabanında toplanan küçük iç ve alt plevral pulmoner nodüllere duyarlı olması için eğitilmiştir. Sistem performansı, 75 iç ve 27 alt plevral nodül içeren 39 BT veri setinde değerlendirilmiştir. Bu veri setinde elde edilen FROC eğrisi, hasta başına kabul edilebilir düzeyde yanlış pozitif bulgular düzeyinde 10-13 sayıda ve akciğer nodüllerine% 80-85 oranında yüksek hassasiyet değerleri tespit edilmiştir [19].

Lee ve arkadaşları (2010), otomatik bir akciğer nodülü tespit sisteminin BT akciğer görüntülerinde akciğer anormalliklerinin tespit edilmesine yardımcı olabildiğini ve bu tespitin şablon bazlı, segmentasyon bazlı ve sınıflandırma bazlı yöntemler kullanılarak gerçekleştirilebileceğini savunmuşlardır. Otomatik akciğer nodülü tespitinin performansını artırmak için, kümeleme yöntemiyle desteklenen bir sınıflandırma yöntemini önermişlerdir. Bu yöntem ile rastgele orman algoritmasından yararlanmışlardır. Nodül yerleşimlerinin uzman radyologlar tarafından işaretlendiği 5721 görüntü dahil 32 hastanın akciğer taraması kullanılarak gerçekleştirilmiştir. Genel olarak, önerilen sistem için% 98,33 olan en yüksek hassasiyet ve% 97,11 özgüllük kaydedilmiştir [20].

Coppini ve arkadaşları (2003), yaptıkları çalışmada biyolojik olarak esinlendirilmiş filtreleri belirgin görüntü özelliklerini geliştirmek için kullanmışlardır. Nodüllerin şekli ve arka plan yapısı hakkındaki bilginin etkin bir şekilde kullanılması için yapay sinir ağlarını kullanmışlardır. Yapılan çalışmada 247 görüntü sistemi kurmak ve test etmek için kullanılmıştır. Çalışma sonucunda duyarlılık% 60 ile% 75 arasında değişen görüntü başına yanlış alarm sayısının 4-10 arasında, doğruluk oranının % 95,7 - 98,0 arasında olduğunu gözlemlemişlerdir [21].

Kakar ve arkadaşları (2009), yaptıkları çalışmada BT ile akciğerler için tam otomatik doku tabanlı bir segmentasyon ve tanıma sistemini sunmuşlardır. Gabor'un görüntüleri filtreleyerek segmentasyon bölümü ile ilgili doku özelliklerini çıkarmışlar ve Fuzzy C Means (FCM) kümelenmesini kullanarak ve genetik algoritma kullanılarak yapmışlardır. Tanıma aşaması için korteks benzeri bir yapı kullanarak şekil tabanlı özelliklerin yanı sıra istatistiksel özellikleri çıkarmışlardır. Sonuç olarak lezyon tespit oranı % 89.04 olduğunu doğrulamışlardır. Ayrıca sınıflandırıcının ortalama duyarlılığı oranı ise % 89.48 olarak bulunmuştur [22].

Farag ve arkadaşları (2004), benzer nodüller arasında, değişebilen 2 ve 3 boyutlu şablonlar kullanarak, nodül belirleme ilgili bir algoritma öne sürmüşlerdir. Öne sürülen algoritma belirleme algoritmasıdır ve bu algoritma bir çeşit algoritma olan genetik algoritmayı kullanıp çapraz-korelasyon şablon eşleştirme ve Bayesian sınıflandırma sonrası işlemlerini harmanlaştırmıştır. Böylece belirleme algoritması ile 130 nodülün 107'sinin doğru bir şekilde sınıflandırılması sağlanmıştır [23].

Bae ve arkadaşları (2005), üç boyutlu volumetrik verilerden yararlanarak otomatik bir pulmoner nodül tespit programı geliştirmişlerdir. Pulmoner olan 20 hastanın bilgisayarlı tomografi görüntülerini kullanarak nodül tespitlerini yapmışlardır. 164 adet nodülü tespit ederek, bir bilgisayar destekli algılama programını değerlendirmek için tespit edilen nodülleri kullanmışlardır. Yaptıkları çalışmaların sonucunda nodül tespitine yönelik genel duyarlılık % 95.1 orandır. Bu oranda 164 nodülün 156 tanesinin doğru hesaplandığını göstermektedir [24].

Armato ve arkadaşları (2002), yaptıkları çalışmada bilgisayarlı tomografi taramaları ile akciğer nodüllerinin saptanması için tam otomatik bilgisayarlı yöntemin uygunluğunu değerlendirmişlerdir. Bir kanser tanıma programından 50 akciğer nodülü ile 38 düşük dozlu bilgisayarlı tomografi taramasının bir veritabanı oluşturmuşlardır. Gri seviye eşikleme tekniklerini kullanan bir bilgisayar algılama yöntemini uygulamışlardır. Nodül adayları setine kural tabanlı bir şema ve kademeli otomatik sınıflandırıcı uygulayarak gerçek nodülleri normal anatomi alanlarından ayırmışlardır. Kullandıkları yöntem ile ortalama 1,0 hatalı pozitif saptama ile % 80'lik bir orana yani 50 nodülün 40 tanesinin doğru bulunma hassasiyetine ulaşmışlardır [25].

Hardie ve arkadaşları (2008), Akciğer grafilerinde pulmoner nodüllerin tespiti için bilgisayar destekli yeni bir algılama sistemi kullanmışlardır. Bunun için halka açık bir veritabanı kullanan bir performans analizi sunmuşlardır. Toplam 181 akciğer nodülü ile 167 göğüs grafisi içeren bağımsız bir veri setini kullanmışlardır. Halka açık test seti ise standart dijital görüntü veri tabanı tarafından oluşturulmuştur. Her aday için 114 adet özellik kümesi hesaplanmıştır. Sonuç olarak sistemin testteki nodüllerin %78,1'ini ve görüntü başına ortalama 4.0 yanlış pozitif tespit ettiğinin göstermişlerdir [26].

1.3. Çalışmanın Organizasyonu ve Bölümleri

Bu çalışma aşağıdaki bölümlerden oluşmaktadır:

Birinci bölümde, çalışmanın genel olarak bir görüş sahibi olabilmek için temel bilgiler verilmiştir. Ayrıca akciğer kanserinin dünyada ve ülkemizde ne kadar sıklıkta yaşandığını öğrenebilmek için akciğer kanseri yaygınlık oranları incelenmiştir. Akciğer kanserinde tanının koyulmasında kanserinin teşhisinin önemi anlatılmıştır. Literatür çalışmalarının değerlendirilmesi sağlanmıştır.

İkinci bölümde, kanser, akciğer kanseri, akciğer kanseri belirtileri, akciğer kanseri risk faktörleri, akciğer tümörü çeşitleri, akciğer görüntüleme yöntemleri, akciğer tümörlerine ait histopatoloji görüntüleri, akciğer tümörü histopatoloji görüntülerinin elde edilmesi, ışık mikroskobundan histopatoloji görüntülerinin elde edilmesi konuları anlatılmıştır.

Üçüncü bölümde, tezde kullanılan boyut indirgeme yöntemleri, özellik ekleme için kullanılan yöntemler, sınıflandırma yöntemleri anlatılmıştır.

Dördüncü bölümde, akciğer kanserinin makine öğrenme teknikleri ile analizi için geliştirilen yöntemlere yer verilmiştir. Uygulanan yöntemlerin sonucunda elde edilen teşhis doğruluk oranları açıklanmıştır. Elde edilen başarı oranları kıyaslanmıştır.

Beşinci bölümde elde edilen sonuçlar ışığında yapılabilecek katkılar anlatılmıştır. İleride yapılabilecek benzer çalışmalar için kullanılacak yöntemler açıklanmıştır. Ayrıca bu çalışmalar sonucunda yapılabilecek uygulamalardan bahsedilmiştir.

2. GENEL BİLGİLER

2.1. Kanser

İnsan vücudunda bir çok hücreler mevcuttur. Bu hücrelerin hepsinin belirli bir görevi bulunmaktadır. Bu hücreler çeşitli görevleri yerine getirebilmek için bölünmektedirler. Yıprandıkları zaman ve ya hasar gördüklerinde ölmektedirler. Bu ölen hücrelerin yerinide yeni hücreler almaktadır. Bu tür durumlarda hücreler bölünmeye başlarlar. Hücreler bölündükçe de büyümektedirler. Hücrelerin kontrolsüz bir şekilde sürekli büyümesi ile kanser başlamaktadır. Bu kanser oluşumu vücudun bir yerinde başladığı zaman, başladığı kısımda sorunlara sebep olur [27].

Bir çok kanser türü bulunmaktadır. Kanser akciğerde, memede, kanda ve kolonda başlayabilmektedir. Kanserler bir açıdan bakıldığında aynıdır. Ama büyüme ve yayılma durumlarına göre farklıdır. Bazı kanserler hızlı büyür ve yayılırlar. Hatta vücudun farklı yerlerinede yayılabilirler. Örneğin; akciğerdeki kanser hücreleri kemiğe geçip, orada büyüyebilirler. Ancak kemiğe geçen akciğer kanseri hücresi değişime uğramaz. Çünkü o hala akciğer kanser hücresidir. Kanser türleri arasında farklılık olarak da bazı kanser türleri en iyi ameliyatla tedavi edilebilirken bazı kanserler kemoterapi ile tedavi edilmektedir [27].

Çoğu kanser türünde tümör olarak adlandırılan bir yumru oluşmaktadır. Ancak vücutta oluşan her yumru kanser değildir. Doktorlar tümörden biyopsi ile bir parça çıkartırlar. Çıkartılan parçalar çeşitli yöntemlerle incelenir. Kanser şüphesiyle alınan parça iyi huylu ise benign, kötü huylu ise malign olarak adlandırılır. Ancak bazı kanser türlerinde tümör oluşmamaktadır. Buna örnek olarak lösemi (Kan kanseri) verilebilir. Lösemi, vücudun kan hücrelerinde büyürler. Böylece kan hücrelerinde tümör oluşmamaktadır [27].

Kanser tedavileri için kullanılan yöntemler bulunmaktadır. Bu yöntemlerden en yaygın olanları cerrahi yöntem, kemoterapi, radyasyon yöntemidir. Cerrahi yöntemde kanserin etkilediği vücut bölümünün bir kısmı veya tamamı cerrahi yöntemle alınır. Bütün kanser türlerinde bu yöntem kullanılmamaktadır. Hastalığın olduğu kısma göre değişmektedir. Kemoterapi yönteminde hastalığı oluşturan kanser hücresini öldürmek veya yavaşlatmak için çeşitli ilaçlar kullanılmaktadır. Radyasyon yönteminde ise hastalığı oluşturan kanser hücresi radyasyona maruz bırakılarak, bu hücrenin ölmesi veya büyümesinin engellenmesi sağlanmaktadır [27].

2.2.Akciğer Kanseri

Akciğer kanseri, hücrelerin kontrol dışı çoğalmasıyla oluşan kötü huylu tümörün akciğerde meydana getirdiği kanser türüdür. Bu oluşan kötü huylu tümör akciğere hasar vermesinin yanında daha ileriki evrelerinde başka organlara da hasara yol açıyor. Hatta bu organlar akciğere daha uzak mesafede yer alan organlar bile olabilmektedir. Ayrıca akciğerde oluşan tümörler iyi yada kötü huylu olabilmektedir.

2.2.1. Akciğer Kanseri Belirtileri

Akciğer kanseri yayılmaya başlayıncaya kadar belirti göstermezken, erken semptomlar şeklinde belirti gösterebilir. Bu semptomlar; ses kısıklığı, iştah kaybı, açıklanamayan kilo kaybı, nefes darlığı, geçmeyen öksürük, kanlı tükürük, öksürme ile gelen göğüs ağrısı, hırıltılı nefes şeklinde semptomları vardır [28].

Akciğer kanserinin vücudun başka bölümlerine yayılmasıyla oluşan belirtiler; cildin ve gözlerin sararması, kemik ağrısı, baş ağrısı, baş dönmesi, denge sorunları, boyundaki ve köprücük kemiği üzerindeki lenf bezlerinin şişmesi şeklindedir [28].

2.2.2. Akciğer Kanseri Risk Faktörleri

Bir insanın kanser gibi bir hastalığa yakalanma ihtimalini arttıran her şeye risk faktörü denir. Akciğer kanserinde de risk faktörü vardır. Bunlar sigarayı bırakma gibi değiştirilebilen, aile öyküsü gibi değiştirilemeyen olmak üzere ikiye ayrılan risk faktörleridir [29].

Değiştirilebilen risk faktörleri;

1) Tütün Dumanı: Risk faktörü olarak akciğer kanseri oluşumunda birincidir. Akciğer kanseri sebebiyle ölenlerin %80'i bu risk faktöründen dolayı ölmektedir.

2) Pasif İçicilik: Sigara içilmezse bile başkalarının içtiği sigara veya tütün ürünün dumanının solunmasıyla oluşan bir risk faktörüdür.

3) Radona Maruz Kalma: Radon; toprakta ve kayalarda uranyum elementinin parçalanmasından kaynaklanan oluşan gaz türüdür. Bu gaz türüne maruz kaldığı zaman akciğerler radyasyona maruz kalır. Bu da akciğer kanseri için bir risk faktörü oluşturmaktadır [29].

4) Asbeste Maruz Kalma

5) İçme Suyunda Arsenik Olması: İçme suyunda yüksek miktarda arsenik bulunan bölgelerde yapılan araştırmalar; akciğer kanserinin bu sebebe bağlı olarak yüksek oranda

olduđu tespit edilmiştir. Bu sebeple içme suyunda arsenik oranının artması da bir risk faktörüdür [29].

Değiştirilemeyen risk faktörleri;

1) Akciğerlerin Başka Bir Sebep Ten Dolayı Radyasyona Maruz Kalması: Bazı hastalıkların tedavisi ve tanısı için radyasyon kullanılmaktadır. Bu kullanılan radyasyon akciğere etki ettiğinde ve sigara içen bir hastaysa akciğer kanserine yakalanma riski yüksektir. Bu şekilde bir risk faktörü değiştirilemeyen bir risk faktörüdür.

2) Hava Kirliliği: Yaşanan şehirlerde havanın temiz olmaması ve kirli olması da risk faktörü oluşturmaktadır.

3) Akciğer Kanserinin Kişisel ve Ailesel Sebepleri: Kardeşler, ebeveynler, yakın akrabalar akciğer kanseri tanısı aldıysa, kişide de akciğer kanseri olma ihtimali bulunmaktadır [29].

2.2.3 Akciğer Tümörü Tipleri

2.2.3.1. Akciğer İyi Huylu Tümörler (Benign)

İyi huylu tümörler, kanser tümörleri arasında değillerdir. Yavaş ve sınırlı olarak büyümektedirler. Bazen kendi kendilerine tedavi gerektirmeden de küçülebilirler. Hatta benign tümörler yok olabilir veya büyümeleri de durabilir [30].

2.2.3.2. Akciğer Kötü Huylu Tümörler (Malign)

Akciğer kötü huylu tümörleri akciğer kanserini oluşturmaktadır. Akciğerde meydana gelen kanser iki tiptir. Bunlar küçük hücreli olmayan akciğer kanseri ile küçük hücreli akciğer kanseridir [31].

2.2.3.2.1. Küçük Hücreli Olmayan Akciğer Kanseri

Küçük hücreli olmayan akciğer kanseri, akciğer kanserinin %80-85'ini oluşturmaktadır. Küçük hücreli olmayan akciğer kanserinde alt grupları vardır. Bu alt gruplar; adenokarsinom, büyük hücreli karsinom, skuamöz hücreli karsinomdur. Bu küçük hücreli olmayan akciğer kanserinin tedavileride yaklaşık olarak benzerdir [31].

1) **Adenokarsinom:** Adenokarsinomların başlaması normalde mukus gibi maddeler salgılayan hücrelerde olur. Adenokarsinom tipi akciğer kanseri normalde eskiden sigara içmiş veya halen sigara içen kişilerde görülür. Ancak sigara içmeyenlerde de en sık

görülen bir türdür. Adenokarsinom akciğerin dış kısımlarında genelde görünmektedir. Yayılım yapmadan önce görülmesi daha kolaydır [31].

2) Büyük Hücreli Karsinom: Akciğerin herhangi bir kısmında görülebilir. Yayılım ve büyüme hızlarından dolayı tedavi edilmesi zorlaşmaktadır. Büyük hücreli karsinomun bir alt tipi küçük hücreli akciğer kanserine benzeyen bir yapıda olup hızlı gelişen bir türdür [31].

3) Skuamöz Hücreli Karsinom: Skuamöz hücrelerinde başlamaktadır. Akciğerin hava yollarının düzgün olmasını sağlayan yassı hücrelerdir. Bronşların yakınında bulunmaktadır [31].

2.2.3.2.2. Küçük Hücreli Akciğer Kanseri

Akciğer kanseri türlerinin %10-15'ini oluşturmaktadır. Diğer bir ismi yulaf hücresi kanseridir. Bu tip akciğer kanseri küçük hücreli olmayan akciğer kanseri türünden daha hızlı yayılmakta ve büyümektedir. Küçük hücreli akciğer kanseri tespit edildiği zaman akciğer kanserine sahip insanların yaklaşık %70'inde kanser yayılmıştır. Bu kanser türü hızlı yayılmasına rağmen kemoterapi ve radyasyon tedavisine iyi bir ölçüde yanıt vermektedir. Ancak çoğu insan için bu kanser türü bir noktada tekrar oluşacaktır [31].

Bütün bu akciğer kanseri türleri arasında bu çalışma için küçük hücreli olmayan akciğer kanseri türlerinden adenokarsinom türü incelenmiştir. Ayrıca adenokarsinom türüne ait histopatolojik görüntüler bu çalışmada kullanılmıştır.

2.2.3.3. Küçük Hücreli Olmayan Akciğer Kanseri Evrelendirme

Akciğer kanserinde tanıya uygun tedavi yönteminin seçilmesi için evrelendirme önem arz etmektedir. Küçük hücreli olmayan akciğer kanserinin evrelendirilmesinde TNM sistemi kullanılmaktadır. T primer tümörünü, N lenf bezlerini, M uzak metastazı göstermektedir. 1997 yılında yapılan bir küçük hücreli olmayan akciğer kanseri sınıflaması kullanılmaktadır [32]. Bu sınıflama aşağıdaki Tablo 2.1.'de gösterilmiştir.

