

Akciğerlerdeki Farklı Dokuların Otomatik Tespit ve Sınıflandırılması

Ali Çınar

**YÜKSEK LİSANS TEZİ**

Elektrik Elektronik Mühendisliği Anabilim Dalı

Mayıs 2019

The Automatic Detection and Classification of Various Tissues in Lungs

Ali Çınar

**MASTER OF SCIENCE THESIS**

Department of Electrical and Electronics Engineering

May 2019

Akciğerlerdeki Farklı Dokuların Otomatik Tespit ve Sınıflandırılması

Ali Çınar

Eskişehir Osmangazi Üniversitesi  
Fen Bilimleri Enstitüsü  
Lisansüstü Yönetmeliği Uyarınca  
Elektrik Elektronik Mühendisliği Anabilim Dalı  
Telekomünikasyon - Sinyal İşleme Bilim Dalında  
YÜKSEK LİSANS TEZİ  
Olarak Hazırlanmıştır

Danışman: Doç. Dr. Semih Ergin

Mayıs 2019

## ONAY

Elektrik Elektronik Mühendisliđi Anabilim Dalı Yüksek Lisans öđrencisi Ali Çınar'ın YÜKSEK LİSANS tezi olarak hazırladıđı “Akciđerlerdeki Farklı Dokuların Otomatik Tespit ve Sınıflandırılması” başlıklı bu çalıřma, jürimizce lisansüstü yönetmeliđin ilgili maddeleri uyarınca deđerlendirilerek oybirliđi ile kabul edilmiřtir.

**Danıřman** : Doç. Dr. Semih Ergin

**İkinci Danıřman** : --

**Yüksek Lisans Tez Savunma Jürisi:**

**Üye** : Doç. Dr. Semih Ergin

**Üye** : Prof. Dr. M. Bilginer Gülmezođlu

**Üye** : Dr. Öğr. Üyesi İdil Iřıklı Esener

Fen Bilimleri Enstitüsü Yönetim Kurulu'nun .....tarih ve  
.....sayılı kararıyla onaylanmıřtır.

Prof. Dr. Hürriyet ERŐAHAN  
Enstitü Müdürü

## ETİK BEYAN

Eskişehir Osmangazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna göre, Doç. Dr. Semih Ergin danışmanlığında hazırlamış olduğum “Akciğerlerdeki Farklı Dokuların Otomatik Tespit ve Sınıflandırılması” başlıklı YÜKSEK LİSANS tezimin özgün bir çalışma olduğunu; tez çalışmamın tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; tezimde verdiğim bilgileri, verileri akademik ve bilimsel etik ilke ve kurallara uygun olarak elde ettiğimi; tez çalışmamda yararlandığım eserlerin tümüne atıf yaptığımı ve kaynak gösterdiğimi ve bilgi, belge ve sonuçları bilimsel etik ilke ve kurallara göre sunduğumu beyan ederim. 02/05/2019

Ali Çınar

## ÖZET

Bu tezde, akciğer x-ray röntgenleri kullanılarak akciğerdeki şüpheli dokular otomatik olarak tespit edilmiş ve şüpheli dokular iyi huylu nodül, kötü huylu nodül ve nodülsüz olarak üç sınıfa ayrılmıştır. İlk önce akciğer dokusu dışındaki dokular eşikleme yöntemi ile çıkarılmıştır. Akciğer röntgen imgelerinde kaburga kemikleri nodüllü bölge ile örtüşebilir. Bu durum nodülün saptanmasını zorlaştıracığından kaburga kemiklerinin saptanması ve bastırılması gerekmektedir. Şablon eşleme yöntemi ile kaburga kemiklerinin yerleri saptanmış ve bu yerlere Gabor filtresi uygulanarak kemikler bastırılmıştır. Bu aşamadan sonra Kenarsız Chan-Vese algoritması kullanılarak akciğerdeki şüpheli dokular bulunmuş ve bu dokulardan öznitelikler çıkarılmıştır. İstatistiksel öznitelikler, Yönlendirilmiş Gradyanların Histogramı öznitelikleri, Yerel İkili Örüntü öznitelikleri, Geometrik öznitelikler, Gri Seviye Eş-Oluşum Matrisi öznitelikleri ve Ölçekten Bağımsız Yoğun Öznitelik Dönüşümü olmak üzere altı farklı öznitelik çıkarımı yapılmıştır. Daha sonra bu öznitelikler kullanılarak sınıflandırma aşamasına geçilmiştir. İstatistiksel öznitelikler, Yönlendirilmiş Gradyanların Histogramı öznitelikleri ve Yerel İkili Örüntü öznitelikleri kullanılarak öznitelik vektörü oluşturulduğunda daha iyi sınıflandırma sonuçları gözlemlenmiştir. Sınıflandırma sonuçları hassasiyet, kesinlik ve doğruluk üzerinden değerlendirilmiştir. Sınıflandırmada,  $k$ -En Yakın Komşu, Karar Ağacı, Rastgele Orman, Lojistik Regresyon Sınıflandırıcısı, Destek Vektör Makineleri, Fisher Doğrusal Ayırtaç Analizi ve Naive Bayes sınıflandırıcıları kullanılmıştır. Rasgele orman sınıflandırıcısı %56 hassasiyet, %65 doğruluk ve %81 kesinlik ile en iyi sonuçları vermiştir.

**Anahtar Kelimeler:** Akciğer kanseri, kaburga kemiklerini bastırmak, bilgisayar destekli teşhis, şüpheli doku saptamak, x-ray röntgen

## SUMMARY

Suspicious regions in lung x-rays were detected automatically and these regions were classified into three groups like malignant nodule, benign nodule and healthy. Firstly, using thresholding method non-lung areas were removed. Ribs may overlap with nodules in lung x-rays. Since this will make the detection of the nodule difficult, it is necessary to detect and suppress the ribs. The location of the rib bones was determined by the template matching method and the bones were suppressed by applying Gabor filter. After this stage, suspicious tissues in the lung were found by using the Chan-Vese without edge algorithm and features were extracted from these tissues. Six different attributes were extracted: Statistical, Histogram of Oriented Gradients, Local Binary Pattern, Geometric, Gray Level Co-Occurrence Matrix and Dense Scale Invariant Feature Transform. Then classification stage was started using these features. Best classification result was obtained using statistical, local binary pattern and histogram of gradients features. Classification results were evaluated with precision, accuracy and accuracy.  $k$ -Nearest Neighbor, Decision Tree, Random Forest, Logistic Linear Classifier, Support Vector Machines, Fisher's Linear Discriminant Analysis and Naive Bayes were used in the classification. The random forest classifier gave the best results with 56% sensitivity, 65% accuracy and 81% specificity.

**Keywords:** Lung cancer, suppression of ribs, computer aided diagnosis, detection of suspicious regions, x-ray

## TEŐEKKÜR

Yüksek Lisans çalışmam boyunca bana öğretmiş olduđu değerli bilgilerden dolayı danışmanım Doç. Dr. Semih Ergin'e teşekkür ederim. Çalışmamda bana yardımları dokunan arkadaşım Bengisu Topuz'a teşekkürü bir borç bilirim. Ayrıca bu süreçte bana hep destek olan aileme teşekkürlerimi sunarım.



## İÇİNDEKİLER

### Sayfa

<b>ÖZET</b> .....	<b>vi</b>
<b>SUMMARY</b> .....	<b>vii</b>
<b>TEŞEKKÜR</b> .....	<b>viii</b>
<b>İÇİNDEKİLER</b> .....	<b>ix</b>
<b>ŞEKİLLER DİZİNİ</b> .....	<b>xi</b>
<b>ÇİZELGELER DİZİNİ</b> .....	<b>xiii</b>
<b>SİMGELER VE KISALTMALAR DİZİNİ</b> .....	<b>xiii</b>
<b>1. GİRİŞ VE AMAÇ</b> .....	<b>1</b>
<b>2. LİTERATÜR ARAŞTIRMASI</b> .....	<b>3</b>
2.1. Kaburga Kemiklerinin Bastırılması.....	3
2.2. Akciğer Bölütlenmesi.....	5
2.3. Akciğer kanseri sınıflandırılması üzerine yapılan çalışmalar .....	6
<b>3. AKCİĞER RÖNTGEN İMGELERİNİN ÖN İŞLENMESİ</b> .....	<b>8</b>
3.1. Akciğer Bölütlenmesi için Yapılan Önışleme .....	8
3.1.1. Eşikleme .....	8
3.2. Kaburga Kemiklerinin Bastırılması için Yapılan Önışleme.....	8
3.2.1. Medyan Filtre .....	8
3.2.2. İkileştirme .....	9
3.2.3. Açma İşlemi .....	9
3.2.4. Kapama İşlemi.....	10
3.2.5. Şablon Eşleme .....	10
3.2.6. Gabor Filtresi.....	10
3.2.7. Gauss Filtresi.....	11
<b>4. AKCİĞER RÖNTGEN İMGELERİNDE ŞÜPHELİ BÖLGE TESPİTİ</b> .....	<b>12</b>
4.1. Mumford ve Shah Minimizasyon Problemi .....	12
4.2. Kenarsız Chan-Vese Aktif Kontur Modeli.....	13
<b>5. AKCİĞER RÖNTGEN İMGELERİNDE TESPİT EDİLEN ŞÜPHELİ DOKULARDAN ÖZİNİTELİK ÇIKARTIMI</b> .....	<b>16</b>
5.1. İstatistiksel Özinitelikler.....	16
5.2. Yönlendirilmiş Gradyanların Histogramı Özinitelikleri.....	17
5.3. Yerel İkili Örüntü Özinitelikleri.....	19

## İÇİNDEKİLER (devam)

### Sayfa

5.4.	Geometrik Öznitelikler.....	20
5.5.	Gri Seviye Eş-Oluşum Matrisi Öznitelikleri.....	21
5.6.	Ölçekten Bağımsız Öznitelik Dönüşümü Öznitelikleri.....	25
5.6.1.	Ölçek Uzayı Uç Noktaları Saptama .....	25
5.6.2.	Düşük Kontrastlı Bölgelerin ve Kenarların Elenmesi .....	26
5.6.3.	Dönme Açısı Belirlenmesi .....	26
5.6.4.	Hücrelerin Betinlenmesi.....	27
5.7.	Ölçekten Bağımsız Yoğun Öznitelik Dönüşümü Öznitelikleri .....	27
<b>6.</b>	<b>AKCİĞER RÖNTGEN İMGELERİNİN SINIFLANDIRILMASI.....</b>	<b>28</b>
6.1.	k-En Yakın Komşu Sınıflandırıcısı .....	28
6.2.	Karar Ağacı Sınıflandırıcısı .....	28
6.3.	Rasgele Orman Sınıflandırıcısı .....	29
6.4.	Fisher Doğrusal Ayırtaç Analizi Sınıflandırıcısı.....	29
6.5.	Destek Vektör Makineleri Sınıflandırıcısı .....	30
6.6.	Naive Bayes Sınıflandırıcısı.....	31
6.7.	Lojistik Doğrusal Sınıflandırıcısı .....	32
<b>7.</b>	<b>MATERYAL VE YÖNTEM .....</b>	<b>33</b>
7.1.	Kullanılan Veritabanı .....	33
7.2.	Akciğer Röntgen İmgelerinin Önışlenmesi .....	33
7.3.	Akciğer Röntgen İmgelerinde Şüpheli Doku Tespiti .....	37
7.4.	Öznitelik Vektörünün Oluşturulması .....	39
7.5.	Şüpheli Olarak Tespit Edilen Dokuların Sınıflandırılması .....	40
<b>8.</b>	<b>BULGULAR VE TARTIŞMA .....</b>	<b>41</b>
<b>9.</b>	<b>SONUÇ VE ÖNERİLER .....</b>	<b>45</b>
	<b>KAYNAKLAR DİZİNİ.....</b>	<b>47</b>

## ŞEKİLLER DİZİNİ

### Sekil

### Sayfa

4. 1. Chan-Vese algoritmasının çalışma prensibi .....	14
5. 1. HOG Yönteminin Uygulanması .....	18
5. 2. GLCM matrisinin oluşumu.....	22
6. 1. SVM sınıflandırıcısı çalışma prensibi .....	30
7. 1. JSRT veritabanından örnek bir imge .....	34
7. 2. Bölütlenmiş imge.....	34
7. 3. Medyan filtrelenmiş ve bölütlenmiş imge farkı .....	34
7. 4. İkileştirme uygulanmış imge .....	34
7. 5. Kapama işlemi uygulanmış imge .....	35
7. 6. Açma işlemi uygulanmış imge .....	35
7. 7. Şablon eşleme uygulanmış imge .....	36
7. 8. Filtrenin büyüklük tepkisi.....	36
7. 9. İmgelerin toplamı .....	36
7. 10. Yeni kemiklerin orijinal imgede yerine konması .....	36
7. 11. Medyan filtre uygulanmış imge.....	37
7. 12. Açma işlemi uygulanmış imge .....	37
7. 13. Gauss filtre uygulanmış imge .....	37
7. 14. Başlangıç konturu.....	38
7. 15. Siyah piksellerin gri yapılması .....	38
7. 16. Şüpheli bölgenin bölütlenmesi .....	38

## ÇİZELGELER DİZİNİ

### Cizelge

### Sayfa

5. 1. İstatistiksel öznitelikler ve matematiksel ifadeleri .....	16
5. 2. GLCM öznitelikleri ve matematiksel ifadeleri .....	22
7. 1. Değerlendirme kriterleri ve matematiksel ifadeleri .....	33
8. 1. Sadece istatistiksel öznitelikler kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri .....	41
8. 2. Sadece GLCM öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri .....	41
8. 3. Sadece LBP öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri .....	41
8. 4. Sadece geometrik öznitelikler kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri .....	42
8. 5. Sadece HOG öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri .....	42
8. 6. Sadece DSIFT öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri .....	42
8. 7. İstatistiksel ve HOG öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri .....	43
8. 8. İstatistiksel ve DSIFT öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri .....	43
8. 9. İstatistiksel ve LBP öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri .....	43
8. 10. İstatistiksel, DSIFT, HOG öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri .....	44
8. 11. İstatistiksel, LBP, HOG öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri .....	44

## SİMGELER VE KISALTMALAR DİZİNİ

<u>Simgeler</u>	<u>Açıklama</u>
C	Fisher Doğrusal Ayırtaç Analizi Sınıflandırıcısında kovaryans matrisi
$d$	GLCM matrisi oluşturulurken kullanılan uzaklık parametresi
D	Geometrik özniteliklerden biri olan çap
O	Geometrik özniteliklerden biri olan ovallik
Y	Geometrik özniteliklerden biri olan yuvarlaklık
$\beta$	Fisher Doğrusal Ayırtaç Analizi Sınıflandırıcısında doğrusal model katsayıları
$\sigma$	Gauss filtresinde standart sapma parametresi
$\phi$	Gabor filtresinde, Yönlendirilmiş Gradyanların Histogramı algoritmasında, Gri Seviye Eş-Oluşum Matrisi algoritmasında ve Ölçekten Bağımsız Öznitelik Dönüşümü Algoritmasında Yön parametresi
$\theta$	Gabor filtresinde uzamsal frekans parametresi
$\ominus$	Açma ve kapama morfolojik işlemlerinde genişletme operatörü
$\oplus$	Açma ve kapama morfolojik işlemlerinde aşındırma operatörü

## SİMGELER VE KISALTMALAR DİZİNİ (devam)

<b><u>Kısaltmalar</u></b>	<b><u>Açıklama</u></b>
DOG	Difference of Gaussian (Gauss Dağılımlı Farklar)
DSIFT	Dense Scale Invariant Feature Transform (Ölçekten Bağımsız Yoğun Öznitelik Dönüşümü)
GLCM	Gray Level Co-Occurrence Matrix (Gri Seviye Eş-Oluşum Matrisi)
HOG	Histogram of Gradients (Yönlendirilmiş Gradyanların Histogramı)
JSRT	Japanese Society of Radiological Technology (Japon Radyolojik Teknoloji Derneği)
k-NN	k-Nearest Neighbor (k-En Yakın Komşu)
LBP	Local Binary Pattern (Yerel İkili Örüntü)
LLC	Logistic Linear Classifier (Lojistik Regresyon Sınıflandırıcısı)
MTANN	Massive Training Artificial Neural Network (Çok Çözünürlüklü Kitlesel Eğitim Yapay Sinir Ağı)
RCGA	Real-coded genetic algorithm (Gerçek kodlu genetik algoritma)
ROI	Region of Interests (İlgi Bölgeleri)
RIU2	Rotation Invariant Uniform-2 (Dönme Açısından Bağımsız Düzgün)
SIFT	Scale Invariant Feature Transform (Ölçekten Bağımsız Öznitelik Dönüşümü)

## 1. GİRİŞ VE AMAÇ

Akciğer kanseri, dünyada ölüme en çok sebep olan kanser türlerinden biridir (Stewart ve Wild, 2015). Yapısal olarak normal akciğer dokusundan olan hücrelerin kontrol dışı çoğalması sonucu akciğer içinde bir kitle oluşmasına dayanmaktadır. Akciğer üzerindeki bu nodül yapısı genellikle bir belirti yerine akciğer radyolojik görüntülenmesi sırasında fark edilirler. Nodülün şekli, yapısı gibi etkenler nodülün iyi veya kötü huylu olabileceğini kabaca gösterir. Ancak kesin tanı için nodülün çıkarılması ve patolojik incelemelerin yapılması gereklidir. Günümüzde akciğer üzerindeki bulunan ve çeşitli nedenler sonucu oluşan nodüller yapılarına ve dokudaki yerlerine göre farklı görüntüleme metotları (x-ray, tomografi, MRI vb.) kullanılmaktadır. Ancak bu nodül yapısının genel olarak hastadan hastaya göre değişiklik göstermesi, tespitinin zor ve zahmetli olması nedeniyle görüntüleme sistemlerinde yeni yaklaşımlara ihtiyaç duyulmaktadır.

