

Meme Kanseri için Şüpheli Bölgelerin Mamografi İmgeleri Üzerinde Belirlenmesi ve
Meme Kanser Türünün Sınıflandırılması

İdil Işıklı Esener

DOKTORA TEZİ

Elektrik Elektronik Mühendisliği Anabilim Dalı

Haziran 2017

The Identification of Suspicious Regions on Mammography Images for Breast Cancer and
the Classification of Breast Cancer Type

İdil Işıklı Esener

DOCTORAL DISSERTATION

Department of Electrical and Electronics Engineering

June 2017

Meme Kanseri için Şüpheli Bölgelerin Mamografi İmgeleri Üzerinde Belirlenmesi ve
Meme Kanser Türünün Sınıflandırılması

İdil Işıklı Esener

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

Danışman: Doç. Dr. Semih ERGİN

Haziran 2017

ONAY

Elektrik Elektronik Mühendisliği Anabilim Dalı Doktora öğrencisi İdil IŞIKLI ESENER'in DOKTORA tezi olarak hazırladığı "Meme Kanseri için Şüpheli Bölgelerin Mamografi İmgeleri Üzerinde Belirlenmesi ve Meme Kanseri Türünün Sınıflandırılması" başlıklı bu çalışma, jürimizce lisansüstü yönetmeliğin ilgili maddeleri uyarınca değerlendirilerek oybirliği ile kabul edilmiştir.

Danışman : Doç. Dr. Semih Ergin



İkinci Danışman : Yrd. Doç. Dr. Tolga Yüksel

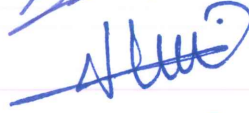


Doktora Tez Savunma Jürisi:

Üye : Prof. Dr. M. Bilginer Gülmezoğlu



Üye : Doç. Dr. Semih Ergin



Üye : Doç. Dr. Serkan Günel



Üye : Yrd. Doç. Dr. Mehmet Koç



Üye : Yrd. Doç. Dr. Kemal Keskin



Fen Bilimleri Enstitüsü Yönetim Kurulu'nuntarih ve
.....sayılı kararıyla onaylanmıştır.

Prof. Dr. Hürriyet ERŞAHAN
Enstitü Müdürü

ETİK BEYAN

Eskişehir Osmangazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna göre, Doç. Dr. Semih Ergin ve Yrd. Doç. Dr. Tolga Yüksel danışmanlığında hazırlamış olduğum “Meme Kanseri için Şüpheli Bölgelerin Mamografi İmgeleri Üzerinde Belirlenmesi ve Meme Kanseri Türünün Sınıflandırılması” başlıklı DOKTORA 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. 14/06/2017

İdil IŞIKLI ESENER

ÖZET

Tez çalışmasının ilk bölümünde, IRMA veritabanı kullanılarak meme kanseri teşhisi yapabilen bir CAD sistemi önerilmiştir. Çalışmada, LCP öznitelikleri, istatistiksel ve frekans-düzlemi öznitelikler ile birleştirilerek 108-boyutlu LCP-tabanlı öznitelik vektörleri elde edilmiştir. Kanser teşhisi, veritabanının yalnızca yağlı doku tipine sahip imgeleri üzerinde 3-sınıflı ve veritabanındaki tüm imgeler ele alınarak 12-sınıflı olarak gerçekleştirilmiştir. 12-sınıflı sınıflandırma çalışması için 1- , 2- ve 3-aşamalı sınıflandırma süreçleri önerilmiş ve bu veritabanı için 12-sınıflı, 3-aşamalı sınıflandırma sürecinin %93,52 doğruluk oranı ile en başarılı performansı sağladığı görülmüştür. İstatistiksel, frekans-düzlemi ve LCP öznitelik vektörleri literatürde kullanılmakta iken, bu üç vektörün birleştirilmesi ile elde edilen ve veriyi ayırt ediciliği daha yüksek olan 108-boyutlu LCP-tabanlı öznitelik vektörlerinin oluşturulması çalışmanın özgünlüğünü göstermektedir.

Tez çalışmasının ikinci bölümünde ise MIAS veritabanı kullanılarak meme doku tespiti, meme kanseri tespit ve teşhisi yapabilen bir CAD sistemi önerilmiştir. Çalışmada, pektoral kas tespiti için, bölge büyütme ve doğru oturtma algoritmaları birlikte kullanılarak özgün bir yöntem geliştirilmiş ve %4,15 ortalama FPR, %3,28 ortalama FNR değerleri elde edilmiştir. Meme doku tipi tespiti için GLCM matrislerinden çıkartılan doku özniteliklerinin 1- ve 2-aşamalı sınıflandırma süreçleri sonucunda %82,48 doğruluk oranı elde edilmiştir. Chan-Vese aktif kontur modelleme yöntemi kullanılarak ROI tespiti gerçekleştirilmiş ve ROI'lerden 11-boyutlu mamografik öznitelikler çıkartılarak indekslenmiştir. İndekslenen özniteliklerin sınıflandırılması sonucunda %83,25 doğruluk oranı ile meme kanseri teşhisi başarılmıştır.

Anahtar Kelimeler: Meme kanseri, bilgisayar destekli teşhis, öznitelik çıkarımı, pektoral kas

SUMMARY

In the first part of the thesis study, a CAD system implementing breast cancer diagnosis using IRMA database is proposed. In the study, 108-dimensional LCP-based feature vectors are obtained by concatenating the LCP features with statistical and frequency-domain features. Cancer diagnosis is realized as 3-class classification using only the fatty tissue type of the database and 12-class classification by considering all images in the database. For the 12-class classification study, 1-, 2- and 3-stage classification processes are proposed and it is observed that 12-class, 3-stage classification process performs best with 93.52% accuracy rate for this database. Although the statistical, frequency-domain and LCP feature vectors are used in the literature, construction of higher discriminative 108-dimensional LCP-based feature vectors obtained by concatenating these three vectors shows the originality of this study.

In the second part of the thesis study, a CAD system implementing breast tissue type detection, breast cancer detection and diagnosis using MIAS database is proposed. In the study, an original method is developed for pectoral muscle detection by using region growing and line fitting algorithms together, and 4.15% mean FPR, 3.28% mean FNR values are obtained. 82.48% accuracy rate is obtained concluding by 1- and 2-stage classification of the tissue features extracted from the GLCM matrices for breast tissue type detection. ROI detection is performed by using Chan-Vese active contour modeling method, and 11-dimensional mammographic features of the ROIs are extracted and indexed. In consequence of classification of the indexed features, breast cancer diagnosis is achieved with 83.25% accuracy.

Keywords: Breast cancer, computer aided diagnosis, feature extraction, pectoral muscle

TEŞEKKÜR

Doktora çalışmalarım süresince, danışman hocam Doç. Dr. Semih ERGİN engin bilgilerini hiçbir zaman benden esirgememiştir. Hem verdiği bilgilerden ve desteklerden dolayı, hem de tez çalışmam süresince sabırla beni motive etmesinden dolayı kendisine teşekkürü bir borç bilirim.

İkinci danışman hocam Yrd. Doç. Dr. Tolga YÜKSEL'e doktora çalışmam süresince tez çalışmasına verdiği desteklerden ve lisansüstü eğitimim süresince verdiği motivasyon desteğinden ötürü teşekkür ederim.

Tez çalışmasındaki sonuçların ve veri tabanlarının radyoloji açısından yorumlanmasında sunduğu önemli katkılardan ötürü Prof. Dr. Cüneyt ÇALIŞIR'a teşekkürü borç bilirim.

Prof. Dr. M. Bilginer GÜLMEZOĞLU ve Doç. Dr. Serkan GÜNAL'ın tez çalışmasının izleme süresince göstermiş oldukları destek, ilgi ve değerli katkılarından dolayı kendilerine teşekkür ederim.

Doktora çalışmam süresince tükenmez sabırları ile bana destek olan sevgili eşim Emre ESENER ve değerli kızım Öykü ESENER'e, ayrıca hayatım boyunca maddi ve manevi desteklerini esirgemeyen ve her zaman yanımda olduklarını hissettiren kıymetli aileme teşekkürlerimi bir borç bilirim.

İÇİNDEKİLER DİZİNİ

	<u>Sayfa</u>
ÖZET	vi
SUMMARY	vii
İÇİNDEKİLER DİZİNİ	ix
ŞEKİLLER DİZİNİ	xii
ÇİZELGELER DİZİNİ	xvi
SİMGELER VE KISALTMALAR DİZİNİ	xviii
1. GİRİŞ VE AMAÇ	1
2. LİTERATÜR ARAŞTIRMASI	5
2.1. Mikrokalsifikasyon Tespiti Yapılan Çalışmalar	5
2.2. Kitle (Anomali) Tespiti Yapılan Çalışmalar	9
2.3. Anomali Tipi Sınıflandırması Yapılan Çalışmalar	15
2.4. Pektoral Kas Tespiti Yapılan Çalışmalar	21
2.5. Kanser Riski Tespit Edilen Çalışmalar	23
2.6. Diğer Çalışmalar	24
3. MAMOGRAFİ İMGELERİNİN ÖN İŞLENMESİ	26
3.1. Kontrast İyileştirme	26
3.2. Mamografi İmgelerinin İyileştirilmesi.....	26
3.2.1. Yerel-olmayan ortalama filtresi.....	27
3.2.2. Medyan filtresi.....	28
3.2.3. Merkez-ağırlıklı medyan filtresi	28
3.2.4. Frost filtresi	28
3.2.5. Bilateral filtre.....	29
3.3. Mamografi İmgelerinin Pektoral Kaslardan Arındırılması.....	30
3.3.1. Bölge büyütme algoritması.....	30
3.3.2. Pektoral kas tespiti doğruluk ölçütü	31
4. MAMOGRAFİ İMGELERİNDE MEME DOKU TİPİ TESPİTİ	34
5. MAMOGRAFİ İMGELERİNDE ŞÜPHELİ BÖLGE TESPİTİ	35
5.1. Aktif Kontur Model – Yılan	35

İÇİNDEKİLER DİZİNİ (devam)

	<u>Sayfa</u>
5.2. Chan-Vese Aktif Kontur Modeli	36
6. MAMOGRAFİ İMGELERİNDEN ÖZİNTELİK ÇIKARTIMI.....	40
6.1. Yerel Konfigürasyon Örüntüsü.....	40
6.2. İstatistiksel Öznitelikler	43
6.3. Frekans-Düzlemi Öznitelikleri	45
6.4. Doku Öznitelikleri	46
6.5. Mamografik Öznitelikler	50
7. MAMOGRAFİ İMGELERİNİN SINIFLANDIRILMASI	57
7.1. k -En Yakın Komşu Sınıflandırıcısı.....	57
7.2. Karar Ağacı Sınıflandırıcısı	57
7.3. Rastgele Orman Sınıflandırıcısı.....	58
7.4. Naïve Bayes Sınıflandırıcısı	59
7.5. Lojistik Doğrusal Sınıflandırıcı	60
7.6. Doğrusal Ayırtaç Sınıflandırıcısı	61
7.7. Fisher Doğrusal Ayırtaç Analizi Sınıflandırıcısı	61
7.8. Destek Vektör Makineleri Sınıflandırıcısı	63
8. MATERYAL VE YÖNTEM	65
8.1. IRMA Veritabanı Kullanılarak Gerçekleştirilen Çalışmalar	66
8.1.1. IRMA veritabanı.....	66
8.1.2. Mamografi imgelerinin ön işlenmesi.....	66
8.1.3. Öznitelik vektörlerinin oluşturulması.....	67
8.1.4. Meme kanseri teşhisi	68
8.1.4.1. <u>3-sınıflı meme kanseri teşhis çalışması</u>	68
8.1.4.2. <u>12-sınıflı meme kanseri teşhis çalışması</u>	69
8.2. MIAS Veritabanı Kullanılarak Gerçekleştirilen Çalışmalar.....	71
8.2.1. MIAS veritabanı	71
8.2.2. Mamografi imgelerinin ön işlenmesi.....	71
8.2.3. Mamografi imgelerinde meme doku tipi tespiti	78

İÇİNDEKİLER DİZİNİ (devam)

	<u>Sayfa</u>
8.2.4. Mamografi imgelerinde ROI tespiti	80
8.2.5. Öznitelik vektörlerinin oluşturulması	82
8.2.6. Meme kanseri teşhisi	82
9. BULGULAR VE TARTIŞMA	84
10. SONUÇ VE ÖNERİLER	110
KAYNAKLAR DİZİNİ	114
ÖZGEÇMİŞ	122

ŞEKİLLER DİZİNİ

<u>Sekil</u>	<u>Sayfa</u>
1. 1. Tez çalışmasının iş akış şeması	4
3. 1. <i>The Area Normalized Error</i> - Yanlış-pozitif tanımı	32
3. 2. <i>The Area Normalized Error</i> - Yanlış-negatif tanımı	33
5. 1. Kontur gelişimi	38
5. 2. Chan-Vese algoritması çalışma prensibi	39
6. 1. LCP algoritması iş akış şeması	40
6. 2. Gri seviye bir imgenin LBP gösterimi	41
6. 3. GLCM gösterimi	47
6. 4. GLCM oluşumu	47
6. 5. Çeşitli yoğunluklara sahip ROI bölgeleri	51
6. 6. Çeşitli kontur yapısına sahip ROI bölgeleri	51
6. 7. Çeşitli şekil yapısına sahip ROI bölgeleri	51
6. 8. Mamografik öznetelik çıkarımında kullanılan imge örnekleri	52
6. 9. Örnek bir mamografi imgesi üzerinde tüm meme ve ROI alan sınırları	52
6. 10. Yoğunluk özneteliği hesabı için elde edilen bölütleme sonuçları:	53
6. 11. Yağlılık derecesi özneteliği hesabı için gerçekleştirilen bölütleme işlemi	54
6. 12. Örnek bir mamografi imgesinin majör, minör ve yarıçap eksenleri	55
6. 13. Örnek bir ROI imgesini çevreleyen konveks alan	56
7. 1. <i>k</i> -NN sınıflandırıcısı çalışma prensibi	57
7. 2. Karar ağacı sınıflandırıcısı çalışma prensibi	58
7. 3. LDC sınıflandırıcısı çalışma prensibi	62
7. 4. FLDA sınıflandırıcısı çalışma prensibi	63
7. 5. SVM sınıflandırıcısı çalışma prensibi	64
8. 1. IRMA veritabanına ait örnek imgeler	66
8. 2. IRMA veritabanına ait ön işleme aşaması adımları	67
8. 3. 3-sınıflı meme kanseri sınıflandırma sürecinin iş akış şeması	69
8. 4. 1-aşamalı meme kanseri sınıflandırma sürecinin iş akış şeması	70

ŞEKİLLER DİZİNİ (devam)

<u>Sekil</u>	<u>Sayfa</u>
8. 5. IRMA veritabanı kullanılarak 2-aşamalı meme kanseri sınıflandırma sürecinin iş akış şeması	70
8. 6. IRMA veritabanı kullanılarak 3-aşamalı meme kanseri sınıflandırma sürecinin iş akış şeması	71
8. 7. MIAS veritabanına ait örnek imgeler	72
8. 8. Örnek mamografi imgesi	73
8. 9. Bölge büyütme yöntemi için kullanılan birinci algoritmanın iş akış şeması.....	74
8. 10. Bölge büyütme yöntemi için kullanılan ikinci algoritmanın iş akış şeması	75
8. 11. Tohum noktası seçimi ve algoritmaların benzerlik koşullarını tanımlamakta kullanılan eşik değerinin hesaplanması için ele alınan bölge	75
8. 12. Bölütleme işlemi sonucunda pektoral kasi kesilmiş bir mamografi imgesi örneği... 76	76
8. 13. 1-aşamalı meme doku tipi sınıflandırma sürecinin iş akış şeması.....	79
8. 14. 2-aşamalı meme doku tipi sınıflandırma çalışması, “önce yağlı – sonra diğerleri” sürecinin iş akış şeması	79
8. 15. 2-aşamalı meme doku tipi sınıflandırma çalışması, “önce glandüler – sonradığerleri” sürecinin iş akış şeması.....	80
8. 16. 2-aşamalı meme doku tipi sınıflandırma çalışması, “önce yoğun – sonradığerleri” sürecinin iş akış şeması	80
8. 17. Örnek bir mamografi imgesi için bir ve iki iterasyon sonucu bölütlenen olası ROI alanları	81
8. 18. Örnek bir mamografi imgesi için ROI tespiti aşamaları.....	82
8. 19. MIAS veritabanı kullanılarak gerçekleştirilen meme kanseri teşhis çalışmasının iş akış şeması	83
9. 1. 108-boyutlu LCP-tabanlı öznitelik vektörlerinin 3-sınıflı sınıflandırılma çalışmasından elde edilen sonuçlar	84
9. 2. LCP 1,4-1 kodlu 108-boyutlu öznitelik vektörlerinin FLDA sınıflandırıcısı ile 3-sınıflı sınıflandırılma çalışmasına ait performans değerlendirmesi	85
9. 3. Tüm öznitelik gruplarının FLDA sınıflandırıcısı ile 3-sınıflı sınıflandırılması çalışmasından elde edilen sonuçlar	86
9. 4. 12-sınıflı 1-aşamalı sınıflandırma çalışmasından elde edilen sonuçlar	87

ŞEKİLLER DİZİNİ (devam)

<u>Sekil</u>	<u>Sayfa</u>
9. 5. LLC sınıflandırıcısı ile 12-sınıflı 1-aşamalı sınıflandırma çalışmasına ait performans değerlendirmesi	87
9. 6. 12-sınıflı 2-aşamalı sınıflandırma çalışmasından elde edilen sonuçlar	88
9. 7. FLDA sınıflandırıcısı ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait performans değerlendirmesi	89
9. 8. Çoğunluk oylama tekniği ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait performans değerlendirmesi	91
9. 9. 12-sınıflı 3-aşamalı sınıflandırma çalışmasından elde edilen sonuçlar	91
9. 10. FLDA sınıflandırıcısı ile 12-sınıflı 3-aşamalı sınıflandırma çalışmasına ait performans değerlendirmesi	92
9. 11. Çoğunluk oylama tekniği ile 12-sınıflı 3-aşamalı sınıflandırma çalışmasına ait performans değerlendirmesi	93
9. 12. Gürültü giderimi sonucu elde edilen mamografi imgesi örnekleri	94
9. 13. Gürültü ve yapaylık giderimi sağlanmış ve arka plandan arındırılmış mamografi imgesi örnekleri	95
9. 14. Bölütleme işlemi sonucunda pektoral kas bulundurmadığına karar verilen mamografi imgesi örnekleri	95
9. 15. Bölge büyütme yöntemi sonucu elde edilen mamografi imgeleri	97
9. 16. Örnek mamografi imgelerinde pektoral kas bölütleme ve iyileştirme adımları	100
9. 17. Birinci öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları	103
9. 18. İkinci öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları	104
9. 19. Üçüncü öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları	104
9. 20. Dördüncü öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları	105
9. 21. Beşinci öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları	105
9. 22. Altıncı öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları	106

ŞEKİLLER DİZİNİ (devam)

<u>Şekil</u>	<u>Sayfa</u>
9. 23. Yedinci öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları	106
9. 24. Yedinci öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarının performans değerlendirmesi	107
9. 25. MIAS veritabanı meme kanseri teşhisi çalışmasından elde edilen sonuçlar	108
9. 26. LLC sınıflandırıcısı ile MIAS veritabanı meme kanseri teşhisi çalışmasının performans değerlendirmesi	108

ÇİZELGELER DİZİNİ

<u>Çizelge</u>	<u>Sayfa</u>
3. 1. Bölge büyütme algoritmasının çalışma prensibi	31
3. 2. <i>The Area Normalized Error</i> – Değerlendirme Kriterleri	33
6. 1. İstatistiksel öznitelikler ve matematiksel ifadeleri	45
6. 2. Doku öznitelikleri ve matematiksel ifadeleri	48
8. 1. Değerlendirme kriterleri ve matematiksel ifadeleri	65
8. 2. LCP dönüşümü ile öznitelik vektörleri oluşturulurken alt bantların çarpıldığı katsayılar ve onlara ait LCP kodu	67
8. 3. 108-boyutlu öznitelik vektörlerinin oluşturulması	68
9. 1. LCP 1,4-1 kodlu LCP vektörleri tabanlı 108-boyutlu öznitelik vektörlerinin FLDA sınıflandırıcısı ile 3-sınıflı sınıflandırılma çalışmasına ait toplam karmaşıklık matrisi 85	
9. 2. LLC sınıflandırıcısı ile 12-sınıflı 1-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi	88
9. 3. FLDA sınıflandırıcısı ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi	89
9. 4. LLC sınıflandırıcısı ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi	90
9. 5. LDC sınıflandırıcısı ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi	90
9. 6. Çoğunluk oylama tekniği ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi	91
9. 7. FLDA sınıflandırıcısı ile 12-sınıflı 3-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi	92
9. 8. Çoğunluk oylama tekniği ile 12-sınıflı 3-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi	93
9. 9. Gürültü giderimi için kullanılan filtrelerin parametre ve başarı oranları	95
9. 10. Bölge büyütme yöntemi sonucu elde edilen başarı oranları	96
9. 11. Histogram eşitleme akabinde bölge büyütme yöntemi sonucu elde edilen başarı oranları	98
9. 12. Bölge büyütme yöntemi ardından T_I eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları	98

ÇİZELGELER DİZİNİ (devam)

<u>Çizelge</u>	<u>Sayfa</u>
9. 13. Histogram eşitleme akabinde bölge büyütme yöntemi uygulanması ardından T_1 eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları	99
9. 14. Bölge büyütme yöntemi ardından T_2 eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları	99
9. 15. Histogram eşitleme akabinde bölge büyütme yöntemi uygulanması ardından T_2 eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları	99
9. 16. Mamografi imgelerinde pektoral kas tespiti değerlendirme sonuçları	102
9. 17. LLC sınıflandırıcısı ile MIAS veritabanı meme kanseri teşhisi çalışmasına ait karmaşıklık matrisi.....	109

SİMGELER VE KISALTMALAR DİZİNİ

<u>Simgeler</u>	<u>Açıklama</u>
AA	Birinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Alçak/Alçak alt bandı
AY	Birinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Alçak/Yüksek alt bandı
AAAA	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Alçak/Alçak/Alçak/Alçak alt bandı
AAAY	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Alçak/Alçak/Alçak/Yüksek alt bandı
AAYA	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Alçak/Alçak/Yüksek/Alçak alt bandı
AAYY	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Alçak/Alçak/Yüksek/Yüksek alt bandı
AYAA	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Alçak/Yüksek/Alçak/Alçak alt bandı
AYAY	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Alçak/Yükse/Alçak/Yüksek alt bandı
AYYA	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Alçak/Yüksek/Yüksek/Alçak alt bandı

SİMGELER VE KISALTMALAR DİZİNİ (devam)

<u>Simgeler</u>	<u>Açıklama</u>
AYYY	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Alçak/Yüksek/Yüksek/Yüksek alt bandı
LCP 1-1	{1 - 1 - 1 - 0} katsayıları ile ağırlıklandırma sonucu elde edilen LCP öznelik vektörleri
LCP 1,4-1	{1,4 - 1 - 1 - 0} katsayıları ile ağırlıklandırma sonucu elde edilen LCP öznelik vektörleri
LCP 1,8-1	{1,8 - 1 - 1 - 0} katsayıları ile ağırlıklandırma sonucu elde edilen LCP öznelik vektörleri
YA	Birinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Yüksek/Alçak alt bandı
YY	Birinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Yüksek/Yüksek alt bandı
YAAA	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Yüksek/Alçak/Alçak/Alçak alt bandı
YAAY	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Yüksek/Alçak/Alçak/Yüksek alt bandı
YAYA	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Yüksek/Alçak/Yüksek/Alçak alt bandı

SİMGELER VE KISALTMALAR DİZİNİ (devam)

<u>Simgeler</u>	<u>Açıklama</u>
YAYY	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Yüksek/Alçak/Yüksek/Yüksek alt bandı
YYAA	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Yüksek/Yüksek/Alçak/Alçak alt bandı
YYAY	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Yüksek/Yüksek/Alçak/Yüksek alt bandı
YYYA	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Yüksek/Yüksek/Yüksek/Alçak alt bandı
YYYY	İkinci seviye 2-boyutlu ayrık dalgacık dönüşümünden elde edilen Yüksek /Yüksek /Yüksek /Yüksek alt bandı
d	GLCM matrisi oluşturulurken kullanılan uzaklık parametresi
P	LCP algoritmasında kullanılan dairesel komşu sayısı
R	LCP algoritmasında kullanılan dairesel komşuluğun yarıçapı
R_{sim}	NLM algoritmasında kullanılan komşuluk penceresi boyutu
R_{win}	NLM algoritmasında kullanılan arama penceresi boyutu
S_R	Pektoral kas tespiti için kullanılan başlangıç tohum bölgesi

SİMGELER VE KISALTMALAR DİZİNİ (devam)

<u>Simgeler</u>	<u>Açıklama</u>
S_R	Pektoral kas tespiti iyileştirmesinde kullanılan başlangıç tohum bölgesi
T_1	Pektoral kas tespiti iyileştirmesinin birinci aşamasında kullanılan eşik değeri
T_2	Pektoral kas tespiti iyileştirmesinin birinci aşamasında kullanılan eşik değeri
Q	Bölge büyütme algoritmasında kullanılan benzerlik ölçütü
σ	NLM algoritmasında kullanılan gürültü seviyesi parametresi
θ	GLCM matrisi oluşturulurken kullanılan yön parametresi
<u>Kısaltmalar</u>	<u>Açıklama</u>
ACC	Accuracy (Doğruluk)
AM	Amplitude Modulation (Genlik Modülasyonu)
AUC	Area Under Curve (Eğri Altında Kalan Alan)
BI-RADS	Breast Imaging-Reporting and Data System (Meme Görüntüleme-Raporlama ve Bilgi Sistemi)

SİMGELER VE KISALTMALAR DİZİNİ (devam)

<u>Kısaltmalar</u>	<u>Açıklama</u>
CAD	Computer Aided Diagnosis (Bilgisayar Destekli Teşhis)
CADe	Computer Aided Detection (Bilgisayar Destekli Tespit)
CC	Cranio-Caudal (Kranio-Kaudal)
CCA	Connected Component Analysis (Bağlı Bileşen Analizi)
CLAHE	Contrast-Limited Adaptive Histogram Equalization (Sınırlı-Kontrast Adaptif Histogram Eşitleme)
CWM	Center-Weighted Median Filter (Merkez-Ağırlıklı Medyan Filtresi)
DDSM	Digital Database for Screening Mammography (Mamografi Görüntüleme Dijital Veritabanı)
DWT	Discrete Wavelet Transform (Ayrık Dalgacık Dönüşümü)
ELM	Extreme Learning Machine (Aşırı Öğrenen Makine)
FCRN	Fully Complex-Valued Relaxation Neural Network (Tam Kompleks-Değerli Gevşetme Sinir Ağı)
FDR	False Discovery Rate (Yanlış Keşif Oranı)

SİMGELER VE KISALTMALAR DİZİNİ (devam)

<u>Kısaltmalar</u>	<u>Açıklama</u>
FFDM	Full Field Digital Mammography (Tüm Alan Dijital Mamografi)
FLDA	Fisher's Linear Discriminant Analysis (Fisher Doğrusal Ayırtaç Analizi)
FM	Frequency Modulation (Frekans Modülasyonu)
FN	False-Negative (Yanlış-Negatif)
FN _m	Mean False-Negativity (Ortalama Yanlış-Negatiflik)
FNR	False Negative Rate (Yanlış Negatif Oranı)
FOR	False Omission Rate (Yanlış İhmal Oranı)
FP	False-Positive (Yanlış-Pozitif)
FP _m	Mean False-Positivity (Ortalama Yanlış-Pozitiflik)
FPR	False Positive Rate (Yanlış Pozitif Oranı)
GAM	Generalized Additive Model (Genelleştirilmiş Toplamsal Model)
GBVS	Grafik-Tabanlı Görsel Çıkıntı (Graph-Based Visual Saliency)

SİMGELER VE KISALTMALAR DİZİNİ (devam)

<u>Kısaltmalar</u>	<u>Açıklama</u>
GLCM	Gray-Level Co-Occurence Matrix (Gri-Seviye Eş-Oluşum Matrisi)
GMM	Gaussian Mixture Model (Gauss Karışım Modeli)
<i>k</i> -NN	<i>k</i> -Nearest Neighbor Classifier (<i>k</i> -EnYakın Komşu Sınıflandırıcısı)
LBG	Linde Buzo-Gray
LBP	Local Binary Pattern (Yerel İkili Örüntü)
LCP	Local Configuration Pattern (Yerel Konfigurasyon Örüntüsü)
LDA	Linear Discriminant Analysis (Doğrusal Ayırtaç Analizi)
LDC	Linear Discriminant Analysis (Doğrusal Ayırtaç Sınıflandırıcısı)
LLC	Logistic Linear Classifier (Lojistik Regresyon Sınıflandırıcısı)
LRM	Local Range Modification (Yerel Aralık Modifikasyonu)
mRMR	Minimum-Redundancy Maximum-Relevance (Minimum-Fazlalık Maksimum-İlişki)

SİMGELER VE KISALTMALAR DİZİNİ (devam)

<u>Kısaltmalar</u>	<u>Açıklama</u>
MIAS	Mammographic Image Analysis Society (Mamografik İmge Analizi Topluluğu)
MLO	Mediolateral Oblique (Mediolateral Oblik)
MLP	Multi-Layer Perceptron (Çok-Katmanlı Yapay Sinir Ağları)
NLM	Non-Local Means Filter (Yerel-Olmayan Ortalama Filtresi)
NPV	Negative Predictive Value (Negatif Tahmin Değeri)
PCA	Principal Component Analysis (Temel Bileşen Analizi)
PPV	Positive Predictive Value (Pozitif Tahmin Değeri)
RBFNN	Radial Based Function Neural Network (Radyal Tabanlı Fonksiyon Sinir Ağları)
ROC	Receiver Operating Characteristic (İşlem Karakteristik Eğrisi)
ROI	Region of Interest (İlgi Alanı)
SCNN	Soft Clustering Neural Network (Yumuşak Kümeleme Sinir Ağı)
SFM	Screening-Film Mammography (Tarama-Film Mamografi)
SNS	Sensitivity (Hassasiyet)

SİMGELER VE KISALTMALAR DİZİNİ (devam)**Kısaltmalar****Açıklama**

SPC

Specificity (Kesinlik)

SVM

Support Vector Machines (Destek Vektör Makineleri)

1. GİRİŞ VE AMAÇ

Meme kanserinin kadın ölümlerinin başlıca nedenlerinden olmasına rağmen erken teşhis ile meme kanseri nedenli ölüm oranı azaltılabilmektedir (Jemal vd., 2011). Meme kanseri tarama yöntemleri arasında kendi kendine muayene, klinik meme muayenesi, mamografi, ultrason ile görüntüleme, manyetik rezonans ile görüntüleme ve doku aspirasyonu gelmektedir.

Bu yöntemlerden tez çalışması kapsamında incelenecek olan mamografi, memenin sıkıştırılarak, düşük doz radyasyon uygulanmasıyla Kranio-Kaudal (Cranio-Caudal – CC) ve Mediolateral Oblik (Mediolateral Oblique – MLO) graflerinin elde edilmesi ile meme kanserinin teşhisinde kullanılan görüntüleme tekniklerinden biridir. CC mamografi, meme ucunun hemen arkası ve çevresini görüntülemek amacı ile memenin aşağıdan ve yukarıdan sıkıştırılması sonucunda çekilmektedir. MLO mamografi ise tüm meme bölgesini görüntülemek amacı ile memenin yanlardan sıkıştırılması sonucunda çekilmektedir. Çekim şekli nedeniyle, MLO mamografilerde yalnızca meme bölgesi değil pektoral kaslarda görüntülenmektedir. Mamografi, meme muayenesinden sonra belirtileri saptanmış kişilere tanısal amaçlı yapıldığı gibi, hiçbir belirtisi bulunmayan kişilere tarama amaçlı da uygulanmaktadır.

Mamografi imgelerinde küçük parlak noktalar olarak görülen mikrokalsifikasyonlar, belirgin olarak görülen kitleler, iki meme arasında görülen asimetri, meme dokusunda görülen yapısal bozulmalar, cilt ve trabekülasyonda izlenen değişiklikler ve aksiller lenf nodu patolojileri uzmanlar tarafından dikkat edilen kanser belirtileridir (C. Çalışır, 2015, sözlü görüşme).

Kalsifikasyonların şekil, dağılım ve boyutlarına göre iyi ya da kötü huylu kanser öncüsü olduğuna karar verilmektedir. Literatürde çokça üzerinde durulan mikrokalsifikasyonlar 1 mm' den daha küçük radyoopak (yağdan fakir – beyaz) yapılardır. Mikrokalsifikasyonlar çoğunlukla kanserli olmamakla beraber kanser öncüsü olarak bilinmektedir. Düzenli dağılım gösteren, düzenli şekle sahip mikrokalsifikasyonlar iyi huylu, dağınık ve düzensiz şekilli mikrokalsifikasyonlar ise kötü huylu kanser hücresi

olarak düşünülmektedir (C. Çalışır, 2015, sözlü görüşme). Boyutları 1mm' den daha büyük kalsifikasyonlar ise kaba kalsifikasyon olarak tanımlanmakta ve iyi huylu kanser öncüsü olarak bilinmektedir. Gerek kaba kalsifikasyon gerekse mikrokalsifikasyon duktal sistemde gerçekleşmiş ise kötü huylu kanser öncüsü olarak nitelendirilmektedir (C. Çalışır, 2015, sözlü görüşme).

Bir diğer kanser belirtisi olan kitlelerin tespitinde ise yoğunluk, şekil, kontur, boyut, eşlik eden bulgular ve eski filmlere göre olan değişiklik göz önünde bulundurulmaktadır. Kitle ne kadar çok radyolüsen (yağdan zengin – siyah) ise o kadar iyi huylu, ne kadar radyopak ise o kadar kötü huylu olarak tanımlanmaktadır. Kitlenin şekli düzensizleştikçe, boyutu arttıkça ya da konturları tam sınırlandırılmıyor ise kötü huylu olma olasılığı artmaktadır. Ancak ne kadar şekli düzensiz ya da konturları belirsiz olursa olsun, kitle yağ içeriyorsa, iyi huylu olarak düşünülmektedir (C. Çalışır, 2015, sözlü görüşme).

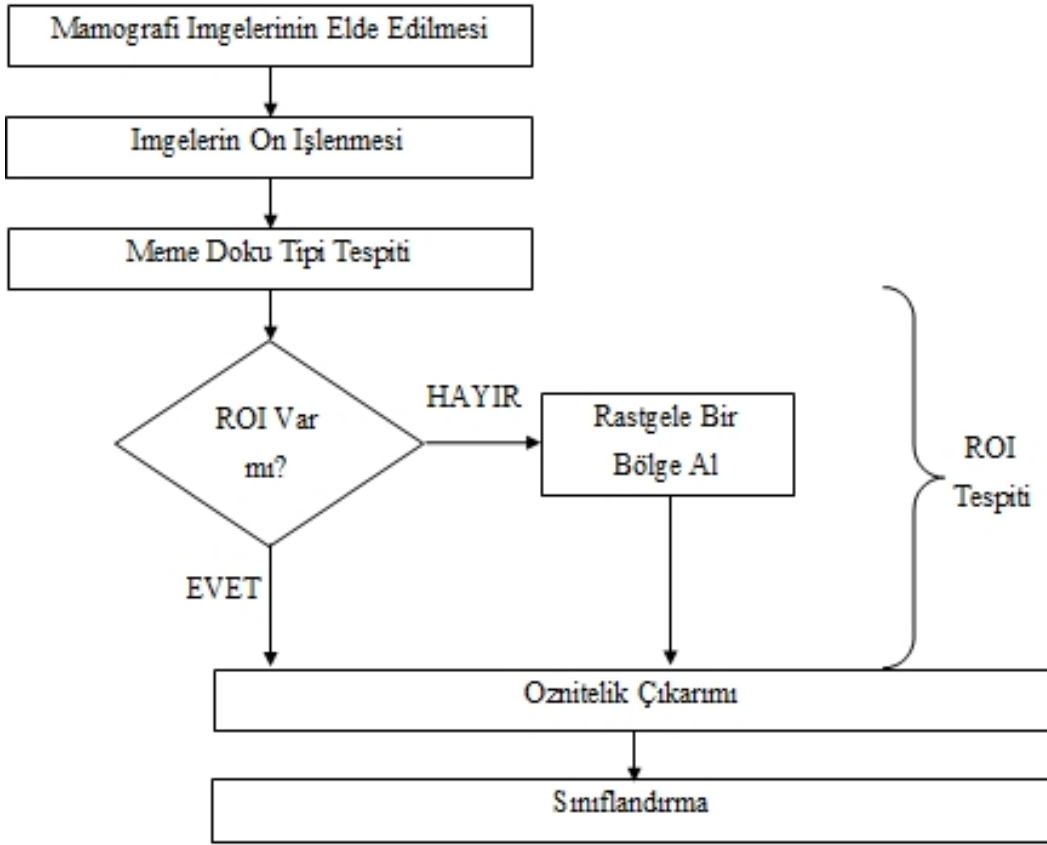
Bu belirtilerin uzmanlar tarafından ne kadar kolay tespit edilebileceği meme doku tipi ile doğrudan ilişkilidir. Meme Görüntüleme-Raporlama ve Bilgi Sistemi (Breast Imaging-Reporting and Data System – BI-RADS) kategorisine göre meme dokuları yağlı, yağlı-glandüler, heterojen yoğun ve çok-yoğun olarak sınıflandırılmaktadır (Vallez vd., 2013). Yağlı meme doku tiplerinde belirtiler rahatlıkla görülebilmekte iken yoğun doku tiplerinde belirtilerin tespit edilmesi zorlaşmaktadır (C. Çalışır, 2015, sözlü görüşme).