Tablo 2.1. Küçük hücreli olmayan akciğer kanserinin TNM sınıflandırılması [32]

	Primer Tümör (T)
Tx	Primer tümörün belirlenememesi, malign hücrelerin saptanmasına karşın görüntüleme yöntemleri ile tümörün gösterilememesi durumu
T0	Primer tümörün belirti göstermemesi
Tis	Karsinoma insitu
T1	En geniş çapı 3 santimetreden küçük veya eşit olan, akciğer ve visseral plevra ile çevrili bir durumda, bronkoskopik olarak lob bronşundan daha proksimale invazyon göstermeyen tümör
T2	Tümörün en geniş çapı 3 santimetreden büyük olan veya ana bronşa invaze fakat ana karinaya uzaklığı 2 santimetreden büyük veya eşit olan veya visseral plevra invazyonu var veya hiler bölgeye uzanan ancak tüm akciğeri kapsamayan atelektazi veya obstrüktif pnömoni var ise
T3	Tümörün herhangi bir büyüklükte olup göğüs duvarı, diyafragma, mediastinal plevra, parietal perikard gibi yapılardan birine direkt yayılım göstermesi veya karinaya 2 santimetreden daha yakın olup karinayı tutmayan ana bronştaki tümör veya bütün bir akciğeri kapsayacak atelektazi veya obstrüktif pnömoni ile birlikte olan tümör
T4	Tümörün herhangi bir büyüklükte olup mediasten, kalp, büyük damarlar, trakea, özafagus, vertebral kolon, karina gibi yapılardan birine yayılması veya malign plevral ya da perikardial efüzyon ile birlikte olan tümör ya da tümörle aynı lob içinde olan satellit nodül
	Lenf Bezleri(N)
Nx	Bölgesel lenf bezlerinin değerlendirilmeye alınamaması
N0	Bölgesel lenf bezine tümör sızması bulunmamakta
N1	Aynı taraf peribronşiyal veya aynı taraf hiler lenf bezlerine tümörün sızması ve primer tümörün direkt yayılması ile intrapulmoner bezlerin tutulması
N2	Aynı taraf mediastinal veya subkarinal lenf bezlerine tümörün sızması
N3	Karşı taraf mediastinal veya hiler, aynı veya karşı taraf supraklaviküler veya skalen lenf bezlerine tümör sızması
	Uzak Metastaz
Mx	Uzak metastaz varlığının değerlendirilememesi
M0	Uzak metastaz yok
M1	Uzak metastaz var

Aşağıda verilen Tablo 2.2.'de yukarıda verilen sınıflandırmaya karşılık gelen küçük hücreli olmayan akciğer kanseri evreleri verilmiştir.

Tablo 2.2. Küçük hücreli olmayan akciğer kanserinin evreleri [32]

EVRE 0	TisN0M0
EVRE 1A	T1N0M0
EVRE 1B	T2N0M0
EVRE 2A	T1N1M0
EVRE 2B	T2N1M0, T3N0M0
EVRE 3A	T3N1M0, T1-3N2M0
EVRE 3B	T4N0-3M0, T1-4N3M0
EVRE 4	Herhangi bir T M1, herhangi bir N M1

2.2.4. Akciğer Görüntüleme Yöntemleri

Akciğer görüntüleme yöntemleri ile akciğere ait kanserlerin görüntülenmesi sağlanmaktadır. Akciğer kanserinin tanı ve tespiti için oldukça önemli bir adımdır. Akciğer görüntüleme çeşitleri; PET, BT, X ışınları, MRG, bronkoskopi, sputum sitoloji yöntemleridir. Bu yöntemler içerisinde sık kullanılanlar MRG,BT,PET, ve X ışınlarıdır.

2.2.4.1. Akciğer Radyografileri

Akciğer grafisi çekiminde X ışınları kullanılmaktadır. Hastalarda önce ön ve arka akciğer grafisi ve gerekli durumlarda sol yan akciğer grafisi çekilir. Ön ve arka akciğer grafisi çekilirken X ışınları sırttan girer ve ön kısımdan çıkarak filme düşmektedir. Sol yan akciğer grafisinde hastanın yan tarafından X ışınları girer diğer yan tarafından çıkarak filme düşer. Bu işlem yapılırken hastadan nefesinin tutulması istenir. X ışını ile film arasındaki mesafe 185 santimetredir. Hastaya göre uygun doz belirlenmelidir. Akciğer alt loblarının damarları görünmelidir. Çünkü bu sayede ayarlanan dozun uygunluğu tespit edilmektedir [33,34,35]. Şekil 2.1.'de akciğer radyografisi örneği gösterilmektedir.



Şekil 2.1. Akciğer radyografisi örneği [36].

2.2.4.2 Bilgisayarlı Tomografi

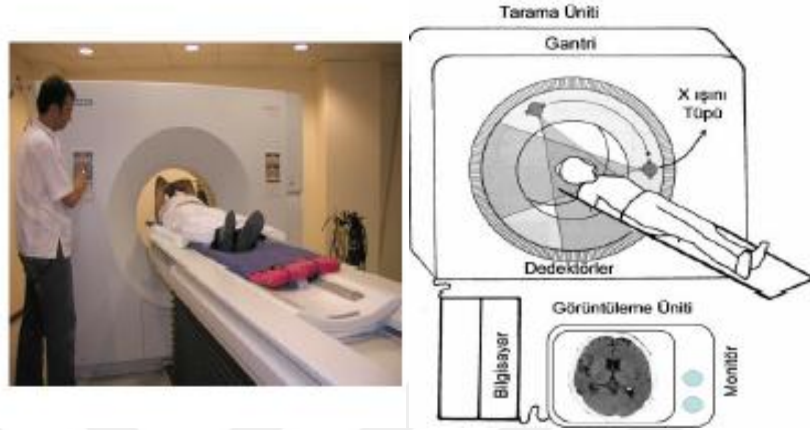
Bilgisayarlı tomografi yöntemini ilk olarak 1963 yılında Cormak teorik olarak ortaya koymuştur. 1972 yılında ise Hounsfield tarafından tanı olarak koyulmuştur. Bu ilerleme X ışınlarının keşfinden beri en büyük ilerlemedir [37,38]. Röntgen cihazının ilham olarak alınmasının yanında, farklı bir şekilde X ışınları inceltilerek, vücuda çizgisel biçimde gönderilerek iki boyutlu kesit görüntüleme yöntemidir. Bilgisayarlı tomografide röntgen tüpü ve görüntü alıcılarının hasta etrafında döndürülmesiyle işlem sağlanmaktadır. Dönme hareketiyle matematiksel çözüm yardımıyla iki boyutlu kesit verileri elde edilmektedir.

Bilgisayarlı tomografi cihazı üç parçadan oluşmaktadır. Tarama bölümü, bilgisayar sistemi ve görüntüleme bölümünden oluşmaktadır.

Tarama bölümünde; X ışını tüpü ve dedektörlerin bulunduğu kare şeklinde bir kutu ile hastanın giriş yaptığı bir aralık vardır. Dedektör zinciri denilen yapı ile bu aralık çevrelenir. Kesit alınacağı zaman bu zincirden ışınlar gönderilerek X ışını tüpü sürekli dönüş yapar.

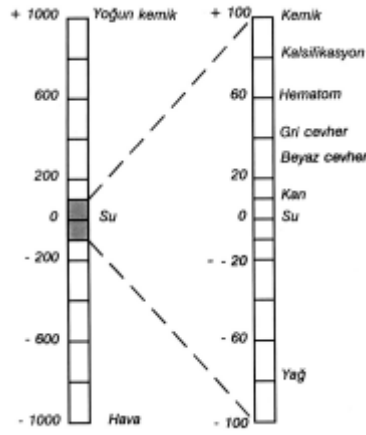
Bilgisayar sistemi dedektörler üzerine düşen ışınların hesaplanarak gönderildiği sistemdir.

Görüntüleme bölümü sayısal işaretlerden oluşan görüntünün ortaya konulduğu ve işlendiği bölümdür. İşlenen görüntüler film üzerine geçirilir [39,38]. Şekil 2.2.'de bilgisayarlı tomografi cihazı ve üniteleri gösterilmektedir.

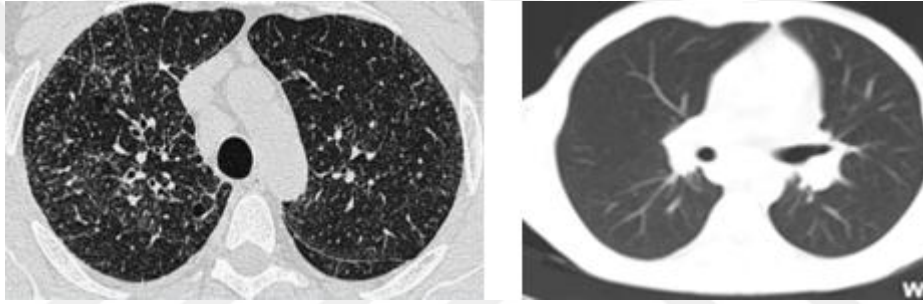


Şekil 2.2. Bilgisayarlı tomografi cihazı ve üniteleri [40].

Bilgisayarlı tomografide görüntü bir kesit görüntüsüdür. İlk olarak X ışın tüpünün kesit düzleminde 360° dönerek bir ışın demeti göndermesidir. X ışını vücuda girmeden önce ve vücuda girdikten sonra ölçülür. Dedektörler tarafından ışını ne kadar soğurduğu hesaplanır. Bilgisayarlı tomografi görüntüleri piksellerden oluşur. Bu piksellere görüntü matriksi denir. Matriks iki kenardaki piksellerin çarpımı şeklinde gösterilir. Bu matriks genelde 512×512 'dir. Sonra dedektörlerin ölçtüğü değerler bilgisayar tarafından X ışınları tutma değerlerine dönüştürülür. Değer dönüşümü Hounsfield cetveline göre yapılır. Şekil 2.3.'te Hounsfield cetveli gösterilmiştir. Bu cetvele göre bütün değerler sayısal değer almaktadır. Daha sonra sistem tarafından aldıkları sayılara göre bütün değerler gri, siyah, beyaz renge boyanmaktadır. Böylece görüntü oluşumu tamamlanmaktadır [40]. Şekil 2.4.'te bilgisayarlı tomografi sonucunda elde edilen görüntülerin örneği verilmiştir.



Şekil 2.3. Hounsfield cetveli [40].



Şekil 2.4. Bilgisayarlı tomografi sonucunda elde edilen görüntüler [40].

2.2.4.3. Manyetik Rezonans Görüntüleme (MRG)

Manyetik rezonans görüntüleme bir çok yönteme karşın üstün olmasına rağmen akciğer görüntüleme için uygun değildir. Çünkü MRG'nin çalışma prensibi sabit manyetik alan içerisinde bulunan dokuya radyo dalgalarının gönderilmesiyle dokuda bulunan hidrojen atomlarının saptırması prensibine dayanır. Ancak akciğer dokusunda hidrojen atomu fazla bulunmamaktadır. Ayrıca çekim sırasında hastanın uzun bir süre hareketsiz kalması gerekmektedir. Bu olanaksız bir durum oluşturmaktadır. Bu sebeplerden dolayı MRG akciğer görüntüleme için uygun değildir [41,38].

2.2.4.4. Pozitron Emisyon Tomografi (PET)

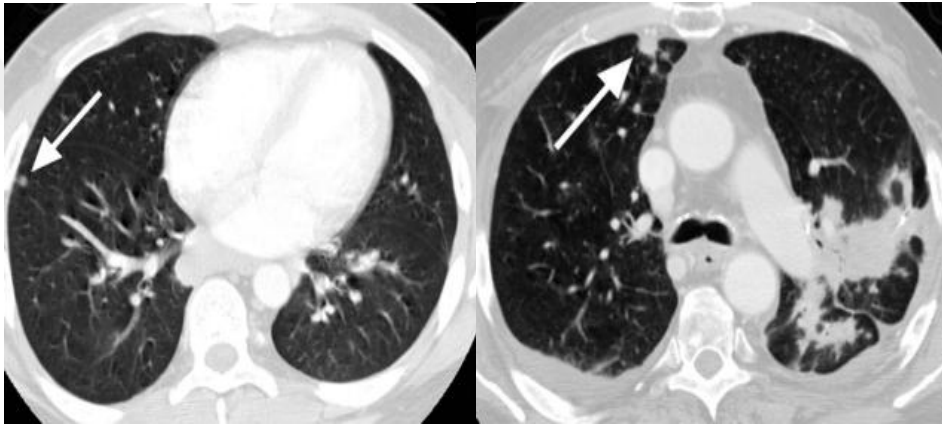
Flor-18 izotopu ile işaretli işaretlenmiş florodeoksiglikoz (FDG) ile yapılan bu tomografi yöntemi malign tümörler görüntülenmektedir. Bu yöntem ile kanserin evrenmesi tedavi takibi doğruluk oranları iyileşmiştir. Akciğer kanseri de yoğun PET uygulamalarının başında gelmektedir [42]. Glikoz metabolizmasının artışıyla beraber

malign tümörlerde FDG tutulması artmaktadır. Böylece PET yöntemiyle malign tümörün belirlenmesi sağlanmıştır [42,43].

Pozitron emisyonunda hücre çekirdeğindeki proton, nötron ve pozitrona dönüşür. Pozitron bir elektrona çarparak iki adet 511 kiloelektronvolt enerjiye dönüşürerek iki foton oluşturur. PET cihazlarında bulunan dedektörler vasıtasıyla oluşmuş olan bu fotonlar algılanarak bilgisayar yardımıyla görüntü haline getirilir [44,45].

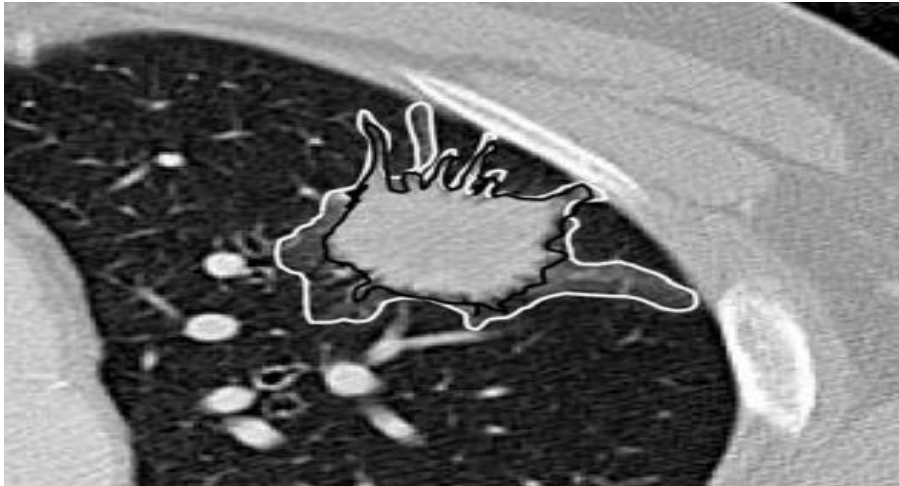
2.2.5. Akciğer Nodüllerini Gösteren Bilgisayarlı Tomografi Görüntüleri

Şekil 2.5.'de a ve b bölümünde akciğer nodüllerini içeren bilgisayarlı tomografi görüntülerine ait örnekler verilmiştir. Oklar ile gösterilen kitleler akciğer nodülleridir.



(a) Akciğer nodülü örneği

(b) Akciğer nodülü örneği



(c) Etrafi çizilmiş akciğer nodülü örneği

Şekil 2.5. Örnek akciğer nodülleri [46].

Şekil 2.5.'in c bölümünde ise iki radyolog tarafından etrafı çizilmiş nodül gösterilmektedir. Bir radyolog beyaz renkte, diğer radyolog siyah renkte çizimi gerçekleştirmiştir.

2.2.6. Akciğer Tümörlerine Ait Histopatoloji Görüntüleri

Mikroskopik histopatoloji patologlar tarafından incelenen tümör görüntülerinin incelenmesiyle akciğer kanserinin değerlendirilmesi ve tanısında çok önemlidir. Çünkü histolojik özellikler sağ kalım sonuçları ile ilişkilidir. Tümöre ait histomorfolojinin dikkatli bir şekilde incelenmesi hastaya için büyük bir önem taşımaktadır. Ayrıca ön görü belirteçleri için kişiselleştirilmiş kanser yönetiminin de ana unsurlardan biridir [47].

Günümüzde akciğer kanseri örneklerinin histolojik özellikleri için ışık mikroskobu kullanılmaktadır. Işık mikroskobu vasıtasıyla alınan görüntülerin kalitesi tanı için çok önemlidir. Çünkü histopatoloji görüntülerinin görüntü kalitesinin düşük olması sonucunda, görüntülerin yanlış değerlendirilmesi bir çok hastada yaşam kalitesinin düşmesine yol açan kötü tercihlere yol açabilmektedir [47].

Bilgisayarlı görüntü işleme yöntemiyle histopatoloji görüntülerinin değerlendirmesinde doğruluk ve tutarlılık arttırımıyla tanısal tutarlılığı sağlamak için karar desteği sağlamaktadır. Otomatik histopatolojik görüntü analiz sistemleriyle çeşitli malign tümörlerin ön görüleri için çok değerlidir [47].

2.2.6.1. Akciğer Tümörü Histopatoloji Görüntülerinin Elde Edilmesi

Akciğer hücresi histopatoloji görüntülerini ışık mikroskobu yardımıyla elde edebilmek için hastalara ait akciğer dokularından tarama ve tanı için hastalardan parça alınması işlemi ilk yapılması gerektirir. Hastalardan parça alınarak elde edilen numuneden histopatoloji görüntüleri elde edebilmek için bazı adımlar izlenmelidir. Bu adımlar;

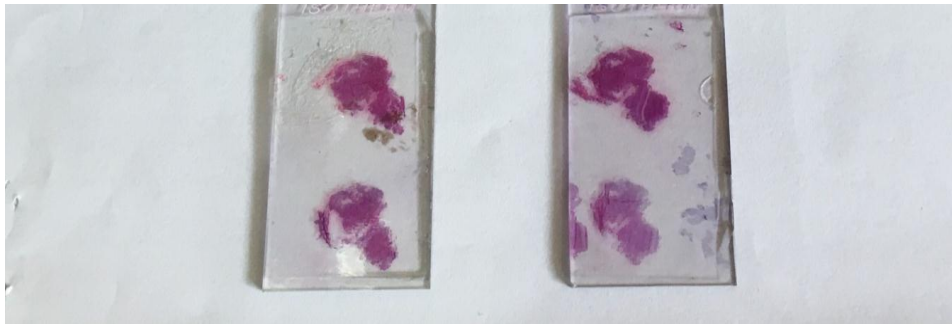
Tespit (Fiksasyon): Dokular insan vücudundan ayrılırsa bile canlıdır. Ayrıca eğer dokular bir hastalığa sahipse taşıdıkları hastalığın bulgularını barındırırlar. Tespit sayesinde dokuların andaki görünümü kötü etkilerden koruyarak bozulması önlenmektedir. Çünkü eğer tespit işlemi bir dokuya yapılmazsa doku, bakterilerin etkisiyle tanısal işlemlerde kullanılmayacak duruma gelir. Tespit işlemi için dokular tespit sıvısı içerisine konurlar. Bu sıvı seyreltilmiş formaldehit solüsyonudur. Tespit işlemi dokunun türüne ve kalınlığına göre değişen sürelerde tespit sıvısında bekletilir. Solüsyonda bekletilen dokular uzun zaman sonra bile dokunun kullanılmasına olanak sağlamaktadır [48].