Akciğer kanserini yenmek için diğer kanser türlerinde olduğu gibi erken teşhis çok önemlidir. X-ray görüntüleme sistemi dünya çapında yaygın olarak kullanılan en eski teşhis sistemlerinden birisidir. Bu sistemden elde edilen görüntülerde kemiklerin siyah, yumuşak dokunun ise daha açık renkli görünür. Bu kontrast farkı sayesinde kemikler ve yumuşak doku içerisinde bulunan nodül kolayca ayırt edilebilmektedir. Bunun yanı sıra x-ray'in düşük maliyetli ve kolay ulaşılabilir olması, az radyasyon yayması ve nodül gelişimin daha sık takibine imkan sağlaması diğer görüntüleme sistemlerine göre daha çok tercih edilmesine neden olmuştur (Garfinkel vd., 1995).

Bu çalışmanın amacı, akciğer röntgen imgelerindeki farklı dokuların otomatik tespiti ve sınıflandırılmasıdır. Çalışma sonunda , akciğer röntgen imgelerinden iyi huylu nodül, kötü huylu nodül ve nodülsüz olmak üzere üç sınıflı sınıflandırma yapılmıştır. Bu kapsamda Japon Radyolojik Teknoloji Derneği tarafından hazırlanan veritabanı kullanılmıştır (Shiraishi vd., 2000). Nodüller akciğer röntgen imgelerinde küresel formda görünürler. Kabaca tanımlarsak, genel olarak iyi huylu nodüller düzgün bir biçime sahiptir ve kötü huylu

nodüller düzensiz bir biçime sahiptir. Nodülün çapı arttıkça kötü huylu olma olasılığı da artmaktadır. Kullanılan veritabanındaki nodüllerin çapı 1 mm'den 60 mm'ye değişkenlik gösterir. 31 nodül 0-10 mm. aralığında, 52 nodül 11-15 mm. aralığında, 36 nodül 16-20 mm. aralığında, 14 nodül 21-25 mm. aralığında, 17 nodül 26-30 mm. aralığında, 4 nodül ise 31-60 mm. aralığındadır (Shiraishi vd., 2000). Nodüllerin bulunduğu bölgeler sol ve sağ akciğerde 3 anatomik lokasyona ayrılmıştır. Bu bölgeler yukarı, orta ve aşağı akciğerdir (Shiraishi vd., 2000). 42 nodül sağ yukarı akciğerde, 30 nodül orta sağ akciğerde, 23 nodül aşağı sağ akciğerde, 20 nodül yukarı sol akciğerde, 21 nodül orta sol akciğerde, 18 nodül ise aşağı sol akciğerde bulunur (Shiraishi vd., 2000).

Röntgenlerde kaburga kemikleri nodüllü bölge ile örtüşebilir. Bu durum istenmeyen durumdur ve nodüllü bölgenin sağlıklı bir şekilde saptanmasını engelleyebilir. Bu yüzden kaburga kemiklerinin bastırılması için ilkönce imgeye önışleme yapılmıştır. Önışleme, imge işlemede çok önemli bir yer tutar, imgenin karakteristiklerini iyileştirir ve daha iyi sınıflandırma sonuçlarının alınmasını sağlar. Önışleme aşamasından sonra kaburga kemiklerinin yeri saptanmış ve daha sonra bu kemikler bastırılmıştır. Daha sonra akciğerde şüpheli bölgeler bulunmuş ve bu şüpheli bölgelerden öznitelikler çıkarılarak sınıflandırma aşamasına geçilmiştir. Farklı öznitelik çıkarma yöntemleri denenmiş ve hangi özniteliklerin iyi huylu nodülü, kötü huylu nodülü veya nodülsüz bölgeyi iyi ifade ettiği sınıflandırma sonuçlarına bakılarak görülmüştür. Sınıflandırma aşamasında birçok farklı sınıflandırıcı kullanılmıştır ve hangi sınıflandırıcının daha iyi sınıflandırma yaptığı görülmüştür. Sınıflandırma sonuçları, bu çalışmada kullanılan veritabanını kullanan literatürdeki diğer çalışmalarda olduğu gibi doğruluk, hassasiyet ve kesinlik değerleri üzerinden verilmiştir. Çalışma sonunda elde edilen sonuçların adil bir şekilde literatürdeki diğer çalışmalarla karşılaştırılması için sadece bu veritabanını kullanan çalışmalar göz önünde tutulmuştur.



## 2. LİTERATÜR ARAŞTIRMASI

Literatür araştırması 3 ana başlık altında yapılmıştır. Bunlar kaburga kemiklerinin bastırılması, akciğer bölütlenmesi ve akciğer kanseri sınıflandırılması üzerine yapılan çalışmalardır.

### 2.1. Kaburga Kemiklerinin Bastırılması

Suzuki vd. (2006), çoklu çözünürlüklü bir kitlesel eğitim yapay sinir ağı (massive training artificial neural network - MTANN) vasıtasıyla akciğer radyograflarındaki kaburga kemiklerinin kontrastını bastırmak için bir görüntü işleme tekniği geliştirmişlerdir. Bir MTANN, giriş akciğer grafileleri ve bunlara karşılık gelen "öğretme" görüntüleri kullanılarak eğitilebilen son derece doğrusal olmayan bir filtredir. Öğretim görüntüleri olarak çift enerjili bir çıkarma tekniğinin kullanılmasıyla elde edilen "kemik" görüntülerini kullanmışlardır. Çeşitli uzamsal frekanslara sahip olan kaburga kemiklerinin etkili bir şekilde bastırılması için, çoklu çözünürlük ayrıştırma / kompozisyon teknikleri ve üç farklı çözünürlüklü görüntü için üç MTANN'dan oluşan çok çözünürlüklü bir MTANN geliştirmişlerdir. Giriş akciğer radyografileleri ve ilgili çift enerjili kemik görüntüleri ile eğitimden sonra, çok çözünürlüklü MTANN, öğretim kemik görüntülerine benzeyen "kemik görüntüsü benzeri" görüntüler sunmuştur. Kemik görüntüsü benzeri görüntüleri ilgili akciğer grafilelerinden çıkartarak, kaburga kemiklerinin büyük ölçüde bastırıldığı "yumuşak doku görüntüsü benzeri" görüntüler üretilmiştir. Teknikleri eğitimsiz akciğer radyografilelerine uygulandığı zaman, akciğer grafilelerinde kaburga kemikleri önemli ölçüde baskılanırken, nodül ve akciğer damarlarının görünürlüğü korunmuştur.

Soleymanpour vd. (2011), kaburga kemiklerinin konumunu belirlemek için dikey yönelime dayalı uzamsal Gabor filtresi kullanmışlardır. Uzamsal Gabor filtresi, uzamsal frekans bileşenlerine önceden tanımlanmış yönlerde yanıt verir ve frekans bileşenlerini

başka yönlerde bastırır. Ayrıca kaburga yapısının akciğer radyografisinde diğer yapılardan daha yüksek frekans bileşenine sahip olduğu ve yatay olarak uzandığı bilinmektedir. Bu yüzden, uzamsal Gabor filtresi kaburga bastırma için seçilmiştir. İmgeler üzerinde dikey olarak yönlendirilmiş bir uzamsal Gabor filtresi uygulanmıştır ve akciğer bölgesinde yatay çizgiler bulunan yerlerde(kaburga gibi) yüksek tepkiler vermiştir. Orijinal görüntüden Gabor filtresi uygulanmış görüntü çıkartıldığında kaburga kemikleri bastırılmış görüntü elde edilmiştir.

Lee vd. (2012), ikili enerji çıkarma görüntülerine ihtiyaç duymadan sadece bir standart akciğer radyografisi kullanarak kaburgaların bastırılması için yeni bir yaklaşım önermiştir. Bu yaklaşımda, doğrudan incelenen görüntüden bir kaburga şablonu oluşturan yerel örnekleme şeması ile parametrik olmayan bir kaburga-sınır modellemesi geliştirmişlerdir. Bu şema, sistemin, kaburga sınırlarının doğru algılanabileceği şekilde uygun bir kaburga şablonu üretmesini sağlar. Ek olarak, kaburga yoğunluğu dağılımını doğrusal bir model olarak modellemişlerdir ve ilgili parametreler gerçek kodlu genetik algoritma (real coded genetic algorithm - RCGA) kullanılarak tahmin edilmiştir. Kaburga gölgesi, kaburga yoğunluğunu orijinal radyograftan çıkartarak başarıyla bastırılmıştır.

Wang vd. (2017), Poisson modeline dayalı bir kaburga baskılama yöntemi geliştirmişlerdir. Görüntü üzerinde kontur dönüşümü yapmışlar ve görüntü üzerinde doku iyileştirmesi yapmak için farklı ölçekler arasında dönüşüm katsayısı korelasyonunu kullanmışlardır. Görüntüye Hessian matrisinin özdeğerine göre şerit tipi algılama filtresi tasarlamışlardır ve kaburgaların bulunduğu alanı tespit etmişlerdir. İyileştirilmiş doku ve kaburga alanı bilgisini birleştirmiş ve kaburga baskılama Poisson modelini kurmuşlardır. Anizotropi ve kontur dönüşümü çok yönlülük özelliği kullanılmış, ölçek ve katsayılar yön bilgisi birleştirilmiş ve doku ile gürültü arasındaki ayrım derecesi artmıştır. Gürültüyü sınırlandırırken doku iyileştirilmiş, kaburga üzerinde doğru segmentasyon yapması gerekmeyen Poisson modelini kullanarak kaburga baskılmasını gerçekleştirmişlerdir. Kaburga baskılanmasından kaynaklanan kenarların doğal olmayan geçiş problemini önlemiş ve kaburgaların etkili bir şekilde bastırılmışlardır.

## 2.2. Akciğer Bölütlenmesi

Armato vd. (2002), lateral akciğer grafiğinde akciğer bölgelerinin otomatik bölütlenmesi için bir yöntem ve sistem geliştirmişlerdir. Gri seviye eşik analizine dayalı, gelişmiş ve otomatik bir görüntü segmentasyon yöntemi geliştirmişlerdir. Global gri seviye histogramının özelliklerine dayanarak yinelemeli bir global gri seviye eşikleme yöntemi uygulanmıştır. Her iterasyonda oluşturulan ikili bir görüntüdeki bölgelerin özellikleri tanımlanmış ve akciğer alanının dışındaki bölgeleri hariç tutmak için değiştirilmiş bir analize tabi tutmuşlardır. Bu global işlemde elde edilen ilk akciğer bölgesi konturu, yeni bir uyarlanabilir lokal gri seviye eşikleme yöntemini kolaylaştırmak için kullanılmıştır. İlgi bölgeleri (region of interest - ROI) ilk kontur boyunca yerleştirilmiştir. Tek tek ROI'ler içindeki piksellere uygulanacak olan tek gri seviye eşikini belirlemek için bir prosedür uygulamışlardır. Kompozit bir ikili görüntü oluşturulduktan sonra “açık” bölgeleri içine alan son bir kontur oluşturulmuştur ve kontura polinom eğrileri eklenmiştir.

Ginneken vd. (2006), standart arka-ön akciğer grafiğinde akciğer alanlarını, kalbi ve klavikülleri segmente etmişlerdir. Üç denetimli bölümlendirme yöntemi karşılaştırılmıştır. Bunlar aktif şekil modelleri, aktif görünüm modelleri ve çok çözünürlüklü piksel sınıflandırma yöntemidir. Aktif şekil modelleri için bir parametre optimizasyonu sunulmuştur ve bu optimizasyonun performansı önemli ölçüde arttırdığı gösterilmiştir. Standart aktif görünüm model şemasının düşük performans gösterdiği, ancak objelerin dışındaki alanları modele dahil ederek büyük gelişmeler elde edilebileceği gösterilmiştir. Akciğer alanı bölütlenmesi için, tüm yöntemler iyi performans göstermiştir ama piksel sınıflandırması en iyi sonuçları vermiştir.

Candemir vd. (2014), akciğer sınırlarını tespit eden bir akciğer segmentasyon yöntemi sunmuşlardır. Yöntem üç ana aşamadan oluşmaktadır. Birincisi, kısmi bir Radon dönüşümü ve Bhattacharyya şekli benzerliği ölçüsü kullanılarak kullanılan hasta radyografisine en çok benzeyen eğitim görüntülerini tanımlamak için içeriğe dayalı bir görüntü alma yaklaşımıdır. İkincisi, eğitim maskelerinin hasta radyografisine deforme edilebilir çakıştırma için SIFT-akışı kullanılarak hastaya özgü anatomik akciğer şekli

modelinin yaratılmasıdır. Üçüncü ve son olarak ise uyarlanmış bir enerji fonksiyonuyla grafik kesim optimizasyon yöntemini kullanarak akciğer sınırlarının çıkartılmasıdır.

Wu vd. (2015), radyografi akciğer görüntülerinden akciğerleri çıkarmak için özelleştirilmiş bir aktif şekil modeli geliştirmişlerdir. İlk olarak, ortalama aktif şekil modeli, gri ölçekli projeksiyon ve afin dönüşümü, başlangıçtaki akciğer hatlarını elde etmek için kullanılmıştır. İkinci olarak, aktif şekil modelinin köşelerini gerçek akciğer kenarına itmek, köşeleri mide gazı bölgelerinden çıkarmak ve köşelerin daha dengeli bir mesafe dağılımına sahip olmak için, mesafe ve kenar sınırlarına sahip yeni bir objektif işlev önerilmiştir. Son olarak, hızlı optimizasyon sağlamak için çoklu çözünürlüklü gösterim ve optimizasyon kullanılmıştır.

