



**EGE ÜNİVERSİTESİ**

**YÜKSEK LİSANS TEZİ**

**AKCİĞER TOMOGRAFİLERİ KULLANILARAK  
YAPAY ZEKA VE GÖRÜNTÜ İŞLEME  
TEKNİKLERİNE DAYALI OTOMATİK NODÜL  
BÖLGE TESPİT YÖNTEMİ GELİŞTİRİLMESİ**

**Arif Erdal TAŞCI**

**Tez Danışmanı : Doç. Dr. Aybars UĞUR**

**Bilgisayar Mühendisliği Anabilim Dalı**

**Bilim Dalı Kodu : 619.01.00**

**Sunuş Tarihi : 01.08.2013**

**Bornova-İZMİR**

**2013**

**EGE ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ**

**(YÜKSEK LİSANS TEZİ)**

**AKCİĞER TOMOGRAFİLERİ KULLANILARAK**

**YAPAY ZEKA VE GÖRÜNTÜ İŞLEME**

**TEKNİKLERİNE DAYALI OTOMATİK NODÜL**

**BÖLGE TESPİT YÖNTEMİ GELİŞTİRİLMESİ**

**Arif Erdal TAŞCI**

**Tez Danışmanı : Doç. Dr. Aybars UĞUR**

**Bilgisayar Mühendisliği Anabilim Dalı**

**Bilim Dalı Kodu : 619.01.00**

**Sunuş Tarihi : 01.08.2013**

**Bornova-İZMİR  
2013**



Arif Erdal TAŞCI tarafından YÜKSEK LİSANS tezi olarak sunulan “Akciğer Tomografileri Kullanılarak Yapay Zeka ve Görüntü İşleme Tekniklerine Dayalı Otomatik Nodül Bölge Tespit Yöntemi Geliştirilmesi” başlıklı bu çalışma E.Ü. Lisansüstü Eğitim ve Öğretim Yönetmeliği ile E.Ü. Fen Bilimleri Enstitüsü Eğitim ve Öğretim Yönergesi’nin ilgili hükümleri uyarınca tarafımızdan değerlendirilerek savunmaya değer bulunmuş ve **01.08.2013** tarihinde yapılan tez savunma sınavında aday ~~oybirliği/oyçokluğu~~ ile başarılı bulunmuştur.

**Jüri Üyeleri:**

**İmza**

**Jüri Başkanı :** Doç. Dr. Aybars UĞUR

.....

**Raportör Üye:** Prof. Dr. M. Serdar KORUKOĞLU

.....

**Üye :** Doç. Dr. Muhammed G. CİNSDİKİCİ

.....





**ÖZET****AKCİĞER TOMOGRAFİLERİ KULLANILARAK YAPAY ZEKA VE  
GÖRÜNTÜ İŞLEME TEKNİKLERİNE DAYALI OTOMATİK  
NODÜL BÖLGE TESPİT YÖNTEMİ GELİŞTİRİLMESİ**

Taşcı, Arif Erdal

Yüksek Lisans Tezi, Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı: Doç. Dr. Aybars UĞUR

Ağustos 2013, 104 Sayfa

Akciğer kanseri, dünyada en yüksek ölüm oranına sahip kanser türlerinden olup, erken tanı ve teşhis durumunda hastaların hayatta kalma oranı önemli derecede artmaktadır. Bilgisayar destekli akciğer nodül tespit - tanılama sistemleri süreci otomatikleştirerek hızlandırması, uzman radyologlara yararlı ve yardımcı olması, tanılama kalitesinin artırılması açısından oldukça önemlidir.

Tez çalışmasında, akciğer BT (Bilgisayarlı Tomografi) kesit görüntüleri üzerinden juxtaopleural nodül bölgelerinin otomatik tespiti ve tanılamasını sağlayan bir yöntem ve sistem geliştirilmiştir. Proje, önileme ve akciğer bölütleme, aday nodül bölgelerinin tespiti, öznelik çıkarma ve sınıflandırma olmak üzere 4 ana aşamadan oluşmaktadır.

Bu çalışmada, alpha hull yöntemi ile belirlenmiş nodül adaylarından çıkarılan özgün 7 adet öznelik ile tüm özneliklerin istatistiksel ölçütler kullanılarak derecelendirilmesi ve seçimi literatüre yapılan katkılar arasındadır.

Tez kapsamında, akciğer BT kesit görüntülerini içeren LIDC verisetinden yararlanılmış, 24 hasta için 138 kesit görüntüsü üzerinden 1410 aday nodül bölgesi ve 40 öznelik çıkartılmıştır. Yapılan deneysel çalışmalar, en iyi sınıflandırma performansının 22 öznelik için 0.9679'lık AUC (Area Under Curve) değeriyle genelleştirilmiş doğrusal model regresyon üzerinde elde edildiğini ve literatürdeki en başarılı sonuçlardan birine ulaşıldığını göstermektedir.

**Anahtar sözcükler:** Bilgisayar Destekli Tespit (BDT), Nodül Tanıma, Akciğer Bölütleme, Öznelik Çıkarma, Örüntü Tanıma, Görüntü İşleme, Makine Öğrenmesi, Akciğer Kanseri, Tıbbi Görüntü Analizi, BT.





**ABSTRACT****DEVELOPMENT OF AUTOMATIC NODULE REGION DETECTION  
METHOD BASED ON ARTIFICIAL INTELLIGENCE AND IMAGE  
PROCESSING TECHNIQUES USING LUNG CTs**

TAŞCI, Arif Erdal

MSc in Computer Engineering

Supervisor: Assoc. Prof. Dr. Aybars UĞUR

August 2013, 104 pages

Lung cancer is one of the types of cancer with highest mortality rate in the world and in case of early detection and diagnosis, the survival rate of patients significantly increases. Computer aided detection and diagnosis of lung nodule systems are quite important in terms of accelerating process by automating, being useful and helpful to expert radiologists, improving the quality of diagnosis.

In this thesis, a novel method and system that provides automatic detection and diagnosis of juxta-pleural nodule region has been developed from cross-sectional images of lung CT (Computerized Tomography). Project, consists of 4 main stages called preprocessing and lung segmentation, detection of candidate nodule regions, feature extraction and classification.

In this study, novel 7 features which are extracted from nodule candidates determined by means of alpha hull method, ranking and selection of all of the features using statistical criterias are the contributions to the literature.

In thesis scope, LIDC dataset containing cross-sectional images of lung CT has been utilized, 1410 candidate nodule regions and 40 features have been extracted from 138 cross-sectional images for 24 patients. Experimental studies show that best classification performance has been obtained for 22 features with 0.9679 AUC value on generalized linear model regression and being reached one of the most successful results in the literature.

**Keywords:** Computer Aided Detection (CAD), Nodule Recognition, Lung Segmentation, Feature Extraction, Pattern Recognition, Image Processing, Machine Learning, Lung Cancer, Medical Image Analysis, CT.



## **TEŞEKKÜR**

Tez çalışması süresince, her türlü desteği esirgemeyerek, bilimsel bilgi birikimi, yol göstericiliği, deneyimi ve önerileri ile bu çalışmanın tamamlanmasına yardımcı olan, tez danışmanım Doç. Dr. Aybars Uğur'a, yüksek lisans çalışmamı 2210 - Yurt İçi Yüksek Lisans Burs Programı kapsamında destekleyen TÜBİTAK BİDEB'e ve her zaman maddi – manevi desteklerini esirgemeyen aileme en içten teşekkürlerimi sunarım.

## İÇİNDEKİLER

### Sayfa

ÖZET .....	v
ABSTRACT.....	vii
İÇİNDEKİLER .....	x
ŞEKİLLER DİZİNİ.....	xiv
ÇİZELGELER DİZİNİ .....	xvii
KISALTMALAR DİZİNİ.....	xix
1. GİRİŞ .....	1
2. AKCİĞER KANSERİNE GENEL BAKIŞ .....	6
2.1 Kanser Nedir?.....	6
2.2 Akciğer Kanseri Nedir? .....	7
2.2.1 Akciğerler.....	7
2.3 Akciğer Kanseri Türleri.....	8
2.4 Akciğer Kanseri İlişkin İstatistikler.....	10
2.5 Küçük Hücreli Dışı Akciğer Kanseri Evrelendirme .....	11
2.6 Küçük Hücreli Dışı Akciğer Kanseri Evrelere Göre Yaşam Oranları .....	11
2.7 Akciğer Nodülü Türleri .....	12
2.7.1 Düzgün sınırlı nodül.....	13
2.7.2 Plevraya yakın nodül.....	13
2.7.3 Damara yakın nodül .....	14
2.7.4 Pleural tail nodül .....	14
2.8 Akciğer Görüntüleme Yöntemleri.....	14
2.8.1 Akciğer radyografileri.....	15
2.8.2 Bilgisayarlı tomografi .....	15
2.8.3 Pozitron emisyon tomografi.....	22
2.8.4 Manyetik rezonans görüntüleme .....	22
2.9 Bilgisayar Destekli Tespit Sistemleri .....	22
2.10 Bilgisayar Destekli Tespit Sisteminin Yararları.....	24
2.11 Bilgisayar Destekli Tespit Sistemlerine İlişkin Problemler .....	25
2.12 Özet .....	25

2.13	Akciğer Nodüllerini İçeren Örnek CT Görüntüleri .....	26
3.	DICOM GÖRÜNTÜ TÜRÜNÜN İNCELENMESİ .....	27
3.1	DICOM Standardı .....	27
3.2	DICOM Yapısı .....	27
3.3	Piksel Değerlerinden HU Birimine Dönüşüm .....	29
4.	ÖNCEKİ ÇALIŞMALAR .....	30
4.1	Görüntünün Elde Edilmesi .....	31
4.2	Önişleme Aşaması .....	31
4.3	Akciğer Bölütleme .....	33
4.4	Akciğer Nodül Tespiti .....	35
4.5	Yanlış Pozitiflerin Giderilmesi .....	36
5.	GELİŞTİRİLEN SİSTEM VE ÖNİŞLEME .....	37
5.1	Geliştirilen Sistemin Mimarisi .....	37
5.2	Önişleme ve Akciğer Bölütleme .....	38
5.2.1	Önişleme .....	38
5.2.1.6	Göğüs kafesinin çıkartılması .....	41
5.2.2	Akciğer bölütleme .....	42
5.3	Tartışma ve Sonuçlar .....	43
6.	PLEVRAYA YAKIN ADAY NODÜL BÖLGELERİNİN TESPİTİ .....	44
6.1	Giriş .....	44
6.2	Alpha Hull Yöntemi ile Juxtapleural Aday Nodül Bölgelerinin Bölütlenmesi .....	44
6.3	Alpha Hull Parametre Değerleri .....	47
6.4	Aday Nodül Bölgelerinin Çıkarılması .....	48
6.5	Gerçek Nodüllerin Etiketlenmesi .....	51
6.6	Tartışma ve Sonuçlar .....	51
7.	ÖZNİTELİK ÇIKARMA – YANLIŞ POZİTİFLERİN GİDERİLMESİ .....	53
7.1	Giriş .....	53
7.2	Şekil Bilgisine Dayalı Öz nitelikler .....	53
7.2.1	Kapsam (Span) .....	54
7.2.2	Derinlik (Depth) .....	54
7.2.3	Sınır uzunluğu (Boundary length) .....	55
7.2.4	Derinliğin kapsama oranı (Depth over span) .....	55
7.2.5	Sınır uzunluğunun kapsama oranı (Bound over span) .....	55

7.2.6	Ana eksen uzunluğunun sınır uzunluğuna oranı (Major axis length over boundary length) .....	55
7.2.7	İkincil eksen uzunluğunun sınır uzunluğuna oranı (Minor axis length over boundary length) .....	56
7.2.8	Kapsamın ana eksen uzunluğuna oranı (Span over major axis length) .....	56
7.2.9	Ana eksen uzunluğu ile ikincil eksen uzunluğunun çarpımının sınır uzunluğuna oranı (Mj_MnAxis over boundary length) .....	56
7.2.10	Alan (Area) .....	56
7.2.11	Yarıçap (Radius) .....	56
7.2.12	Dairesellik oranı (Circularity ratio).....	57
7.2.13	Dışmerkezlilik (Eccentricity).....	57
7.2.14	Katılık (Solidity) .....	57
7.2.15	Dışbükey kabuğun alanı (Convex area) .....	57
7.2.16	Kapsam (Extent) .....	58
7.3	Doku Bilgisine Dayalı Öznitelikler .....	58
7.3.1	Birinci dereceden istatistiksel öznitelikler .....	58
7.3.2	İkinci dereceden istatistiksel öznitelikler .....	62
7.3.3	Şekil ve doku bilgisine dayalı öznitelikler .....	64
7.4	Tartışma ve Sonuçlar .....	65
8.	ÖZNİTELİK SEÇİMİ .....	66
8.1	Giriş .....	66
8.2	Girdi ve Çıktı Matrisinin Oluşturulması .....	66
8.3	Öznitelik Değerleri Üzerinde Normalizasyon İşlemi .....	67
8.4	Özniteliklerin Derecelendirilmesi ve Seçimi .....	68
8.4.1	Hipotez testi : t testi .....	69
8.4.2	Kullback – Leibler ıraksama (Göreceli entropi) .....	69
8.4.3	Bhattacharyya uzaklığı.....	69
8.4.4	Wilcoxon (Mann Whitney U) testi.....	70
8.4.5	ROC (Receiver operating characteristic) değeri .....	70
8.5	Öznitelik Sayısının Belirlenmesi.....	71
8.6	Özniteliklerin Analizi .....	72
8.7	Öznitelik Seçimi Kullanıldığında Elde Edilen Sonuçlar .....	74

8.8	Öznitelik Seçimi Olmadan Elde Edilen Sonuçlar .....	75
8.9	Tartışma ve Sonuçlar .....	76
9.	NODÜL TANIMA .....	78
9.1	Giriş.....	78
9.2	Sınıflandırma Modelleri (LDC, QDC, KNN, MLP, PNN, SVM, CT, GLMR, Ensemble Classifiers).....	78
9.2.1	Doğrusal diskriminant analizi.....	79
9.2.2	İkinci dereceden diskriminant analizi.....	79
9.2.3	K en yakın komşu sınıflandırıcısı .....	79
9.2.4	Çok katmanlı algılayıcı.....	79
9.2.5	Olasılıksal yapay sinir ağı.....	79
9.2.6	Destek vektör makinesi.....	80
9.2.7	Sınıflandırma ağacı .....	80
9.2.8	Naive bayes sınıflandırıcı .....	81
9.2.9	Doğrusal regresyon .....	81
9.2.10	Grup sınıflandırıcılar .....	81
9.3	Test Ekipmanları .....	81
9.4	Kullanılan Veri Seti .....	81
9.5	Kullanılan Performans Ölçütleri .....	84
9.5.1	Sınıflandırma doğruluk oranı.....	84
9.5.2	ROC eğrisi .....	84
9.6	Sınıflandırıcılara Göre Performans Sonuçları.....	85
9.7	Modeller Üzerinde k-katlı Çapraz Doğrulama Uygulanması .....	87
9.8	Örüntü Vektörlerinden Eğitim, Doğrulama ve Test Kümelerinin Oluşturulması.....	88
9.9	Seçilen Modelin Değerlendirilmesi .....	88
9.10	Tartışma ve Sonuçlar .....	92
10.	GENEL DEĞERLENDİRME, SONUÇ VE ÖNERİLER .....	94
10.1	Genel Değerlendirme .....	94
10.2	Sonuç.....	95
10.3	Öneriler ve Gelecekteki Çalışmalar .....	96
	KAYNAKLAR DİZİNİ.....	98
	EKLER .....	

## ŞEKİLLER DİZİNİ

<u>Şekil</u>	<u>Sayfa</u>
2.1 Solunum Sisteminin İç Yapısı (a) Renkli Görünüm (American Lung Association, 2013), (b) Genel Görünüm (American Cancer Society, 2013). ....8	
2.2 Akciğer Kanseri Türleri (American Lung Association, 2013) .....9	
2.3 Düzgün sınırlı nodül (a) Nodül modeli, (b) 5-mm düzgün sınırlı akciğer nodülü.....13	
2.4 Plevraya yakın nodül (a) Nodül modeli, (b) 5-mm plevraya yakın akciğer nodülü (Kostis et al., 2003). .....13	
2.5 Damara yakın nodül (a) Nodül modeli, (b) 7-mm damara bağlantısı olan akciğer nodülü (Kostis et al., 2003). .....14	
2.6 Pleural Tail nodül (a) Nodül modeli, (b) 9-mm plevraya ince bağlantısı olan akciğer nodülü (Kostis et al., 2003). .....14	
2.7 Sol köprücüğün altında gizlenmiş nodülü gösteren akciğer radyografisi (Hayashi et al., 2009) .....15	
2.8 72 yaşında akciğer kanserli bir kadının CT görüntüsü (Ginsberg et al., 2004) .....16	
2.9 CT görüntüleme ünitesi ve CT aygıtı (TÜMRAD-DER, 2013) .....17	
2.10 Farklı oryantasyona sahip Akciğer CT kesit örnekleri (a) Vücut uzayları, (b) Sagittal kesit, (c) Koronal kesit, (d) Aksiyal kesit (Sluimer, 2005). .....18	
2.11 Hounsfield Ölçeği (TÜMRAD-DER, 2013) .....19	
2.12 Pencere değerlerine göre renk dönüşüm işlemi (Colombia Center for New Media Teaching and Learning, 2013) .....20	
2.13 Farklı pencereleme değerleri (a) Mediasten penceresi (b) Parankima penceresi (TÜMRAD-DER, 2013) .....21	
2.14 Akciğer nodüllerinin tespiti için genel bir CAD sisteminin aşamaları (Mughal and Ikram, 2004). .....24	
2.15 Örnek akciğer nodülü CT görüntüleri (a), (b), (c) ve (d) Juxtepleural nodül örnekleri (LIDC, 2012). .....26	
3.1 DICOM yapısı, etiketleri ve değerleri (LIDC, 2012).....28	
5.1 Geliştirilen sistemin genel mimarisi .....37	
5.2 Önışleme aşamaları .....38	



5.3 Kullanılacak akciğer kesitlerinin seçimi (a) Düzgün sınırlı nodül, (b) Plevraya yakın nodül.....	39
5.4 Kesin referans bölgesinin ve nodül bilgilerinin gösterimi (LIDC).....	40
5.5 Pencere değişkenlerinin belirlenmesi (a) Orijinal akciğer CT kesiti, (b) Pencere düzey ve genişliğinin ayarlanması. ....	41
5.6 Göğüs kafesinin çıkartılması (a) Orijinal akciğer kesiti, (b) Boşlukların doldurulması, (c) Otsu eşikleme İşlemi, (d) Morfolojik aşındırma, (e) Göğüs kafesinin maskelenmesi. ....	42
5.7 Akciğer bölütlemesi (a) Göğüs kafesini içeren akciğer görüntüsü, (b) Otsu eşikleme, (c) Morfolojik açma, (d) Boşlukların doldurulması, (e) Akciğer maskesi, (f) Akciğer sınırlarının çıkartılması. ....	43
6.1 Dışbükey kabuk .....	45
6.2 Alpha hull (Frank, T.K., Lorient, S. and Yvinec, M., 2004).....	45
6.3 $\alpha$ yarıçap değerinin değişimine göre oluşan şekil örnekleri (Pless, 2011) .....	46
6.4 Alpha hull örneği (a) Bölütlenmiş akciğer görüntüsü, (b) $\alpha$ -hull yöntemi ile elde edilen sınırlar.....	47
6.5 Aday nodül bölgeleri (a) Akciğer sınırları, (b) Numaralandırılmış aday nodül bölgeleri. ....	51
7.1 İçbükeylik öznitelikleri (a) Kapsam, (b) Derinlik (Sensakovic, Starkey and Armato, 2008).....	54
7.2 Bir nokta ile doğru arasındaki en kısa uzaklığın gösterimi .....	55
7.3 GLCM örnekleri (a) GLCM matrisinin oluşturulması, (b) Konum ve yön tanımlama (Mathworks Inc., 2011). ....	63
8.1 Girdi matrisinin genel görünümü.....	67
8.2 Çıktı matrisi örnekleri (a)Çift satırlı çıktı matrisi, (b) Tek satırlı çıktı matrisi. .....	67
8.3 Öznitelik sayısının değişimi ve AUC değerlerine göre çizilen çizgi grafiği ...	71
8.4 Öznitelik ve hedef sınıftan oluşan eşli Pearson doğrusal katsayı değerlerine göre oluşturulan Heat Map matrisi .....	74
8.5 Öznitelik seçiminin nodül tanıma performansına etkisinin gösterimi.....	76
9.1 Sınıflandırma modelinin genel işleyişi .....	78
9.2 PNN modeli örneği (Specht, 1992). ....	80
9.3 Hata matrisi ve bu matristen çıkarılan performans ölçütleri (a) Hata matrisi, (b) Performans ölçütleri (Fawcett, 2006).....	85

9.4 Sınıflandırıcılara ve performans ölçütlerine göre elde edilen sütun grafiği.....	86
9.5 GLMR üzerinde öznitelik seçim kriteri ve sayılarının AUC ile ilişkisini gösteren sütun grafiği.....	91
9.6 GLMR üzerinde öznitelik seçim kriteri ve sayılarının sınıflandırma doğruluk oranı ile ilişkisini gösteren sütun grafiği.....	92

## ÇİZELGELER DİZİNİ

<u>Çizelge</u>	<u>Sayfa</u>
2.1 Akciğer Kanserine İlişkin İstatistikler .....	10
2.2 Küçük Hücreli Dışı Akciğer Kanseri Evreleri (National Cancer Institute, 2012) .....	11
2.3 Akciğer Kanseri Evreleri ve 5 Yıllık Yaşam Oranları (American Cancer Society, 2013) .....	12
3.1 DICOM Dosya Yapısı (Ulaş ve Boyacı, 2007). ....	29
6.1 $\alpha$ -hull yönteminde kullanılan yarıçap değerleri .....	48
6.2 $ m  < 1.0$ için Bresenham çizgi çizme algoritması (Hearn, Baker and Carithers, 2011) .....	50
8.1 Öznitelik sayısının belirlenmesi .....	72
8.2 Pearson korelasyon değerinin yorumlanması (Evans, 1996).....	73
8.3 Öznitelik seçimi yapıldığında elde edilen sonuçlar .....	75
8.4 Öznitelik seçimi olmadan elde edilen sonuçlar .....	75
9.1 Kullanılan veri setine ilişkin özellikler .....	82
9.2 CT indislerine göre kesit ve aday nodül sayıları .....	83
9.3 Eğitim kümesine ilişkin istatistiksel değerler .....	84
9.4 Sınıflandırıcılara ve performans ölçütlerine göre elde edilen sonuçlar .....	86
9.5 GLMR üzerinde öznitelik seçim kriteri ve sayısına göre 10 iterasyon sonucunda elde edilen en iyi performans değerleri .....	89
9.6 GLMR üzerinde öznitelik seçim kriteri ve sayısına göre elde edilen öznitelik indisleri .....	89
9.7: Öznitelik indisleri ve karşılık gelen öznitelik isimleri.....	90
9.8: Çizelge 9.6'daki 5 farklı öznitelik seçim kriteri ve sayısı için indislerin toplamda tekli frekanslarının sayısı .....	90



## KISALTMALAR DİZİNİ

<u>Kısaltmalar</u>	<u>Açıklama</u>
AAM	Active Appearance Model
ACR	American College of Radiology
AJCC	American Joint Committee on Cancer
ASM	Active Shape Model
CAD	Computer Aided Detection and Diagnosis
CADe	Computer Aided Detection
CADx	Computer Aided Diagnosis
CPU	Central Processing Unit
CT	Classification Tree
CT	Computerized Tomography
DICOM	Digital Imaging and Communications in Medicine
ECT	Ensemble Classifier
FN	False Negative
FP	False Positive
GLCM	Gray Level Co-occurrence Matrix
GLMR	Generalized Linear Model Regression

## KISALTMALAR DİZİNİ (devam)

<u>Kısaltmalar</u>	<u>Açıklama</u>
GT	Ground Truth
HU	Hounsfield Unit
JSRT	Japanese Society of Radiological Technology
LBP	Local Binary Pattern
LP	Lung Parenchyma
LIDC	Lung Image Database Consortium
MLP	Multi Layer Perceptron
MRI	Magnetic Resonance Imaging
NBC	Naive Bayes Classifier
NEMA	National Electrical Manufacturers Association
NLST	National Lung Screening Tutorial
PET	Positron Emission Tomography
RASM	Robust Active Shape Model
SIFT	Scale Invariant Feature Transform
SURF	Speeded up Robust Features
SVM	Support Vector Machine

## KISALTMALAR DİZİNİ (devam)

<u>Kısaltmalar</u>	<u>Açıklama</u>
SVM-RFE	Support Vector Machine - Recursive Feature Elimination
TN	True Negative
TP	True Positive
UICC	International Union Against Cancer
WHO	World Health Organization
3D	3 Dimensional





## 1. GİRİŞ

Akciğer kanseri, akciğer dokularındaki hücrelerin kontrolsüz ve düzensiz bir biçimde çoğaldığı, çevresindeki dokuları kaplayıp, diğer organlara yayılabildiği, dünyada en yüksek ölüm oranına sahip kanser türlerinden birisidir. Dünya Sağlık Örgütü'nün (WHO) raporuna göre akciğer kanseri tüm dünyada kanser türleri arasında, erkeklerde en sık ölüme neden olan birinci, kadınlarda ise ikinci kanser türüdür (WHO, 2004). Tüm dünyada her yıl yaklaşık 1.3 milyon ölüme neden olmaktadır ve bu değer her yıl prostat, göğüs ve kolon kanserlerinden kaynaklanan ölüm sayısının toplamından daha fazladır (Li, 2007). 2013 yılında Amerika'da, 228190 adet yeni akciğer kanseri tanısı ve 159480 adet ölüm ile sonuçlanması beklenirken, erkeklerde %28 ve kadınlarda ise %26 tahmini ölüm oranı ile ilk sırada yer alacağı tahmin edilmektedir (American Cancer Society, 2013). Akciğer kanserinde ölüm oranı, 1991'den beri, Amerika'daki erkekler arasında yılda %1.5 oranında düşüş gösterirken, kadınlarda ise onlarca yıldır yatay bir seviyede kalmıştır. Erkeklerde akciğer kanserinden ölüm oranı Avrupa ülkelerinde de benzer bir eğilim gösterirken, kadınların ölüm oranında son yıllarda düşüş gözlemlenmiştir (Retico et al., 2008, Jemal et al.'dan, 2005).

Akciğer kanserinin tedavisi ve hastanın hayatta kalma oranı, özellikle kanserin erken teşhisine ve hastanın sigara içmemesine bağlıdır. Tütün kullanımı dünyadaki tüm kanser ölümlerinin %22'sine ve akciğer kanseri ölümlerinin %71'ine sebep olmaktadır (WHO, 2013). Akciğer kanseri ileriki evre olgularında analiz edildiğinde, hastanın hayatta kalma oranı yalnızca %15'tir ve son 20 yıldır bu oran önemli ölçüde artırılamamıştır. Buna karşın, akciğer dokularındaki hücrelerin diğer organlara yayılmadan tespiti durumunda ise yaşama oranı %49'lara çıkabilmektedir. Fakat, akciğer kanseri tanılarının sadece %15'i ilk aşamada tespit edilebilmektedir (Armato et al., 2001). Bu durumdan dolayı, akciğer kanseri olma durumuna işaret eden potansiyel nodül bölgelerinin tespit edilmesi, bu hastalıktan kaynaklanan ölüm oranını düşürmede ve akciğer kanserinin tedavisinde oldukça önemli bir rol oynamaktadır.

Akciğer kanserinin ilk aşamada en önemli belirtilerinden biri, kanserli (kötü-huylu) akciğer nodüllerinin tespiti. Akciğer nodülleri, genellikle küresel, dairesel ya da oval biçimde, gri düzeyli renkleri akciğer damarlarına benzer, çapı 3 cm'den küçük, göğüs duvarına ya da damarlara bitişik lezyonlardır (Schneider et al, 2009; Lee et al., 2012). Ayrıca, akciğer nodülleri, konumları ve kendisini kuşatan yapılara yakınlığıyla farklı sınıflara ayrılmaktadır. Bu sınıflar sırasıyla well-circumscribed, juxta-vascular, pleural-tail ve juxta-pleural nodül türleridir.

Akciğer kanserinin erken tespiti ve nodüllerin görüntülenmesi için yaygın olarak düşük radyasyon dozlu CT (bilgisayarlı tomografi), PET (pozitron emisyon tomografi), göğüs radyografisi, ultrasonografi ve X ışınları gibi tıbbi görüntüleme teknikleri kullanılmaktadır. 10 mm'den küçük nodüllerin X ışınları görüntüleme tekniği ile belirlenmesi oldukça zor bir işlemdir (Schneider et al., 2009). Nodüllerin konum, boyut ve doku yoğunluklarının değişiklik gösterebilmesi, göğüs radyografilerinde nodüllerin tespit edilme işlemini 3 kat zorlaştırmaktadır (Gomathi et al., 2010).

Akciğer kanserinin görüntülenmesinde en yüksek hassasiyet ve düşük hata oranına sahip olan, düşük kontrastlı veya küçük boyutlu nodüllerin tespitini sağlayan bilgisayarlı tomografi görüntüleme yöntemi sıklıkla tercih edilmektedir (Li, 2007; Lee et al., 2012). Bilgisayarlı tomografileri ile elde edilen bir dizi kesit görüntüsünde nodül yapılarının daha iyi görselleştirilmesi sağlanırken, uzman radyologların nodül bölgelerinin tespiti ve yorumlaması daha zaman alıcı, güç ve hataya eğilimli bir hale gelmektedir. Ayrıca, uzman doktorlar tarafından belirlenen nodül bölgeleri arasındaki farklılıklar da erken tanı ve tedavi oranını düşürebilmekte, tomografi görüntüleri tekrar tekrar incelenmek zorunda kalabilmektedir.

Bilgisayarlı tomografi görüntüleri üzerinde bilgisayar destekli akciğer nodül bölge tespit yöntemi ve yazılımı geliştirilmesi, akciğer kanserinin erken tanı ve tedavi sürecinde uzman radyologlara yararlı ve yardımcı olması, hastaların hayatta kalma oranını artırması, radyologların nodül tespit ile yorumlama hatalarını düşürmesi, en aza indirmesi, bu işlemler için gereken süreyi azaltması ve gözden kaçabilen kanser özelliği taşıyan nodüllerin tespitini kolaylaştırması

açısından oldukça önem taşımaktadır. Bu ve benzeri birçok yararları içermesi, değişik alanlarda kullanılabilmesi, projenin motivasyonudur ve çalışmanın doğmasına neden olmuştur.

Tez kapsamında geliştirilen sistemin amacı, görüntü işleme ve yapay zeka teknikleri kullanılarak, akciğer tomografileri üzerinden otomatik olarak juxtaopleural nodül bölge tespit yöntemi ve yazılımı geliştirmektir. Juxtaopleural nodüller, plevra yüzeyine belirli derecede bağlantı içermekte, akciğer bölütlemesi aşamasında plevraya yakın Hounsfield birim aralıklarına sahip olmasından dolayı problem çıkarmaktadır. Tez kapsamında, akciğer bölgeleri görüntü işleme teknikleri ile karmaşık arkaplandan ayrılarak bölütlenmiş, alpha hull yöntemi kullanılarak juxtaopleural aday nodül bölgeleri belirlenmiş, bu bölgeler üzerinden morfolojik, istatistiksel ve dokusal öznitelikler çıkartılarak sınıflandırma modelleri yardımıyla asıl nodül bölgesi tespit edilmeye çalışılmıştır. Akciğer tomografi görüntüleri için LIDC veri setinden yararlanılmıştır.

Tez çalışmasında literatüre yapılan temel katkılardan birincisi, alpha hull yöntemi ile belirlenmiş potansiyel nodül adaylarından çıkarılan özgün 7 adet özniteliktir. Bu öznitelikler morfolojik özniteliklerin yanı sıra, morfolojik özniteliklerin ve doku dağılımlarının oranı gibi melez yeni öznitelikleri de kapsamaktadır. Geliştirilen sistemde öznitelik zenginliği sağlanarak nodül adayları arasındaki ayırt edicilik faktörü artırılmaya çalışılmıştır.

Literatüre yapılan ikinci katkı, tez kapsamında özniteliklerin kendi içinde istatistiksel ölçütler kullanılarak derecelendirilmesi ve sınıflandırma başarısını artıracak özniteliklerin seçilmesidir. Öznitelik sayısının performansa etkisi de gözlemlenmiştir.

Tez kapsamında yapılan diğer bir katkı ise, çıkarılan özniteliklerin, literatürde örüntü tanıma problemlerinde kullanılan sınıflandırıcılar üzerinde test edilmesini ve en uygun sınıflandırma modelinin seçilmesini sağlamaktır. Literatürde yapılan çalışmalar, bir tek sınıflandırıcı üzerinde elde edilen sonuçlar üzerinedir. Sınıflandırıcı seçiminde 10 katlı çapraz doğrulama tekniği kullanılarak modelin genelleştirebilme yeteneği ölçülmüştür.

Tez metninin bölümleri şu şekilde özetlenebilir:

Takip eden bölümde, akciğer kanserine ilişkin genel bilgiler ve istatistikler verilmekte, tıbbi görüntüleme yöntemleri açıklanmakta, bilgisayar destekli sistemlerinin yararları ve sınırlamaları üzerinde durularak, nodül türleri örnek görüntülerle açıklanmaktadır.

Tezin üçüncü bölümünde, tıbbi görüntüleme yöntemlerinde kullanılan DICOM görüntü türüne ilişkin genel bilgiler verilmektedir ve temel kavramlar açıklanmaktadır.

Dördüncü bölümde, araştırmacıların bu alanda yapmış oldukları son çalışmalardan bahsedilmektedir. Çalışmalarda sunulan yöntemler ve bu yöntemlerin başarımları üzerinde durulmaktadır.

Beşinci bölümde, tez kapsamında geliştirilen sistemin genel mimarisi ve işleyişi gösterilmekte, gerçekleştirilen ön işleme ve akciğer bölütleme aşamalarından bahsedilmektedir.

Altıncı bölümde, bölütlenen akciğer görüntüleri üzerinden juxtaopleural akciğer aday nodül bölgelerini çıkarmak için kullanılan alpha hull yöntemi açıklanmakta, parametre değerlerine, sınırları belirlemek için kullanılan algoritmaya ve elde edilen sonuçlara ilişkin bilgiler verilmektedir.

Yedinci bölümde, öznitelik çıkarma yöntemleri anlatılmakta, şekle, dokuya ve hem şekle hem de doku dağılımına göre elde edilen özniteliklerden bahsedilmektedir.

Sekizinci bölümde, oluşturulan girdi ve çıktı matrislerinin genel yapısı gösterilmekte, özniteliklerin normalizasyonu, seçimi ve derecelendirilmesi için kullanılan kriterler, özniteliklerin analizi ve sayısının belirlenmesine ilişkin kullanılan yöntem ve yaklaşımlar üzerinde durulmaktadır.

Dokuzuncu bölümde, tez kapsamında geliştirilen nodül tanıma sisteminde hangi veri setindeki görüntülerden yararlanıldığı, kullanılacak sınıflandırma modelinin nasıl seçildiği ve değerlendirildiği, eğitim, doğrulama ve test verilerinin hangi yöntemle oluşturulduğu, hangi başarımlar ölçütleri kullanılarak ne ölçüde başarı sağlandığına ilişkin bilgiler verilmektedir.

Tezin son bölümünde ise literatüre yapılan katkılar, deneysel çalışmalardan çıkarılan sonuçlar, genel değerlendirmeler ve gelecekte bu alanda çalışacak olan araştırmacılara öneriler sunulmaktadır. Kaynaklar dizini ve ekler, bu bölümü takip etmektedir.

## 2. AKCİĞER KANSERİNE GENEL BAKIŞ

### 2.1 Kanser Nedir?

İnsan vücudu, trilyonlarca canlı hücreden oluşur. Normal vücut hücreleri büyür, yeni hücrelere bölünür ve ölür. İnsan yaşamının ilk yıllarında, normal hücreler daha hızlı bölünerek kişinin gelişimini sağlar. Yetişkinlik çağına gelindiğinde hücrelerin çoğu, yıpranmış ya da ölen hücreleri değiştirmek ya da hasarları gidermek amacıyla bölünür (American Cancer Society, 2013).

Vücudun bir bölümünde yer alan hücreler kontrol dışında büyümeye başladığında ve olağandışı değişikliklere neden olduklarında kanser başlamaktadır. Ayrıca, kanserli hücrelerin büyümesi normal hücrelerin büyümesinden farklılık göstermektedir. Kanserli hücreler ölmek yerine sürekli büyüyerek yeni, normal işlevini yerine getirmeyen hücreler oluştururlar, yayılabilecekleri diğer dokular üzerinde dolaşırlar ve dokulara yayılırlar ki bu duruma metastaz adı verilir ve DNA'larına da zarar verirler. Her hücrede yer alan DNA, yapılacak tüm eylemleri düzenler, yönetir. Normal hücrede DNA hasara uğradığında ya onarılır ya da hücre ölümü gerçekleşir. Kanserli hücrede ise hasara uğrayan DNA onarılmaz fakat hücre ölmesi gerektiği halde ölmez, vücut gereksinim duymadığı halde yeni hücreler oluşturmaya devam eder. Yeni oluşan hücreler de hasara uğramış DNA'ya sahiptir. Sigara içilmesi DNA'ya zarar veren etkenlerden birisidir. Kanserli hücreler için kan ve kan yapıcı organlar gereklidir (American Cancer Society, 2013).

Kanserli hücreler kan dolaşımına ve vücuttaki lenf damarlarına girdiğinde metastaz durumu olmakta, diğer doku ve organlara yayılmaktadır. Kanser başlandığı yer neresi olursa olsun, daima başladığı yer ile adlandırılmaktadır. Örneğin, göğüs kanseri karaciğere yayıldığında karaciğer kanseri yerine göğüs kanseri olarak adlandırılmaktadır (American Cancer Society, 2013).

Farklı kanser türleri farklı davranışlar göstermektedir. Akciğer kanseri ve göğüs kanseri farklı yayılım oranı ve tedavi seçeneklerine sahiptir. Tüm tümörler kanserli değildir. Kanserli olmayan tümörler iyi huylu olarak adlandırılır. İyi

huyllu t m rler birtakım problemlere neden olabilmektedir. Bu t m rler,  ok hızlı bir bi imde b y yerek saėlıklı organlar ve dokular  zerinde baskı oluřturabilmektedir. Fakat diėer dokulara ve v cudun diėer b l mlerine yayılmazlar. Bu durumdan dolayı hi bir zaman yařamı tehdit edici bir unsur haline gelmezler (American Cancer Society, 2013).

## **2.2 Akciėer Kanseri Nedir?**

Akciėer kanseri, kanserin akciėerde bařladıėı bir kanser t r d r. Bir kiři akciėer kanseri olduėunda, normal iřlevini yerine getirmeyen h creler denetlenemeyen bir bi imde b y rler ve bir araya gelerek t m r  oluřtururlar, akciėer b lgesinin etrafındaki saėlıklı dokuları yok ederler.

Akciėer kanserini iyi bir bi imde anlayabilmek i in akciėerlerin yapı ve iřlevlerini  ėrenmek gerekmektedir. Bu durumdan dolayı, alt b l mlerde akciėerin genel yapısı ve iřlevleri hakkında bilgiler verilmiřtir.

### **2.2.1 Akciėerler**

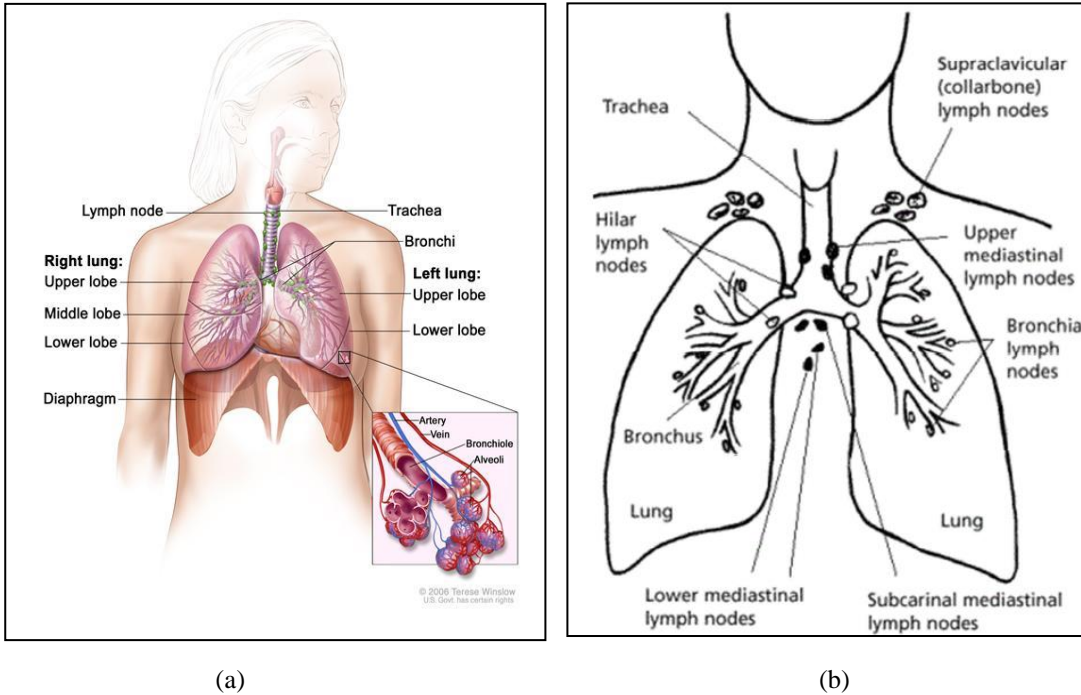
Akciėerler, doku ve organ gruplarının birlikte  alıřarak nefes almamıza yardımcı olduėu, v cudun havalandırma sistemi olarak adlandırılan bir b l md r. Solunum sisteminin temel g revi, oksijen bulunan temiz havayı kan dolařımına ve v cudun i ine, karbondioksit gibi zararlı gazları v cudun dıřına tařıyarak gaz alıřveriřini saėlamaktır (American Lung Association, 2013).