Bilgisayar Destekli Tespit (Computer Aided Detection – CADe) ve Bilgisayar Destekli Teşhis (Computer Aided Diagnosis– CAD) sistemlerinin kullanımı ile gözden kaçabilecek noktaların uzman doktor tarafından görüntülenebilmesi ve tekrar gözden geçirilmesi sağlandığı için, mamografi imgelerinden kanser tespit hassasiyeti artırılabilir. CADe sistemler, mamografi imgelerinden meme kanseri için şüpheli bölgeleri (İlgi Alanı – Region of Interest – ROI) tespit ederken CAD sistemler tespit edilen ROI'lerin iyi ya da kötü huylu olduğuna karar vermektedirler. Bir CAD sistemi genel olarak dört aşamadan oluşmaktadır. Bu aşamalardan ilki imgelerin gürültü giderme, kontrast iyileştirme ya da yapısal iyileştirme işlemlerinin yapıldığı ön işleme aşamasıdır. İkinci aşamada, iyileştirilmiş imgelerden kitle içerme ihtimali olan ROI'ler tespit edilmektedir. Üçüncü aşamada, bu bölgelerden sınıflandırmada kullanılmak üzere öznitelikler çıkartılmakta ve gerekirse öznitelik seçimi yapılmaktadır. Son olarak,

çıkartılan öznelikler kullanılarak ROI'ler normal, iyi ya da kötü huylu kitle olarak sınıflandırılmaktadır. Bu sistemlerde, ROI'ler mikrokalsifikasyonlar, kitleler, yapısal bozukluklar ya da iki meme arasındaki asimetrik bölgeler olarak ele alınabilmektedir. Ayrıca kanser tespiti için meme bölgesinden belli kesitlere bakmak yerine tüm meme bölgesi de incelenebilmektedir (Zhang vd., 2008)

Bu tez çalışması kapsamında, mamografi imgelerinden meme doku tipi tespiti ve meme kanseri tespit ve teşhisi yapabilen bir CAD sistemi önerilmiştir. Önerilen CAD sistemi iki aşamadan oluşmaktadır. Sistemin ilk aşamasında mamografi imgeleri gürültü giderme, kontrast iyileştirme işlemlerine tabi tutularak ön işlenmiş ve meme doku tipi tespiti yapılmıştır. Sistemin ikinci aşamasında mamografi imgelerinin kanser tespit ve teşhisi gerçekleştirilmiştir. Bu amaçla ilk olarak, meme doku tipi belirlenmiş mamografi imgelerinden meme kanseri açısından şüpheli bölgeler çıkarılarak tespit çalışması gerçekleştirilmiştir. Teşhis çalışması amacı ile tespit edilmiş ROI'lerden sınıflandırmada kullanılacak öznelik çıkarımı ve çıkartılan öznelikler kullanılarak imgelerin normal, iyi huylu veya kötü huylu kanserli olarak sınıflandırılması gerçekleştirilmiştir. Böylece tüm meme doku tiplerine yönelik bir CAD sistemi elde edilmiştir. Önerilen sistemin iş akış şeması Şekil 1.1'de verilmiştir.

Buraya kadar tez çalışmasının güdüsü ve ana konusu hakkında bir giriş bilgisine yer verilmiştir. Tez çalışmasının ikinci bölümünde tez konusu ile ilgili literatür araştırması verilmiştir. Bölüm 3'de mamografi imgelerinin ön işlenmesi adımı hakkında genel bilgi verilmiş ve tez çalışmasında bu adım için kullanılan yöntemler anlatılmıştır. Bölüm 4'te meme doku tipi tespiti adımı ile ilgili genel bilgi ve bu adımın önemine yer verilmiştir. Bölüm 5'te mamografi imgelerinden ROI tespiti adımıyla kullanılan yöntem hakkında teorik bilgi verilirken Bölüm 6'da tez çalışmasının çeşitli aşamalarında mamografi imgelerinden çıkartılan öznelikler açıklanmıştır. Bölüm 7'de ise tez çalışmasında mamografi imgelerinin sınıflandırılmasında kullanılan yöntemler hakkında teorik bilgiler verilmiştir. Tez çalışması süresince gerçekleştirilen çalışmalar Bölüm 8'de ayrıntılı olarak anlatılmış ve Bölüm 9'da bu çalışmaların sonuçları verilerek tartışılmıştır. Son olarak Bölüm 10'da tez çalışması hakkında genel bir değerlendirme yapılmış ve ilerleyen dönemlerde bu konuda çalışmak isteyen araştırmacılara önerilerde bulunulmuştur.



Şekil 1. 1. Tez çalışmasının iş akış şeması

2. LİTERATÜR ARAŞTIRMASI

Tez çalışması kapsamında, mamografi imgelerinin işlenmesi ve kanser teşhisi konularında akademik anlamda farklı yöntem ve süreçler kullanılarak yapılmış çalışmalar incelenmiştir. Çalışmalarda genellikle Mamografi Görüntüleme Dijital Veritabanı (Digital Database for Screening Mammography – DDSM) ve Mammographic Image Analysis Society (MIAS) tarafından hazırlanan MIAS veritabanı kullanılmıştır. Bu çalışmalar aşağıda altı alt başlık halinde özetlenmiştir:

2.1. Mikrokalsifikasyon Tespiti Yapılan Çalışmalar

Pal vd. (2008), mamografi imgelerinde mikrokalsifikasyon tespiti için öznelik seçimi yapılan çok-aşamalı bir sistem geliştirmişlerdir. Sistem, MIAS veritabanından alınmış 3 iyi huylu kanserli (benign), 4 kötü huylu kanserli (malignant) ve 1 normal olmak üzere 8 mamografi imgesi ile eğitilmiş, yine MIAS veritabanından alınan 10 abnormal ve 7 normal olmak üzere 17 mamografi imgesi ile test edilmiştir. Test imgelerinden manüel olarak alınmış 3609 kalsifik ve 100000 normal noktanın her biri için 3×3, 5×5, 7×7 ve 9×9'luk pencereler kullanılarak 87 öznelik hesaplanmıştır. Öznelik sayısı Çok-Katmanlı Yapay Sinir Ağları (Multi-Layer Perceptron– MLP) ile her bir nokta için 29'a düşürülmüştür. Daha sonra bu öznelikler, giriş katmanı nöron sayısı 29, gizli katman nöron sayısı 15 olan geriye yayımlı MLP sinir ağları ile sınıflandırılmıştır. Sınıflandırma sonucu elde edilen yanlış-pozitiflik oranını azaltmak amacı ile önce Bağlı Bileşen Analizi (Connected Component Analysis– CCA) ve akabinde elonge bileşen çıkarma yapılmıştır. Geriye kalan mikrokalsifikasyon içeren bölgelerin yerel yoğunluk değerleri hesaplanarak bir eşik değerinden geçirilmiştir. Yoğunluğu fazla olan imgeler kalsifikasyonlu, az olan imgeler ise normal olarak sınıflandırılmıştır.

Lado vd. (2008), yaptıkları çalışmada dijital mamogramlarda mikrokalsifikasyon kümelerini tespit eden bir CAD sistemindeki yanlış-pozitiflik oranını azaltmak için doku unsurlarının etkileşimlerini içeren genişletilmiş bir Genelleştirilmiş Toplamsal Model (Generalized Additive Model – GAM) yöntemi sunmuşlardır. GAM yöntemi, tanımlanan ROI'lerin farklı sürekli unsurlarının doğrusal olmayan etkilerini ölçerek CAD

sistemlerinde normal meme dokusundan lezyonları ayıran istatistiksel bir yaklaşımdır. Bununla beraber, unsurların bu doğrusal olmayan etkileri meme doku tipi gibi bazı faktörlere göre değişkenlik gösterebildiği belirtilmiştir. Çalışmada, GAM yöntemi, meme doku tipi faktör olarak alınarak genişletilmiştir. Meme dokusu, yağlı ve yoğun olmak üzere iki kategoriye ayrılmıştır. Çalışmada İspanya Galicia Community’de bulunan mamografik görüntüleme programından rastgele seçilen 50-64 yaş aralığındaki kadınlara ait 174 mamografi imgesi kullanılmıştır. Meme sınırları komşu piksellerin gri seviye değerleri temel alınarak belirlenmiştir. Çıkarılan meme bölgelerine dalgacık dönüşümü (wavelet transform) uygulanarak olası mikrokalsifikasyon noktaları tespit edilmiş ve gri seviye eşikleme yöntemi ile yanlış tanımların elenmesi amaçlanmıştır. Son olarak yanlış pozitifleri azaltmak için GAM yöntemi uygulanmıştır. Sonuç olarak doku tipi faktörü bulunmayan GAM yönteminde %83,12 hassasiyet ve imge başına 1,46 yanlış-pozitiflik oranı bulunurken, doku tipi faktörünü de içeren genişletilmiş GAM yönteminde hassasiyetin %85,71’e yükseldiğini ve her bir imge için yanlış-pozitiflik oranınının 0,74’e düştüğü belirtilmiştir.

Papadopoulos vd. (2008), mamogramlarda mikrokalsifikasyon tespiti için beş farklı görüntü iyileştirme yöntemini karşılaştırmışlardır. Bu yöntemler; Sınırlı-Kontrast Adaptif Histogram Eşitleme (Contrast-Limited Adaptive Histogram Equalization– CLAHE), Yerel Aralık Modifikasyonu (Local Range Modification– LRM), Fazlalıklı Dalgacık Dönüşümü (Redundant Discrete Wavelet– 2D-RDW) ve dalgacık doğrusal germe (wavelet linear stretching) yöntemleridir. CLAHE yöntemi medikal imgelerin kontrast iyileştirmesinde kullanılan bir yöntemdir. Bu yöntemde, imge içeriksel bölgelere ayrılarak her bir bölgeye histogram eşitleme uygulanmaktadır. LRM, $y = ax + b$ formülünü izleyen lineer bir algoritmadır. Burada y , iyileştirilen resim, x , orijinal gri seviye imge a ve b ise bölgesel kontrasta bağlı parametrelerdir. Bu parametreler imge bloklarını örtüştüren bir enterpolasyon prosedürü ile hesaplanmaktadır. RDW yönteminde, 2-D RDW algoritması kullanılarak imge alt-imgelere ayrıştırılmakta ve temel olarak 2. ve 3. seviyeler kullanılarak yeniden yapılandırılmaktadır. Yöntemde dalgacık (wavelet) katsayılarına iyileştirme yapılmaktadır. Çalışmada kullanılan yöntemlerden CLAHE ve LRM geleneksel görüntü analizi metodolojileridir. Diğerleri ise biortogonal dalgacık dönüşümüne dayanmaktadır. İyileştirilmiş imgelerden 54 tane parlaklık, şekil ve doku öznitelikleri çıkarılmıştır. Öznitelik seçim algoritması ile bu özniteliklerden 22 tanesi seçilmiştir.

Bunlara da Temel Bileşen Analizi (Principal Component Analysis– PCA) uygulanarak boyut 9'a indirilmiştir. Sınıflandırıcı olarak iki gizli katmanlı ileri beslemeli geri yayımlı MLP kullanılmıştır. Bu çalışmada CAD sistemi, görüntü iyileştirme yöntemlerinin mamografi imgelerinde mikrokalsifikasyon tespitine etkilerini doğrulamak için kullanılmıştır. Her yöntemin etkisi, MIAS ve Nijmegen veritabanları üzerinde CAD sisteminin performansının İşlem Karakteristik Eğrisi (Receiver Operating Characteristic – ROC) analizi ile belirlenmesi ile saptanmıştır. ROC eğrisinin altında kalan alan (area under curve – AUC) arttıkça görüntü iyileştirmenin de arttığını belirtmişlerdir. Sonuç olarak LRM (AUC=0,932) ve WSLT (AUC=0,926) yöntemlerinin en iyi performansı gösterdiklerini tespit edilmiştir.

Malar vd. (2012), yaptıkları çalışmada Aşırı Öğrenen Makine (Extreme Learning Machine – ELM) kullanarak dijital mamogramlarda mikrokalsifikasyon tespiti için dalgacık tabanlı doku analizinin verimliliğini incelemişlerdir. Çalışmada MIAS veritabanından alınan 30 normal (10 yağlı, 10 yağlı-glandüler, 10 yoğun-glandüler) ve 25 (13 iyi huylu kanserli, 12 kötü huylu kanserli) mikrokalsifikasyonlu mamografi imgesi kullanılmıştır. İlk olarak, gri seviye mamografi imgeleri global eşikleme kullanılarak ikili formata dönüştürülmüş ve ikili imgelerde etiketlenmiş objeler sonucu oluşan bağlı bileşenlerin alanları hesaplanmıştır. Sonrasında seçilen meme profili, orijinal imge ile çarpılarak sonuç imgesi elde edilmiştir. Son olarak da koyu arka plan, sıfır değerindeki sütunların kesilmesi ile çıkarılmıştır. Bu işlemlerden sonra 32×32'lik ROI'ler manüel olarak elde edilmiştir. Toplamda 55 mamografi imgesinden elde edilen 120 ROI'den 80 tanesini (40 mikrokalsifikasyonlu, 40 normal) eğitimde, 40 tanesini de (20 mikrokalsifikasyonlu, 20 normal) testte kullanılmıştır. ROI'lerin kontrastı gri seviye dilimleme yöntemi ile iyileştirilmiştir. ROI'lerden dalgacık, eş-oluşum matrisi ve Gabor filtresi tabanlı öznelikler seçilmiştir. Dalgacık tabanlı öznelik çıkarımı için ROI'ler Haar dalgacığı ile 4. seviyede ayrıştırılmış, enerji ve ∞ -norm öznelikleri çıkartılmıştır. Eş-oluşum matrisi tabanlı öznelik çıkarımı için $d=2$ uzaklık ve $\theta = \{45^\circ, 90^\circ\}$ yön parametreleri kullanılarak oluşturulmuş matrislerden otokorelasyon, kontrast, korelasyon, asimetri (cluster prominence), çarpıklık (cluster shade), benzersizlik, enerji, entropi, homojenlik, maksimum olasılık, kareler toplamı değişimi, toplam ortalama, toplam değişim, toplam entropi, fark değişimi, fark entropisi, ilişkinin ölçüm bilgisi, ters fark, normalize edilmiş ters fark ve ters fark momenti öznelikleri çıkartılmıştır. Gabor filtresi

tabanlı öznelik çıkarımı için ROI'ler (2, 4, 32, 64) ölçekleri ve (30°, 60°, 90°, 120°) yönlerinde filtrelenmiş ve enerji özneliği çıkartılmıştır. Sonrasında ELM kullanılarak sınıflandırma yapılmıştır. ELM yapısının tek-katmanlı ileri-beslemeli ağlar (single-layer feed-forward network) çatısı altında bulunduğu ve klasik ileri beslemeli ağlara göre oldukça fazla daha hızlı öğrenme kapasitesine sahip olduğu belirtilmiştir. Yapılan çalışmada, ELM'nin performansını, Bayes sinir ağları, Naïve Bayes ve Destek Vektör Makineleri (Support Vector Machines – SVM) gibi diğer bilinen sınıflandırıcılarla da karşılaştırmışlardır. Karşılaştırma kriteri olarak; hassasiyet, doğru pozitif oranı, yanlış pozitif oranı, F-ölçüsü, duyarlılık ve AUC kullanılmıştır. Sonuç olarak ELM yönteminin diğer yöntemlere üstünlük sağladığını tespit etmişlerdir. ELM yönteminin en iyi hassasiyetinin dalgacık tabanlı öznelikler kullanıldığında %94 olduğu belirtilmiştir.

Mamografi imgelerinde mikrokalsifikasyon tespiti, imgedeki normal hücrelerin mikrokalsifikasyonlu hücrelerden daha fazla sayıda olmasından dolayı, dengesiz bir sınıflandırma problemidir. Bu durumda normal hücreler baskın geldiği için başarı oranı azalmaktadır. Bria vd. (2014), bu problemi çözebilmek için beş sınıflandırıcı kaskad bir sistem geliştirmişler ve Tüm Alan Dijital Mamografi (Full Field Digital Mammography – FFDM) veritabanından alınmış 252 mamografi imgesi ile eğittikleri sistemi 1599 mamografi imgesi ile test etmişlerdir. Sistemin ilk aşamasında kuantum gürültüsü içeren mamografi imgeleri düzgün gürültü seviyesi ile yeniden ölçeklenerek imgelerde gürültü eşitlemesi yapılmıştır. Ön işlenmiş mamografi imgelerinin her bir pikseli mikrokalsifikasyon tespiti için sınıflandırma işlemine tabi tutulmuştur. Bunun için mamografi imgelerinden sınıflandırılacak piksel merkezde olacak şekilde 12×12'lik pencereler oluşturulmuş ve bu pencerelerden her bir sınıflandırıcı ile sırasıyla 2, 3, 5, 12, 40 Haar-benzeri öznelikler çıkartılmıştır. Her bir sınıflandırıcı sonucunda mikrokalsifikasyon içermeyen pikseller atılmış ve son sınıflandırıcının sonunda olası mikrokalsifikasyon olarak tanımlanan piksellerin olasılık haritası çıkartılmıştır. Çalışmada, mikrokalsifikasyonlu hücrelerin bağlı bileşenler (connected components) olarak görüldüğü ve sabit pencere boyutu kullanıldığından makrokalsifikasyonların değişik boyutlardaki bileşenlere ayrılarak yanlış-pozitiflik oranını artırdığı belirtilmiştir. Bu nedenle, her bir bağlı bileşeni kapsayan minimum sınırlayıcı dikdörtgen (minimum-bounding rectangle) hesaplanarak 1mm'den büyük olanlar atılarak makrokalsifikasyonlar giderilmiştir. Geri kalan olası mikrokalsifikasyonlu hücrelerin olasılık haritasındaki ikinci en büyük değer bir

eşik değerinin üstünde ise mikrokalsifikasyon olarak kabul edilmiştir. Birbirine en fazla 10 mm uzaklıkta olan mikrokalsifikasyonlu hücreler aralarındaki uzaklığa göre ağırlıklandırılarak kümelenmiş ve 3 mikrokalsifikasyonlu hücreden az sayıda mikrokalsifikasyonlu hücre içeren kümeler sınıflandırmada kullanılmamıştır. Sınıflandırma aşamasında her bir kümeden 145 öznitelik çıkartılmıştır.

Guo vd. (2016), 3 aşamalı bir sistemle mamografi imgelerindeki mikrokalsifikasyonların bulunmasındaki başarımı artırmayı hedeflemişlerdir. İlk olarak bağlı bölge işaretleme yöntemiyle pektoral kasların giderimi sağlamış, gürültüyü giderek Contourlet dönüşümü yardımıyla gerekli öznitelikleri elde etmiş, son aşamada ise bağlantısız basitleştirilmiş darbe-bağımlı sinir ağı (pulse-coupled neural network) kullanarak sınıflandırmayı tamamlamışlardır. Test için veritabanı olarak MIAS veritabanını kullandıkları çalışmada $AUC=0.97$ olarak elde etmişlerdir.

2.2. Kitle (Anomali) Tespiti Yapılan Çalışmalar

Kekre vd. (2009), mamografi imgelerinde kanser tespiti için Linde Buzo-Gray (LBG) algoritmasını kullanmışlardır. LBG algoritması ile her bir imgeden öncelikle 128-boyutlu kod kitapları (codebooks) oluşturulmuş, sonrasında bu kod kitaplarından yine LBG algoritması kullanılarak 8-boyutlu yeni kod kitapları oluşturulmuştur. Kod vektörler (codevectors) ile bölütlenmiş mamografi imgelerinde bulunan her bir bölgenin alanı hesaplanmış ve en fazla alana sahip olan bölge kitle olarak kabul edilmiştir. Yöntem ile elde edilen sonuçlar Gri-Seviye Eş-Oluşum Matrisi (Gray-Level Co-Occurrence Matrix – GLCM) ve Watershed algoritmalarından elde edilen sonuçlar ile karşılaştırılmış, önerilen yöntemin daha verimli olduğu görülmüştür.

Kekre vd. (2010), bir başka çalışmalarında mamografi imgelerini Kekre's Median Fast Codebook Generation algoritmasını kullanarak bölütlemişlerdir. Bölütlenmiş imgelerde bulunan her bir bölgenin alanı hesaplanmıştır. En fazla alana sahip olan bölge kitle olarak kabul edilmiştir. Yöntemin başarı oranı %68,5 olarak belirtilmiştir.

Krishnan vd. (2010), yaptıkları çalışmada yüksek hassasiyet derecesine sahip meme kanseri tespiti için SVM tabanlı bir sınıflandırıcı tasarlamayı amaçlamışlardır. Çalışmada

Kaliforniya Machine Learning Repository'den alınan iki farklı veri seti ile kullanılmıştır. Birinci veri setinde 458 iyi huylu kanserli ve 241 kötü huylu kanserli olmak üzere 699 görüntü mevcuttur. Bu veri setinden küme kalınlığı, hücre boyutu bir-biçimliliği, hücre şekli bir-biçimliliği, sınırsal adhezyonlar, tekil epitelyal hücre boyutu, örtüsüz çekirdek, orta şiddette kromatin, normal çekirdekçik ve mitozlar olmak üzere dokuz öznelik çıkartılmıştır. İkinci veri setinde 357 iyi huylu kanserli ve 212 kötü huylu kanserli olmak üzere toplam 569 görüntü mevcuttur. Bu veri setinde ise yarıçap, doku, çap, alan, düzgünlük, sıklık, konkavlık, konkav noktalar, simetri ve fraktal ölçü olmak üzere 10 tane öznelik çıkartılmıştır. Sınıflandırma işleminden önce seçilen bu özneliklerin sınıfları ayırabilme kapasiteleri istatistiksel t-test kullanılarak ölçülmüştür. Elde edilen verilerin sınıflandırılmasında ise SVM kullanılmıştır. Çalışmanın ana amacı, mamografi imgelerinden çıkartılan özneliklerin en uygun eğitim sürecini belirlemektir. İlk olarak en uygun kernel fonksiyonu seçiminin ve sonrasında uygun eğitim-test ayırımının yapılması hedeflenmiştir. Her bir veri seti için polinomial kernel fonksiyonu ve Gaussian radyal tabanlı kernel fonksiyonunu kullanılmıştır. Birinci veri seti için $\sigma^2 = 1000$ olan radyal tabanlı kernel fonksiyonunun ikinci veri seti için ise $\sigma^2 = 10000$ olan Gaussian radyal tabanlı kernel fonksiyonunun en uygun olduğu tespit edilmiştir. Eğitim-test ayırımında ise, birinci veri seti için en iyi sonuç 400 veri büyüklüğündeki eğitim setinde, ikinci veri seti için ise 350 veri büyüklüğündeki eğitim setinde elde edilmiştir. Çalışmada eğitim ve test ayırımının %60 eğitim, %40 test olarak elde edildiği belirtilmiştir. Sonuç olarak kernel fonksiyonunun ve eğitim-test ayırımının belirlenmesinin SVM tabanlı sınıflandırmaların performanslarını arttırdığı tespit edilmiştir. Çalışmada hassasiyet oranlarının, birinci veri seti için %99,385, ikinci veri seti için ise %93,726 olarak elde edildiği belirtilmiştir.

Biswas ve Mukherjee (2011), mamografi imgelerinde anomali tespiti için iki-katmanlı bir yöntem önermişlerdir. Çalışmada mamografi imgelerinin “texton çantası” olarak karakterize edilebileceği belirtilmiştir. Bu amaçla, ilk katmanda mamogram ROI çok ölçekli bir filtre bankasından geçirilerek doku tanımlayıcılar elde edilmekte ve ikinci katmanda ise doku tanımlayıcılarından Gauss Karışım Modeli (Gaussian Mixture Model – GMM) ile textonlar elde edilmektedir. Elde edilen textonların verilen bir mamografi imgesindeki yoğunluğuna göre anomali olup olmadığına karar verilmektedir. Bu amaçla beklenti maksimizasyonu (expectation-maximization) algoritması kullanılmıştır. Yöntem

MIAS ve DDSM veritabanları kullanılarak yağlı, yağlı-glandüler ve yoğun-glandüler doku tiplerinde test edilmiş ve başarı oranı ortalama %82 olarak belirtilmiştir.

Haider vd. (2011), başlangıç evresindeki meme ve akciğer kanseri tespiti için bir yöntem geliştirmiş ve MIAS veritabanı ile yöntemi test etmişlerdir. Yöntemde ilk olarak imgelere tüm kenar bulma operatörleri uygulanmış ve Sobel operatörünün en uygun sonucu verdiği görülmüştür. Sonrasında aynı kişiye ait farklı açılardan çekilmiş mamografi imgeleri MATLAB 3D grafik programı ile üç boyutlu hale getirilmiş ve bölütlenmiştir. Son olarak, kitlenin boyutu bölütlenmiş imgedeki 2-boyutlu piksellerin ve kitle kalınlığının toplamı olarak hesaplanmıştır. Başarı oranı %100 olarak belirtilmiştir.

Ramos vd. (2012), DDSM veritabanından aldıkları 120 (60 normal, 60 kanserli) CC mamograma eş-oluşum matrisleri, dalgacık ve ridgelet dönüşümleri uygulayarak kitle tespitinde bu yöntemlerin verimliliğini incelemişlerdir. Ayrıca, çalışmada genetik algoritma kullanılarak en iyi öznitelik kümeleri de belirlenmiştir. Çıkarılan özniteliklerin sınıflandırılması için rastgele orman(random forest) algoritması kullanılmıştır. Mamografi imgelerinden çıkartılmış ROI'lerden merkezi anomali merkezi olan 61×61 boyutlarında alt-imgeler alınmış; öznitelik çıkarma ve sınıflandırma işlemleri bu alt-imgeler üzerinde yapılmıştır. Çalışmada, entropi, enerji, toplam ortalama, toplam varyans ve küme eğilimi olmak üzere beş doku tanımlayıcısı kullanılmıştır. GLCM matrisleri $\theta = 0^\circ, 45^\circ, 90^\circ$ ve 135° yönlerinde oluşturulmuş ve her bir matris için beş doku tanımlayıcısı hesaplanarak her ROI için 20 tanımlayıcı çıkartılmıştır. $d = \{1, 3, 6, 9\}$ uzaklık parametresi kullanılarak oluşturulmuş GLCM matrislerinden her ROI için 4×20 boyutlarında öznitelik matrisleri elde edilmiştir. Dalgacık dönüşümü ile öznitelik çıkarımı yapılırken her bir ROI Daubechies3 dalgacığı ile üçüncü seviyede ayrıştırılarak 6 detay katsayı matrisi elde edilmiştir. Her bir detay katsayı matrislerinden beş tanımlayıcı hesaplanarak 30 öznitelik çıkartılmıştır. Toplamda 120 ROI bulunduğu için sınıflandırmada 120×30 'luk öznitelik matrisi kullanılmıştır. Ridgelet dönüşümü ile öznitelik çıkarımı yapılırken her bir ROI Daubechies3 dalgacığı kullanılarak iki seviyede ayrıştırılmış ve her biri 62 sütun içeren iki detay katsayı matrisi elde edilmiştir. Her iki detay matrisinden de beş tanımlayıcı hesaplanmış ve bir ROI için 620 öznitelik çıkartılmıştır. Böylece sınıflandırmada kullanmak üzere 120×60 'lık öznitelik matrisi

oluşturulmuştur. Sınıflandırma performansını artırmak için genetik algoritma ve rastgele orman algoritması birlikte kullanılarak öznelik seçimi yapılmıştır. Çalışma sonucunda en iyi sınıflandırma oranının AUC=0.90 ile dalgacık dönüşümü ile elde edilen özneliklerden genetik algoritma ile seçilenlerin kullanılması ile elde edildiği görülmüştür.

Hachama vd. (2012), mamogram sınıflandırma problemi için imge çakıştırma yöntemini kullanmışlardır. Çalışmada bilinen imge çakıştırma yöntemlerinde benzerlik ölçütünün tek bir genel parlaklık değerine göre belirlendiği fakat medikal görüntüleme olduğu gibi çakıştırılacak imgelerin parlaklık değerleri bağımlılığının lokasyon açısından homojen olmadığı durumlarda işe yaramadığı belirtilmiştir. Bu nedenle, çalışmada benzerlik ölçütü parlaklık bağımlılığı sınıflandırmasına göre belirlenmiştir.

Radovic vd. (2013), mamografi imgelerinde kitle tespiti için öznelik seçimi yapan dört aşamalı bir sistem geliştirmişler ve sistemi MIAS veritabanından alınmış mamografi imgeleri ile test etmişlerdir. Sistemin ilk aşamasında mamografi imgelerinin 1.derece Haar dalgacık dönüşümü katsayıları eşiklenmiş ve ters dalgacık dönüşümü uygulanarak gürültüsüz imgeler elde edilmiştir. Bu imgelerde arka plandan arındırılmış meme bölgesinden pektoral kaslar atılarak ROI tespiti gerçekleştirilmiştir. Sistemin ikinci aşamasında, her bir ROI'den $\theta = 0^{\circ}$ yön ve $d = 1$ uzaklık parametreleri kullanılarak oluşturulan GLCM matrislerinden 20 öznelik çıkartılmıştır. Sınıflandırma aşamasında bu öznelikler, SVM, Naïve Bayes, k -En Yakın Komşu (k -Nearest Neighbor – k -NN), karar ağacı (decision tree), rastgele orman, sinir ağları ve Lojistik Regresyon Sınıflandırıcısı (Logistic Linear Classifier – LLC) kullanılarak sınıflandırılmıştır. En iyi sonuç %76 hassasiyet ile geriye yayımlı çok katmanlı sinir ağlarından MLP sinir ağlarından elde edilmiştir. Sınıflandırma hassasiyetini artırabilmek için Minimum-Fazlalık Maksimum-İlişki (Minimum-Redundancy Maximum-Relevance – mRMR) algoritması ile 20 öznelik arasından 5 tanesi seçilmiştir. Bu durumda en iyi sonuç %79,33 hassasiyetle C4.5 karar ağaçları algoritması ile elde edilmiştir.

Savitha vd. (2013), mamografi imgelerinin kompleks düzlemde incelenmesinin doğru sınıflandırma oranı artıracağını savunmuşlardır. Bu nedenle mamografi imgeleri dairesel dönüşüm ile kompleks düzleme aktarılmış ve Tam Kompleks-Değerli Gevşetme Sinir Ağı (Fully Complex-Valued Relaxation Neural Network – FCRN) olarak

adlandırdıkları sinir ağları ile sınıflandırılmıştır. FCRN doğrusal giriş katmanı, doğrusal olmayan tek gizli katman ve çıkış katmanından oluşmaktadır. Gizli katman aktivasyon fonksiyonu Gauss-benzeri hiperbolik sekant fonksiyonu, çıkış katmanı aktivasyon fonksiyonu olarak üstel fonksiyon kullanılmıştır. İzdüşüm-tabanlı öğrenme algoritması ile logaritmik enerji fonksiyonunu en küçükleyen çıkış ağırlıkları belirlenmiştir. FCRN, MIAS veritabanından alınan 97 imge ile eğitilmiş 11 imge ile test edilmiştir. Sinir ağına kompleks imgeden çıkartılan 9 öznitelik giriş olarak verilmiş ve %97,84 başarı oranı elde edilmiştir.

Ganesan vd. (2014), mamografi imgelerini üçlü öznitelik çıkarımı uygulayarak normal ya da kanserli olarak sınıflandıran bir yöntem geliştirmişlerdir. Yöntem, Singapur Anti-Tuberculosis Association CommHealth'den alınmış 313 (223 normal, 43 iyi huylu kanserli, 47 kötü huylu kanserli) mamografi imgesi üzerinde test edilmiştir. Çalışmada mamografi imgelerinin dönme, ölçekleme ve öteleme gibi doğrusal bozuntulara uğramış bir koordinat sisteminde görüldüğü belirtilmiştir. Bu nedenle ilk olarak imgelere iz dönüşümü (trace transform) uygulanmıştır. Elde edilen 2-boyutlu imge önce çapsal (diametrical) fonksiyonlar ile 1-boyutlu vektöre, sonra da dairesel (circus) fonksiyonlar ile bir noktaya dönüştürülmüştür. Bu noktanın her bir imge için tek ve dönme, öteleme, ölçekleme gibi bozuntulara karşı değişmez olduğu belirtilmiştir. Elde edilen öznitelik sayısı çok fazla olduğu için bir öznitelik seçim algoritması ile sınıflandırmada kullanılacak öznitelikler belirlenmiştir. Çalışmada, SVM, en yakın ortalama, Doğrusal Ayırtaç Sınıflandırıcısı (Linear Discriminant Classifier – LDC) ve GMM olmak üzere 4 farklı sınıflandırıcı kullanılmış ve en iyi sonuç %92,48'lik tanıma oranı ile GMM'den elde edilmiştir.

Agrawal vd. (2014), mamografi imgelerinde kitle tespiti için pektoral kasların varlığından etkilenmeyen bir yöntem geliştirmişler ve MIAS veritabanında test etmişlerdir. Yöntemde ilk olarak mamografi imgeleri, gradyan tabanlı bir yaklaşım kullanılarak meme sınırları belirlenmiş mamografi imgelerine Adaptif global eşikleme ve histogram gerdirme uygulanarak oluşturulan maske ile maskelenmiştir. Maskelenmiş imgelerin kontrastı adaptif histogram eşitleme yöntemi ile artırılmıştır. Maskeleyenme sonucu oluşan keskin kenarlar 7-seviyeli Gauss-Laplace piramit ile yumuşatılarak ön-işlenmiş imgeler elde edilmiştir. Ön-işlenmiş imgeler pektoral kaslardan etkilenmeyecek şekilde Grafik-Tabanlı Görsel Çıkıntı (Graph-Based Visual Saliency – GBVS) algoritması ile bölütlenerek

imgelerin çıkıntı haritası (saliency map) elde edilmiştir. Çalışmada, GBVS yönteminin, bölgelerin çıkıntılığını yerel komşuluklarının yönlü kontrastlarını kullanarak hesapladığı için yoğun dokulardaki kitleleri de tespit edebildiği belirtilmiştir. İmgeler çıkıntı haritasındaki en büyük değerin yarısı ile eşiklenmiş ve ROI'ler çıkartılmıştır. 100 pikselden az piksel içeren bölgeler göz ardı edilmiştir. Bu şekilde 58 kitleden 49'u tespit edilebilmiştir. GBVS ile çıkartılmış ROI'ler normal hücrelerde içerdiğinden, yanlış-pozitiflik oranını azaltmak amacı ile 30-210 piksel boyutluğunda dairesel bölgeler elde edilmiş ve bu bölgelerden parlaklık öznitelikleri, frekans düzlemi öznitelikleri ve dalgacık düzlemi öznitelikleri olmak üzere toplam 154 öznitelik çıkartılmıştır. Sınıflandırma için SVM 3. dereceden polinomial kernel ile kullanılmıştır. En uygun öznitelikler mRMR kriteri ile belirlenmiştir. Sonuç olarak en iyi performansın $AUC=0,891\pm 0,001$ ile Ayırık Dalgacık Dönüşümü (Discrete Wavelet Transform – DWT) ve RDWT Reverse Bi-orthogonal 2.2 dalgacığı ile ilk 3 seviyede elde edilen entropi özniteliklerinin birlikte kullanılması ile elde edildiği belirtilmiştir.

Pareira vd. (2014), CC ve MLO mamogramları birlikte değerlendirerek kitle tespiti yapan bir sistem geliştirmişlerdir. Geliştirilen sistem mamografi imgelerindeki yapaylık giderimi, kontrast iyileştirmesi, bölütleme ve yanlış-pozitiflik oranının azaltılması aşamalarından oluşmaktadır. Çalışmada DDSM veritabanından alınmış 160 (80 iyi huylu kanserli, 80 kötü huylu kanserli) hastanın sağ ve sol meme CC ve MLO mamografi imgeleri kullanılmıştır. Mamografi imgelerindeki artefaktların giderilmesi için imgelere top-hat morfolojik işlemi uygulanmıştır. Orijinal imgeden top-hat algoritması sonucu elde edilen imge çıkartılmış, fark imgesi Otsu eşikleme yöntemi ile eşiklenmiş ve eşiklenen imge orijinal imge ile çarpılarak artefaktlardan arındırılmış mamografi imgesi elde edilmiştir. İmgelerin kontrastını iyileştirmek için dalgacık düzleminde Adaptif Wiener filtre uygulanmıştır. Bu amaçla imgeler Haar, Daubechies ve Coiflets dalgacıkları kullanılarak üç seviyeye ayrıştırılmıştır. Adaptif Wiener filtre ayrışım sonucu elde edilmiş yaklaşım imgelerine uygulanmıştır. En iyi sonucun Coiflets dalgacığı kullanılarak ayrışım yapıldığında elde edildiği belirtilmiştir. Artefaktlardan arındırılmış ve kontrastı iyileştirilmiş mamografi imgelerinde kitle tespiti için Hammouche vd. (2008) tarafından geliştirilmiş bölütleme algoritması kullanılmıştır. Bu algoritmada, imgenin gri-seviye histogramı Daubechies2 dalgacığı ile yaklaşım ve detay bileşenlerine ayrılmaktadır. Yaklaşım bileşenlerine genetik algoritma uygulanarak eşik sayısı ve değerleri

belirlenmekte ve buna göre imge bölütlenmektedir. Bu algoritma sonucunda oluşan yanlış-pozitiflik oranını azaltmak amacı ile eşiklenmiş imgelerden ROI'ler ayrılmış ve bu bölgelerin alan, merkez ve çevre ölçümleri hesaplanmıştır. Gözlemlere dayanarak kitlelerin 1000-40000 piksel kare boyut aralığında olduğu belirlenmiş ve bu boyutun üzerinde olan ROI'lerin kitle olmadığına karar verilmiştir. Bu durumda %95 hassasiyetle kitle tespiti yapılırken imge başına 2.80 yanlış-pozitiflik saptanmıştır. Geriye kalan bölgeler CC ve MLO mamogramlarında bir uzman tarafından karşılaştırılmıştır. İki mamogramda da %95 - %100 alana sahip bölgelerin kitle olduğu kabul edilmiştir. Bu durumda yine %95 hassasiyetle kitle tespiti yapılırken, yanlış-pozitifliğin imge başına 1,37'ye düştüğü görülmüştür. Çalışmada, kompleks-dalgacık dönüşümü, curvelet dönüşümü ve shearlet dönüşümü gibi yöntemlerin kullanılması ile hassasiyetin arttırılabileceği belirtilmiştir.

Anitha vd. (2017), önerdikleri CAD sisteminde iki aşamalı adaptif eşikleme (dual stage adaptive thresholding) kullanmışlardır. Önerilen sistemde ROI tespiti histogram ve yerel pencere eşikleme yöntemiyle elde edilmiştir. Daha sonra tüm imgenin histogram tepe analizi ile global eşikleme gerçekleştirilmiş, eşik ise önerilen eşik seçim kriterini maksimize edecek şekilde elde edilmiştir. Çalışmanın %93 hassasiyet ile gerçekleştirildiği belirtilmiştir.