### **2.3. Akciğer kanseri sınıflandırılması üzerine yapılan çalışmalar**

Wei vd. (2002), dört farklı kategoride öznelik vektörü oluşturmuşlardır. Bunlar geometrik, kontrast, birinci ve ikinci dereceden istatistiklerdir. Geometrik özellikler; Yayılma, Dairesellik, Alan, Eşdeğer yarıçap, Aday noktadan pulmoner hiluma olan uzaklık ve düzlüktür. Genel olarak, nodül bölgesi x-ışını görüntüsündeki arka planından daha parlaktır. Bu nedenle, kontrast bilgisini özellik olarak kullanmışlardır. Birinci dereceden istatistikler olarak histogramdan ortalama, standart sapma, kontrast, eğiklik, kurtosis, enerji ve entropiyi hesaplamışlardır. İkinci dereceden istatistiklerin özelliklerini çıkarmak için birliktelik matrisi yöntemini kullanmışlardır. Haralick dönüşümü kullanılarak elde edilmiştir. Birliktelik matrislerinden on dört skaler istatistiksel özellik hesaplanmıştır. Sınıflandırma da yapay sinir ağları kullanmışlardır ve %80 hassasiyette sınıflandırma gerçekleştirmişlerdir. JSRT veritabanı kullanılarak elde edilen en yüksek sınıflandırma başarısına ulaşmışlardır.

Shiraishi vd. (2006), dört ana kategoride öznelik vektörü oluşturmuşlardır. Bunlar geometrik, gri-seviye, arka plan ve kenar gradyent öznelikleridir. Geometrik öznelikler; normalize edilmiş yatay pozisyon, normalize edilmiş dikey pozisyon, orta hattın oluştuğu

yatay uzaklık, efektif yarıçap ve düzensizliktir. Gri seviye öznitelikleri; ortalama piksel değeri, piksel değerinin standart sapması, iç ve dış bölgelerdeki piksel değerlerinin farklılığı ve kontrasttır. Arka plan öznitelikleri; ortalama piksel değeri, en yüksek piksel alanı ortalaması, akciğer alanı oranı, ROI içindeki piksel değerlerinin standart sapması, ortalama kenar açısı, en yüksek kenar açısı ve en yüksek kenar oranıdır. Kenar gradyent öznitelikleri; dairesel gradyent indeksi, ortalama dairesel gradyent, dairesel gradyentin standart sapması, ortalama kenar gradyenti ve kenar gradyentinin standart sapmasıdır. Sınıflandırma da yapay sinir ağları kullanmışlardır ve %92.5 hassasiyette sınıflandırma gerçekleştirmişlerdir.

Hardie vd. (2008), parlaklık, gradyent ve geometrik olmak üzere 3 farklı kategoride öznitelik çıkarmışlardır. Geometrik öznitelikler olarak eksantriklik, akciğer çevresine olan mesafe, dairesellik, alan ve boyuttur. Parlaklık öznitelikleri olarak maksimum değer, minimum değer, ortalama, kontrast, standart sapma, çarpıklık, basıklıktır ve birinci momentten yedinci momente kadar olan momentlerdir. Gradyent öznitelikleri olarak ortalama gradyent büyüklük, gradyent büyüklük standart sapması, ortalama radyal sapma, radyal sapmanın standart sapması, ortalama radyal gradyent, radyal gradyentin standart sapması, ortalama radyal gradyent çevresidir. Fisher Doğrusal Ayırtaç Analizi Sınıflandırıcısı kullanılarak sınıflandırma yapılmıştır ve hassasiyet %72.3 olarak bulunmuştur.

Gonzalez ve Ponomaryov (2016), şekil ve dokusal öznitelikler çıkarmışlardır. Şekil öznitelikleri olarak alan, ekzantriklik, dairesellik ve fraktal boyut özniteliklerini çıkarmışlardır. Doku öznitelikleri olarak GLCM matrisinden elde ettikleri on özniteliği kullanmışlardır. Bu öznitelikler gri seviye değişimlerini gösterir. GLCM matrisinden elde ettikleri öznitelikler ortalama, varyans, entropi, enerji, kurtosis, ters fark momenti, çarpıklık, kontrast, korelasyon ve düzgünlüktür. Sınıflandırmada Destek Vektör Makinelerini kullanmışlar ve %84.93 hassasiyette sınıflandırma gerçekleştirmişlerdir.

### 3. AKCİĞER RÖNTGEN İMGELERİNİN ÖN İŞLENMESİ

Önişleme, imge işlemede çok önemli bir yer tutar. İmgenin karakteristiklerini iyileştirir ve daha iyi sonuçlar alınmasında etkindir. Akciğer nodülleri röntgenlerde küresel biçimde görünür. Nodüller kaburga kemikleriyle örtüşebilir. Bu durumda nodüllü bölgeyi etkili bir biçimde saptayabilmek için kaburga kemiklerinin bastırılması gereklidir. Bu çalışmada kaburga kemiklerini bastırmak ve akciğer bölgesini bölütlemek için önişleme yapılır.

#### 3.1. Akciğer Bölütlenmesi için Yapılan Önişleme

##### 3.1.1. Eşikleme

Akciğer radyagrofisi, akciğer dokusundan başka dokularda içerir. Yalnızca akciğer bölgesini bölütlemek için eşikleme yöntemi kullanılmıştır. Akciğer dokusunun dışında kalan dokuların piksel değeri, akciğer dokusuna göre belirgin olarak farklılık gösterir. Bu farklılık kullanılarak akciğer bölgesi bölütlenir. Bir piksel değeri eşik değeri olarak belirlenir. Bu eşik değerin üzerindeki piksel değerleri aynı kalırken, eşik değerin altındaki piksel değerleri 0 yapılarak akciğer olmayan dokular kaybedilir.

#### 3.2. Kaburga Kemiklerinin Bastırılması için Yapılan Önişleme

##### 3.2.1. Medyan Filtre

Genellikle imgeden gürültüyü gidermek için kullanılan doğrusal olmayan bir dijital filtreleme tekniğidir (Gallagher ve Wise, 1981). Bu tür gürültü azaltma, daha sonraki

işlemlerin sonuçlarını iyileştirmek için yapılan bir ön işleme adıdır. Medyan filtreleme, dijital imge işlemede çok yaygın olarak kullanılır çünkü belirli koşullar altında, gürültü giderirken kenarları korur. Belli bir boyutta bir pencere belirlenir. Pencere her piksel üzerinde kaydırılır ve her kaydırıldığı piksel değeri penceredeki piksel değerlerinin medyanı ile değiştirilir.

### 3.2.2. İkileştirme

Görüntü ikileştirme, imge işlemede temel bir araştırma teması ve imge tanıma ve kenar / sınır algılamasında önemli bir ön işleme yöntemidir. Daha sonra yapılacak şablon eşleme yöntemi için imge ikileştirilmelidir. Bu amaçla sadece kemik yapılarını ortaya çıkarmak bir piksel değeri eşik değeri olarak belirlenir ve eşik değerin üzerindeki piksel değerleri 1 olurken, eşik değerin altındaki piksel değerleri 0 yapılarak, kemik dışındaki dokular büyük ölçüde kaybedilir ve ikileştirme yapılan imge elde edilir.

### 3.2.3. Açma İşlemi

Matematiksel olarak açma işlemi Eşitlik 3.1'deki gibi ifade edilir.  $\ominus$  genişletmeyi ve  $\oplus$  aşındırmayı ifade eder.

$$A \circ B = (A \ominus B) \oplus B \quad (3.1)$$

Açma, imge işlemede morfolojik gürültünün giderilmesinde önemli bir yer tutar. İmgenin ön planından küçük delikleri kaldırır. (Anuar ve Jambek, 2016).

### 3.2.4. Kapama İşlemi

Matematiksel olarak kapama işlemi Eşitlik 3.2'deki gibi ifade edilir.  $\ominus$  genişletmeyi ve  $\oplus$  aşındırmayı ifade eder.

$$A \cdot B = (A \oplus B) \ominus B \quad (3.2)$$

Kapama, imge işlemede morfolojik gürültünün giderilmesinde önemli bir yer tutar. İmgenin ön planındaki küçük delikleri doldurur (Anuar ve Jambek, 2016).

### 3.2.5. Şablon Eşleme

Sol kaburga kemiklerinden bir kemik ve sağ kaburga kemiklerinden bir kemik şablon olarak kullanılır. Akciğerde bulunan en küçük kemik yapısının alanı eşik değeri olarak belirlenir ve sol taraftan seçilen şablon sol kaburga üzerinde gezdirilir. Şablon ve kemik örtüşmesi belirlenen eşik değerinin üstünde ise kemik yeri belirlenir ve bu yer daha sonra kemik bastırma için kullanılır. Aynı algoritma sağ kaburga için de kullanılır ve sağ kaburgadaki kemiklerin yerleri de belirlenir.

### 3.2.6. Gabor Filtresi

Gabor filtresi, kemikleri bastırmak için kullanılır. Uzamsal bir Gabor filtresi, bir sinüsoidün bir Gauss ile modüle edilmesiyle elde edilir (Gabor, 1946). Sinüzoidi Gauss ile birleştirerek, bir Gabor filtresi elde edilir.  $\theta$  uzamsal frekans,  $\sigma$  standart sapma ve  $\phi$  yön olmak üzere,  $g(x,y,\theta,\phi)$  orijin  $(x,y)$ 'de merkezlenmiş uzamsal Gabor filtresini tanımlayan bir fonksiyon olsun. Uzamsal Gabor filtresi fonksiyonu Eşitlik 3.3'de verilmiştir.



$$g(x, y, \theta, \phi) = \exp(-((x^2 + y^2)/(\sigma^2))) * \exp(2\pi\theta i(x\cos\phi + y\sin\phi)) \quad (3.3)$$

### 3.2.7. Gauss Filtresi

Gauss Filtresi, dürtü cevabı Gauss fonksiyonu olan bir filtredir (Haddad ve Akansu, 1991). İmgede, gürültüyü ve ayrıntıları giderme amaçlı kullanılır. Gauss filtresi Eşitlik 3.4'de verilmiştir.

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2 + y^2}{2\sigma^2}\right)} \quad (3.4)$$

Eşitlikte x, yatay eksenin orijine olan uzaklığını, y, dikey eksenin orijine olan uzaklığı ve  $\sigma$ , Gauss dağılımının standart sapmasını verir.

#### 4. AKCİĞER RÖNTGEN İMGELERİNDE ŞÜPHELİ BÖLGE TESPİTİ

Nodüllerin sınıflandırılması aşamasında daha doğru sonuçlar alabilmek için akciğer üzerindeki şüpheli bölgelerin bulunur ve bu şüpheli bölgelerden öznelik çıkartımı yapılır. Bu çalışmada Kenarsız Chan-Vese Aktif Kontur yöntemi kullanılmıştır. Kenarsız Chan-Vese Aktif Kontur yöntemi Mumford ve Shah tarafından formüle edilen segmentasyon problemine dayanır (Chan ve Vese, 2001).

##### 4.1. Mumford ve Shah Minimizasyon Problemi

$u_0$  imgesi  $C$  konturu ile ikiye ayrılır ve Eşitlik 4.1’de verilen  $F^{MS}(u,C)$  minimizasyon problemi tanımlanır (Mumford ve Shah, 1989).

$$\inf\{F^{MS}(u, C) = \int_{\Omega}(u - u_0)^2 dx dy + \mu \int_{\Omega} |\nabla u|^2 dx dy + v|C|\} \quad (4.1)$$

Eşitlik 4.1’de her bağlı bileşen  $\Omega_i$  içinde  $u$ =sabit  $c_i$  alındığında, problem Eşitlik 4.2’de verilen “minimal bölünmüş problem” olarak adlandırılır (Mumford ve Shah, 1989).  $c_i$ ,  $u_0$  imgesindeki piksellerin ortalamasını ifade eder ve  $v$ ,  $C$ ’ye bağlı bir sabittir.

$$E^{MS}(u, C) = \sum_i \int_{\Omega} (u - c_i)^2 dx dy + v|C| \quad (4.2)$$

$$c_i = \text{ortalama}(u_0), u_0 \subset c_i$$

#### 4.2. Kenarsız Chan-Vese Aktif Kontur Modeli

Minimal bölünmüş problem yardımı ile Eşitlik 4.3'de  $c_1$ ,  $c_2$  ve  $C$ 'ye göre enerjinin minimize edildiği fonksiyon verilir (Chan ve Vese, 2001).  $c_1$ ,  $C$  konturunun içindeki her şeyin ortalama değeri ve  $c_2$ ,  $C$  çizgisinin dışındaki her şeyin ortalama değeri olarak yorumlanabilir.  $\Omega_1$ , kontur içindeki alanı ve  $\Omega_2$ , kontur dışındaki alanı ifade eder.

$$F(c_1, c_2, C) = \int_{\Omega_1=w} (u_o(x-y) - c_1)^2 dx dy + \int_{\Omega_2=\Omega-w} (u_o(x-y) - c_2)^2 dx dy + \nu |C| \quad (4.3)$$

$$C = \partial w, w \subset \Omega$$

Chan ve Vese (2001), makalede Eşitlik 4.4'deki ilk iki terimi iki kuvvet olarak yorumlamışlardır. İlk terim konturu küçültmeye zorlar. İkinci terim ise konturu genişletmeye zorlar. Bu iki kuvvet, kontur ilgilenilen nesnenin sınırına ulaştığında dengelenir. Algoritmanın çalışma mantığı Şekil 4.1'de dört durum üzerinden gösterilmiştir.

$$F(c_1, c_2, C) = \int_{i\check{c}(C)} |u_0 - c_1|^2 dx + \int_{dış(C)} |u_0 - c_2|^2 dx = F_1(C) + F_2(C) \quad (4.4)$$

$$F_1(C) > 0, F_2(C) \approx 0 \quad F_1(C) \approx 0, F_2(C) > 0$$



(a)



(b)

$$F_1(C) > 0, F_2(C) > 0 \quad F_1(C) \approx 0, F_2(C) \approx 0$$



(c)



(d)

**Şekil 4. 1.** Chan-Vese algoritmasının çalışma prensibi (Chan ve Vese, 2001)

Şekil 4.1'deki imgelerde siyah olan kısımlar -1 değeri ile gösterilir ve gri olan kısımlar 1 ile gösterilir. Şekil 4.1'de (a) durumu incelendiğinde ilk kontur tüm nesneyi (-1) ve bazı gri bölgeleri (+1) kapsar. Bu nedenle,  $c_1$  yaklaşık 0'dır ve  $c_2$  1'dir. Kontur dışında kalan imgeden  $c_2$  çıkarıldığında, 0 elde edilir. Dolayısıyla  $F_2$  terimi sıfırdır.  $c_1$  sıfıra yaklaştığı için, kontur içinde kalan imgeden  $c_1$  çıkartıldığında ve formülün gösterdiği gibi karelerin toplamı bulunduğunda, büyük bir pozitif sayıya ulaşılır. Bu nedenle  $F_1 > 0$  çıkar.  $F_1 > 0$  ve  $F_2 = 0$  için kontur bir sonraki adımda kendini daraltacaktır. Şekil 4.1'deki diğer durumlarda bu şekilde yorumlanabilir.