Takip (Doku İşleme): Tespitten sonraki adımların hemen hemen hepsi dokuların otomatik makine ile yapılan işlemlerden geçmesiyle yapılmaktadır. Burada ilk aşama tespit sıvısının ve dokunun içinde bulunan sıvının dokudan uzaklaştırılmasıdır. Buna dehidratasyon denilmektedir. Çünkü dehidratasyon işlemi yapılmadığı zaman doku yumuşak olmaktadır. Yumuşak olan dokudan ince kesit ve ya kesitler alınmasına olanak yoktur. Daha sonra susuz dokuya sırasıyla alkolle ve ksilolle infiltre olan aralıklarla ısıtılan sıvı parafinin girmesi sağlanır. Takip işlemleri oda sıcaklığı ile 60 °C sıcaklıkları arasında yapılmaktadır. Vakum uygulaması ile dokular daha iyi ve kısa sürede işlenmektedir. Ayrıca bu işlemlerin süresini kısaltmak için özel mikrodalga fırınlar kullanılabilir [48].

Bloklama: Parafin uygulanmış dokular kalıplara konularak üzerlerine ısıtılmış parafinin dökülmesinden sonra soğutularak bloklar elde edilir. Bloklar sayesinde dokuların çok ince bir şekilde kesilmesine olanak sağlanır [48].

Kesme: Parafin bloklar mikrotom adı verilen cihaz ile 4-5 mikron kesilir. Elde edilen kesitler lamlar üzerine ılık su banyosunda alınır. Sonra elde edilen kesitlerden parafin maddesi uzaklaştırılır. Böylece istenilen boya rengine göre kesitlerin boyanması işlemine geçilir [48].

Boyama: Kullanılan boya çeşidi hematoksilin ve eosindir. Hücrelerin çekirdekleri hematoksilin ile diğer bölgeler eosin ile boyanmaktadır. Hastalıkların kesin teşhisi için boyanan preparatların kullanılması yeterlidir [48]. Şekil 2.6.'da Fırat Üniversitesi Tıp Fakültesi Hastanesi Patoloji Laboratuvarından alınmış boyalı preparat örneği gösterilmiştir.



Şekil 2.6. Fırat Üniversitesi Tıp Fakültesi Hastanesi Patoloji Laboratuvarından alınmış boyalı olan preparat örneği

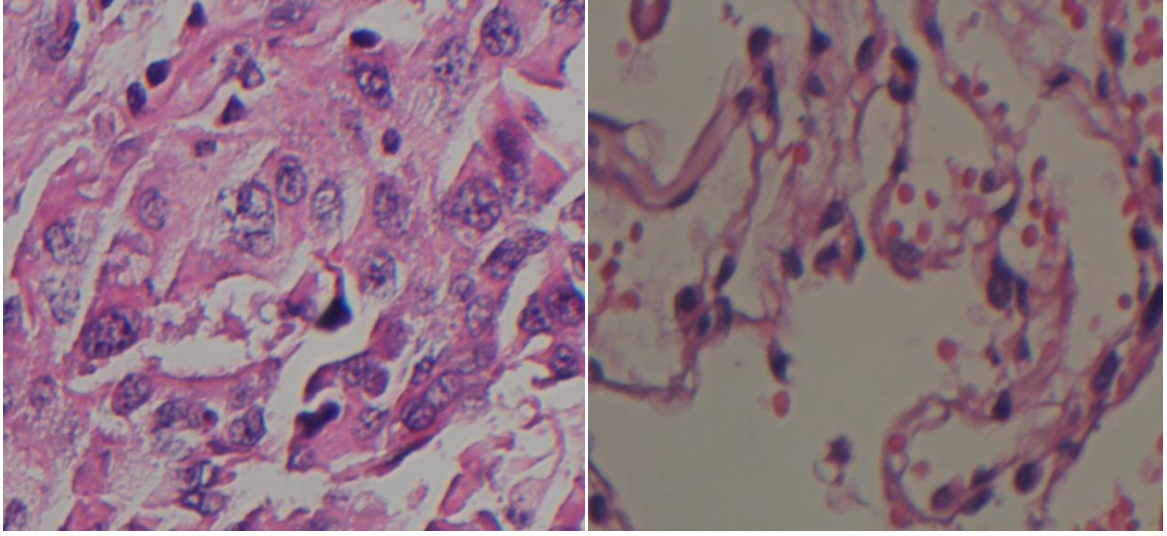
Işık Mikroskobu: Patoloğun sık kullandığı mikroskop olan ışık mikroskobu ile boyanmış preparatlara bakılmaktadır. Patolog baktığı preparatlara morfolojik olarak

inceleme ve sonucunda deęerlendirme yapmaktadır. Işıık mikroskobuyla ışığıın dalga boyundan dolayı en fazla 1000x kadar yakınlaşma dokuya yapılmaktadır [48]. Fırat Üniversitesi Tıp Fakóltesi Hastanesi Patoloji Laboratuvarında ışık mikroskobu aşığıdaki Şekil 2.7.'de gösterilmektedir.



Şekil 2.7. Fırat Üniversitesi Tıp Fakóltesi Hastanesi Patoloji Laboratuvarında kullanılan ışık mikroskobu

Işıık mikroskobuyla bakılan preparatlardan görüntü elde edebilmek için kamera kullanılmaktadır. Bunun için Fırat Üniversitesi Patoloji Laboratuvarında Olympus DP 73 marka kamera kullanılmaktadır. Bu kamera vasıtasıyla alınan histopatolojik görüntü bir bilgisayara gönderilmektedir. Bilgisayar vasıtasıyla hem görüntü kaydedilmekte hem de inceleme yapılabilmektedir. Bu çalışmada kullanılacak olan histopatolojik görüntüler ışık mikroskobundan 400x yakınlaştırmayla elde edilmiştir. Histopatolojik görüntüler Fırat Üniversitesi Tıp Fakóltesi Hastanesi Patoloji Laboratuvarından elde edilmiştir. Elde edilen görüntüler Şekil 2.8.'de gösterilmektedir. Şekil 2.8'e ait a bölümündeki görüntü malign (Kötü Huylu) tümöre ait histopatolojik görüntüdür. Şekil 2.8'e ait b görüntüsü benign (İyi Huylu) tümöre ait histopatolojik görüntüdür. Bu çalışmadaki histopatolojik görüntüler için 60 adet hastadan alınmıştır. Ayrıca her bir hasta için alınmış olan dokudan 4 adet farklı kesit görüntüsü alınmıştır. 120 adet malign ve 120 adet benign olmak üzere toplam 240 adet histopatolojik görüntü elde edilmiştir.



(a) Benign tümör

(b) Malign tümör

Şekil 2.8. Fırat Üniversitesi Tıp Fakültesi Hastanesi Patoloji Laboratuvarından elde edilen görüntülerin örnekleri

3. MATERYAL VE METOT

Bu çalışmada Fırat Üniversitesi Tıp Fakültesi Radyoloji bölümünden tomografi görüntüleri, Fırat Üniversitesi Tıp Fakültesi Patoloji bölümünden ışık mikroskobu yardımıyla histopatoloji görüntüleri akciğer kanseri teşhisi için elde edilmiştir. Elde edilen bu görüntüler akciğer kanserinin erken tanısı ve teşhisi için bilgisayarlı metotların kullanılmasına olanak vermektedir. Bu sebeple görüntüler bazı aşamadan geçirilerek analiz yapılmaktadır. Bu aşamalar aşağıda yazılmıştır:

1) **Özellik Ekleme:** Bu aşamada görüntülere uygulanan çeşitli yöntemlerle görüntü özellik noktalarının tespiti sağlanmıştır. Bu yöntemler ile görüntülerin ölçek değişimlerinden etkilenilmeyecektir.

2) **Boyut İndirgeme:** Bu aşama ile görüntülerin veri boyutunun indirgenmesi sağlanmaktadır. Bu aşama sayesinde görüntülere normalizasyon işlemi yapılmaktadır. Sistemin performansı için çok önemli bir aşamayı oluşturmaktadır.

3) **Sınıflandırma Aşaması:** Bu aşamada boyutu indirgenen görüntülerin belirli gruplara ayrılması sağlanmaktadır. Bu gruplara ayrılarak belirli kriterlere göre ayrılma işlemi söz konusudur. Bu sınıflandırma aşaması için çeşitli yöntemler kullanılmaktadır.

3.1. Özellik Ekleme İçin Kullanılan Yöntemler

3.1.1. Maksimum Durağan Ekstrem Bölgeler (MDEB)

MDEB algoritması, görüntülerde daire veya elips benzeri şekiller (lekeler) bulmak için kullanılan bir algoritmadır. Algoritma, bu şekilleri hesaba katan kilit noktaları seçerek bu kilit noktadaki özniteliklerini hesaplamaktadır [49]. Bir MDEB bölgesi [50], kalıcılığı değişen eşik değerine bağlı olan belirli bir eşik değer üzerinde birbirine bağlı bir dizi noktadan oluşur. Yani, seçilen bölge bir eşik değerleri setine bağlı olmayan yerel bir ikili formdur. MDEB bölgelerini bulmak havza metoduna benzer olarak şekilde çalışmaktadır [51]. Eşik değeri [0-255] arasında değişirken, birbirine bağlanan ve tüm sahnelerde çok fazla değişmeyen bölgeler, MDEB bölgeleri olarak tanımlanmaktadır. Yöntemin uygulanmasında; görüntü noktaları parlaklık değerine göre sıralanır ve noktalar görüntüye artan veya azalan bir şekilde yerleştirilir. Bu işlem sırasında, birleştirme tespit algoritmasıyla bulunan bileşen bölgeleri ve alanları bir listede tutulur. Eşikleme işleminin

her aşamasında, iki bölgenin birleşmesi ile küçük bölge geniş bölgeye dahil edilir ve küçük bölge listeden kaldırılır. Zon alanı değişiminin, eşik değerinin artırılması / azaltılmasıyla en aza indirildiği eşik değerler, maksimum stabil ekstremal zonları üreten eşik aralıkları olarak seçilir. Başka bir deyişle, bağlı bileşenlerin tüm eşik değerlerinde oluşturduğu uç bölgelerin sınırları $A_1, \dots, A_{i-1}, A_i, \dots$ bir dizi bitişik bölge olarak gösterilir. Bu dizi $A_i \subset A_{i+1}$ koşulunu sağlar. Dizinin A_i^* uç bölgesini maksimum kararlı olarak seçmek için [51]; Yerel minimum olması, $f(i) = |A_i \pm \Delta|/|A_i|$ ifadesinin i^* değerinde gereklidir. $| \cdot |$ bölgenin alanını belirtir ve Δ değer, yöntemin bir parametresini belirtir. \pm işareti ayrıca hem azalan hem de artan eşik değerleri için yerel minimum olduğunu gösterir. Bu işlem, maksimum ekstremal stabil bölgeleri elde etmek için ekstrem bölge setindeki tüm bölge dizilerine uygulanır [51]. MDEB yönteminde, benzer bir renge sahip olan komşu görüntü noktaları, kümelenmiş kümelenmeye dayalı kümelemeye tabi tutulur. Kümeleme işlemi için, birbirine bitişik dört veya sekiz noktadaki renk mesafeleri ilişkili bir listede tutulur. Algoritmanın her aşamasında, $b \in [0 \dots B]$, görüntü noktaları aşamalı olarak etiketlenir. Görüntü noktalarının koordinat alanı $\Omega = [1 \dots K] \times [1 \dots M] \subset Z^2$ etiket kümesi olarak belirtilirse, her adım $E_b: \Omega \rightarrow N$ eşlemesi olarak ifade edilir. Etiketlemenin bir sonucu olarak, aynı noktaları içeren bağlantı noktaları görüntüdeki aşırı bölgeleri belirtir. Görüntü ekstremal bölgesinin tüm komşu noktaları arasındaki mesafe, söz konusu adım için hesaplanan bir eşik değeri olan $d_{bhr}(b)$ değerinden düşük olmalıdır. Renk uzayındaki görüntü noktalarının mesafesi ki-kare mesafesi kullanılarak hesaplanır. Başlangıçta, B_0 etiket görüntüsündeki tüm değerler 0 olarak etiketlenir. Etiket görüntüsünde, $d_{bhr}(b)$ 'dan daha kısa mesafedeki tüm komşu noktalar yeni bölge olarak etiketlenir ve E_{b+1} etiket görüntüsü elde edilir [51].

Görüntü noktaları arasındaki uzamsal ilişki nedeniyle, tüm komşu noktaların mesafeleri tekdüze bir dağılım göstermez. Mesafelerin büyük çoğunluğu küçük değerlere sahiptir ve çok az sayıda büyük mesafe değerleri vardır. Bu nedenle, her adımda eşik değeri doğrusal olarak artırılır, bu da başlangıçta çok hızlı bir etiket değişikliği ile sonuçlanır. Böylece, adımların sonuna kadar birçok noktadaki etiketler değişir [25]. Her adımda eşit sayıda görüntü noktasının etiketini değiştirmek için, görüntüdeki tüm komşu noktalar arasındaki mesafe rastgele bir değişken olarak alınır ve eşik değerler, kümülatif dağılım fonksiyonunun (KDF) sıralı tersine göre değiştirilir [25]. Renkli görüntüler için ki-

kare KDF, Denklem 3.1'de olduğu gibi hesaplanır, burada örnek kümenin ortalaması hesaplanarak $\mu = \frac{3\lambda}{2}$ kestirilebilir.

$$H_3(c) = \sqrt{\frac{4c}{\lambda\pi}} e^{-3c/2\pi} + \text{erf}(\sqrt{3c/2\mu}) \quad (3.1)$$

Kestirme sonucunda, Denklem 3.2, ortalama tahminden sonra eşik değerlerini bulmak için kullanılır.

$$d_{bhr}(b) = H^{-1}\left(\frac{b}{B}\right), b \in [0 \dots B] \quad (3.2)$$

Sonra, $d_{bhr}(b)$ değerinin artırılmasıyla tespit edilen ekstremal bölgelerin alan değişiklikleri, her adımda kontrol edilerek maksimum sabit olanlar tespit edilir. Ek olarak, belirli bir değerden daha küçük olanlar maksimum stabil ekstremal bölgelerden elimine edilir [51].

3.1.2. Hızlandırılmış Sağlam Özellikler (HSÖ)

HSÖ yöntemi ölçek ve dönüklük değişimlerinden etkilenmemektedir. Ayrıca HSÖ yöntemi bir tanımlayıcıdır. İlk olarak 2006 yılında Herbert Bay tarafından ortaya konulmuştur. HSÖ yöntemi, iki boyutlu olan Haar dalgacıklarının toplamı ve integral görüntülerin kullanımlarına dayandığı açıklanmaktadır. Tamamlayıcı özelliklere sahip görüntüler, dikdörtgen bir alanın hızlı olarak hesaplanmasına olanak sağlamaktadır. Konumu $I = (a, b)$ olan $I_{\Sigma}(a)$ tamamlayıcı özelliğe sahip, görüntü merkezinde şekillenen dikdörtgen bir alan içerisindeki I giriş görüntüsündeki piksellerin toplamları olarak ifade edilmektedir. Denklem 3.3'de bu ifade gösterilmektedir. Tamamlayıcı görüntüler hesaplandıktan sonra dikdörtgen alan içerisindeki yoğunlukların toplanması için üç işlem yapılmaktadır. Böylece hesaplama zamanı boyuttan bağımsız olmaktadır [52,53].

$$I_{\Sigma}(a) = \sum_{k=0}^{k \leq a} \sum_{l=0}^{l \leq b} I(k, l) \quad (3.3)$$

HSÖ tanımlayıcısı, detektör olarak Hessian matrislerinin determinantını kullanır. Çünkü bu hesaplama sonucunda determinantın maksimum olduğu konum görüntü bölgelerinin elde edilmesi sağlanmaktadır. Ölçek ve konum seçimi için bu determinant kullanılmaktadır. Determinant, görüntünün ikinci dereceden türevi yardımıyla görüntüye ait maksimum ve minimum noktaların bulunmasında kullanılmaktadır. Görüntü noktası $I = (a, b)$ olan bir görüntü için Hessian matrisi Denklem 3.4'te ifade edilmektedir.

$$H(I(a, b)) = \begin{bmatrix} \frac{\partial^2 I}{\partial a^2} & \frac{\partial^2 I}{\partial b \partial a} \\ \frac{\partial^2 I}{\partial a \partial b} & \frac{\partial^2 I}{\partial b^2} \end{bmatrix} \quad (3.4)$$

Denklem 3.4'te bulunan Hessian matrisinin determinanı Denklem 3.5'de gösterilmektedir.

$$\det(H(I(a, b))) = \frac{\partial^2 I}{\partial a^2} \frac{\partial^2 I}{\partial b^2} - \left(\frac{\partial^2 I}{\partial a \partial b}\right)^2 \quad (3.5)$$

Görüntünün ikinci dereceden türevinin alınmasıyla ölçek uzayında değişmeyen kilit noktalar elde edilmektedir. I görüntüsünün σ ölçeğinde $F(a, b)$ noktasındaki Hessian matrisi Denklem 3.6'da hesaplanmaktadır.

$$H(F, \sigma) = \begin{bmatrix} D_{aa}(F, \sigma) & D_{ab}(F, \sigma) \\ D_{ab}(F, \sigma) & D_{bb}(F, \sigma) \end{bmatrix} \quad (3.6)$$

$D_{aa}(F, \sigma)$, I görüntüsündeki F noktasının Gaussun ikinci dereceden türevinin $\frac{\partial^2}{\partial a^2} g(G)$ alınması ile oluşturulmaktadır. $D_{ab}(F, \sigma)$ ve $D_{bb}(F, \sigma)$ değerleri de aynı şekilde hesaplanmaktadır [52,53].