Saė akciėer 3 loba (b l me), sol akciėer ise 2 loba b l nmektedir. V cudun sol tarafında kalbin yer almasından dolayı, sol akciėer saė akciėer b lgesine g re daha k   kt r.

Nefes alındıėı zaman, aėız ve burun yoluyla solunan hava, nefes borusu  zerinden akciėerlere gitmektedir. Nefes borusu kendi i inde bronřlara, bronřlar da daha k   k sonlandırıcı dallara, bronřiyollere b l n r. Bronřiyollerin en u  noktalarında alveoller olarak bilinen, i inden ince kan damarlarının ge tiėi hava kesecikleri bulunmaktadır. Solunan temiz hava v cuda girerken, karbondioksit gazı v cuttan alveollere ge erek dıřarı atılır (American Cancer Society, 2013).

Plevra adında, akciğerin dışını kaplayan iki yapraklı zardan oluşan ince bir katman yer alır. Akciğer loplarını kuşatarak göğüs duvarından akciğerleri ayırır, akciğerlerin nefes alımı için hareketine yardımcı olur.

Solunum sisteminin iç yapısı Şekil 2.1’de gösterilmektedir. Akciğer bölgesinin altında yer alan diyagram, göğüs boşluğunu karın bölgesinden ayırmaktadır. Nefes alış veriş sırasında aşağıya ve yukarıya hareket ederek solunumda görev alır.



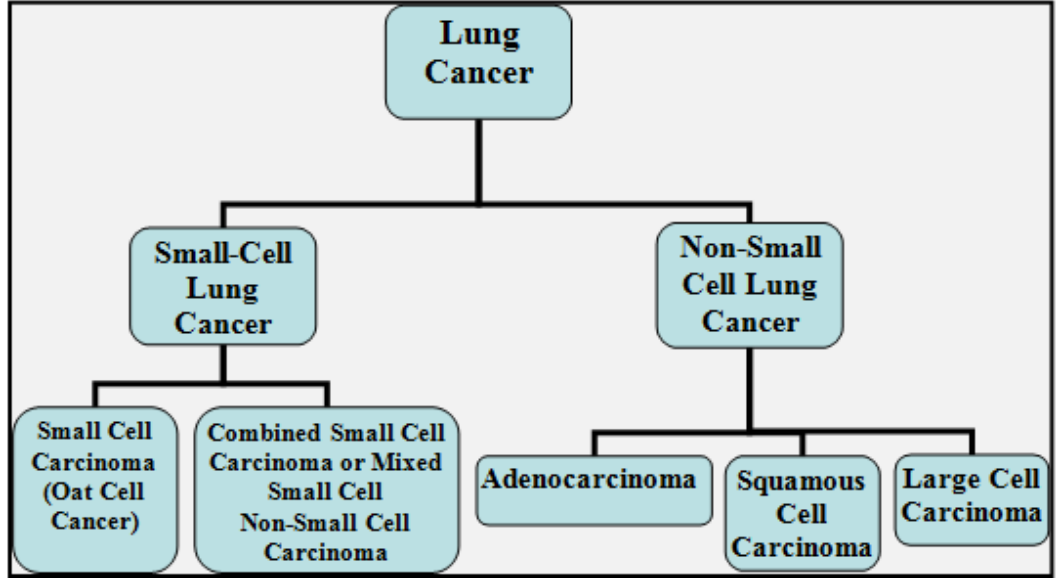
Şekil 2.1 Solunum Sisteminin İç Yapısı (a) Renkli Görünüm (American Lung Association, 2013), (b) Genel Görünüm (American Cancer Society, 2013).

Kanserli hücreler kan dolaşımına ve vücuttaki lenf damarlarına girdiğinde metastaz durumu olmakta, diğer doku ve organlara yayılmakta, bronş etrafındaki lenf düğümlerinde gelişimine devam etmektedir. Kanserlin hangi aşamada olduğu, tedavi için hangi yöntemlerin uygun bulunacağına ilişkin kararlar, kanserli hücrelerin lenf damarlarına girip girmeme duruma göre belirlenmektedir (American Lung Association, 2013).

### 2.3 Akciğer Kanseri Türleri

İki tür akciğer kanseri bulunmaktadır. Bu kanser türleri, kendi içerisinde de alt gruplara ayrılmaktadır (Şekil 2.2).





Şekil 2.2 Akciğer Kanseri Türleri (American Lung Association, 2013)

Tüm akciğer kanserlerinin yaklaşık %20'si **küçük hücreli akciğer kanseridir (Small-Cell Lung Cancer)**. Sigara içimi ile ilişkisinin güçlü, kanserli hücrelerin hızlı büyüme oranına ve en yüksek ölüm oranına sahip olduğu, ilk aşamalarında bile tedavisi çok zor yapılan bir akciğer kanseridir. Erkeklerde kadınlara oranla daha fazla rastlanır.

**Küçük hücreli dışı akciğer kanseri (Non-Small Cell Lung Cancer)**, en yaygın olan akciğer kanseridir. Tüm akciğer kanserlerinin %85 - %90'lık bir kısmını oluşturmaktadır. Küçük hücreli akciğer kanserine göre kanserli hücreler daha yavaş büyür ve yayılır. Erken aşamalarında tespit edildiğinde hastanın hayatta kalma oranı yüksektir. Küçük hücreli dışı akciğer kanseri, mikroskop ile görüntü farklılıkları, şekil ve boyut özelliklerine göre üç türe ayrılmaktadır.

**Adenokarsinom (Adenocarcinoma)** : Akciğerin dış bölgelerinde rastlanır, akciğer kanserinin %40'lık bir bölümünü oluşturur. Kadınlarda ve gençlerde görülme durumu daha fazladır. Sigara içenlerde olmasına rağmen, sigara içmeyen hastalarda en yaygın olan kanser türüdür (American Cancer Society, 2013).

**Skuamöz Hücreli Karsinom (Squamous Cell Carcinoma)** : Tüm akciğer kanserinin %25'lik ve %30'luk kısmını oluşturur. Akciğerin içerisindeki

bronşlara yakın olan skuamöz hücrelerinde rastlanır. Sigara içenlerle ilişkilidir (American Cancer Society, 2013).

**Büyük Hücreli Karsinom (Large Cell Carcinoma)** : Tüm akciğer kanserinin %10 - %15’lik bölümünü oluşturur. Akciğerin herhangi bir bölümde rastlanılabilir. Kanserli hücrelerin, çok hızlı bir biçimde büyüme ve yayılma özelliği ile küçük hücreli akciğer kanserine benzerlik göstermektedir (American Cancer Society, 2013).

## 2.4 Akciğer Kanserine İlişkin İstatistikler

Akciğer kanserine ilişkin istatistikler, hem küçük hücreli hem de küçük hücreli dışı akciğer kanserini kapsamaktadır. Akciğer kanseri, dünyada en yaygın olan kanserler sıralamasında ikinci sıradadır. Erkeklerde prostat kanseri, kadınlarda göğüs kanseri daha sık görülmektedir. Ayrıca, prostat, kolon ve göğüs kanserinden kaynaklanan ölüm oranlarının birleşiminden daha fazla bir oranı ile akciğer kanseri, en ölümcül kanser olma özelliğini korumaktadır.

American Cancer Society’e (2013) göre Amerika’da 2013 yılında akciğer kanserine ilişkin yapılan istatistikler Çizelge 2.1’de gösterilmektedir.

Çizelge 2.1 Akciğer Kanserine İlişkin İstatistikler

Cinsiyet / Durum	Yeni Akciğer Tanısı	Tahmini Ölüm Sayısı
<b>Erkek</b>	118080	87260
<b>Kadın</b>	110110	72220

Çizelge 2.1’e göre 2013 yılında toplam 228190 adet yeni akciğer tanısı ve 159480 kişinin akciğer kanserinden öleceği tahmin edilmektedir. Bununla birlikte, akciğer kanseri hastalarının %67’si 65 yaş ve üzerindeki kişilerden oluşmaktadır. 45 yaş altındaki kişilerde görülme olasılığı yaklaşık %2’dir. Ortalama yaş sınırı 71 olarak belirlenmiştir. Bu kanserin erkeklerde yaşamı boyunca görülme olasılığı %1-%13 arasında değişmekle birlikte, kadınlarda bu oran %1 ile %16 arasındadır. Sigara içen kişilerde görülme olasılığı artmaktadır. Genel anlamda kadınlarda akciğer kanseri oranı daha düşük olmasına rağmen, erkeklerde son 2 yılda bu oranda düşme gözlenirken, kadınlarda son günlerde düşerken aradaki fark kapanmaya yaklaşmıştır (American Cancer Society, 2013).

## 2.5 Küçük Hücreli Dışı Akciğer Kanseri Evrelendirme

Akciğer kanserinde ölüm oranı, tedavi yöntemleri ve uygulanacak yaklaşımlar kanserin hangi evrede olduğuna göre değişiklik göstermektedir. Bunun için Amerika Kanser Birliği (AJCC) (American Joint Committee on Cancer) ve Uluslararası Kanser Savaş Örgütü (UICC) (International Union Against Cancer) tarafından ana tümörün büyüklüğü ve yayılımı (T), bölgesel lenf bezi tutulumu (N) ve uzak metastaz ile diğer organlara yayılımının varlığı (M) dikkate alınarak TNM evrelendirme sistemi önerilmiştir. Bu evreler Çizelge 2.2'den incelenebilir.

Çizelge 2.2 Küçük Hücreli Dışı Akciğer Kanseri Evreleri (National Cancer Institute, 2012)

<b>Occult Karsinom</b>	Tümör hücreleri sputumda bulunur. CT taramaları ve diğer görüntüleme yöntemleri akciğer tümörünü göstermez.
<b>Evre 0</b>	Kanserli hücreler, akciğerin en içteki katmanında bulunur. Carcinoma in situ olarak da bilinir. Büyüme ve yayılma durumunu göstermez.
<b>Evre I</b>	<b>Evre IA:</b> Tümörün çapı < 3 cm'dir, akciğeri kuşatan zarlara ulaşmaz, bronşun ana dallarını etkilemez. Lenf düğümlerine ve diğer organlara yayılmaz. <b>Evre IB:</b> Tümörün çapı < 5 cm'dir, bronşların dallarına ve plevraya doğru büyüme kısmen hava yollarını tıkama görülür.
<b>Evre II</b>	Tümörün çapı < 7 cm'dir, kanser hücreleri lenf düğümlerinin yakınında göğüs duvarı, diyafram, plevra, bronş ve kalbi kuşatan dokulara yayılabilir, akciğerin bir bölümünde birden fazla tümör bulunabilir.
<b>Evre III</b>	Nodül herhangi bir boyutta olabilir, akciğerin bir bölümünde birden fazla nodül olabilir, kalp, soluk borusu gibi diğer organlara yayılabilir.
<b>Evre IV</b>	Kötü huylu tümörler her iki akciğerde bulunabilir, beyin, karaciğer ve kemiğe yayılma durumu görülebilir. Kanserli hücreler plevranın iki katmanı arasında akışkan bir biçimde bulunur.

## 2.6 Küçük Hücreli Dışı Akciğer Kanseri Evrelere Göre Yaşam Oranları

Akciğer kanserinde erken tanı ve tedavi, hastanın hayatta kalması için oldukça önemli bir paya sahiptir. Bu nedenle, küçük hücreli dışı akciğer kanserinde evrelere göre yaşam oranlarını belirlemek oldukça önem taşır. Çizelge 2.3'te hastalıklar her bir evre için 5 yıllık yaşam oranları gösterilmiştir. Evre

IV'teki bir hastanın hayatta kalabilme olasılığının oldukça düşük olduğu söylenebilir.

Çizelge 2.3 Akciğer Kanseri Evreleri ve 5 Yıllık Yaşam Oranları (American Cancer Society, 2013)

EVRE	5 YILLIK YAŞAM ORANI
Evre IA	% 49
Evre IB	% 45
Evre IIA	% 30
Evre IIB	% 31
Evre IIIA	% 14
Evre IIIB	% 5
Evre IV	% 1

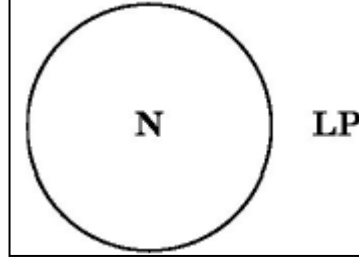
## 2.7 Akciğer Nodülü Türleri

Akciğer nodülleri, küresel/elipsoidal/oval biçimde olan, damla şekline benzer, çevresi akciğer dokusu ile sarılı, arkaplandaki yoğunluktan daha açık ve yarısaydam bir yoğunluğa sahip, karmaşık yapıli akciğer lezyonlarıdır (Diciotti et al., 2008; Kartaloğlu, 2008). Eski adı “coin lezyon” olarak da bilinmektedir (Kartaloğlu, 2008). Lezyon 3 cm'den büyük ise kitle olarak tanımlanmaktadır ve lezyonun kötü huylu olma olasılığı yükselmektedir (Kartaloğlu, 2008). Küçük akciğer nodüllerinin çapı genellikle 1 cm'den küçüktür. Nodülün yoğunluğu, kendini kuşatan akciğer dokusundan önemli ölçüde yüksektir. Akciğer nodülleri, damarlar ve plevra yüzeyi gibi kendini kuşatan yapılara bağlantılar içerebilmektedir.

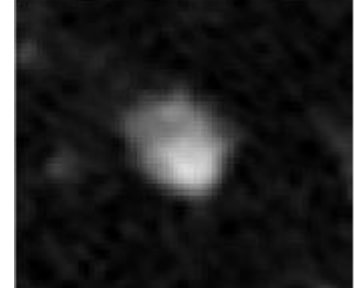
Akciğer nodülleri bulundukları konumlara ve çevresini kuşatan yapılara göre 4 farklı türe ayrılmaktadır (Kostis et al., 2003). Bu nodül türleri sırasıyla düzgün sınırlı (isolated) (well-circumscribed), plevraya yakın (juxtapleural), damara yakın (juxta-vascular), pleural-tail nodülleridir. Akciğer parankimasi (Lung Parenchyma) (LP), plevra yüzeyi (PS), göğüs duvarı (TW), damarlar (V) ve nodül bölgesinin (N) durumlarına göre tanımlanmış nodül türleri alt bölümlerde açıklanarak, örnek görüntüler gösterilmiştir.

### 2.7.1 Düzgün sınırlı nodül

Düzgün sınırlı nodül, akciğer içerisinde olan, çevresindeki damar, plevra gibi yapılara herhangi bir bağlantı içermeyen nodül türüdür. Nodülün CT görüntüsündeki yoğunluk değeri, çevresinde yer alan akciğer parankimasının yoğunluk değerinden yüksektir (Kostis et al., 2003) (Şekil 2.3).



(a)

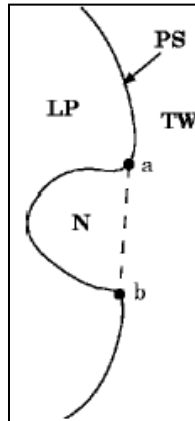


(b)

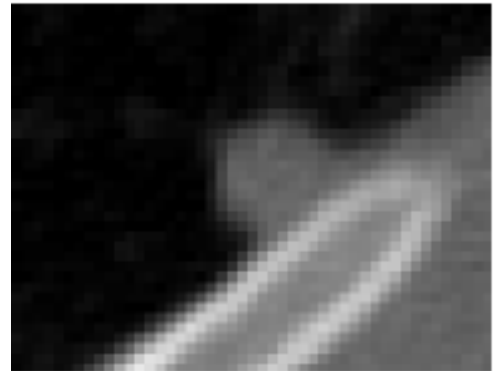
Şekil 2.3 Düzgün sınırlı nodül (a) Nodül modeli, (b) 5-mm düzgün sınırlı akciğer nodülü.

### 2.7.2 Plevraya yakın nodül

Plevraya yakın nodüllerde, nodülün önemli bir bölümünün plevra yüzeyine bağlantısı bulunmaktadır. Bu nedenle, plevranın ve nodülün sınırlarının tanımlanması, oldukça güç bir problem olarak ele alınmaktadır. Genel bir bölütleme yöntemi ile nodülün akciğer parankimasından ayrılması kolay iken nodülün göğüs bölgesinden ve plevradan ayrılması hem nodülün hem de plevranın CT yoğunluklarının yüksek olmasından dolayı zorlaşmaktadır. Bu durumun çözülmesi için, plevranın eğriselliği gibi birtakım özellikler kullanılmaktadır (Kostis et al., 2003) (Şekil 2.4).



(a)

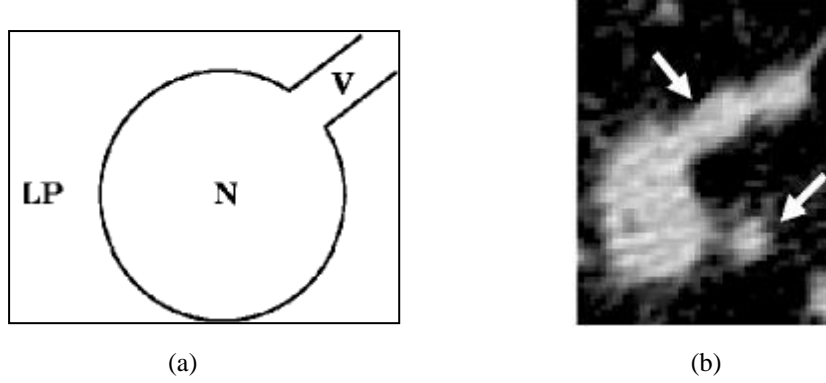


(b)

Şekil 2.4 Plevraya yakın nodül (a) Nodül modeli, (b) 5-mm plevraya yakın akciğer nodülü (Kostis et al., 2003).

### 2.7.3 Damara yakın nodül

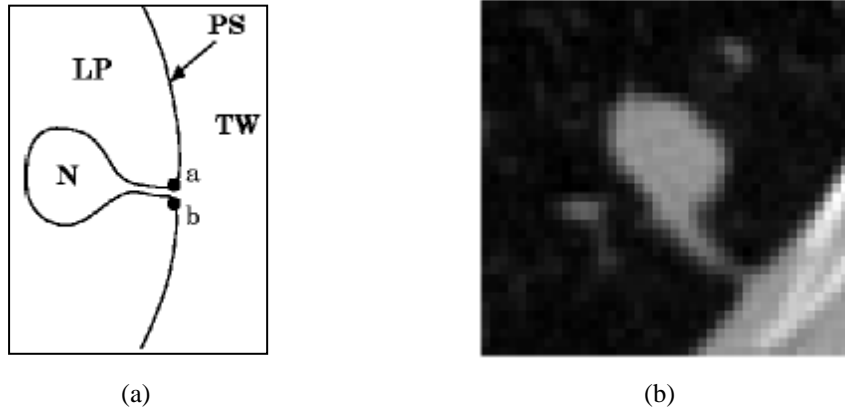
Damara yakın nodül, akciğer içerisinde yer alan, nodül ve damar arasında bağlantının bulunduğu nodül türüdür. Nodül ve nodülün yakınında ya da bitişiğinde yer alan damarların, CT yoğunluk değerlerinin birbirine benzer olması, bölütleme sırasında problemler ortaya çıkarmaktadır (Kostis et al., 2003) (Şekil 2.5).



Şekil 2.5 Damara yakın nodül (a) Nodül modeli, (b) 7-mm damara bağlantısı olan akciğer nodülü (Kostis et al., 2003).

### 2.7.4 Pleural tail nodül

Pleural-tail nodüllerde, plevra yüzeyi ile nodül arasında ince bir bağlantı bulunmaktadır. Bölütleme sırasında plevranın nodüle bağlantılı kısmı da, nodülün hacminin gerçekte olması gerekenden daha az bir değerde hesaplanmaması için dikkate alınmalıdır (Diciotti et al., 2008) (Şekil 2.6).



Şekil 2.6 Pleural Tail nodül (a) Nodül modeli, (b) 9-mm plevraya ince bağlantısı olan akciğer nodülü (Kostis et al., 2003).

## 2.8 Akciğer Görüntüleme Yöntemleri

Akciğer kanserinin tanılanmasında ve nodüllerin tespitinde CT, göğüs X ışınları, PET, bronkoskopi ve sputum sitoloji gibi çeşitli radyolojik görüntüleme

yöntemleri kullanılmaktadır. Bu görüntüleme yöntemlerinden en yaygın olanları, CT, X ışınları, MRI ve PET olup, alt bölümlerde açıklanmıştır.

### 2.8.1 Akciğer radyografileri

Akciğer radyografiler, göğüs hastalıklarında X ışınlarından yararlanılarak, ilk ve temel tanının yapıldığı, 2 yönlü çekilen ucuz bir görüntüleme yöntemi olarak kullanılmaktadır (Şekil 2.7).

Akciğer grafi için hastanın sırtı röntgen tüpüne dönük, elleri her iki skapula görüntüleme alanı dışında kalacak şekilde yanda bulunur. Lateral grafi için eğer yön belirtilmediyse sol yan grafi çekilir. Hastanın sol yanı kasete, sağ yanı tüpe dönük olarak pozisyon verilir. Hastanın boyu, kilosu beden yapısı göz önüne alınarak uygun doz belirlenir. Hastaya nefes aldırılıp tutturularak çekim gerçekleştirilir (TRD Yeterlilik Kurulu, Rehberler ve Standartlar Komitesi, 2007).

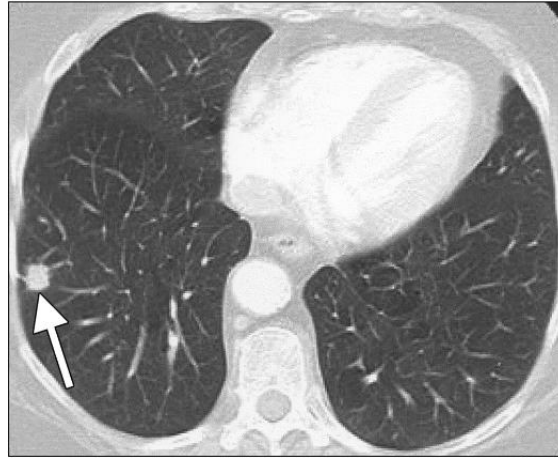


Şekil 2.7 Sol köprücüğün altında gizlenmiş nodülü gösteren akciğer radyografisi (Hayashi et al., 2009)

### 2.8.2 Bilgisayarlı tomografi

Tomografi, 1979 yılında aynı yıl içinde Nobel fizyoloji ve tıp ödülünü kazanan Godfrey N. Hounsfield ve Professor Allan M. Cormack tarafından icat edilmiştir (Gonzalez and Woods, 2002). Bilgisayarlı tomografi, X ışınlarını kullanarak, hastanın iki boyutlu, çok sayıda farklı açılarda görüntüsünün alınmasıyla, akciğer bölgesinin içyapısının ve yumuşak dokuların 3 boyutlu görüntüsünün elde edilmeye çalışıldığı bir tıbbi görüntüleme yöntemidir. CT’de, ring biçiminde hastayı kuşatan bir makine ile X ışın kaynağı arasında duran düz bir masa üzerine uzanan hastanın, farklı açılarda kesit görüntüleri alınır (Gonzalez and Woods, 2002). Hastanın, vücudundaki yapıların daha iyi görülebilmesi için hastaya bir kontrast solüsyonu içirilebilir. Hastaya herhangi bir cerrahi müdahale uygulanmaz.

CT görüntüleri X ışınları gibi hastayı belli bir miktar radyasyona maruz bırakır ve X ışınlarına göre daha hassas bir görüntüleme imkânı sunar. Düşük dozda görüntüleme yapıldığında radyasyon oranının düşmesine, nodüllerin ayrıntılı biçimde görüntülenebilmesine imkan sağlaması açısından akciğer kanserinin erken tanı ve tedavisinde yaygın olarak kullanılmaktadır (Şekil 2.8). Nodülün şekli, boyutu ve konumu, lenf düğümlerine yayıp yayılmadığı gibi bilgiler CT ile elde edilir (American Cancer Society, 2013).



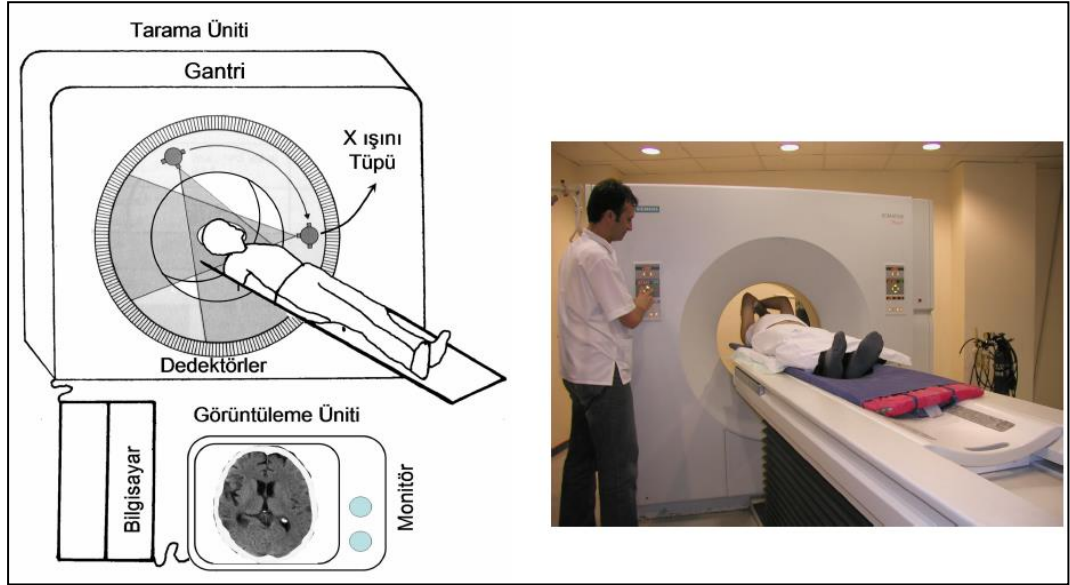
Şekil 2.8 72 yaşında akciğer kanserli bir kadının CT görüntüsü (Ginsberg et al., 2004)

Bir CT aygıtında tarama bölümü, bilgisayar sistemi ve görüntüleme bölümü olmak üzere 3 bölüm bulunmaktadır (Şekil 2.9). **Tarama bölümü**, kare şeklinde gantri'den ve hasta masasından oluşur. Gantri içerisinde X ışını tüpü ve dedektörler, ortasında hastanın girebileceği gantri açıklığı yer alır. CT görüntüsü alınırken, hastanın uzandığı masa kesit kalınlığına ve kesitler arasındaki mesafeye göre hareket eder (TÜMRAD-DER, 2013). X ışını tüpünün farklı açılarda dönmesiyle yaydıkları foton demetleri, hastanın dokuları tarafından soğrularak dedektörler aracılığıyla filmin üzerine düşer. Bu ışının miktarı ölçülerek sayısallaştırılır.

CT aygıtındaki ikinci bölüm, sayısal verileri alan, işleyen kesit görüntülerini ve 3 boyutlu görüntü elemanları olan voksel değerlerini oluşturan **bilgisayar sistemidir** (TÜMRAD-DER, 2013). Bilgisayar sistemi kendi içerisinde CPU (Central Processing Unit), hafıza birimi, aritmetiksel ve mantıksal birimden oluşmaktadır. Rekonstruksiyon işlemi ile tarama alanını ifade eden piksel değerleri elde edilir.



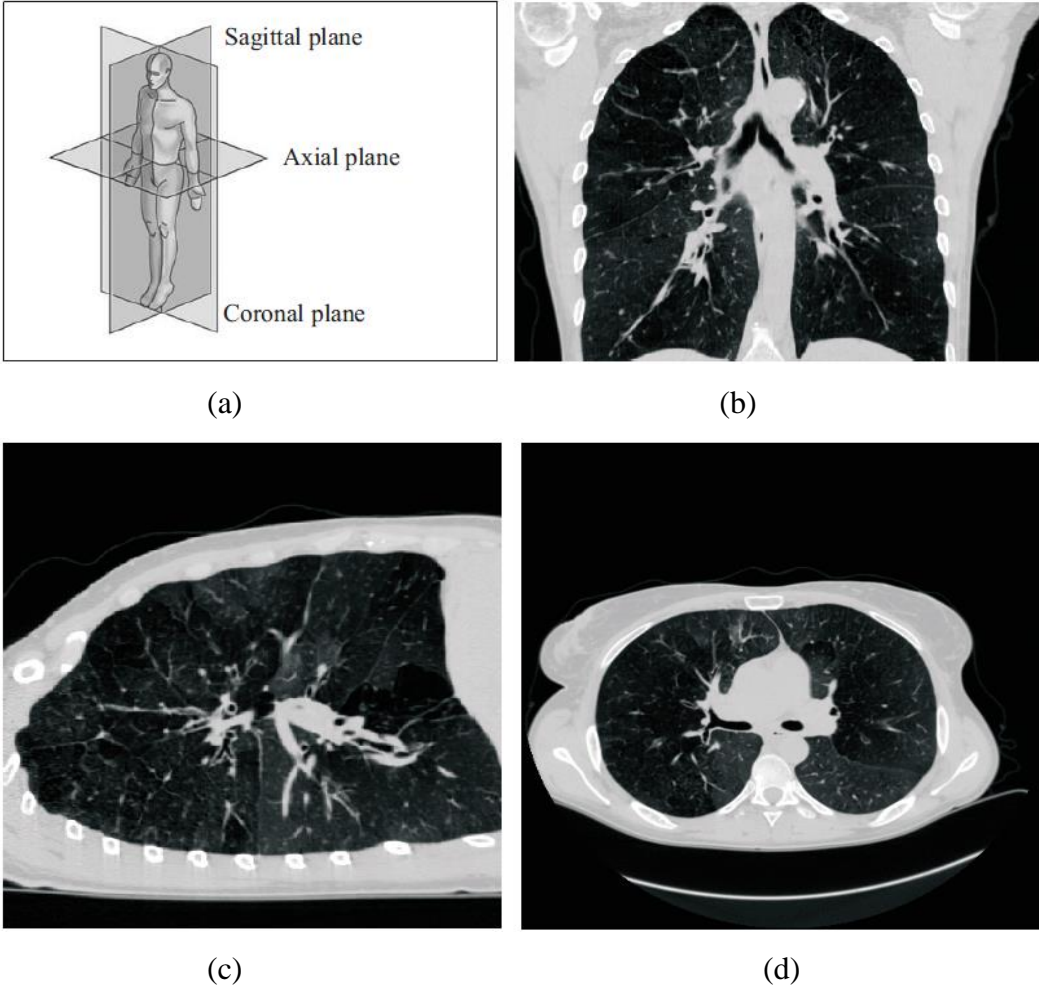
CT aygıtındaki üçüncü bölüm, **görüntüleme bölümüdür**. Rekonstruksiyon işlemi sonucunda elde edilen sayısal değerlerin işlenerek gri ölçekli renk tonlarına sahip görüntünün oluşturulduğu bölümdür. Bu bölüm aynı zamanda sistemin komuta bölümüdür (TÜMRAD-DER, 2013).



Şekil 2.9 CT görüntüleme ünitesi ve CT aygıtı (TÜMRAD-DER, 2013)

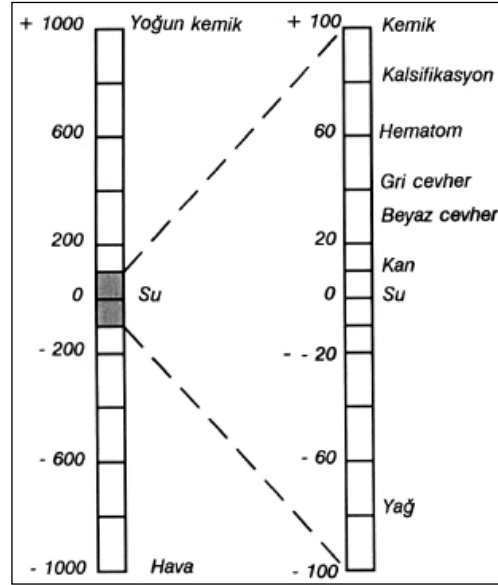
Günümüzde CT görüntüleri  $256 * 256$ ,  $512 * 512$ 'lik matris boyutlarındadır. Matris boyutu arttıkça görüntüdeki ayrıntı oranı ve kalite de artmaktadır. CT aygıtları bir hasta için 400'e yakın kesit görüntüsü çıkartabilmekte, bu görüntülerden 3D (3 Dimensional) görüntüler elde edilebilmektedir.

CT görüntüleme yöntemi ile hastanın konumu değiştirilmeden farklı kesit planları ile farklı yönelime sahip görüntüler alınabilmektedir. Bu görüntüleme yöntemi, multiplanar görüntüleme olarak adlandırılmaktadır. Kesit planları, sagittal, koronal ve aksiyal eksenlerde olabilmektedir. Vücudun farklı kesit planlarına göre oluşturulan görüntüleri, Şekil 2.10'dan incelenebilir.



Şekil 2.10 Farklı oryantasyona sahip Akciğer CT kesit örnekleri (a) Vücut uzayları, (b) Sagittal kesit, (c) Koronal kesit, (d) Aksiyal kesit (Sluimer, 2005).

CT görüntülerinde dedektörler tarafından elde edilen sayısal değerler, görüntülenen dokunun fiziksel özelliğinin ölçümleridir ve Hounsfield Ünitesi (Hounsfield Unit) (HU) ya da CT değeri olarak adlandırılır. Bu değerler, radyasyon atenuasyon (zayıflama) katsayısı olan  $\mu$  değeri ile ilişkilidir (Sluimer, 2005). Atenuasyon, X ışını demetlerinin hastanın üzerinden geçerken dokuları tarafından soğurulması ve sapması sonucu şiddetinde meydana gelen azalmayı temsil eder (Akgül, 2013).  $\mu$  değeri radyasyonun dalga boyuna ve gücüne göre değişiklik gösterir.



Şekil 2.11 Hounsfield Ölçeği (TÜMRAD-DER, 2013)

HU değeri hesaplanırken suyun standart basınç ve sıcaklıkta X ışını atenüasyon katsayısı 0 olarak kabul edilir ve her bir vokselin değerinin suyun  $\mu$  değerine göre ilişkisi ölçülerek doğrusal dönüşümleri ile HU değerleri belirlenir. HU değerlerine dönüşüm denklemi, Eşitlik 2.1’de verilmiştir. HU değerleri -1000 ile +1000 değerleri arasında değişmektedir. Bu aralık Hounsfield Cetveli ya da Hounsfield Ölçeği olarak tanımlanır (TÜMRAD-DER, 2013) (Şekil 2.11). Hounsfield Ünitesinin değeri sudan yüksek olan yerler 0’dan büyüktür. Kemik, kalsifikasyon (kireçlenme), hematoma gibi. Atenüasyon arttıkça piksel daha beyaz bir renge sahip olur. Yağ ve havada ise HU değerleri 0’dan küçüktür.

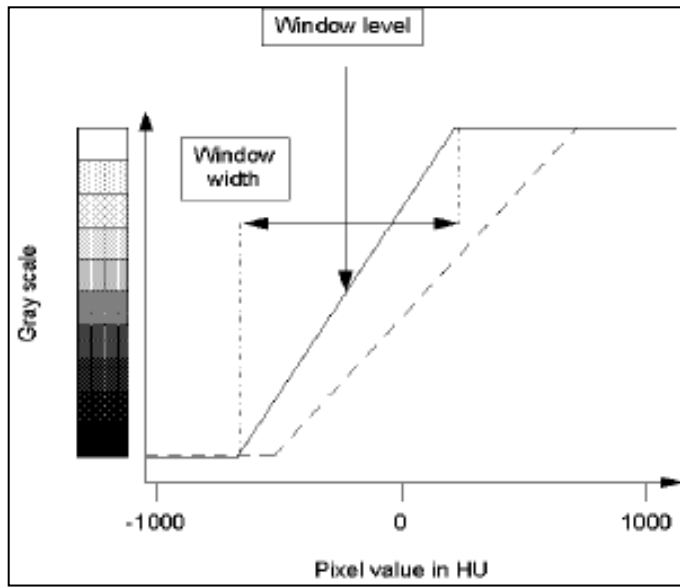
$$\text{CT değeri} = (\mu - \mu_{\text{water}}) / \mu_{\text{water}} \times 1000 \text{ HU} \quad (2.1)$$

### **Pencereleme (Windowing)**

CT görüntüleri oluşturulduğunda siyah ve beyaz renkleri arasında çok sayıda beyaz renk içermektedir. İnsan gözünün algılayabileceği renk sayısı 30-90 arasındadır. CT görüntülerinde odaklanılması gereken yapıya ilişkin HU değerlerinin gözlemlenmesi, görüntülenmek istenmeyen bölgelerin siyah ve beyaz renklere dönüştürülmesi hem bölütleme açısından hem de nodül bölgelerinin daha iyi görüntülenebilmesi için oldukça önem taşımaktadır. Bu kapsamda, CT görüntüleri üzerinde pencereleme yapılarak görüntülenmek istenen en yüksek HU

değerinin üstündeki değerler beyaz, en düşük HU değerinin altında kalan değerler ise siyah yapılır. Görüntülenmek istenen HU aralıkları arasındaki bölgede renk değerleri için ise doğrusal dönüşüm işlemi uygulanarak, siyah beyaz renk aralığına çekilir.

Pencereleme işlemi için 2 temel kavram tanımlanmıştır. Bu kavramlar **pencere genişliği (window width)** ve **pencere düzeyi (window level)**' dir. Şekil 2.12' de pencere genişliği ve düzeyine göre belirlenen HU aralığına göre renk değerlerinin doğrusal dönüşüm grafiği gösterilmiştir.



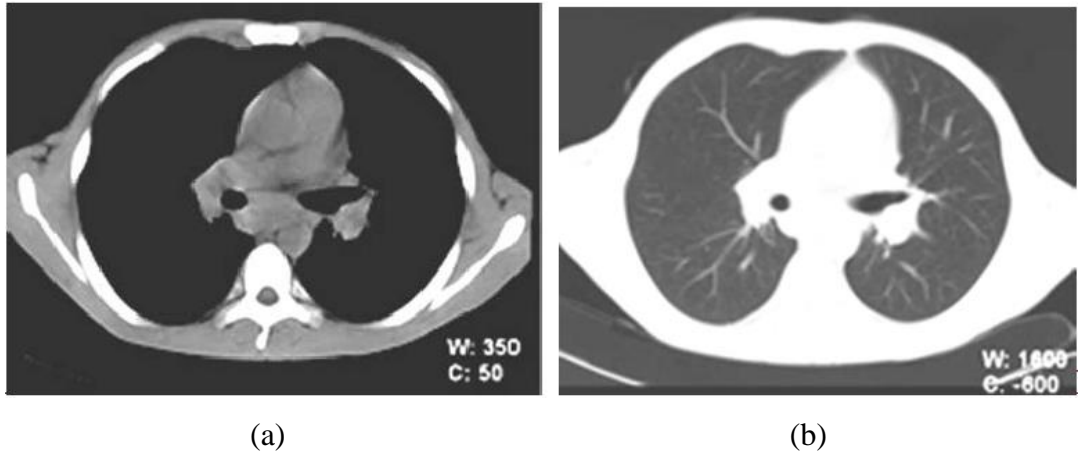
Şekil 2.12 Pencere değerlerine göre renk dönüşüm işlemi (Colombia Center for New Media Teaching and Learning, 2013)

**Pencere genişliği**, seçilen HU değerlerinin en üst ve en alt değerleri arasındaki farktır. Her bir gri renk tonu başına düşen doku yoğunluğunu değiştirmek için kullanılır. Pencere genişliğinin azalması, renk çeşitliliği az olan yapıların daha ayrıntılı olarak görüntülenmesini, farklılıkların açığa çıkarılmasını sağlar. Ayrıca, her bir renk tonu başına düşen doku yoğunluğunun azalması durumunda yüksek kontrastın sağlanması için kullanılmaktadır. Düşük pencere genişliği ile çok çeşitli yapıların görüntülenmesi alan dışında kalarak gözden kaçırılması gibi durumlar da ortaya çıkabilmektedir (Önal, 2008).

Pencere genişliğinin artması durumunda ise görüntülenecek yapıların çeşitliliği artacak, renk değişimleri azalacak ve küçük yapılar arasındaki önemli değişikliklerin farkına varılamayacaktır. Bunun için görüntülenmek istenen yapı veya organın HU değer aralıklarının seçilmesi pencere genişliği açısından uygun olacaktır.