Kenarsız Chan-Vese Aktif Kontur algoritmasında düzey kümesi fonksiyonu  $\phi(x,y)$  kontur değerlerini gösterir. C kontur eğrisinin matematiksel ifadesi Eşitlik 4.5'te verilmiştir.

$$C = \{(x, y): \phi(x, y) = 0\}, \forall (x, y) \in u_0 \quad (4.5)$$

Konturun düzey kümesi fonksiyonuna göre zamanla değişimi Eşitlik 4.6'da verilmiştir.

$$\frac{\partial C}{\partial t} = \frac{\partial \phi(x, y)}{\partial t} \quad (4.6)$$

## 5. AKCİĞER RÖNTGEN İMGELERİNDE TESPİT EDİLEN ŞÜPHELİ DOKULARDA ÖZNETELİK ÇIKARTIMI

İstatistiksel öznitelikler, Yönlendirilmiş Gradyanların Histogramı (Histogram of Oriented Gradients - HOG) öznitelikleri, Yerel İkili Örüntü (Local Binary Pattern – LBP) öznitelikleri, Geometrik öznitelikler, Gri Seviye Eş-Oluşum Matrisi (Gray Level Co-Occurrence Matrix – GLCM) öznitelikleri ve Ölçekten Bağımsız Yoğun Öznitelik Dönüşümü (Dense Scale Invariant Feature Transform – DSIFT) olmak üzere altı farklı öznitelik çıkarımı yapılmıştır.

### 5.1. İstatistiksel Öznitelikler

Birçok makale örüntü sınıflandırmada istatistiksel özniteliklerin başarısını bahsetmiştir. Bu çalışmada, akciğerin şüpheli bölgelerinden enerji, ortalama, varyans, maksimum, minimum, standart sapma, çarpıklık, basıklık, alan tanımlayıcısı, ortalama enerji, enerji varyansı ve entropi olmak üzere 12 farklı istatistiksel öznitelik çıkartılmış ve Çizelge 5.1’de verilmiştir (Esener vd., 2017).

**Çizelge 5. 1.** İstatistiksel öznitelikler ve matematiksel ifadeleri

İstatistiksel Öznitelikler	Matematiksel İfadeleri
Enerji	$\sum_{i=1}^N x_i^2$
Ortalama	$\mu = \frac{1}{N} \sum_{j=1}^N x_j$
Varyans	$\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2$
Maksimum	$\text{maks}\{x_i, 1 \leq i \leq N\}$

Minimum	$\min\{x_i, 1 \leq i \leq N\}$
Standart Sapma	$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2}$
Çarpıklık	$\frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^3}{\sigma^3}$
Basıklık	$\frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^4}{\sigma^4}$
Alan Tanımlayıcısı (Ahonen vd., 2006)	$\sigma / \mu$
Ortalama Enerji	$\mu_{Energy} = \frac{1}{N} \sum_{i=1}^N x_i^2$
Enerji Varyansı	$\frac{1}{N} \sum_{i=1}^N (X_i^2 - \mu_{Energy})^2$
Entropi	$\sum_{i=1}^N p(X_i) \log_2 p(X_i)$

## 5.2. Yönlendirilmiş Gradyanların Histogramı Öznitelikleri

HOG öznitelikleri, vücutta şüpheli bölgelerden yapılan örüntü tanıma problemlerinde başarılı sonuçlar vermiştir. (Song vd., 2012). Bu çalışmada, akciğerdeki şüpheli bölgelerden HOG öznitelikleri çıkartılmıştır.

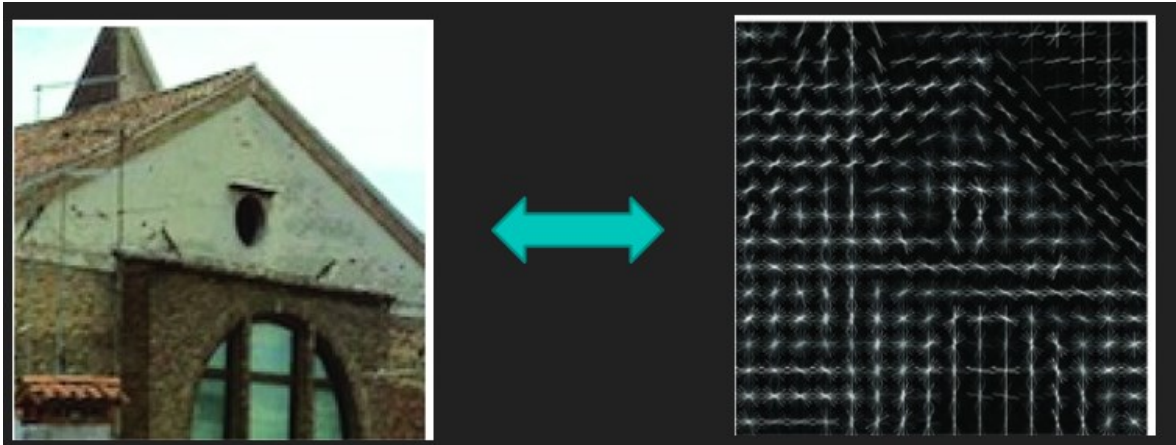
Görüntü gradyanı vektörü, her bir piksel için, hem x eksenini hem de y eksenindeki piksel parlaklık değişikliklerini içeren bir metrik olarak tanımlanır. Tanım, tüm değişkenlerin kısmi türevlerinin bir vektörü olan sürekli bir çok değişkenli fonksiyonun gradyanı ile verilir.  $F(x, y)$  konumunun  $(x, y)$  konumundaki pikselin parlaklık değerini kaydettiğini varsayalım, pikselin gradyan vektörü  $(x, y)$  Eşitlik 5.1'deki gibi tanımlanır:

$$\nabla f(x, y) = \begin{pmatrix} g_x \\ g_y \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} f(x+1, y) - f(x-1, y) \\ f(x, y+1) - f(x, y-1) \end{pmatrix} \quad (5.1)$$

$\partial f/\partial x$  ifadesi, hedef pikselin solu ve sağına komşu pikseller arasındaki parlaklık farkı olarak hesaplanan +x yönünde kısmi türevidir. Benzer şekilde,  $\partial f/\partial y$  ifadesi, hedef pikselin üzeri ve altında komşu pikseller arasındaki parlaklık farkı olarak hesaplanan +y yönünde bir kısmi türevidir. Büyüklük vektörün L2 normudur ve Eşitlik 5.2'deki gibi hesaplanır. Oryantasyon, iki yöndeki kısmi türevlerin birbirine oranının arktanjanıdır ve Eşitlik 5.3'deki gibi hesaplanır.

$$|\nabla f(x, y)| = \sqrt{g_x^2 + g_y^2} \quad (5.2)$$

$$\theta = \arctan\left(\frac{g_y}{g_x}\right) \quad (5.3)$$



(a)

(b)

**Şekil 5. 1.** HOG Yönteminin Uygulanması a) Örnek İmge b) HOG Yönteminin Örnek İmge Üzerinde Uygulanması Sonucu Elde Edilen İmge



### 5.3. Yerel İkili Örüntü Öznitelikleri

İmgelerdeki şüpheli bölgelerden LBP öznitelikleri çıkarılmıştır. LBP, her pikselin komşuluğunu eşikleyerek bir görüntünün piksellerini etiketleyen ve sonucu ikili bir sayı olarak gören basit ama çok verimli bir doku operatörüdür (Wang ve He, 1990). Ayırmacı gücü ve hesaplamalı sadeliği nedeniyle, LBP doku operatörü çeşitli uygulamalarda popüler bir yaklaşım haline gelmiştir. Gerçek dünyadaki uygulamalarda LBP operatörünün en önemli özelliği, örneğin aydınlatma varyasyonları nedeniyle oluşan monotonik gri ölçekli değişikliklere karşı sağlamlığıdır.

Özgün LBP işleci, her piksel etrafında yerel yapıyı kodlayan LBP kodları olarak adlandırılan ondalık sayılarla bir görüntünün piksellerinin değerini değiştirir. Her bir merkezi piksel sekiz komşusu ile karşılaştırılır. Merkezi pikselin değerinden daha küçük olan komşular bit 0'a sahip olur ve merkezi pikselin değerine eşit veya daha büyük olan diğer komşular bit 1'e sahip olur. Verilen her bir merkezi piksel için, bu ikilik bitlerin tümünü, sol üstteki komşusunun birinden başlayarak, saat yönünde birleştirilerek elde edilen bir ikili sayı üretilir. Oluşturulan ikili sayının sonuçtaki ondalık değeri, merkezi piksel değerinin yerini alır.

Bir imgenin LBP gösterimi Eşitlik 5.4'deki gibi hesaplanır. Eşitlikte R dairenin yarıçapını, P komşuluktaki piksel sayısını ifade eder, u birim basamak fonksiyonunu,  $g_i$  i. komşuluk pikselinin yeğinlik (intensity) değerini ve  $g_c$  komşuluktaki merkez pikselin yeğinlik değerini ifade eder.

$$LBP(P, R) = \sum_{i=0}^{P-1} u(g_i - g_c) \cdot 2^i \quad (5.4)$$

İkili örüntü en fazla iki 0-1 veya 1-0 geçişleri içeriyorsa, yerel ikili örüntü biçimli olarak adlandırılır. Dönme açısından bağımsız örüntüler her bir bitin dairesel olarak

minimum değere döndürülmesiyle elde edilir. Dönme açısından bağımsız örüntüler Eşitlik 5.5'deki gibi hesaplanır. Eşitlikteki U ifadesi biçimlilik ölçütünü ifade eder.

$$LBP^{riu2}(P, R) = \begin{cases} \sum_{i=0}^{P-1} u(g_i - g_c) & U(LBP(P, R)) \leq 2 \\ P + 1 & U(LBP(P, R)) > 2 \end{cases} \quad (5.5)$$

#### 5.4. Geometrik Öznitelikler

İmgelerdeki şüpheli bölgelerden geometrik öznitelikler çıkarılmıştır. Şüpheli bölgede birbirine bağlı en büyük doku bulunur ve buradan oryantasyon, major eksen uzunluğu, minor eksen uzunluğu, eksantriklik, katılık, doluluk oranı, çap, konveks alan, alan, yuvarlaklık, ovallık ve çevre öznitelikleri çıkarılır.

Alan, şüpheli doku içindeki toplam piksel sayısını ifade eder. Konveks alan, şüpheli dokuyu içine alacak en küçük konveksin çizilmesi ve bu konveks içindeki toplam piksel sayısının bulunması ile bulunur. Katılık ise şüpheli doku içindeki toplam piksel sayısının(alan) konveks alana bölümü ile bulunur. Doluluk oranı, şüpheli dokuyu içine alacak en küçük dikdörtgenin çizilmesi ve şüpheli doku içindeki toplam piksel sayısının dikdörtgen içindeki toplam piksel sayısına bölümü ile bulunur. Major ve minor eksen uzunluğu bulunurken şüpheli dokuyu içine alacak şekilde bir elips çizilir. Major eksen elipsin x ekseninin piksel cinsinden uzunluğunu ifade eder. Minor eksen ise elipsin y ekseninin piksel cinsinden uzunluğunu ifade eder. Oryantasyon, x ekseni ile elipsin x ekseni arasında kalan açıdır. Eksantriklik, elipsin odakları ile majör eksen uzunluğu arasındaki mesafenin oranıdır. Çevre, şüpheli dokunun çevresindeki toplam piksel sayısını verir.

Çap, şüpheli doku ile aynı alana sahip olan bir dairenin çapı olarak Eşitlik 5.6'daki gibi hesaplanır.

$$D = \sqrt{\frac{4 * Alan}{\pi}} \quad (5.6)$$

Ovallık, şüpheli dokunun oval şekline ne kadar yakın olduğunu gösterir ve Eşitlik 5.7'deki gibi hesaplanır (Esener, 2017).

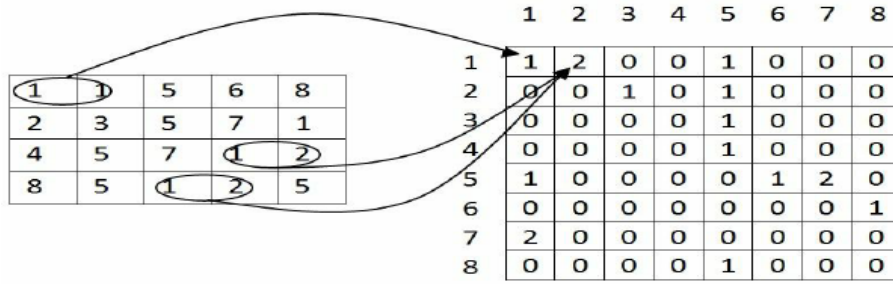
$$O = 2 * \frac{\text{major eksen uzunluğu} - \text{minor eksen uzunluğu}}{\text{major eksen uzunluğu} + \text{minor eksen uzunluğu}} \quad (5.7)$$

Yuvarlaklık, Eşitlik 5.8'deki gibi hesaplanır. r, major eksen uzunluğunun minor eksen uzunluğuna bölümüdür (Esener, 2017).

$$Y = \frac{Alan}{\pi r^2} \quad (5.8)$$

### 5.5. Gri Seviye Eş-Oluşum Matrisi Öznitelikleri

Gri Seviye Eş-oluşum Matrisi (GLCM) doku analizi için kullanılır (Haralick vd., 1973). Kare matris olan GLCM, doku görüntüsündeki gri seviyelerin uzamsal dağılımı hakkında belirli özellikleri ortaya çıkarır. Parlaklık değeri i ile referans piksel olarak bilinen bir piksel değerinin, j parlaklık değeri ile komşu piksel olarak bilinen bir piksel değeri ile belirli bir ilişkide ne sıklıkta ortaya çıktığını gösterir. GLCM matrisinin her elemanı (i,j), birbirine  $\Theta$  yönünde d uzaklıkta olan i ve j piksel değerlerinin tekrar sayısıdır. GLCM matrisin nasıl oluşturulduğuna dair örnek Şekil 5.2'de verilmiştir.



**Şekil 5. 2.** GLCM matrisinin oluşturulması

Şekil 5.2’de  $d=1$  ve  $\Theta=0^\circ$  için  $i$  ve  $j$  piksel değerlerinin ne sıklıkta tekrar ettiğinin örneği verilmiştir.  $i$  ve  $j$  piksel değerleri 1 seçildiğinde, matrisde başka  $i$  ve  $j$  pikselinin 1’e eşit olduğu eleman olmadığı için GLCM matrisinde 1. satır 1. sütuna 1 yazılır. Aynı şekilde matrisde, piksel değeri  $i=1$  ve  $j=2$  olan 2 tane piksel çifti olduğu için GLCM matrisinde 1. satır 2. sütuna 2 yazılır.

Çalışma kapsamında, Haralick vd. (1973), Soh ve Tsatsoulis (1999) ve Clausi (2002) tarafından önerilen öznelikler GLCM matrislerinden çıkartılmıştır. Bu öznelikler ve matematiksel ifadeleri Çizelge 5.2’de verilmiştir. GLCM matrisinden toplamda 22 farklı öznelik çıkartılmıştır (Esener vd., 2016). Çizelgede verilen öznelikler, çizelgenin başında verilen eşitlikler kullanılarak hesaplanmıştır.