Gaussun laplası yaklaşımı ile hesaplanan Gaussun ikinci dereceden kısmi türevlerinin a, b, ab yönünde ağırlıklandırılmış kutu filtre yaklaşımı (D_{aa}, D_{ab}, D_{bb}) ile gösterilmektedir. Gauss kullanılarak Hessian determinanı Denklem 3.7'deki gibi hesaplanmaktadır. $F(a, b, \sigma)$ konumundaki görüntü bölgesine determinant denilmektedir. w değeri 0,9'dur [52,53].

$$\det(H_{approx}) = D_{aa}D_{bb} - (wD_{ab})^2 \quad (3.7)$$

3.1.2.1. Ölçek Uzayı Oluşturulması

Eşleştirilmesi istenen görüntüler aynı ölçekte olabileceği gibi farklı ölçekte de olabilmektedir. Bir ölçek uzayı bütün ölçekler arasında maksimum ölçeği bulabilmek için kullanılan bir fonksiyondur. Ölçek uzayı görüntü piramidine benzemektedir. Görüntüler sürekli Gaussian ile yumuşatılarak piramidin en üst seviyesine ulaşılabilir. Böylece boyut azaltma sağlanmaktadır [52,53].

3.1.2.2. İlgili Noktası Tanımlayıcısı Ve Eşleme

HSÖ tanımlayıcısı, Fast-Hessian detektörü tarafından belirlenen ilgi noktasının ölçüğe bağlı komşu içerisinde piksel yoğunluklarının nasıl dağıtıldığını göstermektedir. Haar dalgacık olarak bilinen filtreler sayesinde kullanılan görüntülerin hesaplanma süresi azalmakta ve sağlamlığı artmaktadır. Bu dalgacıklar a ve b gradyentlerini bulmak için kullanılır [52,53].

Tanımlayıcıların elde edilmesi için iki aşama kullanılmaktadır. Birinci aşama her bir ilgi noktasının görüntü dönüklüğüne karşı değişmezliğini sağlamak için yön atamaktır. Sonra seçilen yöne atanan bir kare bölge elde edilmektedir. HSÖ bu kareden oluşmaktadır. Son olarak da elde edilen özellikler iki görüntüde eşlenmektedir [52,53].

3.1.2.3 Haar Dalgacık Yanıtlarının Toplamına Dayanan Tanımlayıcılar

Tanımlayıcıların oluşturulmasındaki ilk aşama ilgi noktaları etrafında merkezlenmiş kare bir pencere oluşturma ve yön atama işleminde sonucunda oluşan yöne yönlendirilmesidir. Tanımlayıcı penceresi daha küçük 4×4 kare alt bölgelere düzenli olarak bölünmektedir. Bu alt bölgelerin her birinde 2σ boyutlu Haar dalgacıkları 25 dağıtılmış örnek nokta hesaplanmaktadır. Her alt bölgeye ait 5×5 'lik örnek noktalarda Haar dalgacık yanıtları hesaplanmaktadır. Haar dalgacık yanıtları yatay yönde d_a , dikey yönde d_b olarak isimlendirilmiştir. 25 noktanın toplam yanıtı Denklem 3.8'de gösterilmiştir.

$$v_{alt-bölge} = (\sum d_a, \sum d_b, \sum |d_a|, \sum |d_b|) \quad (3.8)$$

Bütün 4×4 alt bölgeler için dört boyutlu tanımlayıcı ve rektörü birleştirilirse tanımlayıcı vektörü $4 \times 4 \times 4 = 64$ boyutunda bir vektör haline gelir [52,53].

3.1.3 Yönlendirilmiş Gradyanlar Histogramı (YGH)

İmgedeki piksellerin yönelim ve büyüklük değerlerinin karakteristiği olarak da adlandırılan YGH yönteminin kullanımı son yıllarda oldukça artmıştır. Nesne ve örüntü tanıma için kullanılmaya başlanan bu yöntem farklı koşullar altında yüksek bir başarı oranı sağlamaktadır. Bu yöntemi ilk defa kullanımını öneren Shashua [54] ve Dalal [55] olmuştur. Bu yöntemin kullanımında temel amaç; imgeyi bir grup temel histogramlar olarak tanımlamaktır. Bu histogramlar imgenin lokal bölgesindeki gradyanların

yönelimlerinin sayısını içermektedir [56,57]. YGH yönteminin oluşumundaki işlem basamakları aşağıda sıralanmıştır:

- Gri seviyeye sahip bir imgenin dikey ve yatay gradyan değerleri elde edilir.
- Yatay ve dikey gradyan değerleri kullanılarak gradyan yönelimi (Θ) ve gradyan büyüklüğü (m) değerleri bulunur.
- Gradyan yönelimleri imgesi [$m \times n$] boyutunda hücrelere bölünür. Her bir hücre içerisindeki gradyan yönelimleri 45° 'lik 8 farklı durumla etiketlenir.
- Belirlenen her bir blok için gradyan büyüklüğü ve yönelimleri ile lokal histogramlar oluşturulur.
- Oluşan lokal histogramlar normalize edilir [58].

Bir imgenin YGH değerlerinin çıkarılması için gerekli hesaplamalar Denklem 3.9, Denklem 3.10, Denklem 3.11, Denklem 3.12'de verilmiştir. İmgeye Sobel filtreleri uygulanarak I_x ve I_y olmak üzere imgenin kenarları belirlenir. Gradyan (G) ve gradyan yöneliminin (Θ) hesaplamaları için aşağıda bulunan formüller kullanılır.

$$I_x = I * S_y \quad (3.9)$$

$$I_y = I * S_d \quad (3.10)$$

$$|G| = \sqrt{I_x^2 + I_y^2} \quad (3.11)$$

$$\Theta = \arctan \frac{I_x}{I_y} \quad (3.12)$$

YGH algoritması uygulanırken oluşturulan histogramda piksellerin yönelim açılarının gruplanması ile daha anlaşılabilir sonuçlar elde edilebilmektedir. Bu gruplama olayı 0° - 360° aralığında yer alan açısız değerlerin istenilen bir aralığa çekilmesi ile mümkün olmaktadır [57]. Daha sonra imgeye histogram seleleme işlemi uygulanır. Histogram seleleri eşit bölünerek oluşturulur. Böylece imge [36×36] boyutundaki hücrelere bölünmektedir [58].

3.2. Boyut İndirgeme İçin Kullanılan Yöntemler

3.2.1. Temel Bileşenler Analizi (TBA)

Temel bileşenler analizi çalışmaları 1901 yılında Karl Pearson tarafından başlatılmıştır, 1933 yılında Hotelling tarafından bu yöntem geliştirilmiştir [59]. TBA çok sayıda birbiri ile ilişkili değişkenler içeren veri setinin boyutlarını değişimlerin korunmasıyla daha az boyuta indirgenmesi ile oluşan bir dönüşüm tekniğidir [60].

Dönüşüm sonucunda ortaya çıkan değişkenler ilk değişkenlerin temel bileşenleri olarak isimlendirilir. Varyans değeri en büyük olan bileşen ilk temel bileşendir. Diğer bileşenlerde varyans değerleri azalarak sıralanır. Gürültüye karşı düşük etkilenme, bellek ve kapasite ihtiyaçlarının azalması, küçük boyutlu uzaylarda daha etkin çalışması TBA'nın temel avantajları arasında sıralanabilir [61].

Eğitim kümesi veri matrisi (A), M tane $N \times 1$ 'lik gözlem vektöründen oluştuğu kabul edilirse Denklem 3.13 ve Denklem 3.14'de olduğu gibi bir sonuç elde edilir.

$$A = [a^1 | a^2 | \dots | a^M] \quad (N \times M) \quad (3.13)$$

$$A = \begin{bmatrix} a_1^1 & a_1^2 & \dots & a_1^M \\ a_2^1 & a_2^2 & \dots & a_2^M \\ \vdots & \vdots & \ddots & \vdots \\ a_N^1 & a_N^2 & \dots & a_N^M \end{bmatrix} \quad (3.14)$$

Çok değişkenli analizde değişkenlerin ölçü birimleri birbirinden farklı olabilmektedir. Fakat daha iyi sonuçların elde edilebilmesi için verilerin ölçü birimleri aynı olmalıdır. Bu sebeple değişken değerlerinin aynı birimde olması sağlanır. Bu da veri ortalamasının 0 yapılmasıyla mümkün olmaktadır. Eğitim vektörlerinin ortalaması m olarak ifade edilir ve Denklem 3.15'de gösterilmektedir. Hesaplanan m ortalama vektörü bütün gözlem vektörlerinden çıkartıldığı zaman değişkenler sıfır ortalamaya sahip olmaktadır. Ortalaması çıkartılmış veri matrisi \tilde{A} olarak ifade edilirse ortalaması çıkarılmış gözlem vektörleri Denklem 3.16'da gösterilir. Sıfır ortalamalı gözlem matrisi ise Denklem 3.17 ve Denklem 3.18'de olduğu gibi ifade edilir.

$$m = \frac{1}{M} \sum_{i=1}^M a^i = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_N \end{bmatrix} \quad (3.15)$$

$$\tilde{A}^i = a^i - m, \forall i \quad (3.16)$$

$$\tilde{A} = [\tilde{a}^1 | \tilde{a}^2 | \dots | \tilde{a}^M] \quad (N \times M) \quad (3.17)$$

$$\tilde{A} = \begin{bmatrix} a_1^1 - m_1 & a_1^2 - m_1 & \dots & a_1^M - m_1 \\ a_2^1 - m_2 & a_2^2 - m_2 & \dots & a_2^M - m_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_N^1 - m_N & a_N^2 - m_N & \dots & a_N^M - m_N \end{bmatrix} \quad (3.18)$$

Bir sonraki adımda sıfır ortalamalı gözlem verisinin kovaryans matrisi Denklem 3.19'da bulunarak Denklem 3.20 elde edilir.

$$C = \tilde{A}\tilde{A}^T \quad (3.19)$$

$$C = \begin{bmatrix} (a_1^1 - m_1)^2 & (a_1^2 - m_1)(a_1^1 - m_2) & \dots & (a_M^1 - m_1)(a_N^1 - m_N) \\ (a_2^1 - m_2)(a_1^1 - m_1) & (a_2^2 - m_2)^2 & \dots & (a_M^2 - m_2)(a_N^2 - m_N) \\ \vdots & \vdots & \ddots & \vdots \\ (a_N^1 - m_N)(a_1^M - m_1) & (a_N^2 - m_N)(a_2^M - m_2) & \dots & (a_N^M - m_N)^2 \end{bmatrix} \quad (3.20)$$

Köşegen üzerindeki değerler varyans, diğerleri kovaryanstır. Varyans diğer boyutlardan bağımsız bir şekilde sadece bir boyuttaki verinin ortalama etrafındaki dağılımını gösterirken, kovaryans ise bir değişkenin diğer değişkenlerle ikili oluşturması ile bir değişim hakkında bilgi verir [62]. Değişkenlerden birinin değeri ile diğerinin arasında doğru orantı varsa bu kovaryans pozitif, ters bir orantı var ise bu kovaryans negatiftir. Değişkenler arasında herhangi bir orantı söz konusu değilse bu kovaryansın sıfır olduğunu gösterir [63].

TBA için kovaryans matrisine öz değer-öz vektör ayrımı Denklem 3.21 kullanılarak sağlanır. C , $N \times N$ boyutlu bir matris, λ herhangi bir skaler ve v sıfırdan farklı bir sütun vektör olmak üzere, Denklem 3.21'i sağlayan λ sayısı C 'nin öz değeri ve v de λ ile ilişkili öz vektörüdür.

$$Cv = \lambda v \quad (3.21)$$

Öz değerler büyükten küçüğe doğru sıralandığında, sıralanmış öz değerlere karşılık gelen öz vektörlerden ilk P tanesi bir matrisin sütunlarını oluşturacak şekilde sıralandığında en iyi izdüşümü gerçekleştiren W iz düşüm matrisi Denklem 3.22'de ki gibi elde edilir. Veri örneklerinin öznitelikleri de öz vektörlerin belirlediği uzay üzerine iz düşüm alınarak Denklem 3.23'de verildiği gibi hesaplanır.

$$W = [w_1 w_2 \dots w_P] \quad (3.22)$$

$$y^i = W^T a^i, i = 1, 2, \dots, M \quad (3.23)$$

Öznitelikler belirlenerek eğitim aşaması sona ermektedir. Sonraki adım, sınıflandırma aşamasıdır. Sınıflandırma aşaması, eğitimde kullanılmayan bir test görüntüsü verildiği durumda, eğitim aşamasındaki öznitelikler kullanılarak test vektörünün, eğitim aşamasındaki sınıflardan birine atanması durumudur. TBA temeline uyumlu sınıflandırıcı, en yakın komşuluk sınıflandırıcısıdır. Bu sınıflandırıcıda hangi sınıfa ait olduğu

bilinmeyen test vektörünün özniteliği hesaplandıktan sonra, eğitim aşamasında belirlenen özniteliklerle kıyaslama yapılır. Test vektörünün özniteliğine en yakın olan eğitim örneğinin sınıfı, test vektörünün sınıfı olarak belirlenmektedir [64].

3.2.2. Genelleştirilmiş Diskriminant Analizi (GDA)

GDA, durumları sınıflandırmak için kategorik bağımlı değerleri genellikle bir ikilik şeklinde kullanılır. Eğer GDA kullanılarak bir veri kümesinin kuvvetli bir şekilde doğru ve yanlış tahminlerin sınıflandırma tablosu sonucu isteniyorsa GDA yüksek bir verim yüzdesi ile doğru sonuç verecektir. GDA'da, çoklu nicel öznitelik, iki veya daha fazla doğal olarak oluşan gruplar arasında ayırım yapmak için kullanılır. GDA örneklerin istatistiksel sınıflandırmasını sağlar ve bunu önceden bilinenler ile yapmaktadır. Ayrıca, GDA grup örneklerinde ortak özellik paylaşımında yardımcı olmaktadır.

GDA tekniği üzerinde çalışan her gruba ait ham veri için GDA fonksiyonu, aşağıdaki Denklem 3.24'de olduğu gibidir [65, 66, 67]:

$$f(C)_m = n_m + \sum_{l=1}^b h_{ml}d_{ml} \quad (3.24)$$

Grupların (C) numarası m ile gösterilmektedir. n_m ise her bir gruba ait bir sabittir. Bir veri kümesini belirli bir gruba sınıflandırmak için kullanılan parametre sayısı b'dir. h_l GDA tarafından belirli bir parametreye (d_l) atanan ağırlıklı katsayıdır. Ağırlık katsayısı değişken kriterler arasındaki mesafeyi maksimize eder. Sınıflandırma tablosu; atama veya tahmin matrisi veya tablosu, GDA'nın performansını değerlendirmek için kullanılır. Tahmin mükemmel olduğunda, tüm durumlar köşegen üzerinde bulunmaktadır. Köşegen üzerindeki durumların yüzdesi doğru sınıflandırmaları göstermektedir [68].

3.2.3. Doğrusal Diskriminant Analizi (DDA)

DDA, veri nesnelерinin sınıf bilgisini de kullanarak sınıfların en iyi ayrıştığı uzayda yer alan vektörleri bulmaya amaçlayan bir lineer boyut indirgeme yöntemidir [69]. DDA en sık kullanılan yöntemler arasında bulunmaktadır. DDA yöntemi için, boyut indirgeme işleminin yapılabilmesi için sınıflar arası saçılım matrisi S_M ve sınıf içi saçılım matrisi S_F değerleri formüle eklenir. b tane sınıfa sahip bir veri kümesi için \bar{a} veri nesnelерinin ortalama değeri olarak gösterilirse, bu matrisler Denklem 3.25 ve Denklem 3.26'da ki gibi hesaplanır:

$$S_M = \sum_{i=1}^b (\bar{a}_i - \bar{a})(\bar{a}_i - \bar{a})^T \quad (3.25)$$

$$S_F = \sum_{i=1}^b \sum_{j=1}^{N_i} (a_{i,j} - \bar{a})(a_{i,j} - \bar{a})^T \quad (3.26)$$

Buradaki \bar{a}_i , i.sınıfın ortalama değerini ifade etmektedir. i.sınıfın j.örneği $a_{i,j}$ ile gösterilmektedir. N_i , i.sınıfındaki örnek sayısı olarak ifade edilmektedir. S_M ve S_F matrislerin bulunmasından sonra Denklem 3.27'deki genelleştirilmiş öz değerler probleminin sonuçlandırılması gösterilmektedir:

$$S_M \alpha = \lambda S_F \alpha \quad (3.27)$$

Bu genelleştirilmiş öz değerler probleminin sonucu olarak en büyük r tane öz değere karşılık gelen r tane öz vektör ile P dönüşüm matrisini oluşturur [70].

3.2.4. Yerel Doğrusal Gömme Analizi (YDGA)

YDGA, Sam T. Roweis ve Lawrence K. Saul tarafından yapılan 2000 yılındaki çalışmasında önerilen doğrusal olmayan boyutsallık azaltma yöntemidir [71].

Veri kümeleri genellikle n-boyutlu bir özellik alanında temsil edilebilir ve her bir boyut belirli bir özellik için kullanılır. YDGA algoritması boyutsallığın azaltılması için denetimsiz bir yöntemdir. Orijinal doğrusal olmayan özellik yapısının geometrik özelliklerini korumaya çalışırken bu n-boyutları azaltmaya çalışır. A_1 verileri için B boyutunun olduğunu ve bu verileri b boyutuna indirgeyerek A_2 verileri Denklem 3.28'de elde edilmektedir:

$$A_1 \in B \rightarrow A_2 \in b \quad (3.28)$$

YDGA ilk önce noktaların en yakınına komşularını bulmaktadır. Daha sonra, her veri vektörüne en yakın komşularının ağırlıklı doğrusal bir kombinasyonu olan bir yaklaşım izlemektedir. Son olarak, vektörleri komşularından en iyi şekilde yeniden yapılandırılan ağırlıkları hesaplamaktadır, daha sonra bu ağırlıklar tarafından en iyi şekilde yeniden yapılandırılan düşük boyutlu vektörleri üretmektedir [72].

YDGA algoritmasının bir avantajı, ayarlanması gereken K'nın veya kümenin bir parçası olarak değerlendirilecek en yakın komşuların sayısı olan yalnızca bir parametrenin olmasıdır. K çok küçük veya çok büyük olarak seçilirse, orijinal verinin şeklini oluşturamaz. Verilen her veri noktası için böylece en yakın komşular K hesaplanmaktadır. Yeni bir nokta oluşturabilmek için her bir noktanın komşularının ağırlıklı toplaması yapılır.