**Pencere düzeyi** ise görüntülenecek gri düzey renk aralığının, seçilen pencere genişliğinin orta noktasını ya da seçilen HU ölçeğinin merkezini temsil eder. Aynı pencere genişliği içerisinde pencere düzeyinin artması HU daha yüksek olan, beyaz renk tonunun fazla olduğu mediasten, toraks gibi yapıların görüntülenmesini sağlarken, yumuşak dokuların farklılıklarının anlaşılması zorlaşacaktır. Pencere düzeyinin azalması durumunda daha düşük HU değer aralığına sahip akciğer, havayolu gibi yapıların daha net görüntülenmesi sağlanacaktır.

Göğüs CT görüntüsünde mediastenın etkili biçimde görüntülenebilmesi için pencere genişliği 350 ve pencere düzeyi +50 olarak belirlenir. Bu değerler ile görüntülenecek HU değer aralıkları  $50 \pm 175$  olup -125 HU ile 225 HU arasındadır. Akciğer bölgesinin görüntülenmesi için pencere düzeyi -400 ile -600 arasında ve pencere genişliği 1500 olarak belirlenir (TÜMRAD-DER, 2013). Pencere genişliği ve düzeyine göre farklı değerlerin alındığı CT görüntüleri Şekil 2.13'te gösterilmiştir.



Şekil 2.13 Farklı pencereleme değerleri (a) Mediasten penceresi (b) Parankima penceresi (TÜMRAD-DER, 2013)

### 2.8.3 Pozitron emisyon tomografi

PET taraması, nükleer tıp görüntüleme yöntemleri arasında en gelişmiş olanıdır. PET ile görüntüleme yapabilmek için pozitron emisyonu yapan radyoizotoplarla bağlanmış olan radyofarmasötikler kullanılır. Radyoaktif bozunmanın bir türü olan pozitron emisyonunda hücre çekirdeğindeki proton, nötron ve pozitif yüklü elektron (pozitron)'a dönüşür. Nükleusdan salınan pozitron, çevredeki bir elektrona çarparak annihilasyon (yok etme) olayına yol açar ve pozitron ve elektronun kütlesi enerjiye dönüşür ( $E=mc^2$ ), iki adet 511 kiloelektronvolt enerjiye sahip foton oluşmasını sağlar. Bu iki foton birbirlerine 180 derece ters yönde hareket ederler. PET cihazlarında bulunan dedektörler yardımı ile aynı anda ve aynı ekseninde oluşmuş olan iki foton algılanarak gelişmiş elektronik yapı ve bilgisayarlar aracılığı ile görüntü haline dönüştürülür (Uğur, 2004).

PET taraması vücuttaki dokuların nasıl çalıştığını ve görüldüğünü, hücrelerin kanserli olup olmadığını, akciğer kanserinin aşamasına göre hangi tedavi seçeneklerinin uygulanacağını gösterir. Bunun için CT'den farklı olarak hastanın kan dolaşımına radyoaktif bir ilaç enjekte edilerek vücutta yayılımı incelenir. En yaygın olarak kullanılan ilaç FDG-18 olarak bilinen glikozun radyoaktif olan türüdür. Kanserli hücreler glikozu enerji için kullanacaklarından vücutta kullanılan yerlere bakılarak kanser teşhisi yapılır (Cancer Research UK, 2013).

### 2.8.4 Manyetik rezonans görüntüleme

Manyetik Rezonans Görüntüleme'de hasta güçlü bir mıknatıs içine konulur. Gönderilen radyo dalgalarının atımlarına karşı vücuttan yayılan radyo dalgalarının gücü ve konumu bilgisayar tarafından belirlenerek hastanın kesit görüntüleri elde edilir (Gonzalez and Woods, 2002). CT görüntülerine göre daha uzun sürebilir. Kanserli hücrelerin beyne ya da omuriliğe yayılıp yayılmadığını belirlemek için kullanılan bir görüntüleme yöntemidir (American Cancer Society, 2013).

## 2.9 Bilgisayar Destekli Tespit Sistemleri

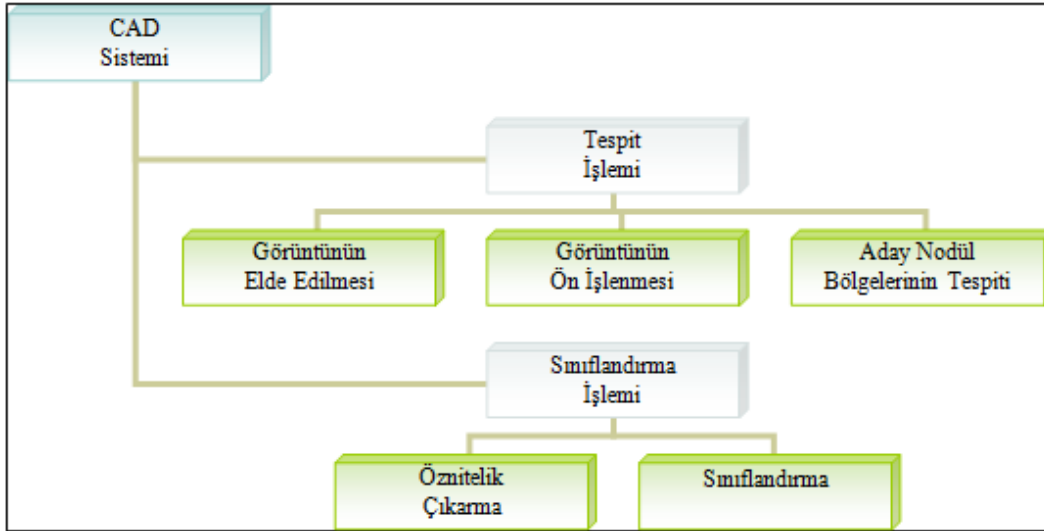
Radyolojide, tıbbi görüntülerde (CT, MRI, PET vb.) yer alan lezyonların bilgisayarlar tarafından otomatik olarak tespiti ve sınıflandırılmasına dayanan

sistemler, bilgisayar destekli tespit (Computer Aided Detection) (CADE) ya da bilgisayar destekli tanılama (Computer Aided Diagnosis) (CADx) sistemleri olarak tanımlanmaktadır. Bilgisayar destekli tespit sistemleri hem görüntü işleme hem de yapay zekâ tekniklerini içeren bir dizi prosedürden, fizik, matematik, istatistik yöntemlerinin etkili bir biçimde uygulanmasından oluşmaktadır.

Teknolojinin ilerlemesiyle birlikte yıllar boyunca artan veri miktarlarıyla orantılı olarak tıbbi görüntülerde de önemli ölçüde artış gözlemlenmiştir. Uzman radyologların, bu görüntüler üzerinden nodül bölgelerini belirlemeleri ve nodülleri yorumlaması oldukça zaman alıcı bir işlem halini almıştır. Ayrıca, doktorların dikkatinden kaçabilen nodüller de olabilmektedir. Nodül konumlarının belirlenmesinde uzman doktorlar arası farklılıkların oluşmasının yanı sıra, akciğer kanserinin erken tanı ve tedavi süreçlerinin gecikmesi gibi etkenlerden dolayı, hastanın 5 yıllık yaşam oranı oldukça azalmaktadır. Bu gibi nedenlerden dolayı, CAD sistemleri tıp alanında oldukça önemli, gerekli ve değerlidir.

CAD sistemlerinin çalışma prensibi, radyologların tıbbi görüntülerden lezyonların çıkarılması, iyi ya da kötü huylu olup olmadıklarının belirlenmesi gibi durumlarda, göstermiş olduğu yaklaşımlara dayanır. CAD sistemlerinin başarısı, günümüzde dünyadaki pek çok hastanede kullanılmasıyla açıklanabilir. Pek çok şirket, geliştirdikleri CAD sistemlerine patent alarak ticarileştirmekte, ürünlerini küresel bir tıp pazarına sunmaktadır (Doi, 2007).

CAD sistemleri genel olarak 2 aşamadan oluşmaktadır. Bu aşamalar, aday nodül bölgelerinin görüntü üzerinde tespiti ve gerçekte nodül olup olmadığının sınıflandırılmasından oluşur. Tespit aşaması, tıbbi görüntüleme kaynaklarından (CT, MRI, PET) görüntünün elde edilmesi, ön işleme yaklaşımlarının uygulanarak görüntünün istenilen duruma getirilmesi ve akciğer bölütleme aşamasını içerir. Sınıflandırma aşamaları ise tespit aşaması sonucunda elde edilen aday bölgelerden özneteliklerin çıkartılması ve sınıflandırıcı yardımıyla gerçek nodül bölgelerinin bulunmasına dayanır (El-Baz et al., 2011; Mughal and Ikram, 2004). Genel bir CAD sisteminin yapısı Şekil 2.14'te gösterilmiştir.



Şekil 2.14 Akciğer nodüllerinin tespiti için genel bir CAD sisteminin aşamaları (Mughal and Ikram, 2004).

## 2.10 Bilgisayar Destekli Tespit Sisteminin Yararları

CAD sistemleri tıp alanında akciğer, göğüs, kolon kanserleri gibi durumlarda kullanımı oldukça yaygın ve önemlidir. CAD sistemlerinin sonuçlarının ikinci bir çıktı olarak uzman radyologlara sunulması sayesinde, radyologların kanserin erken tanı, tedavi ile lezyon hakkında son kararı verme süreçlerinde yardımcı olunması sağlanacaktır.

Radyologların sonuçları arasındaki farklılıklardan dolayı aynı görüntülerin çok sayıda değerlendirilmesi gerekebilmektedir (National Cancer Institute, 2013). CAD sistemleri ile radyologların iş yükü azaltılarak, tıbbi görüntülerin 2 ya da 4 kez okunmasının önüne geçilecek, daha hızlı tanılama işlemlerinin yapılmasına yol açacaktır. CAD sisteminin ürettiği çıktıların doktorun belirttiklerinden daha iyi bir sonuç üretmesi yerine, doktorun bilgileriyle bütünleşerek hem insan hem de bilgisayar özelliklerinin kullanılarak tanılama kalitesinin artırılması amaçlanmaktadır. Bu nedenle, CAD sistemleri akciğer kanserinin erken tespitinde başarı ve performans oranını artırması, gereksiz biyopsilerin, girişimsel yaklaşımların önüne geçmesi ve tedavi başarısının yükselmesinde önemli bir rol oynamaktadır.



## 2.11 Bilgisayar Destekli Tespit Sistemlerine İlişkin Problemler

CAD sistemleri zaman kazancı sağlaması, kanserin erken tanı ve tedavisine yardımcı olması gibi yararlarının yanı sıra dikkate alınması gereken birtakım durumları da meydana getirmektedir. Radyologların sonuçlarında farklılıklarının olması gibi CAD sistemlerinin sonuçlarında da aday nodüllerin tespitinde hatalar çıkabilmektedir. CAD sistemlerinin sonuçlarında, gerçekte nodül olması gereken bölgeyi nodül olarak belirlememesi (FN) (False Negative) ya da nodül olmayacak bölgeyi nodül olarak belirlemesi (FP) (False Positive) gibi durumlardan dolayı uzman doktorların karar verme süreçleri zorlaşabilmektedir. Bu nedenle, CAD sistemlerinin ürettiği sonuçlar alanında uzman doktorlar tarafından doğrulanmalıdır. Ayrıca, CAD sistemlerinin çıktılarının başarımını ölçebilmek amacıyla sınıflandırma doğruluk oranı, FP oranı, hız, otomatikleştirme düzeyi, duyarlılık ve özgüllük değerleri gibi ölçütlerin kullanılması gerekmektedir.

## 2.12 Özet

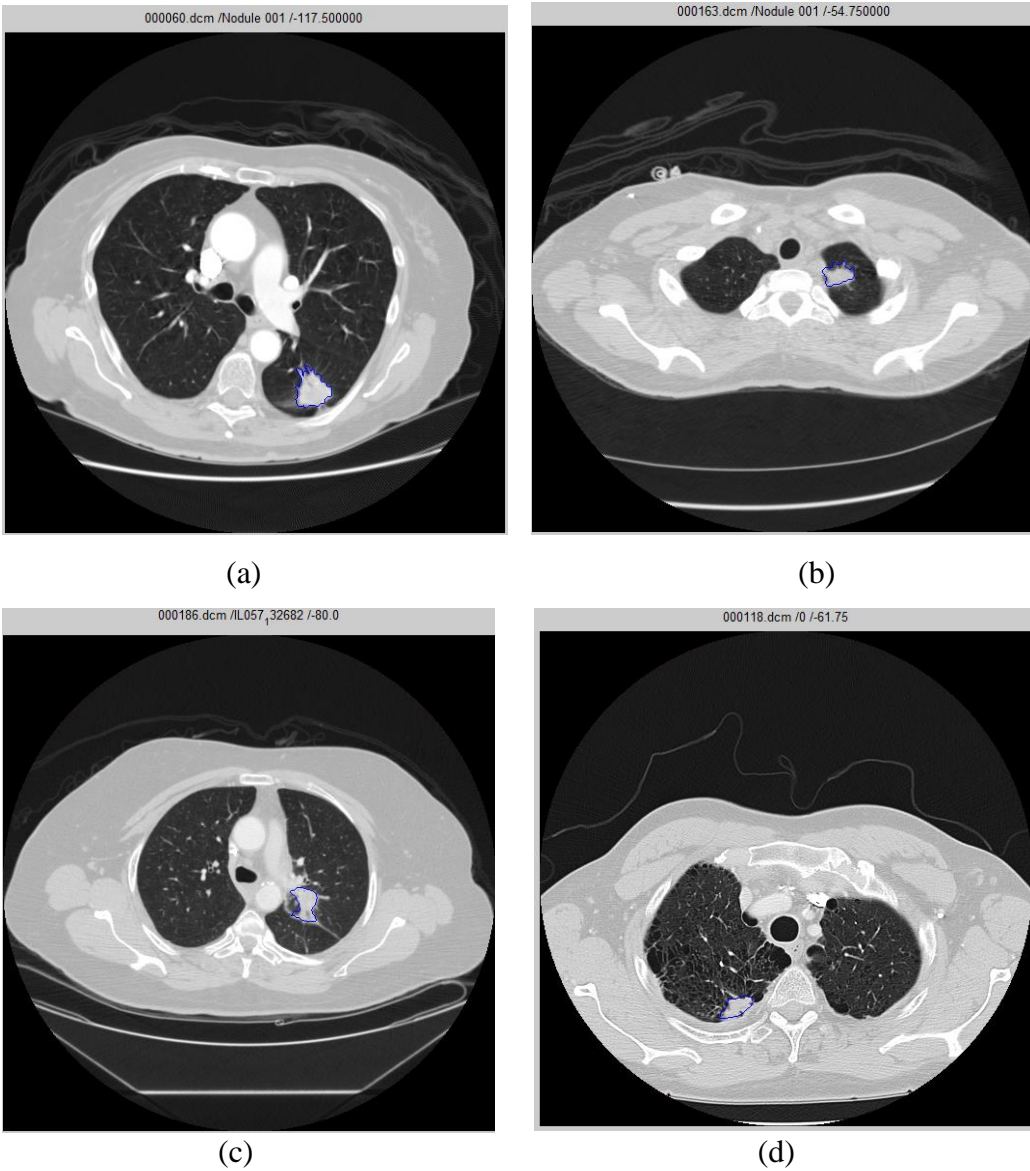
Akciğer kanseri günümüzde, dünyada en yüksek ölüm oranına sahip olan kanserlerden biri olmaya devam etmektedir. Akciğer kanserinden kaynaklanan ölüm oranının azaltılması için hastalığın erken tanı ve tedavisi oldukça önem taşımaktadır. Küçük hücreli dışı akciğer kanserinde, kanserli hücrelerin diğer doku ve organlara yayılmadan Evre I'de tespit edilmesi hastaların 5 yıllık hayatta kalım oranlarını önemli ölçüde artırmaktadır.

Uzman radyologlar tarafından en çok tercih edilen akciğer kanseri görüntüleme yöntemi CT'dir. CT diğer görüntüleme yöntemlerine kıyasla daha duyarlı ve ayrıntılı görüntüleme olanağı sağlamaktadır. Radyologların CT görüntülerinde akciğer kanseri olma özelliği gösterecek nodül bölgelerini tespit edebilmesi, artan veri ve görüntü miktarlarına bağlı olarak giderek zorlaşmakta ve uzman doktorlar arasındaki sonuçların farklılıklarından dolayı ciddi zaman kaybına neden olmaktadır. Bu ve benzeri durumlarda CAD sistemlerine büyük bir ihtiyaç duyulmaktadır. Günümüzde CAD sistemlerine ilişkin hem bilimsel hem de endüstriyel olmak üzere çok sayıda çalışmalar yapılmakta, geliştirilen sistemlerin

performansları artırılmaya çalışılmaktadır. CAD sistemlerinin oluşturduğu sonuçların ikinci bir çıktı olarak doktorlara sunulması sayesinde, onlara yardımcı olarak olası problemlerin önüne geçilmiş olunacak, kanserin erken tanı ve teşhisi sayesinde hastaların ölüm oranı azalacaktır.

### 2.13 Akciğer Nodüllerini İçeren Örnek CT Görüntüleri

Şekil 2.15'te hastaların akciğer nodüllerini içeren örnek CT görüntüleri gösterilmektedir (Şekil 2.15). Görüntüler üzerinde uzman doktorların belirlemiş oldukları referans sınırları da çizdirilmiştir. CT görüntüleri LIDC (Lung Image Database Consortium, 2003) veri setinden elde edilmiştir.



Şekil 2.15 Örnek akciğer nodülü CT görüntüleri (a), (b), (c) ve (d) Juxtepleural nodül örnekleri (LIDC, 2012).

### 3. DICOM GÖRÜNTÜ TÜRÜNÜN İNCELENMESİ

Bu bölümde tez kapsamında kullanılan DICOM görüntü türü ile temel kavramlar açıklanmış, DICOM görüntü standardının özellikleri hakkında bilgiler verilmiştir.

#### 3.1 DICOM Standardı

DICOM (Digital Imaging and Communication in Medicine), tıbbi görüntülerin sayısal olarak depolanması, işlenmesi, iletilmesi ve yazdırılması için National Electrical Manufacturers Association (NEMA) ve American College of Radiology (ACR) tarafından geliştirilmiş bir standarttır. Aynı zamanda tıbbi görüntülerin bilgisayarlar üzerinde iletişime geçmesini sağlayan TCP/IP iletişim protokolünü içerir. Bu görüntü formatına sahip farklı yerlerde bulunan ve farklı donanımlara sahip olan hastanelerdeki doktorlar, bilgisayarları aracılığıyla DICOM protokolüne uygun olarak hızlı ve kolay bir biçimde iletişime geçebilirler. Ayrıca, tıbbi cihaz üreticilerinin aynı görüntü formatı ile ortak bir platform üzerinde çalışacak aygıtları üretmesi ve tıbbi görüntüyü işleyecek cihazın zorunlu tuttuğu yazılımlara bağlı kalmamak açısından önemli derecede bir rekabet sağlamaktadır (Ulaş ve Boyacı, 2007).

DICOM, günümüzde dünyada hemen hemen her hastanede tıbbi görüntüleme standardı olarak kullanılmaktadır. Hem görüntü kalitesinin yüksek olması hem de kendi içinde hastaya, hastaneye, görüntüleme yöntemine ilişkin çeşitli bilgiler bulundurması nedeniyle doktorlar tarafından özellikle tercih edilmektedir.

#### 3.2 DICOM Yapısı

DICOM dosyası iki kısımdan oluşmaktadır. İlk kısım metinsel verilerin kaydedildiği bölüm olup, hastanın adı, yaşı, cinsiyeti, doğum tarihi, hastahanesinin, doktorun adı, tarihi, tıbbi görüntüleme yöntemi, kesit kalınlığı, kesitin konumu gibi özelliklerin tutulduğu yapıların yanı sıra ön ek gibi DICOM dosyasına özgü alanlar da bulunmaktadır.

DICOM dosyasındaki ikinci kısımda görüntü bilgileri yer almaktadır. Görüntü bilgileri yalnızca “Pixel Data” adı verilen alanda ikili veriler halinde tutulmaktadır. Bu nedenle, DICOM dosyaları, hem metinsel hem de ikili verileri barındırmakta ve özelliği itibarıyla veritabanına benzeyen bir işlevselliğe sahiptir. LIDC veri setinden alınan akciğer CT görüntülerindeki DICOM dosyasında yer alan etiketler ve bu etiketlere ilişkin bilgiler Şekil 3.1’de gösterilmiştir. DICOM dosyalarının genel yapısı Çizelge 3.1’den incelenebilir.

Tag ID	VR	V	L...	Description	Value
(0002,0000)	UL	1	4	Group Length	194
(0002,0001)	OB	1	2	File Meta Information Version	00 01
(0002,0002)	UI	1	26	Media Storage SOP Class UID	1.2.840.10008.5.1.4.1.1.2
(0002,0003)	UI	1	64	Media Storage SOP Instance UID	1.3.6.1.4.1.14519.5.2.1.6
(0002,0010)	UI	1	20	Transfer Syntax UID	1.2.840.10008.1.2.1
(0002,0012)	UI	1	16	Implementation Class UID	1.2.40.0.13.1.1
(0002,0013)	SH	1	14	Implementation Version Name	dm4che-1.4.27
(0008,0005)	CS	1	10	Specific Character Set	ISO_IR 100
(0008,0008)	CS	3	22	Image Type	ORIGINAL\PRIMARY\AXIAL
(0008,0016)	UI	1	26	SOP Class UID	1.2.840.10008.5.1.4.1.1.2
(0008,0018)	UI	1	64	SOP Instance UID	1.3.6.1.4.1.14519.5.2.1.6
(0008,0020)	DA	1	8	Study Date	20000101
(0008,0021)	DA	1	8	Series Date	20000101
(0008,0022)	DA	1	8	Acquisition Date	20000101
(0008,0023)	DA	1	8	Content Date	20000101
(0008,0024)	DA	1	8	Overlay Date	20000101
(0008,0025)	DA	1	8	Curve Date	20000101
(0008,00...	DT	1	8	Acquisition Datetime	20000101
(0008,0030)	TM	0	0	Study Time	
(0008,0032)	TM	0	0	Acquisition Time	
(0008,0033)	TM	0	0	Content Time	
(0008,0050)	SH	1	16	Accession Number	2819497684894126
(0008,0060)	CS	1	2	Modality	CT
(0008,0070)	LO	1	18	Manufacturer	GE MEDICAL SYSTEMS
(0008,0090)	PN	1	2	Referring Physician's Name	
(0008,1090)	LO	1	16	Manufacturer's Model Name	LightSpeed Plus
(0008,1155)	UI	1	64	Referenced SOP Instance UID	1.3.6.1.4.1.14519.5.2.1.6
(0010,0010)	PN	1	2	Patient's Name	
(0010,0020)	LO	1	14	Patient ID	LIDC-IDRI-0001
(0010,0030)	DA	0	0	Patient's Birth Date	
(0010,0040)	CS	0	0	Patient's Sex	
(0010,1010)	AS	0	0	Patient's Age	
(0010,21...	DA	1	8	Last Menstrual Date	20000101
(0012,0062)	CS	1	4	Unknown	YES
(0012,0063)	LO	1	46	Unknown	DCM:113100/113105/1131
(0013,0010)	LO	1	4	Unknown	CTP

Tag ID	VR	V	L...	Description	Value
(0013,1010)	UN	1	10	Unknown	LIDC-IDRI
(0013,1013)	UN	1	8	Unknown	62796001
(0018,0010)	LO	1	2	Contrast/Bolus Agent	IV
(0018,0015)	CS	1	6	Body Part Examined	CHEST
(0018,0022)	CS	1	12	Scan Options	HELICAL MODE
(0018,0050)	DS	1	8	Slice Thickness	2.500000
(0018,0060)	DS	1	4	KVP	120
(0018,0090)	DS	1	10	Data Collection Diameter	500.000000
(0018,1020)	LO	1	26	Software Version(s)	LightSpeedApps2.4
(0018,1100)	DS	1	10	Reconstruction Diameter	360.000000
(0018,1110)	DS	1	10	Distance Source to Detector	949.075012
(0018,1111)	DS	1	10	Distance Source to Patient	541.000000
(0018,1120)	DS	1	8	Gantry/Detector Tilt	0.000000
(0018,1130)	DS	1	10	Table Height	144.399994
(0018,1140)	CS	1	2	Rotation Direction	CW
(0018,1150)	IS	1	4	Exposure Time	570
(0018,1151)	IS	1	4	X-ray Tube Current	400
(0018,1152)	IS	1	4	Exposure	4684
(0018,1160)	SH	1	12	Filter Type	BODY FILTER
(0018,1170)	IS	1	6	Generator Power	48000
(0018,1190)	DS	1	8	Focal Spot(s)	1.200000
(0018,1210)	SH	1	8	Convolution Kernel	STANDARD
(0018,5100)	CS	1	4	Patient Position	FFS
(0020,00...	UI	1	64	Study Instance UID	1.3.6.1.4.1.14519.
(0020,000E)	UI	1	64	Series Instance UID	1.3.6.1.4.1.14519.
(0020,0010)	SH	0	0	Study ID	
(0020,0011)	IS	1	8	Series Number	3000566
(0020,0013)	IS	1	2	Instance Number	42
(0020,0032)	DS	3	36	Image Position (Patient)	-166.000000\ -171.4
(0020,0037)	DS	6	54	Image Orientation (Patient)	1.000000\0.000000
(0020,0052)	UI	1	64	Frame of Reference UID	1.3.6.1.4.1.14519.
(0020,0200)	UI	1	64	Synchronization Frame of Ref...	1.3.6.1.4.1.14519.
(0020,1040)	LO	1	2	Position Reference Indicator	SN
(0020,1041)	DS	1	12	Slice Location	-112.500000
(0028,0002)	US	1	2	Samples per Pixel	1
(0028,0004)	CS	1	12	Photometric Interpretation	MONOCHROME2

(0028,0010)	US	1	2	Rows	512
(0028,0011)	US	1	2	Columns	512
(0028,0030)	DS	2	18	Pixel Spacing	0.703125\0.703125
(0028,0100)	US	1	2	Bits Allocated	16
(0028,0101)	US	1	2	Bits Stored	16
(0028,0102)	US	1	2	High Bit	15
(0028,0103)	US	1	2	Pixel Representation	1
(0028,0120)	US	1	2	Pixel Padding Value	63536
(0028,0303)	SH	1	8	Unknown	MODIFIED
(0028,1050)	DS	1	4	Window Center	-600
(0028,1051)	DS	1	4	Window Width	1600
(0028,1052)	DS	1	6	Rescale Intercept	-1024
(0028,1053)	DS	1	2	Rescale Slope	1
(0038,0020)	DA	1	8	Admitting Date	20000101
(0040,0002)	DA	1	8	Scheduled Procedure Step Sta...	20000101
(0040,0004)	DA	1	8	Scheduled Procedure Step End...	20000101
(0040,0244)	DA	1	8	Performed Procedure Step Sta...	20000101
(0040,2016)	LO	0	0	Placer Order Number / Imagin...	
(0040,2017)	LO	0	0	Filler Order Number / Imaging ...	
(0040,A0...	PN	1	14	Verifying Observer Name	Removed by CTP
(0040,A1...	PN	1	14	Person Name	Removed by CTP
(0040,A1...	UI	1	64	UID	1.3.6.1.4.1.14519.5
(0070,0084)	PN	1	2	Content Creator's Name	
(0088,0140)	UI	1	64	Storage Media File-set UID	1.3.6.1.4.1.14519.5
(3006,0024)	UI	1	64	Referenced Frame of Referen...	1.3.6.1.4.1.14519.5
(3006,00...	UI	1	64	Related Frame of Reference UID	1.3.6.1.4.1.14519.5
(7FE0,0010)	O..	1	5...	Pixel Data	00 FC 00 FC 00 FC 0

Şekil 3.1 DICOM yapısı, etiketleri ve değerleri (LIDC, 2012)

Çizelge 3.1 DICOM Dosya Yapısı (Ulaş ve Boyacı, 2007).

<b>Önsöz - Preamble</b>	İlk 128 byte açıklama alanıdır ve boştur.
<b>Önek - Prefix</b>	Dosyanın DICOM dosyası olduğunu belirtir.
<b>Veri elemanları Data Elements</b>	DICOM etiketlerini ve bu etiketlere ilişkin verileri tutar.
<b>Piksel Verisi Pixel Data</b>	Resim Bilgisinin ikili veri olarak saklandığı alandır.

### 3.3 Piksel Değerlerinden HU Birimine Dönüşüm

DICOM dosyalarında “Pixel Data” alanında tutulan veriler ile görüntüleme yapıldığında, oluşan görüntüler görüntüleme aygıtı tarafından anlaşılır biçimde yani piksel türündedir. Tıbbi cihazlara ve görüntüleme yöntemine göre belirlenen birimler ise farklılık gösterebilmektedir. Örneğin, CT görüntülerinde dokuların radyolojik yoğunluğunu ölçmek amacıyla HU kullanılmaktadır. Ayrıca, görüntülenmek istenen organ üzerinde pencereleme işlemi yapılarak HU değer aralıklarına uygun bir görüntülenme sağlanması, hem bölütleme işlemi hem de ilgilenilen yapıya odaklanması açısından önem taşımaktadır. Bu nedenle, DICOM dosyalarında piksel değerleri üzerinde görüntülenecek birime dönüşüm yapmak için bazı alanlar konulmuştur.

DICOM etiketleri içerisinde yer alan “Rescale Slope” ve “Rescale Intercept” alanları ile DICOM dosyalarında saklanan piksel değerlerinden HU birimine doğrusal dönüşüm işlemi gerçekleştirilmektedir (NEMA, 1999). Bu dönüşüm Eşitlik 3.1’de gösterilmiştir. Doğrusal olmayan dönüşüm işlemlerinde ise “Lookup Table” üzerinden değerlere erişilebilmektedir.

$$\text{CT değeri} = m * SV + b ; \quad (3.1)$$

**m** = Rescale Slope

**SV** = Stored Values

**b** = Rescale Intercept olmak üzere

#### 4. ÖNCEKİ ÇALIŞMALAR

Bu bölümde, akciğer kanseri üzerinde yapılan CAD çalışmalarında kullanılan yöntemler ana hatlarıyla anlatılmaktadır. Çalışmalarda yararlanılan veri seti, sağlanan başarı ölçüsü, ele alınan nodül türü gibi konular hakkında bilgi verilmektedir.

Akciğer nodüllerinin bilgisayar destekli tespiti için çok sayıda çalışma yapılmıştır. Bu çalışmaların bir kısmı bilimsel yayınları oluştururken, diğer bir kısmı ise ticari değere sahip olan ve günümüzde hemen hemen her hastanede kullanılan CAD sistemlerinin patentlerinden oluşmaktadır. Akciğer kanserinin hem dünyada en çok ölüm oranına sahip kanser türü olması hem de erken tanı, tedavi süreçlerinin bu hastalıktan kaynaklanan ölüm oranının azaltılmasında önemli bir rol oynaması nedeniyle, bilim insanları ve sanayi kuruluşları için yıllardır üzerinde çalışılan önemli bir alan olmuştur.

Akciğer kanserinde CAD sistemleri için yapılan çalışmalar farklı tıbbi görüntüleme yöntemlerini içermektedir. Literatürde önerilen CAD sistemlerinin değerlendirilmesi güvenilirlik, hız, tanılama doğruluğu, otomasyon düzeyi gibi etkenlerle belirlenmektedir (El-Baz et al., 2013). CT'nin en duyarlı biçimde görüntüleme imkânı sunması, düşük kontrasta sahip yapıları görüntüleyebilmesi, farklı kesit planlarında görüntüleri elde edebilmesi bakımından özellikle tercih edilmektedir. CT görüntüleri üzerinde CAD sistemleri ile akciğer nodülü tespiti için 1989 yılından beri çalışmalar yapılmaktadır (Preteux, 1989). Bu çalışmalar otomatik ve yarı otomatik yöntemlerden oluşmaktadır (Lee et al., 2012).

Akciğer kanserinde kullanılan CAD sistemleri genellikle 4 ana bölümden oluşmaktadır. Bu aşamalar, görüntünün elde edilmesi, ön işleme aşaması ve akciğer bölgelerinin bölütlenmesi, aday nodül bölgelerinin tespiti ve FP oranının indirgenmesidir. Her bir aşama için literatürde çeşitli yöntem ve yaklaşımlar geliştirilmiştir. Bu yöntemlerden ilgili ve önemli olanları bu bölümde incelenmektedir. Alt bölümlerde her bir aşama için sunulan yaklaşımlar açıklanmıştır.

#### 4.1 Görüntünün Elde Edilmesi

Görüntünün elde edilmesi aşaması tıbbi görüntüleme yöntemlerinden akciğer nodülü içeren görüntülerin elde edilmesine dayanır. Bunun için akciğer radyografileri, MRI, CT, PET, CT-PET gibi görüntüleme yöntemleri kullanılmaktadır. CT görüntülerinin daha doğru tanılama sonucu sağlamasından dolayı bu alanda çalışanlar tarafından özellikle tercih edilen bir tıbbi görüntüleme yöntemidir.

Bazı kurumlar, organizasyonlar ya da araştırma merkezleri tarafından hastalardan alınan akciğer CT görüntüleri toplanıp veri setleri oluşturulmakta, bilim insanlarının hizmetine sunulmaktadır. Bunun için akciğer CT görüntülerinden oluşan LIDC (Lung Image Database Consortium, 2003), ELCAP (Cornell University, 2007), Public Database to Address Drug Response (Cornell University, 2008), NLST (National Lung Screening Tutorial, 2009) gibi veri setlerinin dışında, X ışınları görüntüleme yönteminin kullanıldığı JSRT (Japanese Society of Radiological Technology, 1998) gibi veri seti de bulunmaktadır. Ayrıca, araştırmacıların işbirliği içerisinde oldukları hastane ve doktorlar tarafından sağlanan akciğer CT görüntüleri ve bu görüntüler üzerindeki kesin referans (Ground Truth) (GT) değerlerine göre de yapılan çalışmalar bulunmaktadır.

#### 4.2 Önışleme Aşaması

Bu aşama, akciğer nodüllerini içeren görüntülerin nodül tespiti yapılmadan önce kalitesinin artırılarak bilgisayarda işlenmeye uygun ve doktorlar tarafından daha anlaşılır duruma getirilmesi için yapılan bir dizi işleme dayanır. Bu işlemler, temel görüntü işleme teknikleri olan filtreleme, gürültülerin ve istenmeyen bölgelerin giderilmesi ile açma, kapama, aşındırma ve genişletme gibi yöntemleri içeren morfolojik işlemlerdir. Özellikle kullanılan filtrelerde, nodül biçimindeki yapıları daha belirgin hale getirmek ve diğer yapıların etkisini azaltmak amaçlanmıştır.

Akciğer tomografi görüntülerinde düzgün sınırlı nodüllerin tespiti için matematiksel morfolojiye dayanan **Quoit (Q) filtresi Yamamoto et al. (1996)** tarafından uygulanmıştır. Bu çalışmada Q filtresi, Q transform ve Q-inverse transform olmak üzere iki işlemten oluşmaktadır. Q-transform işlemi ile matched filtresine benzer biçimde görüntülerden öznelilikler çıkartılmış ve Q inverse transform işlemi ile de çıkartılan öznelilik parametrelerine göre nodül görüntüsü kurtarılmıştır. Küresel biçimdeki şekilleri iyileştirmek için Hessian matrisine dayanan **“blob-enhancement filter” Frangi et al. (1999)** tarafından geliştirilmiştir. **Arimura et al. (2004)** tarafından yapılan bir çalışmada **matched filtresi** ile iyileştirilmiş nodül yapılarını içeren görüntülerden, **ring average filtresi** kullanılarak nodül yapılarının bastırıldığı görüntülerin çıkarılması ile aday nodül bölgeleri tespit edilmeye çalışılmıştır. **Matsumoto et al. (2006)** tarafından **quantized convergence filtresini** uygulayarak yuvarlak biçimdeki nodül vb. yapıların iyileştirilmesi sağlanmıştır. Bu filtrenin akciğer nodülünün doğru tespit oranını artırmasının dışında, ayırt edici öznelilikleri çıkartma sürecinde de başarılı sonuçlar alındığı belirtilmiştir. **Ochs et al. (2007)** ve **Paik et al. (2004)** çalışmalarında **küresellik yapı iyileştirme filtresi (sphericity structure enhancement filter)** uygulanarak CT görüntülerinde nodüllerin iyileştirilmesi sağlanmıştır (**Lee et al., 2012**).

**Retico et al. (2008)** tarafından küresel biçimdeki yapıların iyileştirilmesi ve nodül bölgelerinin yanlış tespit oranının azaltılması için **nokta iyileştirme filtresi (dot enhancement filter)** uygulanmıştır. Bu filtrenin duyarlılığının artırılması için ise farklı ölçeklerde Gaussian yumuşatma işlemi yapılarak, çapı 5 mm’i aşan nodülleri belirlemede başarı oranı artırılmıştır. **Li et al., (2008)** tarafından, eş zamanlı olarak hem küresel yapıya sahip nodüllerin iyileştirilmesi hem de kan damarları, hava yolu duvarı gibi anatomik yapıların bastırılması için görüntülere **seçici iyileştirme filtresi (selective enhancement filter)** uygulanmıştır. Bu filtre ile yüksek duyarlılık ve düşük FP oranına sahip sonuçlar elde edilmiş, yüksek performans oranına sahip nodül tespit sistemi geliştirilmiştir.

Yapılan bir başka çalışmada ise **3D iris filtresi** uygulanmıştır (**Suárez-Cuenca et al., 2009**). 3D iris filtresi nodül bölgelerindeki renk yoğunluklarına göre gradyan vektörlerinin elde edilmesine dayanır. Nodülün merkezine işaret



eden gradyan vektörlerinin artması filtrenin daha yüksek bir çıktı değerine sahip olmasına ve nodülün iyileşmesine yol açmaktadır. Iris filtresinin kullanılmasıyla sistemin duyarlılık oranı aynı tutularak FP oranının indirgenmesi sağlanmıştır.

### 4.3 Akciğer Bölütleme

Akciğer bölütleme aşaması, görüntüden akciğer bölgesinin ayrılarak diğer kısımların görüntüden çıkarılması için yapılan birtakım işlemleri içermektedir. Sağlıklı akciğer dokuları, akciğerin içinde ve etrafında yer alan yapılara göre daha karanlık bölgeleri oluşturmaktadır. Akciğer dokularının bu özelliği akciğerlerin bölütlenmesi için kullanılan en temel yöntemlerden biridir. Düzgün sınırlı veya damara yakın nodüllerin bölütlenmesinde bir problem yaşanmazken plevraya yakın, bitişik ya da bağlantısı bulunan nodüllerin renk, HU yoğunluklarının plevranın yoğunluğu ile benzer olmasından kaynaklanan problemler ortaya çıkabilmektedir. Bu tarz durumlarda farklı bölütleme yaklaşımlarına ihtiyaç duyulmaktadır. Ayrıca, akciğerler içerisindeki damarlar, bronş ve bronşiyoller gibi yapıların doğru ve etkili biçimde bölünmesi de önem taşımaktadır. Bu kapsamda eşikleme (thresholding), çizge tabanlı vb. yaklaşımlar ile bölütleme işlemi gerçekleştirilmeye çalışılmıştır.

Akciğer bölütlemeye gri düzey eşikleme tabanlı yaklaşımlara örnek çalışmalar **Armato et al. (1999)**, **Hu et al. (2001)**, **Ko and Betke (2001)** tarafından kullanılmıştır. **Armato et al. (1999)**'ın çalışmasında gri düzey eşiklemeye ek olarak rolling ball filtresi bölütlenen akciğer görüntüleri üzerinde uygulanırken, **Ko and Betke (2001)**'in çalışmasında akciğer sınırlarının eğrisellik değerlerindeki değişikliğin yüksek olduğu yerlerde sınır düzeltme işlemi yapılmış, juxtapleural nodüllerin belirlenebilmesi sağlanmıştır. **Hu et al. (2001)**'in çalışmasında ise optimal eşikleme yöntemi uygulanmış, eşikleme sonucu elde edilen akciğer bölgeleri dinamik programlama ile sol ve sağ akciğer bölgelerine ayrılmış, akciğer bölgelerindeki düzgün olmayan sınırlar bir dizi morfolojik işlemler uygulanarak giderilmiştir. **Antonelli et al. (2005)**'in çalışmasında da temel görüntü işleme teknikleri olan eşikleme, morfolojik açma - kapama, sınır tespiti, inceltme (thinning), sınırların yeniden oluşturulması (border reconstruction) ve bölge dolgulama (region filling) yöntemi kullanılmıştır. Diğer

bir çalışmada ise Otsu eşikleme yöntemi kullanılmıştır (**Ross et al., 2009; El-Baz et al., 2013**).

**Armato et al. (1999)**'in kullanmış olduğu eşikleme yöntemine dayanan bölge dolgulama yaklaşımı **Pu et al. (2009)** tarafından, akciğerdeki çatlakların belirlenmesi işleminde arama uzayının boyutunu azaltmak için kullanılmıştır. **Pu et al. (2008)** tarafından yapılan bir başka çalışmada ise adım uzunluğuna göre akciğer sınırları üzerinde dolaşılmasını sağlayan adaptive border marching yöntemi kullanılmış, juxtaopleural nodüllerin akciğer bölütlemeye dikkate alınması sağlanmıştır. Bu çalışmada akciğer içerisindeki yapıların bir kısmının bölütlenememesi (undersegmentation), akciğer dışındaki yapılarda ise olması gerekenden daha fazla bölütlenmesi (oversegmentation) gibi problemleri çözmek amaçlanmıştır.