2.3. Anomali Tipi Sınıflandırması Yapılan Çalışmalar

Vallez vd. (2013), mamografi imgelerinde lezyon tespit ve sınıflandırması öncesinde meme doku tipinin belirlenmesinin başarı oranını artıracağını savunmuşlardır. Bu amaçla geliştirdikleri üç aşamalı sistemin ilk aşamasında mamografi imgesi arka plandan arındırılmış, ikinci aşamasında meme doku tipi belirlenmiş ve üçüncü aşamada lezyon tespiti gerçekleştirilmiştir. Yöntem, Tarama-Film Mamografi (Screening-Film Mammography – SFM) ve FFDM veritabanlarından alınan 1459 mamogram üzerinde test edilmiştir. Ön işlem aşamasında meme bölgesine ait bir noktanın bulunmasından sonra eşik değeri Otsu eşikleme yöntemi ile belirlenmiş 8-yönlü zincir kodu kullanılarak meme bölgesi arka plandan arındırılmıştır. Sınıflandırma için arka plandan arındırılmış tüm meme bölgesinden 298 öznitelik çıkartılmıştır. Bu özniteliklerin 241 tanesi imge histogramlarından ve $d = \{1, 3, 5\}$ uzaklık ve $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ yön parametreleri ile

GLCM matrislerinden elde edilmiş ve bu özneliklerin homojen meme doku tiplerinde etkili olduğu belirtilmiştir. Diğer 57 öznelik ise yerel ikili örüntülerin ortalama, varyans, basıklık (kurtosis), asimetri değerleri ve 24 Gabor filtre Chebyshev momentleri, ortalama ve varyans değerleridir ve bu özneliklerin heterojen meme doku tiplerinde etkili olduğu belirtilmiştir. Çıkarılan özneliklere Kolmogorov-Smirnov testi, Levene testi, varyans analizi ve Kruskal-Wallis testleri uygulanarak meme doku tipinden etkilenmeyen öznelikler bulunmuş ve bunların dışındaki öznelikler seçilmiştir. Seçilen özneliklere PCA, Doğrusal Ayırtaç Analizi (Linear Discriminant Analysis – LDA), öznelik sıralama ve öznelik seçme algoritmaları uygulanarak boyut indirgeme yapılmış ve en uygun boyut indirgeme yönteminin LDA olduğu tespit edilmiştir. Meme doku tipi sınıflandırması için SVM, doğrusal Bayes, Naïve Bayes, k -NN ($k=1$), karesel (quadratic) sınıflandırıcı, en yakın ortalama, Fisher ve parzen sınıflandırıcıları ağırlıklandırılarak kullanılmıştır. Sınıflandırma için birinci seviyede yağlı ve yoğun dokuları, ikinci seviyede homojen ve heterojen dokuları belirleyen iki seviyeli bir ağaç yapısı oluşturulmuştur. Yöntem ile birinci seviyede SFM veritabanında %99, FFDM veritabanında %96,93, ikinci seviyede ise SFM ve FFDM veritabanlarında sırasıyla %99 ve %91,57 tanıma oranları sağlanmıştır. Lezyon tespiti için uyarlamalı filtreleme, Daubechies20 dalgacığı ile dalgacık dönüşümü ve bulanık- k -ortalama (fuzzy- k -means) kullanılmış ve FFDM veritabanından her dokudan eşit sayıda örnek olacak şekilde 724 mamografi imgesi üzerinde test edilmiştir. Yoğun doku tiplerinde yöntemlerin parametrelerinin daha yüksek olduğu belirtilmiştir. Mikrokalsifikasyon tespitinde uyarlamalı filtre, bozuntu tespitinde dalgacık dönüşümü ve kitle tespitinde bulanık- k -ortalama yönteminin daha yüksek tanıma oranı sağladığı görülmüştür. Sonuç olarak, meme doku tipi sınıflandırması yapılmadan önce yaklaşık %78 olan başarı oranı meme doku tipi sınıflandırmasından sonra %91'e yükseltilmiştir. Çalışmada sonraki çalışmalarda başka doku özneliklerinin değerlendirilebileceği ve "kelime çantası (bag-of-words)" yönteminin kullanılabileceği belirtilmiştir.

Guliatto vd. (2008), poligon modelleme yönteminin gürültü etkisini azaltırken önemli öznelikleri koruduğu gerekçesi ile mamografi imgelerinde kanser tespiti konusunda daha etkili olacağını belirtmişlerdir. Yöntem, SFM ve MIAS veritabanlarından temin edilmiş 65 iyi huylu kanserli 46 kötü huylu kanserli olmak üzere 111 kontur üzerinde test edilmiştir. Yöntemde ilk olarak her bir konturun doğrusal bölütleri belirlenmiş ve her komşu iki bölütün uzunluklarına ve açılara bakılarak daha fazla nokta

içeren daha az sayıda bölütler elde edilmiştir. Konturların iyi huylu ya da kötü huylu olmasının tespiti için her bir spikülün uzunluğu ve açısı kullanılarak spikülasyon indeksi hesaplanmış ve bir eşik değerinden geçirilmiştir. Spikülasyon indeksinin yüksek olduğu konturların kötü huylu, düşük olduğu konturların ise iyi huylu olduğuna karar vermişlerdir. Çalışmada yöntemin başarı oranı %94 olarak belirtilmiştir.

Karahaliou vd. (2008), mamografi imgelerinde mikrokalsifikasyon içeren bölgeyi çevreleyen 128×128'lik ROI bölgesinin doku özneliklerini kullanarak mikrokalsifikasyonun iyi huylu ya da kötü huylu olduğuna karar vermişlerdir. Çalışmada mamografi imgelerin öncelikle dalgacık-tabanlı uzamsal adaptif (wavelet-based spatially adaptive) bir yöntem ile kontrastı artırılmıştır. Ön işlenmiş imgelerde 128×128'lik ROI bölgeleri belirlenmiş ve bu bölgelerden gri-seviye ve dalgacık doku öznelikleri çıkartılmıştır. Gri-seviye ve dalgacık doku öznelikleri, olasılıksal sinir ağları (probabilistic neural networks) ile ayrı ayrı sınıflandırma işlemine tabi tutulmuştur. İki sınıflandırma sonuçları çoğunluk oylama ile sonuçlandırılmıştır. Yöntem, DDSM veritabanından yağlı-glandüler ve yoğun glandüler meme doku tipine sahip 54'ü kötü huylu mikrokalsifikasyon içeren 100 imge üzerinde test edilmiş ve %94'lük bir tanıma oranı sağlanmıştır.

Verma vd. (2009), yaptıkları çalışmada dijital mamogramlarda tespit edilen kitlelerin iyi ya da kötü huylu olarak sınıflandırılması için bir Yumuşak Kümeleme Sinir Ağı (Soft Clustering Neural Network–SCNN) tekniği sunmuşlardır. Çalışma, mamografi imgelerinin işlenerek olası kitle bölgelerinin çıkarılması, bu bölgelerin özneliklerinin belirlenmesi ve SCNN yöntemi ile sınıflandırma yapılması aşamalarını içermektedir. Yapılan çalışmada eğitim ve test için DDSM veri tabanından alınan 100 kötü huylu kanserli ve 100 iyi huylu kanserli mamografi imgesi kullanılmıştır. Mamografi imgelerinden ROI'lerin tespiti Kumar vd. (2006) ve Panchal ve Verma (2006) çalışmalarında geliştirilen program ile gerçekleştirilmiştir. ROI'lerden hasta yaşı, küme belirginliği ve BI-RADS tanımlayıcı öznelikleri (yoğunluk, kütle şekli, kütle sınırı ve anormallik değerlendirme derecesi) çıkarılmıştır. SCNN yönteminin eğitimi, bir hata matrisi ile zayıf kümeleri ve küme bulunmayan durumları giriş verisinden çıkardıktan sonra gerçekleştirilmiştir. Sonuç olarak 10-katlamalı (ten-fold) çapraz doğrulama(cross-

validation) metodolojisi kullanılarak yapılan sınıflandırmada SCNN'nin %94 hassasiyet ile kitle tipini belirleyebildiği saptanmıştır.

Keleş vd. (2011), yaptıkları çalışmada meme kanseri tespiti yapan bir uzman sistem geliştirmişlerdir. Mamografik kitle veri seti UCI Machine Learning Repository'den alınmıştır. Çalışmada BI-RADS derecelendirmesi (0: iyi huylu kanserli- 5: kötü huylu kanserli), hasta yaşı, kütle şekli, kütle sınırı ve yoğunluk öznitelikleri kullanılmıştır. Çalışmada sinirsel-bulanık (neuro-fuzzy) yöntemi ile elde edilen fuzzy kurallarına sahip bir uzman sistem oluşturmuşlardır. Sinirsel-bulanık yöntemi sezgisel öğrenme algoritması kullanan 3 katmanlı ileri beslemeli bir yapıya sahiptir. Burada giriş katmanında BI-RADS derecelendirmesi, kütle şekli ve kütle sınırı, gizli katmanda ise bulanık kuralları yer almaktadır. Çıkış katmanındaki iki nöron ise kitlenin iyi huylu ya da kötü huylu olma durumlarıdır. Bu yapıyı Visual C#'de oluşturarak ve veri tabanı platformu olarak SQL Server kullanarak bir arayüze dönüştürmüşlerdir. Arayüzde, hastanın numarası, adı-soyadı, medikal değerleri, doktor teşhisi ve uzman sistemin teşhisi gibi her hasta için detaylı bilgiler yer almaktadır. Çalışmanın sonucunda bu sistemin yaklaşık %76 hassasiyette çalıştığı belirtilmiştir.

Maitra vd. (2011), MIAS veritabanını kullandıkları çalışmada, yağlı, yağlı-glandüler ve yoğun-glandüler meme doku tiplerindeki anomalileri (kalsifikasyon, tanımlanmış, tanımlanamamış kitle ve diğer hastalıklı kitleler) belirlemek için iki aşamalı bir yöntem önermişlerdir. Yöntemin ilk aşamasında dijital mamografi imgelerinin kontrastı bir eşik değeri ile artırılarak meme bölgesi çıkartılmıştır. Yöntemin ikinci aşamasında, arka plandan arındırılmış mamografi imgelerinden 8×8'lik homojen bloklar elde edilerek anomalili kitle belirginleştirilmiştir. Elde edilen homojen blokların renk basamaklama yöntemi ile bölütlenmesi sonucu mamografi imgelerinde anomali olup olmadığı açıkça görülebildiği belirtilmiştir. Segmente edilmiş imgelerin $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ yön ve $d = 1$ uzaklık parametreleri kullanılarak GLCM matrisleri oluşturulmuş; bu matrislerden kontrast, enerji, homojenlik ve korelasyon öznitelikleri çıkartılmıştır. Bu şekilde anomali tipinin tespit edilebildiğinin ifade edildiği çalışmada başarı oranı belirtilmemiştir.

Eltoukhy vd. (2012), dijital mamografi görüntülerinde meme kanseri teşhisi için bir sistem geliştirmişler ve sistemin MIAS veritabanından alınan 322 mamografi imgesi üzerinde test etmişlerdir. Çalışmada, kırpma operasyonu sonucu arka plandan arındırılan mamografi imgelerinden 128×128'lik ROI'ler manüel olarak elde edilmiştir. Öznitelik çıkarmak için mamografi imgelerinden çıkartılmış ROI'lere dalgacık ve curvelet dönüşümleri uygulanmıştır. Sonrasında özniteliklerin kapasitesini belirlemek için bir öznitelik listesi oluşturmuş ve dinamik eşikleme uygulanarak en önemli öznitelikleri belirlenmiştir. Veri seti %70 eğitim ve %30 test olarak ayrıştırılmıştır. Eğitim seti SVM sınıflandırmasını oluşturmak için, test ise sınıflandırıcının hassasiyetini belirlemek için kullanılmıştır. Bu yöntem, en az sayıda katsayı kullanarak en yüksek hassasiyeti elde edene kadar tekrarlanmıştır. Yöntem iki tip sınıflandırma için uygulanmıştır. İlk adımda mamogramın normal ya da anomalili olduğu tespit edilirken, ikinci adımda anomalili mamogramın iyi ya da kötü huylu olduğunun tespiti yapılmaktadır. İlk olarak öznitelik çıkarmak için dalgacık dönüşümü uygulanmıştır. Bu durumda maksimum hassasiyetin ilk sınıflandırmada 1238 katsayı ile %95,84, ikinci sınıflandırmada ise 150 katsayı ile %96,56 olduğu belirtilmiştir. İkinci olarak da öznitelik çıkarmak için curvelet dönüşümü uygulanmıştır. Burada da maksimum hassasiyetin ilk sınıflandırmada 5663 katsayı ile %95,98, ikinci sınıflandırmada ise 333 katsayı ile %97,30 olduğu belirtilmiştir. Sonuçları doğrulamak için optimize edilmiş eşik değerinde 5-katlamalı çapraz doğrulama metodolojisi uygulanmıştır. Sonuç olarak her iki yöntemin de verimli olduğu tespit edilmiştir. Çalışmada dalgacık dönüşümünün katsayı miktarını 25129'dan birinci sınıflandırma için 1238'e, ikinci sınıflandırma için 150'ye; curvelet dönüşümünün ise 46080'den birinci sınıflandırma için 5663'e, ikinci sınıflandırma için 333'e düşürdüğü belirtilmiştir.

Dhanalekshmi ve Phadke (2013), mamografi imgelerini güç spektrum analizi (power spectral analysis) tabanlı oransal kırılma (fractal) analizi ile sınıflandırmışlardır. Mamografi imgelerinde ROI manüel olarak belirlenmiştir. ROI'ler Hanning penceresi ile pencerelenmiş, pencerelerin Fourier dönüşümü alınmış ve büyüklük cevabının karesi alınarak oransal kırılma boyutları (fractal dimension) hesaplanmıştır. Yöntem MIAS veritabanından alınan 22 mamografi imgesi üzerinde test edilmiştir. İyi huylu kitlelerde kırılma boyut değerinin [0,8 1,1], kötü huylu kitlelerde [1,2 1,6] aralığında, diğer anomali

durumlarında ise bu aralıkların dışında olduğu tespit edilmiş ve %82 başarı oranı elde edilmiştir.

Casti vd. (2013), mamografi imgelerinde görülen kitleleri iyi huylu ve kötü huylu olarak sınıflandıran ve bu süreçte kitle sınırlarını düzgün tespit etme zorunluluğu olmayan bir sistem geliştirmişlerdir. Bu amaçla ilk olarak FFDM veritabanından alınmış 120 iyi huylu kanserli, 26 kötü huylu kanserli kitle içeren 146 ROI imgesinin kontrastı Look-Up-Table dönüşümü ile iyileştirilmiştir. İyileştirilen imgelerden merkezi kitlenin merkezi olan dairesel bölgelerden öznitelik çıkartılarak kitle sınırlarını düzgün tespit etme gerekliliği ortadan kaldırılmıştır. Bu dairesel bölgelerden ve bu bölgelerin Gabor filtre cevaplarından radyal korelasyon ve radyal eğilim değerleri hesaplanmıştır. Çalışmada düşük radyal korelasyon ve eğilim değerlerinin iyi huylu kitleleri, yüksek değerlerin ise kötü huylu kitleleri betimlediği belirtilmiştir. Her bir dairesel bölge için hesaplanan dört değer Fisher Doğrusal Ayırtaç Analizi (Fisher's Linear Discriminant Analysis – FLDA) Bayes sınıflandırıcı, SVM ve Radyal Tabanlı Fonksiyon Sinir Ağları (Radial Based Function Neural Network – RBFNN) kullanılarak sınıflandırılmıştır. FLDA, BAYES, SVM ve ANN-RBF sınıflandırma sonuçları sırası ile %82, %76, %85 ve %93 olarak belirtilmiştir.

Casti vd. (2016) bir başka çalışmalarında, çok aşamalı bir yaklaşımla memenin Gauss eğikliğini çıkarmış, dairesel ROI'leri çıkarmış, piksel değerlerinin yerel devamsızlığı ve uzamsal bağımsızlığına bakarak konturdan bağımsız öznitelik tanımlayıcıları tanımlamışlardır. Bu çok aşamalı yaklaşımın içinde aynı anda ROI tespiti ve yönelimli bileşenlerin çıkarımını yaparak dairesel ROI'lerin çıkarım başarımını artırmıştır. Lezyonları ve kötü huylu ROI'leri ayrı ayrı sınıflandırarak iyi huylu ve normal ROI'leri bir sınıf, kötü huylu ROI'leri bir sınıf olarak düşünmüş, sınıflandırmayı FLDA ile gerçeklemiştir. Önerilen sistemi MIAS, DDSM ve FFDM veritabanlarına uygulamış ve %92 hassasiyetli kötü huylu tümör tespitini başarmışlardır.

Swiderski vd. (2017) DDSM veritabanına odaklanarak negatif olmayan matris faktörizasyonu yöntemiyle oluşturulan tanımlayıcılar ile mamogram sınıfları arasındaki ayrımı genişletmeyi amaçlamışlardır. Yazarlar çalışmalarında %83 hassasiyet, %80,2 doğruluk ve AUC=0,83 değerleri elde ettiklerini belirtmişlerdir.

Choi vd. (2016) önerdikleri CAD sisteminde farklı özniteliklerden türetilmiş yeni öznitelik kümesi ile daha ayrıntılı sınıflandırma elde etmiş ve öznitelik sınıflandırıcı seçiminde özgün bir seçim yöntemi önermişlerdir. Kullanılan öznitelikler literatürde mevcut öznitelikler olmakla beraber bu özniteliklerden türetilen öznitelikler AdaBoost ve ağırlıklandırma ile elde edilmiştir. Sınıflandırıcı seçiminde ise daha iyi cevap veren sınıflandırıcıları seçen yerel öğrenme hipotezini kullanan bir yol izlemişlerdir. Çalışmada önerilen CAD sisteminin AUC=0,932 değeriyle literatürdeki temel sınıflandırma yaklaşımlarının önünde olduğu belirtilmiştir.

Beura vd. (2015) iki boyutlu dalgacık dönüşümü ve normalize edilmiş GLCM matrisleri ile oluşturulan öznitelikleri F-test ve t-teste sokarak öznitelik kümesinden seçerek geriye yayımlı sinir ağı kullanan bir sınıflandırıcı ile yeterli başarımla elde etmişlerdir. Önerdikleri sistem ile MIAS veritabanında sağlıklı – kanserli doku için %98, iyi huylu kanserli – kötü huylu kanserli doku için %94,2 ve DDSM veritabanında ise sağlıklı – kanserli doku için %98,8 ve iyi huylu kanserli – kötü huylu kanserli doku için %97,4 doğruluk elde etmişlerdir.

Muramatsu vd. (2016) çalışmalarında bölütlemeye ihtiyaç duymamasından dolayı desen tipinde özniteliklere odaklanmışlardır. Burada deseni elde etmek için üçlü yerel radyal örüntülerini kullanmışlar, bu örüntülerden elde edilen özniteliklerin başarımlarını üçlü yerel örüntüleri, dönmeden bağımsız bir biçimli (uniform) üçlü yerel örüntüleri, GLCM ve dalgacık öznitelikleriyle oluşturulmuş öznitelik kümelerinin MLP ve SVM sınıflandırıcılarıyla başarımlarını gözlemlemişlerdir. Öznitelikleri kendi araştırma merkezlerinde elde ettikleri mamogramlar üzerinde denemiş, literatürde geçerli veritabanlarında denemeler yapmamışlardır. Bu veritabanı için üçlü yerel radyal yerel örüntüleri için AUC=0,90 değeri elde etmişlerdir.

2.4. Pektoral Kas Tespiti Yapılan Çalışmalar

Mamografi imgelerinin pektoral kaslardan arındırılması konusunda literatürde yer alan çalışmalar genel olarak yeşinlik (intensity)-tabanlı ve dalgacık-tabanlı yaklaşımlar olup, bu yaklaşımlar doğru tespiti (line detection) yöntemleri, istatistiksel yöntemler ve diğer yöntemler olmak üzere üç ana başlıkta incelenmektedir (Ganesan vd., 2013).

Yeğlilik-tabanlı yaklaşımlar, pektoral kas bölgesi yeğlilik seviyelerinin meme dokusu yeğlilik seviyelerine göre daha yüksek olacağı ön bilgisini kullanmaktadırlar. Bu yaklaşımlar doğrudan piksel yeğliliklerini (Saltanat vd., 2010; Roshan ve Harada, 2007; Nagi vd., 2010; Liu vd., 2010; Liu vd., 2012), imgelerin histogram (Thangavel ve Karnan, 2005; David vd., 2005; Subashini vd., 2010) ya da gradyan (Chakraborty ve Mukhopadhyay, 2012) bilgilerini kullanarak uygulanabileceği gibi imgelerin gradyanlarına da (Camilus, 2011) uygulanabilmektedir. Ayrıca, uzamsal düzlem yerine dalgacık düzlemde pektoral kas giderimi amaçlanan çalışmalar da literatürde mevcuttur (Mustra vd., 2009; Ferrari vd., 2004; Ma ve Manjunath, 2000).

Mamografi imgelerinden pektoral kas giderimi konusunda literatürde sıklıkla karşılaşılan doğru tespiti yöntemleri pektoral kasların mamografi imgelerinde meme yönüne bağlı bir köşede üçgen geometrisine sahip karakteristikte olduğu ön bilgisine dayanmakta ve üçgenin hipotenüsünü bulmayı amaçlamaktadır. Bu amaçla, düz doğru kestirimi (straight line estimation) (Nagi vd., 2010; Chakraborty ve Mukhopadhyay, 2012; Karssemeijer, 1998; Kwok vd., 2001, 2004; Mustra ve Grgic, 2013; Tzikopoulos vd., 2011; Vaidehi ve Subashini, 2013), Hough dönüşümü (Ferrari vd., 2004; Yam vd., 2001) ve eğri uydurma (Weidong vd., 2007) yöntemleri kullanılmaktadır. Üçgen geometrisine benzetilmiş pektoral kas bölgesinin hipotenüs kenarı tam bir doğru karakteristiğinden ziyade eğrisel bir yapı gösterdiğinden elde edilen doğruyu yöntem başarısını artırabilmek adına dinamik programlama yöntemleri ile eğrileştiren çalışmalar da literatürde mevcuttur (Karssemeijer, 1998; Kwok vd., 2001, 2004; Yam vd., 2001).

İstatistiksel yöntemler kullanılarak yapılan çalışmalar incelendiğinde, son yıllarda graf-temelli yaklaşımlar (graph theory approach) mamografi imgelerinden pektoral kas giderimi amaçlı kullanılmaya başlandığı görülmektedir (Domingues vd., 2010; Camilus vd., 2010; Abdellatif vd., 2012). Bir başka çalışmada ise Moayedı vd. pektoral kas giderimi için piksel enerjilerinin logaritmasını kullanmışlardır (Moayedı vd., 2010).

Bu yöntemlere ek olarak, ortalama-kaydırma bölütleme (mean-shift segmentation) (Sultana vd., 2010), bağlı bileşen etiketleme (connected component labeling) (Li, 2012) ve bölge büyütme algoritması (region growing algorithm) (Maitra vd., 2012; Talha ve Sulong, 2012) gibi yöntemlerde literatürde yer almaktadır.

2.5. Kanser Riski Tespit Edilen Çalışmalar

Li vd. (2013), meme doku tipine bağlı meme kanseri risk tespitini arka plan parlaklığından bağımsız doku öznelikleri kullanarak yapmışlardır. Yazarlar, BI-RADS kategorisine göre yağlı ve yağlı-glandüler olarak sınıflandırılan meme doku tiplerinin düşük, heterojen yoğun ve çok yoğun olarak sınıflandırılan meme doku tiplerinin ise yüksek kanser riski taşıdığını belirtmişlerdir. Çalışmada DDSM veritabanından 3 farklı veri seti oluşturulmuştur. Birinci veri setinde her meme doku tipinden 40 (20 eğitim için, 20 test için) tane olmak üzere 160 CC mamografi imgesi bulunmaktadır. İkinci veri setinde her meme doku tipinden 40 (20 eğitim için, 20 test için) tane olmak üzere 160 MLO mamogram ve üçüncü veri setinde her meme doku tipinden 20 CC ve 20 MLO (10 CC& 10 MLO eğitim için, 10 CC & 10 MLO test için) olmak üzere 160 sağ meme mamogramından oluşmaktadır. Mamografi imgelerinden artifaktların, pektoral kasların atılması ve meme sınırlarının belirlenmesi işlemleri ImageJ yazılım ile manüel olarak yapılmıştır. Arka plan parlaklığından bağımsız doku özneliklerinin çıkartılabilmesi için imgeler normalize edilmiştir. İlk olarak, normalize edilmiş imgelerin her bir pikseli 3x3 komşuluğundaki 8 pikselden oluşan bir vektör ile temsil edilmiştir. İkinci aşamada, dört meme doku tipinin her biri için vektör topluluklarına k -ortalama kümeleme ($k=5$) uygulanarak alt-texton sözlükleri oluşturulmuştur. Sınıflandırmada dört alt-texton sözlüğün birleştirilmesi ile oluşturulmuş tek bir texton sözlüğü kullanılmıştır. Üçüncü aşamada, 8-boyutlu vektörler en yakın küme merkezine göre etiketlenerek her imgenin texton haritası çıkartılmış ve imgeler bu textonların histogramı ile temsil edilmiştir. Sınıflandırma aşamasında birinci, ikinci ve üçüncü veri setleri için sırasıyla $k=2$, $k=2$ ve $k=7$ olarak alınan k -NN sınıflandırıcısı kullanılmıştır. Birinci veri seti kullanıldığında yağlı, yağlı-glandüler, heterojen-yoğun ve çok yoğun doku sınıflandırma hassasiyetleri sırasıyla %75, %25, %100 ve %65 olarak elde edilmiştir. İkinci veri setinde bu oranlar %75, %5, %60 ve %70, üçüncü veri setinde ise %70, %0, %80 ve %35 olarak elde edilmiştir. Yağlı ve yağlı-glandüler doku tipleri birinci sınıf, heterojen-yoğun ve çok yoğun doku tipleri ikinci sınıf olarak kabul edilip iki-sınıflı sınıflandırma problemi çözüldüğünde birinci verisetinde ilk sınıfı ayırma hassasiyetinde düşüş gözlenirken diğer tüm oranlarda artış gözlenmiştir.

Li vd. (2014), bir başka çalışmalarında bir memedeki kanser riskini aynı kişinin diğer memesinin mamogramına bakarak tespit etmeye çalışmışlardır. Yazarlar, kansersiz

bir memenin kanser riskinin diğer meme kansersiz ise düşük, kanserli ise yüksek olduğunu belirtmişlerdir. Çalışmada DDSM veritabanından alınmış her bir BI-RADS sınıfından 20 düşük, 20 yüksek risk taşıyan mamografi imgesi eğitimde, aynı özellik ve sayıda mamografi imgesi de testte kullanılmıştır. İmgelerden artifaktlar ve pektoral kaslar manüel olarak kaldırılmış, meme bölgesi ImageJ yazılımı ile çıkartılmıştır. Göğsün farklı bölgelerinin risk değerlendirmesine katkısını incelemek için meme bölgesinden altı farklı kesit alınmıştır. Bunlardan ilk üçü merkezi meme ucu olan dairesel kesitlerdir. Diğer kesitler ise ilk üç kesitin birleşimi, tüm meme bölgesi ve tüm meme bölgesinden ilk üç kesitin çıkarılması ile elde edilen kesitler olarak tanımlanmıştır. Normalize edilmiş mamografi imgelerinden doku öznelikleri iki farklı sınıftan çıkartılmıştır. Bunlardan birincisi 3×3 , 5×5 ve 7×7 'lik komşuluklarındaki piksellerin parlaklığına göre çıkartılan texton öznelikleri. Bu özneliklerin çıkartılmasında Verma vd. (2009) tarafından belirtilen yol izlenmiştir. Sınıflandırma aşamasında k -NN ($k=3$), doğrusal SVM ve FLDA sınıflandırıcı kullanılmıştır. İkinci sınıf doku öznelikleri ise meme bölgelerine 10 yönlü Gabor filtre uygulanarak elde edilmiştir. Filtrelenmiş imgeler eşiklenmiş ve bağlı bileşenler elde edilmiştir. Bu bileşenlerden geometrik merkezin meme ucuna olan uzaklığı, geometrik merkez ile meme ucunu birleştiren doğrunun yatayla yaptığı açı, bileşenin ana eksenini ile geometrik merkezi meme ucuna birleştiren doğru arasındaki açı ve bileşenin alanı olmak üzere dört öznelik çıkartılmıştır. Bileşenin ana ekseninin yönü bağlı bileşenin piksel koordinatlarına PCA uygulanarak bulunmuştur. Sınıflandırma aşamasında k -NN ($k=9$), doğrusal SVM ve FLDA sınıflandırıcısı kullanılmıştır. Tüm sonuçlar değerlendirildiğinde, en verimli sonucun tüm meme bölgesinden 3×3 'lük komşuluklara bakılarak çıkartılan texton özneliklerinin SVM kullanılarak sınıflandırılması ile %71,3 hassasiyetle elde edildiği görülmüştür. Bu şartlar altında düşük riskli mamogramlar %85, yüksek riskli mamogramlar ise %57,5 hassasiyetle sınıflandırılmaktadır.

2.6. Diğer Çalışmalar

Petroudi vd. (2013), meme doku tipi sınıflandırmasında anlık genlik, frekans ve faz bilgilerini verdiği için Genlik Modülasyonu (Amplitude Modulation – AM) - Frekans Modülasyonu (Frequency Modulation – FM) ayrışımı ile elde edilebilecek doku özneliklerinin kullanılabileceğini belirtmişlerdir. Bu amaçla, 6-ölçek 8-yönlü bir Gabor filtre bankası ile AM-FM ayrışımı gerçekleştirilmiştir. Baskın bileşen analizi (dominant

component analysis) yöntemi ile tüm ölçek ve frekanslarda en fazla değerlendirilen yerel genlik bileşeni ile meme dokusu betimlenmiştir. Yöntem MIAS veritabanından alınan 62 yağlı, 62 yağlı-glandüler ve 62 yoğun-glandüler mamografi imgesi üzerinde test edilmiştir. Sınıflandırma için k -NN ($k=5$) sınıflandırıcısı kullanılmıştır. Çalışmada yağlı, yağlı-glandüler ve yoğun-glandüler meme tipleri sırası ile %81,25, %83,87 ve %86,49 hassasiyetle sınıflandırılmıştır. Tüm meme tipleri ele alındığında yöntemin %84 hassasiyetle çalıştığı belirtilmiştir. Yazarlar, gelecek çalışmalarında AM-FM ayrışımı için adaptif filtre bankası tasarımının yapılacağını belirtmişlerdir.

Li vd. (2014), kernel-tabanlı öğrenme algoritmalarının performansını artırabilmek için optimizasyon denklemleri Fisher ve maksimum mesafe (maximum margin) kriterlerine göre oluşturulmuş veriye-bağlı kernel optimizasyonu önermişlerdir. Önerdikleri optimizasyon yöntemini farklı veritabanlarında kernel PCA ve kernel LDA yöntemleri üzerinde test etmişlerdir. MIAS veritabanından alınmış 12 yağlı, 14 yağlı-glandüler ve 16 yoğun-glandüler mamografi imgesinden istatistiksel öznelikler çıkartılıp optimize edilmiş kernel LDA kullanarak %96,46 hassasiyet ile sınıflandırıldığı belirtilmiştir.

3. MAMOGRAFİ İMGELERİNİN ÖN İŞLENMESİ

Mamografi imgelerinin incelenmesinde temel olarak iki sorun radyoloji uzmanlarının yorumlama sürecini zorlaştırmaktadır. Bu sorunlardan biri ROI üzerindeki düşük kontrastlı özelliklerin kemik, pektoral kas gibi diğer dokular tarafından görülmesinin engellenmesidir. Bu sorunun aşılması için imgelere bir kontrast iyileştirme işleminin uygulanması gerekmektedir. Bir diğer sorun ise görüntüleme sırasında veya etiketleme sonucu imgelerde gürültü oluşmasıdır. Gürültü giderimi içinse imgelerin uygun filtreleme işlemlerine tabi tutulmaları gerekmektedir.

3.1. Kontrast İyileştirme

Histogram, bir imgedeki yeğinlik değerlerinin görülme sıklığını gösteren bir grafiğdir. Herhangi bir imgenin histogramına bakarak o imgenin kontrastı hakkında bilgi edinilebilmektedir. Örneğin, sola dayalı bir histogram karanlık bir imgeyi gösterirken sağa dayalı bir histogram aydınlık bir imgeyi göstermektedir. Bir imgenin karanlık veya aydınlık olması, bir başka deyimle histogramının dar olması o imgenin kontrastının düşük olduğu anlamına gelmektedir. Histogram eşitleme, her bir yeğinlik seviyesinde eşit sayıda piksel olacak şekilde imgelerin histogramını açma işlemidir. Böylece, imgenin histogramı genişlemiş yani kontrastı yükseltilmiş olmakta ve imgeden daha fazla bilgi alma olanağı sağlanmaktadır. Bu nedenle, tez çalışması kapsamında mamografi imgelerine kontrast iyileştirme süreci için histogram eşitleme işlemi uygulanmıştır.

3.2. Mamografi İmgelerinin İyileştirilmesi

Gürültüler imgelerde yüksek frekanslı bileşenler olarak ortaya çıktığından, gürültü giderimi için imgelere bir yumuşatma filtresinin uygulanması gerekmektedir. Yaygın olarak kullanılan yumuşatma filtrelerinden bir türü ortalama filtreleridir. Aritmetik ortalama, ağırlıklı ortalama, geometrik ortalama ve harmonik ortalama bilinen ortalama filtrelerindedir. Ortalama filtrelerinin çalışma prensibi imge üzerinde gezdirilen kayan pencere içerisindeki piksellerin ortalamasını pencerenin merkez pikseline atamaktır. Gürültü gidermeye çalışan yumuşatma filtreleri yüksek frekans bileşeni olan gürültüyü

giderirken yine yüksek frekans bileşeni olan kenar ve detay bilgilerinin de kaybına neden olmaktadır. Bu durumda, gürültüyü giderirken detay bilgisini de koruyan adaptif filtrelerin kullanımı daha uygun olmaktadır. Adaptif filtreler, davranışı (pencere büyüklüğü, vb.) yerel karakteristiklere göre değişen filtrelerdir. Tez çalışması kapsamında, mamografi imgelerine uygulanan filtreler alt bölümler 3.2.1, 3.2.2, 3.2.3, 3.2.4 ve 3.2.5'te anlatılmıştır.

3.2.1. Yerel-olmayan ortalama filtresi

Yerel-Olmayan Ortalama Filtresi (Non-Local Means Filter – NLM), herhangi iki pikselin komşuluklarının birbirine benzerliğine göre pencere boyutunu değiştirirken arama penceresinde türevleri dikkate alarak kenar bilgisini koruyacak ağırlıklandırma fonksiyonunu oluşturmaktadır (Buades vd., 2005). NLM filtresi R_{sim} komşuluk penceresi boyutu, R_{win} arama penceresi boyutu ve σ gürültü seviyesi parametreleri ile çalışmaktadır. NLM filtresinden geçirilmiş iki boyutlu bir $v = \{v(i) | i \in I\}$ imgesi Eşitlik 3.1 ile elde edilmektedir. Eşitlikteki i pikseli R_{win} arama penceresinde bulunan piksellerdir.

$$NL[v](i) = \sum_{j \in I} w(i, j) \cdot v(j) \quad (3.1)$$

Eşitlikteki $w(i, j)$ ifadesi i ve j piksellerinin benzerliğine göre belirlenmiş ağırlık katsayısıdır. Yöntemde benzerlik ölçütü olarak Eşitlik 3.2'de gösterildiği şekilde sıfır ortalama ve σ standart sapmalı Gauss (G_σ) ağırlıklı Öklid uzaklığı $d(i, j)$ kullanılmaktadır. Eşitlikteki $u(N_i)$ ve $u(N_j)$ ifadeleri R_{sim} penceresindeki i ve j piksellerinin yeğinlik değerlerini ifade etmektedir. Benzerliği yüksek piksellerin ağırlığı fazla iken düşük benzerliğe sahip piksellerin ağırlığı daha az olmaktadır.

$$d(i, j) = G_\sigma \cdot \|u(N_i) - u(N_j)\|^2 \cdot R_{sim} \quad (3.2)$$

$d(i, j)$ benzerliğine göre $w(i, j)$ ağırlığı Eşitlik 3.3'te gösterildiği şekilde hesaplanmaktadır. Eşitlikteki h ve Z_i ifadeleri sırasıyla üstel bozulma parametresi ve

normalizasyon sabitini ifade etmektedir. Z_i normalizasyon sabiti Eşitlik 3.4'te gösterildiği şekilde hesaplanmaktadır.

$$w(i, j) = \frac{1}{Z_i} \cdot e^{-\frac{d(i, j)}{h^2}} \quad (3.3)$$

$$Z_i = \sum_j e^{-\frac{d(i, j)}{h^2}} \quad (3.4)$$

3.2.2. Medyan filtresi

Medyan filtresi (median filter) doğrusal olmayan filtreler arasında gürültü giderimi için sıklıkla kullanılan bir filtre türüdür (Gallagher ve Wise, 1981). Bu filtreleme işlemi ile imge üzerinde gezdirilen kayar pencerenin merkez pikseline o penceredeki yeğinlik seviyelerinin ortanca değerinin atanması gerçekleştirilmektedir. Bu şekilde komşuluktaki pikseller arasındaki büyük farklar kaldırılırken kenar bilgisi de korunmaktadır.

3.2.3. Merkez-ağırlıklı medyan filtresi

Rassal değerli dürtü gürültülerinin giderimi için medyan filtrenin geliştirilmesi sonucu elde edilen Merkez-Ağırlıklı Medyan Filtresi (Center-Weighted Median Filter – CWM) ile kayar pencerenin merkez pikselini istenilen oranda ağırlıklandırıldıktan sonra, merkez pikseline o penceredeki yeğinlik seviyelerinin ortanca değerinin atanması gerçekleştirilmektedir (Chen ve Wu, 2001).