Denklem 3.29’da F noktasına ait A_i verileri için en yakın komşuluğunun bulunduğu boyutu en aza indirgeme işlemi gerçekleştirilmektedir:

$$E(W) = \sum_i |A_i - \sum_F W_{iF} A_F|^2 \quad (3.29)$$

Sonra yeni vektör uzayı Y tanımlanarak, Y’nin boyutu en aza indirgenir. Bu durum Denklem 3.30’da gösterilmektedir [71].

$$H(Y) = \sum_i |Y_i - \sum_F W_{iF} Y_F|^2 \quad (3.30)$$

3.2.5. Çok Boyutlu Ölçekleme Analizi (ÇBÖA)

ÇBÖA, k boyutlu bir uzayda gösterilebilen nesnelere orijinal konumlarına çok yakın bir biçimde daha az boyutlu kavramsal bir uzayda belirterek, nesnelere arasındaki ilişkileri belirlemeyi sağlar. Analiz yönteminin amacı, mümkün olduğunca az boyutta olacak şekilde nesnelere yapısını orijinal şekle yakın bir biçimde ortaya koymaktır [73]. Bu sebeple nesnelere ya da birimler arasındaki uzaklıklar hesaplanır. Bu orijinal uzaklıklara göre uygun olan ve daha az boyuta sahip geometrik gösterim ortaya konulur [74]. ÇBÖA’nde metrik ve metrik olmayan modeller bulunmaktadır. ÇBÖA’nde, çok boyutlu gerçek şekil ile indirgenmiş uzayda kestirilen şekil arasındaki farklılığın bir ifadesi olan stress değeri (yani orijinal uzaklıklar ile gösterim uzaklıkları arasındaki uygunluğun ölçüsü) hesaplanmaktadır. Metrik olmayan ölçekleme için stress değeri Denklem 3.31’de gösterilmektedir [75,76].

$$Stress = \sqrt{\frac{\sum \sum (d_{ij} - \widehat{d}_{ij})^2}{\sum d_{ij}^2}} \quad (3.31)$$

d_{ij} : i. ve j. bireyler arasındaki konfigürasyon uzaklığını temsil etmekte, \widehat{d}_{ij} : i. ve j. bireyler arasındaki veri uzaklığını temsil etmektedir [75,76]. Stress değerlerinin büyüklüklerine göre konfigürasyon uzaklıklarının orijinal uzaklıklara uyumluluğu; stress değeri 0,20 ise zayıf uyum, stress değeri $0,10 < 0,20$ ise orta uyum; stress değeri $0,05 < 0,10$ ise iyi uyum; stress değeri $0,025 < 0,05$ ise mükemmel uyum; stress değeri $0,00 < 0,025$ ise tam uyum şeklinde değerlendirilmektedir. ÇBÖA çözümlerinde sıfıra yakın olan stress değerini veren boyut çözümleri uygun olarak nitelendirilen bir çözümdür [74,77].

3.2.6. Komşuluk Koruyan Gömme Analizi (KKGA)

KKGA, verinin yerel komşuluk yapısını korumayı amaçlayan lineer bir yöntemdir [78]. KKGA yönteminde, d boyutuna sahip c sayısı kadar örnekli bir veri grubu $A = [a_1, a_2, \dots, a_c]$ için, c düğümlü komşuluk grafiğinin oluşturulması ile başlamaktadır. Komşuluk grafiği oluşturulurken bir veri örneğinin ε -komşuluğu veya k en yakın komşuları kullanılır. Oluşturulan komşuluk grafiği kullanılarak, simetrik $c \times c$ ağırlık matrisi W bulunur. Eğer komşuluk grafiğinde a_i ve a_j değerleri birbirine bağımlı değil ise bu komşuluklar için $W_{ij} = 0$ olur. Eğer komşuluk grafiğinde a_i ve a_j değerleri birbirine bağlı ise, bu W_{ij} ağırlık değerleri Denklem 3.32 ve Denklem 3.33’de verilen amaç fonksiyonunu minimize eden değerler olur [79]:

$$\min \sum_i \|a_i - \sum_j W_{ij} a_j\| \quad (3.32)$$

$$\sum_j W_{ij} = 1, j=1,2,\dots,m \quad (3.33)$$

Sonra Denklem 3.32’deki genelleştirilmiş öz değerler problemi çözülür:

$$AMA^T \alpha = \lambda AA^T \alpha \quad (3.34)$$

Denklem 3.34’ün çözümü olan en küçük değere sahip r tane öz değerlere karşılık gelen r tane öz vektör ile P dönüşüm matrisini oluşturulur [79].

3.2.7. Stokastik Yakınlık Gömme Analizi (SYGA)

Yakınlık verilerinden anlamlı boyutlar elde etmek için kendi kendini düzenleyen algoritma olan SYGA yöntemi kullanılmaktadır. SYGA yöntemi, bir dizi gözlem arasındaki benzerlikleri en iyi koruyan düşük boyutlu Öklid gömmeleri üretmektedir. SYGA yöntemi bir başlangıç konfigürasyonu ile başlar ve rastgele nesne çiftlerini tekrar tekrar seçerek ve haritadaki mesafelerinin ilgili yakınlıkları ile daha yakından eşleşmesi için koordinatlarını ayarlayarak yinelemeli olarak sentezler. Klasik çok boyutlu ölçeklendirme yönteminin ve doğrusal olmayan eşleme yönteminin aksine, SYGA yöntemi, örneklem boyutuna göre doğrusal olarak ölçeklenmektedir ve geleneksel gömme prosedürleriyle anlaşılmasın çok büyük veri kümelerine uygulanabilmektedir. SYGA yöntemi algoritma olarak basit, sağlam ve yakınsaktır. Ayrıca

keşifsel veri analizi ve görselleştirmeyi içeren çok çeşitli bilimsel sorunlara uygulanabilmektedir [80].

SYGA yönteminin analizi için aşağıdaki adımlar izlenir:

1) N tane noktanın D-boyutlu koordinatlarını sıfırlanır, $\{x_{ik}; i = 1, 2, \dots, N; k = 1, 2, \dots, D\}$. Bir kesme mesafesi r_c ve bir başlangıç öğrenme hızı $\lambda > 0$ seçilir.

2) Rastgele bir i noktası seçilir. Her $j \neq i$ noktası için, i giriş alanının r_{ij} 'ye olan yakınlığını öğrenin ve D-boyutlu haritadaki $d_{ij} = \|x_i - x_j\|$ öklid mesafesini hesaplanır. Eğer $r_{ij} \leq r_c$ ise veya $r_{ij} > r_c$ ve $d_{ij} < r_{ij}$ ise, x_j koordinatlarını Denklem 3.35 ile güncellenir.

$$x_j \leftarrow x_j + \lambda \frac{r_{ij} - d_{ij}}{d_{ij} + \varepsilon} (x_j - x_i) \quad (3.35)$$

ε , 0'a bölünmesini önlemek için kullanılan küçük bir sayıdır.

3) λ öğrenme oranı öngörülen bir $\delta\lambda$ ile azaltılır.

4) C çevrimi için öngörülen sayıda 2. ve 3. adım tekrar edilir [81].

3.3. Sınıflandırma İçin Kullanılan Yöntemler

Bu yüksek lisans tezinde aşağıda maddeler halinde yazılmış olan sınıflandırma yöntemleri kullanılmıştır.

3.3.1. Karar Ağaç (KA) Sınıflandırıcısı

Karar ağaç (KA) sınıflandırıcısı yöntemi literatürde genellikle kullanılan bir sınıflandırma yöntemi olmasının yanında aynı zamanda örüntü tanımlama algoritması olarak da belirtilmektedir. Bu yöntem ağaç oluşumunda kullanılan yöntemlerin anlaşılabilir ve sade olmasından dolayı etkin olarak kullanılmaktadır. Bu yöntem çok kademeli ve ardışık yaklaşımlarla oluşmaktadır. Çok karışık sınıflandırma problemlerini aşamalı bir şekilde basit bir karar verme işlemi gerçekleştirmektedir [82].

Bir karar ağacı üç kısımdan oluşmaktadır. Bunlar düğüm, yaprak ve dal olarak isimlendirilmektedir. Dallar ve yapraklar ağacın diğer elemanlarını temsil etmesine karşın, düğüm ağaç yapısında bir öz niteliği temsil eder. Bir ağaçta en ilk kısım yaprak, en son kısım ise köktür. Bu yapılar arasında kalan kısım ise daldır [83]. Ağaç; verileri içeren bir

kök düğümünden ve dallar ile yapraklardan oluşur. Eğitim verilerine ait nitelik bilgilerinden yararlanılarak bir karar ağacı yapısı oluşturulmasında oluşan prensip, verilere ait sorular sorulması ve elde edilen cevaplara göre hareket edilerek hızlı bir şekilde sonuçlanması olarak tanımlanabilir. Bu şekilde karar ağacı sorulara aldığı cevapları toplayarak karar kuralları oluşturur. Karar ağacı aldığı cevapları toplayarak karar kuralları oluşturur. Kök düğümünde verilerin sınıflandırılması ile sorular sorularak ağaç yapısı oluşturulur. Boşta bir dalları olmayan düğümler ve yaprak kalmayana kadar bu şekilde işlem devam etmektedir [82]. Ağaç oluşumunda kullanılan değişkenlerin sayısına bağlı olarak bir değişkenli veya çok değişkenli karar ağaç yapıları bulunabilmektedir [84]. Tek değişkenli karar ağacının her bir iç düğümün de tek bir özneliğin test edilmesi ile birden fazla alt kümeye ayrılması durumu olabilir [85]. Bu şekilde karar ağaç tekrarlı bir şekilde giriş verisine ait bir yaprak düğümüne ulaşıncaya kadar ve yapraktaki sınıf etiketi gözlemlere atanıncaya kadar bölünme devam eder. Tek değişkenli bir karar ağacında karar sınırlarının özel değerleri eğitim verisinden deneysel olarak tahmin edilmektedir. Sürekli verilerin olması durumunda, eğitim verisinden her bir iç düğüm noktasında öz nitelik değeri olan a_i 'nin veri uzayındaki bir özelliği gösterdiği ve b 'nin a_i 'nin gözlenen aralığında bir eşik değeri olduğu $a_i > b$ şeklindeki mantıksal değerlendirme gerçekleştirilir. Farklılıkların maksimum hale getirilmesi veya iniş düğümlerindeki benzerliğin düşük hale getirilmesi gibi bazı koşullar kullanılarak eşik değeri b belirlenebilir [86]. Bir veri setinin B_1, B_2, \dots, B_n şeklinde birkaç sınıftan oluştuğu ve F 'nin sınıf değerlerini gösterdiği düşünülürse, bir sınıfa ait ihtimal $M_i = (B_i/|F|)$ olur ve sınıflara ait entropi Denklem 3.36'da;

$$\text{Entropi}(F) = -\sum_{i=1}^n M_i \log_2(M_i) \quad (3.36)$$

şeklinde hesaplanır. Veri setindeki H özneliğine göre F sınıf değerleri F_1, F_2, \dots, F_n şeklinde alt kümelere ayrıldığı göz önüne alınsın. H öznelik değerleri kullanılarak F sınıf değerlerinin bölünmesi sonucunda elde edilecek kazanç, Denklem 3.37'de,

$$\text{Kazanç}(H, F) = \text{Entropi}(F) - \sum_{i=1}^n \frac{|F_i|}{|F|} \text{Entropi}(F_i) \quad (3.37)$$

eşitliği ile hesaplanır. F kümesi için H özneliğinin değerini belirlenmesinde Denklem 3.38'de bölümlenme bilgisi kullanılır [86].

$$\text{Bölünme Bilgisi (H)} = - \sum_{i=1}^n \frac{|F_i|}{|F|} \log_2 \frac{|F_i|}{|F|} \quad (3.38)$$

Bu durumda kazanç oranı Denklem 3.39'da,

$$\text{Kazanç Oranı} = \frac{\text{Kazanç (H,F)}}{\text{Bölünme Bilgisi (H)}} \quad (3.39)$$

şeklinde hesaplanır. Bu eşitlik, sınıflandırma işleminde kullanılacak ayırma ile elde edilen bilgi oranını verir. Bu kıstas kullanılarak, ağacın her bir düğümünde kazanç oranı maksimum olacak şekilde F eğitim kümesi tekrarlı bir şekilde ayrılmaktadır. Bu işleme her bir yaprak düğümü sadece bir sınıfa ait gözlem değerleri içerene kadar tekrar edilir. Eğer tek bir yaprağı olan alt ağacın veya bu ağacın en çok kullanılan dalının budanması beklenen hata oranını düşürecekse ağaç budanır. Alt dallardaki hata oranı azaldığından tüm ağaç için hata oranı azalacaktır. Budama işlemi sonunda hata oranının minimum hale getirildiği bir ağaç elde edilmektedir [86].

3.3.2. Rastgele Orman (RO) Sınıflandırıcısı

RO, ağaç tipi sınıflandırıcılardan oluşmaktadır [87]. RO, günümüzde ki algoritmalar arasında doğruluğu en yüksek sınıflandırıcıdır. Ayrıca hızlı ve belirli kriterleri olmayan bir yöntemdir [88]. RO sınıflandırıcıda ilk olarak gerçek veri setinden eğitim veri seti elde edilir. Sonra oluşturulan eğitim setinden bir ağaç oluşturulmaktadır [89, 87]. RO sınıflandırıcısı ile ağaç üretebilmek için dışarıdan iki tane parametre girilmelidir. Bu parametreler en iyi çoğalmayı sağlayabilmek için düğümde kullanılacak değişkenlerin sayısı m , geliştirilecek ağaç sayısı N 'dir. m değeri hatalara göre artırılıp azaltılabilen ve rastgele girilen bir değerdir. m değerinin artırılıp azaltılmasıyla en iyi m değeri bulunur ve sınıflandırma duyarlılığı artmaktadır [90].

T eğitim veri setinden T_K yer değiştirmeli eğitim veri seti elde edilmektedir. Yeni eğitim veri setinden $h(x, T_K)$ sınıflandırıcısı elde edilir [87].

RO en büyük ölçüde ağaç oluşturabilmek için bir düğümde belirli kriterlerde bölünme işlemi gerçekleştirir [91]. Bu düğümler bölünme için homojen sınıf dağılımına sahip olmalıdırlar. Düğümlerin homojenliğini ölçmek için bu sınıflandırıcı tarafından Gini indeksi kullanılmaktadır. Verilen T eğitim veri seti için C_i sınıfına ait bir örnek seçilerek Gini indeksi Denklem 3.40'da gösterilmektedir:

$$\sum \sum_{j \neq i} (f(C_i, T) \div |T|) * (f(C_j, T) \div |T|) \quad (3.40)$$

$(f(C_i, T) \div |T|)$ ifadesi seçilmiş olan örneğin C_i sınıfına ait olma ihtimalini göstermektedir [92]. Gini indeksi değeri düştükçe homojenlik artmaktadır. Bir düğümün başarılı olma şartı bir üst düğümden Gini indeks değerinin düşük olmasıyla belirlenir. Gini indeks değeri sıfır olunca da ağaç dallanma işlemi bitmektedir [93]. Ağaç gelişimi bittikten sonra da girdi verisi sınıflandırmaya başlanır. Bu sınıflandırma işleminde RO sınıflandırıcı tarafından girdi verisi, ormandaki bütün ağaçlara yerleştirilir. Sonra ağaçlar arasında bir oylama yapılmaktadır. Oylama en yüksek oyu alan ağaç sınıfa atanmaktadır [94].

3.3.3. Naive Bayes (NB) Sınıflandırıcısı

Sınıflandırma, makine öğrenmesinde ve veri madenciliğinde temel bir konudur. Sınıflandırmada bir algoritma öğrenmenin amacı, sınıf etiketli bir dizi veri örneği hazırlayan bir sınıflandırıcı oluşturmaktır. Bir öznitelik değerleri dizisiyle temsil edilen (a_1, a_2, \dots, a_n) F bir örnektir ve a_i, A_i özelliğinin değeridir. B, sınıflandırma değişkeni olsun ve B'nin değeri b olsun. Pozitif ve negatif sınıf olmak üzere iki sınıf vardır. Örneğe sınıf etiketi atanan değere sınıflandırıcı denir [95]. Olasılık olarak Bayes kuralına göre, bir örnekleme olasılığı $F = (a_1, a_2, \dots, a_n)$, b sınıfı olmaktadır. Bu da Denklem 3.41'de gösterilmektedir.

$$p(b | F) = \frac{p(F | b) p(b)}{p(F)} \quad (3.41)$$

F, B = + olarak sınıflandırılırsa ve eğer sadece

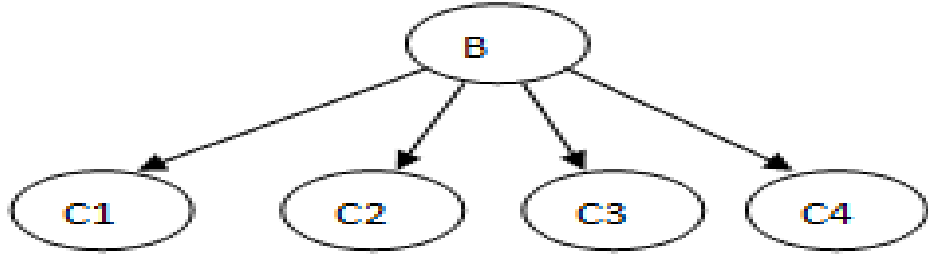
$$f_b(F) = \frac{p(B = + | F)}{p(B = - | F)} \geq 1 \quad (3.42)$$

ise buradaki $f_b(F)$, Bayesian sınıflandırıcı olarak Denklem 3.42'de adlandırılır. Sınıf değişkeninin değer verilen tüm özelliklerinin bağımsız olduğunu düşünelim; yani,

$$p(F | b) = p(a_1, a_2, \dots, a_n | b) = \prod_{i=1}^n p(a_i/b) \quad (3.43)$$

Denklem 3.43'ün sonucunda ortaya çıkan sınıflandırıcı Denklem 3.44'te elde edilir:

$$f_{nb}(F) = \frac{p(B = +)}{p(B = -)} \prod_{i=1}^n \frac{p(a_i | B = +)}{p(a_i | B = -)} \quad (3.44)$$