Gri düzey eşikleme ve piksel tabanlı yaklaşımların dışında aktif şekil modeli (active shape model) (ASM) (snakes), aktif görünüm modeli (active appearance model) (AAM), biçim değiştirebilir model (deformable model), level set gibi farklı yaklaşımlar da kullanılmaktadır (**El-Baz et al., 2013**). Popülasyon tabanlı, hastaya özgü şekil istatistiklerini ve değiştirilmiş SIFT (Scale Invariant Feature Transform) tanımlayıcısını kullanan, biçim değiştirilebilir modelden oluşan akciğer bölütleme sistemi **Shi et al. (2008)** tarafından gerçekleştirilmiştir. **Van Ginneken et al. (2006)** tarafından yapılan bir çalışmada aktif şekil modeli, aktif görünüm modeli, çok çözünürlüklü piksel sınıflandırma yöntemi akciğer bölütlenmesinde karşılaştırılmış, çok çözünürlüklü piksel sınıflandırılmasının veya üç yöntemin birleşiminin daha iyi sonuçlar verdiği belirtilmiştir. Ayrıca 3D bölge büyütme tabanlı yarı otomatik bir yaklaşım **Diciotti et al. (2008)** tarafından, robust active shape model (RASM) tabanlı bir yaklaşım da **Sun et al. (2012)** tarafından geliştirilmiştir.

Yapılan çalışmaların dışında **Sensakovic et al. (2008)** tarafından, tıbbi görüntülerin bölütlenmesi sırasında ortaya çıkan içbükey bölgeler arasından uygun olmayanların tespiti ve onarılması için hiyerarşik alpha hull yöntemi önerilmiş ve rolling ball filtresi, morfolojik kapama işlemi ile önerilen yöntem karşılaştırılmıştır. Ayrıca, **Nunzio et al. (2011)** tarafından yapılan çalışmada ise

morfolojik kapama ve alpha hull yöntemlerinin juxtaopleural nodüllerin tespitindeki başarı oranı karşılaştırılmıştır. Alpha hull yönteminde yanlış pozitiflerin sayısını azaltmak amacıyla geometrik ve desensel öznitelikler çıkartılarak YSA (Yapay Sinir Ağı)'na verilmiştir.

#### 4.4 Akciğer Nodül Tespiti

Akciğer nodül tespit aşaması, akciğer üzerinde nodülün yer alıp almadığının eğer varsa hangi türde olduğunun tespit edilmesini içermektedir. (Lee et al., 2012). Akciğer nodül tespiti, yoğunluklarına dayalı ya da model tabanlı tespit yöntemlerine göre sınıflandırılmaktadır. Yoğunluk değerlerine dayalı nodül tespit yöntemi akciğer nodüllerinin, akciğer dokusundan daha yüksek yoğunluğa sahip olduğu fikrinden yola çıkmaktadır. Literatürde, eşikleme yöntemleri (Armato et al., 2001), kümeleme algoritmaları (Gurcan et al., 2002), YSA (Suzuki et al., 2003), matematiksel morfoloji (Awai et al., 2004) gibi yöntemler ile yoğunluk değerleri dikkate alınarak akciğer nodülleri tespit edilmeye çalışılmıştır (Ye et al., 2008). 3D küresel ve silindirik filtreler kullanılarak bu tür yapılara benzer nodüllerin tespitine ilişkin çalışmalarda bulunmaktadır (Takizawa et al., 2003; Li et al., 2004; El-Baz et al., 2013). Li et al. (2008) tarafından yapılan çalışmada seçici iyileştirme filtresi uygulanması sonucunda nodülün iyileştirildiği görüntülerden 3D bölge büyütme algoritması kullanılarak nodül bölgelerinin bölütlenmesi sağlanmış, otomatik kural tabanlı sınıflandırıcı yardımıyla yanlış pozitifler giderilmiştir.

Model tabanlı tespit yöntemlerinde ise şablon eşleme (Lee et al., 2001; Farag et al., 2004), nesne tabanlı deformasyon (object based deformation) (Ge et al., 2005), anatomi tabanlı genel model (anatomy based generic model) (Brown et al., 2001) gibi yöntemler kullanılmış, küresel yapıdaki nodüllerin tespiti sağlanmıştır (Ye et al., 2008). (Lee et al., 2001) tarafından yapılan çalışmada genetik algoritma tabanlı şablon eşleme ile akciğer nodülleri %72'lik doğruluk oranı ile tespit edilmiştir. Ayrıca, volumetrik şekil indeksi ve çok ölçekli noktasal iyileştirme filtresinin kullanıldığı bir çalışmada çıkartılan yapısal tanımlayıcılara göre aday nodüllerin tespiti sağlanmıştır (Ye et al., 2008).

Yoğunluk ve model tabanlı yöntemlerin birlikte kullanılarak akciğer nodül tespitinin daha doğru oranda yapılmasını sağlamaya yönelik çalışmalar da bulunmaktadır. **(Paik et al., 2004)**'in çalışmasında bir yüzeyden ya da yakınından geçen yüzey normallerinin sayısı ile orantılı bir skor elde edilmiş, görüntünün gradyanı çıkartılıp yüzey normallerinin yönelimi belirlenmiş aday akciğer nodül bölgeleri ile kolon polipleri tespit edilmiştir **(Ye et al., 2008)**.

#### 4.5 Yanlış Pozitiflerin Giderilmesi

Bu aşama, gerçek pozitiflik oranını ve duyarlılığı en yüksek değere çıkarmak amacıyla kullanılmaktadır. Nodül tespiti sonucunda yanlış pozitiflerin elenerek, oranın azaltılması sağlanır **(Lee et al., 2012)**.

YSA ile voksel tabanlı bir yaklaşımın kullanıldığı bir çalışmada, komşu kesitlerde yer alan bölgelerden özneliliklerin çıkartılması ile yanlış pozitiflerin giderilmesi sağlanmıştır **(Retico et al., 2008)**. Ayrıca,  $\alpha$ -hull yöntemi ile çıkartılan aday juxtaopleural nodül bölgelerinden, yanlış pozitif bölgelerin çıkarılması için ileri yayılım ve momentuma sahip, öğreticili öğrenmeye dayanan çok katmanlı YSA kullanılmıştır **(Nunzio et al., 2011)**.

Bu alanda yapılan diğer çalışmalar ve ayrıntılı deneysel sonuçlar **(Lee et al., 2012)**, **(Elbaz et al., 2013)**'dan incelenebilir.

## 5. GELİŞTİRİLEN SİSTEM VE ÖNİŞLEME

Bu bölümde, tez kapsamında geliştirilen plevraya yakın akciğer nodülü bölge tespit yöntemi ve sistemi hakkında bilgiler verilmekte, geliştirilen sistemin mimarisi gösterilmekte ve temel aşamaları özetlenmektedir.

### 5.1 Geliştirilen Sistemin Mimarisi

Tez kapsamında geliştirilen sistem, literatürdeki diğer çalışmalarda genel olarak kullanılan CAD aşamalarını içermektedir (Şekil 5.1). Bu aşamalar, görüntülerin elde edilmesi, önışleme, akciğer bölütleme, aday nodül bölgelerinin tespiti, öznitelik çıkarma ve sınıflandırmadır.



Şekil 5.1 Geliştirilen sistemin genel mimarisi

Görüntünün elde edilmesi aşamasında, MRI, CT, PET, CT-PET, radyografi gibi tıbbi görüntüleme yöntemleri ile akciğer nodülü içeren görüntülerin kurum, kuruluş, organizasyon, araştırma merkezi ya da hazır veri setlerinden sağlanmasıdır. Tez çalışmasında LIDC veri seti kullanılmıştır.

Önışleme aşaması, görüntüler üzerinde gürültülerin giderilmesi, görüntünün iyileştirilmesi, istenilen duruma getirilebilmesi için uygulanan görüntü işleme yaklaşımlarını kapsamaktadır.

Akciğer bölütleme aşaması, görüntülerden akciğer bölgelerinin çıkartılarak kesitteki diğer bölümlerden ayrılması için uygulanan yöntemleri içermektedir.

Aday nodül tespiti aşaması, bölütlenen akciğer görüntüleri üzerinden plevraya yakın aday nodül bölgelerinin tespit edilmesi için uygulanan yöntem ve yaklaşımlardan oluşmaktadır.

Öznitelik çıkarma aşamasında, aday nodül bölgeleri üzerinden tanımlayıcılar çıkartılmaktadır. Nodül tanıma aşaması ise bu tanımlayıcıları dikkate alarak sınıflandırıcılar yardımıyla aday nodül bölgesinin gerçek nodül bölgesi olup olmadığını belirlemektedir.

İlerleyen bölümlerde bu aşamalar ayrıntılı olarak anlatılmış, her biri için kullanılan yöntem ve yaklaşımlar ayrıntılı olarak açıklanmıştır.

## 5.2 Önışleme ve Akciğer Bölütleme

Önışleme ve akciğer bölütleme aşamasında kullanılan yöntemler ve yapılan işlemler bu bölümde açıklanmıştır.

### 5.2.1 Önışleme

Önışleme aşaması, veri setinden uygun görüntülerin seçilmesi, DICOM görüntü türüne özgü işlemler, kesin referans bölgelerinin gösterimi, görüntü kalitesinin artırılması, görüntü birimleri arasındaki dönüşüm, pencere düzeyi ve genişliğinin ayarlanması, göğüs kafesinin çıkarılması gibi işlemleri içermektedir (Şekil 5.2).

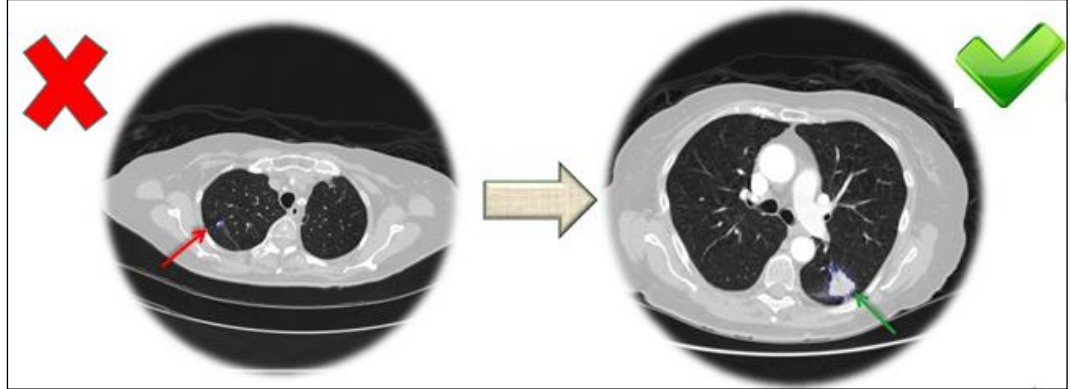


Şekil 5.2 Önışleme aşamaları

### 5.2.1.1 Verisetinden uygun görüntülerin seçilmesi

LIDC veri setinden elde edilen akciğer kesitleri üzerinde yalnızca juxta-pleural nodül içeren görüntüler belirlenmiş, diğer nodül türlerinin (isolated, juxta-vascular, pleural-tail) yer aldığı akciğer kesitleri veri setinden çıkartılmıştır.

Kullanılacak akciğer kesitlerine ilişkin örnek bir görüntü Şekil 5.3'ten incelenebilir.



Şekil 5.3 Kullanılacak akciğer kesitlerinin seçimi (a) Düzgün sınırlı nodül, (b) Plevraya yakın nodül.

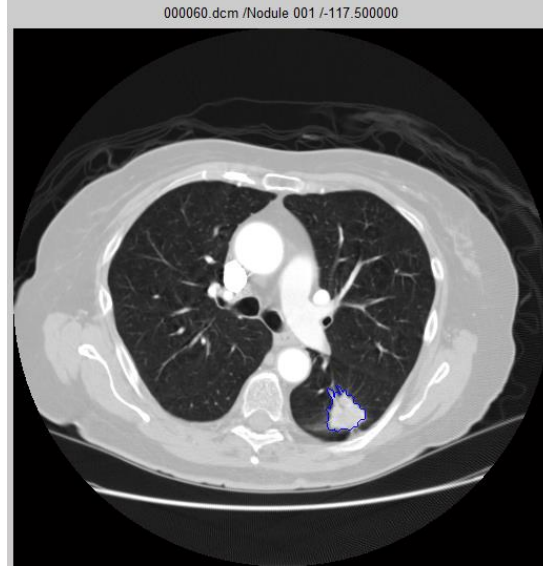
### 5.2.1.2 DICOM görüntü türüne özgü işlemler

Tıbbi görüntüleme yöntemlerine göre elde edilen görüntüler DICOM görüntü türündedir. DICOM görüntü türünde hem metinsel hem de ikili veriler bulunmaktadır. Metinsel veriler içerisinde hastanın, görüntüleme yönteminin, kesit görüntülerinin özelliklerine erişilebilmektedir.

Tez çalışmasında, metinsel veriler kısmında yer alan etiketlere göre ilgili kesit numarası belirlenmekte, hastanın tüm akciğer kesitlerini oluşturan 3D taraması üzerinden ilgili kesit konumlarına ulaşılmaktadır.

### 5.2.1.3 Kesin referans bölgelerinin gösterimi

Akciğer nodülü içeren görüntüler üzerinde uzman radyologlar tarafından sınırları belirlenmiş nodül bölgelerinin kesitler üzerinde gösterimi gerçekleştirilmiştir. Kesin referans bölgeleri, her bir hastanın tüm kesit görüntülerine ilişkin bilgileri içeren xml dosyasında tutulmaktadır. Nodül ve kesit konumu ile nodül ismi, xml dosyasından okunmakta, ilgili kesit numarası ile birlikte gösterilmektedir (Şekil 5.4).



Şekil 5.4 Kesin referans bölgesinin ve nodül bilgilerinin gösterimi (LIDC)

#### **5.2.1.4 Double formatına çevrim**

Görüntülerin değerleri üzerinde yapılacak işlemlerin doğruluğu açısından piksel değerleri double formatına dönüştürülmüştür. Bu sayede, Hounsfield birimi ile piksel değerleri arasındaki dönüşüm işlemleri ve pencere düzeyi ve genişliğinin belirlenmesi sırasında ortaya çıkacak yuvarlama problemlerinin önüne geçilmiştir.

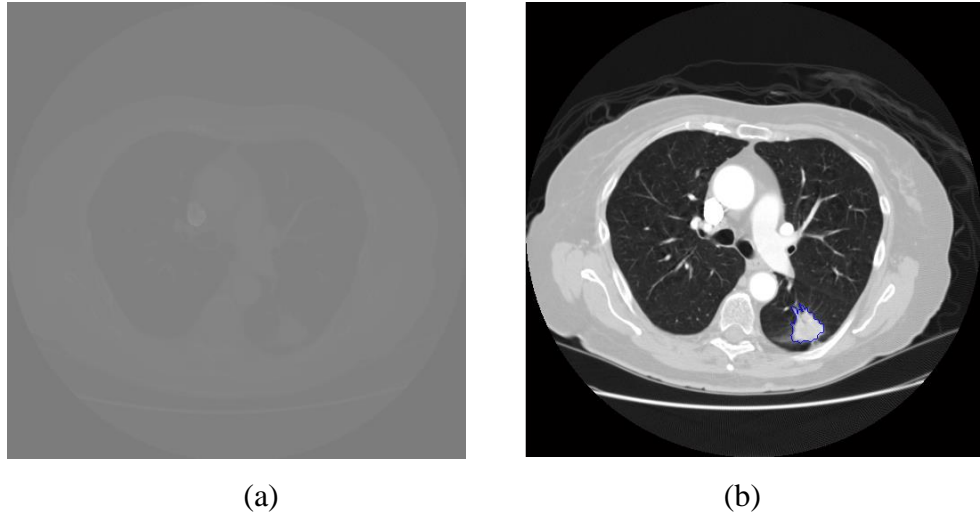
#### **5.2.1.5 Pencere düzeyi ve genişliğinin belirlenmesi**

DICOM içinde tutulan görüntü bilgileri ikili veriler biçiminde piksel değerlerinden oluşmaktadır. Akciğer bölgesine uygun pencere düzeyi ve genişliğinin belirlenebilmesi amacıyla, piksel değerlerinin radyolojik yoğunluk birimi olan Hounsfield ünitesine dönüştürülmesi gerekmektedir. DICOM dosyasındaki Rescale Slope ve Rescale Intercept değerlerine göre doğrusal dönüşüm işlemi yapılmıştır (Bkz. Eşitlik 3.1). Görüntülerdeki **Rescale Slope** değişkeni için **1** ve **Rescale Intercept** değişkeni için ise **-1024** değeri kullanılmıştır.

Akciğer CT'sinde kullanılan aralık değeri-pencere genişliği 1500 ve pencere düzeyi ise -400'tür. Bu pencere değerleri ile  $-400 \pm 1500$  yani  $[-1150, 350]$  arasındaki HU değerleri etkili bir biçimde gösterilebilmektedir. Bu aralığın üst sınırını aşan HU değerleri beyaz, alt sınırı aşan HU değerleri ise siyah renkte gösterilmektedir. Ayrıca, görüntü işleme yöntemlerinin



uygulanması açısından 16 bitlik görüntü 8 bitlik görüntü türüne dönüştürülmüştür (Şekil 5.5).

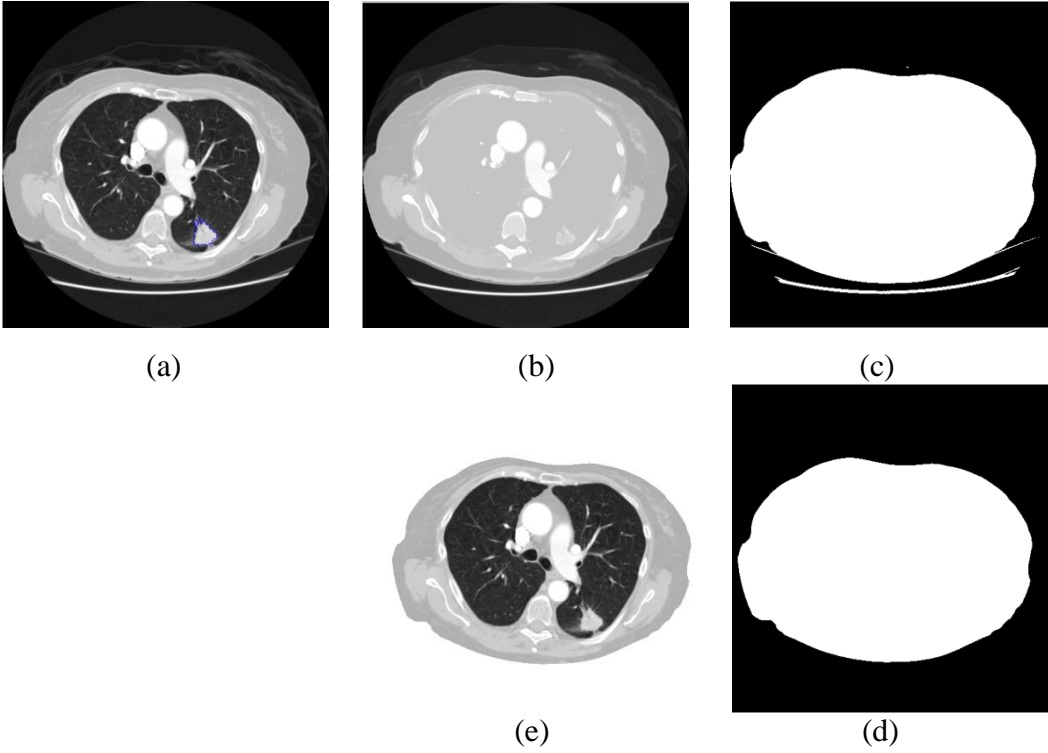


Şekil 5.5 Pencere değişkenlerinin belirlenmesi (a) Orijinal akciğer CT kesiti, (b) Pencere düzey ve genişliğinin ayarlanması.

#### **5.2.1.6 Göğüs kafesinin çıkartılması**

Göğüs kafesinin çıkartılması, 8 bitlik akciğer kesit görüntüsü üzerinden ilgilenilmeyen yapıların çıkartılarak göğüs kafesinin elde edilmesi için morfolojik işlemler gibi görüntü işleme tekniklerine dayanır. Bunun için ilk olarak göğüs bölgesinin içi boşluklarla doldurulur. Bu işlem ile akciğerler, göğüs kafesinin akciğer bölgeleri dışında kalan kısmının rengi ile kaplanmaktadır. Ardından, Otsu eşikleme yöntemi kullanılarak kesit görüntüsü üzerinde göğüs kafesinin ve alt kısmında yer alan çizgilerin maskesi çıkartılmaktadır. Otsu eşikleme yöntemi, doku yoğunlukları arasındaki sınıf içi varyansı en az indirgeyerek, önplan ve arkaplan renklerinin siyah-beyaz renklerinde en uygun biçimde ayrılabilmesi için gereken eşik değerini belirlemede kullanılmaktadır.

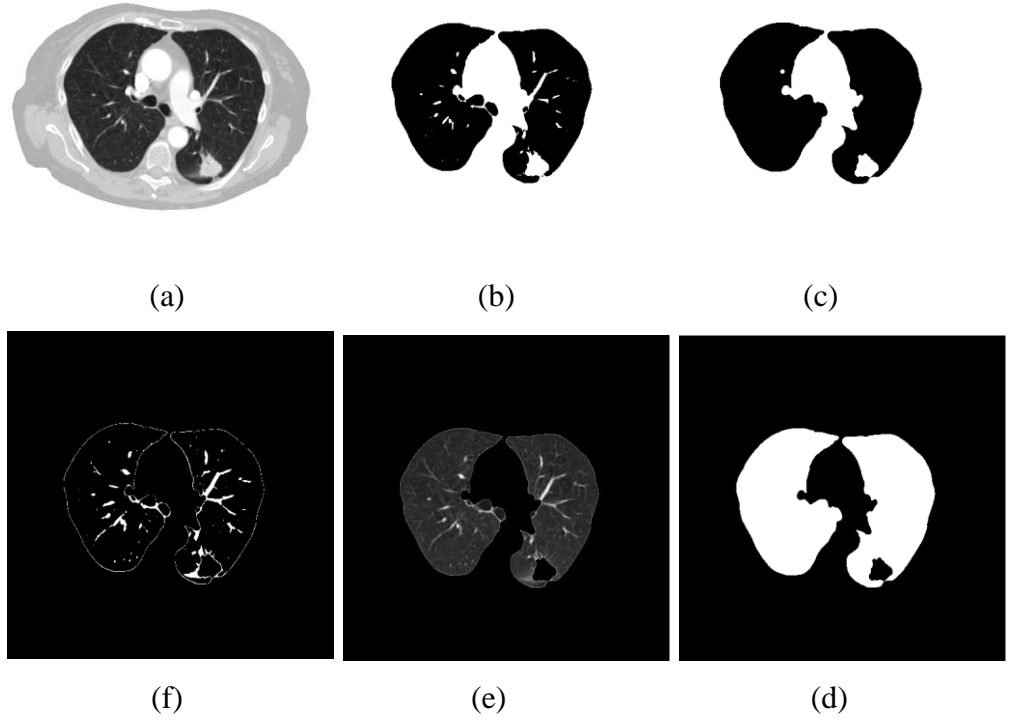
Göğüs kafesinin dışında kalan parçaları temizleyebilmek amacıyla 10 birim yarıçaplı disk şeklinde yapısal eleman oluşturularak görüntü üzerinde aşındırma işlemi yapılmıştır. Aşındırma sonucunda elde edilen göğüs kafesini içeren maskenin dışında kalan alanlar asıl kesit görüntüsünde beyaz renk ile kaplanmış ve akciğer bölgesini içine alan göğüs kafesi elde edilmiştir. Yapılan işlemlere ilişkin bir örnek Şekil 5.6'dan incelenebilir.



Şekil 5.6 Göğüs kafesinin çıkartılması (a) Orijinal akciğer kesiti, (b) Boşlukların doldurulması, (c) Otsu eşikleme İşlemi, (d) Morfolojik aşındırma, (e) Göğüs kafesinin maskelenmesi.

### 5.2.2 Akciğer bölütleme

Görüntüden göğüs kafesi çıkartıldıktan sonraki aşama akciğerlerin bölütlenmesidir. Bu aşamada yalnızca göğüs kafesi üzerindeki alanda Otsu eşikleme yöntemi uygulanarak akciğerler bölütlenmiştir. Bölütleme sonucunda akciğerlerin dışında birtakım parçalar oluşabilmektedir. Bu durumun çözülebilmesi amacıyla 4 birim yarıçaplı disk biçiminde yapısal eleman ile bölütlenen akciğer görüntüsü üzerinde morfolojik açma işlemi gerçekleştirilmiştir. Ayrıca, akciğer bölgesi içerisindeki boşluklar doldurularak, akciğer maskesi üzerinde de Otsu eşikleme yöntemi ile akciğer sınırları çıkarılmıştır. Yapılan işlemler sonucunda sınır çizgilerinden oluşan bölgelerde en yüksek alana sahip iki bölge belirlenerek akciğer bölütleme işlemi tamamlanmıştır. Akciğer bölütleme sürecine ilişkin bir örnek Şekil 5.7’den incelenebilir.



Şekil 5.7 Akciğer bölütlemesi (a) Göğüs kafesini içeren akciğer görüntüsü, (b) Otsu eşikleme, (c) Morfolojik açma, (d) Boşlukların doldurulması, (e) Akciğer maskesi, (f) Akciğer sınırlarının çıkartılması.

### 5.3 Tartışma ve Sonuçlar

Akciğer CT görüntüleri üzerinde ön işleme aşamasında gerçekleştirilen pencere düzeyi ve genişliğinin ayarlanması, göğüs kafesinin çıkartılması gibi işlemlerde kullanılan yöntem ve yaklaşımların, parametrelerin doğruluğu, bu aşamayı takip eden akciğer bölütleme sürecinin başarısında da oldukça önemli bir rol oynamaktadır.

Akciğer bölütleme aşamasında kullanılan Otsu eşikleme yöntemlerine alternatif olarak sabit eşikleme, çok düzeyli eşikleme, optimal eşikleme, Ridler & Calvard eşikleme, Kittler & Illingworth eşikleme gibi yöntemler de kullanılmaktadır. Bu eşikleme yöntemleri geliştirilen sistem üzerinde denenmiş, damar çıkarma ve nodül bölgelerinin sınırlarında birtakım bölütleme problemleri ile karşılaşmıştır. Bu nedenle Otsu eşikleme yöntemi tercih edilmiştir. Eşikleme yöntemleri dışında dışbükey kabuk (convex hull) yaklaşımlarına dayanan adaptif border marching, alpha hull gibi model tabanlı yöntemler ile de akciğer bölütlemesi gerçekleştirilebilir.

## 6. PLEVRA YA YAKIN ADAY NODÜL BÖLGELERİNİN TESPİTİ

Bu bölümde, akciğer bölütleme aşaması sonucunda elde edilen akciğer görüntüleri üzerinde, alpha hull yöntemi kullanılarak juxta-pleural nodül bölge tespiti için gerçekleştirilen çalışma açıklanmaktadır.

### 6.1 Giriş

Akciğer nodülleri için geliştirilen CAD sistemlerinin amaçlarından biri nodül bölgelerinin otomatik biçimde tespitini sağlayarak radyologlara kanserin erken tanı ve teşhisi aşamasında yardımcı olmaktır.

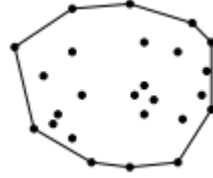
Juxta-pleural nodüller, plevraya belirli bir derece bağlantıya sahiptir. Doku yoğunlukları plevra ve akciğer bölgesini kuşatan göğüs kafesininkine yakın, akciğer dokusundan ise farklıdır. Bu nedenle, nodül bölgesinin tespit edilebilmesi amacıyla doku yoğunluklarına göre bölütleme işleminin yapılması başarısız olacaktır.

Tez kapsamında, dışbükey kabuk yönteminin genelleştirilmesine dayanan alpha hull yöntemi ile aday nodül bölgeleri tespit edilmiştir. Bu yöntemin özelliklerine, elde edilen sonuçlara, kullanılan parametrelere ve alternatif yöntemlere ilişkin bilgiler verilmiştir.

### 6.2 Alpha Hull Yöntemi ile Juxta-pleural Aday Nodül Bölgelerinin Bölütlenmesi

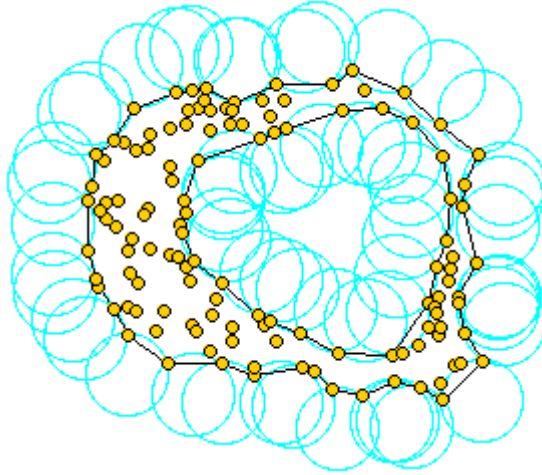
Alpha shape (alfa şekil), verilen sonlu noktalar kümesi için oluşturulan dışbükey kabuğun genelleştirilmesini sağlayan bir yaklaşımdır. Alfa şekiller, ayrıca delaunay üçgenlemenin bir altgrafıdır. Genelleştirilme işlemi için ayrıntı düzeyini belirleyen alpha parametresi kullanılır. İlk olarak **Edelsbrunner, Kirkpatrick and Seidel (1983)** tarafından ortaya çıkarılmıştır. Yüzey oluşturmada, geometrik modellemede, moleküler yapının modellenmesinde, sınıflandırmada, görselleştirmede, ızgara üretiminde, tıbbi görüntü analizinde, deprem verilerinin görselleştirilmesinde kullanılmaktadır (**Marjan, 2006**).

Dışbükey kabuk, verilen tüm noktaları kapsayan dışbükey çokgen ve tüm kapalı yarıdüzlemlerin kesişimi olarak tanımlanır. Geometrik algoritmalar alanında zaman karmaşıklığı az dışbükey kabuk yöntemlerinin geliştirilmesi üzerine yaygın olarak çalışmalar yapılmaktadır (Şekil 6.1).



Şekil 6.1 Dışbükey kabuk

**$\alpha$ -hull** ise noktalar kümesini içeren  $1/\alpha$  yarıçaplı tüm kapalı disklerin kesişimi ile ifade edilmektedir (Edelsbrunner, Kirkpatrick and Seidel, 1983).  **$\alpha$ -hull**'da diskler içerisinde herhangi bir nokta bulunmamakta, noktalar yalnızca disklerin sınır çizgileri üzerinde bulunmaktadır (Şekil 6.2).  $\alpha$  değeri küçüldükçe ayrıntı düzeyi artarken, büyüdükçe azalmaktadır (Şekil 6.3). Alfa şekli oluşturmak için  $\alpha$  yarıçaplı tüm diskler silinir. Her bir diskin temas ettiği iki noktayı birleştiren bir kenar çizilerek alfa şekli elde edilir. Alpha hull yönteminde,  $\alpha$  parametresinin uygun biçimde belirlenebilmesi, istenilen şeklin ve geometrik modelin oluşturulmasında oldukça önemlidir.

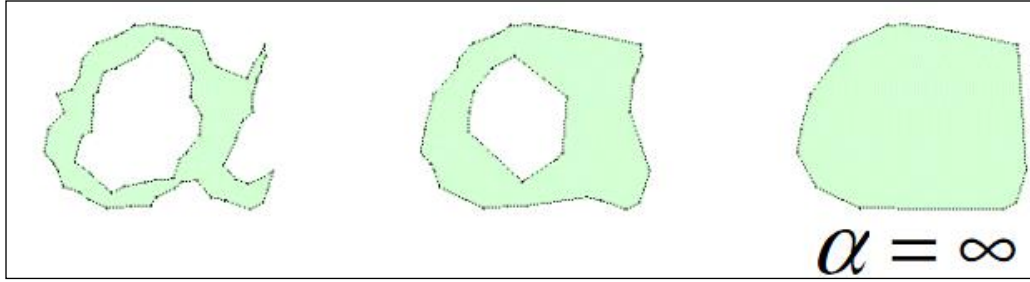


Şekil 6.2 Alpha hull (Frank, T.K., Lorient, S. and Yvinec, M., 2004)

Her reel  $\alpha$  değeri için  $1/\alpha$  yarıçaplı genelleştirilmiş disk kavramı Eşitlik 6.1'de gösterilmektedir.  $\alpha$  değeri  $-\infty$  ile  $+\infty$  arasında değişebilmektedir.  $\alpha$  değeri yeteri kadar büyük ise tüm uzayı,  $1/\alpha$  yarıçap değerine eşit olduğunda noktalar kümesini kapsayan en küçük daireyi,  $\alpha = 0$  ise noktalar kümesinin dış bükey

kabuğunu ve  $\alpha$  yeteri kadar küçük ise noktalar kümesinin kendisini temsil etmektedir (Edelsbrunner, Kirkpatrick and Seidel, 1983).

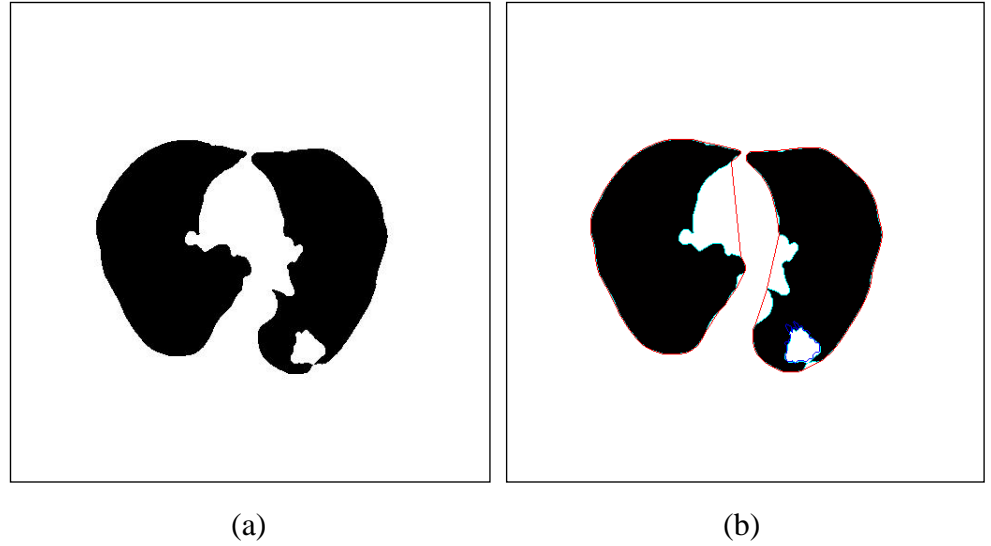
$\alpha > 0$  ise  $1/\alpha$  yarıçaplı kapalı disk olmak üzere,  
 $\alpha = 0$  ise kapalı yarıdüzlem olmak üzere,  
 $\alpha < 0$  ise  $-1/\alpha$  yarıçaplı kapalı diskin tümleyeni  
 (6.1)  
 olarak tanımlanmaktadır.



Şekil 6.3  $\alpha$  yarıçap değerinin değişimine göre oluşan şekil örnekleri (Pless, 2011)

Tez kapsamında, bölütlenen akciğer görüntüleri üzerinde  $\alpha$ -hull yöntemi kullanılarak aday juxtapleural nodül bölgelerinin belirlenmesi sağlanmıştır. Bunun için farklı kesit görüntüleri üzerinde değişen  $\alpha$  parametre değerlerine göre aday nodül bölgeleri çıkartılmıştır.  $\alpha$  değerinin belirlenmesi aşamasında kesin referans bölgelerinin, çıkartılan aday nodül bölgeleri içerisinde yer alma durumu dikkate alınmıştır. Ayrıca, toplam, sol ve sağ akciğerdeki aday nodül bölgelerinin sayısı akciğer görüntülerine ve  $\alpha$  değerine göre de değişiklik gösterebilmektedir.

Şekil 6.4'te bölütlenmiş akciğer görüntüsü ve  $\alpha$ -hull yöntemi ile elde edilen sınırlara ilişkin bir örnek gösterilmektedir.  $\alpha$ -hull yöntemi sonucunda akciğer bölgelerinin sınırları kırmızı renkteki çizgilerle, kesin referans bölgesi mavi renkle gösterilmiştir.



Şekil 6.4 Alpha hull örneği (a) Bölütlenmiş akciğer görüntüsü, (b)  $\alpha$ -hull yöntemi ile elde edilen sınırlar.

### 6.3 Alpha Hull Parametre Değerleri

Tez çalışmasında,  $\alpha$ -hull yöntemi ile juxta-pleural nodül bölgelerinin bölütlenmesi sırasında kullanılan parametre değerleri Çizelge 9.4’te gösterilmiştir.  $\alpha$ -hull yöntemi için kullanılan algoritma ve kod örneği <http://www.mathworks.com/matlabcentral/fileexchange/28851-alpha-shapes/content/alphavol.m> adresinden elde edilmiştir.

$\alpha$ -hull yöntemi kapsamında sol ve sağ akciğer görüntüleri üzerindeki yarıçap değerleri 45 birim ile 80 birim arasında 5’er birim farklılıkla ayrı ayrı denenmiştir. Her bir CT görüntüsü için alınan en iyi sonuç değerleri parametre olarak belirlenmiştir. Yarıçap değerleri seçilirken,  $\alpha$ -hull yöntemi sonucunda oluşan çizgilerin nodül sınırlarını doğru biçimde kapsaması dikkate alınmıştır.

Çizelge 6.1  $\alpha$ -hull yönteminde kullanılan yarıçap değerleri

CT Kişi İndisi	Sol Bölgenin Yarıçapı	Sağ Bölgenin Yarıçapı
1	60	55
2	60	55
3	60	55
4	60	55
5	60	40
6	70	50
7	65	65
8	60	45
9	60	45
10	50	40
11	60	60
12	55	60
13	75	70
14	75	70
15	80	75
16	80	75
17	70	55
18	80	60
19	60	50
20	50	50
21	50	50
22	50	60
23	60	50
24	50	60

CT kişi indisi, her bir hasta için toplanan kesit görüntülerinin indisini temsil etmektedir. Sol ve sağ bölgenin yarıçapları geliştirilen sistemin ekran çözünürlüğüne göre değişiklik gösterebilmektedir. Tez kapsamında, 1366 \* 768'lik ekran çözünürlüğü kullanılmıştır.

#### 6.4 Aday Nodül Bölgelerinin Çıkarılması

Plevraya yakın aday nodül bölgelerini de kapsayan yeni akciğer sınırları  $\alpha$ -hull yöntemi ile belirlendikten sonraki aşamada, bu sınırların önceden bölütlenmiş akciğer görüntüsünün sınırlarına temas eden noktaları başlangıç ve bitiş noktaları olarak kabul edilmiştir. Başlangıç ve bitiş noktaları arasında yer alan ilk kısım, bölütlenmiş akciğer görüntüsünden diğer kısım ise  $\alpha$ -hull yöntemi ile bölütlenen yeni akciğer bölgesinin sınırlarından alınarak aday nodül bölgelerinin sınırları belirlenmiştir.



Aday nodül bölgelerin oluşturulmasından önce hem gereksiz işlemleri azaltmak hem sistemin performansını artırmak amacıyla, kesin olarak gerçek nodül olamayacağı bilinen aday nodüllerin ihmal edilmesini sağlamak için sisteme sınırların uzaklığına dayalı bir ön koşul eklenmiştir. Bu durum, bölgenin eni ya da boyundaki başlangıç ve bitiş noktaları arasındaki uzaklığın 10 birimden fazla olmasını gerektirmektedir. LIDC veri setinde en küçük 3mm'lik nodüllerin bulunması sisteme bu koşulun koyulmasında etkili olmuştur.

Tez kapsamında,  $\alpha$ -hull yöntemi ile belirlenen yeni nodül bölge sınırlarının çizilmesinde birtakım özel durumlarla karşılaşılmıştır. Bu durumlardan ilki, aday nodül bölgelerinin sınırlarını birleştiren çizgi çizilmesi sırasında  $\alpha$ -hull yöntemi ile elde edilen konumların tamsayı olmamasından kaynaklanmaktadır.

Sayısal görüntülerin 2 boyutlu tamsayı koordinatlarına sahip piksellerden oluşmasından dolayı, çizgi çizme algoritmalarının verdiği konumlar tam sayı olmalıdır. Eğimden yararlanarak çizme, Digital Differential Analyzer, Bresenham çizgi çizme algoritmaları bilgisayar grafiklerinde yaygın olarak kullanılmaktadır. Çizgi çizme algoritmalarının performansı, zaman maliyetine, tamsayılar üzerinde işlem yapmasına ve yuvarlama durumuna göre değişmektedir. Koordinat değerlerinin yuvarlanması merdiven biçiminde görüntülerin oluşmasına neden olabilmektedir (Hearn and Baker, 2004).

Çizgi çizme algoritmaları arasında en etkili, doğru ve en çok kullanılan **Bresenham** çizgi çizme algoritmasıdır. Bresenham algoritması, sadece artırımsal tamsayı yöntemlerine dayanır. İdeal x ve y değerine yakınlığı dikkate alarak x ve y değerlerinin artırılıp artırılmayacağını belirler. Daire ve eğri çizmek için de uyarlanabilir. Kayan noktalı işlemlerin daha yavaş ve maliyetli olmasından dolayı tamsayılı işlemlere nazaran daha az tercih edilmektedir. Bu nedenle, tez kapsamında nodül bölgelerinin sınırlarının çiziminde Bresenham çizgi çizme algoritması kullanılmıştır.