3.2.4. Frost filtresi

Frost filtresi (Frost filter), adaptif ve üstel ağırlıklı ortalama alan bir filtre türüdür (Shi ve Fung, 1994). Frost filtreleme ile imge üzerinde gezdirilen kayar pencerenin üstel ağırlığı Eşitlik 3.5'te verildiği şekilde, merkez pikselinin orijine olan uzaklığı ve üstel ağırlık katsayısı kullanılarak hesaplanmaktadır. Üstel ağırlık katsayısı ise Eşitlik 3.6'da verildiği şekilde, yerel standart sapma ve yerel ortalama değerlerinin oranlanması ile hesaplanmaktadır. Eşitliklerdeki (x, y) noktası merkez pikselin koordinatlarını gösterirken

$expWeight$, σ , μ ve α terimleri sırasıyla merkez pikselin üstel ağırlığını, standart sapmasını, ortalamasını ve üstel ağırlık katsayısını ifade etmektedir.

$$expWeight = e^{-\alpha\sqrt{x^2+y^2}} \quad (3.5)$$

$$\alpha = \frac{\sigma}{\mu} \quad (3.6)$$

Bu durumda, filtrelenen pikselin alacağı yeni değer o pikselin filtre merkezine olan uzaklığı, yerel ortalama ve yerel standart sapma değerine göre Eşitlik 3.7 kullanılarak hesaplanmaktadır.

$$m'(t) = K_I \cdot e^{-\alpha|t|} \quad (3.7)$$

$$|t| = \sqrt{x^2 + y^2} \quad (3.8)$$

Eşitlikteki $m'(t)$ ve K_I terimleri sırasıyla pikselin filtreleme sonucu alacağı değeri ve normalizasyon sabitini ifade etmektedir.

3.2.5. Bilateral filtre

Bilateral filtre (Bilateral filter) ağırlıklı ortalama alan bir başka filtre türüdür. Bu filtre ile ağırlık katsayısı filtrelenecek pikselin komşu piksellerine olan uzamsal uzaklık ve yeğlilik seviyesine göre hesaplanmaktadır (Tomasi ve Manduchi, 1998). Herhangi bir I imgesindeki p pikselinin σ_S boyutundaki G_{σ_S} Gauss komşuluğunda yer alan q pikseline göre bilateral filtrelenmiş yeni değeri I_p^{bf} Eşitlik 3.9 ile hesaplanmaktadır (Paris ve Durand, 2006). Eşitliklerdeki W_p^{bf} , I_p ve I_q terimleri sırasıyla normalizasyon sabiti, p pikselinin yeğlilik seviyesi ve q pikselinin yeğlilik seviyesini ifade ederken σ_r komşu pikselin yeğlilik seviyesi farkına göre ne kadar ağırlıklandırıldığını kontrol eden bir parametredir.

$$I_p^{bf} = \frac{I}{W_p^{bf}} \cdot \sum_{q \in S} G_{\sigma_s} \cdot (\|p - q\|) \cdot G_{\sigma_r} \cdot (|I_p - I_q|) \cdot I_q \quad (3.9)$$

$$W_p^{bf} = \sum_{q \in S} G_{\sigma_s} \cdot (\|p - q\|) \cdot G_{\sigma_r} \cdot (|I_p - I_q|) \quad (3.10)$$

3.3. Mamografi İmgelerinin Pektoral Kaslardan Arındırılması

Tüm meme bölgesini içeren mamografi imgelerinden meme kanseri için ROI'lerin çıkartılması, CAD sistemlerin hassasiyetini doğrudan etkileyen önemli bir aşamadır. Mamografi imgelerinde, meme kanseri için ROI'ler diğer bölgelere göre daha parlak olduğundan, bu bölgelerin tespitinde yeğlilik seviyeleri esas alınmaktadır. Bu durumda, benzer yeğlilik seviyelerine sahip olduklarından, MLO çekim mamografi imgelerinde gözükten pektoral kasların varlığı ROI tespitini zorlaştıran en önemli unsurdur. Bu nedenle, ROI tespiti öncesinde mamografi imgelerinin mutlaka pektoral kaslardan arındırılması gerekmektedir. Tez çalışmasında, mamografi imgelerinde pektoral kas tespiti bölge büyütme algoritması kullanılarak gerçekleştirilmiştir.

3.3.1. Bölge büyütme algoritması

Bölge büyütme algoritması, bölge tabanlı bölütleme yöntemlerinden olup, büyütme işlemi için önceden tanımlanmış benzerlik koşulu göz önünde bulundurularak imgedeki tüm pikselleri doğrudan bölgelere ayıran bir algoritmadır (Gonzalez ve Woods, 2014). Bu algoritmanın esası benzer özelliğe sahip piksellerin bir araya getirilmesi ile bölgelerin büyütülmesine dayanmaktadır. Bu amaçla, ilk olarak bir tohum noktası veya tohum noktaları kümesi ve bir benzerlik ölçütü tanımlanmaktadır. Tanımlanan tohum kümesi başlangıç bölgesi olarak kabul edilmekte ve bu kümedeki her bir pikselin 4 veya 8-komşuluğundaki pikseller benzerlik koşulu açısından değerlendirilmektedir. Benzerlik koşulunu sağlayan komşu piksellerin bölgeye dâhil edilmesi ve bu piksellerin komşuluklarının değerlendirilmesi ile devam eden algoritma değerlendirilmeyen hiçbir piksel kalmayınca sonlanmaktadır. Algoritmanın çalışma prensibi Çizelge 3.1'de verilmiştir.

Bölge büyütme algoritmasının ilk adımı olan tohum noktası/kümesinin seçimi algoritmanın başarısını doğrudan etkileyen faktörlerden biridir. Algoritmanın uygulanacağı imge ile ilgili herhangi bir ön bilgi olmadığında tohum noktası/kümesinin rastgele seçimi mümkün olduğu gibi imgede aranan bölge ile ilgili yeğlilik, renk, boyut, konum vb.

bilinen özellikler doğrultusunda tohum noktası/kümesi seçimi de gerçekleştirilebilmektedir.

Çizelge 3. 1. Bölge büyütme algoritmasının çalışma prensibi

İfade	Açıklama
$\bigcup_{i=1}^n R_i = R$	Herhangi bir bölgeye atanmamış hiçbir piksel kalmamaktadır.
$R_i : i . \text{bağlı bölge olarak tanımlanmaktadır.}$ $i = 1, 2, \dots, n$	Bir bölgenin içinde yer alan tüm bölgeler önceden tanımlanmış benzerlik koşuluna göre benzerdir.
$\forall i \text{ için } R_i \cap R_j = \phi$	Herhangi iki bölge birbirinden ayrıktır.
$\forall i \text{ için } P(R_i) = \text{DOĞRU}$	İmgede bölütlenen bölgeler benzerlik koşulunu DOĞRU olarak sağlayan bölgelerdir.
$\forall R_i \text{ ve } R_j \text{ komşu bölgeleri için}$ $P(R_i \cup R_j) = \text{YANLIŞ}$	Herhangi iki komşu bölge birbirinden farklıdır.

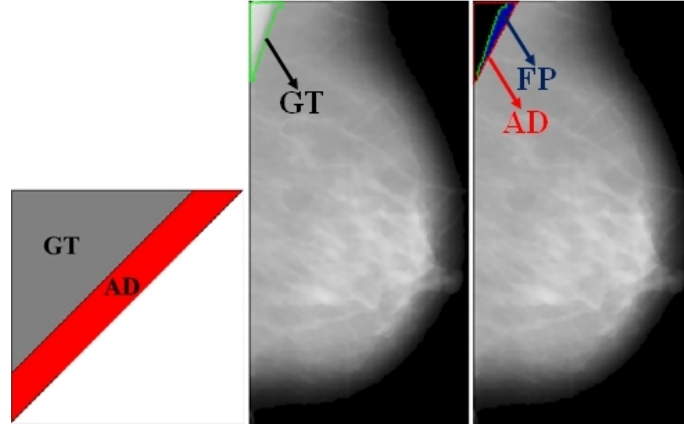
Algoritmanın başarısını etkileyen faktörlerden bir diğeri ise benzerlik koşulunun belirlenmesidir. Gri seviye imgelerde piksel yeğlilikleri veya uzamsal özellikleri benzerlik ölçütü olarak kullanılabilir.

Tez çalışması kapsamında, bölge büyütme algoritması tabanlı özgün bir algoritma oluşturularak mamografi imgelerinde pektoral kas tespiti gerçekleştirilmiştir. Pektoral kas tespiti çalışması alt bölüm 8.2.2’de ayrıntılı olarak anlatılmıştır.

3.3.2. Pektoral kas tespiti doğruluk ölçütü

Literatürde yer alan çalışmalarda iki farklı değerlendirme ölçütünün yer aldığı görülmüştür (Abdellatif vd., 2012; Camilus vd., 2011). Bu ölçütlerden ilki, *The Area Normalized Error*, algoritma sonucu tespit edilen pektoral kas bölgesi ile referans pektoral kas bölgesi arasındaki yanlış-pozitiflik ve yanlış-negatiflik oranlarını incelemektedir.

Ölçüte göre, Şekil 3.1’de gösterildiği gibi, algoritmanın pektoral kas olarak belirlediği piksel eğer gerçekten pektoral kas değil ise, o piksel yanlış-pozitif (false-positive – FP) olarak değerlendirilmektedir.



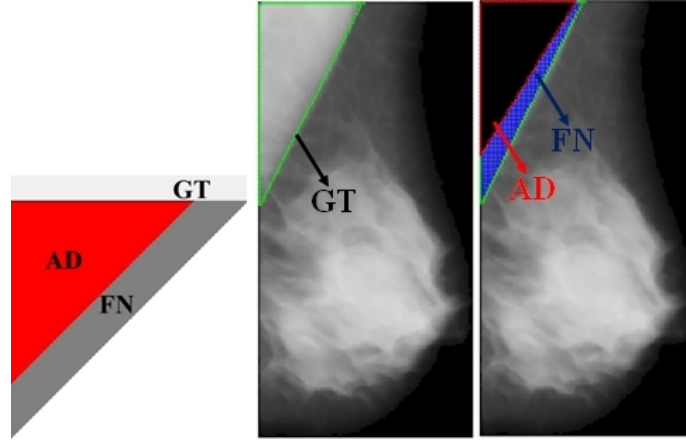
Şekil 3. 1. *The Area Normalized Error* - Yanlış-pozitif tanımı

Şekil 3.2’den de görüldüğü üzere, gerçekten pektoral kas olan bir piksel algoritma tarafından tespit edilememiş ise, o piksel yanlış-negatif (false-negative – FN) olarak değerlendirilmektedir. Şekillerdeki AD ve GT ifadeleri sırasıyla algoritmanın pektoral kas olarak belirlediği pikselleri ve referans bilgiye göre pektoral kas olan pikselleri göstermektedir.

The Area Normalized Error kriterine göre tanımlanan yanlış-pozitif ve yanlış-negatif tanımlarının matematiksel ifadeleri sırasıyla Eşitlik 3.11 ve 3.12’de verilmiştir. Eşitliklerde A , p , $B_{alg}(i)$ ve $B_{gro}(i)$ terimleri sırasıyla referans pektoral kas bölgesinin alanını, pektoral kas satır sayısını, algoritma sonucu tespit edilen pektoral kas bölgesinin i . satırındaki sınır noktasını ve referans pektoral kas bölgesinin i . satırındaki sınır noktasını ifade etmektedir.

$$FP = \frac{I}{A} \cdot \sum_{i=1}^p \max[0, B_{alg}(i) - B_{gro}(i)] \quad (3.11)$$

$$FN = \frac{I}{A} \cdot \sum_{i=1}^p \max[0, B_{gro}(i) - B_{alg}(i)] \quad (3.12)$$



Şekil 3. 2. *The Area Normalized Error- Yanlış-negatif tanımı*

Bu tanımlamalar doğrultusunda, *The Area Normalized Error* ölçütüne göre Çizelge3.2’de verilen ifadeler değerlendirilmektedir.

Çizelge 3. 2. *The Area Normalized Error – Değerlendirme Kriterleri*

The Area Normalized Error – Değerlendirme Kriterleri

- 1.FP_m (ortalama yanlış-pozitiflik)
 - 2.FN_m (ortalama yanlış-negatiflik)
 3. FP < 0,05 & FN < 0,05
 - 4.min(FP, FN) < 0,05 & 0,05 < max(FP, FN) < 0,10
 - 5.min(FP, FN) < 0,05 & max(FP, FN) > 0,10
 6. 0,05 < FP < 0,10 & 0,05 < FN < 0,10
 7. 0,05 < min(FP, FN) < 0,10 & max(FP, FN) > 0,10
 8. FP > 0,10 & FN > 0,10
-

Bu ölçütlerden ikincisi, Hausdorff uzaklığı, tespit edilen pektoral kas bölgesi (A) ile referans pektoral kas bölgesinin (B) her bir pikseli için, diğer bölgede minimum uzaklığı sağlayan pikseli bulup, daha sonra bu uzaklık arasından en büyük olanı seçmektedir (Dik ve Aktaş, 2014).

$$H(A, B) = \max\{h(A, B), h(B, A)\} \quad (3.13)$$

$$h(A, B) = \max_{a \in A} \{\min_{b \in B} \{d(a, b)\}\} \quad (3.14)$$

4. MAMOGRAFİ İMGELERİNDE MEME DOKU TİPİ TESPİTİ

Radyoloji uzmanları meme doku tipinin meme kanseri riski ile doğrudan ilişkili olduğunu ve meme doku yoğunluğunun artması ile kanser riskinin de arttığını belirtmektedirler. Meme doku yoğunluğu arttıkça mamografi imgesi daha parlak olarak gözükmetedir. Bu durumda, CAD sistemleri gerek pektoral kas tespiti gerekse ROI tespiti için yeğlilik-tabanlı yaklaşımlar kullandıklarından, sistem performansını yükseltmek amacı ile meme doku tipine bağımlı CAD sistemlerinin tasarlanması gerekmektedir. Böylece, her bir doku tipine sahip imgeler kanser teşhisi için kendi aralarında değerlendirilecek ve doku tipi farklılıklarından kaynaklı hataların en aza indirgenmesi sağlanabilecektir. Ayrıca, her bir doku tipi için kendi tiplerine uygun pektoral kas tespiti, ROI tespiti ve öznitelik çıkarımı algoritmaları kullanarak, CAD sistemlerinin hassasiyeti daha da arttırılabilecektir.

Bu bilgiler doğrultusunda, tez çalışması kapsamında hazırlanan CAD sisteminde, mamografi imgelerinin ön işlenmesi aşamasından sonra meme doku tipi tespiti gerçekleştirilmiştir. Öznitelik çıkarımı aşaması tüm doku tipleri için aynı olup, meme doku tipi yalnızca alt bölüm 8.1.4’de anlatıldığı şekilde sınıflandırma aşamasında kullanılmıştır.

MIAS veritabanı kullanılarak gerçekleştirilen meme doku tipi tespiti çalışması için çıkartılan öznitelikler alt bölüm 6.4’te, bu özniteliklerin sınıflandırılması süreçleri ise alt bölüm 8.2.3’te ayrıntılı olarak anlatılmıştır.

5. MAMOGRAFİ İMGELERİNDE ŞÜPHELİ BÖLGE TESPİTİ

Radyoloji uzmanları mamografi imgelerinde görülen anomalileri, anomalinin yoğunluğunu, şeklini, kontur yapısını, boyutunu, anomali varlığına eşlik eden bulguları ve eski filmlere göre memede görülen değişiklikleri göz önünde bulundurarak incelemekte ve hastalık sınıflandırmasına karar vermektedirler. Bir mamografi imgesi meme dokusunun röntgen görüntüsüdür. Meme dokusundaki farklılıklar, mamografi imgesinin gri tonlamasındaki farklılık olarak tanımlanmaktadır. Mamografi imgelerinde tüm meme görüntüsü yerine sadece anomali içerme şüphesi taşıyan bölgelerin CAD sistemlerinde kullanımı öznelik çıkarımı aşamasında gerek güvenilirlik gerekse zaman açısından daha verimli olacaktır. Tez çalışmasında şüpheli bölge tespiti için kullanılan Chan-Vese Aktif Kontur Modeli (Active Contour Model Without Edge) ve bu yöntemin temelini oluşturan Aktif Kontur Model – Yılan (Active Contour Model – Snake) takip eden alt bölümlerde anlatılmıştır.

5.1. Aktif Kontur Model – Yılan

Aktif kontur modelleme, imge üzerinde ilgi alanlarının sınırlarına gelindiğinde duran kapalı elastik bir şerit oluşturulması olarak ifade edilmektedir (Kass vd., 1988). Aktif kontur modelleme algoritmasına göre, (x, y) koordinat değerleri ile gösterilen, parametrik tanım kümesi elemanı (s) ve zamana (t) bağlı olarak modellenen iki boyutlu bir kontur (yılan) Eşitlik 5.1 ile ifade edilmektedir (Kass vd., 1988).

$$\vec{v}[s] = (x(s, t), y(s, t)) \quad , \quad 0 \leq s \leq l \quad (5.1)$$

Aktif kontur modelleme algoritmasında, ilgi alanının iç veya dış kısmına bir yılan çizilmekte ve yılan, ilgi alanının şeklini alana kadar iterasyonlar yapılarak hareket ettirilmektedir. Algoritmada, yılanın imge üzerindeki hareketi enerji minimizasyonu ile sağlanmaktadır (Kass vd., 1988; Filho vd., 2014). Bu durumda, yılanın imge üstündeki konumu Eşitlik 5.2’de gösterildiği gibi enerjilerin toplamı olarak ifade edilmektedir.

$$E(\vec{v}) = E_{img}(\vec{v}) + E_{int}(\vec{v}) + E_{ext}(\vec{v}) \quad (5.2)$$

Eşitlikte, $E_{img}(\vec{v})$ görüntü enerjisidir ve yılanı kenar gibi detaylara yönlendirmektedir. $E_{int}(\vec{v})$, yılanın şeklini kontrol eden iç enerji iken $E_{ext}(\vec{v})$ ise yılanı ilgi alanında yerel minimuma yönlendiren dış enerjidir ve imge üzerinden gelen imge gradyan şiddetine bağlıdır (Işıkçı ve Duru, 2015). Yüksek imge gradyanının olduğu ilgi alanı sınırlarına ulaştığında yılanın hareketi durmakta ve algoritma sonlanmaktadır.

5.2. Chan-Vese Aktif Kontur Modeli

Chan ve Vese tarafından önerilen bölgesel aktif kontur modelleme yönteminin amacı gradyan bilgisini kullanmadan bir imgedeki ilgi alanını arka plandan ayırmaktır (Chan ve Vese, 2001). Bu yöntem ile başlangıç konturu (kapalı eğri) olarak bir maske seçilmekte ve enerji minimizasyonu sağlanacak şekilde maskenin küçültülmesi ve genişletilmesi ile sonuca ulaşılmaktadır. Chan ve Vese yöntemlerinde enerji fonksiyonunun oluşturulmasında Mumford-Shah yöntemi ve düzey kümesi analizini esas almışlardır (Chan ve Vese, 2001; Osher ve Sethian, 1988). Mumford-Shah yöntemine göre, C konturu ile ikiye ayrılmış olan u_0 resmine Eşitlik 5.3'te verilen $F^{MS}(u, C)$ enerji fonksiyonu tanımlanmaktadır (Mumford ve Shah, 1989).

$$\inf \left\{ F^{MS}(u, C) = \int_{\Omega} (u - u_0)^2 dx dy + \int_{\frac{u}{C}} |\nabla u|^2 dx dy + v|C| \right\} \quad (5.3)$$

Eşitlikte $u = \forall \Omega_i$ bölgesi için sabit c_i alındığında, enerji fonksiyonu Eşitlik 5.4'te gösterildiği hale indirgenmektedir. (Mumford ve Shah, 1989).

$$E^{MS}(u, C) = \sum_i \int_{\Omega} (u - c_i)^2 dx dy + v|C| \quad (5.4)$$

$$c_i = ortalama(u_0), u_0 \subset c_i$$

Mumford-Shah yöntemi doğrultusunda, Chan-Vese algoritmasında C başlangıç konturu ile ikiye ayrılmış u_0 imgesine tanımlanan $F(c_1, c_2, C)$ enerji fonksiyonu Eşitlik 5.5'de verilmiştir (Chan ve Vese, 2001).

$$\begin{aligned}
F(c_1, c_2, C) &= \mu \cdot \text{Uzunluk}(C) + \nu \cdot \text{Alan}(i\check{c}(C)) + \\
&\lambda_1 \int_{i\check{c}(C)} |u_0(x, y) - c_1|^2 dx dy + \\
&\lambda_2 \int_{dis(C)} |u_0(x, y) - c_2|^2 dx dy
\end{aligned} \tag{5.5}$$

Eşitlikte, μ, λ_1 ve λ_2 terimleri pozitif sabitler ve ν terimi C 'ye bağlı bir sabit olup, nümerik hesaplamalarda $\lambda_1 = \lambda_2 = 1$ ve $\nu = 0$ olarak kullanılmaktadır. Bu durumda, $F(c_1, c_2, C)$ enerji fonksiyonu Eşitlik 5.6'ya indirgenmektedir. Konturun içinde kalan alanın ortalaması c_1 ve konturun dışındaki alanın ortalaması c_2 Eşitlik 5.7 ile hesaplanmaktadır.

$$\begin{aligned}
F(c_1, c_2, C) &= \mu \cdot \text{Uzunluk}(C) + \int_{i\check{c}(C)} |u_0(x, y) - c_1|^2 dx dy + \int_{dis(C)} |u_0(x, y) - c_2|^2 dx dy \\
&= \mu \cdot \text{Uzunluk}(C) + F_1(C) + F_2(C)
\end{aligned} \tag{5.6}$$

Eşitlik 5.6'da verilen $F(c_1, c_2, C)$ enerji fonksiyonunda ilk terim, en son bulunan konturun gürültü nedeniyle küçük bir alan ile birleşmesini engelleyen regülarizasyon terimidir. İkinci terim, konturu küçülmeye zorlayan, üçüncü terim ise konturu genişlemeye zorlayan kestirim terimleridir.

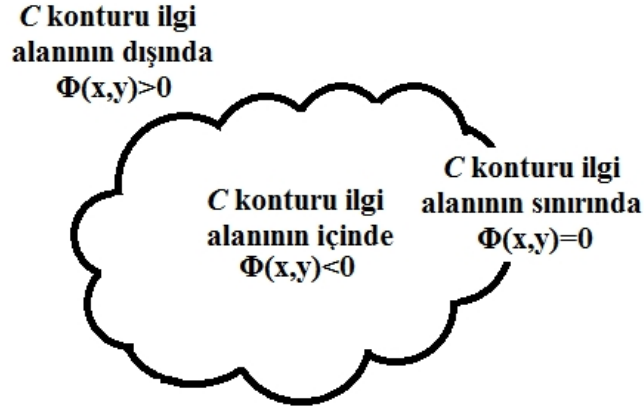
Chan-Vese algoritmasında kullanılan düzey kümesi $\phi(x, y)$ fonksiyonu kontur değerlerini göstermek için uygulanmakta ve C kapalı kontur eğrisi Eşitlik 5.7'de gösterildiği şekilde formüle edilmektedir.

$$C = \{(x, y) : \phi(x, y) = 0\}, \forall (x, y) \in u_0 \tag{5.7}$$

Düzyer kümesi fonksiyonuna göre konturun zamanla gelişimi Eşitlik 5.8 ile ifade edilmektedir.

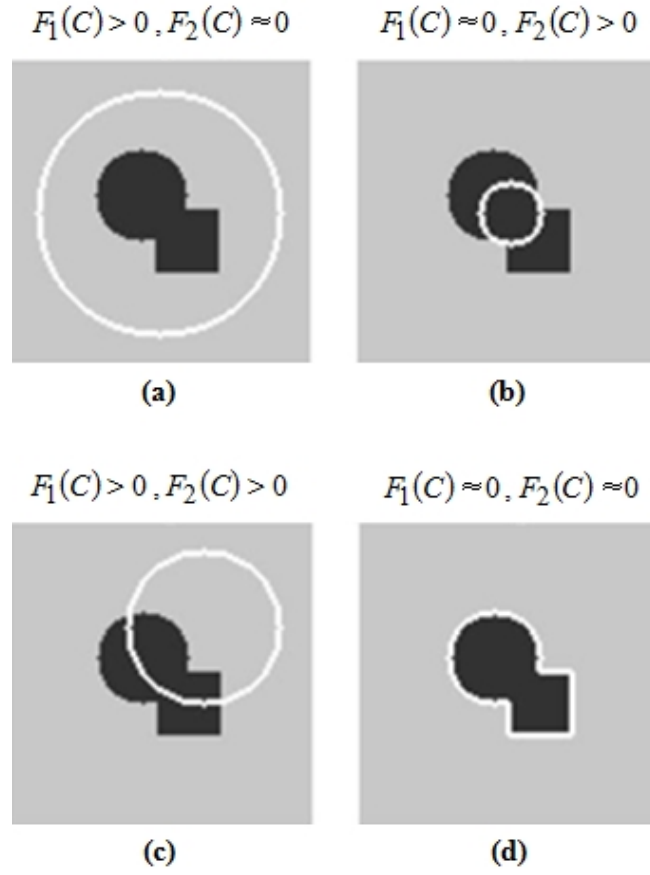
$$\frac{\partial C}{\partial t} = \frac{\partial \phi(x, y)}{\partial t} \tag{5.8}$$

İlgi alanı ile C konturunun birbirine göre konumsal durumuna göre düzey kümesi fonksiyonu pozitif, negatif veya sıfır olabilmektedir. Şekil 5.1'den de görüleceği gibi $\phi(x,y) = 0$ iken C konturunun ilgi alanının sınırlarında, $\phi(x,y) < 0$ iken C konturunun ilgi alanının içinde ve $\phi(x,y) > 0$ iken C konturunun ilgi alanının dışında olduğu anlaşılmaktadır.



Şekil 5. 1. Kontur gelişimi

Chan-Vese algoritmasının çalışma prensibi Şekil 5.2'de görselleştirilmiştir. Şekilde siyah bölgeler ilgi alanını göstermekte ve (-1) değeri ile temsil edilmektedir. Gri bölgeler, imgede ilgi alanı dışında kalan bölgeleri göstermekte ve (+1) değeri ile temsil edilmekte ve beyaz eğri ise konturu belirtmektedir. Şekilde gösterilen 1. durumda, kontur ilgi alanının dışında olduğundan, Eşitlik 5.6 kullanılarak $F_1(C) > 0$ ve $F_2(C) \approx 0$ hesaplanmaktadır. Bu durumda, Chan-Vese algoritması konturu küçülmeye zorlamaktadır. 2. durumda, kontur ilgi alanının içinde olduğundan $F_1(C) \approx 0$ ve $F_2(C) > 0$ hesaplanmakta ve kontur genişlemeye zorlanmaktadır. 3. durumda ise kontur hem ilgi alanını hem de dışını kapsadığı için $F_1(C) > 0$ ve $F_2(C) > 0$ hesaplanmaktadır. Bu durumda, Chan-Vese algoritması konturu hem küçülmeye hem de genişlemeye zorlamaktadır. Son olarak, 4. durumda, kontur ilgi alanı sınırlarında olduğundan $F_1(C) \approx 0$ ve $F_2(C) \approx 0$ hesaplanmakta ve Chan-Vese algoritması sonlandırılmaktadır. İlk üç durumda enerji fonksiyonu dengelenmediği yani sıfır olmadığı için algoritma konturu küçültürken ya da genişleterek çalışmaya devam etmekte; son durumda ise $F(c_1, c_2, C) = 0$ dengesi sağlandığından algoritma sonlandırılmaktadır. İterasyonlar ile ilerleyen algoritma $F(c_1, c_2, C) = 0$ dengesi sağlanmadığı takdirde maksimum iterasyon sayısına ulaşıldığında sonlandırılmaktadır.



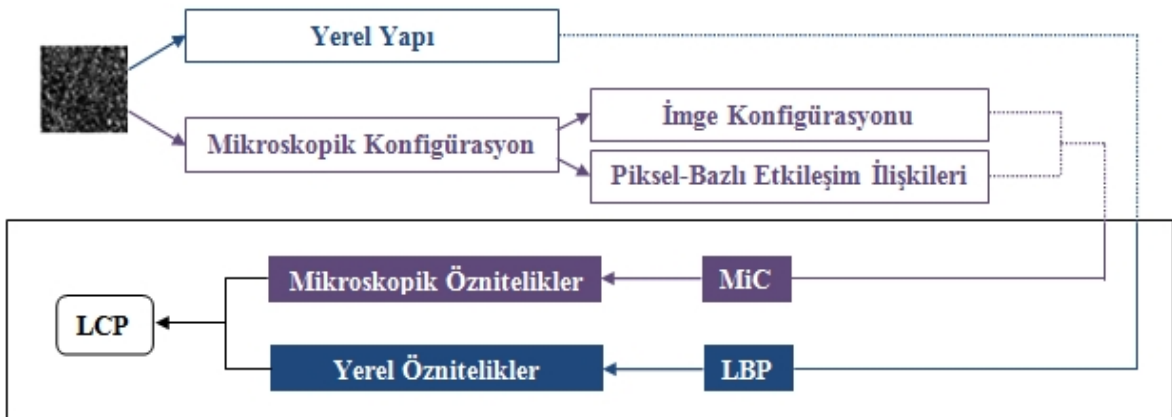
Şekil 5. 2. Chan-Vese algoritması çalışma prensibi (Chan ve Vese, 2001): (a) 1. durum (b) 2. durum (c) 3. durum (d) 4. Durum

6. MAMOGRAFİ İMGELERİNDEN ÖZİNİTELİK ÇIKARTIMI

Öznitelik çıkarımı aşamasında sınıf içi değişimi en az sınıflar arası değişimi en fazla tutan, en az sayıda özellik çıkarmak esas alınmaktadır. Sınıflandırma başarısının ne kadar yüksek olacağı, çıkartılan öznitelik vektörlerinin sınıfları ne kadar iyi temsil edebildiğine bağlı olduğundan bu aşama oldukça önem arz etmektedir. Mamografi imgelerinden öznitelik vektörlerinin elde edilmesi sürecinde istatistiksel yöntemler, model-tabanlı yöntemler ve piksel karakteristiği ya da imgenin frekans spektrumuna göre doku öznitelikleri hesaplayan sinyal işleme yöntemleri kullanılabilirdiği gibi şekil, dağılım, boyut, kontur, yoğunluk, parlaklık gibi mamografik özniteliklerden de faydalanılabilmektedir. Bu doğrultuda, tez çalışmasının çeşitli aşamalarında Yerel Konfigürasyon Örüntüsü (Local Configuration Pattern – LCP) algoritması ile elde edilen öznitelikler, bazı istatistiksel ve frekans-düzlemi öznitelikleri, doku öznitelikleri ve mamografik öznitelikler kullanılmıştır.

6.1. Yerel Konfigürasyon Örüntüsü

LCP, Şekil 6.1’de gösterildiği gibi iki boyutlu bir imgenin yerel bölge bilgisi ile örüntü sıklıklarının harmanlayan bir algoritma kullanarak öznitelik vektörü oluşturmaktadır [13].



Şekil 6. 1. LCP algoritması iş akış şeması

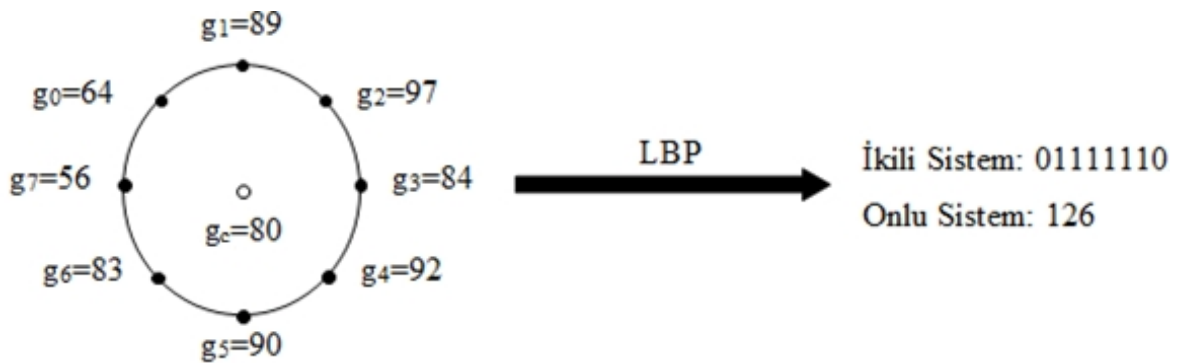
LCP algoritmasında yerel bölge bilgisi, Yerel İkili Örüntü (Local Binary Pattern – LBP) algoritması ile elde edilmektedir. Ojala vd. (2002) tarafından önerilen LBP, son yıllarda öznelik çıkarımı aşamasında (Ahonen vd., 2006; Zhang vd., 2010; Zhang vd., 2005) sıklıkla kullanılan gri seviye, dönme açısından bağımsız bir imge gösterim yöntemidir.

Herhangi bir I imgesinin LBP gösterimi, Eşitlik 6.1’de gösterildiği şekilde I imgesinin her bir pikselinin, R yarıçaplı dairesel komşuluğundaki P tane piksel ile karşılaştırılması ile elde edilmektedir. Eşitlikteki g_i ve g_c ifadeleri sırası ile i . komşuluk pikselinin ve komşuluktaki merkez pikselin yeğinlik değerleri olarak tanımlanmaktadır. Burada $u(x)$ birim basamak fonksiyonudur.

$$LBP(P, R) = \sum_{i=0}^{P-1} u(g_i - g_c) \cdot 2^i \quad (6.1)$$

$$u(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Gri seviye bir imgenin LBP gösteriminin elde edilmesi Şekil 6.2’de örneklendirilmiştir.



Şekil 6. 2. Gri seviye bir imgenin LBP gösterimi

Dairesel ikili gösterimde en fazla iki bit değişimi olan örüntüler bir biçimli olarak tanımlanmaktadır. Dönme açısından bağımsız örüntüler, her bir bitin dairesel olarak, en büyük değer son eleman olacak şekilde, minimum değere döndürülmesi ile elde

edilmektedir. Bu bilgiler doğrultusunda, I imgesinde LBP algoritması kullanılarak elde edilen dönme açısından bağımsız bir biçimli iki boyutlu ikili örüntüler Eşitlik 6.2 ile hesaplanmaktadır. Eşitlikteki U ifadesi bir biçimlilik ölçütünü belirtmektedir.

$$LBP^{riu2}(P, R) = \begin{cases} \sum_{i=0}^{P-1} u(g_i - g_c) & , \quad U(LBP(P, R)) \leq 2 \\ P+1 & , \quad U(LBP(P, R)) > 2 \end{cases} \quad (6.2)$$

LBP-tabanlı yöntemler ile imge gösterimi yeğlilik değişimlerine karşı başarı gösterdikleri halde, gri-seviye yeğlilik değerlerini ikili seviyeye nicemleme aşamasında imge görünümü ve piksel ilişkilerini kodlama kapasitesi azalmaktadır. Bu durumu ortadan kaldırmak amacı ile Eşitlik 6.3 kullanılarak komşu piksellerinin ortalaması μ olan her bir bölgenin yerel varyansı hesaplanmaktadır. Sonuç olarak dönme açısından bağımsız bir biçimli yerel öznitelikleri (O) Eşitlik 6.4 ile elde edilmektedir.

$$VAR = \frac{1}{P} \cdot \sum_{i=0}^{P-1} (g_i - \mu)^2 \quad (6.3)$$

$$O = \frac{LBP^{riu2}}{VAR} \quad (6.4)$$

LBP vektörleri aynı çıkmış örüntülerin varyanslarının da eşit olması durumunda, örüntülerden, dokusal özelliklerine bağlı olarak farklı konfigürasyonda da olsa, aynı öznitelik vektörleri elde edilmektedir. Bu durumu ortadan kaldırmak amacı ile örüntülerin mikroskobik konfigürasyonu modellenmektedir. Bu amaçla, öncelikle, komşu piksel yeğlilik değerleri ağırlıklandırılarak her bir örüntünün merkez pikselini tanımlayacak doğrusal modeller oluşturulmaktadır. Modelleme hatasını en aza indirgeyen en uygun ağırlık vektörü (A_L) en küçük kareler yöntemi ile hesaplanmaktadır. Ağırlık vektörünü dönme açısından bağımsız kılmak için 1-boyutlu Fourier dönüşümü (H_L) alınmakta ve dönüşümün büyüklüğü ($|H_L|$) mikroskobik öznitelik vektörü olarak hesaplanmaktadır. Sonuç olarak, q tane örüntüye sahip I imgesinin LCP öznitelik vektörü Eşitlik 6.5'te gösterildiği şekilde yerel öznitelikler ve mikroskobik öznitelikler birleştirilerek elde edilmektedir.

$$LCP = \left[\left[H_0 | O_0 \right]; \left[H_1 | O_1 \right]; \dots; \left[H_{q-1} | O_{q-1} \right] \right] \quad (6.5)$$

Bu şekilde imgeler hem doku özelliklerine hem de örüntü sıklıklarına göre değerlendirilmekte ve öznitelik vektörü çıkartılmaktadır.

6.2. İstatistiksel Öznitelikler

Doğrudan imgelerden elde edilen istatistiksel özniteliklerin sınıflandırma başarısındaki olumlu etkisi bilinmektedir (Woods vd., 1993). Ayrıca, LCP algoritmasının da mamografi imgelerinden ayırt ediciliği yüksek öznitelikler çıkarttığı belirtilmiştir (Ergin ve Kılınc, 2014). Bu nedenle, doğrudan imgelerden istatistiksel öznitelikler çıkarmak yerine, LCP öznitelik vektörlerinden istatistiksel öznitelikler elde edilmiştir.

İngeler, nesnelere yansıyan enerjinin algılanması ile oluşmaktadır. Dolayısıyla, sınıflandırma çalışmalarında imgelerin enerjisi ayırt edilebilirliği artırıcı etki göstermektedir. Bu nedenle, LCP öznitelik vektörlerinin enerjileri, ortalama enerjileri ve enerji varyansları hesaplanmıştır.

Bir dağılımın ortalaması, o dağılımdaki her bir elemanı temsil eden tek bir orta değeri tek bir sayı ile ifade eden bir istatistik ölçüsü olduğundan, LCP öznitelik vektörlerinin ortalamaları hesaplanmıştır. Yanı sıra, LCP öznitelik vektörlerinin maksimum ve minimum değerleri de kullanılmıştır.