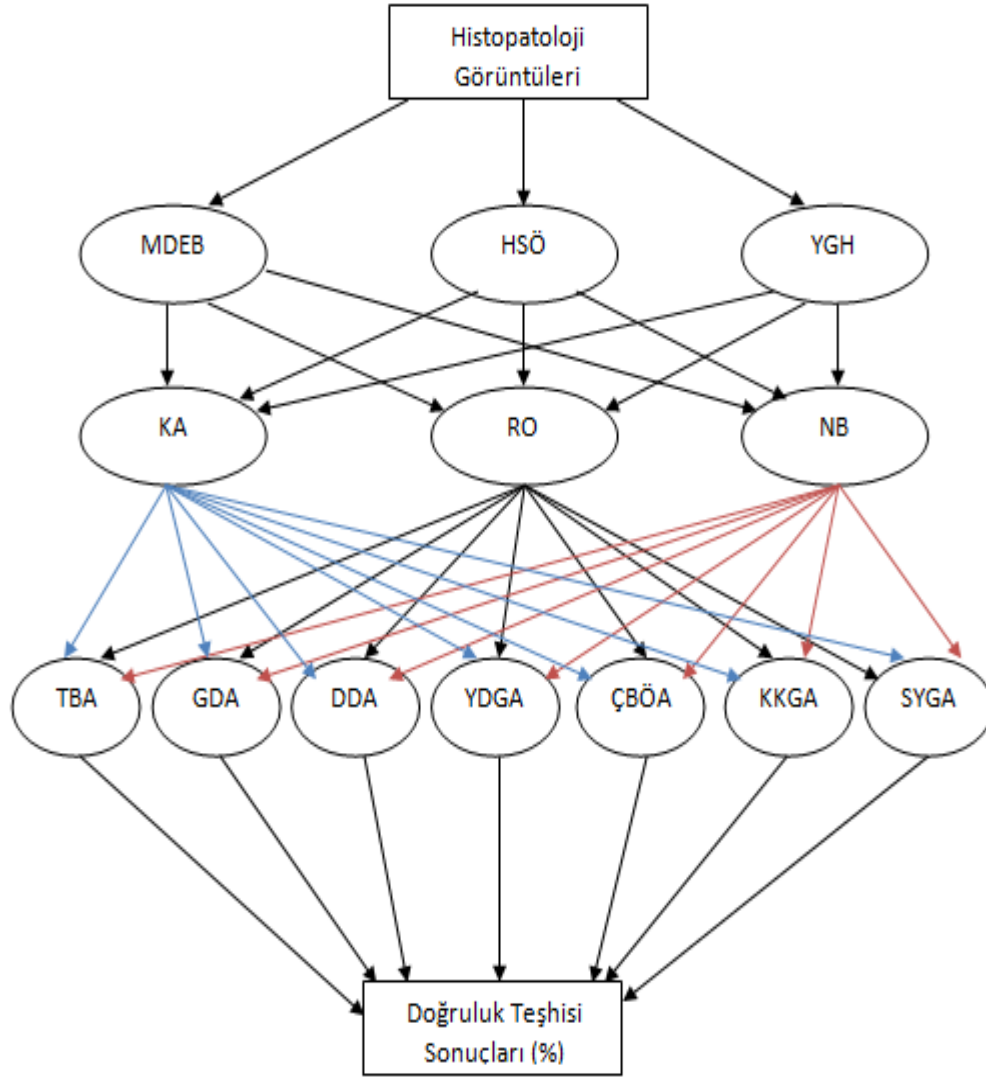


Şekil 3.1. Naive Bayes sınıflandırıcı örneği [95].

f_{nb} (F) fonksiyonuna Naive Bayes (NB) sınıflandırıcı da denilmektedir. Şekil 3.1’de Naive Bayes örneği gösterilmiştir. Naive Bayes’te, her bir öznelik düğümünün sınıf düğümü dışında bir atası yoktur. Naive Bayes, Bayesian ağının en sade şeklidir. Çünkü bütün özelliklerin bağımsız olduğu sınıf değişkenine sahiptir [95].

4. AKCİĞER KANSERİ TEŞHİSİ İÇİN GELİŞTİRİLEN YÖNTEMLERİN ANALİZ BULGULARI

Akciğer kanserinin erken tanısı için bazı yöntemler geliştirilmiştir. Bu yöntemlerin oluşum aşamalarını göstermek için Şekil 4.1 kullanılmıştır.

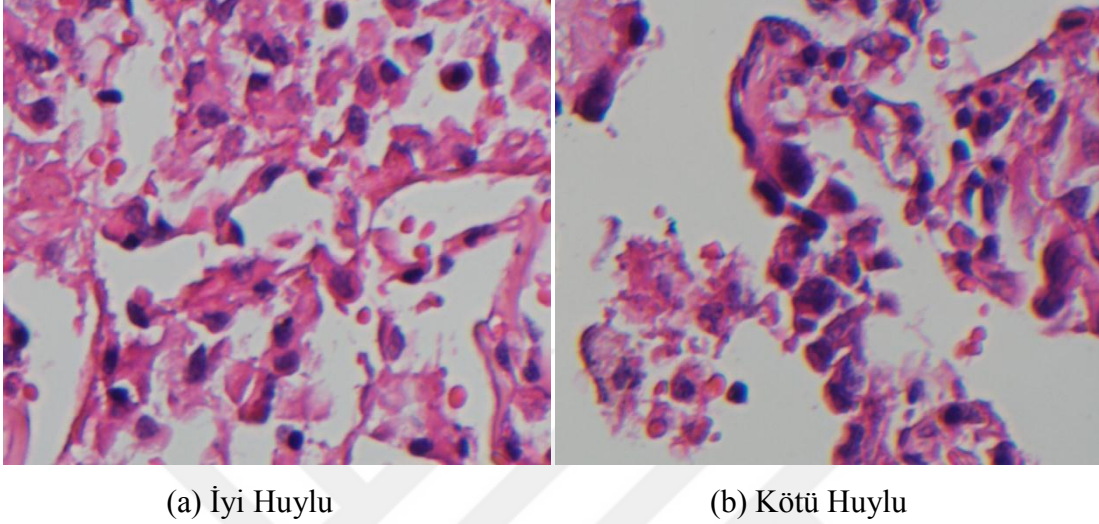


Şekil 4.1. Akciğer kanseri teşhisi için geliştirilen yöntemlerin akış şeması

Şekil 4.1 incelendiğinde bir sınıflandırma yöntemi, bir boyut azaltma yöntemi, bir özellik ekleme için kullanılan yöntem beraber kullanılarak analiz gerçekleştirilmiştir. Analizin sonucunda doğruluk teşhis oranı bulunmuştur.

Bu çalışmada 120 adet iyi huylu ve 120 adet kötü huylu tümörlere ait olmak üzere toplamda 240 adet ışık mikroskobundan elde edilen histopatoloji görüntüsü kullanılmıştır.

Şekil 4.2 incelendiğinde iyi huylu tümör ile kötü huylu tümör arasındaki fark histopatoloji görüntülerinde görülebilmektedir. İyi huylu tümördeki hücreler daha düzgün ve dağılmamış bir şekilde görülmektedir. Kötü huylu tümörde ise hücrelerin dağıldığı ve oval şeklinin bozulduğu görülmektedir.



Şekil 4.2. Akciğere ait tümörün histopatoloji görüntü örnekleri

240 adet akciğer histopatoloji görüntüsünün 120 adedi eğitim amaçlı, 120 adedi test amaçlı kullanılmıştır. Yani iyi huylu 60 adet görüntü, kötü huylu 60 adet görüntü toplamda 120 adet görüntü yöntemlerde eğitim amaçlı kullanılmıştır. 60 adet iyi huylu görüntü, 60 adet kötü huylu görüntü toplamda 120 adet görüntü yöntemlerde test amaçlı kullanılmıştır.

4.1. Histopatoloji Görüntüleri İçin HSÖ – Boyut İndirgeme - NB Yöntemi

Bu yöntem için özellik ekleme yöntemi olan HSÖ, boyut indirgeme yöntemleri olan TBA-GDA-DDA-YDGA-ÇBÖA-KKGA-SYGA yöntemleri, sınıflandırıcı olarak NB kullanılarak elde edilen yöntem kullanılmıştır. Bu yöntem sonucunda elde edilen sonuçlar Tablo 4.1’de gösterilmiştir. Tablo 4.1 incelendiğinde en yüksek doğruluk sonucunu veren yöntemin doğruluk oranı koyu renkte yazılmıştır. Ayrıca doğruluk teşhisi sonuçlarının bulunması için kullanılan özellik sayısı da görülmektedir.

Tablo 4.1. HSÖ-Boyut İndirgeme-NB yöntemi doğruluk teşhisi sonuçları

Özellik Sayısı	Kullanılan Boyut Azaltma Yöntemleri Ve Doğruluk Oranı (%)						
	TBA (%)	GDA (%)	DDA (%)	YDGA (%)	ÇBÖA (%)	KKGA (%)	SYGA (%)
5	85,833	100,000	79,166	90,000	85,833	79,166	86,666
10	75,833	100,000	78,333	85,833	75,833	89,166	86,666
15	77,500	100,000	77,500	85,000	77,500	90,000	88,333
20	76,666	100,000	75,833	87,500	76,666	90,000	89,166
25	79,166	100,000	75,833	86,666	79,166	90,000	85,000
30	74,166	100,000	70,833	85,833	74,166	85,833	85,000
35	75,000	100,000	73,333	83,333	75,000	80,833	83,333
40	77,500	100,000	75,833	81,666	77,500	77,500	86,666
45	75,833	100,000	75,833	78,333	75,833	73,333	82,500
50	75,833	100,000	73,333	75,833	75,833	71,666	85,833

Tablo 4.1'e göre en yüksek doğruluk oranını kullanılan boyut indirgeme yöntemleri arasında bulunan GDA boyut indirgeme yöntemi vermiştir. GDA boyut indirgeme yöntemi ile %100,000 oranında bir doğruluk oranı bulunmuştur. Ayrıca GDA yöntemi bütün kullanılan özellik sayısı değerlerinde de %100,000 doğruluk oranını vermiştir. Bu da HSÖ-GDA-NB yöntemi ile kişilerin akciğer kanseri tümörünü tanımda kullanılan histopatoloji görüntülerinde %100,000 oranında doğruluk bulunduğunu göstermektedir.

4.2. Histopatoloji Görüntüleri İçin HSÖ – Boyut İndirgeme - RO Yöntemi

Bu yöntem için özellik ekleme yöntemi olan HSÖ, boyut indirgeme yöntemleri olan TBA-GDA-DDA-YDGA-ÇBÖA-KKGA-SYGA yöntemleri, sınıflandırıcı olarak RO kullanılarak elde edilen yöntem kullanılmıştır. Bu yöntem sonucunda elde edilen sonuçlar Tablo 4.2'de gösterilmiştir. Tablo 4.2 incelendiğinde en yüksek doğruluk sonucunu veren yöntemin doğruluk oranı koyu renkte yazılmıştır. Ayrıca doğruluk teşhisi sonuçlarının bulunması için kullanılan özellik sayısı da görülmektedir.

Tablo 4.2. HSÖ-Boyut İndirgeme-RO yöntemi doğruluk teşhisi sonuçları

Özellik Sayısı	Kullanılan Boyut Azaltma Yöntemleri Ve Doğruluk Oranı (%)						
	TBA (%)	GDA (%)	DDA (%)	YDGA (%)	ÇBÖA (%)	KKGA (%)	SYGA (%)
5	95,000	93,333	90,833	86,666	95,000	85,833	88,333
10	92,500	91,666	91,666	90,833	92,500	87,500	90,833
15	89,166	91,666	89,166	92,500	89,166	87,500	90,833
20	88,333	91,666	89,166	92,500	88,333	88,333	89,166
25	89,166	91,666	88,333	91,166	89,166	88,333	88,333
30	87,500	91,666	88,333	91,166	87,500	85,833	89,166
35	85,000	91,666	86,666	91,166	85,000	83,333	90,833
40	84,166	91,666	85,833	90,000	84,166	82,500	95,000
45	86,666	91,666	85,000	88,333	86,666	82,500	90,000
50	86,666	99,166	85,833	89,166	86,666	83,333	90,000

Tablo 4.2'e göre en yüksek doğruluk oranını kullanılan boyut indirgeme yöntemleri arasında bulunan GDA boyut indirgeme yöntemi vermiştir. GDA boyut indirgeme yöntemi ile %99,166 oranında bir doğruluk oranı bulunmuştur. Ayrıca GDA yöntemi 50 özellik sayısı değerinde %99,166 doğruluk oranını vermiştir. Bu da HSÖ-GDA-RO yöntemi ile kişilerin akciğer kanseri tümörünü tanımda kullanılan histopatoloji görüntülerinde %99,166 oranında doğruluk bulunduğunu göstermektedir.

4.3. Histopatoloji Görüntüleri İçin HSÖ – Boyut İndirgeme - KA Yöntemi

Bu yöntem için özellik ekleme yöntemi olan HSÖ, boyut indirgeme yöntemleri olan TBA-GDA-DDA-YDGA-ÇBÖA-KKGA-SYGA yöntemleri, sınıflandırıcı olarak KA kullanılarak elde edilen yöntem kullanılmıştır. Bu yöntem sonucunda elde edilen sonuçlar Tablo 4.3'de gösterilmiştir. Tablo 4.3 incelendiğinde en yüksek doğruluk sonucunu veren yöntemin doğruluk oranı koyu renkte yazılmıştır. Ayrıca doğruluk teşhisi sonuçlarının bulunması için kullanılan özellik sayısı da görülmektedir.

Tablo 4.3. HSÖ-Boyut İndirgeme-KA yöntemi doğruluk teşhisi sonuçları

Özellik Sayısı	Kullanılan Boyut Azaltma Yöntemleri Ve Doğruluk Oranı (%)						
	TBA (%)	GDA (%)	DDA (%)	YDGA (%)	ÇBÖA (%)	KKGA (%)	SYGA (%)
5	89,166	91,666	90,000	79,166	89,166	76,166	75,833
10	89,166	91,666	90,000	87,500	89,166	80,833	87,500
15	85,000	91,666	85,000	84,166	85,000	80,833	82,500
20	85,000	91,666	85,000	82,500	85,000	80,833	79,166
25	87,500	91,666	85,000	80,000	87,500	80,833	75,000
30	87,500	91,666	85,000	80,000	87,500	80,833	74,166
35	87,500	91,666	85,000	80,000	87,500	80,833	78,333
40	82,500	91,666	85,000	80,000	82,500	80,833	78,333
45	83,333	91,666	85,000	80,000	83,333	80,833	81,666
50	83,333	91,666	85,000	79,166	83,333	80,833	72,500

Tablo 4.3'e göre en yüksek doğruluk oranını kullanılan boyut indirgeme yöntemleri arasında bulunan GDA boyut indirgeme yöntemi vermiştir. GDA boyut indirgeme yöntemi ile %91,666 oranında bir doğruluk oranı bulunmuştur. Ayrıca GDA yöntemi tüm özellik sayısı değerinde %91,666 doğruluk oranını vermiştir. Bu da HSÖ-GDA-KA yöntemi ile kişilerin akciğer kanseri tümörünü tanımda kullanılan histopatoloji görüntülerinde %91,666 oranında doğruluk bulunduğunu göstermektedir.

4.4. Histopatoloji Görüntüleri İçin MDEB – Boyut İndirgeme - NB Yöntemi

Bu yöntem için özellik ekleme yöntemi olan MDEB, boyut indirgeme yöntemleri olan TBA-GDA-DDA-YDGA-ÇBÖA-KKGA-SYGA yöntemleri, sınıflandırıcı olarak NB kullanılarak elde edilen yöntem kullanılmıştır. Bu yöntem sonucunda elde edilen sonuçlar Tablo 4.4'de gösterilmiştir. Tablo 4.4 incelendiğinde en yüksek doğruluk sonucunu veren yöntemin doğruluk oranı koyu renkte yazılmıştır. Ayrıca doğruluk teşhisi sonuçlarının bulunması için kullanılan özellik sayısı da görülmektedir.

Tablo 4.4. MDEB-Boyut İndirgeme-NB yöntemi doğruluk teşhisi sonuçları

Özellik Sayısı	Kullanılan Boyut Azaltma Yöntemleri Ve Doğruluk Oranı (%)						
	TBA (%)	GDA (%)	DDA (%)	YDGA (%)	ÇBÖA (%)	KKGA (%)	SYGA (%)
5	67,500	100,000	73,333	61,666	67,500	69,166	75,000
10	66,666	100,000	65,833	69,166	66,666	80,000	80,833
15	70,833	100,000	63,333	77,500	70,833	79,166	71,666
20	75,000	100,000	69,166	83,333	75,000	78,333	85,833
25	76,666	100,000	71,666	80,000	76,666	81,666	81,666
30	75,000	100,000	72,500	79,166	75,000	84,166	81,666
35	77,500	100,000	72,500	76,166	77,500	84,166	76,666
40	79,166	100,000	76,666	75,000	79,166	86,666	85,833
45	80,833	100,000	75,833	72,500	80,833	83,333	77,500
50	82,500	100,000	77,500	69,166	82,500	81,666	81,666

Tablo 4.4'e göre en yüksek doğruluk oranını kullanılan boyut indirgeme yöntemleri arasında bulunan GDA boyut indirgeme yöntemi vermiştir. GDA boyut indirgeme yöntemi ile %100,000 oranında bir doğruluk oranı bulunmuştur. Ayrıca GDA yöntemi tüm özellik sayısı değerinde %100,000 doğruluk oranını vermiştir. Bu da MDEB-GDA-NB yöntemi ile kişilerin akciğer kanseri tümörünü tanımda kullanılan histopatoloji görüntülerinde %100,000 oranında doğruluk bulunduğunu göstermektedir.

4.5. Histopatoloji Görüntüleri İçin MDEB –Boyut İndirgeme - RO Yöntemi

Bu yöntem için özellik ekleme yöntemi olan MDEB, boyut indirgeme yöntemleri olan TBA-GDA-DDA-YDGA-ÇBÖA-KKGA-SYGA yöntemleri, sınıflandırıcı olarak RO kullanılarak elde edilen yöntem kullanılmıştır. Bu yöntem sonucunda elde edilen sonuçlar Tablo 4.5'de gösterilmiştir. Tablo 4.5 incelendiğinde en yüksek doğruluk sonucunu veren yöntemin doğruluk oranı koyu renkte yazılmıştır. Ayrıca doğruluk teşhisi sonuçlarının bulunması için kullanılan özellik sayısı da görülmektedir.