**Çizelge 5. 2.** GLCM öznelikleri ve matematiksel ifadeleri

GLCM Öznelikleri	Matematiksel İfadeleri
$P(i, j): GLCM$	$= \begin{bmatrix} P(1,1) & \dots & P(1, N_g) \\ \vdots & \ddots & \vdots \\ P(N_g, 1) & \dots & P(N_g, N_g) \end{bmatrix}$
	$P_x(i) = \sum_{j=1}^{N_g} P(i, j) \quad , \quad P_y(i) = \sum_{i=1}^{N_g} P(i, j)$

$P_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P_{i+j=k}(i, j) , k = 2, 3, \dots, 2N_g$	
$P_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P_{ i-j =k}(i, j) , k = 0, 1, \dots, N_g - 1$	
$\mu_x = \sum_i \sum_j i \cdot P(i, j) \quad , \quad \mu_y = \sum_i \sum_j j \cdot P(i, j)$	
$\sigma_x = \sum_i \sum_j (i - \mu_x)^2 \cdot P(i, j) \quad , \quad \sigma_y = \sum_i \sum_j (j - \mu_y)^2 \cdot P(i, j)$	
Otokorelasyon (Soh ve Tsatsoulis, 1999)	$\sum_i \sum_j (i \cdot j) \cdot P(i, j)$
Kontrast (Haralick vd., 1973; Soh ve Tsatsoulis, 1999)	$\sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i, j) \right\}_{ i-j =n}$
Korelasyon (MATLAB R2016a)	$\frac{\sum_i \sum_j (i - \mu_x) \cdot (j - \mu_y) \cdot P(i, j)}{\sigma_x \cdot \sigma_y}$
Korelasyon (Haralick vd., 1973; Soh ve Tsatsoulis, 1999)	$\frac{\sum_i \sum_j (i \cdot j) \cdot P(i, j) - \mu_x \cdot \mu_y}{\sigma_x \cdot \sigma_y}$
Asimetri (Soh ve Tsatsoulis, 1999)	$\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{i + j - \mu_x - \mu_y\}^4 \cdot P(i, j)$
Çarpıklık (Soh ve Tsatsoulis, 1999)	$\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{i + j - \mu_x - \mu_y\}^3 \cdot P(i, j)$
Benzersizlik (Soh ve Tsatsoulis, 1999)	$\sum_i \sum_j  i - j  \cdot P(i, j)$

Enerji (Haralick vd., 1973 ; Soh ve Tsatsoulis, 1999)	$\sum_i \sum_j \{P(i, j)\}^2$
Entropi (Soh ve Tsatsoulis, 1999)	$- \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P(i, j) \cdot \log\{P(i, j)\}$
Homojenlik (MATLAB R2016a)	$\sum_i \sum_j \frac{1}{1+ i-j } \cdot P(i, j)$
Homojenlik (Soh ve Tsatsoulis, 1999)	$\sum_i \sum_j \frac{1}{1+(i-j)^2} \cdot P(i, j)$
Maksimum Olasılık (Soh ve Tsatsoulis, 1999)	$\max_{i, j} P(i, j)$
Varyans (Haralick vd., 1973)	$\sum_i \sum_j (i - \mu)^2 \cdot P(i, j)$
Toplam Ortalaması (Haralick vd., 1973)	$\sum_{i=2}^{2N_g} i \cdot P_{x+y}(i)$
Toplam Varyansı (Haralick vd., 1973)	$\sum_{i=2}^{2N_g} (i - f_{14})^2 \cdot P_{x+y}(i)$
Toplam Entropisi (Haralick vd., 1973)	$- \sum_{i=2}^{2N_g} P_{x+y}(i) \cdot \log\{P_{x+y}(i)\}$
Fark Varyansı (Haralick vd., 1973)	variance of $P_{x-y}$
Fark Entropisi (Haralick vd., 1973)	$- \sum_{i=2}^{2N_g} P_{x-y}(i) \cdot \log\{P_{x-y}(i)\}$

<p>Korelasyon Bilgi Ölçeği 1 (Haralick vd., 1973)</p>	$\frac{HXY - HXY1}{\max\{HX; HY\}}$ $HXY = -\sum_i \sum_j P(i, j) \cdot \log(P(i, j))$ $HXY1 = -\sum_i \sum_j P(i, j) \cdot \log(p_x(i) \cdot p_y(i))$ <p><math>HX</math> ve <math>HY</math> <math>p_x</math> ve <math>p_y</math>'nin entropi degerleridir.</p>
<p>Korelasyon Bilgi Ölçeği 2 (Haralick vd., 1973)</p>	$(1 - \exp[-2 \cdot (HXY2 - HXY)])^{1/2}$ $HXY2 = -\sum_i \sum_j P_x(i) \cdot P_y(i) \cdot \log\{P_x(i) \cdot P_y(i)\}$
<p>Inverse Difference Normalized (Clausi, 2002)</p>	<p>Normalize Edilmiş Homojenlik (MATLAB R2016a)</p>
<p>Inverse Difference Moment Normalized (Clausi, 2002)</p>	<p>Normalize Edilmiş Homojenlik (Soh ve Tsatsoulis, 1999)</p>

## 5.6. Ölçekten Bağımsız Öznitelik Dönüşümü

Ölçekten bağımsız öznitelik dönüşümü, görüntü tabanlı eşleme ve tanıma için bir görüntü tanımlayıcısıdır ve Ölçekten Bağımsız Öznitelik Dönüşümü (Scale Invariant Feature Transform – SIFT) yapısı dört bölümde incelenir (Lowe, 1999).

### 5.6.1. Ölçek Uzayı Uç Noktaları Saptama

Potansiyel bölgeleri belirlemek için ölçek uzayında anahtar olabilecek bölgeler araştırılır. Anahtar bölgeleri algılamanın ilk aşaması, aynı nesnenin farklı görünüşleri altında tekrarlanabilir şekilde belirlenen konumları ve ölçekleri tanımlamaktır. Bunun için Difference of Gaussian (Gauss Dağılımlı Farklar , DOG) yöntemi kullanılır. Çeşitli makul

varsayımlar altında bir ölçek uzay fonksiyonu elde etmek için Gauss fonksiyonu kullanılır. Ölçek uzayı Eşitlik 5.9'daki gibi tanımlanır.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (5.9)$$

$I(x, y)$  imgeyi ve  $G(x, y, \sigma)$  Gauss fonksiyonunu ifade eder. Anahtar noktaları daha verimli bir şekilde saptamak için Eşitlik 5.10 kullanılır. Eşitlikte  $k$  sabit çarpımsal faktörü ifade eder.

$$D(x, y, \sigma) = ((G(x, y, k\sigma) - G(x, y, \sigma))^2 * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (5.10)$$

### 5.6.2. Düşük Kontrastlı Bölgelerin ve Kenarların Elenmesi

DoG, kenarlar için daha yüksek tepkiye sahiptir, bu nedenle kenarların da çıkarılması gerekir. Ana eğriliği hesaplamak için 2x2 Hessian matrisi kullanılır. Kenarlar için bir özdeğerin diğerinden daha büyük olduğu bilindiği için kenarlar için bir eşik değeri belirlenir ve bu eşik değerin üstündeki anahtar bölgeleri atılır. Böylece düşük kontrastlı bölgeler ve kenarlar elenmiş olur.

### 5.6.3. Dönme Açısı Belirlenmesi

Ölçekten bağımsız imge rotasyonunu sağlamak için her anahtar bölgesine dönme açısı uygulanır. Ölçeğe bağlı olarak anahtar bölge konumunun etrafında bir komşuluk alınır ve o bölgede gradyan büyüklüğü ve yönü hesaplanır. Gradyan büyüklüğü  $m(x, y)$  Eşitlik 5.11'de ve dönme açısı  $\Theta(x, y)$  Eşitlik 5.12'de hesaplanır.

$$m(x, y) = (L(x + 1, y) - L(x, y - 1))^2 + (L(x, y + 1) - L(x, y - 1))^2 \quad (5.11)$$

$$\Theta(x, y) = \tan^{-1} \left( \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \quad (5.12)$$

360 dereceye denk gelen, 36 hücre içeren bir dönme açısı histogramı oluşturulur ve gradyan büyüklüğü ve anahtar bölge ölçeğinin 1,5 katına eşit gauss ağırlıklı dairesel pencere



ile ağırlıklandırılır. Histogramdaki en yüksek tepe noktası alınır ve %80'in üzerindeki herhangi bir tepe noktasının da dönme açısını hesapladığı kabul edilir. Aynı konum ve ölçekle fakat farklı yönlerle anahtar bölgeleri oluşturur ve dönme açısından bağımsız öznitelikler çıkarılır.

#### **5.6.4. Hücrelerin Betimlenmesi**

Gradyan büyüklükleri ve dönme açıları her bir anahtar bölgeden hesaplanır. Sonra her bir hücreden dönme açıları histogramı kullanılarak öznitelik vektörü oluşturulur.

#### **5.7. Ölçekten Bağımsız Yoğun Öznitelik Dönüşümü**

Ölçekten ve dönme açısından bağımsız öznitelikler çıkarılmasını sağlayan SIFT çeşitidir. (Wang et al., 2010). SIFT algortiması ile aynı mantıkta çalışır. Tek farkı anahtar noktaları saptanırken imgenin her yerinden eşit aralıklarla ve çakışmalı olarak verilen yarıçapa göre örnekler alınmasıdır.

## 6. AKCİĞER RÖNTGEN İMGELERİNİN SINIFLANDIRILMASI

Şüpheli dokulardan öznelilikler çıkarıldıktan sonra dokuların sınıflandırılması için  $k$ -En Yakın Komşu ( $k$ -Nearest Neighbor –  $k$ -NN), Karar Ağacı, Rastgele Orman, Lojistik Doğrusal Sınıflandırıcısı (Logistic Linear Classifier – LLC), Destek Vektör Makineleri (Support Vector Machines – SVM), Fisher Doğrusal Ayırtaç Analizi (Fisher's Linear Discriminant Analysis – FLDA) ve Naive Bayes sınıflandırıcıları kullanılmıştır.

### 6.1. $k$ -En Yakın Komşu Sınıflandırıcısı

$k$  en yakın komşu sınıflandırıcısı, mevcut tüm verileri depolayan ve yeni verileri benzerlik ölçütüne göre uzaklık fonksiyonları kullanarak sınıflandıran basit bir algoritmadır (Kim et al., 2012). Bir veri, komşularının çoğunluk oyuyla sınıflandırılır. Durum, uzaklık fonksiyonu ile ölçülen en yakın komşuları arasında en yaygın olan sınıfa atanır.  $k = 1$  ise, durum en yakın komşusunun sınıfına atanır. En yaygın kullanılan uzaklık fonksiyonlarından biri öklit uzaklığıdır.  $X$  ve  $Y$  vektör arasındaki Öklit uzaklığı Eşitlik 6.1'deki gibi hesaplanır.

$$\text{Öklit\_Uzaklığı}(X,Y) = \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (6.1)$$

### 6.2. Karar Ağacı Sınıflandırıcısı

Karar ağacı, bir ağaç yapısı şeklinde sınıflandırma veya regresyon modelleri oluşturur. Bir veri kümesini daha küçük alt kümelere ayırırken, aynı zamanda ilişkili bir karar ağacı kademeli olarak geliştirilir (Safavian ve Landgrebe, 1991). Sonuç, karar düğümleri ve yaprak düğümleri olan bir ağaçtır. Bir karar düğümünün iki veya daha fazla dalı vardır. Yaprak düğümü, bir sınıflandırma veya kararı temsil eder. Ağacın en üstündeki karar düğümü, kök düğümü diye adlandırılır. Bu durum örnek ile açıklanırsa, hava karar düğümünü, havanın nasıl olacağı (yağmurlu, güneşli) karar düğümünün dallarını, yaprak düğümü de o hava şartlarında aktivite yapılıp yapılmayacağını belirtir. Bütün örnekler aynı

sınıfa dahi oluncaya kadar ağaç dallanır, sonunda yaprak düğümü oluşur ve sınıflandırma yapılmış olur. Karar ağaçları hem kategorik hem de sayısal verileri ele alabilir.

### 6.3. Rastgele Orman Sınıflandırıcısı

Rastgele orman birden fazla karar ağacı oluşturur ve daha doğru ve istikrarlı bir tahmin elde etmek için onları birleştirir. Rastgele Orman takriben bir karar ağacıyla aynı parametrelere sahiptir. Rastgele Orman, ağaçları büyütürken modele ilave rastgelelik katmaktadır. Bir düğümü bölerken en önemli özelliği aramak yerine, rastgele bir özellik alt kümesi içinde en iyi özelliği arar. Bu, genellikle daha iyi bir modelle sonuçlanan geniş bir çeşitlilikle sonuçlanır. Bu nedenle, Rastgele Orman'da, bir düğümü bölmek için kullanılan algoritma ile özelliklerin sadece rastgele bir alt kümesi dikkate alınır. Ek olarak, mümkün olan en iyi eşikleri normal bir karar ağacı gibi aramak yerine, her özellik için rastgele eşikleri kullanarak ağaçları daha rastgele yapabilir.

### 6.4. Fisher Doğrusal Ayırtaç Analizi Sınıflandırıcısı

FLDA, 1936'da R. A. Fisher tarafından geliştirilen bir sınıflandırma yöntemidir. Basit ve gürbüz bir algoritması olan ve genellikle doğruluğu karmaşık yöntemler kadar iyi olan modeller üretir. Algoritma, iki sınıfı en iyi ayıran değişkenlerin doğrusal bir kombinasyonunu arama kavramına dayanmaktadır (Fisher, 1936). Bunun için Eşitlik 6.2'de verilen fonksiyonu en büyükmeyi amaçlar. Eşitlik 6.2 ve 6.3'te  $\mu_1$  ve  $\mu_2$  vektör ortalamalarını ifade eder.

$$S(\beta) = \frac{\beta^T \cdot \mu_1 - \beta^T \cdot \mu_2}{\beta^T \cdot C \cdot \beta} \quad (6.2)$$

Eşitlik 6.2'de  $\beta$  doğrusal model vektörlerini, C ise kovaryans matrisi ifade eder. Model katsayıları Eşitlik 6.3'deki gibi ve kovaryans matrisi ise Eşitlik 6.4'deki gibi hesaplanır.

$$\beta = C^{-1}(\mu_1 - \mu_2) \quad (6.3)$$

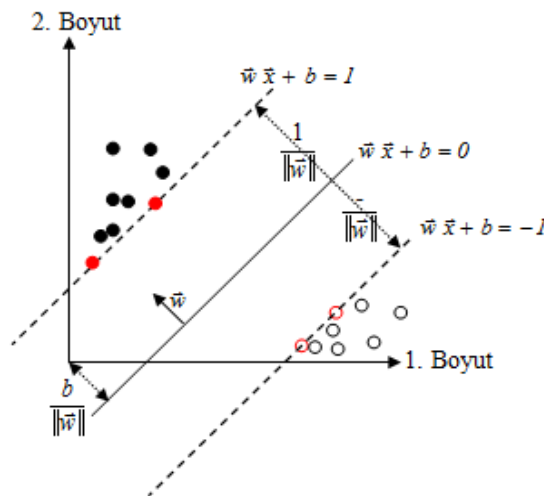
$$C = \frac{n_1 \cdot C_1 + n_2 \cdot C_2}{n_1 + n_2} \quad (6.4)$$

## 6.5. Destek Vektör Makineleri Sınıflandırıcısı

SVM, iki sınıf arasındaki sınırı maksimize eden hiper düzlemi bularak sınıflandırma yapar (Özkan vd., 2015). Hiper düzlemi tanımlayan vektörler (durumlar) destek vektörleridir. Algoritmanın çalışma mantığı şu şekildedir; optimal bir hiper düzlem tanımlanır, sınır maksimize edilir ve veriler, doğrusal karar yüzeyleriyle sınıflandırmanın daha kolay olduğu yüksek boyutlu alana eşlenir. Eşitlik 6.5'de  $\vec{x}_{test}$  vektörü için sınıflandırmanın yapıldığı karar fonksiyonu verilmiştir.

$$f(\vec{x}_{test}) = \sum_{i=1}^M \{\alpha_i \cdot L_i \cdot (\vec{x}_i^T \cdot \vec{x}_{test}) + b\} \quad (6.5)$$

Eşitlik 6.5'de  $\alpha_i$  kuadratik denklemin sıfırdan farklı katsayılarını,  $\vec{x}_i$  vektörleri,  $L_i$  sınıf etiketlerini ve  $\vec{x}_{test}$  test vektörünü belirtir. Sınıflandırıcının çalışma prensibi Şekil 6.1'de verilmiştir. Şekildeki eşitliklerde  $\vec{w}$  hiperdüzleme normal vektörü gösterir ve  $\frac{b}{\|\vec{w}\|}$  hiperdüzlem ile orijin arasındaki en büyüklenecek dik uzaklığı verir (Esener, 2017).