Bresenham çizgi çizme algoritması Çizelge 6.2'den incelenebilir. Eğimin mutlak değerinin birden küçük ve büyük olma durumuna göre x ve y değişkenleri yer değiştirmektedir.

Çizelge 6.2  $|m| < 1.0$  için Bresenham çizgi çizme algoritması (Hearn, Baker and Carithers, 2011)

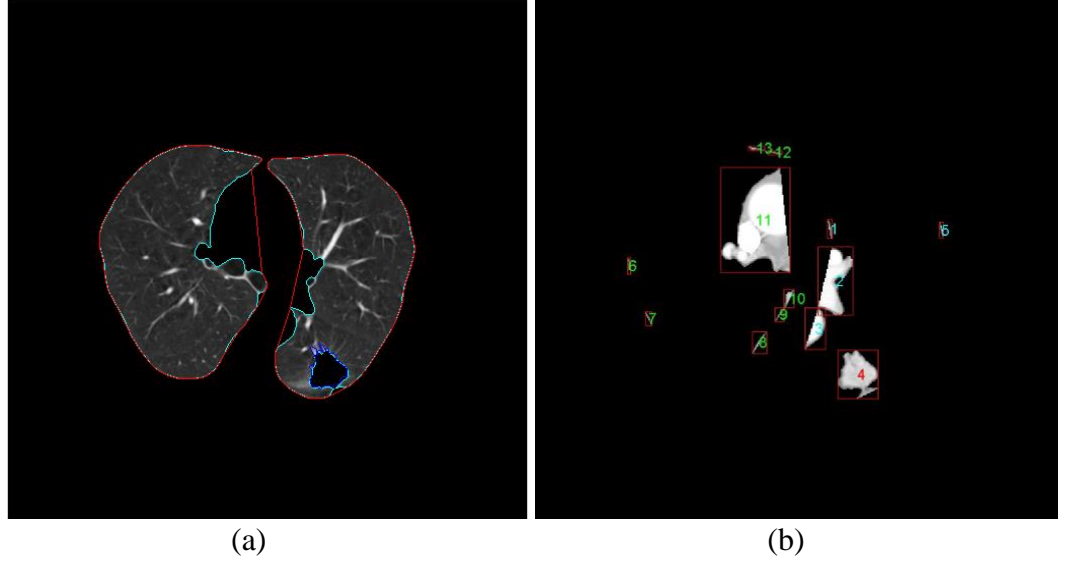
1. Girdi olarak başlangıç ve bitiş noktalarını al, solda bulunan noktayı  $(x_0, y_0)$  olarak sakla.
2. İlk nokta  $(x_0, y_0)$  için rengi ayarla ve boyala.
3.  $\Delta x$ ,  $\Delta y$ ,  $2\Delta y$  ve  $2\Delta y - 2\Delta x$  sabitlerini hesapla, karar parametresi için başlangıç değerini hesapla.
  - a.  $p_0 = 2\Delta y - \Delta x$
4. Çizgi boyunca her  $x_k$  değeri için  $k=0$ 'dan başlayarak aşağıdaki testi gerçekleştir.
  - a. Eğer  $p_k < 0$  ise, sonraki çizilecek nokta  $(x_k + 1, y_k)$  ve
    - $p_{k+1} = p_k + 2\Delta y$
  - a. Aksi takdirde, sonraki çizilecek nokta  $(x_k + 1, y_k + 1)$  ve
    - $p_{k+1} = p_k + 2\Delta y - 2\Delta x$
5. 4. Adımı  $\Delta x - 1$  kere tekrarla.

Karşılaşılan özel durumlardan ikincisi ise nodül adaylarının sınırlarını birleştiren çizgi çizildiğinde bresenham çizgi çizme algoritması sonucu elde edilen piksel konumlarının, bölütlenmiş akciğer görüntüsündeki sınırlarla çakışmasıdır. Bu durumda tek bir aday nodül bölgesi yerine aynı sınırlar içerisinde birden fazla bölge oluşabilmektedir. Bu durumu önlemek amacıyla  $\alpha$ -hull ile belirlenen, akciğer görüntüsünün sınırları dışında kalan ve bresenham algoritması sonucunda elde edilen yeni konumların, varolan diğer konumlar içerisinde yer alıp almadığı kontrol edilerek bulunanlar silinmiştir.

Aday nodül bölgelerinin sınırları çizildikten sonra sınır çizgileri arasında kalan alanların içi **imfill** komutuyla doldurulmuştur. Ayrıca, sağ akciğerden 1 numara ile başlayarak aday nodül bölgeleri sırasıyla numaralandırılmış ve etiketlenmiştir.

Şekil 6.6'da akciğer görüntüsü üzerindeki sınırlara ve  $\alpha$ -hull yöntemi ile çıkarılan aday nodül bölgelerine ilişkin bir örnek gösterilmektedir. Akciğer sınırları yeşil renkte,  $\alpha$ -hull yöntemi ile elde edilen sınırlar kırmızı renkte

gösterilmiştir. Aday nodül bölgeleri, sol akciğer üzerinde yeşil renkte, sağ akciğer üzerinde mavi renkte ve kesin referans bölgesinde kırmızı renkte numaralandırılmıştır.



Şekil 6.5 Aday nodül bölgeleri (a) Akciğer sınırları, (b) Numaralandırılmış aday nodül bölgeleri.

## 6.5 Gerçek Nodüllerin Etiketlenmesi

Geliştirilen sistemde aday nodül bölgelerinden yalnızca 1 tanesi gerçek nodül olmaktadır. Gerçek nodülün belirlenip numaralandırılması için her aday nodül bölgesinin ve kesin referans bölgesinin merkez noktası hesaplanmıştır. Aday nodül bölgelerinin merkez noktası ile kesin referans bölgesinin merkez noktası arasındaki Öklid uzaklığı en az olan bölge gerçek nodül olarak tanımlanmış, kırmızı renkle numaralandırılarak gösterilmiştir (Bkz. Şekil 6.6).

## 6.6 Tartışma ve Sonuçlar

Tez kapsamında,  $\alpha$ -hull yöntemi ile aday nodül bölgeleri belirlenmiştir. Bu yöntemlerin dışında morfolojik kapama, rolling ball filtresi gibi yöntemlerde literatürde kullanılmasına rağmen  $\alpha$ -hull yöntemi daha etkili olduğu ve doğru sonuçlar verdiği için tercih edilmiştir.

$\alpha$ -hull yöntemi ile aday nodül bölgelerinin başlangıç ve bitiş konumları arasındaki sınırlar çizdirilirken bresenham çizgi çizme algoritması, hız ve doğruluk oranından dolayı tercih edilmiştir. Bunun dışında, yüksek eğrisellik değerine sahip çizgilerin çizilmesi durumunda, doğrusal interpolasyon yerine

kübik, hermit kübik ya da yüksek derecede spline eğrileri kullanılması daha uygun olacaktır (Sensakovic, Starkey and Armato, 2008).

$\alpha$  değerinin seçimi aday nodül bölgelerinin doğru biçimde belirlenebilmesi için oldukça önemlidir. Uygun parametre değeri geliştirilen sistemde yapılan çalışmalar sonucunda önceden tanımlanmıştır. Deneysel çalışmalar bölümünde kullanılan parametre değerleri gösterilmiştir. Bu parametre değerinin otomatik olarak belirlenerek ya da var olan parametre değerleri arasında seçim yapılması sağlanarak sistemin performansı artırılabilir.

Geliştirilen sistemde, aday nodül bölgelerinden mutlaka birisi kesin referans bölgesi olmaktadır. Aday nodül bölgeleri arasından kesin referans bölgesi seçilirken birtakım kısıtlamalar getirilerek iyileştirmeler yapılabilir. Ayrıca, kesin referans bölgesinin merkez noktasının aday nodül bölgesinin merkez noktasına Öklid uzaklığının belli bir değeri geçmeme ( $1.5r_{GT}$  gibi) şartı eklenerek  $\alpha$ -hull yönteminin başarısı ölçülebilir.

## 7. ÖZNETELİK ÇIKARMA – YANLIŞ POZİTİFLERİN GİDERİLMESİ

### 7.1 Giriş

Öznitelik çıkarma aşaması, bir önceki aşamada  $\alpha$ -hull yöntemi ile elde edilen aday nodül bölgelerinden yalnızca kesin referans bölgesini içeren aday nodülü gerçek nodül olarak ayırt edebilmek, tüm pikseller yerine daha düşük boyutlu bilgiler içeren özniteliklerle sistemin performansını artırarak sınıflandırmak amacıyla kullanılmaktadır.

Öznitelik çıkarma, girdi verisi olarak alınan görüntülerden anlamlı bilgilerin, değerlerin, tanımlayıcıların çıkarılması işlemi olarak tanımlanmaktadır. Algısal olarak farklı şekiller arasındaki farklılıkların fazla ve benzer şekiller arasındaki farklılıkların az olması, taşıma, döndürme ve ölçekleme işlemlerinden bağımsız değerlerin üretilmesi, koordinat sistemi dönüşümleri, gürültüler, farklı bir şekil tarafından şeklin üstünün örtülmesi gibi durumlardan etkilenmemesi, güvenilir olması, kolay ayırt edilebilmesi, tanımlayıcıların kalitesi açısından oldukça önem taşımaktadır. Öznitelik çıkarma aşaması, tanıma, sınıflandırma, verilen örüntü türüne benzer örüntülerin bulunması gibi problemlerin çözümünde sıklıkla kullanılmaktadır.

Tez kapsamında, nodül bölgelerinin tanınması amacıyla sınıflandırıcının eğitimi ve testi için gerekli olan şekil, renk ve doku bilgisine dayalı öznitelikler çıkartılmıştır. Öznitelikler üzerinde eşik değerleri tanımlanarak gerçek nodül olma şartı sağlanmaktadır. Alt bölümlerde çıkarılan öznitelikler hakkında gerekli bilgiler verilmiştir.

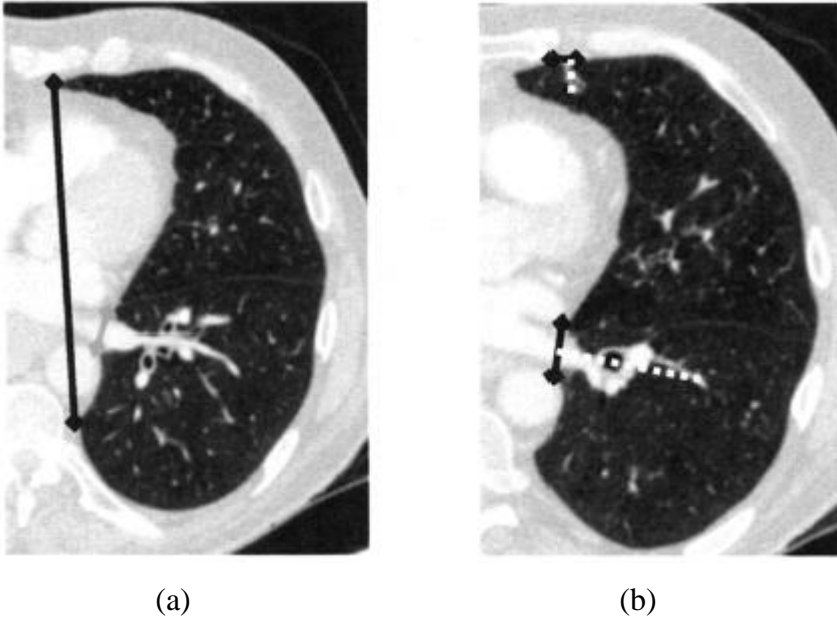
### 7.2 Şekil Bilgisine Dayalı Öznitelikler

Şekil bilgisine dayalı öznitelikler, ilgilenilen görüntü bölgesinden şekil bilgisini içeren ayırt edici etmenlerin çıkarılmasını içermektedir. Bu öznitelikler, morfolojik ya da geometrik öznitelikler olarak da adlandırılmaktadır. İlgilenilen bölgenin renk yoğunlukları ve desen bilgileri bu aşamada dikkate alınmaz.

Yapılan tez çalışmasında, akciğer görüntülerindeki şekillere özgü öznitelikler (depth (derinlik), span (kapsam), boundary length (sınır uzunluğu), radius (yarıçap)) kullanıldığı gibi, (area (alan), circularity ratio (dairesellik oranı), eccentricity (dış merkezlilik), solidity (katılık)) gibi genel özniteliklerden de yararlanılmıştır. Şekil bilgisine dayalı olarak kullanılan özniteliklerin açıklamaları alt bölümlerde açıklanmıştır.

### 7.2.1 Kapsam (Span)

Aday nodül bölgesinin başlangıç ve bitiş konumları arasındaki Öklid uzaklığıdır. Aday iç bükey bölgenin büyüklüğünü temel alarak uygun olmayan bölgenin elenmesini sağlar (Şekil 7.1) (Sensakovic, Starkey and Armato, 2008).



Şekil 7.1 İçbükeylik öznitelikleri (a) Kapsam, (b) Derinlik (Sensakovic, Starkey and Armato, 2008).

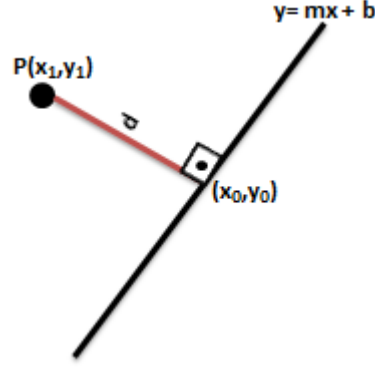
### 7.2.2 Derinlik (Depth)

Aday nodül bölgesinin kapsam çizgisi ile sınır noktaları arasındaki en büyük dikey uzaklıktır (Sensakovic, Starkey and Armato, 2008) (Bkz. Şekil 7.1). Önemli kısıtlardan birisi kapsam çizgisinin doğru olarak değil doğru parçası biçiminde dikkate alınmasıdır. Doğru parçasının dışında kalan noktalarda uzaklık 0 olarak kabul edilmiştir.

İki nokta arasındaki en kısa uzaklık aralarında çizilen çizginin uzaklığıdır. Bir nokta ile doğru arasındaki en kısa uzaklık ise o doğruya çizilen dikme ile elde edilir. Şekil 7.2’de görüldüğü gibi nokta ile doğru arasındaki  $d$  uzaklığı en kısa

uzaklığa eşittir.  $(x_1, y_1)$  noktası ile  $y = mx + n$  doğrusu arasındaki uzaklığa ilişkin formül Eşitlik 7.1’de gösterilmiştir. Nokta ile doğru parçası arasındaki uzaklık hesabında ise ek olarak sınır kontrolleri yapılmıştır.

$$d = \frac{|y_1 - mx_1 - b|}{\sqrt{m^2 + 1}} \quad \text{veya} \quad d = \frac{|Ax_1 + By_1 + c|}{\sqrt{A^2 + B^2}} \quad (7.1)$$



Şekil 7.2 Bir nokta ile doğru arasındaki en kısa uzaklığın gösterimi

### 7.2.3 Sınır uzunluğu (Boundary length)

Aday nodül bölgesinin başlangıç ve bitiş konumları arasındaki nokta sayısına eşittir. Gürültülerin olduğu aday nodül bölgeleri üzerinde en küçük sınır uzunluk eşik değeri tanımlanarak uygun olmayanların ortadan kaldırılmasına yardımcı olmaktadır (Sensakovic, Starkey and Armato, 2008).

### 7.2.4 Derinliğin kapsama oranı (Depth over span)

Derinliğin kapsama oranını ifade eder (depth / span). Şekil bilgisi olarak öznitelik çıkarma aşamasında kullanılmaktadır.

### 7.2.5 Sınır uzunluğunun kapsama oranı (Bound over span)

Aday nodül bölgesinin  $\alpha$ -hull yöntemi ile belirlenen başlangıç ve bitiş konumları arasındaki nokta sayısını ifade eden sınır uzunluğunun, kapsama oranına eşittir (boundary length / span). Şekil bilgisi olarak öznitelik çıkarma aşamasında kullanılmaktadır. Tez kapsamında çıkarılan özgün özniteliklerden biridir.

### 7.2.6 Ana eksen uzunluğunun sınır uzunluğuna oranı (Major axis length over boundary length)

Aday nodül bölgesinin merkez noktasından ve sınırlarından geçen en uzun çizginin uzunluğunun,  $\alpha$ -hull yöntemi ile belirlenen başlangıç ve bitiş konumları arasındaki nokta sayısını ifade eden sınır uzunluğunun oranına eşittir (major axis

length / boundary length). Şekil bilgisi olarak öznitelik çıkarma aşamasında kullanılmaktadır. Tez kapsamında çıkarılan özgün özniteliklerden biridir.

#### **7.2.7 İkincil eksen uzunluğunun sınır uzunluğuna oranı (Minor axis length over boundary length)**

Aday nodül bölgesinin merkez noktasından ve sınırlarından geçen en kısa çizginin uzunluğunun,  $\alpha$ -hull yöntemi ile belirlenen başlangıç ve bitiş konumları arasındaki nokta sayısını ifade eden sınır uzunluğuna oranını ifade eder (minor axis length / boundary length). Ana eksene dikey olan eksen ikincil eksen temsil eder. Şekil bilgisi olarak öznitelik çıkarma aşamasında kullanılmaktadır. Tez kapsamında çıkarılan özgün özniteliklerden biridir.

#### **7.2.8 Kapsamın ana eksen uzunluğuna oranı (Span over major axis length)**

Aday nodül bölgesinin  $\alpha$ -hull yöntemi ile belirlenen başlangıç ve bitiş konumları arasındaki Öklid uzaklığı ile tanımlanan kapsamın değerinin, ana eksen uzunluğunun oranına eşittir (Span / major axis length). Şekil bilgisi olarak öznitelik çıkarma aşamasında kullanılmaktadır. Tez kapsamında çıkarılan özgün özniteliklerden biridir.

#### **7.2.9 Ana eksen uzunluğu ile ikincil eksen uzunluğunun çarpımının sınır uzunluğuna oranı (Mj\_MnAxis over boundary length)**

Ana eksen uzunluğu ile ikincil eksen uzunluğunun çarpımının sınır uzunluğuna oranına eşittir (Major axis length \* Minor axis length / boundary length). Şekil bilgisi olarak öznitelik çıkarma aşamasında kullanılmaktadır. Tez kapsamında çıkarılan özgün özniteliklerden biridir.

#### **7.2.10 Alan (Area)**

Aday nodül bölgesinin başlangıç ve bitiş konumları arasında kalan sınır noktaları ile başlangıç ve bitiş noktalarını birleştiren kapsamdan oluşan kapalı bölgenin toplam nokta sayısını temsil eder. Aday nodül bölgesinin büyüklük değerlerine göre eşik değerinin belirlenmesi ve gerçek nodülün bulunmasını sağlar (Sensakovic, Starkey and Armato, 2008).

#### **7.2.11 Yarıçap (Radius)**

Aday nodül bölgesinin sınır noktalarının merkez noktasına olan ortalama Öklid uzaklığıdır (Nunzio et al., 2011).



### 7.2.12 Dairesellik oranı (Circularity ratio)

Bir şeklin daireye hangi ölçüde benzediğini tanımlar. Şeklin alanının aynı çevreye sahip dairenin alanına oranına eşittir. Dairesellik oranına ilişkin denklem Eşitlik 7.2’de gösterilmiştir.

$$C_1 = \frac{A_s}{A_c} \quad \text{ve} \quad C_1 = \frac{4\pi \cdot A_s}{\theta^2} \quad (7.2)$$

$C_1$  dairesellik oranını,  $A_s$  şeklin alanı,  $A_c$  daire alanı,  $\theta$  dairenin çevresi olmak üzere.

Dairesellik oranı aynı zamanda daire varyansı olarak da tanımlanmaktadır. Daire varyansı ise Eşitlik 7.3’teki gibidir (Mingqiang, Kidiyo and Joseph, 2008).

$$C_{var} = \frac{\sigma_r}{\mu_r} \quad (7.3)$$

$C_{var}$  dairenin varyansı,  $\sigma_r$  ve  $\mu_r$  şeklin merkezi ( $g_x, g_y$ )’den ( $x_i, y_i$ ) sınır noktalarına  $i \in [0, N - 1]$  olan radyal uzaklığın yarıçapının standart sapması ve ortalaması olmak üzere.

### 7.2.13 Dışmerkezlilik (Eccentricity)

En – boy oranıdır. Aday nodül bölgesinin çevresini kuşatan minimum sınırlı dikdörtgenin uzun kenarı olan ana eksen uzunluğunun ( $E_a$ ), kısa kenarı olan ikincil eksen uzunluğuna ( $E_b$ ) oranı ile hesaplanmaktadır.

Diğer bir tanımla, aday nodül bölgesi ile aynı ikinci momente sahip elipsin dışmerkezliliğinin sayısal değeri olarak da tanımlanmaktadır (Mathworks Inc., 2011). Dışmerkezlilik denklemi Eşitlik 7.4’te gösterilmiştir.

$$E = E_a / E_b \quad (7.4)$$

### 7.2.14 Katılık (Solidity)

Aday nodül bölgesinin alanının ( $A_s$ ), şekli çevreleyen dışbükey kabuğun alanının ( $A_{hull}$ ) oranına eşittir. Şeklin iç bükey ya da dış bükey olup olmadığını belirler. Dışbükey bir şeklin katılık değeri her zaman 1’e eşittir (Mingqiang, Kidiyo and Joseph, 2008). Katılığa ilişkin denklem Eşitlik 7.5’te gösterilmiştir.

$$\text{Katılık} = A_s / A_{hull} \quad (7.5)$$

### 7.2.15 Dışbükey kabuğun alanı (Convex area)

Şeklin dışbükey kabuğunda yer alan toplam piksel sayısına eşittir (Mathworks Inc., 2011).

### 7.2.16 Kapsam (Extent)

Aday nodül bölgesinin içindeki toplam piksel sayısının, şekli kuşatan minimum sınırlı dikdörtgensel bölgenin alanının oranına eşittir (Mingqiang, Kidiyo and Joseph, 2008). Kapsam değerinin nasıl hesaplandığına ilişkin denklem Eşitlik 7.6'da gösterilmiştir.

$$E = A_{\text{shape}} / A_{\text{mbb}} \quad (7.6)$$

E kapsam,  $A_{\text{shape}}$  şeklin alanı ve  $A_{\text{mbb}}$  şekli kuşatan minimum sınırlı dikdörtgensel bölgenin alanı olmak üzere.

## 7.3 Doku Bilgisine Dayalı Öznitelikler

Doku veya desen, bir yüzeyde düzenli olarak bir örüntünün, elemanın tekrarlanması, benzerliklerin gruplanması olarak tanımlanmaktadır. Desen bilgileri renk, parlaklık, şekil, boyut gibi özelliklerden oluşmaktadır (Srinivasan and Shobha, 2008).

Görüntü desen analizi ise görüntüden bilgiler çıkartılarak elde edilen konumsal değişiklikleri niteleyen, matematiksel prosedürler ve modellerdir. Doku analizi, desen sınıflandırma, görüntü bölütleme ve şekil belirlemede kullanılmaktadır (Srinivasan and Shobha, 2008).

Birinci dereceden, ikinci dereceden veya daha yüksek dereceden istatistiksel bilgiler kullanılarak dokusal öznitelikler çıkarılabilmektedir. Tez kapsamında, 1. ve 2. dereceden istatistiksel dokusal öznitelikler çıkartılmış, ayrıca Hu momentlerinden de yararlanılmıştır. Alt bölümlerde, çıkarılan öznitelikler ve açıklamaları ile ilgili bilgiler verilmiştir.

### 7.3.1 Birinci dereceden istatistiksel öznitelikler

İstatistiksel özellikler, görüntünün her bir pikselindeki doku yoğunluğunun uzaysal dağılımını analiz ederek, her noktadan yerel öznitelikler çıkarılması, özniteliklerin dağılımından birtakım istatistikler elde edilmesini içerir (Srinivasan and Shobha, 2008).

Birinci dereceden istatistiksel bilgiler yalnızca ilgilenilen piksel değerini dikkate alarak, komşu piksellerle olan ilişkilerden etkilenmez. Bu değerler, histogram tabanlı yaklaşımlardan elde edildiği gibi moment gibi özniteliklerden

de elde edilebilir. Ortalama, varyans, çarpıklık, basıklık gibi momente dayalı öznitelikler tez çalışmasında kullanılmıştır. Kullanılan dokusal öznitelikler şu şekildedir:

#### **7.3.1.1 Aritmetik ortalama (Arithmetic mean)**

İlgilenilen bölgedeki doku yoğunluklarının aritmetik ortalaması hesaplanmaktadır. Aritmetik ortalamaya ilişkin denklem Eşitlik 7.7’de gösterilmiştir.

$$\mu = \frac{1}{N} \sum_{i=1}^N I_i \quad (7.7)$$

$\mu$  aritmetik ortalama,  $I_i$  i.pikselin doku yoğunluğu ve N ilgilenilen bölge içerisindeki toplam piksel sayısı olmak üzere.

#### **7.3.1.2 Geometrik ortalama (Geometric mean)**

İlgilenilen bölgedeki doku yoğunluklarının geometrik ortalaması hesaplanmaktadır. Geometrik ortalamaya ilişkin denklem Eşitlik 7.8’de gösterilmiştir.

$$G = \sqrt[N]{I_1 \cdot I_2 \cdots I_N} \quad (7.8)$$

G geometrik ortalama,  $I_i$  i.pikselin doku yoğunluğu ve N ilgilenilen bölge içerisindeki toplam piksel sayısı olmak üzere.

#### **7.3.1.3 Harmonik ortalama (Harmonic mean)**

İlgilenilen bölgedeki doku yoğunluklarının harmonik ortalaması hesaplanmaktadır. Harmonik ortalamaya ilişkin denklem Eşitlik 7,9’da gösterilmiştir.

$$H = \frac{n}{\frac{1}{I_1} + \frac{1}{I_2} + \cdots + \frac{1}{I_n}} \quad (7.9)$$

H harmonik ortalama,  $I_i$  i.pikselin doku yoğunluğu ve N ilgilenilen bölge içerisindeki toplam piksel sayısı olmak üzere.

#### **7.3.1.4 Standart sapma (Standard deviation)**

İlgilenilen bölgedeki doku yoğunluklarının standart sapması hesaplanmaktadır. Bu öznitelik ile doku yoğunluklarının ortalama değerden ne kadar sapma değeri gösterdiği, kendi içindeki değişim miktarı bilgisi elde edilmektedir. Standart sapmaya ilişkin denklem Eşitlik 7.10’da gösterilmiştir.

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (I_i - \mu)^2} \quad (7.10)$$

s standart sapma,  $I_i$  i.pikselin doku yoğunluğu,  $\mu$  aritmetik ortalama ve n ilgilenilen bölge içerisindeki toplam doku sayısı olmak üzere.

#### **7.3.1.5 Çarpıklık (Skewness)**

İlgilenilen bölgenin doku yoğunluğu histogramının, ortalama değerin etrafındaki asimetrisini ölçmek için kullanılır. Sıfıra, pozitif veya negatif değere eşit olabilir. Normalize edilmiş üçüncü moment olarak da bilinmektedir. Çarpıklık değerinin hesaplanmasına ilişkin denklem Eşitlik 7.11’de gösterilmektedir.

$$g = \frac{m_3}{m_2^{3/2}}, \quad m_3 = \sum (I_i - \mu)^3 / n \quad \text{ve} \quad m_2 = \sum (I_i - \mu)^2 / n \quad (7.11)$$

g çarpıklık,  $m_3$  3. merkezi moment ve  $m_2$  2. merkezi moment olmak üzere.

#### **7.3.1.6 Basıklık (Kurtosis)**

İlgilenilen bölgenin doku yoğunluğu histogramının keskinliğini ölçmek amacıyla kullanılır. Normalize edilmiş dördüncü moment olarak da bilinmektedir. Basıklık değerinin hesaplanmasına ilişkin denklem Eşitlik 7.12’de gösterilmektedir.

$$k = \frac{m_4}{m_2^2}, \quad m_4 = \sum (I_i - \mu)^4 / n \quad \text{ve} \quad m_2 = \sum (I_i - \mu)^2 / n \quad (7.12)$$

k basıklık,  $m_4$  4. merkezi moment ve  $m_2$  2. merkezi moment olmak üzere.

#### **7.3.1.7 Entropi (Bilgi vitimi) (Entropy)**

Entropi, rastsal bir değişkenin belirsizliğinin ölçüsü olarak tanımlanmaktadır. Öznitelik çıkarma aşamasında, ilgilenen görüntü bölgesinin desenini tanımlamak amacıyla, doku yoğunluklarının rastgeleliğini ölçmek için kullanılan istatistiksel bir ölçüttür (Mathworks Inc., 2011). Entropi, Eşitlik 7.13’deki gibi tanımlanır.

$$Entropy = - \sum_{i=1}^n p_i \log_2 p_i \quad (7.13)$$

n toplam durum sayısı,  $p_i$  i. durumun gerçekleşme olasılığı olmak üzere.

#### **7.3.1.8 Çeyrek değerler genişliği (Interquartile range)**

Üçüncü çeyrek değer ile birinci çeyrek değer arasındaki farka eşittir. Doku yoğunluklarının dağılımını tahminlemek ve aykırı değerleri tespit etmek amacıyla

kullanılan istatistiksel özniteliktir (Mathworks Inc., 2011). Çeyrek değerler genişliğinin hesaplanmasına ilişkin denklem Eşitlik 7.14'te gösterilmiştir.

$$IQR = Q3 - Q1 \quad (7.14)$$

IQR çeyrek değerler genişliği, Q3 üçüncü çeyrek değer, Q1 birinci çeyrek değer olmak üzere.

### 7.3.1.9 Hu geometrik momentleri

Momentler taşıma, döndürme ve ölçekleme işlemlerinden bağımsız olarak desen öznitelikleri çıkarmakta kullanılan skaler niceliklerdir. Hu tarafından 1962 yılında 7 farklı moment tanımlanmıştır. Bu momentler arasında 7. Moment çarpıklığa karşı da değişmezdir. Hu momentleri örüntü tanıma problemlerinde sıklıkla kullanılmaktadır. 2 boyutlu sürekli  $f(x,y)$  fonksiyonu için  $r = p + q$  dereceli geometrik moment Eşitlik 7.17'deki gibi tanımlanmaktadır (Flusser, J., Suk, T. and Zitová, B., 2009).

$$M_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y) dx dy \quad (7.17)$$

Eşitlik 7.17 gri tonlamalı görüntüye uyarlandığında, görüntünün momenti elde edilmektedir. Görüntünün momentine ilişkin denklem Eşitlik 7.18'de gösterilmiştir.

$$M_{ij} = \sum_x \sum_y x^i y^j I(x,y) \quad (7.18)$$

Tez çalışmasında Hu momentleri dokusal öznitelik olarak kullanılmıştır. Hu momentlerinin denklemleri Eşitlik 7.19'da gösterilmiştir (Hu, 1962).

$$\phi_1 = \eta_{20} + \eta_{02} \quad (7.19)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\phi_5 = (\eta_{30} - 3\eta_{12}) + (\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (\eta_{21} + \eta_{03})(3\eta_{21} - \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$\phi_6 = (\eta_{20} - \eta_{02}) + \left[ \frac{(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2}{4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})} \right]$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$\eta_{ij}$  normalize edilmiş merkezi moment olmak üzere.

### 7.3.2 İkinci dereceden istatistiksel öznitelikler

Birinci dereceden istatistiksel öznitelikler, görüntünün gri düzey dağılımını dikkate alırken piksellerin ilişkisel konumları ve doku yoğunlukları hakkında bilgi vermemektedir (Theodoridis and Koutroumbas, 2009). İkinci dereceden istatistiksel öznitelikler ise verilen uzaklık değeri ve açıya göre belirlenen komşu pikseller arasındaki ilişkileri dikkate almaktadır.

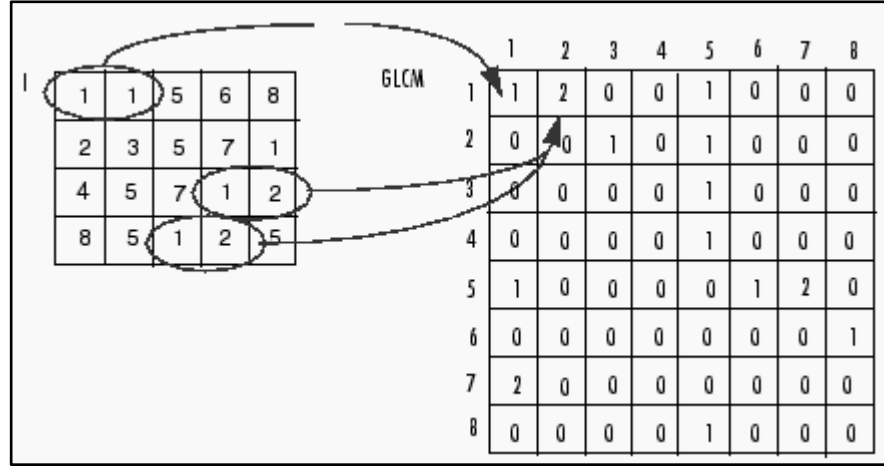
Tez çalışmasında, ikinci dereceden istatistiksel öznitelikler arasında yaygın olarak kullanılan **gri düzey eş oluşum matrisi (gri düzey uzaysal bağımlılık) (Gray level co-occurrence matrix)**'den yararlanılmıştır. Alt bölümde, bu matris ile ilgili gerekli açıklamalar yapılmıştır.

#### 7.3.2.1 Gri düzey eş oluşum matrisi

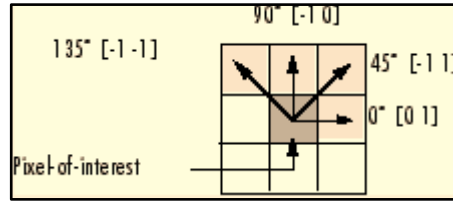
1973 yılında Haralick et al. tarafından tanımlanan gri düzey eş oluşum matrisi,  $i$  renk yoğunluğu değerine sahip olan pikselin  $j$  renk yoğunluğu değerine sahip olan pikselle, verilen uzaklık birimince komşu olduğu durum sayılarını oluşturur. Ayrıca, ikili resimler iki farklı ve gri tonlamalı resimler genellikle 8 farklı gri düzeyine ölçeklenir.

Gri düzey eş oluşum matrisi (GLCM) taşıma işleminden bağımsız iken, döndürme işlemine karşı hassastır. Bunun için 0, 45, 90 ve 135 derecedeki açılara göre 4 farklı yönde gri düzey eş oluşum matrisleri oluşturulur. Bu matrislerin değerlerinin toplamı ya da ortalaması alınarak matris, döndürme işlemine karşı bağımsız hale getirilir. Verilen  $N * N$  boyutlu,  $I$  görüntüsü üzerinde,  $P$  gri düzey eş oluşum matrisi Eşitlik 7.5'deki gibi tanımlanmaktadır (Eleyan and Demirel, 2011). Örnek bir GLCM ve konum – açı ayarlama görüntüsü Şekil 7.3'te gösterilmiştir.

$$P(i, j) = \sum_{x=1}^N \sum_{y=1}^N \begin{cases} 1, & \text{Eğer } I(x, y) = i \text{ ve } I(x + \Delta x, y + \Delta y) = j \\ 0, & \text{değilse} \end{cases} \quad (7.15)$$



(a)



(b)

Şekil 7.3 GLCM örnekleri (a) GLCM matrisinin oluşturulması, (b) Konum ve yön tanımlama (Mathworks Inc., 2011).

Yapılan tez çalışmasında gri düzeyli aday nodül bölgeleri 8 farklı doku düzeyine ölçeklendirilmiştir. Ardından 4 farklı yönde 1 birim uzaklıkta gri düzey eş oluşum matrisleri oluşturulmuş ve bu matrislerin tek bir matris içinde toplanarak döndürme işleminden bağımsız olması sağlanmıştır.

GLCM elde edildikten sonraki aşamada, matrisin tüm elemanlarının toplamı 1 olacak şekilde normalize edilerek ikinci dereceden istatistiksel öznitelikler çıkartılmıştır. Bu öznitelikler, otokorelasyon (autocorrelation), kontrast (contrast), küme önemi (cluster prominence), cluster shade, enerji, entropi ve homojenlik olmak üzere 7 adettir.  $P(i, j)$  normalize edilmiş GLCM matrisindeki  $(i, j)$ . girdi olsun. Matrisin satır ve sütunlarının aritmetik ortalaması ve kullanılan öznitelikler Eşitlik 7.16'da gösterilmiştir. (Soh and Tsatsoulis, 1999).

$$\mu_x = \sum_i \sum_j i \cdot p(i, j) \quad \mu_y = \sum_i \sum_j j \cdot p(i, j) \quad (7.16)$$

$$Autocorrelation = \sum_i \sum_j (ij) \cdot p(i, j)$$

$$Contrast = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \mid |i-j| = n \right\}$$

$$Cluster\ prominence = \sum_i \sum_j (i+j - \mu_x - \mu_y)^4 p(i,j)$$

$$Cluster\ shade = \sum_i \sum_j (i+j - \mu_x - \mu_y)^3 p(i,j)$$

$$Energy\ (Angular\ Second\ Moment) = \sum_i \sum_j p(i,j)^2$$

$$Entropy = - \sum_i \sum_j p(i,j) \log(p(i,j))$$

$$Homogeneity = \sum_i \sum_j \frac{1}{1 + (i-j)^2} p(i,j)$$

Tez çalışmasında dokusal öznitelik olarak toplam 22 farklı tanımlayıcı kullanılmıştır. Bu özniteliklerin 15 tanesi birince dereceden istatistiksel özniteliklerden, kalan 7 tanesi ise GLCM matrisi ile elde edilen ikinci dereceden istatistiksel tanımlayıcılardan oluşmaktadır.

### 7.3.3 Şekil ve doku bilgisine dayalı öznitelikler

Şekil ve doku bilgisine dayalı öznitelikler ayrı ayrı çıkartılmış olmasının dışında, her iki öznitelik bilgisinin de bir arada kullanıldığı 2 adet özgün tanımlayıcı tez kapsamında tanıma kalitesinin artırılması amacıyla geliştirilmiştir. Aday nodül bölgelerinden çıkartılan 2 öznitelik alt bölümlerde açıklanmıştır.

#### 7.3.3.1 Kapsamın varyansa oranı (span over variance)

Aday nodül bölgesinin  $\alpha$ -hull yöntemi ile belirlenen başlangıç ve bitiş konumları arasındaki Öklid uzaklığı ile tanımlanan kapsamın değerinin, ilgilenilen bölgede yer alan doku yoğunluklarının varyansının oranına eşittir (span over variance). Tez kapsamında şekil ve doku bilgisine dayalı olarak çıkartılan özgün özniteliklerden biridir.

#### 7.3.3.2 F1

Aday nodül bölgesinin ana eksen uzunluğunun, ikincil eksen uzunluğuyla çarpımından elde edilen değerin ilgilenilen bölgede yer alan doku yoğunluklarının varyansının oranına eşittir (major axes length \*



minor axes length / variance). Bu özniteliğe F1 ismi verilmiştir. Tez kapsamında şekil ve doku bilgisine dayalı olarak çıkartılan özgün özniteliklerden biridir.

#### **7.4 Tartışma ve Sonuçlar**

Öznitelik çıkarma aşaması tanılama sürecindeki başarı oranının artırılması, performans kazancı sağlanması için örüntü tanıma problemlerinde sıklıkla kullanılan ara aşamalardan birisidir. Çıkarılan özniteliklerin kalitesi, ilgilenilen bölgeyi diğerlerinden daha etkili olarak ayırabilmesini sağlayan bilgi düzeyine, beklenen sınıfa uygunluğuna ve özniteliklerin kendi aralarındaki ilişkilerin zayıf olmasına göre artmaktadır. Bu durumdan dolayı, çıkarılmış olan öznitelikler bir sonraki aşamada özniteliklerin seçimi ve derecelendirilmesi için oldukça önemli bir işleve sahip olacaktır.

Yapılan tez çalışmasında, 16 adet şekle dayalı, 22 adet dokusal ve 2 adet hem şekle hem de dokuya dayalı olmak üzere toplam 40 öznitelik çıkartılmıştır.

## 8. ÖZNİTELİK SEÇİMİ

### 8.1 Giriş

Öznitelik seçimi aşaması, sınıflandırıcı tasarımında yüksek düzeyde ayırt edici bilgiye sahip olan özniteliklerin seçilmesi, düşük düzeyde bilgi sağlayan özniteliklerin ihmal edilmesini amaçlayan yöntem ve yaklaşımları içermektedir.

Öznitelik veya değişken seçimi işleminin, verilerin boyutunu düşürerek sınıflandırıcının tahminleme performansını artırmak ve genelleştirebilme yeteneğini iyileştirmek, eğitim süresini kısaltmak, daha hızlı ve az maliyetli, daha az depolama alanına gereksinim duyan tahminleyiciler üretmek, verilerin görselleştirilmesini ve anlaşılmasını kolaylaştırmak gibi önemli amaç ve katkıları bulunmaktadır. İnternet belgelerinde metin işlenmesinde, gen ifadesi dizi analizinde, kombinasyonel kimya problemlerinde oldukça yoğun bir biçimde çalışmalar yapılmaktadır (Guyon and Elisseeff, 2003).