Tablo 4.5. MDEB-Boyut İndirgeme-RO yöntemi doğruluk teşhisi sonuçları

Özellik Sayısı	Kullanılan Boyut Azaltma Yöntemleri Ve Doğruluk Oranı (%)						
	TBA (%)	GDA (%)	DDA (%)	YDGA (%)	ÇBÖA (%)	KKGA (%)	SYGA (%)
5	95,833	96,666	92,500	71,666	95,833	89,166	91,666
10	95,833	96,666	92,500	80,000	95,833	84,666	83,333
15	93,333	97,500	85,000	84,166	83,333	83,333	91,666
20	92,500	97,500	85,000	80,833	92,500	88,333	86,666
25	88,333	97,500	83,333	82,500	88,333	85,833	93,333
30	88,333	97,500	84,166	82,500	88,333	84,166	92,500
35	84,166	97,500	81,666	79,166	84,166	80,833	84,166
40	85,000	97,500	86,666	79,166	85,000	83,333	81,166
45	84,166	97,500	84,166	80,833	84,166	82,500	86,666
50	82,500	97,500	82,500	79,166	82,500	78,333	85,000

Tablo 4.5'e göre en yüksek doğruluk oranını kullanılan boyut indirgeme yöntemleri arasında bulunan GDA boyut indirgeme yöntemi vermiştir. GDA boyut indirgeme yöntemi ile %97,500 oranında bir doğruluk oranı bulunmuştur. Ayrıca GDA yöntemi 15-20-25-30-35-40-45-50 özellik sayıları değerinde %97,500 doğruluk oranını vermiştir. Bu da MDEB-GDA-RO yöntemi ile kişilerin akciğer kanseri tümörünü tanımda kullanılan histopatoloji görüntülerinde %97,500 oranında doğruluk bulunduğunu göstermektedir.

4.6. Histopatoloji Görüntüleri İçin MDEB –Boyut İndirgeme-KA Yöntemi

Bu yöntem için özellik ekleme yöntemi olan MDEB, boyut indirgeme yöntemleri olan TBA-GDA-DDA-YDGA-ÇBÖA-KKGA-SYGA yöntemleri, sınıflandırıcı olarak KA kullanılarak elde edilen yöntem kullanılmıştır. Bu yöntem sonucunda elde edilen sonuçlar Tablo 4.6'da gösterilmiştir. Tablo 4.6 incelendiğinde en yüksek doğruluk sonucunu veren yöntemin doğruluk oranı koyu renkte yazılmıştır. Ayrıca doğruluk teşhisi sonuçlarının bulunması için kullanılan özellik sayısı da görülmektedir.

Tablo 4.6. MDEB-Boyut İndirgeme-KA yöntemi doğruluk teşhisi sonuçları

Özellik Sayısı	Kullanılan Boyut Azaltma Yöntemleri Ve Doğruluk Oranı (%)						
	TBA (%)	GDA (%)	DDA (%)	YDGA (%)	ÇBÖA (%)	KKGA (%)	SYGA (%)
5	93,333	69,166	86,666	63,333	93,333	80,000	76,666
10	91,666	69,166	86,666	72,500	91,666	73,333	82,500
15	92,500	87,500	85,833	73,333	92,500	73,333	77,500
20	92,500	87,500	86,666	67,500	92,500	71,666	75,000
25	92,500	87,500	86,666	73,333	92,500	72,500	70,833
30	92,500	87,500	86,666	70,833	92,500	70,000	66,666
35	92,500	87,500	86,666	70,000	92,500	70,000	64,166
40	92,500	87,500	78,333	72,500	92,500	70,000	70,833
45	92,500	87,500	78,333	72,500	92,500	70,000	73,333
50	92,500	87,500	78,333	80,000	92,500	70,000	72,500

Tablo 4.6'ya göre en yüksek doğruluk oranını kullanılan boyut indirgeme yöntemleri arasında bulunan TBA-ÇBÖA boyut indirgeme yöntemlerinde vermiştir. TBA-ÇBÖA boyut indirgeme yöntemleri ile %93,333 oranında bir doğruluk oranı bulunmuştur. Ayrıca TBA-ÇBÖA yöntemleri 5 özellik sayısı değerinde %93,333 doğruluk oranını vermiştir. Bu da MDEB-TBA-KA yöntemi ve MDEB-ÇBÖA-KA ile kişilerin akciğer kanseri tümörünü tanımda kullanılan histopatoloji görüntülerinde %93,333 oranında doğruluk bulunduğunu göstermektedir.

4.7. Histopatoloji Görüntüleri İçin YGH – Boyut İndirgeme - NB Yöntemi

Bu yöntem için özellik ekleme yöntemi olan YGH, boyut indirgeme yöntemleri olan TBA-GDA-DDA-YDGA-ÇBÖA-KKGA-SYGA yöntemleri, sınıflandırıcı olarak NB kullanılarak elde edilen yöntem kullanılmıştır. Bu yöntem sonucunda elde edilen sonuçlar Tablo 4.7'de gösterilmiştir. Tablo 4.7 incelendiğinde en yüksek doğruluk sonucunu veren yöntemin doğruluk oranı koyu renkte yazılmıştır. Ayrıca doğruluk teşhisi sonuçlarının bulunması için kullanılan özellik sayısı da görülmektedir.

Tablo 4.7. YGH-Boyut İndirgeme-NB yöntemi doğruluk teşhisi sonuçları

Özellik Sayısı	Kullanılan Boyut Azaltma Yöntemleri Ve Doğruluk Oranı (%)						
	TBA (%)	GDA (%)	DDA (%)	YDGA (%)	ÇBÖA (%)	KKGA (%)	SYGA (%)
5	70,833	48,333	84,166	72,500	70,833	75,000	76,666
10	69,166	48,333	82,500	74,166	70,833	75,000	81,666
15	70,000	48,333	86,666	69,166	70,000	67,500	85,000
20	71,666	48,333	78,333	68,333	71,666	69,166	80,833
25	68,333	48,333	80,000	66,666	68,333	70,000	84,166
30	69,166	48,333	81,666	65,833	69,166	71,666	80,833
35	64,166	48,333	82,500	65,000	64,166	68,333	81,666
40	66,666	48,333	80,833	62,500	66,666	69,166	81,666
45	60,833	48,333	80,833	63,333	60,833	68,333	84,166
50	64,166	48,333	79,166	52,500	64,166	67,500	82,500

Tablo 4.7'ye göre en yüksek doğruluk oranını kullanılan boyut indirgeme yöntemleri arasında bulunan DDA boyut indirgeme yöntemi vermiştir. DDA boyut indirgeme yöntemi ile %86,667 oranında bir doğruluk oranı bulunmuştur. Ayrıca DDA yöntemi 15 özellik sayısı değerinde %86,667 doğruluk oranını vermiştir. Bu da YGH-DDA-NB yöntemi ile kişilerin akciğer kanseri tümörünü tanımda kullanılan histopatoloji görüntülerinde %86,667 oranında doğruluk bulunduğunu göstermektedir.

4.8. Histopatoloji Görüntüleri İçin YGH – Boyut İndirgeme - RO Yöntemi

Bu yöntem için özellik ekleme yöntemi olan YGH, boyut indirgeme yöntemleri olan TBA-GDA-DDA-YDGA-ÇBÖA-KKGA-SYGA yöntemleri, sınıflandırıcı olarak RO kullanılarak elde edilen yöntem kullanılmıştır. Bu yöntem sonucunda elde edilen sonuçlar Tablo 4.8'de gösterilmiştir. Tablo 4.8 incelendiğinde en yüksek doğruluk sonucunu veren yöntemin doğruluk oranı koyu renkte yazılmıştır. Ayrıca doğruluk teşhisi sonuçlarının bulunması için kullanılan özellik sayısı da görülmektedir.

Tablo 4.8. YGH-Boyut İndirgeme-RO yöntemi doğruluk teşhisi sonuçları

Özellik Sayısı	Kullanılan Boyut Azaltma Yöntemleri Ve Doğruluk Oranı (%)						
	TBA (%)	GDA (%)	DDA (%)	YDGA (%)	ÇBÖA (%)	KKGA (%)	SYGA (%)
5	70,833	42,500	70,833	77,500	70,833	75,000	75,000
10	71,666	42,500	70,000	76,666	71,666	75,833	74,166
15	74,166	42,500	76,666	76,666	74,166	76,666	71,666
20	73,333	42,500	75,000	71,666	73,333	74,166	77,500
25	73,333	42,500	75,833	73,333	73,333	74,166	75,833
30	74,166	42,500	75,833	69,166	74,166	73,333	80,833
35	73,333	42,500	75,833	68,333	73,333	71,666	80,833
40	70,000	42,500	75,833	65,000	70,000	70,833	73,333
45	74,166	42,500	75,000	66,666	74,166	72,500	75,000
50	70,000	42,500	71,666	64,166	70,000	73,333	80,000

Tablo 4.8'e göre en yüksek doğruluk oranını kullanılan boyut indirgeme yöntemleri arasında bulunan SYGA boyut indirgeme yöntemi vermiştir. SYGA boyut indirgeme yöntemi ile %80,833 oranında bir doğruluk oranı bulunmuştur. Ayrıca SYGA yöntemi 30-35 özellik sayıları değerlerinde %80,833 doğruluk oranını vermiştir. Bu da YGH- SYGA - NB yöntemi ile kişilerin akciğer kanseri tümörünü tanımda kullanılan histopatoloji görüntülerinde %80,833 oranında doğruluk bulunduğunu göstermektedir.

4.9. Histopatoloji Görüntüleri İçin YGH – Boyut İndirgeme - KA Yöntemi

Bu yöntem için özellik ekleme yöntemi olan YGH, boyut indirgeme yöntemleri olan TBA-GDA-DDA-YDGA-ÇBÖA-KKGA-SYGA yöntemleri, sınıflandırıcı olarak KA kullanılarak elde edilen yöntem kullanılmıştır. Bu yöntem sonucunda elde edilen sonuçlar Tablo 4.9'da gösterilmiştir. Tablo 4.9 incelendiğinde en yüksek doğruluk sonucunu veren yöntemin doğruluk oranı koyu renkte yazılmıştır. Ayrıca doğruluk teşhisi sonuçlarının bulunması için kullanılan özellik sayısı da görülmektedir.

Tablo 4.9. YGH-Boyut İndirgeme-KA yöntemi doğruluk teşhisi sonuçları

Özellik Sayısı	Kullanılan Boyut Azaltma Yöntemleri Ve Doğruluk Oranı (%)						
	TBA (%)	GDA (%)	DDA (%)	YDGA (%)	ÇBÖA (%)	KKGA (%)	SYGA (%)
5	70,000	47,500	71,666	68,333	70,000	80,000	71,666
10	70,000	47,500	68,333	63,333	70,000	82,500	69,166
15	60,000	47,500	77,500	60,833	60,000	79,166	74,166
20	60,000	47,500	76,666	58,333	60,000	77,500	72,500
25	60,000	47,500	74,166	58,333	60,000	77,500	70,000
30	60,000	47,500	70,833	58,333	60,000	75,000	68,333
35	60,000	47,500	70,000	60,000	60,000	75,000	65,000
40	62,500	47,500	67,500	60,000	62,500	75,000	61,666
45	62,500	47,500	63,333	59,166	62,500	75,000	77,500
50	62,500	47,500	67,500	59,166	62,500	75,000	70,000

Tablo 4.9'a göre en yüksek doğruluk oranını kullanılan boyut indirgeme yöntemleri arasında DDA-KKGA-SYGA boyut indirgeme yöntemleri vermiştir. DDA-KKGA-SYGA boyut indirgeme yöntemleri ile %77,500 oranında bir doğruluk oranı bulunmuştur. Ayrıca DDA yöntemi 15 özellik sayısı değerinde, KKGA yöntemi ile 20-25 özellik sayıları değerinde, SYGA yöntemi ile 45 özellik sayısı değerinde %77,500 doğruluk oranını vermiştir. Bu da YGH- DDA –KA, YGH- KKGA –KA, YGH- SYGA- KA yöntemleri ile kişilerin akciğer kanseri tümörünü tanımda kullanılan histopatoloji görüntülerinde %77,500 oranında doğruluk bulunduğunu göstermektedir.

4.10. Histopatoloji Görüntülerinin Doğruluk Teşhisi Oranlarının Kıyaslanması

Akciğer tümörü histopatoloji görüntüleri için bulunan doğruluk teşhis oranları Tablo 4.10'da bulunmaktadır. Tablo 4.10'a göre HSÖ-GDA-NB yöntemi için %100,000, HSÖ-GDA-RO yöntemi için %99,166, HSÖ-GDA-KA için %91,666, MDEB-GDA-NB yöntemi için %100,000, MDEB-GDA-RO yöntemi için %97,500, MDEB-TBA-KA yöntemi için %93,333, MDEB-ÇBÖA-KA yöntemi için %93,333, YGH-DDA-NB yöntemi için %86,667, YGH-SYGA-NB yöntemi için % 80,833, YGH-DDA-KA yöntemi için

%77,500, YGH-KKGA-KA yöntemi için %77,500, YGH-SYGA-KA yöntemi için %77,500 oranlarında doğruluk teşhis oranları bulunmuştur.

Tablo 4.10. Histopatoloji görüntülerinin doğruluk teşhisi oranlarının kıyaslanması için yöntemlerin tek başlarına doğruluk teşhis oranları

	Doğruluk Teşhisi Oranı (%)
Yöntemler	%
HSÖ-GDA-NB	100,000
HSÖ-GDA-RO	99,166
HSÖ-GDA-KA	91,666
MDEB-GDA-NB	100,000
MDEB-GDA-RO	97,500
MDEB-TBA-KA	93,333
MDEB-ÇBÖA-KA	93,333
YGH-DDA-NB	86,667
YGH- SYGA -NB	80,833
YGH- DDA –KA	77,500
YGH- KKGA –KA	77,500
YGH- SYGA- KA	77,500

5. SONUÇLAR VE ÖNERİLER

Akciğer kanserinin erken teşhisi için bilgisayar destekli tanı sistemlerinin kullanılması oldukça önemlidir. Çünkü akciğer kanserinin tespiti için bilgisayar destekli tanı sistemi kullanıldığı zaman daha yüksek doğrulukta ve daha yüksek hızda sonuçlar alınabilmektedir.

Bu çalışmada akciğer tümörünün histopatoloji görüntüleri kullanılmıştır. Bu görüntüleri kullanarak analiz edebilmek için makine öğrenme teknikleri ve görüntü işleme yöntemleri kullanılmıştır. Bu yöntemlerin kullanılmasıyla histopatoloji görüntülerinin sentezi ve sınıflandırılması sağlanmıştır. Kullanılan yöntemlerin sonucunda yüksek teşhis doğruluk oranı bulunması hedeflenmiştir.

Bu çalışmasının laboratuvarında çalışan patoloğlara yardımcı olabilmesi, teşhis hızının artması, yanlış teşhisin önüne geçilmesi için yararı bulunmaktadır. Ayrıca bu çalışma, akciğer kanserinin teşhisi için ön bilgi verme amacıyla sistemi kullanacak doktorlara yardımcı olacaktır. Bu gibi konularda gelişmelerin sağlanabilmesi için yüksek doğruluk oranına sahip yöntemlere ihtiyaç duyulmaktadır.

Bu çalışma incelendiğinde akciğer tümörünün erken tespiti için doğruluk oranının yüksek olduğu görülmektedir. Ancak bu çalışma incelendiğinde yöntemlere göre değişkenlik gösteren teşhis doğruluk oranı göze önüne alınmalıdır. Akciğer tümörü yüksek teşhis doğruluk oranı sayesinde çok iyi bir şekilde ayırım yapan yöntemler bulunmuştur. Bu da kullandığımız akciğer tümörünün histopatoloji görüntülerinin ayırımının çok iyi yapıldığını göstermektedir. Kullandığımız yöntemler sonucunda %100,000'lük bir teşhis doğruluk oranına sahip yöntemler vardır. Bu sebeple iyi huylu ve kötü huylu tümörün histopatoloji görüntülerinin tam olarak ikiye ayrıldığı görülmüştür. Kullandığımız 240 adet akciğer tümörü histopatoloji görüntüsünün tam anlamıyla 120 adet iyi huylu, 120 adet kötü huylu ayırımı yapılmıştır.

HSÖ-GDA-NB, MDEB-GDA-NB yöntemleri %100,000 olan teşhis doğruluk oranıyla en iyi sonuç veren yöntemler olmuştur. Bu da demektir ki bu yöntemler kullanıldığında histopatoloji görüntülerinin tam olarak ayırımı yapılabilmektedir. HSÖ-GDA-RO, HSÖ-GDA-KA, MDEB-GDA-RO, MDEB-TBA-KA, MDEB-ÇBÖA-KA, YGH-DDA-NB yöntemleriyle de yüksek teşhis doğruluk oranı elde edilmiştir. Ancak bu yöntemlerin kullanılmasıyla akciğer tümörü histopatoloji görüntülerinin tam olarak iyi huylu ve kötü

huyllu ayrımı sađlanamamıřtır. Ancak tam ayrım yapan yntemlerle bu yntemlerinin teřhis dođruluk oranı farkının ok az olduđu belirlenmiřtir.

YGH- SYGA –NB, YGH- DDA –KA, YGH- KKGA –KA, YGH- SYGA- KA yntemlerinin kullanımıyla iyi bir teřhis dođruluk oranı sađlanamamıřtır. Bu yntemlerin akciđer tmr histopatoloji grntlerinin ayrımı iin kullanımı uygun deđildir.

Bu alıřmada kullanılan yntemlerin dıřında akciđer tmrnn sınıf ayrımını yapabilmek iin bařka makine đrenme teknikleri de kullanılabilir. Ayrıca bu alıřmada kullanılan boyut indirgeme yntemlerinin dıřında, faktr analizi gibi yntemlerle yeni dođruluk tespit oranı bulunabilir. Bu alıřmada kullanılan yntemlerin dıřında ikili sađlam deđiřmez leklenebilir anahtar noktaları, yapay sinir ađları gibi yntemlerle yeni bir alıřma yapılabilir.

Bu alıřma sonucunda otomatik olarak kanser hcrelerinin sınıflarının ayrımını yapan bir bilgisayar destekli tespit cihazı oluřturulabilir. Ayrıca bařka alanlarda grnt sınıflandırılması iin yine bu alıřmada kullanılan yntemlerden faydalanma sz konusudur.