Bir dağılımın varyansı, ortalama etrafındaki ikinci momenti, o dağılımdaki her bir elemanın dağılım ortalamasından ne kadar farklı olduğunu göstermektedir. Herhangi bir dağılımın varyansının düşük olması o dağılımdaki elemanların birbirine benzer olduğunu, yüksek olması ise elemanların birbirinden farklı olduğunu göstermektedir. Mamografi imgeleri bir dağılım olarak düşünüldüğünde, imgelerdeki kanserli bölgeler ile normal bölgelerin varyansının yüksek olduğu gerekçesi ile LCP öznitelik vektörlerinin varyans değerleri öznitelik olarak kullanılmıştır.

Standart sapma, bir dağılımın elemanlarının dağılımın ortalamasına göre nasıl bir yayılım gösterdiğinin ölçüsüdür. Standart sapmanın düşük çıkması o dağılımdaki

elemanların birbirine benzer olduğunu, yüksek olması ise elemanların birbirinden farklı olduğunu göstermektedir. Ortalamaları eşit olabilecek LCP öznitelik vektörlerinin birbirlerinden ayrılabilirliğinin artırılabilmesi için LCP öznitelik vektörlerinin standart sapmaları öznitelik olarak kullanılmıştır. Standart sapmanın ortalamaya oranı olarak hesaplanan alan-tanımlayıcısı da LCP öznitelik vektörlerinden elde edilen bir başka istatistiksel özniteliktir.

Çarpıklık (skewness), ortalama etrafındaki üçüncü moment, bir dağılımın ortalama etrafındaki asimetri ölçüsü olarak tanımlanmaktadır (Anonim, 2017). Normal dağılımların (ortalama=0, varyans=1) çarpıklık değeri sıfırdır. Negatif çarpıklık dağılımın ortalamasının soluna, pozitif çarpıklık ise dağılımın ortalamasının sağına doğru yayıldığını göstermektedir. Çarpıklığında varyans gibi bir değişinti ölçütü olduğu gerekçesi ile LCP öznitelik vektörlerinin çarpıklık değerleri öznitelik olarak kullanılmıştır.

Basıklık, ortalama etrafındaki dördüncü moment, bir dağılımın uzaksal-yatık olduğunun ölçüsü olarak tanımlanmaktadır (Anonim, 2017). Normal dağılımların basıklık değeri üç olarak hesaplanmaktadır. Uzaksal-yatıklığı normal dağılımdan fazla olan dağılımların basıklığı üçten fazla, az olan dağılımların basıklığı ise üçten az olarak hesaplanmaktadır. Basıklığında varyans ve çarpıklık gibi bir başka değişinti ölçütü olduğu gerekçesi ile LCP öznitelik vektörlerinin basıklık değerleri öznitelik olarak kullanılmıştır.

Radyoloji uzmanları, bir mamografi imgesinde, kanserli bölgelerin sağlıklı bölgelere göre ve iyi huylu kanserli bölgelerin kötü huylu kanserli bölgelere göre daha düzensiz bir dağılım gösterdiklerini belirtmektedirler. Bu durum, istatistik biliminde entropiye karşılık gelmektedir. Bu sebeple mamografi imgelerinin düzensizliğini öznitelik olarak değerlendirmek amacı ile LCP öznitelik vektörlerinin entropi değerleri hesaplanmıştır.

Tez çalışması kapsamında kullanılan istatistiksel öznitelikler ve $N \times 1$ boyutlu X vektöründen elde edilmesinin matematiksel ifadeleri Çizelge 6.1’de verilmiştir.

Çizelge 6. 1. İstatistiksel öznitelikler ve matematiksel ifadeleri

İstatistiksel Öznitelikler	Matematiksel İfadeleri
Enerji	$\sum_{i=1}^N X_i^2$
Ortalama	$\mu = \frac{1}{N} \cdot \sum_{i=1}^N X_i$
Varyans	$var = \frac{1}{N-1} \cdot \sum_{i=1}^N (X_i - \mu)^2$
Maksimum	$maksimum \{X_i i = 1, 2, \dots, N\}$
Minimum	$minimum \{X_i i = 1, 2, \dots, N\}$
Standart Sapma	$\sigma = \sqrt{var}$
Çarpıklık	$\frac{1}{\sigma^3} \cdot \sum_{i=1}^N (X_i - \mu)^3$
Basıklık	$\frac{1}{\sigma^4} \cdot \sum_{i=1}^N (X_i - \mu)^4$
Alan-Tanımlayıcısı (Ahonen vd., 2006)	$\frac{\sigma}{\mu}$
Ortalama Enerji	$\mu_{Energy} = \frac{1}{N} \cdot \sum_{i=1}^N X_i^2$
Enerji Varyansı	$\frac{1}{N-1} \cdot \sum_{i=1}^N (X_i^2 - \mu_{Energy})^2$
Entropi	$\sum_{i=1}^N p(X_i) \cdot \log_2 p(X_i)$

6.3. Frekans-Düzlemi Öznitelikleri

Bir mamografi imgesinde kanser teşhisi açısından önemli bir öznitelik olan parlaklığın imgedeki değişimleri, en iyi şekilde imgenin uzamsal frekansında

gözlemlenmektedir. Bu nedenle, frekans-düzlemi özneliklerinin mamografi imgelerini daha iyi temsil edebileceği gerekçesi ile imgeler Daubechies1 dalgacığı kullanılarak ikinci seviyede 2-boyutlu ayrık dalgacık dönüşümüne tabi tutulmuştur. İmgelerin enerji değerlerinin parlaklık değerleri ile doğrudan ilişkili ve parlaklığın da meme kanseri teşhisinde etkili bir öznelik olduğu gerekçesi ile dalgacık dönüşümü sonucu elde edilen her bir alt bandın enerjisi frekans-düzlemi öznelikleri olarak kullanılmıştır.

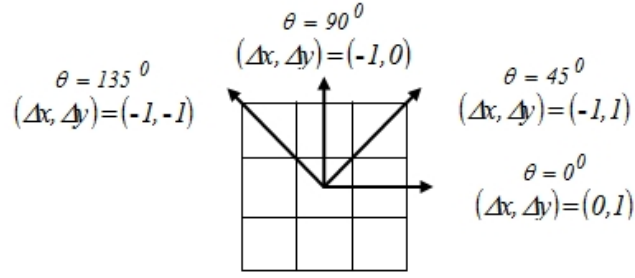
6.4. Doku Öznelikleri

Doku, bir desen ya da desenlerin bir bölge üzerinde tekrarlanmasıdır (Demirhan ve Güler, 2010). Bu durumda, doku, bir bölge içindeki piksel yoğunluğunun istatistiksel özneliklerinin bir kümesi olarak tanımlanabilmektedir (Severoğlu, 2016). Bu bilgiler doğrultusunda, Haralick vd. (1973) imgelerden GLCM matrislerinin oluşturulması ve bu matrislerden çıkartılan istatistiksel özneliklerin doku öznelikleri olarak kullanılmasını önermişlerdir.

GLCM, $P(i, j | d, \theta)$ ikinci dereceden bileşik durum olasılık yoğunluk fonksiyonunun hesaplanması ile oluşturulmaktadır (Haralick vd., 1973). $P(i, j | d, \theta)$ fonksiyonu, N - gri seviyeli bir I imgesinin d uzaklığı ve θ yönünde komşu iki pikseli arasında i gri-seviyesinden j gri-seviyesine geçiş olasılığını vermekte ve Eşitlik 6.7'de gösterildiği şekilde hesaplanmaktadır (Çalışkan ve Ergen, 2014). Eşitlikte, (x, y) ve (Δ_x, Δ_y) ikili terimleri sırasıyla pikselin imgedeki uzamsal koordinatlarını ve d uzaklığının θ açısına göre bileşenlerini ifade etmektedir; $I(x, y)$ referans piksel ve $I(x + \Delta_x, y + \Delta_y)$ de karşılaştırma pikseli olarak tanımlanmaktadır.

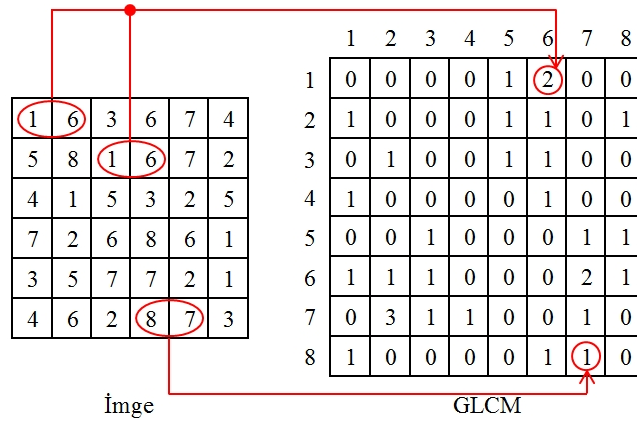
$$P(i, j | d, \theta) = \sum_{i=1}^N \sum_{j=1}^N \begin{cases} 1 & , & I(x, y) = i \text{ ve } I(x + \Delta_x, y + \Delta_y) = j \\ 0 & , & \text{aksi takdirde} \end{cases} \quad (6.7)$$

Eşitlikte de görüldüğü gibi, her bir (d, θ) ikilisi için $N \times N$ boyutlu bir adet GLCM $(P(i, j | d, \theta))$ oluşturulmaktadır. GLCM oluşumu için sıklıkla kullanılan yönler $d = 1$ uzaklığı için Şekil 6.3'te gösterilmiştir.



Şekil 6. 3. GLCM gösterimi

$N = 8$ -gri seviyeli bir imgenin $d = 1$ ve $\theta = 0^\circ$ alınarak hesaplanmış 8×8 boyutundaki eş-oluşum matrisi Şekil 6.4'te gösterilmiştir. Şekilde, imge matrisinde bulunan (1,6) çifti iki kez tekrarlandığından GLCM matrisinin (1,6) koordinatındaki elemanı "2" olarak hesaplanmaktadır. Benzer şekilde, (8,7) çifti imge matrisinde 1 kez tekrarlandığı için GLCM matrisinin (8,7) koordinatındaki elemanı "1" olarak hesaplanmaktadır. GLCM matrisinde "0" olarak görülen elemanlar, ilgili koordinatların imge matrisinde çift olarak yer almadığını göstermektedir. Daha yüksek gri-seviyeye sahip imgelerde, hesap yükünü azaltmak amacı ile imgeler öncelikle 8-gri seviyeye sahip olacak şekilde nicemlenmekte ve sonrasında GLCM oluşturulmaktadır. Bu şekilde, gri-seviye sayısından bağımsız olarak, 8×8 boyutlarında eş-oluşum matrisleri elde edilmektedir.



Şekil 6. 4. GLCM oluşumu

Tez çalışması kapsamında, mamografi imgelerinden elde edilen eş-oluşum matrislerinden Haralick vd. (1973), Soh ve Tsatsoulis (1999) ve Clausi (2002) tarafından önerilen doku öznelikleri çıkartılmıştır. Bu öznelıklar, matematiksel ifadeleri ile birlikte Çizelge 6.2'de verilmiştir.

Çizelge 6. 2. Doku öznelikleri ve matematiksel ifadeleri

Doku Öznelikleri	Matematiksel İfadeleri
$P(i, j): GLCM = \begin{bmatrix} P(1,1) & \dots & P(1, N_g) \\ \vdots & \ddots & \vdots \\ P(N_g, 1) & \dots & P(N_g, N_g) \end{bmatrix}$	
$P_x(i) = \sum_{j=1}^{N_g} P(i, j) \quad , \quad P_y(i) = \sum_{i=1}^{N_g} P(i, j)$	
$P_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P_{i+j=k}(i, j) \quad , \quad k = 2, 3, \dots, 2N_g$	
$P_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P_{ i-j =k}(i, j) \quad , \quad k = 0, 1, \dots, N_g - 1$	
$\mu_x = \sum_i \sum_j i \cdot P(i, j) \quad , \quad \mu_y = \sum_i \sum_j j \cdot P(i, j)$	
$\sigma_x = \sum_i \sum_j (i - \mu_x)^2 \cdot P(i, j) \quad , \quad \sigma_y = \sum_i \sum_j (j - \mu_y)^2 \cdot P(i, j)$	
f1 Otokorelasyon (Soh ve Tsatsoulis, 1999)	$\sum_i \sum_j (i \cdot j) \cdot P(i, j)$
f2 Kontrast (Haralick vd., 1973; Soh ve Tsatsoulis, 1999)	$\sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{\substack{i=1 \\ i-j =n}}^{N_g} \sum_{j=1}^{N_g} P(i, j) \right\}$
f3 Korelasyon (MATLAB R2015a)	$\frac{\sum_i \sum_j (i - \mu_x) \cdot (j - \mu_y) \cdot P(i, j)}{\sigma_x \cdot \sigma_y}$
f4 Korelasyon (Haralick vd., 1973; Soh ve Tsatsoulis, 1999)	$\frac{\sum_i \sum_j (i \cdot j) \cdot P(i, j) - \mu_x \cdot \mu_y}{\sigma_x \cdot \sigma_y}$

f5	Asimetri (Soh ve Tsatsoulis, 1999) (Cluster Prominence)	$\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{i + j - \mu_x - \mu_y\}^4 \cdot P(i, j)$
f6	Çarpıklık (Soh ve Tsatsoulis, 1999) (Cluster Shade)	$\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{i + j - \mu_x - \mu_y\}^3 \cdot P(i, j)$
f7	Benzersizlik (Soh ve Tsatsoulis, 1999)	$\sum_i \sum_j i - j \cdot P(i, j)$
f8	Enerji (Haralick vd., 1973 ; Soh ve Tsatsoulis, 1999)	$\sum_i \sum_j \{P(i, j)\}^2$
f9	Entropi (Soh ve Tsatsoulis, 1999)	$- \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P(i, j) \cdot \log\{P(i, j)\}$
f10	Homojenlik (MATLAB R2015a)	$\sum_i \sum_j \frac{1}{1 + i - j } \cdot P(i, j)$
f11	Homojenlik (Soh ve Tsatsoulis, 1999)	$\sum_i \sum_j \frac{1}{1 + (i - j)^2} \cdot P(i, j)$
f12	Maksimum Olasılık (Soh ve Tsatsoulis, 1999)	$\max_{i, j} P(i, j)$
f13	Varyans (Haralick vd., 1973)	$\sum_i \sum_j (i - \mu)^2 \cdot P(i, j)$
f14	Toplam Ortalaması (Haralick vd., 1973)	$\sum_{i=2}^{2N_g} i \cdot P_{x+y}(i)$
f15	Toplam Varyansı (Haralick vd., 1973)	$\sum_{i=2}^{2N_g} (i - f_{14})^2 \cdot P_{x+y}(i)$

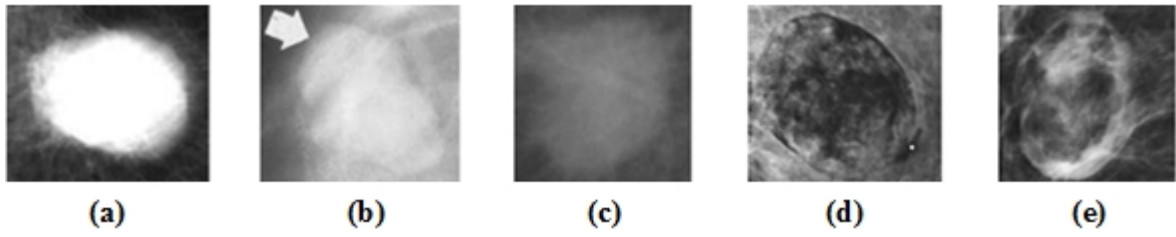
f16	Toplam Entropisi (Haralick vd., 1973)	$- \sum_{i=2}^{2N_g} P_{x+y}(i) \cdot \log\{P_{x-y}(i)\}$
f17	Fark Varyansı (Haralick vd., 1973)	variance of P_{x-y}
f18	Fark Entropisi (Haralick vd., 1973)	$- \sum_{i=2}^{2N_g} P_{x-y}(i) \cdot \log\{P_{x-y}(i)\}$
f19	Korelasyon Bilgi Ölçeği 1 (Haralick vd., 1973) (Information Measure of Correlation 1)	$\frac{HXY - HXY1}{\max\{HX; HY\}}$ $HXY = - \sum_i \sum_j P(i, j) \cdot \log(P(i, j))$ $HXY1 = - \sum_i \sum_j P(i, j) \cdot \log(p_x(i) \cdot p_y(i))$ <p>HX ve HY p_x ve p_y'nin entropi degerleridir.</p>
f20	Korelasyon Bilgi Ölçeği 2 (Haralick vd., 1973) (Information Measure of Correlation 2)	$(1 - \exp[-2 \cdot (HXY2 - HXY)])^{1/2}$ $HXY2 = - \sum_i \sum_j P_x(i) \cdot P_y(i) \cdot \log\{P_x(i) \cdot P_y(i)\}$
f21	Inverse Difference Normalized (Clausi, 2002)	Normalize Edilmiş Homojenlik (MATLAB R2015a)
f22	Inverse Difference Moment Normalized (Clausi, 2002)	Normalize Edilmiş Homojenlik (Soh ve Tsatsoulis, 1999)

6.5. Mamografik Öznitelikler

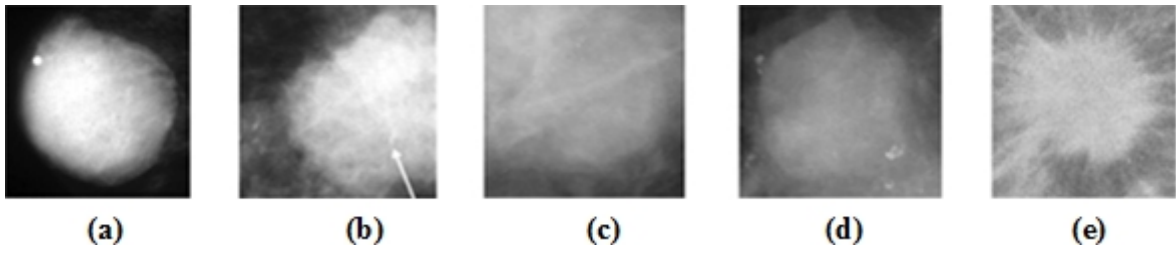
Radyoloji uzmanları mamografi imgelerinde görülen anomalileri, anomalinin yoğunluğunu, şeklini, kontur yapısını, boyutunu, anomali varlığına eşlik eden bulguları ve eski filmlere göre memede görülen değişiklikleri göz önünde bulundurarak incelemekte ve teşhis sürecini planlamaktadırlar.

Uzmanlara göre, anomali ne kadar çok radyolüsen ise o kadar iyi huylu, ne kadar radyopak ise o kadar kötü huylu olarak düşünülmektedir. Ancak, meme doku tipinin yoğun olmasından kaynaklı radyopak yapı gösteren mamografi imgelerinde, anomali tipini sadece yoğunluk bilgisine göre değerlendirmek yanıltıcı olabilmektedir. Bu nedenle, diğer mamografik bulguların da göz önünde bulundurulması gerekmektedir.

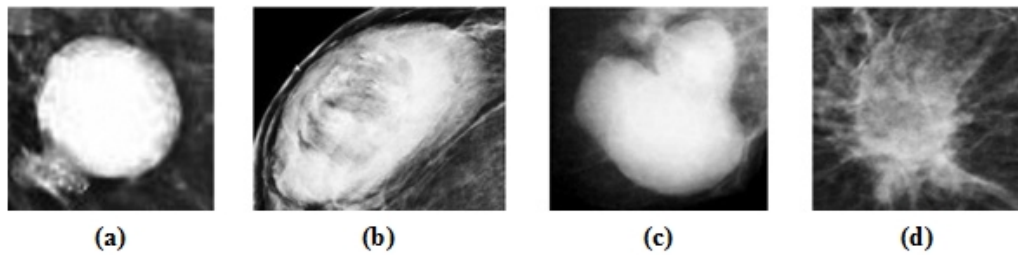
Anomalinin şekli düzensizleştikçe, boyutu arttıkça ya da konturları tam sınırlanmıyorsa ise kötü huylu olma olasılığı artmaktadır. Ancak ne kadar şekli düzensiz ya da konturları belirsiz olursa olsun, anomali yağ içeriyorsa, iyi huylu olarak düşünülmektedir. Çeşitli yoğunluk, kontur ve şekil yapısına sahip ROI bölgeleri sırasıyla Şekil 6.5, Şekil 6.6 ve Şekil 6.7’de örneklendirilmiştir.



Şekil 6. 5. Çeşitli yoğunluklara sahip ROI bölgeleri (Chokri ve Farida’dan, 2017):
(a) yüksek (b) ortalama (c) düşük (d) yağlı (e) karma

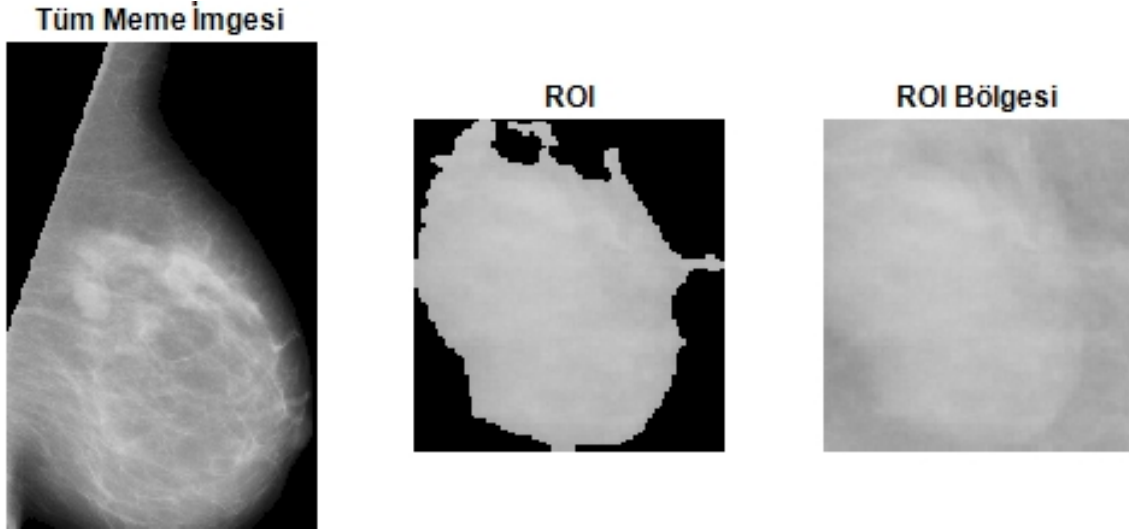


Şekil 6. 6. Çeşitli kontur yapısına sahip ROI bölgeleri (Chokri ve Farida’dan, 2017):
(a) sınırlanmış (b) mikrolobüle (c) gizli (d) belirsiz (e) spiküle



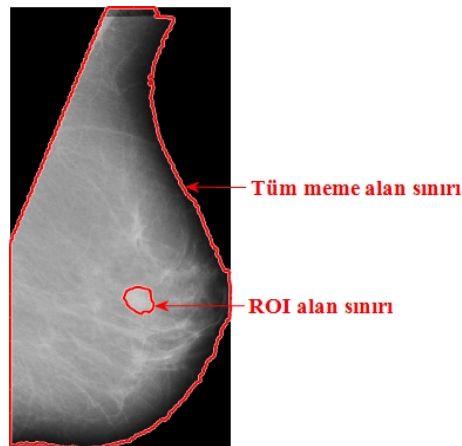
Şekil 6. 7. Çeşitli şekil yapısına sahip ROI bölgeleri (Chokri ve Farida, 2017):
(a) yuvarlak (b) oval (c) lobüle (d) düzensiz

Bu bilgiler doğrultusunda, tez çalışması kapsamında mamografi imgelerinde boyut, yoğunluk, şekil ve kontur tanımlaması yapabilecek öznitelikler çıkartılmıştır. Öznitelik çıkarımında tüm meme imgeleri, ROI imgeleri ve ROI bölgeleri kullanılmıştır. Bu imgeler Şekil 6.8’de örneklendirilmiştir.



Şekil 6. 8. Mamografik öznitelik çıkarımında kullanılan imge örnekleri

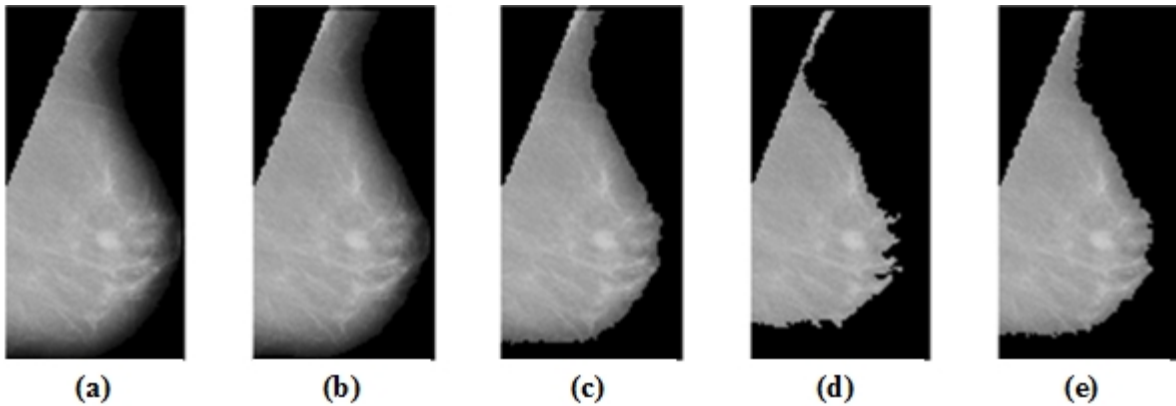
Tespit edilen ROI’lerin boyutlarının tanımlanması için öncelikle alanları, sınırları içerisinde kalan toplam piksel sayısı olarak, hesaplanmıştır. ROI boyutunun tüm meme boyutundan bağımsız olarak değerlendirilmesi için ROI alanı tüm meme alanına oranlanmıştır. Tüm meme ve ROI alan sınırları örnek bir mamografi imgesi üzerinde Şekil 6.9’da gösterilmiştir.



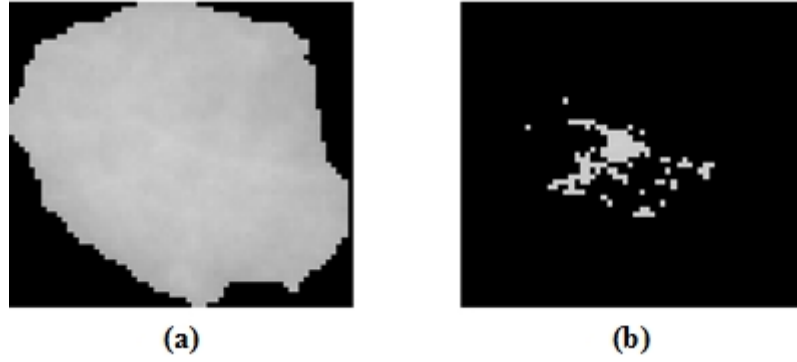
Şekil 6. 9. Örnek bir mamografi imgesi üzerinde tüm meme ve ROI alan sınırları

ROI yoğunluk tanımlaması için meme yoğunluğu, ROI yağlılık derecesi ve ROI yoğunluk derecesi olmak üzere üç farklı öznitelik çıkartılmıştır. Yoğunluk özneliği, tüm meme imgesinde belli bir eşik değerinden daha yüksek yeğlilik seviyesine sahip piksel sayısının meme parenkimi alanına oranlanması ile hesaplanmıştır. Eşik değeri seçilmesi için öncelikle Otsu eşikleme (Otsu, 1979) yöntemi kullanılarak, mamografi imgesini, arka plan görüntüsü, meme parenkimi, pektoral kas tespiti aşamasında kesilememiş kas bölgesinden geriye kalan alan ve ROI bölgesi olarak ayıracak üç eşik değeri belirlenmiştir. Yapılan incelemeler sonucunda, mamografi imgeleri ikinci ve üçüncü eşik değerlerinin ortalama değeri ile eşiklendiğinde meme parenkiminin parlak piksellerinin en uygun şekilde bölümlendiği saptanmış ve bu değer yoğunluk özneliği çıkartımında kullanılacak eşik değeri olarak kabul edilmiştir. Bahsedilen dört eşik değeri ile bölütleme sonuçları Şekil 6.10'da örnek bir mamografi imgesinde gösterilmiştir.

İçerisinde yağ bulunduran ROI'lerin iyi huylu olma olasılığının arttığı ön bilgisine dayanarak, yoğunluk tanımlaması için ROI yağlılık derecesi hesaplanmıştır. ROI yağlılık derecesi, ROI imgesinde belli bir eşik değerinden daha koyu olan piksel sayısını ROI alanına oranlayarak hesaplanmıştır. Çeşitli denemeler sonucunda, en uygun eşik değeri 200 olarak kabul edilmiştir. Yağlılık derecesi özneliği hesabı için gerçekleştirilen bölütleme işlemi Şekil 6.11'da örnek bir mamografi imgesinde gösterilmiştir.



Şekil 6. 10. Yoğunluk özneliği hesabı için elde edilen bölütleme sonuçları: (a) orijinal imge (b) Birinci eşik değeri ile bölütleme sonucu (c) İkinci eşik değeri ile bölütleme sonucu (d) Üçüncü eşik değeri ile bölütleme sonucu (e) İkinci ve üçüncü eşik değerlerinin ortalaması ile bölütleme sonucu



Şekil 6. 11. Yağlılık derecesi özniteliği hesabı için gerçekleştirilen bölütleme işlemi: (a) orijinal imge (b) bölütleme sonucu

Mamografi imgelerinde görülen anomalilerin çevresine göre daha parlak olmasının kötü huylu olma olasılığını artırdığı ön bilgisi doğrultusunda, ROI'lerin yoğunluk derecesi hesaplanmıştır. ROI yoğunluk derecesi ROI bölgesindeki en yüksek yeğinlik seviyesinin en düşük yeğinlik seviyesine oranlanması ile hesaplanmıştır.

ROI kontur tanımlaması, konturun meme parenkimindeki belirginliği ve sınırlılığının hesaplanması ile gerçekleştirilmiştir. ROI'nin belirginliği Şekil 6.6 (d)'de görüldüğü gibi, ROI yoğunluğu ile meme parenkimi yoğunluğu arasındaki benzerlik ile ilişkilidir. Bu nedenle, belirginlik özniteliği, ROI yeğinlik ortalamasının tüm meme imgesinin yeğinlik ortalamasına oranlanması ile hesaplanmıştır.

ROI sınırlılığı ise anomali konturlarının takip edilebilir (Bkz. Şekil 6.6 (a)) ya da (Bkz. Şekil 6.6 (e)) takip edilemez olduğunun bir ölçüsüdür. Bu ölçü için kompaktlık özniteliği kullanılmıştır (Rangayyan vd., 1997). Bir imgenin kompaktlığı Eşitlik 6.8 ile hesaplanmaktadır. Eşitlikte C , a , p ve r terimleri sırasıyla imgenin kompaktlığını, alanını, çevresini ve yarıçapını ifade etmektedir. ROI konturu tamamen takip edilebilir olduğunda kompaktlık sifira eşit olurken, takip edilebilir özelliği azaldıkça kompaktlık değeri artmaktadır.

$$C = 1 - \frac{4\pi a}{p^2}, \quad a = \pi r^2 \quad (6.8)$$

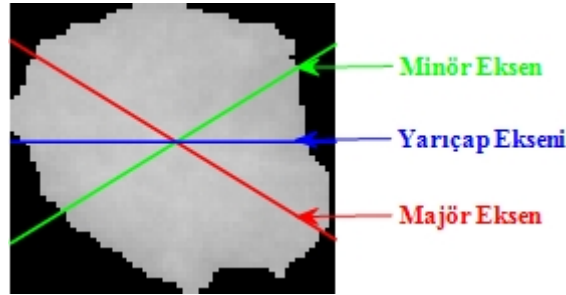
ROI şekil tanımlaması için ise yuvarlaklık, ovallik, eksantriklik, katılık (solidity) ve düzensizlik öznitelikleri kullanılmıştır. Yuvarlaklık özniteliği Eşitlik 6.9'da gösterildiği

şekilde hesaplanmaktadır. Eşitlikte A terimi ROI'nin toplam piksel sayısı cinsinden alanını, r ve R terimleri ise sırasıyla ROI yarıçapını ve yuvarlaklığını temsil etmektedir. ROI imgesi tam bir daire olduğunda yuvarlaklık özneliği bire eşit olmakta, daire şeklinden uzaklaştıkça da bir'den sapmaktadır.

$$R = \frac{A}{\pi r^2} \quad , \quad r = \frac{\text{Major Eksen Uzunugu}}{\text{Minor Eksen Uzunugu}} \quad (6.9)$$

Benzer şekilde, ovallik özneliği (O) ROI bölgesinin oval şekline ne kadar yakın olduğunu göstermekte ve Eşitlik 6.10 ile hesaplanmaktadır. Eksantriklik ise bir şeklin ne kadar elipse benzediğini gösterdiğinden, şekil özneliği olarak kullanılmıştır. Şekil 6.12, örnek bir mamografi imgesinin majör, minör ve yarıçap eksenleri göstermektedir.

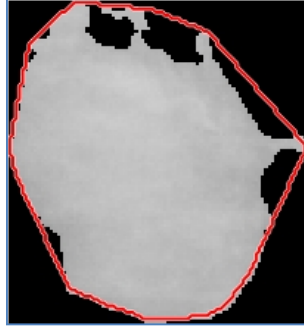
$$O = 2 \cdot \frac{\text{Major Eksen Uzunugu} - \text{Minor Eksen Uzunugu}}{\text{Major Eksen Uzunugu} + \text{Minor Eksen Uzunugu}} \quad (6.10)$$



Şekil 6. 12. Örnek bir mamografi imgesinin majör, minör ve yarıçap eksenleri

Katılık özneliği, ROI şeklinin düzenliliğini tanımlamaktadır. Bu özneliğin hesaplanması için ROI imgesini çevreleyen en küçük konveks alan hesaplanmıştır. ROI imgesinin toplam piksel sayısı cinsinden alanının konveks alana oranı katılık derecesi olarak tanımlanmıştır. Örnek bir ROI imgesini çevreleyen konveks alanın çizilmesi Şekil 6.13'te gösterilmiştir. Şekilde, kırmızı çizgi konveks alanın sınırlarını göstermektedir.

Son olarak, ROI şekil düzensizliği, ROI konturlarındaki yön değişimleri esas alınarak hesaplanmıştır. Yön değişim sayısının ROI toplam sınır piksel sayısına oranı olarak hesaplanan düzensizlik özneliği, Şekil 6.7 (d)'de gösterildiği gibi düzensiz şekle sahip ROI'lerde yüksek, daha düzenli şekillerde ise düşük çıkmaktadır.



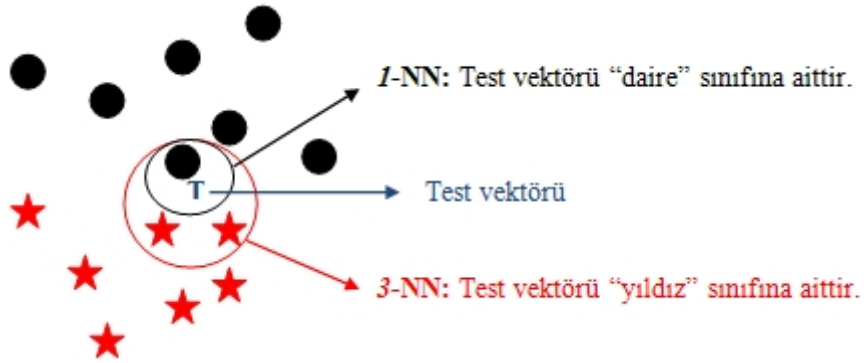
Şekil 6. 13. Örnek bir ROI imgesini çevreleyen konveks alan

7. MAMOGRAFİ İMGELERİNİN SINIFLANDIRILMASI

Sınıflandırma problemlerinin son adımı olan sınıflandırma aşamasında seçilen sınıflandırıcılar da sınıflandırma hassasiyetini oldukça etkilemektedir. Tez çalışması kapsamında, sınıflandırma aşamalarında, k -NN, karar ağacı, rastgele orman, Naïve Bayes, LLC, LDC, FLDA ve SVM sınıflandırıcıları kullanılmıştır.

7.1. k -En Yakın Komşu Sınıflandırıcısı

k -en yakın komşu sınıflandırıcısı, öznitelik vektör uzayındaki her bir öznitelik vektörünün birbirlerine olan uzaklıklarına göre sınıflandırma işlemi yapmaktadır (Fix ve Hodges, 1952). k -NN sınıflandırıcısı ile sınıflandırılacak olan test öznitelik vektörünün daha önceden sınıflandırılmış olan öznitelik vektörlerine olan Öklid uzaklığı hesaplanmakta ve minimum uzaklığı veren k adet öznitelik seçilmektedir. Seçilen k adet öznitelik çoğunlukla hangi sınıfa ait ise test öznitelik vektörü o sınıfa atanmaktadır. k -NN sınıflandırıcısının çalışma prensibi şematiği Şekil 7.1’de verilmiştir.