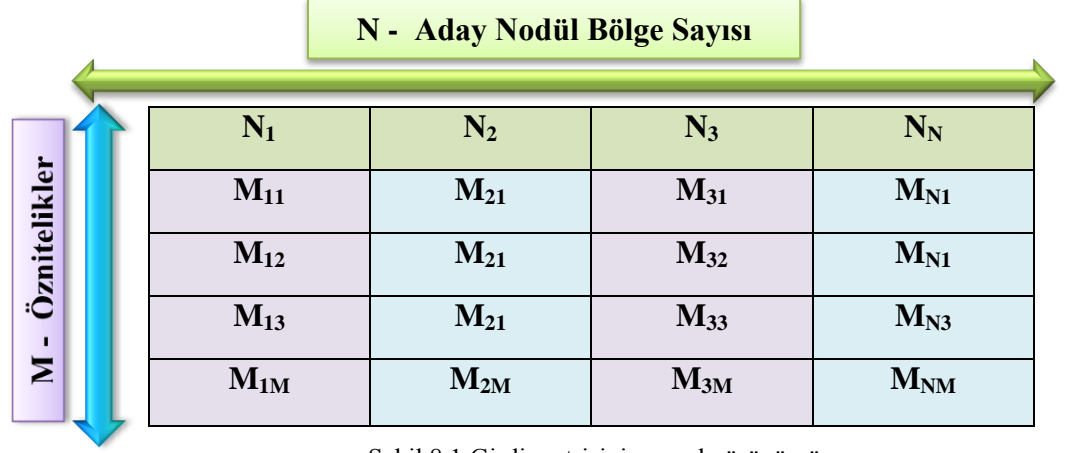
Geliştirilen tez projesi kapsamında aday nodül bölgelerinden çıkartılan öznitelikler arasından önemli olanların seçilmesi amacıyla çalışma yapılmış, öznitelik sayısı ve kalitesinin ölçülmesi için uygulanan yöntem ve yaklaşımlar alt bölümlerde açıklanmıştır. Öznitelik sayısı ve derecelerinin belirlenmesi, tez çalışmasında katkı yapılan bölümlerden biri olması açısından önem taşımaktadır.

### 8.2 Girdi ve Çıktı Matrisinin Oluşturulması

Öznitelik seçim aşamasına geçmeden önce her aday nodül bölgesi için öznitelik vektöründen örüntü vektörü oluşturulmuş ve girdi matrisi olarak tanımlanmıştır. Girdi matrisindeki her sütun bir aday nodül bölgesi için çıkarılan öznitelik değerlerini ve her satır ise ilgilenilen özneliğin farklı aday nodül bölgeleri üzerindeki değerlerini temsil etmektedir.

Girdi matrisinin genel görünümü Şekil 8.1'deki gibidir. Çıktı matrisinin tanımlanmasında ise her aday nodül bölgesi için ilgili sütuna eğer kesin referans bölgesini içeren gerçek nodül ise 1 değil ise 0 olarak atanmıştır. Çıktı matrisinin iki farklı türü bulunmaktadır. İlki yalnızca 1 veya 0 değeri ile tek satırdan

oluşmakta, diğeri ise ilk çıktı matrisinin tümleyeniyile birleşiminden oluşan iki satırlık bir matristir. Çıktı matrisi örneği Şekil 8.2’de gösterilmektedir.



Şekil 8.1 Girdi matrisinin genel görünümü

N <sub>1</sub>	N <sub>2</sub>	N <sub>3</sub>	N <sub>N</sub>
0	1	0	0
1	0	1	1

(a)

N <sub>1</sub>	N <sub>2</sub>	N <sub>3</sub>	N <sub>N</sub>
0	1	0	0

(b)

Şekil 8.2 Çıktı matrisi örnekleri (a)Çift satırlı çıktı matrisi, (b) Tek satırlı çıktı matrisi.

### 8.3 Öznitelik Değerleri Üzerinde Normalizasyon İşlemi

Girdi matrisi oluşturulduktan sonra elde edilen öznitelik değerleri geniş bir aralığı kapsamaktadır. Bu durum sınıflandırıcının eğitimini zorlaştırmakta ve maliyetini artırmaktadır. Bu nedenle, sınıflandırıcı tasarımı ve öznitelik seçim süreçlerine geçmeden önce önileme aşaması kapsamında, her bir öznitelik için elde edilen değerler  $[-1 \ 1]$  aralığına ölçeklendirilmiştir. Ölçekleme işlemi için öznitelik verilerinin değerleri üzerinde doğrusal dönüşüm işlemi gerçekleştiren **min-max normalizasyonu** uygulanmıştır.  $[\min_x \ \max_x]$  aralığındaki değerler  $[\text{new\_min}_x \ \text{new\_max}_x]$  değerlerine ölçeklenmek istendiğinde  $v$  değerinin  $\text{new\_v}$  değerine dönüşümü Eşitlik 8.1’de gösterilmiştir.

$$\text{new\_v} = \frac{v - \min_x}{\max_x - \min_x} \cdot (\text{new\_max}_x - \text{new\_min}_x) + \text{new\_min}_x \quad (8.1)$$

## 8.4 Özniteliklerin Derecelendirilmesi ve Seçimi

Öznitelik değerlerinin seçimi için literatürde 3 temel yöntem kullanılmaktadır. Bunlar, **filtre**, **sarmalayıcı** ve **gömülü** yöntemlerdir. **Filtre yöntemi** matematiksel bir fonksiyon (korelasyon, bilgi kazancı vb.) kullanılarak sınıflandırıcıdan bağımsız biçimde özniteliklerin derecelendirilmesini ve seçimini sağlamaktadır. Hızlıdır fakat sınıflandırıcı dikkate alınmadığından başarı oranı düşebilmektedir. **Sarmalayıcı yöntemler**, sınıflandırıcı başarısına göre öznitelikleri derecelendirmektedir. Başarı oranı daha yüksek olmasına rağmen işlem yükü oldukça fazladır ve yavaştır (Guyon and Elisseeff, 2003).

**Gömülü yöntemler** ise sınıflandırıcının eğitimi sürecinde, verilen sınıflandırıcıya özgü olarak öznitelik seçimini gerçekleştirir. Eğitim kümesinin eğitim ve doğrulama kümesine bölünmeden, sınıflandırıcının birden fazla eğitimi yapılmaksızın özniteliklerin seçimi yapılır. Budama işlemini sağlayan karar ağaçları, SVM-RFE (Support Vector Machine – Recursive Feature Elimination) gibi sınıflandırıcılarda gömülü olarak bulunur. Sarmalayıcı yöntemlere göre daha hızlıdır. Her öznitelik kümesini ayrı ayrı değerlendirme gereksinimini ortadan kaldırır ve verilerin daha etkin kullanımını sağlar (Guyon and Elisseeff, 2003).

Öznitelik seçimi işlemi genel kapsamda iki temel aşamadan oluşmaktadır. İlk aşamada öznitelik derecelendirme veya indirgeme yaklaşımları kullanılarak toplam sayısı  $m$  tane olan tüm özniteliklerin arasından  $n$  tanesi seçilir. Bu aşama, gereksiz özniteliklerin çıkarılması ve işlem yükünün azaltılması için önem taşır ve skaler öznitelik seçim aşaması olarak adlandırılır.

İkinci aşamada ise indirgenmiş  $n$  farklı öznitelik arasından en iyi kombinasyona sahip  $k$  farklı öznitelik altkümesinin seçilmesi sağlanır. Bu aşamada, bir önceki iterasyonda indirgenen öznitelik sayılarıyla hesaplama yükü azaltılır (Theodoridis and Koutroumbas, 2010).

Yapılan tez çalışmasında, ilk aşamada sınıflandırıcıdan bağımsız olarak özniteliklerin derecelendirildiği filtre yönteminden yararlanılmıştır. Öznitelik derecelendirme ölçütleri olarak **istatistiksel t testi**, **Kullback – Leibler İraksama**

(göreceli entropi), bhattacharyya uzaklığı, wilcoxon testi, ROC (receiver operating characteristic) değeri kullanılmıştır. Bu ölçütler alt bölümlerde açıklanmıştır.

#### 8.4.1 Hipotez testi : t testi

Öznitelik seçiminde ilk aşama her bir özniteliğin bağımsız olarak değerlendirilmesi ve önemsiz olanların ihmal edilmesidir. Bunun için istatistiksel t testi kullanılmaktadır. İstatistiksel t testinde öznitelik değerlerinin ortalamasının iki farklı küme içerisinde ne kadar değişip değişmediği, farklılaştığı ölçülerek önem derecesi belirlenir. Eğer bu fark değeri fazla ise alternatif hipotezdir ve öznitelik seçilir. Aksi takdirde sıfır hipotezidir ve öznitelik ihmal edilir (Theodoridis and Koutroumbas, 2010).

#### 8.4.2 Kullback – Leibler ıraksama (Göreceli entropi)

Öznitelik seçiminde, her bir özniteliğin bağımsız olarak değerlendirilmesinin dışında, tüm öznitelik vektörünü dikkate alarak hedef sınıf için ayırt edicilik kriteri yüksek olanların seçilmesinde kullanılan yöntemlerden birisidir.

Kullback – Leibler Iraksaması, nodül tanıma aşamasında aday nodül bölgeleri içerisinde nodül olanların sınıfındaki özniteliklerin değerleri ile nodül olmayanların değerlerinden oluşan sınıfların varyansları arasındaki güçlü bağımlılığı ölçmektedir. Kullback – Leibler Iraksaması aynı rastgele değişken üzerinde iki dağılımın entropisini karşılaştırmak için kullanılır. İki dağılım arasında ıraksama denklemi Eşitlik 8.2’de gösterilmiştir (Theodoridis and Koutroumbas, 2010).

$$d_{1,2} = \frac{1}{2} \text{trace}\{S_1^{-1}S_2 + S_2^{-1}S_1 - 2I\} + \frac{1}{2} (m_1 - m_2)^T (S_1^{-1} + S_2)(m_1 - m_2) \quad (8.2)$$

$d_{1,2}$  ıraksama değerini,  $S_1$  kovaryans matrisini,  $m_i$  , i. sınıfın beklenen ortalama vektörü,  $I$  birim matris olmak üzere.

#### 8.4.3 Bhattacharyya uzaklığı

İki sınıfın dağılımı arasındaki benzerliği ölçmek için kullanılır. Bu uzaklık, Bayes sınıflandırıcısının hatası için yakın bir ilişki gösterir. İki sınıf arasındaki  $B_{1,2}$  Bhattacharyya uzaklığı Eşitlik 8.3’te gösterilmiştir (Theodoridis and Koutroumbas, 2010).

$$B_{1,2} = \frac{1}{8} (m_1 - m_2)^T \left( \frac{S_1 + S_2}{2} \right)^{-1} (m_1 - m_2) + A \quad (8.3)$$

$$A = 0.5 \ln \frac{0.5(|S_1 + S_2|)}{\sqrt{|S_1||S_2|}} \text{ ve } |\cdot| \text{ ilgilenilen matrisin determinantı olmak üzere.}$$

#### 8.4.4 Wilcoxon (Mann Whitney U) testi

Parametrik olmayan istatistiksel hipotez testidir. İki bağımsız değişkene ait dağılımın aynı olup olmadığını karşılaştırmak için kullanılmaktadır. Eşleştirilmiş t testinin parametrik olmayan karşılığıdır.

#### 8.4.5 ROC (Receiver operating characteristic) değeri

ROC eğrisi ile rastgele sınıflandırıcının eğimi arasında kalan alanı temsil etmektedir. Belli bir özniteliğin sınıf ayırt edilebilirlik kriterini ölçmede kullanılır.

Bu öznitelik seçim ölçütleri, öznitelikleri verilen matematiksel fonksiyonlara göre derecelendirmekte fakat öznitelikler arasındaki ilişki düzeyini (korelasyonu) dikkate almamaktadır. Sisteme **çapraz korelayon katsayısı** (cross-correlation coefficient) eklenerek bu durum sağlanmıştır. Bunun için ilk olarak, en iyi dereceye sahip  $i_1$  indisli öznitelik seçilir. İlk öznitelik ile diğer öznitelikler arasındaki çapraz korelasyon katsayısı hesaplanır ve ikinci önemli olan  $i_2$  indisli öznitelik Eşitlik 8.4'e göre belirlenmektedir (Theodoridis and Koutroumbas, 2010).

$$i_2 = \max\{a_1 C_j - a_2 |\rho_{i_1, j}|\} \quad j \neq i_1 \quad (8.4)$$

$C_j$  j. öznitelik için seçim kriterinin değerini ve en iyi öznitelik ( $i_1$ ) ile öznitelik  $j \neq i_1$  arasındaki  $|\rho_{i_1, j}|$  çapraz korelasyon katsayısını birleştirir.  $a_1$  ve  $a_2$  kullanıcı tarafından tanımlanan değerlerdir.

İki öznitelik değeri dışında kalan öznitelikler ise Eşitlik 8.5'e göre seçilmektedir (Theodoridis and Koutroumbas, 2010).

$$i_k = \max\left\{a_1 C_j - \frac{a_2}{k-1} \sum_{r=1}^{k-1} |\rho_{i_1, j}|\right\}, \quad j = i_r, \quad r = 1, 2, \dots, k-1 \quad (8.5)$$

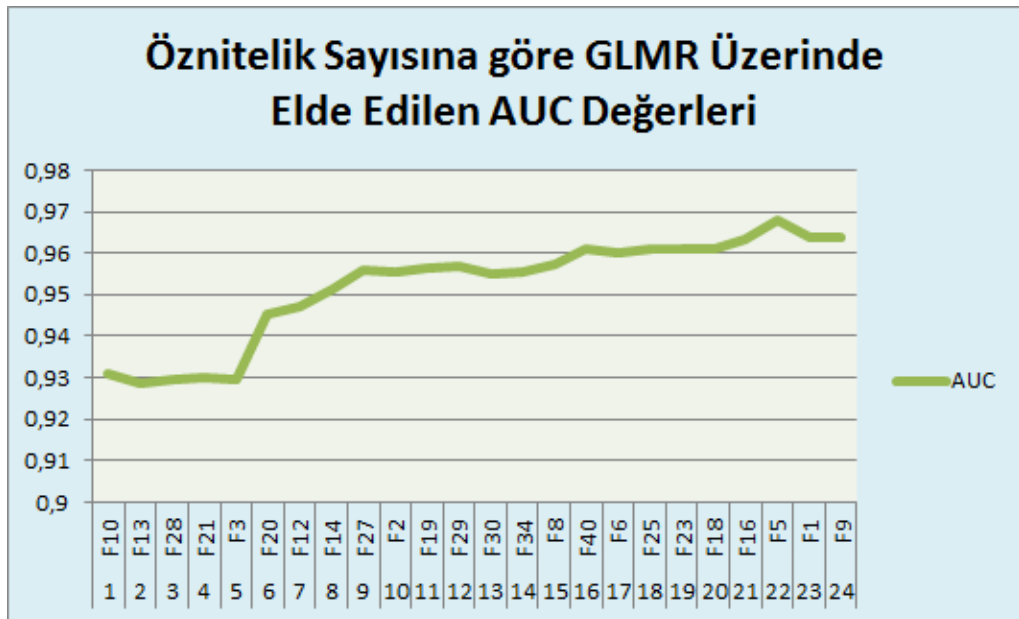
Tez çalışmasında, öznitelik derecelendirme ölçütüne ek olarak çapraz korelasyon katsayısı 1 olarak alınmış ve birbirleriyle ilişkili olan özniteliklerin seçilme olasılığı en aza indirgenmeye çalışılmıştır.

5 farklı öznitelik derecelendirme ölçütü, tez çalışmasında kullanılmıştır. Bir sonraki aşamada bu ölçütlere göre sıralanan öznitelikler arasından uygun sayının ve ölçütün belirlenmesi için gerekli işlemler yapılmaktadır.

### 8.5 Öznitelik Sayısının Belirlenmesi

Tez çalışmasında, tüm öznitelikler 5 farklı öznitelik seçim ölçütüne göre derecelendirilmiş, en iyi 5 sınıflandırıcı üzerinde özniteliklerin performansı ölçülerek sayıları belirlenmiştir. Her bir sınıflandırıcı 10 kez çalıştırılmış ve elde edilen başarı sonuçlarının genel ortalaması alınmıştır. Öznitelik sayısı, her bir öznitelik seçim kriterine göre, en iyi sınıflandırıcının ROC eğrisi altında kalan alanı (AUC) (Area Under Curve) başarı ölçütü olarak esas alınmıştır. Başarı oranının düştüğü ve önemli derecede iyileşme olmadığı yer öznitelik sayısının artırılmasının sonlanma koşulu olarak belirlenmiştir.

Yapılan deneysel çalışmalar sonucunda genelleştirilmiş doğrusal model regresyon üzerinde ‘ROC’ öznitelik seçim kriterine göre ilk 22 öznitelik seçilinceye kadar AUC değerinde önemli derecede performans artışı görülmüş, sonraki öznitelik sayısının artışlarında ise önemli bir artış gözlemlenmemiştir. Bu nedenle, uygun öznitelik sayısı 22 olarak belirlenmiştir. ROC seçim ölçütüne göre öznitelik sayısının artışına bağlı olarak sınıflandırıcılar üzerinde elde edilen deneysel sonuçlar Çizelge 8.1’de, çizgi grafiği Şekil 8.3’te gösterilmiştir.



Şekil 8.3 Öznitelik sayısının değişimi ve AUC değerlerine göre çizilen çizgi grafiği

Çizelge 8.1 Öznitelik sayısının belirlenmesi

Genelleştirilmiş Doğrusal Model Regresyon Sınıflandırıcısında Öznitelik Sayısının Performansa Etkisi					
Derece	Öznitelik İndisi	AUC	Derece	Öznitelik İndisi	AUC
1	F10	0.9311	13	F30	0.9552
2	F13	0.9287	14	F34	0.9554
3	F28	0.9298	15	F8	0.9573
4	F21	0.9302	16	F40	0.9610
5	F3	0.9297	17	F6	0.9602
6	F20	0.9453	18	F25	0.9608
7	F12	0.9470	19	F23	0.9609
8	F14	0.9514	20	F18	0.9610
9	F27	0.9559	21	F16	0.9631
10	F2	0.9556	22	F5	0.9679
11	F19	0.9564	23	F1	0.9639
12	F29	0.9568	24	F9	0.9639

Öznitelik seçiminin ikinci aşaması ise filtre yönteminden elde edilen 22 öznitelik arasından uygun olan öznitelik kombinasyonun seçilmesini içermektedir. Bu kapsamda sarmalayıcı yöntemler içerisinde yer alan ve çalışma anında dinamik olarak öznitelik ekleme ve çıkarılmasına olanak sağlayan **Sequential Floating Forward Selection (SFFS)** algoritması uygulanmıştır. Algoritmanın işlem yükünün oldukça fazla olması, çalışması için gereken zamanı önemli derecede artırmıştır ve performans düşüklüğüne neden olmuştur. Bu nedenle, 10 öznitelik kombinasyonuna kadar öznitelik seçimi yapılmış ve en iyi 5 sınıflandırıcı üzerinde performanslar karşılaştırılmıştır. Bir önceki aşamada elde edilen sonuçlara göre herhangi bir performans artışı olmadığından dolayı SFFS algoritması ve sarmalayıcı yöntemler kullanılmamıştır.

## 8.6 Özniteliklerin Analizi

Bir önceki aşama sonucunda elde edilen özniteliklerin hem kendi aralarında hem de çıktı sınıflarıyla korelasyonu ölçülerek önem dereceleri belirlenmiştir. Özniteliklerin etkili olabilmesi için hedef sınıf çıktı değerleriyle



korelasyonun yüksek, kendi aralarındaki ilişki düzeyinin ise düşük olması beklenmektedir. Bu kapsamda, öznitelik değerlerinin kendisi arasında ve hedef sınıfla olan eşli doğrusal ilişki düzeyi, istatistikte yoğun olarak kullanılan Pearson korelasyonu ile hesaplanmıştır. Pearson doğrusal korelasyon katsayısına ilişkin denklem Eşitlik 8.6’da gösterilmiştir.

$$r_{i,j} = \frac{\sum_i (x_i - \bar{x}_i)(y_i - \bar{y}_i)}{\sum_i (x_i - \bar{x}_i)^2 \sum_j (y_j - \bar{y}_j)^2} \quad veya \quad R(i,j) = \frac{C(i,j)}{\sqrt{C(i,i) C(i,j)}} \quad (8.6)$$

$r_{i,j}$  ve  $R(i,j)$  Pearson doğrusal korelasyon katsayısı,  $x$  değerlerinden oluşan  $X$  özniteliği ve  $y$  değerli  $Y$  sınıfı ve  $C$  kovaryans matrisi olmak üzere.

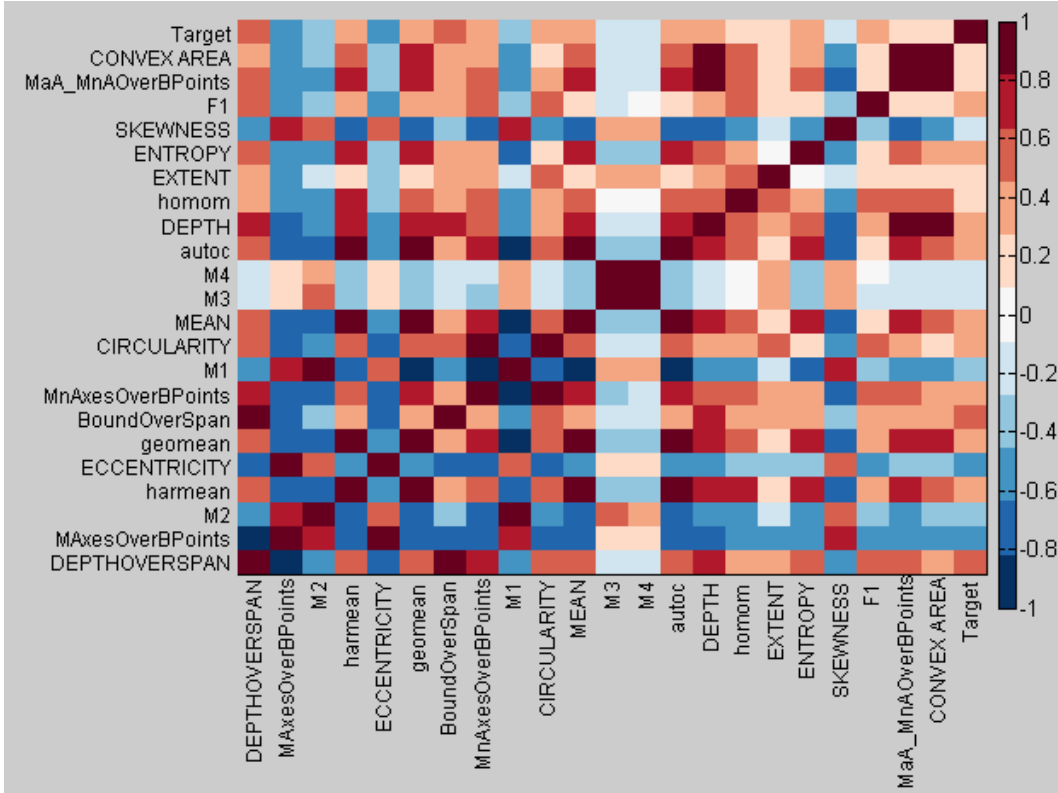
$r_{i,j}$  değerleri özniteliklerin hedef sınıf ile kendi aralarındaki doğrusal bağımlılığını göstermektedir. Bu değer, doğrusal bağımlılığın derecesine ve yönüne göre -1 ile +1 arasında değiştirmektedir ( $-1 \leq r \leq 1$ ). +1 ve -1’e yaklaştıkça değişkenler arasında ilişkinin derecesi güçlenirken 0’a yaklaştığında düşmektedir. Korelasyon değerinin 0 olması iki değişken arasında herhangi bir doğrusal ilişkinin olmadığını göstermektedir.  $r$ ’nin mutlak değerine göre korelasyonun güç derecesi (Evans, 1996) tarafından Çizelge 8.2’deki gibi tanımlanmıştır.

Çizelge 8.2 Pearson korelasyon değerinin yorumlanması (Evans, 1996)

<u>Korelasyonun Gücü</u>	<u> r </u>
<b>Çok zayıf</b>	0 – 0.19
<b>Zayıf</b>	0.20 – 0.39
<b>Orta</b>	0.40 – 0.59
<b>Güçlü</b>	0.60 – 0.79
<b>Çok Güçlü</b>	0.80 – 1.00

Öznitelik seçimi sonucu elde edilen 22 özniteliğin analizinin gerçekleştirilebilmesi amacıyla korelasyon değerlerinden oluşan matrisin renklendirilmesini sağlayan **Heat Map** veri görselleştirme aracı kullanılmıştır.

İnternet sayfaların görüntülenme sıklığında, moleküler biyolojide genlerin ifadesi için kullanılan Heat Map matrisinin farklı türleri de bulunmaktadır. Heat Map matrisi, içindeki değerlerin büyüklüğüne göre renklendirme yaparak özniteliklerin analizi hakkında bilgi vermektedir (Şekil 8.3). Renklerin koyulaşması korelasyonun yüksek olduğunu göstermektedir.



Şekil 8.4 Öznitelik ve hedef sınıftan oluşan eşli Pearson doğrusal katsayı değerlerine göre oluşturulan Heat Map matrisi

## 8.7 Öznitelik Seçimi Kullanıldığında Elde Edilen Sonuçlar

Tez çalışmasında, öznitelik seçimi aşamasında filtre yöntemlerinden 5 farklı öznitelik seçim kriteri (**istatistiksel t testi**, **Kullback – Leibler İraksama**, **bhattacharyya uzaklığı**, **wilcoxon testi**, **ROC değeri**) 40 farklı öznitelik ve en iyi 5 sınıflandırıcı üzerinde uygulanarak, özniteliklerin derecelendirilmesi ve seçimi gerçekleştirilmiştir. En iyi 5 sınıflandırıcı olarak geliştirilmiş doğrusal model regresyon (GLMR), ikinci dereceden diskriminant sınıflandırıcı (QDC), doğrusal diskriminant sınıflandırıcı (LDC), k en yakın komşu sınıflandırıcısı (k-NN) ( $k = 1$ ) ve olasılıksal yapay sinir ağı (PNN) kullanılmıştır.

En iyi başarı sonucu geliştirilmiş doğrusal model regresyon (GLMR) üzerinde ‘ROC’ öznitelik seçim ölçütünde elde edilmiş ve en uygun öznitelik sayısının 22 olduğu gözlemlenmiştir.

Tez çalışmasında özniteliklerin ROC seçim ölçütüne göre derecelendirilmesi sonucunda elde edilen 22 adet öznitelik sırasıyla şunlardır:

DepthOverSpan, MAxesOverBPoints, M2, harmonic mean, eccentricity, geometric mean, bound over span, MnAxesOverBPoints, M1, Circularity, Mean, M3, M4, autocorrelation, depth, homogeneity, extent, entropy, skewness, F1, MaA\_MnAOverBpoints, Convex Area'dır. Sınıflandırıcılar üzerinde elde edilen deneysel sonuçlar Çizelge 8.3'te gösterilmiştir.

Çizelge 8.3 Öznitelik seçimi yapıldığında elde edilen sonuçlar

<b><u>Sınıflandırıcı</u></b>	<b><u>Ortalama AUC Değeri (10 İterasyon)</u></b>
GLMR	<b>0.9679</b>
QDC	0.9034
LDC	0.8793
K-NN (K=1)	0.8666
PNN (S=0.02)	<b>0.8591</b>

Çizelge 8.3 incelendiğinde sistemin 10 iterasyon çalıştırılması sonucunda ilk 5 sınıflandırıcı arasından en iyi sonuç 0.9679 AUC değeri ile geliştirilmiş doğrusal model regresyonda, en düşük sonuç ise 0.8591 AUC değeri ile olasılıksal yapay sinir ağında elde edildiği gözlemlenmiştir.

## 8.8 Öznitelik Seçimi Olmadan Elde Edilen Sonuçlar

Öznitelik çıkarma aşaması sonucunda elde edilen 40 özneliğin öznitelik seçimi işlemine uğramadan doğrudan sınıflandırıcı eğitimi, doğrulanması ve testinde kullanıldığında oluşan sonuçlar Çizelge 8.4'te gösterilmiştir.

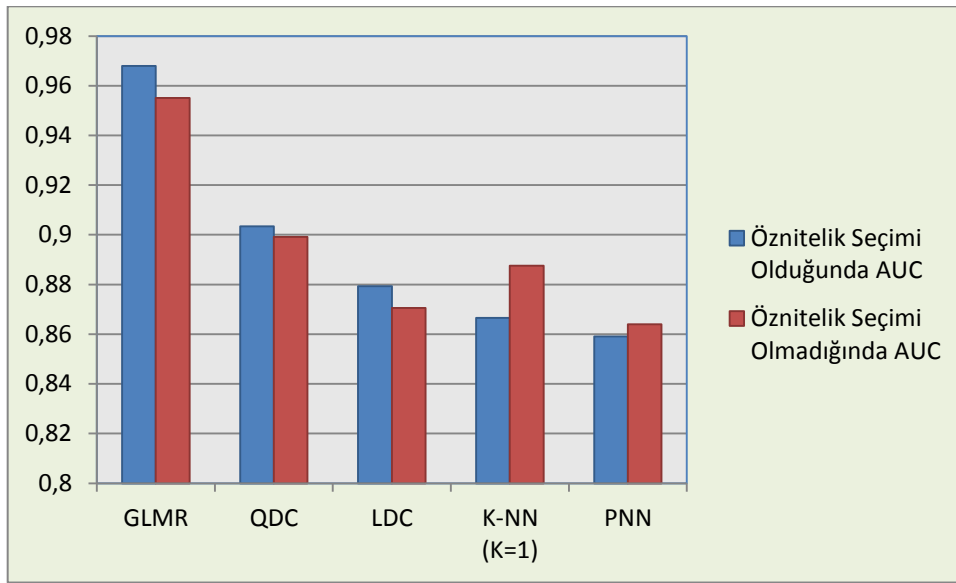
Çizelge 8.4 Öznitelik seçimi olmadan elde edilen sonuçlar

<b><u>Sınıflandırıcı</u></b>	<b><u>Ortalama AUC Değeri (10 İterasyon)</u></b>
GLMR	<b>0.9551</b>
QDC	0.8992
LDC	0.8706
K-NN (K=1)	0.8876
PNN (S=0.02)	<b>0.8640</b>

Çizelge 8.4'te en iyi 5 sınıflandırıcı üzerinde 10 iterasyon yapılarak 40 farklı öznitelik için elde edilen AUC değerlerinin ortalamasını göstermektedir.

Öznitelik seçimi yapıldığında, 22 farklı öznitelikle ortalama 0.9679'luk AUC değerine ulaşılrken, öznitelik seçimi yapılmadığında ise 40 farklı öznitelikle 0.9551 ortalama AUC değerine ulaşılmıştır. Her iki durumda da en iyi sonuç genelleştirilmiş doğrusal model regresyon sınıflandırıcısı ile sağlanmıştır.

Şekil 8.4'te öznitelik seçiminin nodül tanıma performansına etkisi sütun grafiği ile gösterilmiştir. K-NN ve PNN sınıflandırıcıları hariç diğer 3 sınıflandırıcıda öznitelik seçimi sonucunda AUC değerleriyle performans artışı sağlandığı gözlemlenmektedir.



Şekil 8.5 Öznitelik seçiminin nodül tanıma performansına etkisinin gösterimi

Yapılan deneysel çalışmalar sonucunda, öznitelik seçiminin hem sınıflandırıcı performansını hem de sınıflandırıcının eğitimi için gerekli işlem süresini kısalttığı gözlemlenmiştir.

## 8.9 Tartışma ve Sonuçlar

Öznitelik seçimi, öznitelik çıkarma aşaması sonucunda elde edilen özniteliklerden en uygun olanların seçilerek sınıflandırıcı eğitiminin hızlandırılması ve performansının artırılması açısından önemlidir. Doğru özniteliklerin çıkarılması ve seçilmesi sistemin genel başarısını yüksek ölçüde etkileyen faktörlerden biridir.

Tez çalışması sürecinde çıkartılan 40 farklı öznitelik değerinden 22 tanesi ‘ROC’ öznitelik seçim yöntemine göre seçilmiştir. Diğer öznitelik seçim ölçütleri üzerinde elde edilen başarı oranı daha düşük kalmıştır. Ayrıca, sarmalayıcı yöntemlerde de farklı seçim kriterleri denenerek seçim yapılabilir. Fakat, zaman karmaşıklığı ve işlem yükü yoğun olacak, performans düşüklüğü ortaya çıkacaktır.

Tez çalışmasında, öznitelik seçimi işlemi sınıflandırıcıdan bağımsız yapıldığından, her sınıflandırıcı için AUC sonuçları kısmen değişebilmektedir. Bu durum öznitelik seçiminin genel anlamda başarılı fakat sınıflandırıcıya bağımlı olduğunu göstermektedir. Genetik algoritmalar gibi metasezgisel yöntemler de öznitelik seçimi aşamasında uygulanabilir, bu sayede sınıflandırıcının durumu da göz önüne alınmış olacaktır.

## 9. NODÜL TANIMA

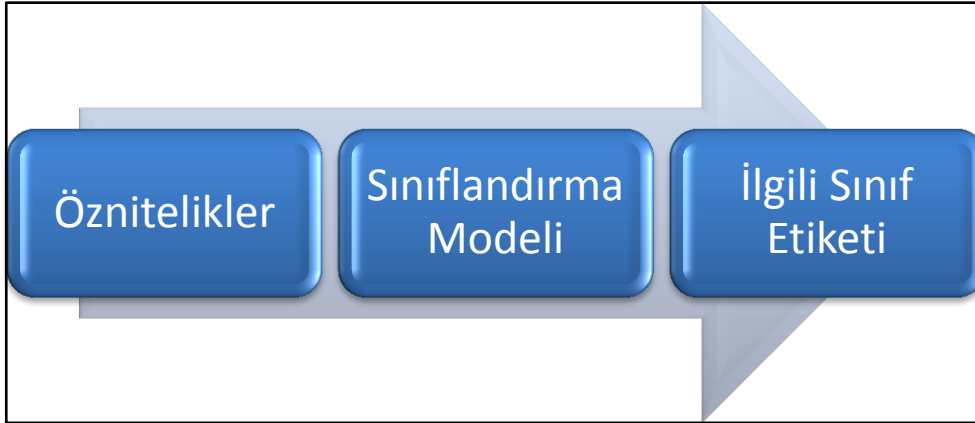
### 9.1 Giriş

Nodül tanıma, aday nodül bölgelerinden çıkartılan özniteliklerden ilgilenilen bölgenin gerçek nodül olup olmadığını sınıflandırıcılar yoluyla belirleme, örüntü tanıma işlemidir.

Bu bölümde, sınıflandırıcı modelleri hakkında genel bilgi verilmiş, tüm aday nodül bölgelerinin öznitelikler vektöründen eğitim, çapraz doğrulama ve test kümelerinin nasıl oluşturulduğu açıklanmıştır. Ayrıca, sınıflandırma modelinin seçimi, değerlendirilmesi ve performans ölçütlerine ilişkin deneysel sonuçlara da yer verilmiştir.

### 9.2 Sınıflandırma Modelleri (LDC, QDC, KNN, MLP, PNN, SVM, CT, GLMR, Ensemble Classifiers)

Sınıflandırma modelleri, varolan nesnelerin tanımlayıcılarına bağlı olarak yeni bir nesneyi önceden belirlenmiş sınıflara atamayı sağlamaktadır (Şekil 9.1).



Şekil 9.1 Sınıflandırma modelinin genel işleyişi

Tez çalışmasında kullanılan sınıflandırma modelleri doğrusal diskriminant analizi, ikinci dereceden diskriminant analizi, k en yakın komşu sınıflandırıcısı, çok katmanlı algılayıcı (MLP), olasılıksal yapay sinir ağı, destek vektör makinesi (SVM), sınıflandırma ağacı (CT), Naive Bayes sınıflandırıcı (NBC), genelleştirilmiş doğrusal model regresyon (GLMR) ve grup sınıflandırıcılarından (ECT) oluşmaktadır. Her sınıflandırıcı ayrı bir başlık içinde kısaca açıklanmıştır.

### 9.2.1 Doğrusal diskriminant analizi

Verileri içerisinde bulunan farklı grupların doğrusal ayırt edilebilirliğini en iyilemeye çalışır. Verilerin Gauss dağılımına sahip olduğu varsayımına dayanır. Tüm veri seti için bir adet kovaryans matrisi bulunmaktadır. Bunun için sınıflar arası varyansın, sınıf içi varyansa oranını maksimize etmek amaçlanır.

### 9.2.2 İkinci dereceden diskriminant analizi

Doğrusal diskriminant analizine benzer fakat ikinci dereceden denklem ve karar yüzeylerine sahiptir. Ayrıca, daha çok parametre almaktadır. Kovaryans matrisi her sınıf için ayrıdır. Bu durum, verilerin daha iyi sınıflandırılmasına yardımcı olur.

### 9.2.3 K en yakın komşu sınıflandırıcısı

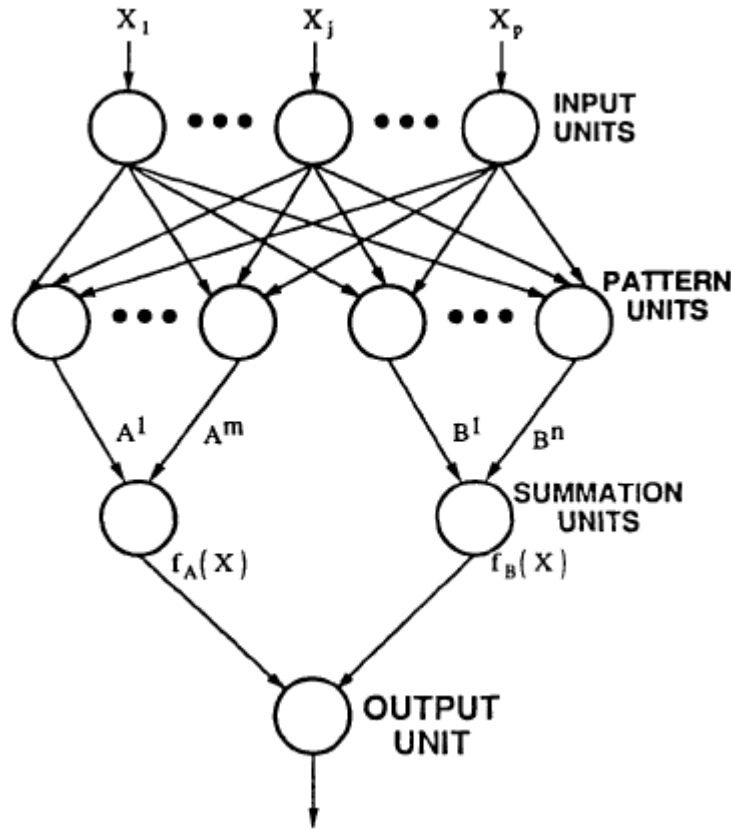
Verilen bir örneğin diğer örneklerle öznitelik uzayında belirlenen ölçüte göre uzaklıklarını karşılaştırarak en yakın olan k durumun sonuçları arasından çoğunluğa sahip grubun seçilmesini sağlamaktadır.  $K=1$  ise en yakın örneğin bulunduğu sınıf çıktı olarak kullanılır. Uzaklık ölçütü olarak Manhattan, Öklid, Minkowski, Mahalanobis uzaklıkları kullanılmaktadır.

### 9.2.4 Çok katmanlı algılayıcı

Doğrusal olmayan problemleri hızlı ve etkili biçimde çözebilmek amacıyla kullanılmaktadır. Girdi katmanı, en az 1 adet olmak üzere ara katman ve çıktı katmanından oluşur. Verilen girdilerin örüntüsünü, çıktı olarak eşlemek için çeşitli öğrenme kuralları uygulanır. Geri yayılım sayesinde yapay sinir ağının ağırlıkları güncellenerek beklenen çıktıya uygun duruma getirilir. Ağın ileri beslemeli olması ise ağırlıkların girdi katmanından çıktı katmanına doğru birbirlerini etkilediğini gösterir.

### 9.2.5 Olasılıksal yapay sinir ağı

Gerilimli yapay sinir ağlarına kıyasla daha hızlı, kolay eğitim ve öğrenme sürecine sahip, gürültülere karşı daha duyarlı olan olasıksal yapay sinir ağı, Specht tarafından 1988 yılında tanımlanmıştır. Örneklerden bir kerede öğrenir, Bayes optimal karar yüzeylerine sahiptir. Girdi, örüntü, toplama ve çıktı birimleri olmak üzere 4 katmandan oluşur (Şekil 9.2). Örüntü tanımada, doğrusal olmayan eşlemede, sınıf üyeliklerinin olasılıklarını tahminlemede kullanılır. En önemli dezavantajı her girdi örüntüsü için bir düğüm veya nöron gerektirmesidir (Specht, 1992).



Şekil 9.2 PNN modeli örneği (Specht, 1992).