KAYNAKLAR

- [1] **Cancer Facts And Figures**, “American Cancer Society”, Atlanta. 2010.
- [2] **Sağlık İstatistikleri Yıllığı**, Türkiye Cumhuriyeti Sağlık Bakanlığı, Ankara, 2009.
- [3] **Parkin, D. M., Pisani, P., & Ferlay, J.**, 1999. Global cancer statistics. CA: a cancer journal for clinicians, 49(1), 33-64.
- [4] <https://www.who.int/en/news-room/fact-sheets/detail/cancer>. 20 Ekim 2019.
- [5] **Gültekin, M., & Boztaş, G.**, 2014. Türkiye kanser istatistikleri. Sağlık Bakanlığı, Türkiye Halk Sağlığı Kurumu, 43.
- [6] **Ceylan, M.**, 2006. Bilgisayar Tomografili Akciğer Görüntülerinin Değerlendirilmesinde Kompleks Değerli Yeni Bir Akıllı Sistem Tasarımı, Doktora Tezi, Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
- [7] **Doi K.**, 2007. “Computer-aided diagnosis in medical imaging: historical review, current status and future potential”, Computerized Medical Imaging and Graphics, 31(4-5), 198-211, .
- [8] **Aydın Kaya, Ali Seydi Keçeli, Ahmet Burak Can.**, 2018. Examination of various classification strategies in classification of lung nodule characteristics, Journal of the Faculty of Engineering and Architecture of Gazi University, <https://doi.org/10.17341/gazimmfd.416530>
- [9] **Sander, S., & Attramadal, A.**, 1968. Hormon Sorumlusu Ve Hiçbilir Olmayan Meme Tümörlerine Ait Radiol'un İn Vivo Yükseltmesi. Açta Pathologica Microbiologica Scandinavica , 74 (2), 169-178.
- [10] **Giger, M. L., Bae, K. T., & MacMahon, H.**, 1994. Computerized detection of pulmonary nodules in computed tomography images. Investigative Radiology, 29(4), 459-465.
- [11] **Xu, X. W., Doi, K., Kobayashi, T., MacMahon, H., & Giger, M. L.**, 1997. Development of an improved CAD scheme for automated detection of lung nodules in digital chest images. Medical Physics, 24(9), 1395-1403.
- [12] **Qian, J., Fan, L., Wei, G. Q., Novak, C. L., Odry, B. L., Shen, H., ... & McGuinness, G.**, 2002. Knowledge-based automatic detection of multitype lung nodules from multidetector CT studies. In Medical Imaging 2002: Image Processing (Vol. 4684, pp. 689-698). International Society for Optics and Photonics.

- [13] **Edis, E. C., & Karhkaya, C.,** 2007. The cost of lung cancer in Turkey. *Tuberkuloz ve Toraks Dergisi*, 55(1), 51-58.
- [14] **Dehmeshki, J., Ye, X., Lin, X., Valdivieso, M., & Amin, H.,** 2007. Automated detection of lung nodules in CT images using shape-based genetic algorithm. *Computerized Medical Imaging and Graphics*, 31(6), 408-417.
- [15] **Akira, H. (Ed.),** 2003. *Complex-valued neural networks: theories and applications (Vol. 5)*. World Scientific.
- [16] **Beigelman-Aubry, C., Raffy, P., Yang, W., Castellino, R. A., & Grenier, P. A.,** 2007. Computer-aided detection of solid lung nodules on follow-up MDCT screening: evaluation of detection, tracking, and reading time. *American Journal of Roentgenology*, 189(4), 948-955.
- [17] **Lee, Y., Seo, J. B., Lee, J. G., Kim, S. S., Kim, N., & Kang, S. H.,** 2009. Performance testing of several classifiers for differentiating obstructive lung diseases based on texture analysis at high-resolution computerized tomography (HRCT). *Computer methods and programs in biomedicine*, 93(2), 206-215.
- [18] **Penedo, M.G., Carreira, M., Mosquera, A., Cabello, D.,** 1998. Computer-aided diagnosis: a neural-network-based approach to lung nodule detection, *IEEE Transactions on Medical Imaging* vol.17, 872–880.
- [19] **Retico, A., Delogu, P., Fantacci, M. E., Gori, I., & Martinez, A. P.,** 2008. Lung nodule detection in low-dose and thin-slice computed tomography. *Computers in biology and medicine*, 38(4), 525-534.
- [20] **Lee, S. L. A., Kouzani, A. Z., & Hu, E. J.,** 2010. Random forest based lung nodule classification aided by clustering. *Computerized medical imaging and graphics*, 34(7), 535-542.
- [21] **Coppini, G., Diciotti, S., Falchini, M., Villari, N., & Valli, G.,** 2003. Neural networks for computer-aided diagnosis: detection of lung nodules in chest radiograms. *IEEE Transactions on Information Technology in Biomedicine*, 7(4), 344-357.
- [22] **Kakar, M., & Olsen, D. R.,** 2009. Automatic segmentation and recognition of lungs and lesion from CT scans of thorax. *Computerized Medical Imaging and Graphics*, 33(1), 72-82.
- [23] **Farag, A. A., El-Baz, A., Gimel'farb, G., & Falk, R.,** 2004. Detection and recognition of lung abnormalities using deformable templates. In *Proceedings of the*

17th International Conference on Pattern Recognition, 2004. ICPR 2004. (Vol. 3, pp. 738-741). IEEE.

- [24] **Bae, K. T., Kim, J. S., Na, Y. H., Kim, K. G., & Kim, J. H.**, 2005. Pulmonary nodules: automated detection on CT images with morphologic matching algorithm—preliminary results. *Radiology*, 236(1), 286-293.
- [25] **Armato, S. G., Li, F., Giger, M. L., MacMahon, H., Sone, S., & Doi, K.**, 2002. Lung cancer: performance of automated lung nodule detection applied to cancers missed in a CT screening program. *Radiology*, 225(3), 685-692.
- [26] **Hardie, R. C., Rogers, S. K., Wilson, T., & Rogers, A.**, 2008. Performance analysis of a new computer aided detection system for identifying lung nodules on chest radiographs. *Medical Image Analysis*, 12(3), 240-258.
- [27] <http://www.cancer.org/cancer/cancer-basics/what-is-cancer.html>. 29 Ekim 2019.
- [28] <http://www.cancer.org/cancer/lung-cancer/detection-diagnosis-staging/signs-symptoms.html>. 31 Ekim 2019.
- [29] <http://www.cancer.org/cancer/lung-cancer/causes-risks-prevention/risk-factors.html>. 31 Ekim 2019.
- [30] https://www.kinderkrebsinfo.de/hastalar/kanser_hakknda_sorular/solid_tuemoer_nedir/index_tur.html. 31 Ekim 2019.
- [31] <http://www.cancer.org/cancer/lung-cancer/about/what-is.html>. 1 Kasım 2019.
- [32] **Bakanlığı, S., & Yılmaz, Ş. V. D. D. A.**, 2008. Küçük Hücreli Dışı Akciğer Kanseri Klinik Ve Patolojik Evrelerin Karşılaştırılması.
- [33] **Murfitt, J., Robinson, P. J. A., & Whitehouse, R. W.**, 2003. The normal chest: methods of investigation and differential diagnosis. *Textbook of radiology and imaging*, seventh ed. Edinburgh: Churchill Livingstone.
- [34] **Fraser, R. S., & Paré, J. P.**, 1994. *Synopsis of diseases of the chest*. WB Saunders company.
- [35] **Arıyürek, M.**, 2009. Konvansiyonel akciğer radyolojisi.
- [36] **Kayhan, S., & Akgunes, A.**, 2011. Histopatolojik Olarak Tanı Konulan Komplike Akciğer Kist Hidatik Olguları. *Türkiye Parazitolojisi Dergisi*, 35, 189-193.
- [37] **Oyar, O., & UK, G.**, 2003. *Tıbbi Görüntüleme Fiziği*, Ankara.
- [38] **Demir, D.**, 2008. Akciğer kitlelerinin görüntülenmesinde düşük doz bilgisayarlı tomografinin rolü.
- [39] **Tuncel E.**, 2002. *Klinik radyoloji*. 2. baskı. İstanbul: Nobel&Güneş, 40-6.

- [40] http://www.tumrad.net/FileUpload/ds58732/File/bilgisayarli_tomograf1_bt_.pdf. 1 Kasım 2019.
- [41] **Başaran Demirkazık F.**, 2005. Bilgisayarlı Akciğer Tomografisinde İndikasyonlar. Türk Toraks Derneği 8. Yıllık Kongresi Özet Kitabı, Antalya.
- [42] **Sönmezoğlu, K.**, 2007. Akciğer Kanserinde Pozitron Emisyon Tomografi (PET) Kullanımı.
- [43] **Gould, M. K., Maclean, C. C., Kuschner, W. G., Rydzak, C. E., & Owens, D. K.**, 2001. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. *Jama*, 285(7), 914-924.
- [44] **Uğur, Ö.**, 2004. Pozitron Emisyon Tomografisi, XIII. TPOG Ulusal Pediatrik Kanser Kongresi,4s.
- [45] **Taşcı, E.**, 2013. Akciğer tomografileri kullanılarak yapay zeka ve görüntü işleme tekniklerine dayalı otomatik nodül bölge tespit yöntemi geliştirilmesi (Doctoral dissertation, Yüksek Lisans Tezi, Ege Üniversitesi, 104p (Yayınlanmamış)).
- [46] **Armato III, S. G., McLennan, G., McNitt-Gray, M. F., Meyer, C. R., Yankelevitz, D., Aberle, D. R., ... & Reeves, A. P.**, 2004. Lung image database consortium: developing a resource for the medical imaging research community. *Radiology*, 232(3), 739-748.
- [47] **Yu, K. H., Zhang, C., Berry, G. J., Altman, R. B., Ré, C., Rubin, D. L., & Snyder, M.**, 2016. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nature communications*, 7, 12474.
- [48] <http://ankarapatolojidernegi.org/2017/09/01/patoloji-nedir/>. 3 Kasım 2019.
- [49] **Extremal, M. S., Matas, J., Chum, O., Urban, M., & Pajdla, T.**, 2002. Robust wide baseline stereo from. In *In British Machine Vision Conference*.
- [50] **Matas, J., Chum, O., Urban, M., & Pajdla, T.**, 2004. Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, 22(10), 761-767.
- [51] **Güney, M., & Arıca, N.**, 2009. DESEN TABANLI İLGİ BÖLGESİ TESPİTİ. *Journal of Naval Sciences and Engineering*, 5(1), 94-106.
- [52] **Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L.**, 2008. Speeded-up robust features (SURF). *Computer vision and image understanding*, 110(3), 346-359.
- [53] **Karakuş, P., & Karabörk, H.**, 2014. Surf Algoritması Kullanılarak Uzaktan Algılama Görüntülerinin Geometrik Kaydı.

- [54] **Shashua, A., Gdalyahu, Y., & Hayun, G.,** 2004. Pedestrian detection for driving assistance systems: Single-frame classification and system level performance. In IEEE Intelligent Vehicles Symposium, 2004 (pp. 1-6). IEEE.
- [55] **Dalal, N., & Triggs, B.,** 2005. Histograms of oriented gradients for human detection.
- [56] **Öztürk, S., Sankur, B., & Ceyhun, B.,** 2010. Karmaşalı Sahnelerde İnsan Bulunması. In IEEE 18th Signal Processing and Communications Applications Conference (SIU).
- [57] **Karakaya, F., Altun, H., & Çavuşlu, M. A.,** 2009. Gerçek zamanlı nesne tanıma uygulamaları için HOG algoritmasının FPGA tabanlı gömülü sistem uyarlaması. In IEEE 17th Signal Processing and Communications Applications Conference (SIU).
- [58] **Alpaslan, N.,** 2013. Gradyan Tabanlı Heterojen Öznitelik Çıkarma Yöntemlerine Yeni Yaklaşımlar (Doctoral dissertation, Yüksek Lisans Tezi, İnönü Üniversitesi, 78s, Ocak).
- [59] **Filiz, Z.,** 2003. Güvenilirlik Çözümlemesi, Temel Bileşenler ve Faktör Çözümlemesi.
- [60] **Çilli, M.,** 2007. İnsan hareketlerinin modellenmesi ve benzeşiminde temel bileşenler analizi yönteminin kullanılması (Doctoral dissertation, Doktora tezi, Hacettepe Üniversitesi, Sağlık Bilimleri Enstitüsü).
- [61] **Sütçüler, E.,** 2006. Gerçek zamanlı video görüntülerinden yüz bulma ve tanıma sistemi.
- [62] **Konak, E. S.,** 2018. Bilgisayar destekli yüz tanıma sistemi tasarımı.
- [63] **Alpar, R.,** 2003. Uygulamalı Çok Değişkenli Yöntemlere Giriş 1, Nobel Basımevi, 2. Baskı, Ankara.
- [64] **Yazar, I., Yavuz, H. S., & Çay, M. A.,** 2009. Temel Bileşen Analizi Yönteminin ve Bazı Klasik ve Robust Uyarlamalarının Yüz Tanıma Uygulamaları. Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi, 22(1), 49-63.
- [65] **Johnson, R. A. & Wichern, D. W.,** 1992. Applied Multivariate Statistical Analysis. Prentice-Hall, Englewood Cliffs, NJ.
- [66] **Wunderlin, D. A., Diaz, M. P., Ame, M. V., Pesce, S. F., Hued, A. C. & Bistoni, M. A.,** 2001. Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia river basin (Cordoba, Argentina). Wat. Res. 35, 2881–2894.

- [67] **Singh, K. P., Malik, A., Mohan, D. & Sinha, S.,** 2004. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India): a case study. *Wat. Res.* 38, 3980–3992.
- [68] **Shrestha, S., Kazama, F., & Nakamura, T.,** 2008. Use of principal component analysis, factor analysis and discriminant analysis to evaluate spatial and temporal variations in water quality of the Mekong River. *Journal of Hydroinformatics*, 10(1), 43-56.
- [69] **Martinez, M., ve Kak, A.C.,** 2001. "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Cilt No. 23, 228-233.
- [70] **Yildiz, E., & Sevim, Y.,** 2016. Comparison of linear dimensionality reduction methods on classification methods. In 2016 National conference on electrical, electronics and biomedical engineering (ELECO) (pp. 161-164). IEEE.
- [71] <https://medium.com/analytics-vidhya/locally-linear-embedding-ile-data-mining-b956616d24e9>. 10 Kasım 2019.
- [72] <https://www.youtube.com/watch?v=scMntW3s-Wk&t=27s>. 10 Kasım 2019.
- [73] **Kalaycı, Ş.,** 2010. SPSS uygulamalı çok değişkenli istatistik teknikleri (Vol. 5). Ankara, Turkey: Asil Yayın Dağıtım.
- [74] **Acar, A. B.,** 2013. İşgücü Piyasası Temel Göstergeleri Bakımından Çok Boyutlu Ölçekleme Analizi İle Türkiye ile diğer OECD Ülkelerinin Karşılaştırılması. İstanbul Üniversitesi İşletme Fakültesi İşletme İktisadı Enstitüsü Yönetim Dergisi, 24(75), 121-144.
- [75] **Alpar, R.,** 2011. Uygulamalı çok değişkenli istatistiksel yöntemler, Ankara. Turkey: Detay yayıncılık.
- [76] **Çelik, Ş.,** 2015. Çok boyutlu ölçekleme analizi ile hayvancılık açısından Türkiye’de illerin sınıflandırılması. Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi, 31(4), 159-167.
- [77] **Ersöz, F.,** 2008. Türkiye ile OECD ülkelerinin sağlık düzeyleri ve sağlık harcamalarının analizi. *İstatistikçiler Dergisi: İstatistik ve Aktüerya*, 1(2), 95-104.
- [78] **Cai, D., He, X., Zhou, K., Han, J., & Bao, H.,** 2007. Locality sensitive discriminant analysis. In *IJCAI* (Vol. 2007, pp. 1713-1726).
- [79] **Yildiz, E., & Sevim, Y.,** 2016. Comparison of linear dimensionality reduction methods on classification methods. In 2016 National conference on electrical, electronics and biomedical engineering (ELECO) (pp. 161-164). IEEE.

- [80] **Agrafiotis, D. K.**, 2003. Stochastic proximity embedding. *Journal of computational chemistry*, 24(10), 1215-1221.
- [81] **Rassokhin, D. N., & Agrafiotis, D. K.**, 2003. A modified update rule for stochastic proximity embedding. *Journal of Molecular Graphics and Modelling*, 22(2), 133-140.
- [82] **Safavian S.R., Landgrebe D.**, 1991. "A survey of decision tree classifier methodology", *IEEE Transactions on Systems Man and Cybernetics*, 21, 660-674.
- [83] **Quinlan J.R.**, 1993. "C4.5: Programs for Machine Learning", Morgan Kaufmann, San Mateo, CA, 302 s.
- [84] **Friedl M.A., Brodley C.E.**, 1997. "Decision tree classification of land cover from remotely sensed data", *Remote Sensing of Environment*, 61, 399-409.
- [85] **Swain P.H., Hauska H.**, 1977. "Decision tree classifier - design and potential", *IEEE Transactions on Geoscience and Remote Sensing*, 15, 142-147.
- [86] **Kavzođlu, T., & ölkesen, İ.**, 2010. Karar ağaçları ile uydu görüntülerinin sınıflandırılması: Kocaeli örneđi. *Harita Teknolojileri Elektronik Dergisi*, 2(1), 36-45.
- [87] **Breiman, L.**, 2001. Random forests. *Machine learning*, 45(1), 5-32.
- [88] **Breiman, L. ve Cutler, A.**, 2005. Random Forest, http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.html. 11 Kasım 2019.
- [89] **Archer, K. J., & Kimes, R. V.**, 2008. Empirical characterization of random forest variable importance measures. *Computational Statistics & Data Analysis*, 52(4), 2249-2260.
- [90] **Özlem, A. K. A. R., & Görmüş, E. T.**, 2019. Göktürk-2 ve Hyperion EO-1 Uydu Görüntülerinden Rastgele Orman Sınıflandırıcısı ile Arazi Kullanım Haritalarının Üretilmesi. *Geomatik*, 4(1), 68-81.
- [91] **Özkan, Y.**, 2008. Veri madenciliđi yöntemleri. Papatya Yayıncılık Eğitim.
- [92] **Pal, M.**, 2005. Random forest classifier for remote sensing classification *International Journal of Remote Sensing*, 26 (1), 217-222.
- [93] **Watts, J. D., Powell, S. L., Lawrence, R. L., & Hilker, T.**, 2011. Improved classification of conservation tillage adoption using high temporal and synthetic satellite imagery. *Remote Sensing of Environment*, 115(1), 66-75.
- [94] **Liaw, A. ve Wiener, M.**, 2002. RandomForest tarafından sınıflandırma ve regresyon. *R haberi*, 2 (3), 18-22.
- [95] **Zhang, H.**, 2004. Saf Bayes'in iyiciliđi. *AA*, 1 (2), 3.

ÖZGEÇMİŞ

Furkan ESMERAY

- 1988 : Elazığ'da doğdu.
- 2005 : Özel Harput Lisesi'nden mezun oldu.
- 2008-2013 : Erciyes Üniversitesi Mühendislik Fakültesi Elektrik-Elektronik Mühendisliği Bölümü'nden mezun oldu.
- 2018 : Fırat Üniversitesi Fen Bilimleri Enstitüsü Biyoteknoloji Anabilim Dalı'nda yüksek lisansa başladı.