Şekil 6. 1. SVM sınıflandırıcısı çalışma prensibi

## 6.6. Naive Bayes Sınıflandırıcısı

Naive Bayes sınıflandırıcı, tahmin ediciler arasındaki bağımsızlık varsayımlarıyla Bayes'in teoremine dayanmaktadır. Naive Bayesian modelinin oluşturulması kolaydır, karmaşık ve tekrarlayan parametre tahmini yapmadığı için çok büyük veri kümeleri için özellikle yararlıdır. Sadeliğine rağmen, Naive Bayesian sınıflandırıcısı çoğu zaman şaşırtıcı derecede iyi sonuç verir ve sıklıkla daha sofistike sınıflandırma yöntemlerinden daha iyi performans gösterdiğinden yaygın olarak kullanılır. Bayes teoremi Eşitlik 6.6'da verilmiştir. Birbirinden bağımsız ve rastgele gerçekleşen iki olaydan birinin gerçekleşmesi durumunda ikincisinin gerçekleşmesi olasılığıdır.

$$P(A | B) = \frac{P(B|A)P(A)}{P(B)} \quad (6.6)$$

Sınıflandırma yapılırken en yüksek olasılıklı sınıf hedef sınıf olarak Eşitlik 6.7 kullanılarak seçilir. Girdi vektörlerinin birden çok olduğu durumlarda hedef sınıf tahmini yaparken tüm özellikler için koşullu olasılıkların çarpımı yapılır ve Bayes teoremi Eşitlik 6.8'deki gibi farklı bir forma dönüşür (Rish, 2001).

$$S_{hedef} = \underset{s_j \in s}{argmax} P(s_j | v) \quad (6.7)$$

$$P(v_1, v_2, \dots, v_n | s_j) = \prod_{i=1}^n P(v_i | s_j) \quad (6.8)$$

Her girdi için paydadaki değer sabit kaldırılır ve sınıflandırma modeli Eşitlik 6.9'daki gibi hesaplanır.

$$S_{hedef} = \underset{s_j \in s}{argmax} (P(s_j) \prod_{i=1}^n P(v_i | s_j)) \quad (6.9)$$

## 6.7. Lojistik Doğrusal Sınıflandırıcısı

LLC, bağımlı ve bağımsız değişkenler arasında bir ilişki kurarak ayırmsama modeli oluşturur (Webb, 2002). Lojistik regresyon analizi kullanılarak bağımlı ve bağımsız değişkenler arasında bir ilişki kurulur ve doğrusal sınıflandırıcı kullanılarak ayırmsama modeli oluşturulur. Herhangi bir  $x$  örneği için oluşturulan LLC modeli Eşitlik 6.10'da verilmiştir.

$$\log \left( \frac{p(\vec{x}|w_1)}{p(\vec{x}|w_2)} \right) = \vec{\beta}^T \cdot \vec{x} + \beta_0 \quad (6.10)$$

Yöntemde loglinear modeli oluşturulur, kovaryans ve sınıf olasılıkları matrislerinin eşit olduğu varsayımı Eşitlik 6.11'deki gibi yapılır.  $\beta_0'$  Eşitlik 6.12'deki gibi tanımlanır. Karar kriteri Eşitlik 6.13 ve 6.14'de verilmiştir.

$$p(w_1 | \vec{x}) = \frac{\exp(\vec{\beta}^T \cdot \vec{x} + \beta_0)}{1 + \exp(\vec{\beta}^T \cdot \vec{x} + \beta_0)} \quad (6.11)$$

$$p(w_2 | \vec{x}) = \frac{1}{1 + \exp(\vec{\beta}^T \cdot \vec{x} + \beta_0)}$$

$$\beta_0' = \beta_0 + \log \left( \frac{p(w_1)}{p(w_2)} \right) \quad (6.12)$$

$$\vec{x} \in \begin{cases} w_1, & \frac{p(w_1|\vec{x})}{p(w_2|\vec{x})} > 1 \\ w_2, & \frac{p(w_1|\vec{x})}{p(w_2|\vec{x})} < 1 \end{cases} \quad (6.13)$$

$$\vec{x} \in \begin{cases} w_1, & \vec{\beta}^T \cdot \vec{x} + \beta_0 > 0 \\ w_2, & \vec{\beta}^T \cdot \vec{x} + \beta_0 < 0 \end{cases} \quad (6.14)$$

## 7. MATERYAL VE YÖNTEM

Çalışmada JSRT (Japanese Society of Radiological Technology - JSRT) veritabanı kullanılmıştır. Çalışma sonunda sınıflandırma başarısını ölçmek için Çizelge 7.1’de verilen metrikler kullanılmıştır.

**Çizelge 7. 1.** Değerlendirme kriterleri ve matematiksel ifadeleri

Değerlendirme Kriterleri	Matematiksel İfadeleri
TP: Doğru Pozitif (True Positive) TN: Doğru Negatif (True Negative) FP: Yanlış Pozitif (False Positive) FN: Yanlış Negatif (False Negative) S: Sınıf Sayısı $V_j = j.$ sınıfa ait veri sayısı	
Hassasiyet (Sensitivity – SNS)	$\% SNS = \frac{TP}{TP+FN} \cdot 100$
Kesinlik (Specificity – SPC)	$\% SPC = \frac{TN}{TN+FP} \cdot 100$
Doğruluk (Accuracy – ACC)	$\% ACC = \frac{\sum_{j=1}^S SNS_j \cdot V_j}{\sum_{j=1}^S V_j}$

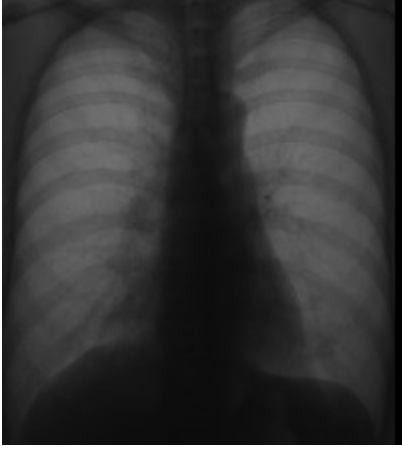
### 7.1. Kullanılan Veritabanı

JSRT veritabanı, 154 x-ray akciğer röntgen imgesi nodüllü ve 93 akciğer röntgen imgesi normal olmak üzere toplamda 247 akciğer imgesinden oluşur (Wang vd., 2016). 154 nodüllü imgeden 100’ü kötü huylu ve 54’ü iyi huyludur. Tüm röntgen görüntülerinin boyutu  $2048 \times 2048$  pikseldir ve gri ölçekli renk derinliği 12 bittir (Wang vd., 2016).

### 7.2. Akciğer Röntgen İmgelerinin Önışlenmesi

JSRT veritabanından örnek bir imge Şekil 7.1’de verilmiştir. Sadece akciğer bölgesini bölütlemek için piksel değeri 45’in altında olan bütün pikseller 0 yapıldı ve akciğer

olmayan dokular kaybedildi. Piksel değeri 45 ve üstü olan bütün pikseller aynı değerde kaldı. Şekil 7.2’de bölütlenme yapılmış akciğer imgesi verilmiştir.



Şekil 7.1. JSRT veritabanından örnek bir imge



Şekil 7.2. Bölütlenmiş İmge

İmgede sadece kemikleri ortaya çıkarmak için bölütlenmiş imgeye 200x200 medyan filtre uygulandı ve medyan filtre uygulanmış imgeden bölütlenmiş imge çıkarıldı (Wang vd., 2016). Şekil 7.3’de bu iki imgenin farkı verilmiştir. Daha sonra kemikleri daha iyi görebilmek için ikilileştirme uygulandı (Wang vd., 2016). Piksel parlaklık değeri 0.0005’in altında olan piksellerin yeni değeri 0, piksel parlaklık değeri 0.0005’in üstünde olan piksellerin yeni değeri ise 1 yapıldı. Şekil 7.4’de ikilileştirme sonucunda elde edilen imge verilmiştir.



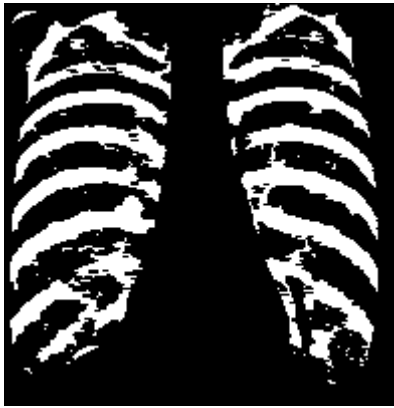
Şekil 7.3. Medyan filrelenmiş ve bölütlenmiş imge farkı



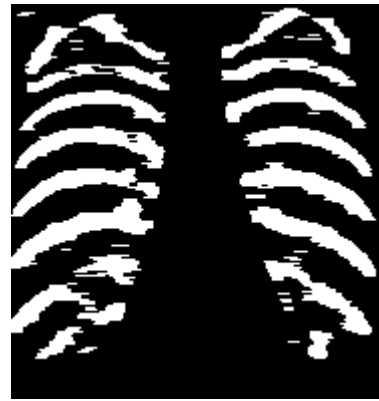
Şekil 7.4. İkilileştirilme uygulanmış imge



Şekil 7.4'de görüldüğü gibi imgede sadece kemikler kaldı ama hala bazı dokular kemiklerle birleşmiş halde bulunmaktadır. Bu dokuları da kaybetmek ve daha düzgün bir kemik yapısı oluşturmak için sırasıyla kapama ve açma morfolojik işlemleri uygulandı. İlk önce kapama işlemi uygulanmıştır. Bu işlem için 20 piksel yarıçapındaki çizgi şeklindeki yapı elemanı oluşturulup, boşluklar Şekil 7.5'deki gibi doldurulmuştur. Daha sonra açma işlemi uygulanmıştır. Bu işlem için 40 piksel yarıçaplı çizgi şeklindeki yapı elemanı oluşturulup, 40 pikselden küçük yarıçapa sahip çizgiler Şekil 7.6'daki gibi çıkarılır.



Şekil 7. 5. Kapama işlemi uygulanmış imge



Şekil 7. 6. Açma işlemi uygulanmış imge

Daha sonra bastırılacak olan kemiklerin lokasyonunun belirlenmesi için şablon eşleme yöntemi kullanılmıştır. Sağ kaburgadan bir kemik sağ kaburga için, sol kaburgadan bir kemik ise sol kaburga için şablon olarak kullanılmıştır. Şablon eşleme yapılmadan önce zamandan tasarruf edebilmek için imgenin boyutu düşürülmüştür. Bir eşik değeri belirlenerek, sağ şablon kemiği sağ kaburga üzerinde gezdirilmiştir. Eğer şablon ve kemik belirlenen eşik değerinin üzerindeyse, kaburgada şablonun bulunduğu lokasyonda kemik olduğu saptanmış ve lokasyonu belirlenmiştir. Bu şekilde sağ kaburgadaki bütün kemikler saptanmıştır. Aynı algoritma sol kaburga içinde sol şablon kemiği kullanılarak uygulanmıştır. Şablon eşleme sonucunda lokasyonu belirlenen kemiklerin piksel değeri Şekil 7.1'deki değeriyle değiştirilmiştir. Şablon eşleme sonucu elde edilmiş imge Şekil 7.7'de verilmiştir. Kemikleri bastırmak için Şekil 7.7'deki imgeye sinüzoidal faktörün dalga boyu 2 olan  $0^\circ$  yönlü Gabor Filtre uygulanmış ve filtrenin büyüklük tepkisi elde edilmiştir (Soleymanpour vd., 2011). Şekil 7.8'deki imgede filtrenin büyüklük tepkisi verilmiştir.



**Şekil 7.7.** Şablon eşleme yapılmış imge

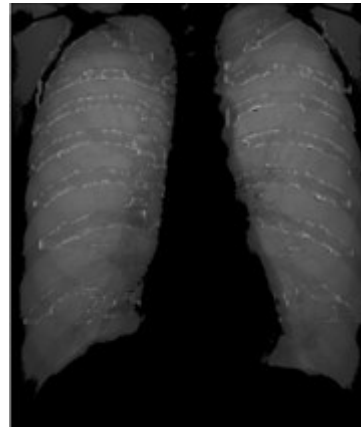


**Şekil 7.8.** Filtrenin büyüklük tepkisi

Şekil 7.7 ve Şekil 7.8'deki imgeler matris toplamı gibi toplanmıştır ve Şekil 7.9'daki imge elde edilmiştir. Şekil 7.9'da elde edilen kemikler Şekil 7.1'deki kemiklerin yerine konmuştur ve Şekil 7.10'daki imge elde edilmiştir.

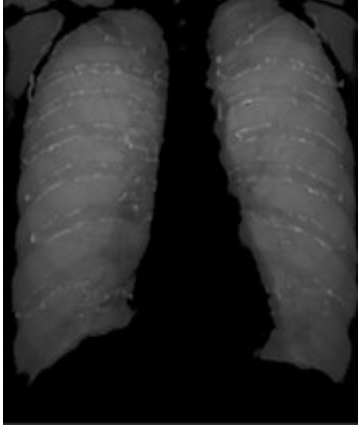


**Şekil 7.9.** İmgelerin toplamı



**Şekil 7.10.** Yeni kemiklerin orijinal imgede yerine konması

Şekil 7.10'deki imgeye 3x3 boyutunda medyan filtre uygulandı ve Şekil 7.11'deki imge elde edildi. Daha sonra açma işlemi uygulanmıştır. Bu işlem için 10 piksel yarıçaplı çizgi şeklindeki yapı elemanı oluşturulup, 10 pikselden küçük yarıçapa sahip çizgiler çıkarılmış ve Şekil 7.12'deki imge elde edilmiştir.

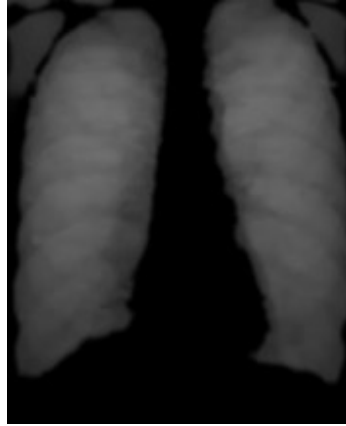


**Şekil 7. 11.** Medyan filtre uygulanmış imge



**Şekil 7. 12.** Açma işlemi uygulanmış imge

Önişlemenin son adımı olarak Şekil 7.12'deki imgeye standart sapması 2 olan Gauss Filtresi uygulanmış ve Şekil 7.13'deki kemiklerin bastırılmış olduğu imge elde edilmiştir (Soleymanpour vd., 2011).