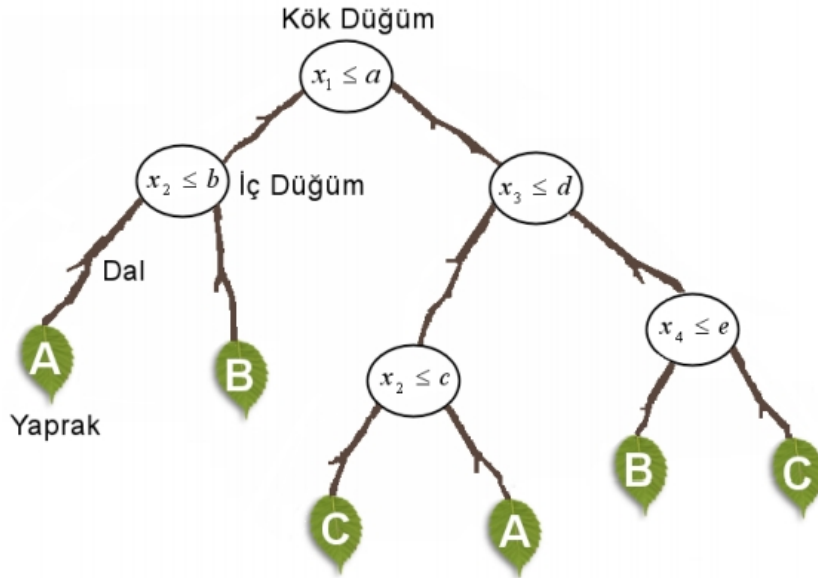


Şekil 7. 1. k -NN sınıflandırıcısı çalışma prensibi

7.2. Karar Ağacı Sınıflandırıcısı

Karar ağacı sınıflandırıcılarının amacı, öznitelik vektörlerinin bir kümeleme algoritması ile grubun tüm elemanları aynı sınıf etiketine sahip olana kadar alt gruplara bölünmesidir (Safavian ve Landgrebe, 1991). Karar ağaçları Şekil 7.2’de gösterildiği gibi tüm veri kümesinin oluşturduğu kök düğüm ile başlamaktadır. Ağacın dalları sınıflandırma

kurallarını, yaprakları ise sınıf etiketlerini ifade etmektedir. Karar ağaçları sınıflandırıcısının eğitim aşamasında, örneklerin hepsi aynı sınıfa ait ise düğüm yaprak olarak sonlanmakta ve sınıf etiketini almaktadır. Örneklerin hepsi aynı sınıfta değil ise örnekleri sınıflara ayıracak en iyi öznelik seçilmekte ve bu işleme örneklerin hepsi aynı sınıfa ait olana kadar devam edilmektedir. En iyi öznelik seçimi için bilgi kazancı (information gain), gini indeksi (gini index) ve entropi gibi kavramlar kullanılmaktadır. Eğitim aşamasında oluşan karar kuralları, test aşamasında sınıflandırılacak örneğe uygulanmaktadır. Sonuç olarak, örneğin ulaştığı yaprağın etiketi o örneğin sınıf bilgisini vermektedir.



Şekil 7. 2. Karar ağacı sınıflandırıcısı çalışma prensibi (Kavzoğlu ve Çölkesen'den, 2010)

7.3. Rastgele Orman Sınıflandırıcısı

Karar ağacı sınıflandırıcısının eğitim aşamasında karar ağacının öğrenmesi, bir başka deyimle karar kurallarının oluşturulması gerçekleştirilmektedir. Karar ağacının öğrenmesi aşamasında rastgele orman (random forest) algoritması, hızlandırılmış ağaçlar (boosted trees), döndürme ağacı (rotation forest), ID3 ve C4.5 algoritmaları kullanılmaktadır (Breiman, 2001). Bu çalışmada kullanılan rastgele orman algoritması,

sınıflandırma işlemi sırasında birden fazla karar ağacı kullanarak sınıflandırma hassasiyetini yükseltmeyi amaçlamaktadır.

7.4. Naïve Bayes Sınıflandırıcısı

Bayes sınıflandırıcılar, test vektörlerinin hangi sınıfa hangi olasılıkla ait olduklarını hesaplamaktadırlar. Bayes karar kuramına göre test örneği, ait olma olasılığı en yüksek olan sınıfa atanmaktadır (Chen vd., 2009). 2-sınıflı bir sınıflandırma probleminde herhangi bir \bar{x} örneği için Bayes karar kuralı Eşitlik 7.1'de verilmiştir.

$$P(w_i | \bar{x}) > P(w_j | \bar{x}), \forall j \neq i \Leftrightarrow \bar{x} \in w_i \quad (7.1)$$

Eşitlikteki $P(w_i | \bar{x})$ ve $P(w_j | \bar{x})$ terimleri \bar{x} örneğinin sırasıyla w_i ve w_j sınıflarına ait olma olasılıklarını göstermektedir ve $P(w_i | \bar{x})$ Eşitlik 7.2'de gösterildiği şekilde hesaplanmaktadır.

$$P(w_i | \bar{x}) = \frac{p(\bar{x} | w_i) \cdot P(w_i)}{p(\bar{x})} \quad (7.2)$$

Eşitlikteki $P(w_i)$, $p(\bar{x} | w_i)$ ve $p(\bar{x})$ terimleri sırasıyla w_i sınıfının var olma olasılığını, w_i sınıfının \bar{x} örneğini içerme olasılığını ve \bar{x} örneğinin olasılık yoğunluk fonksiyonunu göstermektedir. \bar{x} örneğinin 1-boyutlu ve l -boyutlu olduğu durumlarda $p(\bar{x} | w_i)$ ve $p(\bar{x})$ ifadeleri sırasıyla Eşitlik 7.3, 7.4 ve 7.5'te gösterildiği şekilde hesaplanmaktadır. Eşitliklerdeki μ , σ ve Σ ifadeleri sırasıyla özniteliklerin ortalaması, 1-boyutlu özniteliklerin varyansı ve l -boyutlu özniteliklerin kovaryans matrisini göstermektedir.

$$p(\bar{x} | w_i) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(\bar{x} - \mu)^2}{2\sigma^2}} \quad (7.3)$$

$$p(\bar{x} | w_i) = \frac{1}{(2\pi)^{l/2} \cdot |\Sigma|^{l/2}} \cdot e^{-\frac{1}{2}(\bar{x} - \mu)^T \cdot \Sigma^{-1} \cdot (\bar{x} - \mu)} \quad (7.4)$$

$$p(\vec{x}) = \sum_{i=1}^2 p(\vec{x} | w_i) \cdot P(w_i) \quad (7.5)$$

Naïve Bayes sınıflandırıcılar, 1 veya l -boyutlu öznitelikleri oluşturan tüm elemanların birbirlerinden istatistiksel olarak bağımsız yani korelasyonsuz olduğunu kabul ederek Eşitlik 7.1’de verilen karar kriteri ile sınıflandırma yapmaktadır (Chen vd., 2009).

7.5. Lojistik Doğrusal Sınıflandırıcı

Lojistik doğrusal sınıflandırıcısı, veri kümesindeki en az değişkeni kullanarak en iyi uyuma sahip olacak şekilde bağımlı ve bağımsız değişkenler arasındaki ilişkiyi tanımlayarak bir ayırmsama modeli oluşturmaktadır (Webb, 2002 b). Yöntemde bağımlı ve bağımsız değişkenler arasındaki ilişki lojistik regresyon analizi ile belirlenmekte ve ayırmsama modeli doğrusal sınıflandırıcı ile tayin edilmektedir. Lojistik regresyon analizinde bağımlı değişkenin alacağı değerlerden birinin gerçekleşme olasılığı kestirilmektedir. 2-sınıflı bir sınıflandırma probleminde herhangi bir \vec{x} örneği için oluşturulan LLC modeli Eşitlik 7.6’da verilmiştir.

$$\log\left(\frac{p(\vec{x} | w_1)}{p(\vec{x} | w_2)}\right) = \vec{\beta}^T \cdot \vec{x} + \beta_0 \quad (7.6)$$

Yöntemde sınıf olasılıkları ve kovaryans matrislerinin eşit olduğu, sınıflar arasında loglinear model oluşturulabileceği varsayımları yapılmaktadır. Bu varsayımlar Eşitlik 7.7 ile gösterilmektedir.

$$p(w_1 | \vec{x}) = \frac{\exp(\vec{\beta}^T \cdot \vec{x} + \beta_0)}{1 + \exp(\vec{\beta}^T \cdot \vec{x} + \beta_0)} \quad (7.7)$$

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

Eşitlikte $\beta'_0 = \beta_0 + \log\left(\frac{p(w_1)}{p(w_2)}\right)$ olarak tanımlanmaktadır. LLC karar kriteri Eşitlik

7.8 ve 7.9’da verilmiştir.

$$\vec{x} \in \begin{cases} w_1 & , \quad \frac{p(w_1 | \vec{x})}{p(w_2 | \vec{x})} > 1 \\ w_2 & , \quad \frac{p(w_1 | \vec{x})}{p(w_2 | \vec{x})} < 1 \end{cases} \quad (7.8)$$

$$\vec{x} \in \begin{cases} w_1 & , \quad \vec{\beta}^T \cdot \vec{x} + \beta_0 > 0 \\ w_2 & , \quad \vec{\beta}^T \cdot \vec{x} + \beta_0 < 0 \end{cases} \quad (7.9)$$

7.6. Doğrusal Ayırtaç Sınıflandırıcısı

Doğrusal ayırtaç sınıflandırıcısı ile eğitim öznitelik vektörlerinin doğrusal birleşimlerini bularak birbirinden ayıran doğrusal bir hiperdüzlem, $g(\vec{x})$, oluşturup sınıflandırma yapmayı amaçlanmaktadır (Webb, 2002 a). Bu hiperdüzlem doğrusal ayırtaç fonksiyonu olarak tanımlanıp Eşitlik 7.10 ile ifade edilmektedir.

$$g(\vec{x}) = \vec{w}^T \cdot \vec{x} + w_0 \quad (7.10)$$

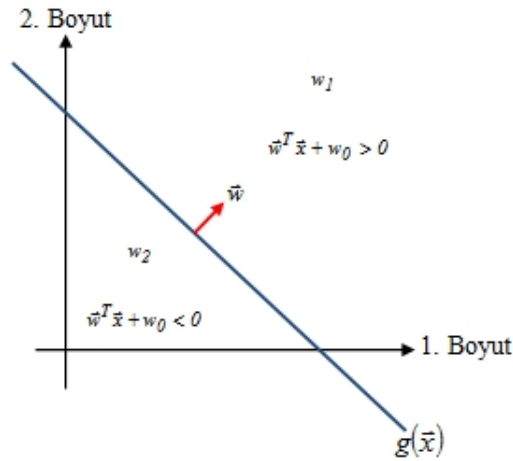
2-boyutlu bir sınıflandırma problemi düşünüldüğünde, LDC test aşamasında, test vektörü \vec{x} doğrusal ayırtaç fonksiyonuna ($g(\vec{x})$) iz düşürülerek Eşitlik 7.11'de verilen karar kriterine göre sınıflandırılmaktadır.

$$\vec{x} \in \begin{cases} w_1 & , \quad \vec{w}^T \cdot \vec{x} + w_0 > 0 \\ w_2 & , \quad \vec{w}^T \cdot \vec{x} + w_0 < 0 \end{cases} \quad (7.11)$$

LDC sınıflandırıcısının çalışma prensibi Şekil 7.3'te gösterilmektedir.

7.7. Fisher Doğrusal Ayırtaç Analizi Sınıflandırıcısı

Fisher doğrusal ayırtaç sınıflandırıcısı, eğitim verisini sınıf-içi varyansı minimum, sınıflar-arası varyansı maksimum yapacak düşük-boyutlu alt uzaya indirgeyerek sınıfları ayıran bir hiperdüzlem oluşturmaktadır (Fisher, 1936). Bu sebeple, FLDA, Eşitlik 7.12'de verilen fonksiyonu en büyükmeyi amaçlamaktadır.



Şekil 7.3. LDC sınıflandırıcısı çalışma prensibi

$$J(\bar{w}) = \frac{\bar{w}^T \cdot S_B \cdot \bar{w}}{\bar{w}^T \cdot S_W \cdot \bar{w}} \quad (7.12)$$

Eşitlikte \bar{w} eğitim kümesindeki N -boyutlu doğrusal bağımsız öznitelikleri, S_B ve S_W ise sırasıyla sınıflar-arası ve sınıf-içi saçılım matrislerini göstermekte ve $\bar{\mu}_i$ ($i = 1, 2$) ortalamalı 2-sınıflı sınıflandırma problemi için Eşitlik 7.13 ve 7.14'te gösterilmektedir.

$$S_i = \sum_{j=1}^N (\bar{x}_j - \mu_i) \cdot (\bar{x}_j - \mu_i)^T \quad (7.13)$$

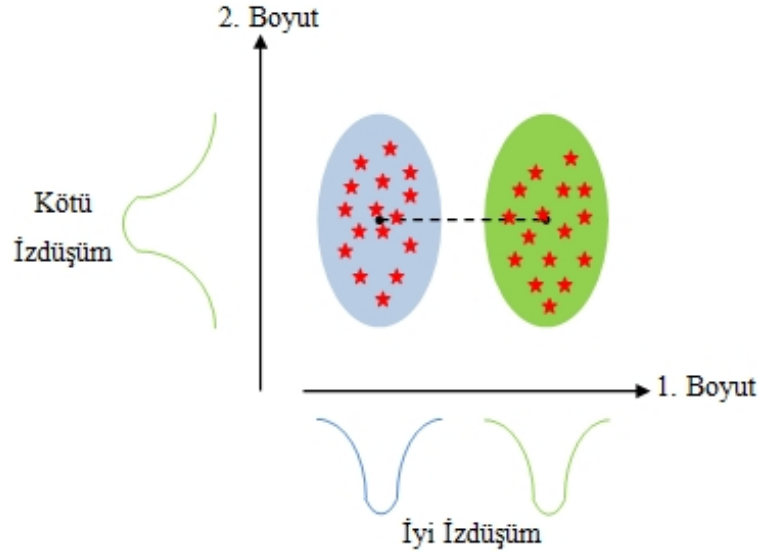
$$S_W = S_1 + S_2$$

$$S_B = (\bar{\mu}_1 - \bar{\mu}_2) \cdot (\bar{\mu}_1 - \bar{\mu}_2)^T \quad (7.14)$$

Eğitim aşamasında FLDA sınıflandırıcısı, $J(\bar{w})$ fonksiyonunu en büyükmek amacı ile $(S_W^{-1} \cdot S_B)$ 'nin özdeğer/özvektör ayrışımı gerçekleştirilmektedir. Eğitim kümesindeki öznitelikler, sıfır olmayan ya da en büyük m özvektöre karşılık gelen özdeğerler ile oluşturulan izdüşüm matrisine izdüşürülerek sınıfları birbirinden ayıran hiperdüzlem elde edilmektedir (Özkan vd., 2015).

Test aşamasında, \bar{x}_{test} vektörü de izdüşüm matrisine izdüşürülmekte ve izdüşürülmüş eğitim vektörleri ile arasındaki uzaklık hesaplanmaktadır. Sonuç olarak,

\bar{x}_{test} vektörü arasındaki uzaklığın en az çıktığı vektörün ait olduğu sınıfa atanmaktadır. FLDA sınıflandırıcısının çalışma prensibi Şekil 7.4'te gösterilmektedir.



Şekil 7. 4. FLDA sınıflandırıcısı çalışma prensibi

7.8. Destek Vektör Makineleri Sınıflandırıcısı

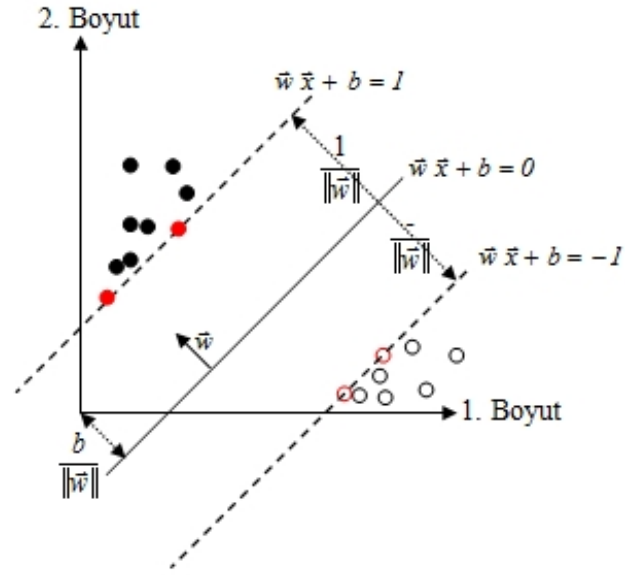
Destek vektör makineleri sınıflandırıcısı, eğitim kümesindeki öznitelik vektörlerini ayıran en uygun hiperdüzlemi bulmayı amaçlamaktadır (Chapelle vd., 1999). SVM sınıflandırıcısı en uygun hiperdüzlem olarak o hiperdüzleme en yakın öznitelik vektörlerinin arasındaki uzaklığı maksimum yapan hiperdüzlemi seçmektedir (Özkan vd., 2015). Bu vektörler, destek vektörler olarak bilinmektedir.

SVM, doğrusal olarak ayrılabilen verileri sınıflandırabildiği gibi doğrusal olmayan verileri de uygun bir çekirdek fonksiyonu kullanarak daha yüksek boyutlarda doğrusal olarak ayrılabilir hale getirip sınıflandırabilmektedir.

Eğitim vektörleri $TS = \{(\bar{x}_1, L_1), (\bar{x}_2, L_2), \dots, (\bar{x}_M, L_M)\}$ olan 2-sınıflı bir sınıflandırma probleminde \bar{x}_i vektörleri L_i sınıf etiketlerini temsil etmektedir. Bu durumda \bar{x}_{test} vektörü için karar fonksiyonu Eşitlik 7.15'te verilmektedir.

$$f(\bar{x}_{test}) = \sum_{i=1}^M \left\{ \alpha_i \cdot L_i \cdot (x_i^T \cdot x_{test}) + b \right\} \quad (7.15)$$

Eşitlikteki α_i kuadratik denklemin sıfırdan farklı katsayılarını, $\left(\frac{|b|}{\|\vec{w}\|}\right)$ hiperdüzlem ile orijin arasındaki en büyülenecek dik uzaklık ve \vec{w} hiperdüzleme normal vektörü göstermektedir. \vec{x}_{test} vektörünün sınıfı karar fonksiyonun işaretine göre belirlenmektedir. SVM sınıflandırıcısının çalışma prensibi Şekil 7.5'te gösterilmektedir.



Şekil 7. 5. SVM sınıflandırıcısı çalışma prensibi

8. MATERYAL VE YÖNTEM

Tez çalışması kapsamında gerçekleştirilen çalışmalarda, IRMA ve MIAS veritabanları kullanılmıştır. Gerçekleştirilen çalışmaların performans değerlendirmesi kullanılan metrikler ve matematiksel ifadeleri Çizelge 8.1’de verilmiştir.

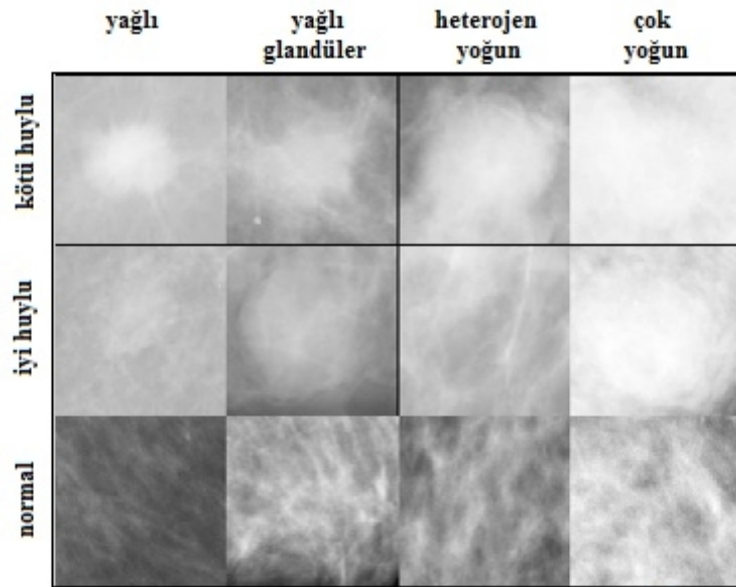
Çizelge 8. 1. Değerlendirme kriterleri ve matematiksel ifadeleri

Değerlendirme Kriterleri	Matematiksel İfadeleri
TP: Doğru Pozitif (True Positive)TN: Doğru Negatif (True Negative) FP: Yanlış Pozitif (False Positive)N: Yanlış Negatif (False Negative)	
C : Sınıf Sayısı N_i : i .sınıfa ait veri sayısı	
Hassasiyet(Sensitivity – SNS)	$\% SNS = \frac{TP}{TP + FN} \cdot 100$
Kesinlik(Specificity – SPC)	$\% SPC = \frac{TN}{TN + FP} \cdot 100$
Pozitif Tahmin Değeri(Positive Predictive Value – PPV)	$\% PPV = \frac{TP}{TP + FP} \cdot 100$
Negatif Tahmin Değeri(Negative Predictive Value – NPV)	$\% NPV = \frac{TN}{TN + FN} \cdot 100$
Yanlış Pozitif Oranı(False Positive Rate – FPR)	$\% FPR = \frac{FP}{FP + TN} \cdot 100$
Yanlış Negatif Oranı(False Negative rate – FNR)	$\% FNR = \frac{FN}{TP + FN} \cdot 100$
Yanlış Keşif Oranı(False Discovery Rate – FDR)	$\% FDR = \frac{FP}{TP + FP} \cdot 100$
Yanlış İhmal Oranı(False Omission Rate – FOR)	$\% FOR = \frac{FN}{TN + FN} \cdot 100$
Doğruluk(Accuracy – ACC)	$\% ACC = \frac{\sum_{i=1}^C SNS_i \cdot N_i}{\sum_{i=1}^C N_i}$

8.1. IRMA Veritabanı Kullanılarak Gerçekleştirilen Çalışmalar

8.1.1. IRMA veritabanı

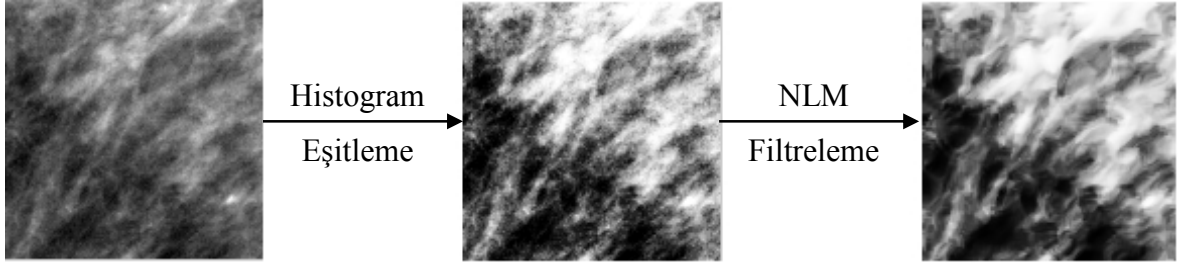
IRMA veritabanı, BI-RADS tarafından tanımlanmış tüm meme doku tiplerine (yağlı, yağlı-glandüler, heterojen-yoğun, çok yoğun) ait 233 normal (sağlıklı), 233 iyi ve 233 kötü huylu lezyon içeren mamografi imgesi olmak üzere toplamda 12 sınıf ve 2796 imge içermektedir (Deserno vd., 2011). Veritabanındaki imgeler 130×130 boyutlarında, 8 bit/piksel çözünürlüğünde, .png formatında, tüm meme bölgesinden normal ve hastalıklı bölgelerin çıkartılmış olduğu imge parçalarıdır. Veritabanında bulunan her bir sınıfa ait birer örnek imge Şekil 8.1’de gösterilmiştir.



Şekil 8. 1. IRMA veritabanına ait örnek imgeler

8.1.2. Mamografi imgelerinin ön işlenmesi

Mamografi imgelerinde en önemli özniteliklerden biri olan kontur bilgisini korurken gürültü giderme olanağı sağladığından imge iyileştirme aşamasında, histogramı eşitlenerek kontrastı artırılmış mamografi imgelerine NLM filtresi uygulanmıştır belirtilmiştir (Ergin ve Kılınc, 2014). Filtre parametreleri $R_{sim} = 5$, $R_{win} = 3$ ve $\sigma = 15$ olarak seçilmiştir. Ön işleme aşamaları örnek mamogram üzerinde Şekil 8.2’de verilmiştir.



Şekil 8. 2. IRMA veritabanına ait ön işleme aşaması adımları

8.1.3. Öznitelik vektörlerinin oluşturulması

Öznitelik çıkartımı için ön işleme tabi tutulmuş imgeler öncelikle Daubechies1 dalgacığı kullanılarak birinci seviyede 2-boyutlu ayrık dalgacık dönüşümüne tabi tutulmuştur. Dalgacık dönüşümü sonucu elde edilen her bir alt banda (Alçak/Alçak – AA, Alçak/Yüksek – AY, Yüksek/Alçak – YA, Yüksek/Yüksek – YY) $P = 8$ - komşuluk ve $R = 2$ yarıçap parametreleri ile LCP dönüşümü uygulanarak 81×1 -boyutlu vektörler elde edilmiştir. Elde edilen vektörlerin 81. boyutundaki değerler diğer boyutlardaki değerlerden çok büyük olduğu için 81. boyutlar atılmış ve her bir alt banttan elde edilen vektörler Çizelge 8.2’de verilen katsayılar ile ağırlıklandırılarak toplanmıştır. Bu süreç veritabanındaki tüm imgelere uygulanmış ve her birine ait 80-boyutlu öznitelik vektörleri elde edilmiştir belirtilmiştir (Ergin ve Kılınç, 2014).

Çizelge 8. 2. LCP dönüşümü ile öznitelik vektörleri oluşturulurken alt bantların çarpıldığı katsayılar ve onlara ait LCP kodu (Ergin ve Kılınç’tan, 2014)

LCP kodu	AA-AY-YA-YY katsayıları
LCP 1-1	1-1-1-0
LCP 1,4-1	1,4-1-1-0
LCP 1,8-1	1,8-1-1-0

Öznitelik vektörlerinin imgeleri temsil edebilirlik kapasitesini yükseltmek amacı ile, Ergin ve Kılınç’ın (2014) çalışmalarında kullandıkları katsayı kümelerinden en başarılı olarak tespit ettikleri LCP 1,4-1 kodlu 80-boyutlu öznitelik vektörlerine alt bölüm 6.2’de anlatılan istatistiksel ve alt bölüm 6.3’te anlatılan frekans-düzlemi öznitelikleri eklenerek vektör uzayı 108-boyuta genişletilmiştir. Oluşturulan 108-boyutlu öznitelik vektörü Çizelge 8.3’te özetlenmiştir.

Çizelge 8. 3. 108-boyutlu öznitelik vektörlerinin oluşturulması

LCP Öznitelik Vektörü	İstatistiksel Öznitelikler	Frekans-Düzlemi Öznitelikleri	108-Boyutlu Öznitelik Vektörü
80×1	12×1	16×1	108×1
LCP 1,4-1	LCP 1,4-1: Enerji	AAAA: Enerji	108-Boyutlu LCP-Tabanlı Öznitelik Vektörü
	LCP 1,4-1: Ortalama	AAAY: Enerji	
	LCP 1,4-1: Varyans	AAYA: Enerji	
	LCP 1,4-1: Maksimum	AAYY: Enerji	
	LCP 1,4-1: Minimum	AYAA: Enerji	
	LCP 1,4-1: Standart Sapma	AYAY: Enerji AYYA: Enerji	
	LCP 1,4-1: Çarpıklık	AYYY: Enerji	
	LCP 1,4-1: Basıklık	YAAA: Enerji	
	LCP 1,4-1: Alan- Tanımlayıcısı	YAAAY: Enerji YAYA: Enerji	
	LCP 1,4-1: Ortalama Enerji	YAYY: Enerji YYAA: Enerji	
	LCP 1,4-1: Enerji Varyansı	YYAY: Enerji YYYA: Enerji	
	LCP 1,4-1: Entropi	YYYY: Enerji	

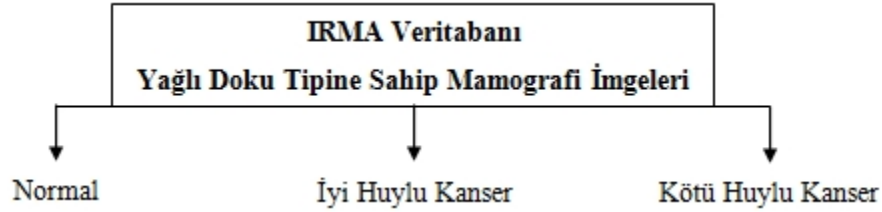
8.1.4. Meme kanseri teşhisi

IRMA veritabanı kullanılarak yapılan meme kanseri teşhisi çalışması 108-boyutlu LCP-tabanlı öznitelik vektörlerinin FLDA, LDC, LLC, SVM, k -NN ($k=5$), Naïve Bayes, karar ağacı ve rastgele orman sınıflandırıcıları kullanılarak 3- ve 12-sınıflı sınıflandırılması ile gerçekleştirilmiştir. Sınıflandırma çalışmasında çapraz-doğrulama tekniği kullanılmıştır.

8.1.4.1. 3-sınıflı meme kanseri teşhis çalışması

3-sınıflı meme kanseri teşhis çalışmasında veritabanının yalnızca yağlı doku tipine sahip mamografi imgeleri sağlık durumlarına göre *normal*, *iyi huylu kanser* ya da *kötü*

huylu kanser olmak üzere üç sınıfa ayrılmıştır. Çalışmada, her bir sınıfa ait 233 imgenin 210 tanesi (%90) olmak üzere toplamda 630 imge eğitim aşamasında ve 23 tanesi (%10) olmak üzere toplamda 69 imge test aşamasında kullanılmıştır. 1-aşamalı meme kanseri sınıflandırma sürecinin iş akış şeması Şekil 8.3'de gösterilmiştir.



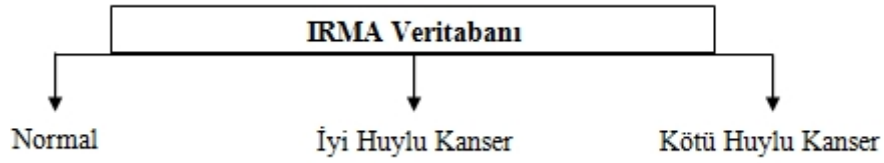
Şekil 8. 3. 3-sınıflı meme kanseri sınıflandırma sürecinin iş akış şeması

8.1.4.2. 12-sınıflı meme kanseri teşhis çalışması

12-sınıflı meme kanseri teşhis çalışmasında veritabanındaki tüm imgeler 1-, 2- ve 3-aşamalı meme kanseri sınıflandırma süreçlerine tabi tutulmuştur. Çalışmada, her bir sınıfa ait 233 imgenin 210 tanesi (%90) olmak üzere toplamda 2520 imge eğitim aşamasında ve 23 tanesi (%10) olmak üzere toplamda 276 imge test aşamasında kullanılmıştır.

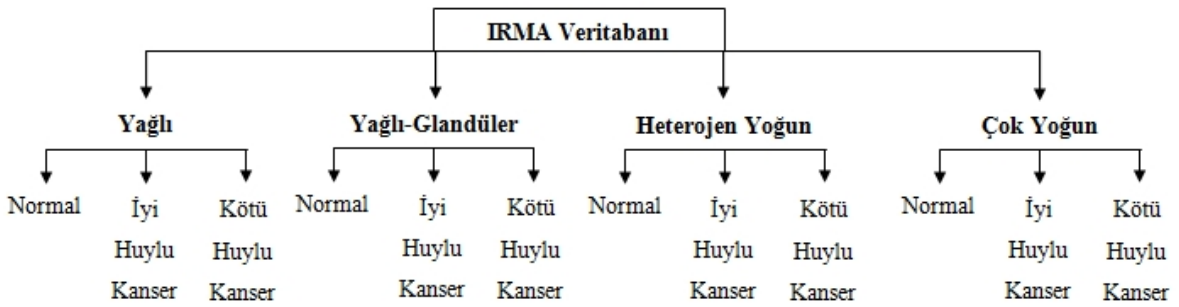
1-aşamalı meme kanseri sınıflandırma sürecinde, mamografi imgeleri meme doku tipi dikkate alınmaksızın direk olarak sağlık durumlarına göre *normal*, *iyi huylu kanser* ya da *kötü huylu kanser* olmak üzere üç sınıfa ayrılmıştır. 1-aşamalı meme kanseri sınıflandırma sürecinin iş akış şeması Şekil 8.4'te gösterilmiştir.

2-aşamalı meme kanseri sınıflandırma süreci mamografi imgelerinin meme doku tipi tespiti ve meme kanseri teşhisinden oluşmaktadır. Bu sürecinin ilk aşamasında mamografi imgeleri meme doku tiplerine göre *yağlı*, *yağlı-glandüler*, *heterojen yoğun* ve *çok yoğun* olmak üzere dört sınıfa ayrılmaktadır. Sürecin ikinci aşamasında ise her bir meme doku tipine sınıflandırılmış mamografi imgeleri sağlık durumlarına göre *normal*, *iyi huylu kanser* ya da *kötü huylu kanser* olmak üzere üç sınıfa ayrılmıştır.



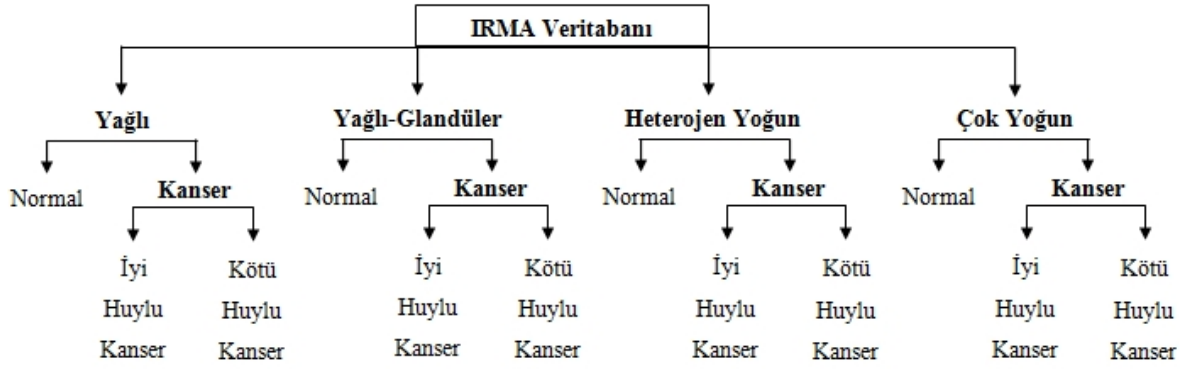
Şekil 8. 4. 1-aşamalı meme kanseri sınıflandırma sürecinin iş akış şeması

IRMA veritabanı kullanılarak gerçekleştirilen 2-aşamalı meme kanseri sınıflandırma sürecinin iş akış şeması Şekil 8.5’de gösterilmiştir.



Şekil 8. 5. IRMA veritabanı kullanılarak 2-aşamalı meme kanseri sınıflandırma sürecinin iş akış şeması

3-aşamalı meme kanseri sınıflandırma süreci mamografi imgelerinin meme doku tipi tespiti ve meme kanseri teşhisinden oluşmakla beraber bu çalışmada meme kanseri teşhisi iki aşamada gerçekleştirilmektedir. Bu sürecinin ilk aşamasında mamografi imgeleri meme doku tiplerine göre *yağlı*, *yağlı-glandüler*, *heterojen yoğun* ve *çok yoğun* olmak üzere dört sınıfa ayrılmaktadır. Sürecin ikinci aşamasında, her bir meme doku tipine sınıflandırılmış mamografi imgeleri sağlık durumlarına göre *normal* ya da *kanser* olmak üzere iki sınıfa ayrılmıştır. Son olarak, üçüncü aşamada, *kanserli* olduğu belirlenmiş mamografi imgeleri kanser tiplerine göre *iyi huylu kanser* ya da *kötü huylu kanser* olarak sınıflandırılmaktadır. IRMA veritabanı kullanılarak gerçekleştirilen 3-aşamalı meme kanseri sınıflandırma sürecinin iş akış şeması Şekil 8.6’da gösterilmiştir.



Şekil 8. 6. IRMA veritabanı kullanılarak 3-aşamalı meme kanseri sınıflandırma sürecinin iş akış şeması

8.2. MIAS Veritabanı Kullanılarak Gerçekleştirilen Çalışmalar

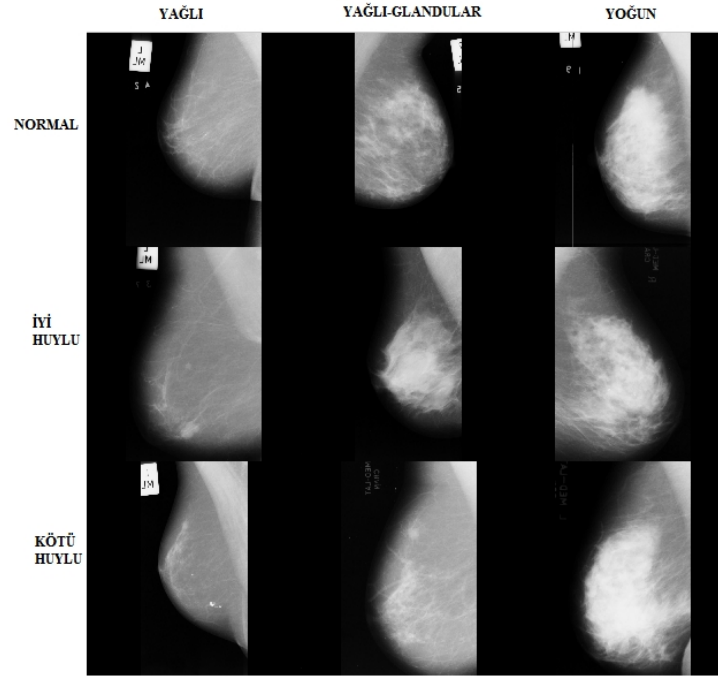
8.2.1. MIAS veritabanı

MIAS veritabanı, 106 tane yağlı, 104 tane glandüler ve 112 tane yoğun doku tiplerine sahip toplam 161 kişiye ait 322 mamografi imgesi ve bu imgelere ait 207 tane normal, 69 tane iyi huylu ve 54 tane kötü huylu kanser olmak üzere toplam 330 tane içermektedir (Suckling, 1994).

Veritabanındaki imgeler 1024×1024 boyutlarında, 8 bit/piksel çözünürlüğünde, .pgm formatında MLO çekim-tüm meme mamografi imgeleridir. Veritabanında sağ/sol göğüs, göğüs dokusu yoğunluk derecesi, anomali olup olmadığı, var olan anomalinin tipi ve yeri bilgileri verilmiştir. Veritabanında bulunan her bir sınıftan birer örnek imge Şekil 8.7’de gösterilmiştir.

8.2.2. Mamografi imgelerinin ön işlenmesi

MIAS veritabanındaki mamografi imgeleri tüm memeyi kapsadığından, kanser teşhisi için bu imgelerden ROI’lerin tespit edilmesi gerekmektedir. ROI bölgelerinin meme parenkimine göre daha parlak olduğu gerekçesi ile piksellerin yeğlilik seviyeleri farklılıkları dikkate alınarak ROI tespiti gerçekleştirilmektedir.