### 9.2.6 Destek vektör makinesi

İstatistiksel öğrenme teorisine dayanan bir sınıflandırma yöntemidir. İkili sınıflandırmada,  $d$  boyutlu girdiyi en uygun biçimde iki gruba ayıran hiper düzlemi bulmayı amaçlar. Sınıfların birbirlerine en yakın örnekleri ile hiper düzleme dik uzaklıklarını maksimize edilmeye, optimum destek vektörlerini bulmaya çalışılır. Doğrusal olarak ayrılamayan veriler ise daha üst boyutlu öznelik uzaylarına eşlenerek yüksek boyutlu bir uzayda görüntülenip, doğrusal SVM problemi olarak hiperdüzlem ile çözülebilir (Boswell, 2002).

### 9.2.7 Sınıflandırma ağacı

Bir veya daha çok tahminleyici değişkenin değerlerine göre verinin bulunduğu sınıfı, bağımlı değişkeni tahminlemek için kullanılmaktadır. Bunun için, kök düğümünden yaprak düğümüne kurallar takip edilir. Yaprak düğümde verinin hangi sınıfa ait olduğu bilgisi tutulmaktadır. Sınıflandırma ağacı, kategorik verileri, regresyon ağacı ise nümerik verileri içermektedir (Mathworks Inc., 2013).



### 9.2.8 Naive bayes sınıflandırıcı

Bayes teoremine dayanan olasılıksal sınıflandırıcıdır. Her özneliğin istatistiksel olarak birbirinden bağımsız olduğu yaklaşımına dayanır. Öznitelikler arasındaki ilişki dikkate alınmaz. Öznitelik vektörü verilen bir örneğin ilgilenilen sınıfa ait olma olasılığı belirlenir. Örüntü tanıma problemlerinde kullanılmaktadır. Eğitimi ve sınıflandırması hızlıdır. Gerçek ve ayık veriyi işleyebilir (Wikipedia, 2013).

### 9.2.9 Doğrusal regresyon

Doğrusal regresyon, girdi ve çıktı değerleri nümerik olan eğitim örneklerine uygun doğrusal fonksiyonu uydurarak değişkenin tahmin edilmesini sağlar. Değişkenler arasındaki analizi ölçmek için de kullanılmaktadır.

Genelleştirilmiş doğrusal model regresyonu ise doğrusal regresyonun genelleştirilmesidir. Yanıt değişkenlerinin normal dağılımdan farklı bir dağılıma sahip olmasına izin verilir.

### 9.2.10 Grup sınıflandırıcılar

Zayıf öğrenme modellerinin birleştirilip yüksek kalitede yeni bir sınıflandırıcının oluşturulması yaklaşımına dayanır. Sınıflandırma kararları her bir sınıflandırıcı için ayrı ayrı değerlendirilir. Her bir sınıflandırıcının sonucundan daha doğru sonuçlar elde edilerek sınıflandırma performansının artırılması amaçlanır (Mathworks, 2012).

## 9.3 Test Ekipmanları

Tez çalışmasında, projedeki tüm verilerin kullanımı ve yazılımlar Matlab R2011a çalışma ortamında Intel i7-2670 QM 2.2 Ghz işlemcili notebook üzerinde gerçekleştirilmiştir.

## 9.4 Kullanılan Veri Seti

Tez çalışmasında, kullanılan veri seti ve özellikleri Çizelge 9.1’de gösterilmiştir. Juxtapleural nodül türünü içeren görüntüler dikkate alınmış, diğer nodül türlerini içeren akciğer CT görüntüleri kullanılmamıştır. Bu görüntüler, nodül bulunan CT kesitleri incelenerek belirlenmiştir. Her kişinin yalnızca 1 CT görüntüsü bulunmakta, her CT görüntüsü çok sayıda kesit içermektedir.

Kullanılan her kesit görüntüsü üzerinde yalnızca 1 tane nodül bulunmaktadır. Görüntülerin bir kısmında, farklı açılarda kesitlerin alınmasından dolayı iki akciğer görüntüsü tam olarak gösterilememekte, ek bölümler de yer alabilmektedir.

Çizelge 9.1 Kullanılan veri setine ilişkin özellikler

<b>Yararlanılan Veri Kümesi</b>	<b>LIDC</b>
<b>Kullanılan Tıbbi Görüntüleme Yöntemi</b>	CT
<b>Nodül Türü</b>	Juxtapleural
<b>Toplam Kişi CT Sayısı</b>	24
<b>Toplam Kesit Sayısı</b>	138
<b>Toplam Aday Nodül Sayısı</b>	1410
<b>Toplam Öznitelik Sayısı</b>	40
<b>Seçilen Öznitelik Sayısı</b>	22
<b>Eğitim Verisinin Boyutu</b>	40 X 1269
<b>Test Verisinin Boyutu</b>	40 X 141
<b>Çapraz Doğrulama için k Değeri</b>	10
<b>Çapraz Doğrulama Bölümleme Türü</b>	Rastgele
<b>Eğitim Verisi – Test Verisi Oranı</b>	90 - 10

Ayrıca, her hasta kişi için CT indislerine göre kaç kesit kullanıldığı ve bu kesitlerden elde edilen toplam aday nodül bölge sayısı Çizelge 9.2’de gösterilmektedir. Toplamda 24 adet kanser hastası için 138 kesit görüntüsü kullanılmış ve 1410 farklı aday nodül bölgesi çıkartılmıştır.

Çizelge 9.2 CT indislerine göre kesit ve aday nodül sayıları

CT İndisi	Toplam Kesit Sayısı	Toplam Aday Nodül Sayısı
1	4	49
2	3	29
3	5	32
4	11	100
5	7	61
6	4	71
7	10	159
8	5	31
9	4	27
10	16	124
11	3	21
12	7	86
13	4	66
14	4	49
15	7	94
16	1	11
17	8	103
18	9	55
19	5	60
20	4	50
21	3	7
22	2	20
23	9	77
24	3	28
<b>TOPLAM</b>	<b>138</b>	<b>1410</b>

## 9.5 Kullanılan Performans Ölçütleri

Tez kapsamında kullanılan performans ölçütlerinin uygunluğu, eğitim setindeki verilerin özelliklerine bağlıdır. Bu nedenle, ilk olarak eğitim kümesinin özellikleri açıklanmaktadır (Çizelge 9.3).

Çizelge 9.3 Eğitim kümesine ilişkin istatistiksel değerler

<b>Toplam Bölge Sayısı</b>	1269
<b>Nodül Olan Bölge Sayısı</b>	124
<b>Nodül Olmayan Bölge Sayısı</b>	1145
<b>Sınıf Çarpıklığı (Pozitif / Negatif)</b>	0.11

Eğitim kümesinin sınıf çarpıklığının (nodül olan bölgelerin sayısının, nodül olmayan bölgelerin sayısına oranının) 1 olmamasından dolayı, tez kapsamında çarpıklığa karşı hassas olmayan ROC eğrisi altında kalan alan AUC performans ölçütü kullanılmıştır. Bunun dışında, genel başarıyı ölçmek için sınıflandırma doğruluk oranı da olmak üzere 2 adet performans ölçütü uygulanmıştır. Bu kriterler alt başlıklarda açıklanmıştır.

### 9.5.1 Sınıflandırma doğruluk oranı

Sınıflandırıcıya verilen örneklerin tahminlenen sınıfı ile gerçek sınıfı arasındaki doğruluk oranının belirlenmesidir. Yani, doğru sonuçların tüm sonuçlara oranına eşittir. Sınıflandırıcının genel başarı performansını belirlemek açısından önemlidir. Sınıflandırma doğruluk oranına ilişkin denklem, Şekil 9.3'te gösterilmektedir.

### 9.5.2 ROC eğrisi

Sınıflandırıcıları performanslarına göre görselleştirmek, düzenlemek ve seçmek için makine öğrenmesi ve veri madenciliğinde kullanılan bir performans kriteridir. Hata (Confusion) matrisi ise gerçek sınıflar ile tahminlenen sınıflar arasındaki ilişkiyi göstermek, pozitif ve negatif durumların sayısına göre performansı ölçmek için kullanılmaktadır. Hata matrisi ve bu matrsten elde edilen ölçütler, Şekil 9.3'te gösterilmektedir.

	T / G	Gerçek Sınıf	
		p	n
Tahminlenen Sınıf	Y	Gerçek Pozitif (GP)	Yanlış Pozitif (YP)
	N	Yanlış Negatif (YN)	Gerçek Negatif (GN)
Sütun Toplamı		P	N

(a)

$$\begin{aligned}
 YP \text{ Oranı (FP Rate)} &= \frac{YP}{N} & GP \text{ Oranı (TP (Hit) Rate)} &= \frac{GP}{P} \\
 Kesinlik (Precision) &= \frac{GP}{GP + YP} & Hassasiyet (Recall) (Sensitivity) &= \frac{GP}{P} \\
 Doğruluk (Accuracy) &= \frac{GP + GN}{P + N} & F - ölçütü &= \frac{2}{\frac{1}{kesinlik} + \frac{1}{hassasiyet}} \\
 Özgüllük (Specificity) &= 1 - YP \text{ Oranı}
 \end{aligned}$$

(b)

Şekil 9.3 Hata matrisi ve bu matristen çıkarılan performans ölçütleri (a) Hata matrisi, (b) Performans ölçütleri (Fawcett, 2006).

Hata matrisinde, eğer örnek pozitifse ve pozitif olarak sınıflandırılmışsa gerçek pozitif, negatif olarak sınıflandırılmışsa yanlış negatif, eğer örnek negatifse ve negatif olarak sınıflandırılmışsa gerçek negatif, pozitif olarak sınıflandırılmışsa yanlış pozitif olarak adlandırılmaktadır (Fawcett, 2006). ROC eğrisi y ekseninde GP oranına, x ekseninde ise YP oranına göre çizilmekte, altında kalan alan (AUC) değerinin 1'e yakın olması sınıflandırıcının başarılı olduğunu göstermektedir.

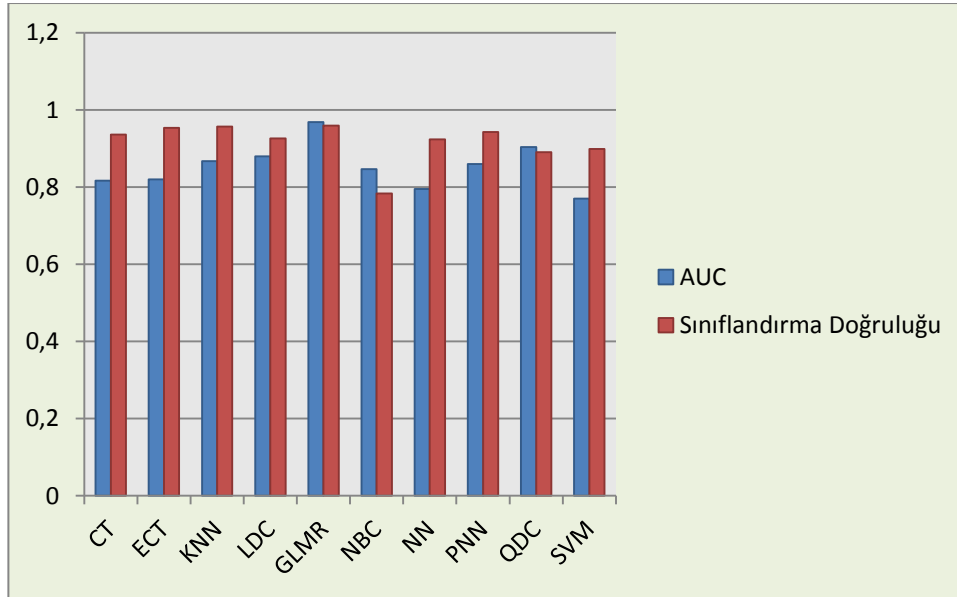
## 9.6 Sınıflandırıcılara Göre Performans Sonuçları

Tez çalışmasında, aday nodül bölgelerinden çıkarılan 40 öznitelikten seçilen 22 öznitelik için ROC öznitelik seçim kriteri, sınıflandırıcılara ve performans ölçütlerine göre 10 iterasyon sonunda elde edilen ortalama sonuçlar Çizelge 9.4'te ve Şekil 9.4'te gösterilmiştir. Sınıflandırma ağacı (CT), grup sınıflandırıcı (ECT), K en yakın komşu sınıflandırıcı (k=1), doğrusal diskriminant sınıflandırıcı (LDC), genelleştirilmiş doğrusal model regresyon (GLMR), Naive Bayes Sınıflandırıcı (NBC), çok katmanlı algılayıcı (MLP), olasılıksal yapay sinir ağı (PNN), ikinci dereceden diskriminant sınıflandırıcı (QDC), destek vektör

makinesi (SVM) sınıflandırıcıları üzerinde AUC ve sınıflandırma doğruluk oranı performans ölçütleri kullanılmıştır.

Çizelge 9.4 Sınıflandırıcılara ve performans ölçütlerine göre elde edilen sonuçlar

Sınıflandırıcı / Performans Ölçütü	AUC	Sınıflandırma Doğruluk Oranı %
CT	0,8162	93,59
ECT	0,8196	95,32
K-NN (K=1)	0,8666	95,69
LDC	0,8793	92,6
GLMR	<b>0,9679</b>	<b>95,88</b>
NBC	0,8465	78,35
MLP	0,795	92,3
PNN (S=0.02)	0,8591	94,24
QDC	0,9034	89,06
SVM	0,7701	89,84



Şekil 9.4 Sınıflandırıcılara ve performans ölçütlerine göre elde edilen sütun grafiği

Şekil 9.4'te görüldüğü gibi, AUC ve sınıflandırma doğruluk oranına göre en iyi sonuç GLMR üzerinde elde edilmiştir. Bu sonucu AUC kriteri için sırasıyla QDC, LDC, K-NN, PNN sınıflandırma doğruluk oranı için ise K-NN, ECT, PNN, CT sınıflandırıcıları izlemektedir.

Yapılan denemeler sonucunda K-NN sınıflandırıcısı için  $k$  değeri 1, PNN için yayılma parametresi 0.02, çok katmanlı yapay sinir ağında ara katman nöron sayısı 5, eğitim - test verisi oranı 80 - 20 olarak belirlenmiş, grup sınıflandırıcılarda ise bagging yöntemi ile sınıflandırma ağaçları 20 adet öğrenme döngüsü yapılarak gerçekleştirilmiştir. GLMR için ise binomial dağılım kullanılmıştır.

Kullanılan diğer sınıflandırıcılardan NBC, SVM ve CT için ise Matlab'de yer alan varsayılan (**default**) parametreler kullanılmıştır. Bu parametreler, **NBC** için Gaussian normal dağılımı (**'Distribution' = 'Normal'**), sınıfların öncelikli olasılıkları eğitim matrisindeki sınıfların ilişkisel frekanslarından belirlenmesi için (**'Prior' = 'Empirical'**), Kernel Smoothing pencere bantgenişliğinin varsayılan değeri Gaussian dağılımı için optimal olan değer kullanılarak her öznelik ve sınıf kombinasyonu için otomatik olarak seçilmesi ile belirlenmiş, (**'KSSupport = unbounded'**) ve (**KSType = normal**)'dir. **SVM sınıflandırıcısı** için varsayılan parametre değerleri sırasıyla **autoscale** = 'true', **boxconstraint** = 1, **kernelcachelimit** = 500, **kernel\_function** = linear, **kktviolationlevel** = 0, **method** = 'SMO', **mlp\_params** = [1 -1], **options** yapısı için Display = off ve MaxIter = 15000, **polyorder** = 3, **rbf\_sigma** = 1, **showplot** = false, **tolkkt** =  $1e-3$ 'tür. **CT** için ise varsayılan parametreler ve değerleri, **CategoricalPredictors** = [], **ClassNames** = hedef sınıf etiketlerinin isimleri (0 veya 1), **Cost** = [Cost(i,j)=1 eğer  $i \neq j$  ve Cost(i,j)=0 eğer  $i=j$ ], **crossval** = off, **kfold** = 10, **MergeLeaves** = 10, **MinLeaf** = 1, **MinParent** = 10, **NVarToSample** = all, **PredictorNames** = {'x1','x2',...}, **prior** = empirical, **Prune** = on, **PruneCriterion** = error, **ResponseName** = Response, **ScoreTransform** = none, **SplitCriterion** = gdi, **Surrogate** = off, **weights** = ones(size(X,1),1) 'dir.

## 9.7 Modeller Üzerinde k-katlı Çapraz Doğrulama Uygulanması

Yapılan tez çalışmasında, veri madenciliğinde ve örüntü tanıma problemlerinde sınıflandırma modellerinin seçimi amacıyla sıklıkla kullanılan k-katlı çapraz doğrulama yöntemi kullanılmıştır.  $k$  değeri varyans ve hata değerini dengelemek için 10 olarak belirlenmiştir. Eğitim kümesinin onda biri doğrulama kümesi olarak ayrılmıştır.

K katlı çapraz doğrulama yönteminde eğitim kümesi eş boyutlu ve rastgele k parçaya ayrılmaktadır. K iterasyonun her aşamasında sırasıyla k. parça test verisi olarak bırakılmakta, diğer veriler ise eğitim verisi olarak belirlenmektedir. (k-1) parçanın eğitime göre uydurulan sınıflandırıcının performansı diğer doğrulama parçasındaki verilere göre belirlenir. k iterasyon için tahminleyicinin hata ortalaması alınarak sınıflandırıcının eğitim kümesi üzerinde genel tahminleme hatası ölçülmüş ve sınıflandırıcı seçimi gerçekleştirilmiştir. Model seçimi sonucunda 0.9551'lik AUC değeri ile en iyi sonuç genelleştirilmiş doğrusal model regresyon üzerinde elde edilmiştir.

## **9.8 Örüntü Vektörlerinden Eğitim, Doğrulama ve Test Kümelerinin Oluşturulması**

Her aday nodül bölgesi üzerinden çıkartılan öznitelik değerlerinden oluşan vektör, örüntü adını almaktadır. Tüm nodül bölgeleri için elde edilen örüntü vektörlerinden, sınıflandırma modellerinin eğitimi, doğrulanması ve değerlendirilmesi için veri setinin bölümlere ayrılması gerekmektedir.

Eğitim kümesi, örüntü vektörlerine göre sınıflandırma modelinin uydurulması, doğrulama kümesi sınıflandırma modelinin seçimi ve tahminleyicinin hatasının belirlenmesi, test kümesi ise seçilen modelin genelleştirme performansını belirleyerek yeni veriler üzerinde tahminleyici hatasının değerlendirilmesi için kullanılmaktadır (Hastie, Tibshirani and Friedman, 2009).

Tez çalışmasında, eğitim verisinin test verisine oranı 90-10'dur. 1410 aday nodül bölgesi için 1269 adet eğitim verisi, 141 adet test verisi örneği bulunmaktadır. Ayrıca, 10 katlı çapraz doğrulama yöntemi ile her sınıflandırıcı için 10 iterasyonun her birinde 1269 verinin 127'ser tanesi doğrulama verisi, diğer bölümleri ise eğitim verisi olarak kullanılmıştır (Bkz. Çizelge 9.1).

## **9.9 Seçilen Modelin Değerlendirilmesi**

10 katlı çapraz doğrulama yöntemiyle, eğitim ve doğrulama kümesindeki verilere göre en iyi performansı gösteren genelleştirilmiş doğrusal model regresyon, sınıflandırıcı modeli olarak seçilmiştir.



Seçilen modelin yeni veriler üzerindeki genelleştirme performansını belirleyebilmek amacıyla test verileri üzerinde seçilen model, AUC ve sınıflandırma doğruluk oranına göre değerlendirilmiştir. Genelleştirilmiş doğrusal model regresyona göre öznitelik sayısı ve seçim kriteri etkenleri dikkate alındığında, 10 iterasyon sonucunda elde edilen en iyi performans sonuçları Çizelge 9.5’te gösterilmiştir.

Çizelge 9.5 GLMR üzerinde öznitelik seçim kriteri ve sayısına göre 10 iterasyon sonucunda elde edilen en iyi performans değerleri

Öznitelik Seçim Kriteri	Öznitelik Sayısı	AUC	Sınıflandırma Doğruluk Oranı %
<b>Ttest</b>	4	0.8713	91.48
<b>Bhattacharyya</b>	9	0.8522	89.57
<b>ROC</b>	7	0.8871	91.88
<b>Entropy</b>	10	0.8722	89.63
<b>Wilcoxon</b>	12	0.8526	91.06
<b>ROC</b>	<b>22</b>	<b>0.7190</b>	<b>85.67</b>

Çizelge 9.5’teki özniteliklerin seçim kriteri ve sayısına göre belirlenen öznitelik indisleri Çizelge 9.6’da gösterilmektedir. Çizelge 9.7’de ise her öznitelik indisi ve bu öznitelik indisine karşılık gelen öznitelik isimleri yer almaktadır.

Çizelge 9.6 GLMR üzerinde öznitelik seçim kriteri ve sayısına göre elde edilen öznitelik indisleri

Ttest (4)	Bhattacharyya (9)		ROC (7)	Entropy (10)		Wilcoxon (12)		ROC (22)		
27	31	19	10	31	19	13	23	10	27	6
13	32		13	32	12	35	27	13	2	25
28	30		28	30		15	6	28	19	23
34	13		21	33		39	20	21	29	18
	33		3	28		22		3	30	16
	28		20	29		19		20	34	5
	27		12	27		26		12	8	
	29			13		4		14	40	

Çizelge 9.7 Öznitelik indisleri ve karşılık gelen öznitelik isimleri

İndis	Ad	İndis	Ad	İndis	Ad	İndis	Ad
1	Area	11	Radius	21	HarMean	31	M5
2	Circularity	12	BoundOverSpan	22	Std	32	M6
3	Eccentricity	13	MAxisOver BPoints	23	Skewness	33	M7
4	Solidity	14	MnAxisOver BPoints	24	Kurtosis	34	autoc
5	Convex Hull Area	15	SpanOverMA	25	Entropy	35	contr
6	Extent	16	MaA_MnAOver BPoints	26	IQR	36	cprom
7	Span	17	SpanOverVar	27	M1	37	cshad
8	Depth	18	F1	28	M2	38	energ
9	BoundPoints	19	Mean	29	M3	39	entro
10	DepthOverSpan	20	GeoMean	30	M4	40	homom

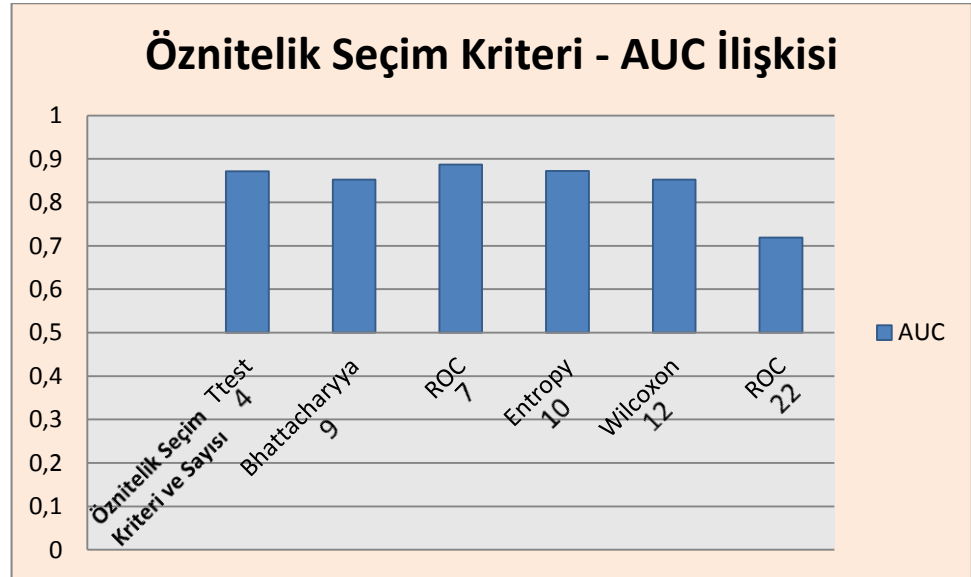
Çizelge 9.6'daki indisler üzerinde Apriori algoritması uygulanarak destek ve güven değerleri ile öznitelik indislerine göre birliktelik kuralları çıkartılmıştır. Bu sayede, her indisin önem derecesi, ortak indisler gibi çeşitli bilgiler elde edilmiştir. Çizelge 9.8'de Apriori algoritmasının çıktısı gösterilmektedir. Apriori algoritması uygulanırken güven değeri 0 olarak belirlenmiştir.

Çizelge 9.8 Çizelge 9.6'daki 5 farklı öznitelik seçim kriteri ve sayısı için indislerin toplamda tekli frekanslarının sayısı

İndis	Adet	İndis	Adet	İndis	Adet	İndis	Adet	İndis	Adet
27	5	33	2	12	3	23	2	18	1
13	6	29	3	35	1	6	2	16	1
28	5	19	4	15	1	14	1	5	1
34	2	10	2	39	1	2	1		
31	2	21	2	22	1	8	1		
32	2	3	2	26	1	40	1		
30	3	20	3	4	1	25	1		

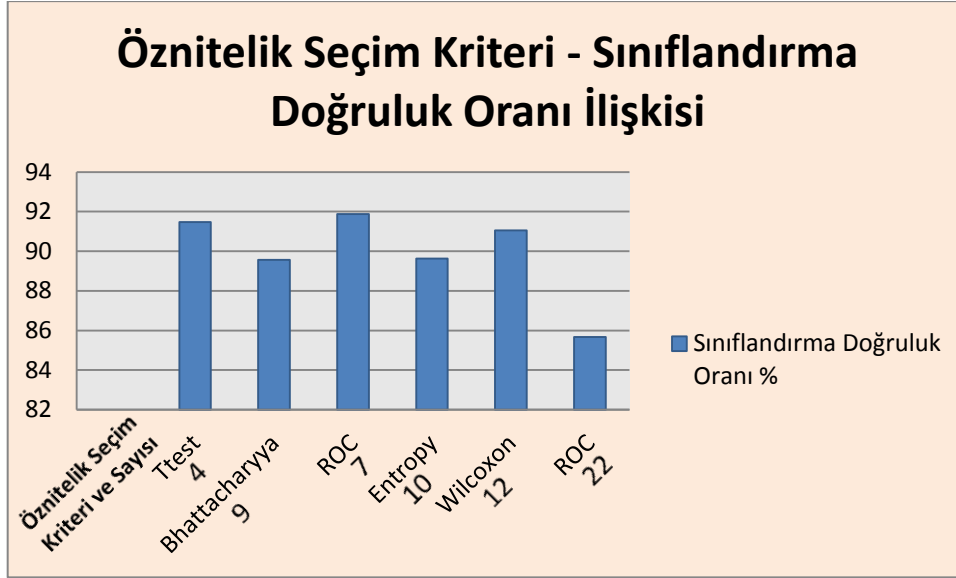
Çizelge 9.7'ye göre, altı durumun hepsinde 13 numaralı indis yer alırken, 13, 27 ve 28 numaralı indisler en az 5 durum içerisinde yer almaktadır. İndisler ikili olarak gruplandırıldığında 5 durumda yer alan indis grupları (27, 13) ile (13, 28)'tir. İndisler üçlü olarak gruplandırıldığında 4 durumda yer alan indis grupları (27, 13, 28) ve (27, 13, 19)'dur. İndisler dördü olarak gruplandırıldığında ise en az 4 durumda yer alan indis grubu bulunmamaktadır. Bu durumdan dolayı, indislerin beşli ve daha yüksek sayıda gruplandırılması ile elde edilen sonuçlar incelenmemiştir.

Çizelge 9.5'teki GLMR sınıflandırıcısı için test kümesi üzerinde, öznitelik seçim kriterleri, en uygun öznitelik sayıları ve AUC değerlerine göre çizilen sütun grafiği Şekil 9.5'te gösterilmiştir. Rastgele sınıflandırıcının AUC performans değerinin 0.5, pozitif bir sınıflandırıcının AUC performansının ise 0.5'ten büyük olması gerektiğinden, 0.5 değerinin üstünde kalan değerler gösterilmiştir. ROC öznitelik seçim kriteri için seçilen 7 öznitelikte 0.8871 AUC değeriyle en iyi sonuç elde edilmiştir. Model seçimi sonucunda belirlenen GLMR ile eğitim kümesi üzerinde öznitelik seçim aşaması sonucunda belirlenmiş 22 öznitelik için en iyi nodül tanıma performansı elde edilmiş olmasına rağmen test kümesi için bu sonuç 0.7190 AUC değerine düşmüştür.



Şekil 9.5 GLMR üzerinde öznitelik seçim kriteri ve sayılarının AUC ile ilişkisini gösteren sütun grafiği

Şekil 9.6'da ise öznitelik seçim kriteri ve sayılarına göre elde edilen sınıflandırma doğruluk oranı ilişkisi, sütun grafiği ile gösterilmiştir.



Şekil 9.6 GLMR üzerinde öznitelik seçim kriteri ve sayılarının sınıflandırma doğruluk oranı ile ilişkisini gösteren sütun grafiği

Test verileri üzerinde, genelleştirilmiş doğrusal model regresyon ile ROC öznitelik seçim kriteri ve 22 öznitelik için 0.7190 AUC değerine ve %85.67 sınıflandırma doğruluk oranına ulaşılmıştır. Eğitim verisi üzerinde elde edilen 0.9679 AUC değerine ve %95.88 sınıflandırma doğruluk oranına göre %25.7 oranında AUC değerinde, %10.6 oranında sınıflandırma doğruluk oranında performans düşüklüğü yaşanmıştır. Bu durumun nedeninin, test verileri için eğitim verilerine göre seçilen öznitelik kümelerinin uygun olmamasından ya da tüm veri kümesinin eğitim ve test kümesine bölünmesi sırasında iki kümedeki aday nodül bölgelerinin yapısal, geometrik ya da dokusal olarak birbirlerinden farklılaşma göstermesinin sonucu olabileceği düşünülmektedir.

## 9.10 Tartışma ve Sonuçlar

Aday nodül bölgelerinden çıkarılmış olan özniteliklerin sınıflandırıcılara göre eğitilmesi ve test edilmesi sonucunda farklı performans sonuçları elde edilmiştir. Farklı türden sınıflandırıcıların problem üzerinde uygulanması sistemin genel başarısını artıran etkenlerden biridir.

Tez çalışmasında, LIDC verisetinde CT görüntüleri üzerinde juxtaopleural akciğer nodülleri için eğitim verilerine ve 22 öznitelik kümesine göre 0.9679 AUC değeriyle en iyi sonuç genelleştirilmiş doğrusal model regresyon sınıflandırıcısında elde edilmiştir.

Eđitim ve test k melerinde yer alan verilerin  zellikleri, her iki k mede aynı sınıflandırıcının farklı AUC ve sınıflandırma dođruluk oranına sahip olmasını sađlamıştır. Bu nedenle, hem seilen sınıflandırıcının yeni verilere g re genelleştirebilme performansı hem de test k mesinin oluřturulması aısından  nem tařımaktadır.

## 10. GENEL DEĞERLENDİRME, SONUÇ VE ÖNERİLER

### 10.1 Genel Değerlendirme

Akciğer kanserinin bilgisayar destekli tespit ve tanılması, en yüksek ölüm oranına sahip kanser türlerinden biri olması, erken tanı ve teşhisi sağlayarak hastaların ölüm oranının azaltılması açısından oldukça önem taşımaktadır.

Bu tez çalışmasında, yapay zeka ve görüntü işleme teknikleri kullanılarak otomatik olarak bilgisayar destekli aday juxtaopleural akciğer nodül bölgelerinin tespitini sağlamaya ve bu aday nodül bölgelerinin tanılanmasına yönelik bir yöntem ve sistem geliştirilmeye odaklanılmıştır. Geliştirilen sistem, önışleme, akciğerlerin bölütlenmesi, aday nodül bölgelerinin çıkartılması, öznitelik çıkarma, seçimi ve sınıflandırma aşamalarından oluşmaktadır.

Geliştirilen sistem ile bilimsel alanda, içbükey aday nodül bölgelerinin alpha hull yöntemi ile belirlenerek, şekil ve doku bilgisine dayalı olarak çıkarılan 7 adet özgün öznitelik ile literatüre yeni bir yöntem sunulmuş, tüm özniteliklerin istatistiksel ölçütler kullanılarak derecelendirilmesi ve seçimiyle performansın artırılması açısından katkılar yapılmıştır.

Tez çalışmasının tıbbi alanda, uzman radyologlara akciğer kanserinin erken tanı ve teşhisi sürecinde yararlı ve yardımcı olması, hastaların sağ kalım oranlarının artırılması, radyologların nodül tespit ve yorumlama hatalarının azaltılması, en aza indirgenmesi, gözden kaçabilen nodüllerin tespitinin kolaylaştırılması, bu işlemler için gereken sürenin kısaltılması ve tanılamanın hızlandırılması, akciğer CT görüntülerinin tekrar kontrolünün ve gereksiz biyopsinin önüne geçilmesi, hem insan hem de bilgisayar özelliklerinin kullanımıyla tanılama kalitesini artırılması gibi çok çeşitli konularda da yararları bulunmaktadır.

## 10.2 Sonuç

Bu tez çalışmasında, aday juxtaopleural nodül bölgelerinin belirlenebilmesi amacıyla akciğer bölgeleri üzerinde  $\alpha$ -hull yöntemi uygulanmıştır. Öznitelik çıkarma aşamasında ise aday nodül bölgelerinden şekil ve/veya doku bilgisine dayalı özgün 7 adet olmak üzere toplamda 40 öznitelik çıkartılmıştır. Bu özniteliklerden 22 tanesi öznitelik seçim aşamasında ROC öznitelik seçim kriteri ile seçilmektedir. Bu sayede, AUC değeri 0.9551'den 0.9679'a çıkarak sistemin genel performansının artırıldığı gözlemlenmiştir. Bu kapsamda, 10 adet sınıflandırıcı için deneysel çalışma yapılmış ve en iyi sonuç, genelleştirilmiş doğrusal model regresyon sınıflandırıcısı üzerinde elde edilmiştir.

Tez kapsamında, gerçekleştirilmiş 8 adet deneysel çalışma bulunmaktadır. Bunlar,  $\alpha$ -hull parametre değerlerinin çıkartılması, öznitelik sayısının sınıflandırıcı performansına etkisinin belirlenmesi, özniteliklerin kendi aralarındaki ve hedef sınıf arasındaki eşli Pearson doğrusal katsayı değerlerine göre Heat Map matrisinin oluşturulması, öznitelik kriterinin ve sayısının seçilen 5 sınıflandırıcı üzerindeki AUC değerlerinin belirlenmesi, öznitelik seçimi olmadan en iyi 5 sınıflandırıcı üzerinde elde edilen AUC değerleri, veri setinden her CT görüntüsü için kullanılan kesit sayısı ve  $\alpha$ -hull yöntemi ile elde edilen toplam aday nodül sayıları, 10 sınıflandırıcı üzerinde 22 öznitelik ve ROC öznitelik seçim kriteri sonucunda elde edilen AUC değeri ile sınıflandırma doğruluk oranı, seçilen GLMR üzerinde öznitelik seçim kriteri ve sayısına göre 10 iterasyon için elde edilen en iyi performans değerleridir.

Gerçekleştirilmiş olan projenin, görüntü işlemede yer alan eşikleme yöntemleri,  $\alpha$ -hull yöntemi, geometrik ve dokusal öznitelikler, öznitelik seçim kriterleri, sınıflandırıcıların karşılaştırılması ve probleme uygunluğu, performans ölçütlerinin belirlenmesi ve uygun olanın kullanılması gibi araştırma alanlarında, araştırmacıya önemli ölçüde katkısı bulunmaktadır.

CAD sistemlerinin tıp ve mühendislik alanlarında giderek önemi artmaktadır. Bu durum, ürüne, yonteme yönelik patentlenebilir çalışma yapılmasını sağlamaktadır.

Aday nodül bölgelerinin çıkarılma yöntemi, özniteliklerin zenginliği ve kalitesi, sınıflandırıcıların performansı açısından önem taşımaktadır.

Tez kapsamının oldukça geniş olması, bu konuda çalışan araştırmacının örüntü tanıma, makine öğrenmesi, yapay zekâ, görüntü işleme, geometrik algoritmalar, radyoloji, istatistik gibi çeşitli çalışma alanlarında bilgili olmasını gerektirmektedir. Ayrıca, tıp alanında mühendislik gerektiren proje bazlı çalışmalarda doktor ve mühendis iletişimi oldukça önem taşımaktadır. Veri kümelerinin elde edilmesi, tıbbi terim ve ifadelerin anlaşılması, gerekli altyapının öğrenilmesi gibi durumlar açısından uzman radyolog ve mühendisin karşılıklı iletişimi gereklidir. Bu durum, çalışmanın genel başarısının artırılması için de oldukça önem taşımaktadır.

### 10.3 Öneriler ve Gelecekteki Çalışmalar

Akciğer bölütleme sırasında kullanılan Otsu eşikleme yöntemi etkili olmasına rağmen diğer eşikleme yöntemlerinin uygulanması ya da metasezgisel algoritmaların kullanılmasıyla daha başarılı sonuçlar elde edilebilir. Aday nodül bölgelerinin belirlenmesi aşamasında  $\alpha$ -hull yöntemi kullanılmıştır. Bu yöntemin dışında içbükey bölgelerin belirlenmesi amacıyla geometrik algoritmalar, morfolojik işlemler uygulanabilir.

Öznitelik çıkarma aşamasında, şekle ve dokuya dayalı özniteliklerin dışında SIFT (Scale Invariant Feature Transform), SURF (Speeded up Robust Features), LBP (Local Binary Pattern), texton gibi yöntemlerde kullanılarak öznitelik zenginliği sağlanabilir ve öznitelik kalitesi artırılabilir.

Öznitelik seçiminde, tez kapsamında filtre yöntemlerinden istatistiksel ölçütler kullanılmasının dışında sarmalayıcı, gömülü yöntemler veya genetik algoritmalar uygulanarak sistemin performansı artırılabilir. Yapay zekâ yöntemlerine dayanan, probleme uygun yeni sınıflandırıcıların tasarlanması da genel başarıyı etkileyecektir.



İleriki çalışmalarda, diğer akciğer nodül türlerini de dikkate alan bir sistem geliştirilerek, çalışmanın kapsamı genişletilebilir. Ayrıca, her aşama için farklı yöntem ve yaklaşımlar, geliştirilmiş sistem üzerinde uygulanarak yöntemlerin kendi aralarında karşılaştırılması sağlanabilir. Akciğer CT görüntüleri hazır verisetleri dışında, yerel hastanelerden de elde edilerek deneysel sonuçlar gerçekleştirilebilir.

## KAYNAKLAR DİZİNİ

- Akgül, E.**, Röntgen Cihazları ve Fizik Prensipleri, <http://lokman.cu.edu.tr/radyodiagnostik/dersler/shmyo/rontgen%20fiz%20ve%20cihaz/9%20attenuasyon.ppt> , (Erişim Tarihi: 18 Şubat 2013).
- Awai, K., Murao, K., Ozawa, A., Komi, M., Hayakawa, H., Hori, S. and Nishimura, Y.**, 2004, Pulmonary nodules at chest CT: effect of computer-aided diagnosis on radiologists' detection performance, *Radiology*, 230(2):347-352pp.
- Antonelli, M., Lazzerini, B. and Marcelloni, F.**, 2005, Segmentation and reconstruction of the lung volume in CT images. In *Proceedings of the ACM Symposium on Applied Computing*. 255–259pp.
- Armato III, S. G., Giger, M. I., Moran, C. J., Blackburn, J. T., Doi, K., and Macmahon, H.**, 1999, Computerized detection of pulmonary nodules on CT scans, *Radiographics*, 19(5): 1303–1311pp.
- Armato, S. G., Giger M. L. and MacMahon, H.**, 2001, Automated detection of lung nodules in CT scans: Preliminary results, *Medical Physics*, 28(8):1552–1561pp.
- American Cancer Society**, 2013, “Cancer Facts & Figures 2013”, <http://www.cancer.org/acs/groups/content/@epidemiologysurveillance/documents/document/acspc-036845.pdf>, (Erişim Tarihi: 18 Şubat 2013).
- American Cancer Society**, 2013, “Lung Cancer (Non-Small Cell)”, <http://www.cancer.org/acs/groups/cid/documents/webcontent/003115-pdf.pdf>, (Erişim Tarihi: 18 Şubat 2013).
- American Lung Association**, 2013, “Knowing the Basics”, <http://www.lung.org/lung-disease/lung-cancer/learning-more-about-lung-cancer/understanding-lung-cancer/knowing-the-basics.html>, (Erişim Tarihi: 18 Şubat 2013).
- Arimura, H., Katsuragawa, S. and Suzuki, K., Li, F., Shiraishi, J., Sone, S. and Doi, K.**, 2004, “Computerized scheme for automated detection of lung nodules in low dose computed tomography images for lung cancer screening, *Academic Radiology*, 11(6): 617–629pp.
- Boswell, D.**, 2002, Introduction to Support Vector Machines, <http://www.work.caltech.edu/~boswell/IntroToSVM.pdf>, (Erişim Tarihi 28 Mayıs 2013).
- Brown, M. S., McNitt-Cray, M. F., Golldin, J. G., Suh, R. D., Sayre, J. W. and Aberle, D. R.**, 2001, Patient-specific models for lung nodule detection and surveillance in CT Images, *Medical Imaging (IEEE Transactions on)*, 20(12):1242-1250pp.
- Cancer Research UK**, 2013, PET Scan, <http://www.cancerresearchuk.org/cancer-help/about-cancer/tests/pet-scan> (Erişim Tarihi: 18 Şubat 2013).