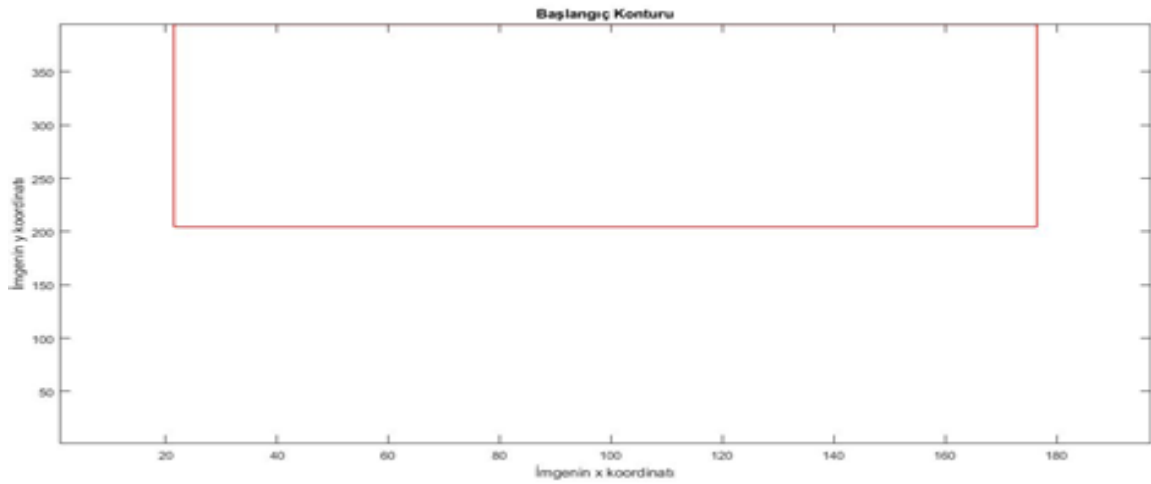


**Şekil 7. 13.** Gauss Filtre uygulanmış imge

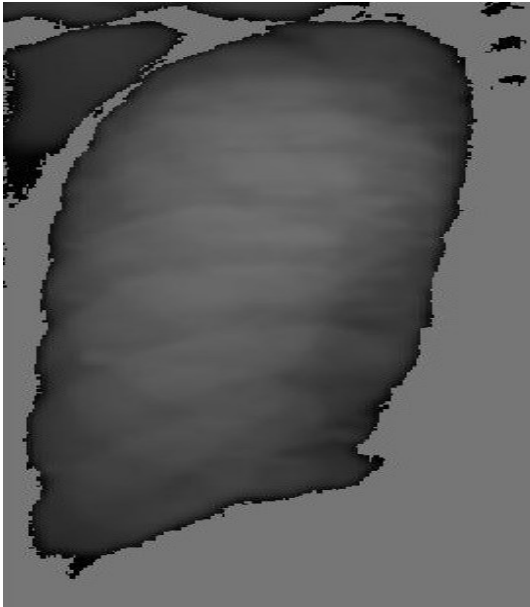
### 7.3. Akciğer Röntgen İmgelerinde Şüpheli Doku Tespiti

Bütün imgelerde Kenarsız Chan-Vese Aktif Kontur algoritması aynı sayıda iterasyon yaparak şüpheli bölgeler bulmuştur. Her imge için sol akciğer ve sağ akciğer olmak üzere ayrı ayrı şüpheli bölgeler bulunmuştur ve başlangıç konturu aynı şekilde uygulanmıştır.

Başlangıç konturu Şekil 7.14'deki gibi alınmıştır. Şüpheli bölgeler bulunmadan önce her imgedeki siyah piksellerin piksel değeri 118 ile değiştirilmiştir. Sol akciğer için; kemikleri bastırılmış bir imgedeki siyah piksellerin 118 piksel değeri ile değişmesi Şekil 7.15'de verilmiştir. Chan-Vese algoritması 1500 iterasyon yaptırılarak şüpheli bölgeler bölütlenmiştir. Şekil 7.16'deki beyaz bölgeler algoritmanın şüpheli bulduğu yerleri göstermektedir. Sağ akciğer içinde aynı işlemler yapılmıştır.



Şekil 7. 14. Başlangıç Konturu



Şekil 7. 15. Siyah piksellerin gri yapılması



Şekil 7. 16. Şüpheli bölgenin bölütlenmesi

#### 7.4. Öznitelik Vektörlerinin Oluşturulması

Her imgede bulunan şüpheli bölgelerden 32x32 boyutunda parça alınıp, bu parçalardan öznitelik vektörleri oluşturulmuştur. Her imge için; ilk önce alt bölüm 5’de anlatılan özniteliklerin hepsi ayrı ayrı çıkartılıp öznitelik vektörü oluşturulmuş ve her öznitelik vektörünün sınıflandırma başarısına bakılmıştır. Daha sonra başarısı yüksek olan öznitelikler birlikte kullanılarak öznitelik vektörü oluşturulmuş ve yine sınıflandırma başarılarına bakılmıştır. Böylece en yüksek başarıyı veren özniteliklerin oluşturduğu öznitelik vektörü elde edilmiştir. Yukarıda anlatılan tüm işlemler Chan-Vese algoritması uygulanmadan direk nodüllü bölgeden 32x32 boyutunda parça alınması ile tekrar edilmiştir.

32x32 boyutundaki matrisin satırları sütun haline getirilmiş ve 1024x1 boyutundaki vektörden alt bölüm 5.1’de anlatılan istatistiksel öznitelikler çıkarılmıştır. Sadece istatistiksel özellikler kullanılarak elde edilen öznitelik vektörü 12x1 boyutludur. 32x32 boyutundaki bölgeye 16x16 hücre boyutlu HOG uygulanmış ve öznitelikler elde edilmiştir. Sadece HOG kullanılarak elde edilen öznitelik vektörü 36x1 boyutludur. 32x32 boyutundaki bölgeden dönme açısından bağımsız olacak şekilde LBP öznitelikleri çıkartılmıştır. Dairenin yarıçapı 1 ve komşuluktaki piksel sayısı 8 olarak alınmıştır. Sadece LBP kullanılarak elde edilen öznitelik vektörü 10x1 boyutludur. 32x32 boyutundaki bölgeden alt bölüm 5.4’de anlatılan geometrik öznitelikler çıkarılmıştır. Sadece geometrik öznitelikler kullanılarak elde edilen öznitelik vektörü 12x1 boyutludur. 32x32 boyutundaki bölgeden GLCM matrisi oluşturulmuştur. GLCM matrisi alt bölüm 5.5’de anlatıldığı gibi  $d=2$  ve  $\Theta=0$  alınarak oluşturulmuştur. Bu GLCM matrisinden tablo 5.2’de verilen öznitelikler çıkartılmıştır. Sadece GLCM matrisinden elde edilen öznitelikler kullanılarak oluşturulan öznitelik vektörü 22x1 boyutludur. 32x32 boyutlu bölgeden DSIFT öznitelikleri çıkartılmıştır. Sadece DSIFT öznitelikleri kullanılarak oluşturulan öznitelik vektörü 64x1 boyutludur.

İstatistiksel ve HOG öznitelikleri birlikte kullanılıp 48x1 boyutunda öznitelik vektörü elde edilmiştir. İstatistiksel ve LBP öznitelikleri birlikte kullanılıp 22x1 boyutunda öznitelik vektörü elde edilmiştir. İstatistiksel ve DSIFT öznitelikleri birlikte kullanılıp 76x1 boyutunda öznitelik vektörü elde edilmiştir. İstatistiksel, HOG ve DSIFT öznitelikleri

birlikte kullanılıp 112x1 boyutunda öznitelik vektörü elde edilmiştir. . İstatistiksel, HOG ve LBP öznitelikleri birlikte kullanılıp 58x1 boyutunda öznitelik vektörü elde edilmiştir.

### **7.5. Şüpheli Olarak Tespit Edilen Dokuların Sınıflandırılması**

Sağlıklı, iyi huylu nodüllü ve kötü huylu nodüllü olmak üzere üç sınıf vardır ve sınıflandırma üç sınıf üzerinden gerçekleştirilmiştir. Sınıflandırmada 10 katlamalı çapraz doğrulama yöntemi kullanılmıştır. Her sınıftaki imgelerin %90'ı eğitim ve %10'u test amaçlı kullanılmıştır. Sınıflandırma, 10 aşamada tamamlanmıştır. Her aşamada her sınıfın farklı %10'u test aşamasına girmiş ve sınıflandırma başarısı hassasiyet, doğruluk ve kesinlik üzerinden hesaplanmıştır. Her aşamada bulunan doğruluk, hassasiyet ve kesinlik değerleri toplanıp ortalaması alınmış ve sistemin hassasiyeti, doğruluğu ve kesinliği bulunmuştur. Kullanılan sınıflandırıcılar k-NN(k=5), rastgele orman(ağaç sayısı=100), karar ağacı, naive bayes, LLC, SVM ve FLDA'dır.

## 8. BULGULAR VE TARTIŞMA

Chan-Vese algoritmasının şüpheli bölgeyi bulma hassasiyeti %73 olarak hesaplanmıştır. Şüpheli bölgelerden çıkarılan özniteliklerin sınıflandırma başarıları doğruluk, hassasiyet ve kesinlik değerleri üzerinden verilmiştir.

Sadece 12x1 boyutlu istatistiksel öznitelikler kullanılarak elde edilen doğruluk, hassasiyet ve kesinlik değerleri Çizelge 8.1’de verilmiştir. Sadece 22x1 boyutlu GLCM öznitelikleri kullanılarak elde edilen doğruluk, hassasiyet ve kesinlik değerleri Çizelge 8.2’de verilmiştir. Sadece 10x1 boyutlu LBP öznitelikleri kullanılarak elde edilen doğruluk, hassasiyet ve kesinlik değerleri Çizelge 8.3’de verilmiştir.

**Çizelge 8. 1.** Sadece istatistiksel öznitelikler kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri.

	k-NN	FLDA	Rastgele Orman	Karar Ağacı	SVM	Naive Bayes	LLC
<b>Doğruluk</b>	53%	59%	60%	56%	33%	59%	43%
<b>Hassasiyet</b>	50%	49%	51%	50%	26%	51%	33%
<b>Kesinlik</b>	76%	78%	79%	78%	63%	79%	67%

**Çizelge 8. 2.** Sadece GLCM öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri.

	k-NN	FLDA	Rastgele Orman	Karar Ağacı	SVM	Naive Bayes	LLC
<b>Doğruluk</b>	39%	43%	44%	38%	48%	45%	49%
<b>Hassasiyet</b>	36%	37%	39%	34%	39%	43%	40%
<b>Kesinlik</b>	70%	71%	71%	68%	71%	73%	71%

**Çizelge 8. 3.** Sadece LBP öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri.

	k-NN	FLDA	Rastgele Orman	Karar Ağacı	SVM	Naive Bayes	LLC
<b>Doğruluk</b>	42%	50%	54%	46%	52%	45%	52%
<b>Hassasiyet</b>	41%	41%	47%	42%	42%	40%	43%
<b>Kesinlik</b>	71%	72%	75%	72%	73%	71%	73%

Sadece 12x1 boyutlu geometrik öznitelikler kullanılarak elde edilen doğruluk, hassasiyet ve kesinlik değerleri Çizelge 8.4'de verilmiştir. Sadece 36x1 boyutlu HOG öznitelikleri kullanılarak elde edilen doğruluk, hassasiyet ve kesinlik değerleri Çizelge 8.5'de verilmiştir. Sadece 64x1 boyutlu DSIFT öznitelikleri kullanılarak elde edilen doğruluk, hassasiyet ve kesinlik değerleri Çizelge 8.6'da verilmiştir.

**Çizelge 8. 4.** Sadece geometrik öznitelikler kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri.

	k-NN	FLDA	Rastgele Orman	Karar Ağacı	SVM	Naive Bayes	LLC
<b>Doğruluk</b>	40%	48%	46%	44%	48%	47%	47%
<b>Hassasiyet</b>	38%	40%	39%	41%	39%	40%	39%
<b>Kesinlik</b>	70%	71%	71%	72%	71%	72%	70%

**Çizelge 8. 5.** Sadece HOG öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri.

	k-NN	FLDA	Rastgele Orman	Karar Ağacı	SVM	Naive Bayes	LLC
<b>Doğruluk</b>	43%	49%	53%	43%	55%	45%	55%
<b>Hassasiyet</b>	38%	43%	44%	40%	46%	39%	45%
<b>Kesinlik</b>	72%	73%	74%	71%	75%	72%	75%

**Çizelge 8. 6.** Sadece DSIFT öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri.

	k-NN	FLDA	Rastgele Orman	Karar Ağacı	SVM	Naive Bayes	LLC
<b>Doğruluk</b>	50%	45%	58%	45%	47%	51%	51%
<b>Hassasiyet</b>	45%	41%	48%	41%	37%	43%	41%
<b>Kesinlik</b>	75%	72%	77%	72%	69%	74%	72%

Yukarıda verilen çizelgeler incelendiğinde en yüksek doğruluk, hassasiyet ve kesinlik değerlerini veren özniteliklerin istatistiksel öznitelikler olduğu görülür. LBP, HOG ve DSIFT öznitelikleri de yüksek değerler vermiştir. Bundan sonraki aşamada en yüksek değerleri veren istatistiksel özniteliklerine ayrı ayrı LBP, HOG ve DSIFT öznitelikleri eklenerek sınıflandırma başarılarına bakılmıştır.



İstatistiksel ve HOG öznitelikleri birlikte kullanıldığında oluşan 48x1 boyutlu öznitelik vektörü ile elde edilen doğruluk, kesinlik ve hassasiyet değerleri Çizelge 8.7’de verilmiştir. İstatistiksel ve DSIFT öznitelikleri birlikte kullanıldığında oluşan 76x1 boyutlu öznitelik vektörü ile elde edilen doğruluk, kesinlik ve hassasiyet değerleri Çizelge 8.8’de verilmiştir. İstatistiksel ve LBP öznitelikleri birlikte kullanıldığında oluşan 22x1 boyutlu öznitelik vektörü ile elde edilen doğruluk, kesinlik ve hassasiyet değerleri Çizelge 8.9’de verilmiştir.

**Çizelge 8. 7.** İstatistiksel ve HOG öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri.

	k-NN	FLDA	Rastgele Orman	Karar Ağacı	SVM	Naive Bayes	LLC
<b>Doğruluk</b>	52%	57%	64%	54%	32%	52%	43%
<b>Hassasiyet</b>	50%	51%	55%	49%	25%	47%	33%
<b>Kesinlik</b>	76%	78%	81%	77%	62%	76%	67%

**Çizelge 8. 8.** İstatistiksel ve DSIFT öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri.

	k-NN	FLDA	Rastgele Orman	Karar Ağacı	SVM	Naive Bayes	LLC
<b>Doğruluk</b>	51%	45%	59%	50%	32%	51%	43%
<b>Hassasiyet</b>	48%	39%	49%	46%	25%	43%	34%
<b>Kesinlik</b>	76%	72%	78%	75%	62%	75%	67%

**Çizelge 8. 9.** İstatistiksel ve LBP öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri.

	k-NN	FLDA	Rastgele Orman	Karar Ağacı	SVM	Naive Bayes	LLC
<b>Doğruluk</b>	52%	51%	60%	54%	32%	53%	41%
<b>Hassasiyet</b>	50%	43%	51%	50%	25%	48%	33%
<b>Kesinlik</b>	76%	74%	79%	77%	62%	77%	66%

Son olarak ayrı ayrı istatistiksel, DSIFT, HOG ve istatistiksel, HOG, LBP öznitelikleri kullanılarak sistemin sınıflandırma başarısına bakılmıştır. İstatistiksel, DSIFT ve HOG öznitelikleri birlikte kullanıldığında oluşan 112x1 boyutlu öznitelik vektörü ile elde edilen doğruluk, kesinlik ve hassasiyet değerleri Çizelge 8.10’da verilmiştir. İstatistiksel, HOG ve LBP öznitelikleri birlikte kullanıldığında oluşan 58x1 boyutlu öznitelik vektörü ile elde edilen doğruluk, kesinlik ve hassasiyet değerleri Çizelge 8.11’de verilmiştir.