Şekil 8. 7. MIAS veritabanına ait örnek imgeler

Şekil 8.8’de verilen örnek mamografi imgesinde de görüldüğü gibi, MLO-çekim mamografilerde görüntülenen pektoral kasların, imgelerde görüntüleme esnasında oluşan sıklıkla oluşan benek gürültülerin (speckle noise) ve sağ/sol göğüs, CC/MLO çekim bilgilerini veren etiketleme sonucu arka planda oluşan düşük ve yüksek yeğinlikli yapay parlaklıkların varlığı yeğinlik seviyesi tabanlı ROI tespit işlemini zorlaştırmaktadır.

Mamografi imgelerinde görüntüleme esnasında oluşan gürültülerin giderimi için imgelere ayrı ayrı medyan, CWM, Frost ve bilateral filtreleri uygulanmıştır.

Gürültü giderimi aşamasından sonra imgelere morfolojik işlemler uygulanarak Şekil 8.8’de gösterilen düşük ve yüksek yeğinlikli yapaylıkların giderimi gerçekleştirilmiştir. Bu doğrultuda, mamografi imgeleri ikili seviyeye dönüştürülmüş ve meme bölgesini yapay parlaklıklardan daha büyük olacağı gerekçesi ile en büyük alana sahip bölge belirlenmiştir. Daha sonra imgeler sadece meme bölgesini içerecek şekilde kesilmiş ve arka plandan arındırılmıştır. Son olarak, arka plandan arındırılmış mamografi imgelerinin yön tayini yapılmış ve pektoral kasların imgenin sol üst köşesinde olması amacı ile tüm imgeler sola dayalı hale getirilmiştir. Yön tayini için Nagi vd. (2010) tarafından önerilen algoritma kullanılmıştır. Bu algoritma ile arka plandan arındırılmış ikili

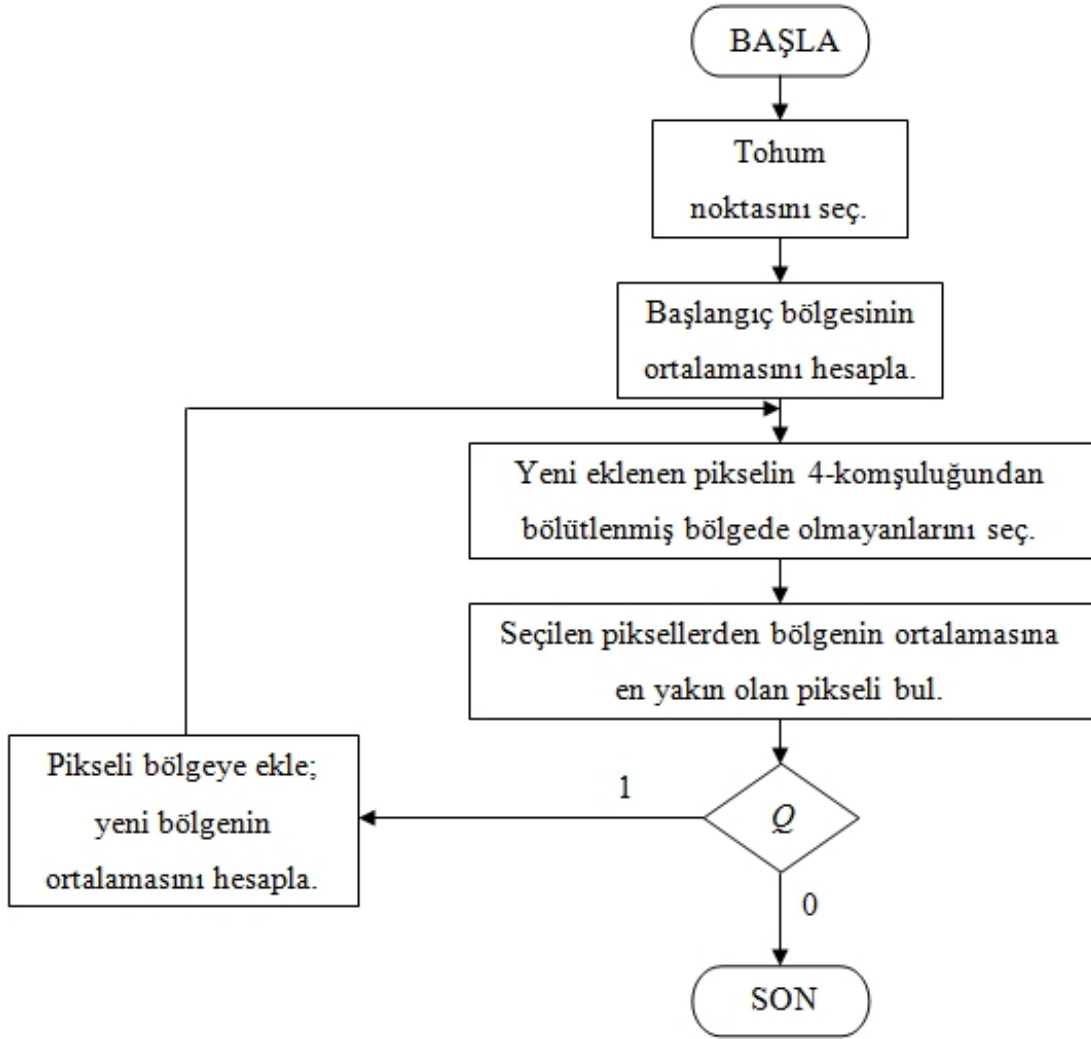
seviye mamografi imgelerinin ilk 5 ve son 5 sütunlar toplamı hesaplanmaktadır. Eğer ilk 5 sütun toplamı son 5 sütundan toplamından küçük ise meme bölgesi sağa dayalı olarak belirlenmekte ve aynama işlemi sonucu sola dayalı olarak düzenlenmektedir.



Şekil 8.8. Örnek mamografi imgesi

Gürültü ve yapaylık giderimi sağlanmış, arka plandan arındırılmış ve sola dayalı hale getirilmiş mamografi imgelerine, pektoral kas tespiti için iki farklı algoritma kullanılarak bölge büyütme yöntemi uygulanmıştır. Kullanılan iki algoritmada da tek bir tohum noktası ile algoritma başlatılmış ve bölge büyütme algoritmasına P_R bölgesine eklenen son pikselin 4-komşuluğunun Q benzerlik koşulu açısından değerlendirilmesi ile devam edilmektedir. Kullanılan algoritmalardan birincisi Q benzerlik koşulu olarak bölge ortalaması (P_R_mean) ile bölgeye en son eklenen pikselin yeğinlik farkını kullanır. Algoritma, bu fark, önceden tanımlanmış maksimum fark değerinden küçük olduğu sürece devam etmekte ve bu koşul sağlandığı sürece 4-komşuluktaki piksellerden yeğinlik seviyesi bölge ortalamasına en yakın olan piksel bölgeye eklenmektedir. 1. algoritmada, I imgesinin r . satır c . sütun pikseli için kullanılan Q benzerlik koşulu Eşitlik 8.1'de ve algoritmanın iş akış şeması Şekil 8.9'da gösterilmiştir. Eşitlikteki T terimi bölge algoritması ile bölgeye eklenen pikseller arasında olması istenen maksimum fark değerini göstermektedir.

$$Q(r,c) = \begin{cases} 1 & , \quad I(r,c) - P_R_mean < T \\ 0 & , \quad \text{aksi takdirde} \end{cases} \quad (8.1)$$



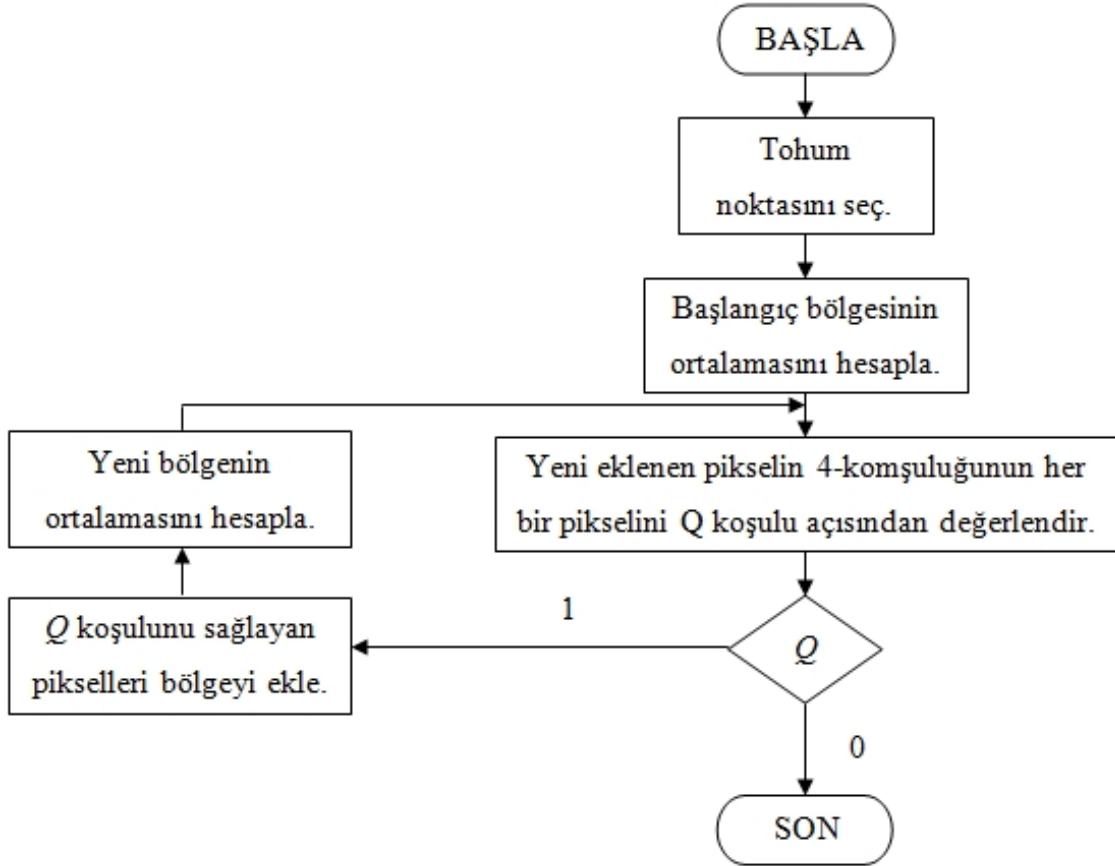
Şekil 8. 9. Bölge büyütme yöntemi için kullanılan birinci algoritmanın iş akış şeması

Kullanılan ikinci algoritmada ise Eşitlik 8.2’de verilen Q benzerlik koşulu kullanılmaktadır. Eşitlikteki P_R_mean ve T terimleri sırasıyla bölütlenmiş imgenin ortalaması ve önceden belirlenmiş eşik değerini ifade etmektedir. Algoritmanın iş akış şeması Şekil 8.10’da gösterilmiştir.

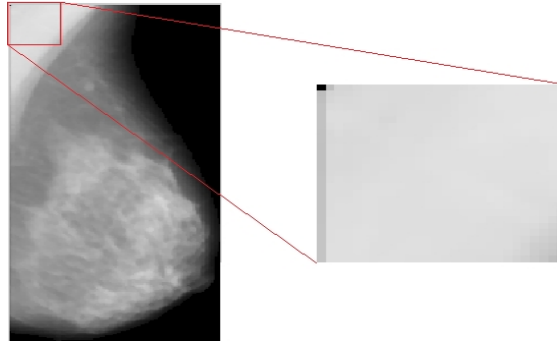
$$Q(r,c) = \begin{cases} 1 & , \quad P_R_mean - T < I(r,c) < P_R_mean + T \\ 0 & , \quad \text{aksi takdirde} \end{cases} \quad (8.2)$$

Her iki algoritmanın da başlangıç adımı olan tohum noktasının seçimi için pektoral kasların ön işlenmiş mamografi imgelerinin sol üst köşesinde olacağı ve kas yeğinlik seviyelerinin diğer bölgelerin yeğinlik seviyelerine göre daha yüksek (daha parlak) olacağı bilgileri kullanılmıştır. Bu doğrultuda, imgelerin sol üst köşesinden sütunlarının %25’i

boyutunda karesel bölge kesilmiş (S_R) ve bu bölgedeki en parlak piksel tohum noktası olarak seçilmiştir. Örnek bir mamografi imgesinde, tohum noktası seçimi ve algoritmaların benzerlik koşullarını tanımlanmakta kullanılan T eşik değerinin hesaplanması için ele alınan bölge Şekil 8.11’de gösterilmiştir.



Şekil 8. 10. Bölge büyütme yöntemi için kullanılan ikinci algoritmanın iş akış şeması

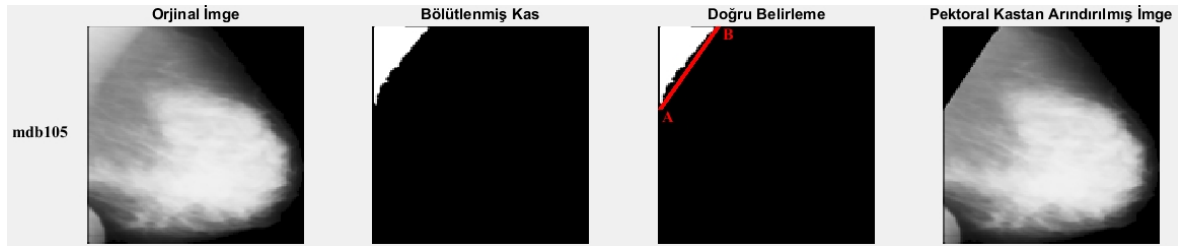


Şekil 8. 11. Tohum noktası seçimi ve algoritmaların benzerlik koşullarını tanımlamakta kullanılan eşik değerinin hesaplanması için ele alınan bölge

Her iki algoritmada da aynı T eşik değeri kullanılmıştır. Bu değer Eşitlik 8.3'te gösterildiği şekilde imgelerden kesilen S_R karesel bölgesinin standart sapmasının (σ_{S_R}) maksimum değerine (tohum noktasına - $\max(S_R)$) oranı olarak hesaplanmıştır. Bu şekilde, muhtemel pektoral kas olan tohum noktasından ancak standart sapma kadar farklı yeğlinliğe sahip piksellerin pektoral kas bölgesine dâhil edilmesi amaçlanmıştır.

$$T = \text{floor}\left(\frac{\sigma_{S_R}}{\max(S_R)}\right) \quad (8.3)$$

Bölge büyütme algoritmaları sonucunda bölütlenen imgelerde, pektoral kas bölgesini bir üçgen olarak düşünerek üçgenin kenar köşelerinin (A ve B noktaları) koordinatları tespit edilmiştir. Son olarak, Şekil 8.12'de gösterildiği üzere A ve B noktaları arasında $y = m \cdot x + b$ doğrusu belirlenmiş, bu doğrunun üstünde kalan alan pektoral kas olarak tanımlanmış ve kesilmiştir.



Şekil 8. 12. Bölütleme işlemi sonucunda pektoral kası kesilmiş bir mamografi imgesi örneği

Bölütleme işlemi sonucunda, pektoral kas bölgesindeki piksel yeğniliklerinin meme dokusundaki piksel yeğniliklerine benzer olduğu durumlarda, algoritmanın gürbüzlüğünü sağlamak amacı ile ön işlenmiş mamografi imgelerine ilk olarak histogram eşitleme uygulanarak zıtlıklar vurgulanmış ve sonrasında yukarıda anlatılan süreç gerçekleştirilmiştir.

Mamografi imgelerinden pektoral kas tespiti çalışmasının performansının artırılması için iyileştirme aşamasının gerekliliği saptanmıştır. Bu nedenle, bölge büyütme yöntemi algoritmaları sonucu pektoral kaslardan arındırılmış mamografi imgeleri öncelikle ikili seviyeye dönüştürülmüştür. İkili seviyeye dönüştürülmüş mamografi imgelerinde hesaplanan en büyük alanın 10.000 pikselden fazla olduğu durumlarda imgelerde pektoral

kas olmadığı kabul edilmiş, diğer durumlarda ise Şekil 8.12’de gösterildiği şekilde doğru belirleme işlemi sonucu belirlenen pektoral kaslar imgelerden çıkartılmıştır.

İyileştirme aşamasının ilk adımı olan imgelerin ikili seviyeye dönüştürülmesi işlemi için iki farklı eşik değeri kullanılmıştır. Eşik değerleri elde edilen pektoral kaslardan arındırılmış mamografi imgelerinin sol üst köşesinden sütunlarının %25’i boyutunda kesilmiş karesel bölgeden (S_{R_2}) elde edilmiştir. Bu adım için kullanılan ilk eşik değeri (T_1) Eşitlik 8.4’te verilmiştir. Eşitlikteki N terimi S_{R_2} bölgesindeki piksel sayısını ifade etmektedir.

$$T_1 = \frac{I}{N} \cdot \sum_{i=1}^N S_{R_2}(i) \quad (8.4)$$

İyileştirme aşamasının ilk adımı için kullanılan ikinci eşik değeri (T_2) Eşitlik 8.5’te verilmiştir. Eşitlikteki N terimi S_{R_2} bölgesindeki piksel sayısını, T_1 ve T terimleri ise sırasıyla ilk adımda kullanılan ilk eşik değerini ve bölge büyütme yöntemi için kullanılan eşik değerini ifade etmektedir.

$$T_2 = \begin{cases} T_1 & , \quad \sum_{i=1}^N S_{R_2}(i) \neq 0 \\ T & , \quad \sum_{i=1}^N S_{R_2}(i) = 0 \end{cases} \quad (8.5)$$

İyileştirme aşamasının ikinci adımında, ilk olarak, sadece pektoral kas bölgesinden ziyade meme dokusundan da kesim yapıldığı aşırı-bölütlenmiş ya da pektoral kas bölgesinin tamamen kesilemediği imgeler belirlenmiştir. Bu imgelerin her biri kendi ortalamaları eşik değeri kabul edilerek ikili-seviyeye dönüştürülmüştür. İkili-seviyeye dönüştürülmüş imgelere *Canny* kenar bulma detektörü uygulanarak meme dokusu konturları çıkartılmıştır. Meme konturları ve içlerinde kalan bölgelerin türevlerine göre kas-doku değişimlerinin olduğu yerler tespit edilmiş ve buralara doğru belirleme işlemi uygulanması sonucu elde edilen pektoral kaslar imgelerden çıkartılmıştır.

8.2.3. Mamografi imgelerinde meme doku tipi tespiti

Mamografi imgelerinin meme doku tipi tespitinin gerçekleştirilmesi için imgelerin $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ yön ve $d = \{1, 2, 3, 4\}$ uzaklık parametreleri ile GLCM matrisleri oluşturulmuş ve bu matrislerden Çizelge 6.2’de verilen 22-boyutlu doku öznitelikleri çıkartılarak yedi farklı öznitelik grubu oluşturulmuştur.

Birinci, ikinci, üçüncü ve dördüncü öznitelik grupları sırasıyla $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ yönleri ve $d = \{1, 2, 3, 4\}$ uzaklık parametreleri kullanılarak oluşturulan dört GLCM matrisinden çıkartılan 22-boyutlu doku özniteliklerinin alt alta dizilmesi ile elde edilen 88-boyutlu özniteliklerdir.

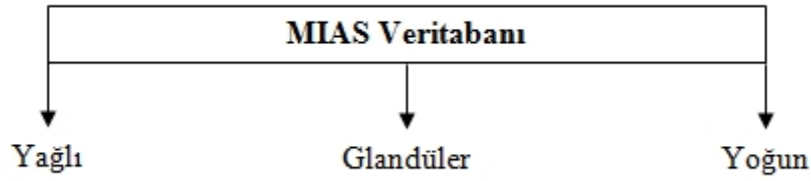
Beşinci öznitelik grubu, sırasıyla $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ yönleri ve $d = \{1, 2, 3, 4\}$ uzaklık parametreleri kullanılarak oluşturulan dört GLCM matrisinden çıkartılan 22-boyutlu doku özniteliklerinin ortalamalarının alınması ile hesaplanan 22-boyutlu ortalama vektörlerinin alt alta dizilmesi ile elde edilen 88-boyutlu özniteliklerdir.

Altıncı öznitelik grubu, sırasıyla birinci, ikinci, üçüncü ve dördüncü grup 88-boyutlu öznitelik vektörlerinin alt alta dizilmesi ile elde edilen 352-boyutlu özniteliklerdir.

Yedinci öznitelik grubu, 108-boyutlu LCP-tabanlı öznitelik vektörünün (Bkz. Çizelge 8.3) MIAS veritabanı için oluşturulması ve 88-boyutlu beşinci grup öznitelik vektörü ile alt alta dizilmesi sonucu elde edilen 196-boyutlu özniteliklerdir.

Oluşturulan tüm öznitelik grupları FLDA ve SVM sınıflandırıcıları kullanılarak 4-ve 5-katlamalı çapraz-doğrulama tekniği ile 1-aşamalı ve 2-aşamalı süreçlerde sınıflandırılmıştır.

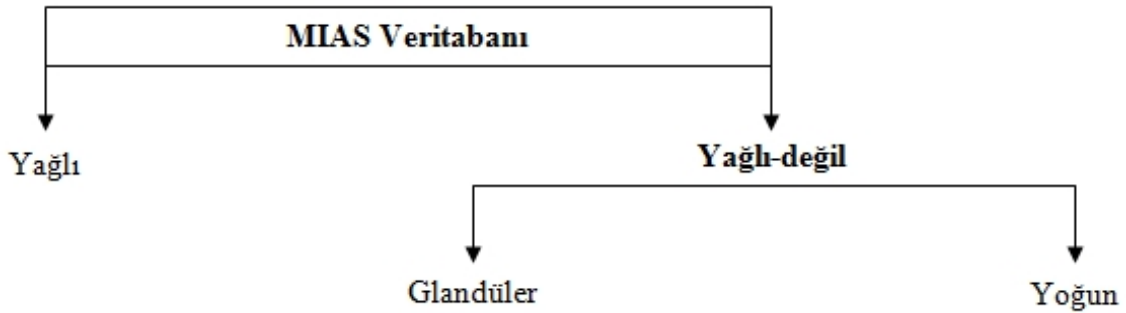
1-aşamalı meme doku tipi sınıflandırma sürecinde mamografi imgeleri doku tiplerine göre *yağlı*, *glandüler* ve *yoğun* olmak üzere üç sınıfa ayrılmıştır. 1-aşamalı meme doku tipi sınıflandırma sürecinin iş akış şeması Şekil 8.13’te gösterilmiştir.



Şekil 8. 13. 1-aşamalı meme doku tipi sınıflandırma sürecinin iş akış şeması

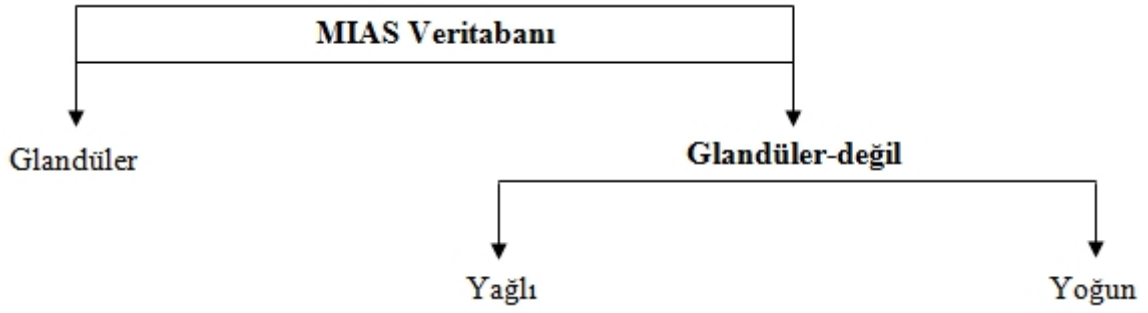
2-aşamalı meme doku tipi sınıflandırma çalışması “önce yağlı – sonra diğerleri”, “önce glandüler – sonra diğerleri” ve “önce yoğun – sonra diğerleri” olmak üzere üç farklı süreç izlenerek gerçekleştirilmiştir.

2-aşamalı meme doku tipi sınıflandırma çalışmasının “önce yağlı – sonra diğerleri” sürecinde, mamografi imgeleri öncelikle *yağlı* ya da *yağlı-değil* olmak üzere iki sınıfa ayrılmıştır. Sonrasında, birinci aşamada *yağlı-değil* olarak sınıflandırılan mamografi imgeleri, *glandüler* ya da *yoğun* olmak üzere tekrar iki sınıfa ayrılmıştır. “Önce yağlı – sonra diğerleri” sürecinin iş akış şeması Şekil 8.14’de gösterilmiştir.



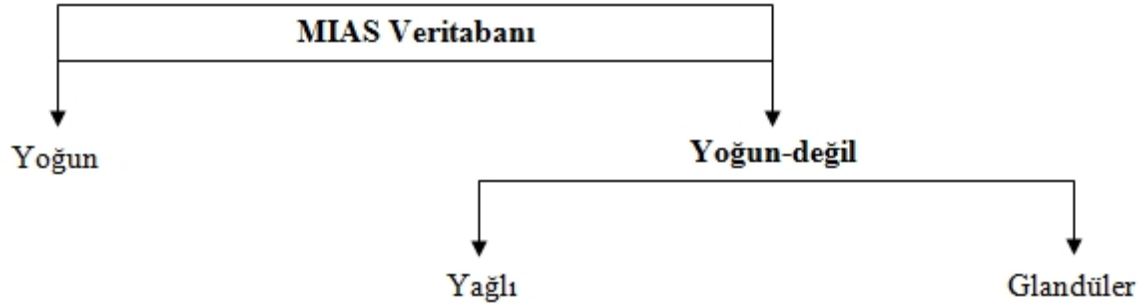
Şekil 8. 14. 2-aşamalı meme doku tipi sınıflandırma çalışması, “önce yağlı – sonra diğerleri” sürecinin iş akış şeması

“Önce glandüler – sonra diğerleri” sürecinin ilk aşamasında mamografi imgeleri *glandüler* ya da *glandüler-değil* olmak üzere iki sınıfa ayrılmıştır. Sürecin ikinci aşamasında ise, birinci aşamada doku tipi *glandüler-değil* olarak sınıflandırılan mamografi imgeleri, *yağlı* ya da *yoğun* olmak üzere iki sınıfa ayrılmıştır. “Önce glandüler – sonra diğerleri” sürecinin iş akış şeması Şekil 8.15’te gösterilmiştir.



Şekil 8. 15. 2-aşamalı meme doku tipi sınıflandırma çalışması, “önce glandüler – sonra diğerleri” sürecinin iş akış şeması

“Önce yoğun – sonra diğerleri” sürecinde ise ilk aşamada *yoğun* ya da *yoğun-değil* olmak üzere iki sınıfa ayrılan mamografi imgelerinden doku tipi *yoğun-değil* olarak sınıflandırılan mamografi imgelerinin *yağlı* ya da *glandüler* olarak sınıflandırılması gerçekleştirilmiştir. “Önce yoğun – sonra diğerleri” sürecinin iş akış şeması Şekil 8.16’da gösterilmiştir.



Şekil 8. 16. 2-aşamalı meme doku tipi sınıflandırma çalışması, “önce yoğun – sonra diğerleri” sürecinin iş akış şeması

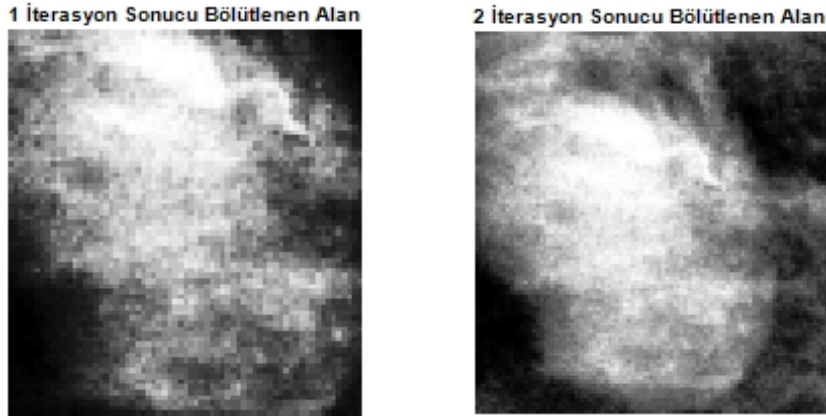
8.2.4. Mamografi imgelerinde ROI tespiti

Mamografi imgelerinden ROI’lerin tespiti için bölgesel tabanlı Chan-Vese aktif kontur modelleme yöntemi kullanılmıştır. Başlangıç konturunun belirlenmesi için kullanıcıdan anomali içerdiğinden şüphelenilen bir noktayı seçmesi istenmiştir. Seçilen bu noktaya Chan-Vese aktif kontur modelleme yöntemi bir ve iki iterasyon gerçekleştirilerek uygulanmış ve ROI alanları bölütlenmiştir. Bölütlenen ROI alanlarından hangisinin seçileceğinin belirlenmesi için bir karar ölçütü oluşturulmuştur. Bu karar ölçütü için, her iki iterasyon sonucunda Eşitlik 8.6’da görüldüğü gibi, bölütlenen alan ile bölütlenen alan

dışında kalan alanın ortalamalarının farklarının bölütlenen alanın ortalamasına göre yüzdelik değişim oranı hesaplanmıştır. Eşitlikte μ_i ve μ_d terimleri sırasıyla bölütlenen alanın ve bölütlenen alan dışında kalan alanın ortalamalarını göstermektedir.

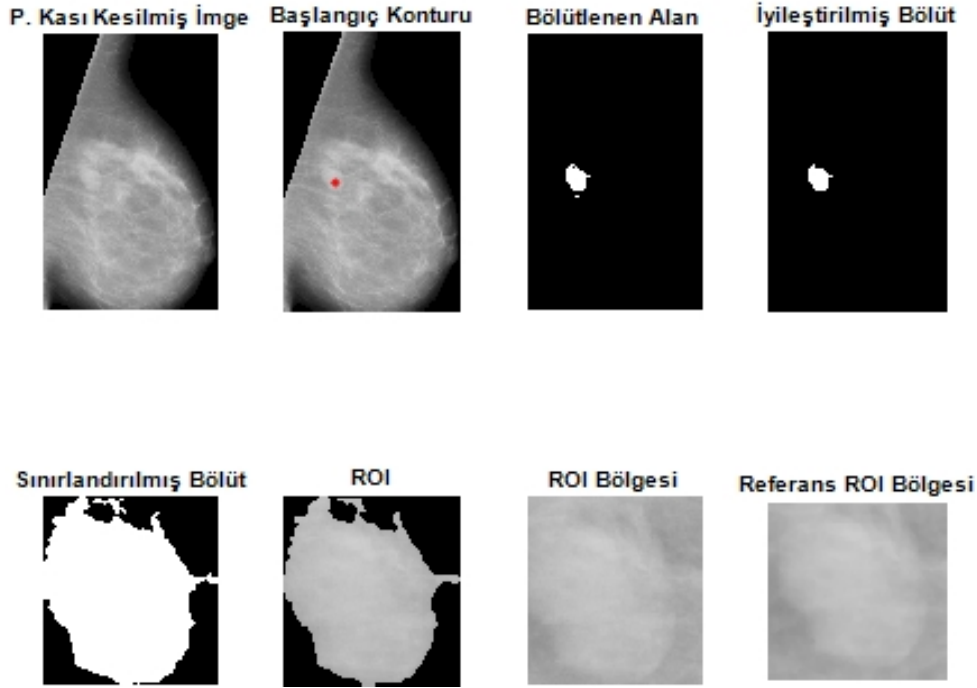
$$\text{Değişim Oranı} = 100 \cdot \frac{\mu_i - \mu_d}{\mu_i} \quad (8.6)$$

Değişim oranının düşük çıkması bölütlenen alan dışında kalan alanda yeğinlik açısından bölütlenen alan ile benzerlik gösteren alanların varlığını göstermektedir. Bu durumda, verimli bir bölütleme işlemi için, iterasyonlara devam edilmesi gerekmektedir. Bu nedenle, bir ve iki iterasyon sonucu elde edilen olası ROI alanlarından değişim oranı daha büyük hesaplanan bölütlenen ROI alanı olarak seçilmiştir. Herhangi bir mamografi imgesi için, bir ve iki iterasyon sonucu bölütlenen olası ROI alanları Şekil 8.17'de örneklendirilmiştir. ROI'lerin daha görülebilir olması için imgelere histogram eşitleme uygulanmıştır.



Şekil 8. 17. Örnek bir mamografi imgesi için bir ve iki iterasyon sonucu bölütlenen olası ROI alanları

Bölütlenen alanlara CCA uygulanarak ilgisiz pikseller atılmış ve alandaki boşluklar doldurularak iyileştirme yapılmıştır. Son olarak, iyileştirilen bölütler imge içinden kesilerek ROI'ler ve ROI'leri barındıran dikdörtgen bölgeler elde edilmiştir. Mamografi imgelerinde ROI tespiti aşamaları örnek bir mamografi imgesi için Şekil 8.18'de gösterilmiştir.



Şekil 8. 18. Örnek bir mamografi imgesi için ROI tespiti aşamaları

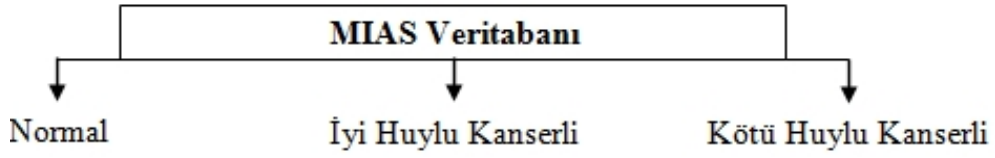
8.2.5. Öznitelik vektörlerinin oluşturulması

Tespit edilen ROI imgelerinin sağlık durumlarına göre sınıflandırılması aşamasında kullanılmak üzere alt bölüm 6.5'te anlatılan boyut, yoğunluk (meme yoğunluğu, ROI yağlılık derecesi, ROI yoğunluk derecesi), kontur (belirginlik, sınırlılık) ve şekil (yuvarlaklık, ovallık, eksantriklik, katılık, düzensizlik) öznitelikleri çıkartılmıştır. Bu öznitelikler, çok büyük değerlerin öznitelik vektör uzayını daraltmasını önlemek amacı ile 1-10 arası değerlere indekslenerek kullanılmıştır.

8.2.6. Meme kanseri teşhisi

MIAS veritabanı kullanılarak yapılan meme kanseri teşhisi çalışması 52 normal, 40 iyi huylu kanser ve 18 kötü huylu kanser içeren toplamda 110 tane seçilmiş ROI üzerinde gerçekleştirilmiştir. Bu aşamada, seçilen ROI'lerden oluşturulan 11-boyutlu mamografik özniteliklerin FLDA ve LLC sınıflandırıcıları kullanılarak 4- ve 5- katlamalı çapraz doğrulama tekniği ile sınıflandırılması gerçekleştirilmiştir.

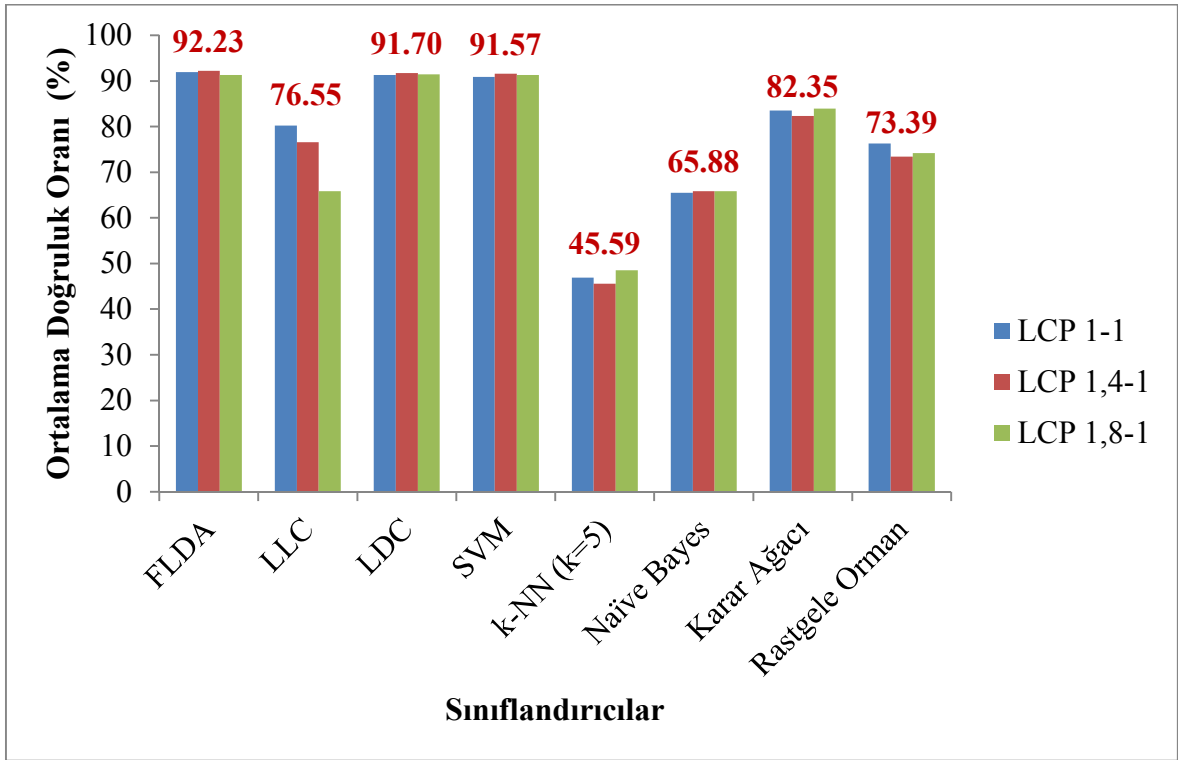
Sınıflandırma sürecinde, mamografi imgeleri meme doku tipi dikkate alınmaksızın direk olarak sağlık durumlarına göre *normal*, *iyi huylu kanser* ya da *kötü huylu kanser* olmak üzere üç sınıfa ayrılmıştır. MIAS veritabanı kullanılarak gerçekleştirilen meme kanseri teşhisi çalışmasının iş akış şeması Şekil 8.19'da gösterilmiştir.