## KAYNAKLAR DİZİNİ (devam)

- Cornell University**. 2007, ELCAP public lung image database, Vision and Image Analysis Group (VIA) and International Early Lung Cancer Action Program (I-ELCAP) Labs, <http://www.via.cornell.edu/lungdb.html>, (Erişim Tarihi: 18 Şubat 2013).
- Cornell University**, 2008, Public lung database to address drug response. Vision and Image Analysis Group (VIA) and International Early Lung Cancer Action Program (I-ELCAP) Labs, <http://www.via.cornell.edu/crpf.html>, (Erişim Tarihi: 18 Şubat 2013).
- Diciotti, S., Picozzi, G., Falchini, M., Mascalchi, M., Villari, N. and Valli, G.**, 2008, 3-D Segmentation Algorithm of Small Lung Nodules in Spiral CT Images, *Information Technology in Biomedicine (IEEE Transactions on)*, 12(1):7-19pp.
- Doi, K.**, 2007, Computer-aided diagnosis in medical imaging: Historical review, current status and future potential, *Computerized Medical Imaging and Graphics*, 31(4-5):198-211pp.
- Edelsbrunner, H., Kirkpatrick D. G., and Seidel, R.**, 1983, On the Shape of a Set of Points in the Plane, *Information Theory (IEEE Transactions on)*, 29(4): 551-559pp.
- El-Baz, A., Beache, G. M., Gimel'farb, G., Suzuki, K., Okada, K., Elnakib, A., Soliman, A. and Abdollahi B.**, 2013, Computer-Aided Diagnosis Systems for Lung Cancer: Challenges and Methodologies, *International Journal of Biomedical Imaging*, 2013, Article ID 942353, 46 pages.
- Eleyan, A. and Demirel, H.**, 2011, Co-occurrence matrix and its statistical features as a new approach for face recognition, *Turk J Elec Eng & Comp Sci*, 19(1) :97-107pp.
- Evans, J. D.**, 1995, Straightforward statistics for the behavioral sciences. Pacific Grove, CA: Brooks/Cole Publishing, 624p.
- Farag, A., El-Baz, A., Gimelfarb, G.G., Falk, R. and Hushek, S.G.**, 2004, Automatic detection and recognition of lung abnormalities in helical CT images using deformable templates, *Lecture Notes in Computer Science*, Springer-Verlag, *Medical Image Computing and Computer-Assisted Intervention*, 3217:856-864pp.
- Flusser, J., Suk, T. and Zitová, B.**, 2009, Moments and Moment Invariants in Pattern Recognition, John Wiley & Sons, Ltd., 312p.
- Frangi, A.F., Niessen, W.J., Hoogeveen R.M., Walsum, T.V. and Viergever, M.A.**, 1999, Model-based quantitation of 3-D magnetic resonance angiographic images, *Medical Imaging (IEEE Transactions on)*, 18(10):946-56pp.
- Frank, T.K., Loriot, S. and Yvinec, M.**, 2004, 3D Alpha Shapes, [http://www.ics.uci.edu/~dock/manuals/cgal\\_manual/Alpha\\_shapes\\_3/Chapter\\_main.html](http://www.ics.uci.edu/~dock/manuals/cgal_manual/Alpha_shapes_3/Chapter_main.html), (Erişim Tarihi 22 Nisan 2013).

## KAYNAKLAR DİZİNİ (devam)

- Ge, Z.Y., Sahiner, B., Chan, H. P., Hadjiiski, L. M., Cascade, P. N., Bogot, N., Kazerooni, E. A., Wei, J. and Zhou, C.,** 2005, Computer-aided detection of lung nodules: False positive reduction using a 3D gradient field method and 3D ellipsoid fitting, *Medical Physics*, 32(8): 2443-2454pp.
- Gonzalez, R., C. and Woods, R. E.,** 2002, Digital Image Processing, Prentice-Hall, New Jersey, 793pp.
- Ginsberg, M. S., Akin, O., Berger, D. M., Zakowski, M. F. and Panicek, D., M.,** 2004, Pulmonary Tumorlets: CT Findings, *American Journal of Roentgenology*, 183(2):293-296pp.
- Gurcan, M., Sahiner B. and Petrick, N., Chan H.P., Kazerooni E. A., Cascade, P. N. and Hadjiiski,** 2002, Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system, *Medical Physics*, 29(11):2552-2558pp.
- Guyon, I. and Elisseeff, A.,** 2003, An Introduction to Variable and Feature Selection, *Journal of Machine Learning Research*, 3:1157-1182pp.
- Haralick, R. M., Shanmugam, K. and Dinstein, I.,** 1973, Textural features for image classification, *Systems, Mans and Cybernetics (IEEE Transactions on)*, 3(6):610-621pp.
- Hastie, T., Tibshirani, R. and Friedman, J.,** 2009, The Elements of Statistical Learning: Data Mining, Inference and Prediction, Second Edition, Springer, 763p.
- Hayashi, H., Ashizawa, K., Uetani, M., Futagawa, S., Fukushima, A., Minami, K., Honda, S. and Hayashi, K.,** 2009, Detectability of peripheral lung cancer on chest radiographs:effect of the size, location and extent of ground-glass opacity, *The British Journal of Radiology*, 82(976) :272-278pp.
- Hearn, D., D., Baker, M. P. and Carithers, W. R.,** 2011, Computer Graphics with OpenGL, Prentice Hall, Fourth edition, 881pp.
- Hu, M-K,** 1962, Visual Pattern Recognition by Moment Invariants, *Information Theory (IEEE Transactions on)*, 8(2):179-187pp.
- Hu, S., Hoffman, E.A. and Reinhardt, J. M.,** 2001, Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images,” *Medical Imaging (IEEE Transactions on)*, 20(6):490–498pp.
- Japanese Society of Radiological Technology,** 1998, JSRT Digital Image Database, <http://www.jsrt.or.jp/jsrt-db/eng.php>.
- Jemal A., Murray T., Ward E., Samuels, A., Tiwari, R., C., Ghafoor, A., Feuer, E., J., Thun, M., J.,** 2005, CA Cancer J. Clin. 55:10–30pp.
- Kartaloğlu, Z.,** 2008, Soliter Pulmoner Nodüle Yaklaşım, *Turkish J Thorac Cardiovasc Surg* 2008;16(4):274-283pp.
- Ko, J. P. and Betke, M.,** 2003, Chest CT: automated nodule detection and assessment of change over time—preliminary experience, *Radiology*., 218(1): 267–273pp.

## KAYNAKLAR DİZİNİ (devam)

- Kostis, W., J., Reeves, A., P., Yankelevitz, D. F. and Henschke, C., I.,** 2003, "Three-dimensional segmentation and growth rate estimation of small pulmonary nodules in helical CT images", *IEEE Trans. Med. Imag.*, 22(10): 1259 -1274pp.
- Lee, S. L. A., Kouzani, A. Z. and Hu E. J.,** 2012, Automated detection of lung nodules in computed tomography images: a review, *Machine Vision and Applications*, 23(1):151-163pp.
- Lee, Y., Hara, T., Fujita, H., Itoh, S. and Ishigaki, T.,** 2001, "Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique," *Medical Imaging (IEEE Transactions on)*, 20(7):595-604pp.
- Li, Q. and Doi, K.,** 2004, New selective nodule enhancement filter and its application for significant improvement of nodule detection on computed tomography, in *Medical Imaging: Imaging Processing*, Proceedings of SPIE, 1–9pp.
- Li, Q., Li, F. and Doi K.,** 2008, Computerized detection of lung nodules in thin-section CT images by use of selective enhancement filters and an automated rule-based classifier, *Acad. Radiology*, 15(2):165 -175pp.
- Marjan, C.,** 2006, "Computational Topology Seminar – Alpha Shapes", Max Planck Institute, <http://www.mpiinf.mpg.de/~jgiesen/tch/sem06/Celikik.pdf>, (Erişim Tarihi 22 Nisan 2013).
- Math Works Inc.,** Matlab R2011a documentation, <http://www.mathworks.com/help/index.html>, (Erişim Tarihi: 30 Nisan 2012).
- MathWorks Inc.,** Classification Trees and Regression Trees, <http://www.mathworks.com/help/stats/classification-trees-and-regression-trees.html>, (Erişim Tarihi 30 Mayıs 2013).
- Matsumoto, S., Kundel, H.L., Gee, J.C., Geftter, W.B. and Hatabu, H,** 2006, Pulmonary nodule detection in CT images with quantized convergence index filter, *Medical Image Analysis*, 10 (3) : 343–352pp.
- Matsumoto, M., Tateno, Y. and Takeshi, I.,** 1996, Quoit filter-a new filter based on mathematical morphology to extract the isolated shadow, and its application to automatic detection of lung cancer in X-ray CT, *Pattern Recognition Proceedings of the 13th International Conference on*, 2:3-7pp.
- Mingqiang, Y., Kidiyo, K. and Joseph, R.,** 2008, A survey of shape feature extraction techniques, *Pattern Recognition Techniques, Technology and Applications*, 43-90pp.
- Mughal, M.N. and Ikram, W.,** 2004, Early Lung Cancer Detection by Classifying Chest CT Images: A Survey, *Proceedings of the 8th International Multitopic Conference, Proceedings of (INMIC' 04)*, 67–72pp.

## KAYNAKLAR DİZİNİ (devam)

- National Electrical Manufacturers Association**, 1999, “DICOM Supplement 33: Grayscale Softcopy Presentation State Storage”, [ftp://medical.nema.org/medical/dicom/final/sup33\\_ft.pdf](ftp://medical.nema.org/medical/dicom/final/sup33_ft.pdf), 87p, (Erişim Tarihi 30 Mayıs 2013).
- National Lung Screening Trial**, 2009, NLST, <https://biometry.nci.nih.gov/cdas/studies/nlst/>.
- National Cancer Institute**, RSNA 2003 LIDC Educational Exhibit: Fundamental Issues for the Creation of a Resource for the Image Processing Research Community, <http://imaging.cancer.gov/programsandresources/reportsandpublications/reportsandpresentations/rsna2003lidceducationalexhibit>, (Erişim tarihi: 2 Eylül 2012).
- National Cancer Institute, Lung Image Database Consortium**, LIDC-IDRI, <https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>, (Erişim Tarihi: 18 Şubat 2013).
- National Cancer Institute**, 2012, “Stages”, <http://www.cancer.gov/cancertopics/wyntk/lung/page6>, (Erişim Tarihi: 22 Şubat 2013).
- Nunzio, G. De., Massafra, A., Cataldo, R., Mitri, I. De, Peccarisi, M., Fantacci, M.E., Gargano and G., Torres, E. L.**, Approaches to juxta-pleural nodule detection in CT images within the MAGIC-5 Collaboration, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 648(1):103-106pp.
- Ochs, R.A., Goldin, J.G., Fereidoun, A., Kim, H.J., Brown, K., Batra, P., Roback D., McNitt-Gray, M.F. and Brown, M.F.**, 2007, Automated classification of lung bronchovascular anatomy in CT using Adaboost, *Medical Image Analysis*, 11(3):315–324pp.
- Paik, D.S., Beaulieu, C.F., Rubin, G.D., Acar B., Jeffrey R.B., Yee J., Dey, J. and Napel S.**, 2004, Surface normal overlap: a computer-aided detection algorithm with application, *Medical Imaging (IEEE Transactions on)*, 23(6):661-675pp.
- Pless, R.**, 2011, “Alpha hulls, shapes and weighted things”, <http://research.engineering.wustl.edu/~pless/546/lectures/lecture22.pdf>, (Erişim Tarihi 22 Nisan 2013).
- Pu, J., Leader J. K., Zheng B., Knollmann, F., Fuhrman, C., Sciurba, F. C. and Gur, D.**, 2009, A computational geometry approach to automated pulmonary fissure segmentation in CT examinations, *Medical Imaging (IEEE Transactions on)*, 28(5):710–719pp.
- Pu, J., Roos, J., Yi, C. A., Napel, S., Rubin, G. D., and Paik, D. S.**, 2008. Adaptive border marching algorithm:automatic lung segmentation on chest CT images. *Computerized Med. Imaging Graph.* 32(6):452–462pp.

## KAYNAKLAR DİZİNİ (devam)

- Pretoux F., Merlet, N., Grenier, P. and Mouehli, M.,** 1989, Algorithms for automated evaluation of pulmonary lesions by high resolution CT via image analysis, in *Proc. Radiol. Soc. Of North America RSNA '89*, 1989(416).
- Retico, A., Delogu, P., Fantacci, M.E., Gori, I. and Martinez, A.P.,** 2008, Lung nodule detection in low-dose and thin-slice computed tomography, *Computers in Biology and Biomedicine*, 38(4):525-534pp.
- Ross, J. C., Estépar, R. S., Diaz A., Westin, C. F., Kikinis, R., Silverman, E. K. and Washko, G. R.,** 2009, Lung extraction, lobe segmentation and hierarchical region assessment for quantitative analysis on high resolution computed tomography images, in *Proceedings of the International Conference on Medical Imaging Computing and Computer-Assisted Intervention (MICCAI'09)*, 5762: 690–698pp.
- Sensakovic, W. F., Starkey, A. and Armato, S.G.,** 2008, A general method for the identification and repair of concavities in segmented medical images, *IEEE Nuclear Science Symposium Conference Record*, 5320-5326pp.
- Shi, Y., Qi, F., Xue, Z., Chen, L., Ito, K., Matsuo, H. and Shen, D.,** 2008, Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics, *Medical Imaging (IEEE Transactions on)*, 27(4): 481–494pp.
- Sluimer, I. C.,** 2005, Automated image analysis of the pathological lung in CT, Phd Thesis, University Medical Center Utrecht, 157p (published).
- Soh, L-K. and Tsatsoulis, C.,** 1999, Texture Analysis of SAR Sea Ice Imagery Using Gray Level Co-Occurrence Matrices, *Geoscience and Remote Sensing (IEEE Transactions on)*, 37(2) : 780-795pp.
- Specht, D. F.,** 1992, Enhancements to Probabilistic Neural Networks, *International Joint Conference on (IJCNN)*, 1:761-768pp.
- Srinivasan, G.N. and Shobha, G.,** 2008, Statistical Texture Analysis, *Proceedings of World Academy of Science, Engineering and Technology*, 36:6p.
- Suárez-Cuenca, J.J., Tahoces, P. G., Souto, M., Lado, M.J., Remy-Jardin, M., Remy, J. and Vidal, J.J.,** 2009, Application of the iris filter for automatic detection of pulmonary nodules on computed tomography images, *Computers in Biology and Medicine*, 39(10): 921-933pp.
- Sun, S., Bauer, C. and Beichel, R.,** 2012, Automated 3-D Segmentation of Lungs With Lung Cancer in CT Data Using a Novel Robust Active Shape Model Approach, *Medical Imaging (IEEE Transactions on)*, 31(2):449-460pp.
- Suzuki, K., Armota, S. G., Li, F., Sone, S. and Doi, K.,** 2003, Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography, *Medical Physics*, 30(7):1602-1617pp.



## KAYNAKLAR DİZİNİ (devam)

- Takizawa, H., Shigemoto, K., Yamamoto S., Matsumoto, T., Tatenno, Y., Iinuma, T. and Matsumoto, M.,** 2003, A recognition method of lung nodule shadows in X-Ray CT images using 3D object models,” *International Journal of Image and Graphics*, 3(4):533–545pp.
- Theodoridis, S. and Koutroumbas, K.,** 2009, Pattern Recognition, Fourth Edition, Academic Press, US, 967p.
- Theodoridis, S. and Koutroumbas, K.,** 2010, An Introduction to Pattern Recognition: A Matlab Approach, Pattern Recognition, First Edition, Academic Press, US, 216p.
- Tüm Radyoloji Teknisyenleri ve Teknikerleri Derneği (TÜMRAD-DER),** 2013, “Bilgisayarlı Tomografi (BT)”, [http://www.tumradder.com/FileUpload/ds58732/File/bilgisayarli\\_tomograf1\\_bt\\_.pdf](http://www.tumradder.com/FileUpload/ds58732/File/bilgisayarli_tomograf1_bt_.pdf), , (Erişim Tarihi: 3 Mart 2013).
- Türk Radyoloji Derneği (TRD) Yeterlilik Kurulu, Rehberler ve Standartlar Komitesi,** 2007, İki yönlü akciğer grafisi, 4s.
- Ulaş, M. and Boyacı, A.,** 2007, DICOM Görüntü Standardı, Akademik Bilişim’07-IX. Akademik Bilişim Konferansı Bildirileri, 2:69-74s.
- Uğur, Ö.,** 2004, Pozitron Emisyon Tomografisi, XIII. TPOG Ulusal Pediatrik Kanser Kongresi,4s.
- Ünal, D.,** 2008, Tıpta Kullanılan Görüntüleme Teknikleri, Gazi Üniversitesi Eğitim Fakültesi, Lisans Tezi, 36s, (yayınlanmamış).
- Van Ginneken, B., Stegmann M. B. and Loog, M.,** 2006, Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database, *Medical Image Analysis*, 10(1):19–40pp.
- Ye, X., Lin, X., Dehmeshki, J., Slabaugh, G. and Beddoe, G.,** 2009, Shape-based computer-aided detection of lung nodules in thoracic CT images, *Biomedical Engineering, (IEEE Transactions on)*, 56(7):1810-1820pp.
- Wikipedia,** 2013, Naive Bayes Classifier, [https://en.wikipedia.org/wiki/Naive\\_Bayes\\_classifier](https://en.wikipedia.org/wiki/Naive_Bayes_classifier), (Erişim Tarihi 20 Haziran 2013).
- World Health Organization,** 2013, “Cancer, Fact Sheet No 297”, <http://www.who.int/mediacentre/factsheets/fs297/en/>, (Erişim Tarihi: 18 Şubat 2013).
- World Health Organization,** 2012, “The World Health Report”, [http://www.who.int/whr/2004/annex/topic/en/annex\\_2\\_en.pdf](http://www.who.int/whr/2004/annex/topic/en/annex_2_en.pdf), (Erişim tarihi: 2 Eylül 2012).



## **EKLER**

- Ek 1 Türkçe İngilizce Sözlük
- Ek 2 İngilizce Türkçe Sözlük
- Ek 3 Öznitelik Seçim Kriteri ve Sayısının Belirlenmesinde Elde Edilen Deneysel Sonuçlar

## EKLER

### EK-1 Türkçe İngilizce Sözlük

Türkçe	İngilizce
Adenokarsinom	Adenocarcinoma
Akciğer Bölütleme	Lung Segmentation
Akciğer Kanseri	Lung Cancer
Akciğer Parankiması	Lung Parenchyma
Aktif Görünüm Modeli	Active Appearance Model
Aktif Şekil Modeli	Active Shape Model
Alan	Area
Alfa Şekil	Alpha Shape
Alıcı İşletim Karakteristiği	Receiver Operating Characteristic
Anatomi Tabanlı Genel Model	Anatomy Based Generic Model
Aritmetik Ortalama	Arithmetic Mean
Basıklık	Kurtosis
Biçim Değiştirebilir Model	Deformable Model
Bilgisayar Destekli Tespit ve Tanı	Computer Aided Detection and Diagnosis
Bilgisayar Destekli Tanı	Computer Aided Diagnosis
Bilgisayar Destekli Tespit	Computer Aided Detection
Bilgisayarlı Tomografi	Computerized Tomography
Bölge Dolgulama	Region Filling
Büyük Hücreli Karsinom	Large Cell Carcinoma
Manyetik Rezonans Görüntüleme	Magnetic Resonance Imaging
Çarpıklık	Skewness
Çeyrek Değerler Genişliği	Interquartile Range
Çok Katmanlı Algılayıcı	Multi Layer Perceptron
Dairesellik Oranı	Circularity Ratio
Damara Yakın Nodül	Juxta-vascular Nodule
Derinliğin Kapsama Oranı	Depth Over Span
Destek Vektör Makinesi	Support Vector Machine
Destek Vektör Makinesi – Özyinelemeli Öznitelik Eleme	Support Vector Machine – Recursive Feature Elimination
Derinlik	Depth
Dışbükey Kabuk	Convex Hull
Dışbükey Kabuğun Alanı	Convex Area

<b>Dışmerkezlilik</b>	Eccentricity
<b>Doğrusal Diskriminant Sınıflandırıcı</b>	Linear Discriminant Classifier
<b>Düzgün Sınırlı Nodül</b>	Isolated Nodule
<b>Entropi, Bilgi Yitimi</b>	Entropy
<b>Genelleştirilmiş Doğrusal Model Regresyon</b>	Generalized Linear Model Regression
<b>Geometrik Ortalama</b>	Geometric Mean
<b>Gerçek Negatif</b>	True Negative
<b>Gerçek Pozitif</b>	True Positive
<b>Görüntü İşleme</b>	Image Processing
<b>Gri Düzey Eş Oluşum Matrisi</b>	Gray Level Co-occurrence Matrix
<b>Grup Sınıflandırıcı</b>	Ensemble Classifier
<b>Güçlü Aktif Şekil Modeli</b>	Robust Active Shape Model
<b>Harmonik Ortalama</b>	Harmonic Mean
<b>Hata Matrisi</b>	Confusion Matrix
<b>Homojenlik</b>	Homogeneity
<b>Hounsfield Ünitesi</b>	Hounsfield Unit
<b>İkinci Dereceden Diskriminant Sınıflandırıcı</b>	Quadratic Discriminant Classifier
<b>İnceltme</b>	Thinning
<b>K-En Yakın Komşu Algoritması</b>	K-Nearest Neighbor Algorithm
<b>K-Katlı Çapraz Doğrulama</b>	K-Fold Cross Validation
<b>Kapsam</b>	Span, Extent
<b>Katılık</b>	Solidity
<b>Kesin Referans</b>	Ground Truth
<b>Kontrast</b>	Contrast
<b>Küçük Hücreli Akciğer Kanseri</b>	Small-Cell Lung Cancer
<b>Küçük Hücreli Dışı Akciğer Kanseri</b>	Non-small Cell Lung Cancer
<b>Küme Önemi</b>	Cluster Prominence
<b>Küresellik Yapı</b>	Sphericity Structure
<b>İyileştirme Filtresi</b>	Enhancement Filter
<b>Makine Öğrenmesi</b>	Machine Learning
<b>Merkezi İşlem Birimi</b>	Central Processing Unit
<b>Naive Bayes Sınıflandırıcı</b>	Naive Bayes Classifier
<b>Nesne Tabanlı Deformasyon</b>	Object Based Deformation
<b>Nodül Tanıma</b>	Nodule Recognition

<b>Nokta İyileştirme Filtresi</b>	Dot Enhancement Filter
<b>Olasılıksal Sinir Ağı</b>	Probabilistic Neural Network
<b>Otokorelasyon</b>	Autocorrelation
<b>Ölçekten Bağımsız Öznitelik Dönüşümü</b>	Scale Invariant Feature Transform
<b>Önek</b>	Prefix
<b>Önsöz</b>	Preamble
<b>Örüntü Tanıma</b>	Pattern Recognition
<b>Öznitelik Çıkarma</b>	Feature Extraction
<b>Pencereleme</b>	Windowing
<b>Pencere Genişliği</b>	Window Width
<b>Pencere Düzeyi</b>	Window Level
<b>Piksel Verisi</b>	Pixel Data
<b>Plevraya Yakın Nodül</b>	Juxtapleural Nodule
<b>Pozitron Emisyon Tomografi</b>	Positron Emission Tomography
<b>Seçici İyileştirme Filtresi</b>	Selective Enhancement Filter
<b>Sınıflandırma Ağacı</b>	Classification Tree
<b>Sınır Uzunluğu</b>	Boundary Length
<b>Sınırların Yeniden Oluşturulması</b>	Border Reconstruction
<b>Skvamöz Hücreli Karsinom</b>	Squamous Cell Carcinoma
<b>Standart Sapma</b>	Standard Deviation
<b>Tıbbi Görüntü Analizi</b>	Medical Image Analysis
<b>Veri Elemanları</b>	Data Elements
<b>Yanlış Negatif</b>	False Negative
<b>Yanlış Pozitif</b>	False Positive
<b>Yarıçap</b>	Radius
<b>Yerel İkili Örüntü</b>	Local Binary Pattern
<b>3 Boyutlu</b>	3 Dimensional

**EK-2 İngilizce Türkçe Sözlük**

<b>İngilizce</b>	<b>Türkçe</b>
<b>Active Appearance Model</b>	Aktif Görünüm Modeli
<b>Active Shape Model</b>	Aktif Şekil Modeli
<b>Adenocarcinoma</b>	Adenokarsinom
<b>Alpha Shape</b>	Alfa Şekil
<b>Anatomy Based Generic Model</b>	Anatomi Tabanlı Genel Model
<b>Area</b>	Alan
<b>Arithmetic Mean</b>	Aritmetik Ortalama
<b>Autocorrelation</b>	Otokorelasyon
<b>Border Reconstruction</b>	Sınırların Yeniden Oluşturulması
<b>Boundary Length</b>	Sınır Uzunluğu
<b>Central Processing Unit</b>	Merkezi İşlem Birimi
<b>Circularity Ratio</b>	Dairesellik Oranı
<b>Classification Tree</b>	Sınıflandırma Ağacı
<b>Cluster Prominence</b>	Küme Önemi
<b>Computer Aided Detection and Diagnosis</b>	Bilgisayar Destekli Tespit ve Tanılama
<b>Computer Aided Detection</b>	Bilgisayar Destekli Tespit
<b>Computer Aided Diagnosis</b>	Bilgisayar Destekli Tanı
<b>Computerized Tomography</b>	Bilgisayarlı Tomografi
<b>Confusion Matrix</b>	Hata Matrisi
<b>Contrast</b>	Kontrast
<b>Convex Area</b>	Dışbükey Kabuğun Alanı
<b>Convex Hull</b>	Dışbükey Kabuk
<b>Data Elements</b>	Veri Elemanları
<b>Deformable Model</b>	Biçim Değiştirebilir Model
<b>Depth</b>	Derinlik
<b>Depth Over Span</b>	Derinliğin Kapsama Oranı
<b>Dot Enhancement Filter</b>	Nokta İyileştirme Filtresi
<b>Eccentricity</b>	Dışmerkezlilik
<b>Ensemble Classifier</b>	Grup Sınıflandırıcı
<b>Entropy</b>	Entropi, Bilgi Yitimi
<b>False Negative</b>	Yanlış Negatif
<b>False Positive</b>	Yanlış Pozitif
<b>Feature Extraction</b>	Öznitelik Çıkarma
<b>Generalized Linear Model Regression</b>	Genelleştirilmiş Doğrusal Model Regresyon
<b>Geometric Mean</b>	Geometrik Ortalama
<b>Gray Level Co-occurrence Matrix</b>	Gri Düzey Eş Oluşum Matrisi

<b>Ground Truth</b>	Kesin Referans
<b>Harmonic Mean</b>	Harmonik Ortalama
<b>Homogeneity</b>	Homojenlik
<b>Hounsfield Unit</b>	Hounsfield Ünitesi
<b>Image Processing</b>	Görüntü İşleme
<b>Interquartile Range</b>	Çeyrek Değerler Genişliği
<b>Isolated Nodule</b>	Düzgün Sınırlı Nodül
<b>Juxtapleural Nodule</b>	Plevraya Yakın Nodül
<b>Juxta-vascular Nodule</b>	Damara Yakın Nodül
<b>K-Fold Cross Validation</b>	K-Katlı Çapraz Doğrulama
<b>K-Nearest Neighbor Algorithm</b>	K-En Yakın Komşu Algoritması
<b>Kurtosis</b>	Basıklık
<b>Large Cell Carcinoma</b>	Büyük Hücreli Karsinom
<b>Linear Discriminant Classifier</b>	Doğrusal Diskriminant Sınıflandırıcı
<b>Local Binary Pattern</b>	Yerel İkili Örüntü
<b>Lung Cancer</b>	Akciğer Kanseri
<b>Lung Parenchyma</b>	Akciğer Parankiması
<b>Lung Segmentation</b>	Akciğer Bölütleme
<b>Machine Learning</b>	Makine Öğrenmesi
<b>Magnetic Resonance Imaging</b>	Manyetik Rezonans Görüntüleme
<b>Medical Image Analysis</b>	Tıbbi Görüntü Analizi
<b>Multi Layer Perceptron</b>	Çok Katmanlı Algılayıcı
<b>Naive Bayes Classifier</b>	Naive Bayes Sınıflandırıcı
<b>Nodule Recognition</b>	Nodül Tanıma
<b>Non-small Cell Lung Cancer</b>	Küçük Hücreli Dışı Akciğer Kanseri
<b>Object Based Deformation</b>	Nesne Tabanlı Deformasyon
<b>Pattern Recognition</b>	Örüntü Tanıma
<b>Pixel Data</b>	Piksel Verisi
<b>Positron Emission Tomography</b>	Pozitron Emisyon Tomografi
<b>Preamble</b>	Önsöz
<b>Prefix</b>	Önek
<b>Probabilistic Neural Network</b>	Olasılıksal Sinir Ağı
<b>Quadratic Discriminant Classifier</b>	İkinci Dereceden Diskriminant Sınıflandırıcı
<b>Radius</b>	Yarıçap
<b>Receiver Operating Characteristic</b>	Alıcı İşletim Karakteristiği
<b>Region Filling</b>	Bölge Dolgulama
<b>Robust Active Shape Model</b>	Güçlü Aktif Şekil Modeli
<b>Scale Invariant Feature Transform</b>	Ölçekten Bağımsız Öznelik Dönüşümü
<b>Selective Enhancement Filter</b>	Seçici İyileştirme Filtresi

<b>Skewness</b>	Çarpıklık
<b>Small-Cell Lung Cancer</b>	Küçük Hücreli Akciğer Kanseri
<b>Solidity</b>	Katılık
<b>Span, Extent</b>	Kapsam
<b>Sphericity Structure Enhancement Filter</b>	Küresellik Yapı İyileştirme Filtresi
<b>Squamous Cell Carcinoma</b>	Skvamöz Hücreli Karsinom
<b>Standard Deviation</b>	Standart Sapma
<b>Support Vector Machine</b>	Destek Vektör Makinesi
<b>Support Vector Machine – Recursive Feature Elimination</b>	Destek Vektör Makinesi – Özyinelemeli Öznitelik Eleme
<b>Thinning</b>	İnceltme
<b>True Negative</b>	Gerçek Negatif
<b>True Positive</b>	Gerçek Pozitif
<b>Window Level</b>	Pencere Düzeyi
<b>Window Width</b>	Pencere Genişliği
<b>Windowing</b>	Pencereleme
<b>3 Dimensional</b>	3 Boyutlu

**EK-3 Öznitelik Seçim Kriteri ve Sayısının Belirlenmesinde****Elde Edilen Deneysel Sonuçlar**

Öznitelik Seçim Kriteri	ROC
Toplam Öznitelik Sayısı	40
İncelenen Öznitelik Sayısı	25
Çapraz Doğrulama k Değeri	10
Performans Kriteri	ROC
Seçilen Sınıflandırıcı Sayısı	5
İterasyon Sayısı	10

Öznitelik Sayısı	GLMR	QDC	LDC	K-NN (K = 1)	PNN (S = 0.02)
25	0.9645	0.9031	0.8720	0.8687	0.8611
24	0.9639	0.9062	0.8726	0.8701	0.8647
23	0.9639	0.9062	0.8726	0.8701	0.8647
22	<b>0.9679</b>	0.9034	0.8793	0.8666	0.8591
21	0.9631	0.9049	0.8745	0.8693	0.8605
20	0.9610	0.8861	0.8716	0.8726	0.8606
19	0.9609	0.8925	0.8670	0.8684	0.8657
18	0.9608	0.8876	0.8689	0.8669	0.8579
17	0.9602	0.8874	0.8705	0.8650	0.8632
16	0.9610	0.8834	0.8697	0.8610	0.8546
15	0.9573	0.8779	0.8604	0.8380	0.8329
14	0.9554	0.8827	0.8406	0.8282	0.8299
13	0.9552	0.8768	0.8424	0.8273	0.8257
12	0.9568	0.8851	0.8454	0.8254	0.8268
11	0.9564	0.8880	0.8469	0.8221	0.8246
10	0.9556	0.8823	0.8220	0.7980	0.7980
9	0.9559	0.8851	0.8224	0.7807	0.7881
8	0.9514	0.8744	0.8226	0.7788	0.7848
7	0.9470	0.8699	0.8074	0.8068	0.8027
6	0.9453	0.8677	0.8240	0.8050	0.8034
5	0.9297	0.8540	0.8112	0.7795	0.7761
4	0.9302	0.8682	0.8074	0.7675	0.7686
3	0.9298	0.8299	0.8050	0.7344	0.7330
2	0.9287	0.8323	0.8079	0.7467	0.7289
1	0.9311	0.8092	0.8057	0.6915	0.8074



**EK-3 Öznitelik Seçim Kriteri ve Sayısının Belirlenmesinde****Elde Edilen Deneysel Sonuçlar**

Öznitelik Seçim Kriteri	TTest
Toplam Öznitelik Sayısı	40
İncelenen Öznitelik Sayısı	25
Çapraz Doğrulama k Değeri	10
Performans Kriteri	ROC
Seçilen Sınıflandırıcı Sayısı	5
İterasyon Sayısı	10

Öznitelik Sayısı	GLMR	QDC	LDC	K-NN (K = 1)	PNN (S = 0.02)
25	0.9635	0.8970	0.8765	0.8801	0.8724
24	0.9634	0.8856	0.8676	0.8763	0.8652
23	<b>0.9649</b>	0.8955	0.8640	0.8773	0.8708
22	0.9639	0.8928	0.8681	0.8668	0.8603
21	0.9612	0.8858	0.8739	0.8618	0.8583
20	0.9616	0.8883	0.8757	0.8649	0.8605
19	0.9595	0.8850	0.8694	0.8651	0.8607
18	0.9594	0.8958	0.8655	0.8673	0.8615
17	0.9592	0.8951	0.8457	0.8656	0.8580
16	0.9583	0.8957	0.8457	0.8668	0.8635
15	0.9585	0.8911	0.8596	0.8738	0.8694
14	0.9590	0.8884	0.8563	0.8765	0.8697
13	0.9590	0.8874	0.8600	0.8573	0.8522
12	0.9596	0.8777	0.8752	0.8701	0.8667
11	0.9604	0.8702	0.8775	0.8644	0.8617
10	0.9560	0.8784	0.8517	0.8421	0.8392
9	0.9552	0.8787	0.8467	0.8471	0.8411
8	0.9554	0.8765	0.8499	0.8125	0.8134
7	0.9449	0.8657	0.8404	0.8242	0.8208
6	0.9449	0.8583	0.8446	0.8036	0.8011
5	0.9399	0.8629	0.8536	0.8050	0.8015
4	0.9339	0.8446	0.8483	0.7891	0.7888
3	0.9260	0.8236	0.8339	0.7307	0.6987
2	0.9266	0.8385	0.8289	0.7177	0.6948
1	0.9165	0.8441	0.8127	0.7272	0.6350

**EK-3 Öznitelik Seçim Kriteri ve Sayısının Belirlenmesinde****Elde Edilen Deneysel Sonuçlar**

Öznitelik Seçim Kriteri	Entropy
Toplam Öznitelik Sayısı	40
İncelenen Öznitelik Sayısı	25
Çapraz Doğrulama k Değeri	10
Performans Kriteri	ROC
Seçilen Sınıflandırıcı Sayısı	5
İterasyon Sayısı	10

Öznitelik Sayısı	GLMR	QDC	LDC	K-NN (K = 1)	PNN (S = 0.02)
25	0.9598	0.8661	0.8657	0.8655	0.8621
24	0.9598	0.8667	0.8573	0.8603	0.8609
23	0.9542	0.8634	0.8497	0.8498	0.8509
22	0.9536	0.8575	0.8517	0.8521	0.8454
21	0.9570	0.8598	0.8526	0.8463	0.8401
20	0.9543	0.8568	0.8522	0.8249	0.8248
19	0.9546	0.8547	0.8520	0.8503	0.8461
18	<b>0.9602</b>	0.8550	0.8512	0.8256	0.8185
17	0.9582	0.8553	0.8503	0.8392	0.8553
16	0.9572	0.8538	0.8550	0.8494	0.8485
15	0.9484	0.8530	0.8433	0.8183	0.8232
14	0.9473	0.8510	0.8467	0.8141	0.8098
13	0.9333	0.8500	0.8463	0.8163	0.8163
12	0.9371	0.8385	0.8385	0.8111	0.8080
11	0.9400	0.8321	0.8224	0.8205	0.8124
10	0.9178	0.8236	0.8190	0.8020	0.7830
9	0.8480	0.8218	0.8480	0.7756	0.7652
8	0.8724	0.8232	0.8324	0.7446	0.7060
7	0.9172	0.8233	0.8371	0.7471	0.6386
6	0.9038	0.8228	0.8107	0.7378	0.4996
5	0.8783	0.8196	0.7903	0.6902	0.4996
4	0.8606	0.7227	0.6545	0.6696	0.5000
3	0.8651	0.7139	0.6544	0.6711	0.5000
2	0.7055	0.7085	0.6188	0.6798	0.5000
1	0.7396	0.6130	0.5387	0.6252	0.5000

**EK-3 Öznitelik Seçim Kriteri ve Sayısının Belirlenmesinde****Elde Edilen Deneysel Sonuçlar**

Öznitelik Seçim Kriteri	Bhattacharyya
Toplam Öznitelik Sayısı	40
İncelenen Öznitelik Sayısı	25
Çapraz Doğrulama k Değeri	10
Performans Kriteri	ROC
Seçilen Sınıflandırıcı Sayısı	5
İterasyon Sayısı	10

Öznitelik Sayısı	GLMR	QDC	LDC	K-NN (K = 1)	PNN (S = 0.02)
25	0.9615	0.8712	0.8562	0.8619	0.8594
24	0.9523	0.8666	0.8587	0.8592	0.8635
23	0.9548	0.8633	0.8486	0.8555	0.8480
22	0.9598	0.8582	0.8496	0.8537	0.8533
21	0.9560	0.8571	0.8515	0.8422	0.8349
20	0.9527	0.8556	0.8563	0.8483	0.8413
19	0.9546	0.8547	0.8520	0.8503	0.8461
18	0.9614	0.8535	0.8513	0.8194	0.8168
17	<b>0.9617</b>	0.8541	0.8488	0.8394	0.8401
16	0.9581	0.8539	0.8530	0.8474	0.8466
15	0.9490	0.8520	0.8445	0.8243	0.8216
14	0.9418	0.8492	0.8413	0.8150	0.8145
13	0.9373	0.8503	0.8480	0.8163	0.8193
12	0.9256	0.8369	0.8348	0.8101	0.8089
11	0.9303	0.8338	0.8238	0.8197	0.8132
10	0.9183	0.8240	0.8220	0.8035	0.7891
9	0.8294	0.8233	0.8480	0.7773	0.7689
8	0.8979	0.8215	0.8316	0.7393	0.7033
7	0.9245	0.8197	0.8308	0.7349	0.6989
6	0.9244	0.8178	0.8299	0.6831	0.6719
5	0.9207	0.7292	0.8317	0.7121	0.6733
4	0.9234	0.7195	0.8320	0.7159	0.6746
3	0.8656	0.7141	0.6544	0.6651	0.5000
2	0.7055	0.7083	0.6190	0.6849	0.5000
1	0.7383	0.6129	0.5383	0.9323	0.5000

**EK-3 Öznitelik Seçim Kriteri ve Sayısının Belirlenmesinde****Elde Edilen Deneysel Sonuçlar**

Öznitelik Seçim Kriteri	Wilcoxon
Toplam Öznitelik Sayısı	40
İncelenen Öznitelik Sayısı	25
Çapraz Doğrulama k Değeri	10
Performans Kriteri	ROC
Seçilen Sınıflandırıcı Sayısı	5
İterasyon Sayısı	10

Öznitelik Sayısı	GLMR	QDC	LDC	K-NN (K = 1)	PNN (S = 0.02)
25	0.9624	0.8891	0.8597	0.8846	0.8667
24	0.9629	0.8909	0.8586	0.8861	0.8705
23	0.9628	0.8781	0.8367	0.8812	0.8650
22	0.9628	0.8752	0.8404	0.8866	0.8741
21	0.9628	0.8705	0.8379	0.8825	0.8695
20	0.9631	0.8731	0.8380	0.8731	0.8702
19	0.9634	0.8725	0.8504	0.8800	0.8706
18	0.9620	0.8773	0.8493	0.8847	0.8718
17	0.9622	0.8808	0.8484	0.8800	0.8755
16	0.9629	0.8813	0.8499	0.8838	0.8763
15	<b>0.9636</b>	0.8826	0.8490	0.8870	0.8789
14	0.9623	0.8903	0.8479	0.8789	0.8771
13	0.9582	0.8746	0.8395	0.8767	0.8667
12	0.9515	0.8801	0.8255	0.8769	0.8711
11	0.9511	0.8882	0.8249	0.8624	0.8603
10	0.9528	0.8856	0.8221	0.8625	0.8572
9	0.9418	0.8775	0.8203	0.8537	0.8482
8	0.9422	0.8732	0.8180	0.8570	0.8566
7	0.9421	0.8784	0.8154	0.8434	0.8455
6	0.9289	0.8636	0.8015	0.8473	0.8500
5	0.9291	0.8520	0.8023	0.8193	0.8251
4	0.9305	0.8538	0.7950	0.7731	0.7719
3	0.9277	0.8640	0.7953	0.7794	0.7853
2	0.9257	0.8545	0.8317	0.7647	0.7165
1	0.9255	0.8319	0.8329	0.7084	0.6727