**Çizelge 8. 10.** İstatiksel, DSIFT ve HOG öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri.

	k-NN	FLDA	Rastgele Orman	Karar Ağacı	SVM	Naive Bayes	LLC
<b>Doğruluk</b>	51%	43%	60%	51%	32%	53%	42%
<b>Hassasiyet</b>	49%	39%	49%	46%	26%	45%	34%
<b>Kesinlik</b>	76%	74%	79%	77%	62%	77%	66%

**Çizelge 8. 11.** İstatiksel, HOG ve LBP öznitelikleri kullanılarak elde edilen doğruluk, kesinlik ve hassasiyet değerleri.

	k-NN	FLDA	Rastgele Orman	Karar Ağacı	SVM	Naive Bayes	LLC
<b>Doğruluk</b>	53%	55%	65%	50%	32%	53%	43%
<b>Hassasiyet</b>	51%	49%	56%	45%	26%	49%	33%
<b>Kesinlik</b>	77%	77%	81%	75%	63%	77%	67%

Yukarıdaki çizelgeler incelendiğinde en iyi sonucun rasgele orman sınıflandırıcısı kullanıldığında % 65 doğruluk, % 56 hassasiyet ve %81 kesinlik ile istatistiksel, HOG ve LBP özniteliklerinin oluşturduğu öznitelik vektörü kullanıldığında elde edildiği görülmüştür.

Chan-Vese algoritması kullanılmadan direk nodüllü bölgeden alınan 32x32 boyutundaki parçalardan çıkarılan özniteliklerin sınıflandırma sonuçları da Chan-Vese algoritması kullanıldığında ortaya çıkan sonuçlarla aynıdır.

## 9. SONUÇ VE ÖNERİLER

Bu çalışmada, akciğer x-ray röntgenlerinde akciğer dokusu üzerindeki şüpheli yapıları otomatik olarak tespit edilebildiği ve bu yapıları iyi huylu, kötü huylu ve normal doku olarak sınıflara ayırabilen bir yaklaşım önerilmiştir. Çalışmada 247 adet akciğer x-ray röntgeni içeren kötü huylu nodüllü, iyi huylu nodüllü ve normal hasta görüntülerinden oluşan bir veritabanı kullanılmıştır. Çalışma üç ana kısımdan oluşturulmuştur. İlk aşamada röntgen üzerinde akciğer dokusu etrafında bulunan diğer doku görüntüleri eşikleme yöntemi ile çıkartılmıştır. Akciğer röntgenlerinde nodül yapısının kemik altında kalması durumunda teşhisi zor olmaktadır. Şüpheli bölge tespitine yönelik çalışmalardaki en önemli adım bu zorluğun giderilmesine yönelik olmalıdır. Çalışma kapsamında ön işleme aşamasında şablon eşleme yöntemi ile kaburga kemiklerinin yerleri saptanmış ve bu yerlere Gabor filtresi uygulanarak kemikler bastırılmıştır.

Çalışmanın ikinci kısmında akciğer röntgen imgelerinde farklı dokuların tespitine yönelik çalışmalar yapılmıştır. Bu amaç doğrultusunda Kenarsız Chan-Vese algoritması kullanılmıştır. Her akciğer röntgeninde sağ ve sol loblar üzerinde bu algoritma denenmiş ve şüpheli bölgeler tespit edilmiştir. Elle yapılan tespit ile Kenarsız Chan-Vese algoritması karşılaştırıldığında algoritmanın şüpheli bölgeyi bulma hassasiyeti %73 olarak hesaplanmıştır. Bu da Chan-Vese algoritmasının şüpheli bölgeleri başarıyla çıkardığını göstermektedir. Elde edilen şüpheli bölgeden altı farklı öznelik çıkarımı yapılmıştır. Bunlar istatistiksel, geometrik, LBP, GLCM, HOG ve DSIFT öznelikleridir.

Tezin son kısmı Kenarsız Chan-Vese algoritması ile elde edilen şüpheli bölgelerin sınıflandırmasına yönelik gerçekleştirilmiştir. Sınıflandırmada,  $k$ -En Yakın Komşu, Karar Ağacı, Rastgele Orman, Lojistik Regresyon Sınıflandırıcısı, Destek Vektör Makineleri, Fisher Doğrusal Ayırtaç Analizi ve Naive Bayes sınıflandırıcıları kullanılmıştır. Sınıflandırma 10 katlamalı çapraz doğrulama yöntemi kullanılarak yapılmıştır. Her sınıftaki imgelerin %90'ı eğitim ve %10'u test amaçlı kullanılmıştır. Sınıflandırma, 10 aşamada tamamlanmıştır. Her aşamada her sınıfın farklı %10'u test aşamasına girmiş ve sınıflandırma başarısı hassasiyet, doğruluk ve kesinlik üzerinden hesaplanmıştır. İlk önce her öznelik

grubunu ayrı ayrı sınıflandırıcı algoritmalara sokup, sınıflandırma başarıları incelenmiştir. En yüksek başarılı sınıflandırma sonucunu veren öznitelikler istatistiksel özniteliklerdir. LBP, HOG ve DSIFT öznitelikleri de başarılı sonuçlar vermiştir. En iyi sonuç veren öznitelik istatistiksel öznitelikler olduğundan, istatistiksel özniteliklere ayrı ayrı HOG, LBP ve DSIFT öznitelikleri eklenip sınıflandırma başarılarına bakılmıştır. İstatistiksel ve HOG öznitelikleri diğerlerine göre daha iyi sonuç vermiştir. Bundan dolayı bu aşamadan sonra İstatistiksel ve HOG özniteliklerine sırasıyla LBP ve DSIFT öznitelikleri eklenerek sınıflandırma sonuçlarına bakılmıştır. İstatistiksel, HOG ve LBP öznitelikleri kullanılarak öznitelik vektörü oluşturulduğunda rasgele orman sınıflandırıcısı ile en iyi sınıflandırma sonuçlarını verdiği görülmüştür. Sınıflandırmada; doğruluk %65, hassasiyet %56 ve kesinlik %81 olarak hesaplanmıştır.

Literatürde JSRT veritabanını kullanan çalışmalar incelendiğinde ve bulunan hassasiyet sonuçlarına bakıldığında bu çalışmanın en iyi sekizinci sonucu verdiği görülmüştür. Öznitelikler Chan-Vese algoritması kullanılmayıp direk nodüllü bölgeden çıkarıldığında elde edilen sınıflandırma sonuçları, Chan-Vese algoritması kullanıp elde edilen sonuçlarla aynı çıkmıştır. Bu da Chan-Vese algoritmasının umut verici sonuçlara ulaştığını göstermektedir. Çalışma kapsamında akciğer kanseri yönünden şüpheli dokuların otomatik tespiti için kullanılan Chan-Vese algoritması ve sınıflandırıcıların tez çalışması kapsamında hedeflenen başarıya ulaştığı ve umut verici bir yaklaşım olduğu görülmektedir. Bu alanda yapılan çalışmalarda kullanılan veritabanlarının kısıtlı olması, geniş imge içerikli veritabanlarının eksikliği ve veritabanları içerisindeki imgelerdeki bozukluklar gibi sorunlar günümüzde hala giderilmesi gereken en önemli problemlerdir. Yapılan bu tez çalışmasının yapılan diğer araştırmalara yol göstermesi, gelecekte x-ray veya diğer görüntüleme teknikleri üzerinde yapılacak çalışmalara ışık tutması ve literatüre katkı sağlanması beklenmektedir. Gelecek çalışmada, daha hassas bir şekilde şüpheli bölge bulan bir algoritma kullanılarak bu bölgeden çıkarılan öznitelikler ile sistemin başarısı arttırılmaya çalışılacaktır. Daha farklı özniteliklerinde çalışmaya dahil edilmesi sınıflandırma başarısını arttırabilir. Ayrıca, diğer çalışmalarda yapıldığı gibi veritabanında bulunan çok bozuk olan imgeler çalışmaya dahil edilmezse, sınıflandırma başarısı artacaktır.

## KAYNAKLAR DİZİNİ

- Ahonen, T., Hadid, A., Pietkäinen, M., 2006, Face description with local binary patterns: Application to face recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28, 2037-2041.
- Armato, S.G., Giger, M.L., Macmahon, H., 2002, Method and system for the segmentation of lung regions in lateral chest radiographs, United States Patent.
- Candemir, S., Jaeger, S., Palaniappan, K., Musco, J.P., Singh, R.K., Zhiyun, X., Karargyris, A., Antani, S., Thoma, G., McDonald, C.J., 2014, Lung Segmentation in Chest Radiographs Using Anatomical Atlases With Nonrigid Registration, *IEEE transactions on medical imaging*, 33, 2, 577-590.
- Chan, T.F., Vese, L.A., 2001, Active contours without edges, *IEEE Transactions on Image Processing*, 10, 2, 266-277.
- Chen, T., Takagi, M., 1993, Image Binarization By Back Propagation Algorithm, *International Archives of Photogrammetry and Remote Sensing*, 29, 345-345.
- Clausi, D. A., 2002, An analysis of co-occurrence texture statistics as a function of grey level quantization, *Canadian Journal of Remote Sensing*, 28, 1, 45-62.
- Esener, İ. I., Ergin, S., Yuksel, T., 2016, A Genuine GLCM-based Feature Extraction for Breast Tissue Classification on Mammograms, *International Journal of Intelligent Systems and Applications in Engineering*, 4, 124-129.
- Esener, İ. I., Ergin, S., Yuksel, T., 2017, A New Feature Ensemble with a Multistage Classification Scheme for Breast Cancer Diagnosis, *Journal of Healthcare Engineering*, 2017, 1-15.
- Esener, İ.I., 2017, Meme Kanseri için Şüpheli Bölgelerin Mamografi İmgeleri Üzerinde Belirlenmesi ve Meme Kanseri Türünün Sınıflandırılması, Doktora Tezi, Eskişehir Osmangazi Üniversitesi, 150 s.
- Fisher, R.A., 1936, The use of multiple measurements in taxonomic problems, *Annals of Eugenics*, 7, 179-188.
- Gabor, D., 1946, Theory of communication, *Journal of the Institution of Electrical Engineers- Part III: Radio and Communication Engineering*; 93, 26, 429-457.
- Garfinkel, L., Murphy, G., Lawrence, W.J., Lenhard, R.J., 1995, *American Cancer Society Textbook of Clinical Oncology*, The Society Press.
- Gallagher N., Wise G., 1981, A theoretical analysis of the properties of median filters, *IEEE Transaction on Acoustic Speech Signal Processing*, 29, 6, 1135-1141.

### KAYNAKLAR DİZİNİ (devam)

- Ginneken, V.B., Stegman, M.B., Loog, M., 2006, Segmentation of Anatomical Structures in Chest Radiographs using Supervised Methods: A Comparative Study on a Public Database, *Medical Image Analysis*, 10, 19-40.
- Gonzalez, E.R., Ponomaryov, V., 2016, Automatic Lung nodule segmentation and classification in CT images based on SVM, 9th International Kharkiv Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves, 1-4.
- Haddad, R.A., Akansu, A.N., 1991, A Class of Fast Gaussian Binomial Filters for Speech and Image Processing, *IEEE Transactions on Acoustics, Speech and Signal Processing*, 39, 3, 723-727.
- Haralick, R. M., Shanmugam, K., Dinstein, I., 1973, Textural features of image classification, *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3, 6, 1973.
- Hardie, RC., Rogers, S.K., Wilson, T., Rogers, A., 2007, Performance analysis of a new computer aided detection system for identifying lung nodules on chest radiographs, *Medical Image Analysis*, 12, 240-258.
- Kim, J., Kim, B.S., Savarese, S., 2012, Comparing image classification methods:K-nearest-neighbor and support-vector-machines, 6. WSEAS International Conference on Computer Engineering and Applications, 133-138.
- Lee, S.J., Wang, J.W., Wuc, H.H., Yuand, M.Z., 2012, A nonparametric-based rib suppression method for chest radiographs, *Computers & Mathematics with Applications*, 64, 5, 1390-1399.
- Lowe, D.G., 1999, Object recognition from local scale-invariant features, 7th International Conference on Computer Vision, 1150-1157.
- Mumford, D., Shah, J., 1989, Optimal approximations by piecewise smooth functions and associated variational problems, *Communication on Pure and Applied Mathematics*., 42, 5, 577-685.
- Özkan, K., Ergin, S., Işık, Ş., Işıklı, İ., 2015, A new classification scheme of plastic wastes based upon recycling labels, *Waste Management*, 35, 29-35.
- Rish, I., 2001, An empirical study of the naive Bayes classifier, *IJCAI Workshop on Empirical Methods in artificial intelligence*, 3, 41-46.
- Said, K.A.M., Jambek, A.B., Sulaiman, N., 2016, A study of image processing using morphological opening and closing processes, *International Journal of Control Theory and Applications*, 9, 31, 15-21.
- Safavian, S.R., Landgrebe, D., 1991, A survey of decision tree classifier methodology, *IEEE Transactions on Systems, Man, and Cybernetics*, 21, 660-674.

### KAYNAKLAR DİZİNİ (devam)

- Shiraishi, J., Abe, H., Li, F., Engelmann, R., MacMahon, H., Doi, K., 2006, Computer-aided diagnosis for the detection and classification of lung cancers on chest radiographs ROC analysis of radiologists performance, *Academic Radiology*, 13, 8, 995-1003.
- Shiraishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T., Komatsu, K., Matsui, M., Fujita, H., Kodera, Y., and Doi, K., 2000, Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists' detection of pulmonary nodules, *AJR* 174, 71-74.
- Soh, L., Tsatsoulis, C., 1999, Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices, *IEEE Transactions on Geoscience and Remote Sensing*, 37, 2, 780-795.
- Song, L., Liu, X., Ma, L., Zhou, C., Zhao, X. and Zhao, Y., 2012, Using HOG-LBP features and MMP learning to recognize imaging signs of lung lesions, 25. *International Symposium On Computer-Based Medical Systems*, 1-4.
- Soleymanpour , E., Pourreza, H.R., Ansariour, E., Sadooghi, M., 2011, Fully Automatic Lung Segmentation and Rib Suppression Methods to Improve Nodule Detection in Chest Radiographs, *Journal of Medical Signals and Sensors*, 1, 3, 191-199.
- Stewart, B., Wild, C.P., 2015, *World Cancer Report 2014*, WHO Press.
- Suzuki, K., Abe, H., MacMahon, H., Doi, K., 2006, Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN), *IEEE Transactions Medical Imaging*, 25, 4, 406-416.
- Wang, C., Elazab, A., Wu, J., Hua, Q., 2017, Lung nodule classification using deep feature fusion in chest radiography, 57, 10-18.
- Wang, L., He, D.C., 1990, Texture Classification Using Texture Spectrum, *Pattern Recognition*, 23, 8, p. 905-910.
- Wang, J., Li, F., Yulin, JI., Liang, Z., 2017, Kind of x-ray chest image rib suppression method based on poisson model, *United States Patent*.
- Wang, J.G., Li, J., Yau, W.-Y. and Sung, E., 2010, Boosting dense SIFT descriptors and shape contexts of face images for gender recognition, *IEEE Computer Society Conference On Computer Vision and Pattern Recognition Workshops*, 96-102.
- Webb, A.R., 2002, *Linear discriminant analysis in Statistical Pattern Recognition*, John Wiley & Sons, New York, p. 123-124.

**KAYNAKLAR DİZİNİ (devam)**

- Wei, J., Hagihara, Y., Shimuzu, A., Kobatake, H., 2002, Optimal image feature set for detecting lung nodules on chest X-ray images, *Computer Assisted Radiology and Surgery*, 706-711.
- Wu, G., Zhang, X., Luo, S., Hu, Q., 2015, Lung Segmentation Based on Customized Active Shape Model from Digital Radiography Chest Images, *Journal of Medical Imaging and Health Informatics*, 5, 2, 184-191.