Şekil 8. 19. MIAS veritabanı kullanılarak gerçekleştirilen meme kanseri teşhis çalışmasının iş akış şeması

9. BULGULAR VE TARTIŞMA

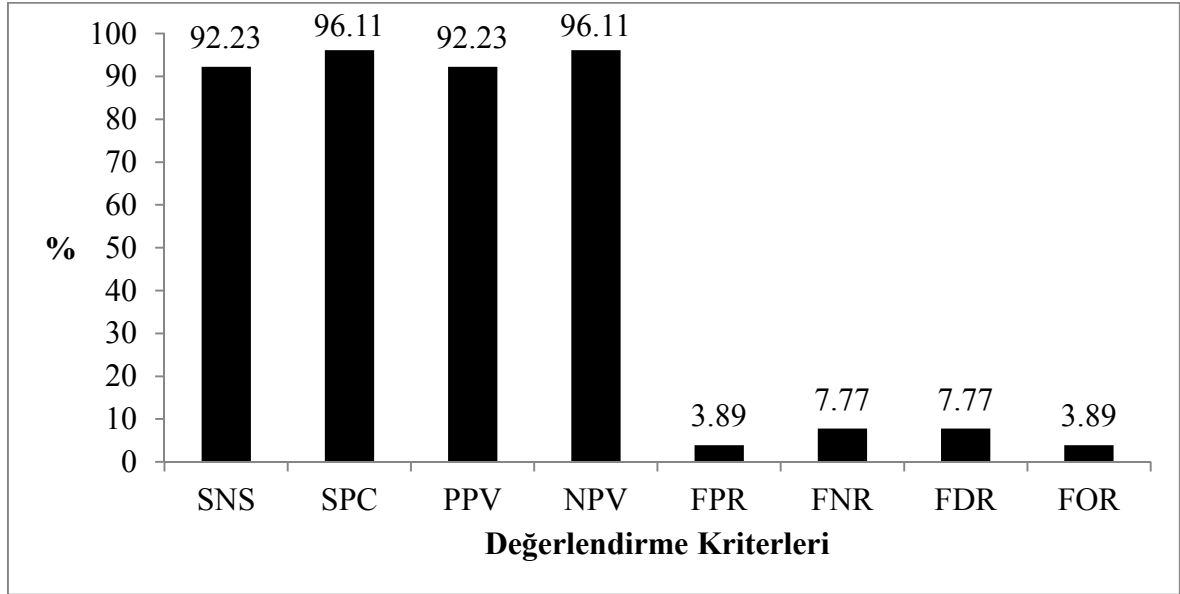
Tez çalışmasının birinci bölümünde IRMA veritabanı kullanılmıştır. Bu bölümde ilk olarak, veritabanının yalnızca yağlı doku tipine sahip mamografi imgeleri kullanılarak, 3-sınıflı sınıflandırma çalışması Bölüm 7’de anlatılan sınıflandırıcılar kullanılarak gerçekleştirilmiştir. Çalışmada her üç katsayı (Bkz. Çizelge 8.2, Çizelge 8.3) ile ağırlıklandırılan LCP-öznitelik vektörleri kullanılarak hesaplanmış 108-boyutlu LCP-tabanlı öznitelik vektörleri kullanılmıştır (Işıklı Esener vd., 2015). Çalışmadan elde edilen ortalama doğruluk oranları Şekil 9.1’de gösterilmiştir.



Şekil 9. 1. 108-boyutlu LCP-tabanlı öznitelik vektörlerinin 3-sınıflı sınıflandırılma çalışmasından elde edilen sonuçlar

108-boyutlu öznitelik vektörleri kullanılarak gerçekleştirilen 3-sınıflı sınıflandırma çalışmasında en yüksek doğruluk oranı (%92,23) LCP 1,4-1 kodlu LCP vektörleri tabanlı 108-boyutlu öznitelik vektörleri sınıflandırıldığında elde edilmiştir. Bu öznitelik vektörlerinin çalışmada en başarılı olan FLDA sınıflandırıcısı ile sınıflandırılmasına ait

performans değerlendirmesi Şekil 9.2’de, oluşturulan toplam karmaşıklık matrisi ise Çizelge 9.1’de verilmiştir.



Şekil 9. 2. LCP 1,4-1 kodlu 108-boyutlu öznelik vektörlerinin FLDA sınıflandırıcısı ile 3-sınıflı sınıflandırılma çalışmasına ait performans değerlendirmesi

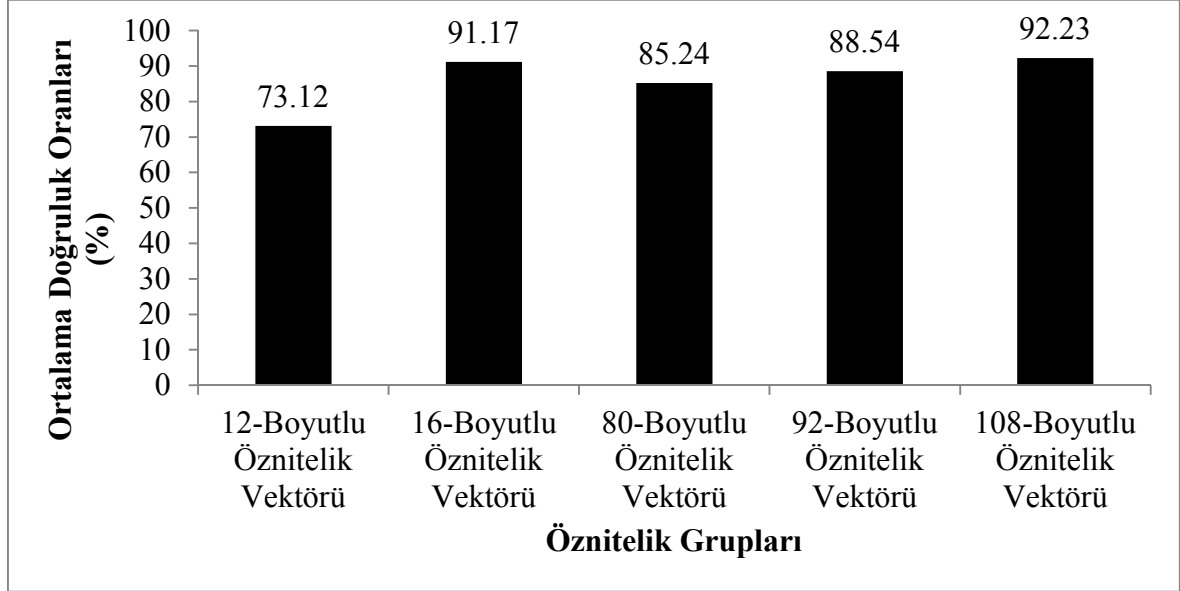
Çizelge 9. 1. LCP 1,4-1 kodlu LCP vektörleri tabanlı 108-boyutlu öznelik vektörlerinin FLDA sınıflandırıcısı ile 3-sınıflı sınıflandırılma çalışmasına ait toplam karmaşıklık matrisi

		Sınıflandırma Sonucu		
		Normal	İyi Huylu	Kötü Huylu
Referans Sınıflar	Normal	237	0	16
	İyi Huylu	0	253	0
	Kötü Huylu	42	1	210

Çizelge 9.1 incelendiğinde veritabanında bulunan iyi huylu kanserli mamografi imgeleri %100 doğruluk ile sınıflandırıldığı, kötü huylu kanser sınıfına ait yanlış-pozitif ve yanlış-negatif sonuçların normal sınıfına ait olduğu gözlenmektedir.

12-boyutlu istatistiksel, 16-boyutlu frekans-düzlemi, LCP 1,4-1 kodlu LCP öznelikleri ve bu üç öznelik grubundan elde edilen 108-boyutlu öznelik vektörlerinin çalışmada en iyi performansı gösteren FLDA sınıflandırıcısı ile sınıflandırıldığında elde edilen ortalama doğruluk oranları Şekil 9.3’te verilmiştir. Bu oranlar incelendiğinde, ilk üç

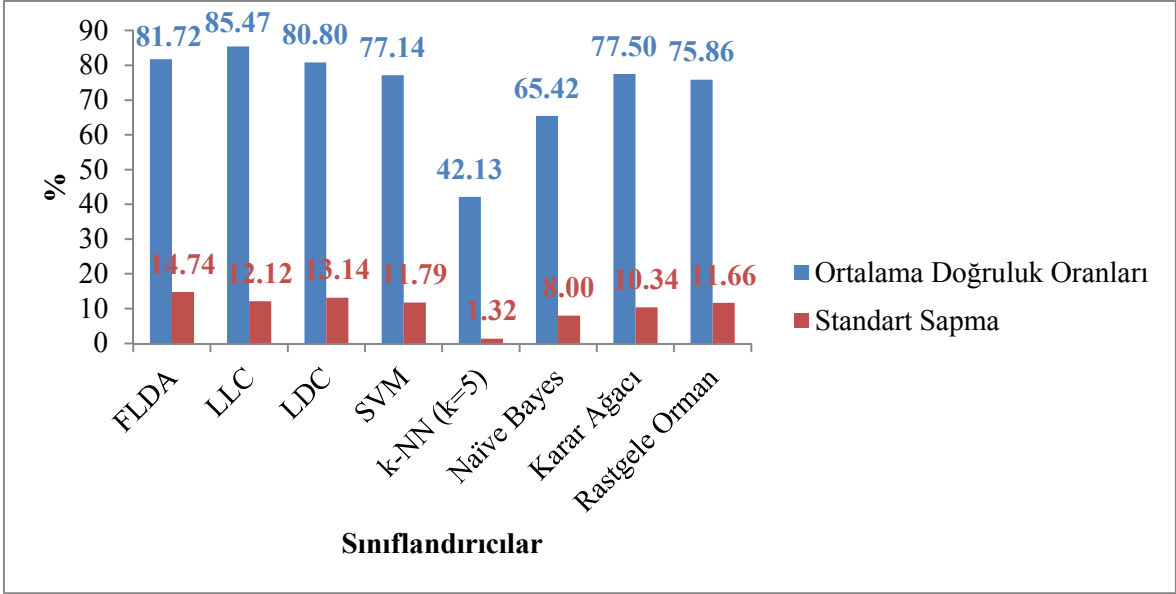
öznitelik grubunun ayırt ediciliğinin beraber kullanıldıklarında daha da arttığı gözlemlenmiştir.



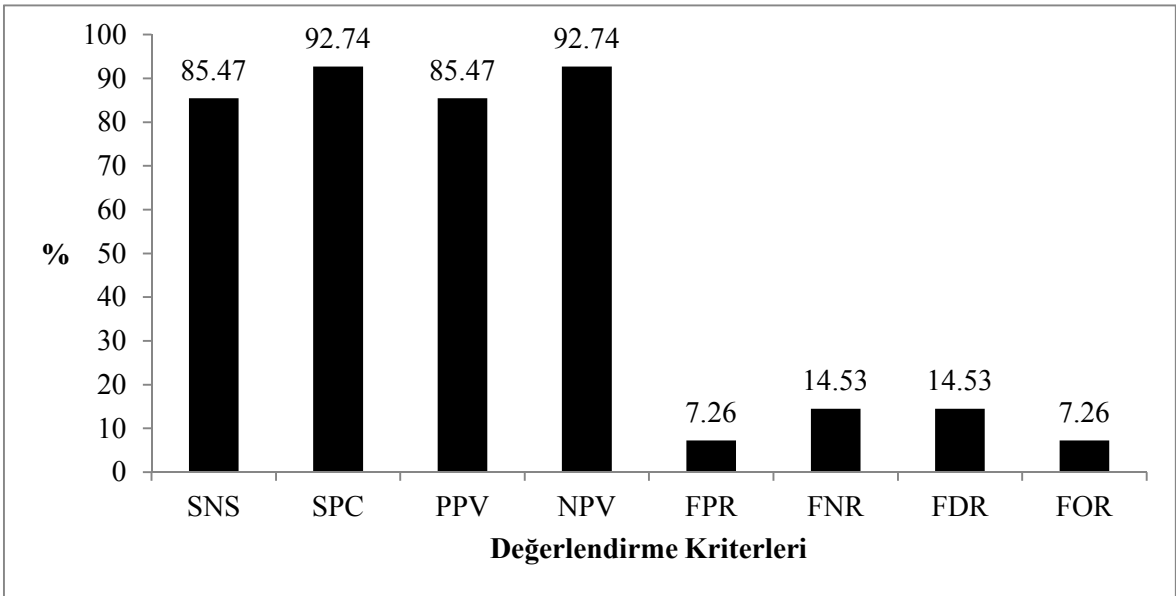
Şekil 9. 3. Tüm öznitelik gruplarının FLDA sınıflandırıcısı ile 3-sınıflı sınıflandırılması çalışmasından elde edilen sonuçlar

Tez çalışmasının birinci bölümünde, 3-sınıflı sınıflandırma çalışması veritabanındaki tüm imgelerin kullanılması ile 12-sınıflı olarak tekrarlanmıştır. 12-sınıflı sınıflandırma çalışmasında, 3-sınıflı sınıflandırma çalışmasında en ayırt edici öznitelik grubu olarak seçilen LCP 1,4-1 kodlu LCP vektörleri tabanlı 108-boyutlu öznitelik vektörleri kullanılmıştır(Işıklı Esener vd., 2017).

12-sınıflı sınıflandırma çalışması alt bölüm 8.1.4.2’de anlatıldığı gibi, 1- , 2- ve 3- aşamalı sınıflandırma süreçleri izlenerek gerçekleştirilmiştir.1-aşamalı sınıflandırma çalışmasından elde edilen ortalama doğruluk oranları ve katlamalar arası standart sapma değerleri Şekil 9.4’te gösterilmiştir. Çalışmada en yüksek doğruluk oranı (%85,47) LLC sınıflandırıcısı kullanıldığında elde edilmiştir. LLC sınıflandırıcısının performans değerlendirmesi Şekil 9.5’te verilmiştir.



Şekil 9. 4. 12-sınıflı 1-aşamalı sınıflandırma çalışmasından elde edilen sonuçlar



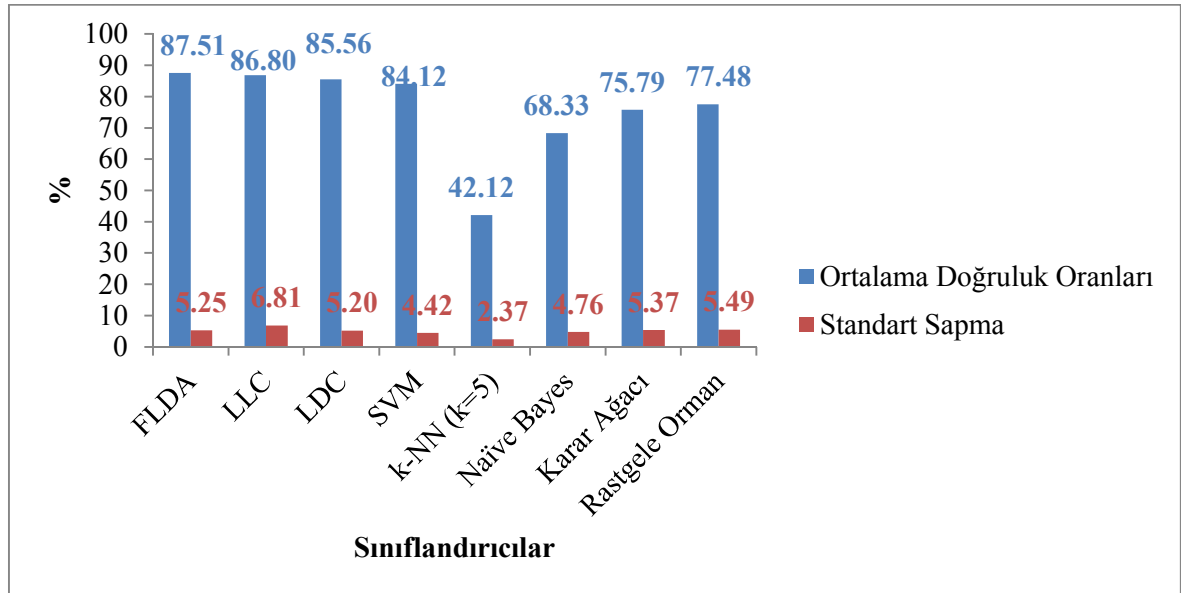
Şekil 9. 5. LLC sınıflandırıcısı ile 12-sınıflı 1-aşamalı sınıflandırma çalışmasına ait performans değerlendirmesi

LLC sınıflandırıcısı kullanılarak oluşturulan toplam karmaşıklık matrisi Çizelge 9.2’de verilmiştir. Çizelge 9.2 incelendiğinde, iyi ve kötü huylu kanser sınıflarından elde edilen yanlış-pozitif ve yanlış-negatif sonuçların normal sınıfına ait olduğu gözlenmektedir. Bu nedenle, sistem hassasiyetini artırmak amacı ile meme doku tipi belirlenmesi ve tüm doku tiplerinin kendi aralarında sağlık durumları açısından değerlendirilmesini sağlayan 2- ve 3-aşamalı sınıflandırma süreçleri uygulanmıştır (Işıklı Esener vd., 2017).

Çizelge 9. 2. LLC sınıflandırıcısı ile 12-sınıflı 1-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi

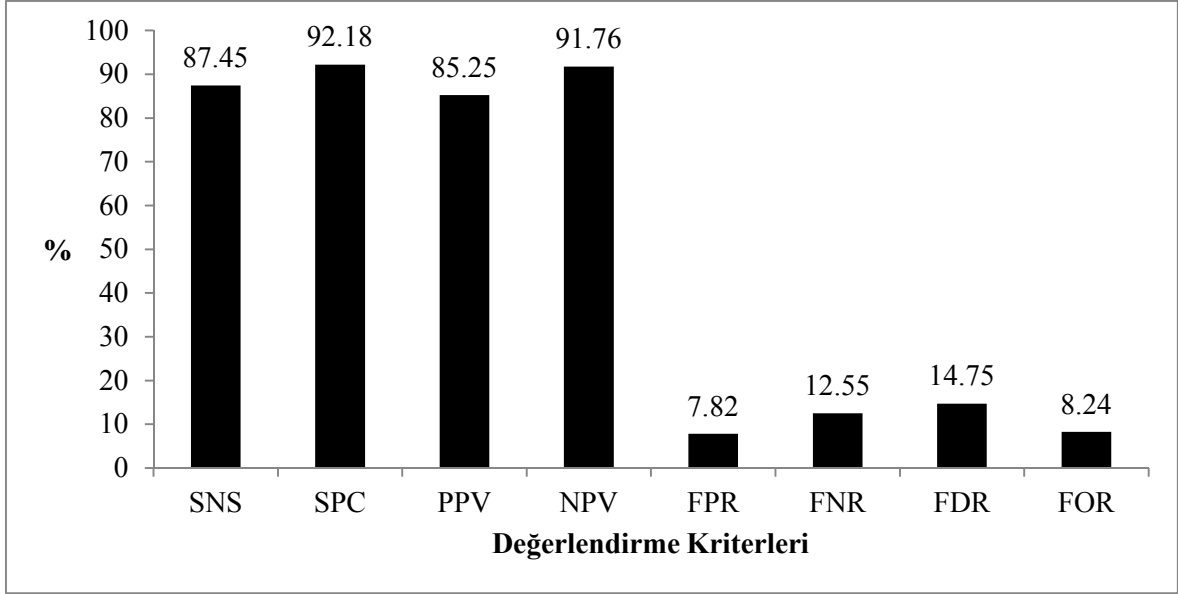
		Sınıflandırma Sonucu		
		Normal	İyi Huylu	Kötü Huylu
Referans Sınıflar	Normal	862	91	59
	İyi Huylu	123	889	0
	Kötü Huylu	166	2	844

2-aşamalı sınıflandırma çalışmasından elde edilen ortalama doğruluk oranları ve katlamalar arası standart sapma değerleri Şekil 9.6'da gösterilmiştir.



Şekil 9. 6. 12-sınıflı 2-aşamalı sınıflandırma çalışmasından elde edilen sonuçlar

Çalışmada en yüksek doğruluk oranı (%87,51) FLDA sınıflandırıcısı kullanıldığında elde edilmiştir. LLC algoritması, eşit ön olasılıklara sahip sınıflar arasında lojistik doğrusal modeller oluşturulabileceği kabulüne dayanmaktadır. 2-aşamalı sınıflandırma sürecinde sınıflardaki örnek sayısı, yani sınıfların ön olasılıkları, eşit olmadığı için LLC sınıflandırıcısının başarımlarının düşmesi olağan kabul edilmiştir. Bu süreçte en başarılı olan FLDA sınıflandırıcısının performans değerlendirmesi ve bu sınıflandırıcı kullanılarak oluşturulan toplam karmaşıklık matrisi sırasıyla Şekil 9.7 ve Çizelge 9.3'te verilmiştir.



Şekil 9. 7. FLDA sınıflandırıcısı ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait performans değerlendirmesi

Çizelge 9. 3. FLDA sınıflandırıcısı ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi

		Sınıflandırma Sonucu		
		Normal	İyi Huylu	Kötü Huylu
Referans Sınıflar	Normal	872	52	88
	İyi Huylu	186	826	0
	Kötü Huylu	142	5	865

Çizelge 9.3 incelendiğinde, iyi ve kötü huylu kanser sınıflarından elde edilen yanlış-pozitif ve yanlış-negatif sonuçların 1-aşamalı sınıflandırma sürecinde olduğu gibi normal sınıfına ait olduğu gözlenmektedir.

2-aşamalı sınıflandırma sürecinin performansını artırmak amacı ile bu süreçte en yüksek başarıyı sağlayan FLDA, LLC ve LDC sınıflandırıcılarına ait toplam karmaşıklık matrisleri incelenmiştir. LLC ve LDC sınıflandırıcılarından elde edilen toplam karmaşıklık matrisleri sırasıyla Çizelge 9.4 ve 9.5'te verilmiştir.

Çizelge 9. 4. LLC sınıflandırıcısı ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi

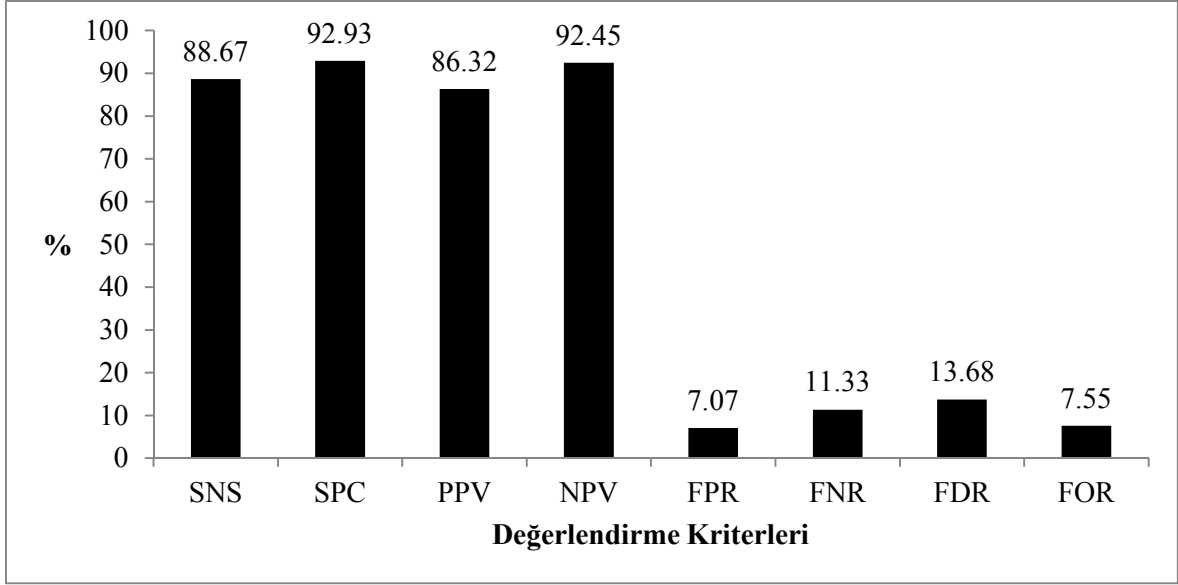
		Sınıflandırma Sonucu		
		Normal	İyi Huylu	Kötü Huylu
Referans Sınıflar	Normal	835	105	72
	İyi Huylu	135	874	3
	Kötü Huylu	153	0	859

Çizelge 9. 5. LDC sınıflandırıcısı ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi

		Sınıflandırma Sonucu		
		Normal	İyi Huylu	Kötü Huylu
Referans Sınıflar	Normal	877	27	108
	İyi Huylu	218	794	0
	Kötü Huylu	228	2	782

Çizelgeler 9.3, 9.4 ve 9.5 birlikte incelendiğinde, iyi huylu kanser / normal, kötü huylu kanser / normal ve iyi huylu kanser / kötü huylu kanser sınıfları arası FPR ve FNR değerlerinin en düşük olduğu sonuçların sırasıyla FLDA, LLC ve LDC sınıflandırıcıları kullanıldığında elde edildiği görülmüştür. Bu sınıflandırıcılar çoğunluk oylama (majority voting) tekniği kullanılarak birleştirilmiş ve ortalama doğruluk oranı %88,79'a yükseltilmiştir (Işıklı Esener vd., 2017). Çalışmanın performans değerlendirmesi Şekil 9.8 ve oluşturulan toplam karmaşıklık matrisi Çizelge 9.6'da verilmiştir.

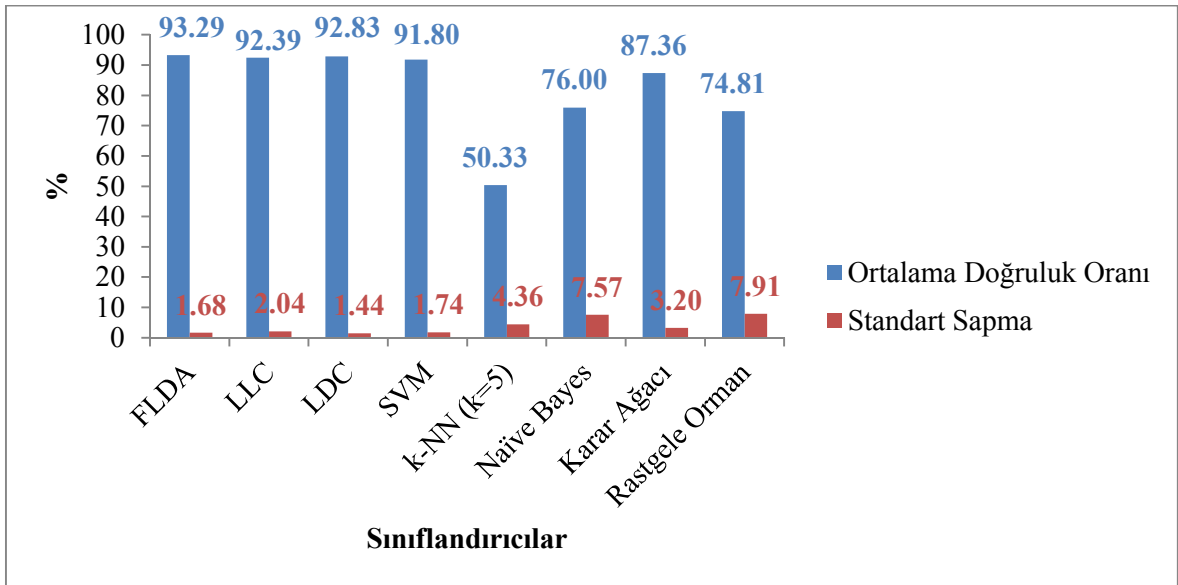
Çoğunluk oylama tekniği ile gerçekleştirilen 12-sınıflı 2-aşamalı sınıflandırma süreci sonucunda, FPR ve FNR değerlerinde azalma sağlanmıştır. Ancak, iyi ve kötü huylu sınıflarının FNR değerlerinin azaltılması gerekliliği Çizelge 9.6'dan açıkça görülmektedir. Bu nedenle 12-sınıflı sınıflandırma çalışmasına 3-aşamalı sınıflandırma süreci ile devam edilmiş ve bu süreçten elde edilen ortalama doğruluk oranları ve katlamalar arası standart sapma değerleri Şekil 9.9'da gösterilmiştir.



Şekil 9. 8. Çoğunluk oylama tekniği ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait performans değerlendirmesi

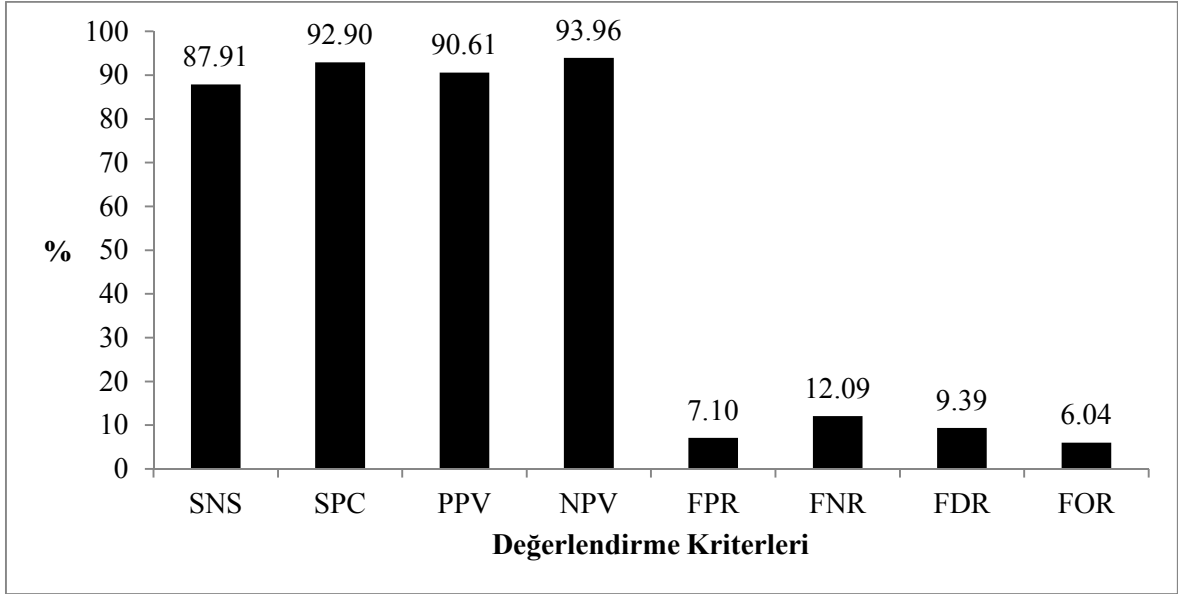
Çizelge 9. 6. Çoğunluk oylama tekniği ile 12-sınıflı 2-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi

		Sınıflandırma Sonucu		
		Normal	İyi Huylu	Kötü Huylu
Referans Sınıflar	Normal	887	41	84
	İyi Huylu	162	850	0
	Kötü Huylu	140	2	870



Şekil 9. 9. 12-sınıflı 3-aşamalı sınıflandırma çalışmasından elde edilen sonuçlar

Çalışmada en yüksek doğruluk oranı (%93,29) FLDA sınıflandırıcısı kullanıldığında elde edilmiştir. FLDA sınıflandırıcısının performans değerlendirmesi Şekil 9.10'da verilmiştir.



Şekil 9. 10. FLDA sınıflandırıcısı ile 12-sınıflı 3-aşamalı sınıflandırma çalışmasına ait performans değerlendirmesi

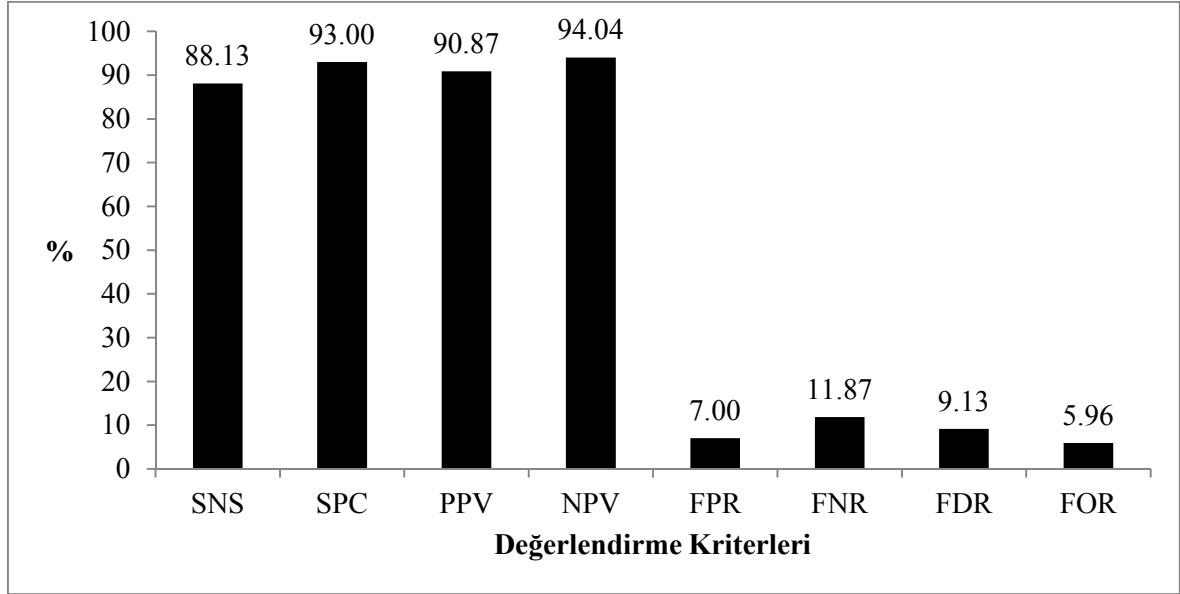
12-sınıflı 3-aşamalı sınıflandırma çalışmasında FLDA sınıflandırıcısı kullanılarak oluşturulan toplam karmaşıklık matrisi Çizelge 9.7'de verilmiştir.

Çizelge 9. 7. FLDA sınıflandırıcısı ile 12-sınıflı 3-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi

		Sınıflandırma Sonucu		
		Normal	İyi Huylu	Kötü Huylu
Referans Sınıflar	Normal	393	445	174
	İyi Huylu	174	629	209
	Kötü Huylu	139	260	613

2-aşamalı sınıflandırma sürecinde olduğu gibi, 3-aşamalı süreçte de en yüksek başarıyı sağlayan FLDA, LLC ve LDC sınıflandırıcıları çoğunluk tekniği kullanılarak birleştirilmiştir (Işıklı Esener vd., 2017). Bu durumda, doğruluk oranı %93,52 değerine yükseltilmiştir.

Çoğunluk oylama tekniği ile gerçekleştirilen 12-sınıflı 3-aşamalı sınıflandırma çalışmasının performans değerlendirmesi Şekil 9.11 ve oluşturulan toplam karmaşıklık matrisi Çizelge 9.8'de verilmiştir.



Şekil 9. 11. Çoğunluk oylama tekniği ile 12-sınıflı 3-aşamalı sınıflandırma çalışmasına ait performans değerlendirmesi

Çoğunluk oylama tekniği ile gerçekleştirilen 12-sınıflı 3-aşamalı sınıflandırma süreci sonucunda, iyi huylu ve kötü huylu kanser sınıflara ait FPR ve FNR değerlerinde azalma sağlanmıştır.

Çizelge 9. 8. Çoğunluk oylama tekniği ile 12-sınıflı 3-aşamalı sınıflandırma çalışmasına ait toplam karmaşıklık matrisi

		Sınıflandırma Sonucu		
		Normal	İyi Huylu	Kötü Huylu
Referans Sınıflar	Normal	393	444	175
	İyi Huylu	174	630	208
	Kötü Huylu	139	257	616

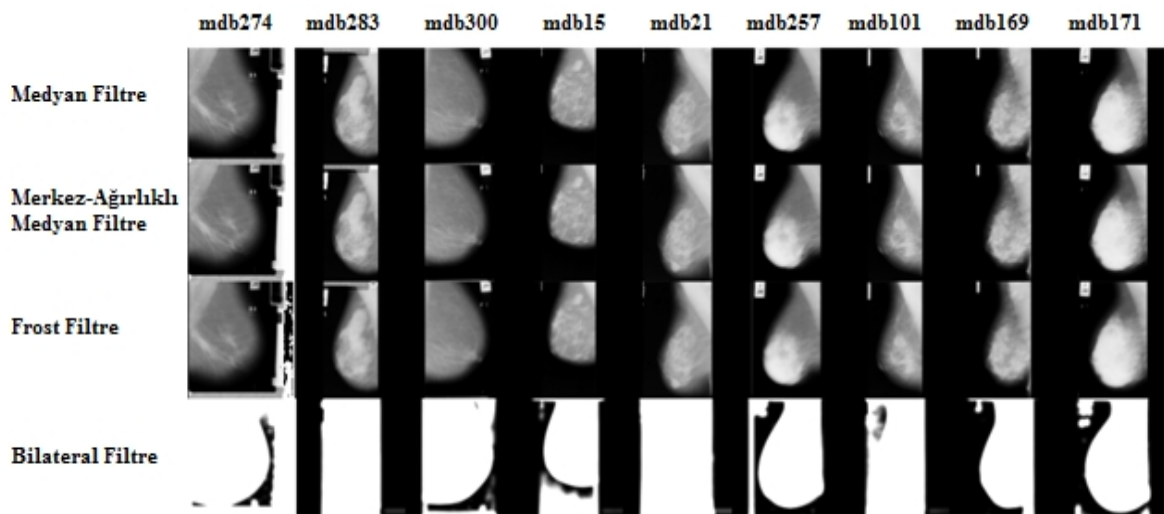
IRMA veritabanının 12-sınıflı sınıflandırma çalışmasında izlenen 1- , 2- ve 3- aşamalı sınıflandırma süreçlerinde en yüksek doğruluk 3-aşamalı sınıflandırma sürecinde elde edilmiştir. Doğruluk oranının yanı sıra, süreçlerde elde edilen standart sapmalar incelendiğinde 1-aşamalı sınıflandırma çalışmasında çok yüksek olan değerlerin aşaması

sayısı arttıkça düştüğü gözlemlenmiştir. Bu durum, 3-aşamalı sınıflandırma sürecinin katlamalarda kullanılan imgelerden bağımsız çalıştığını göstermektedir.

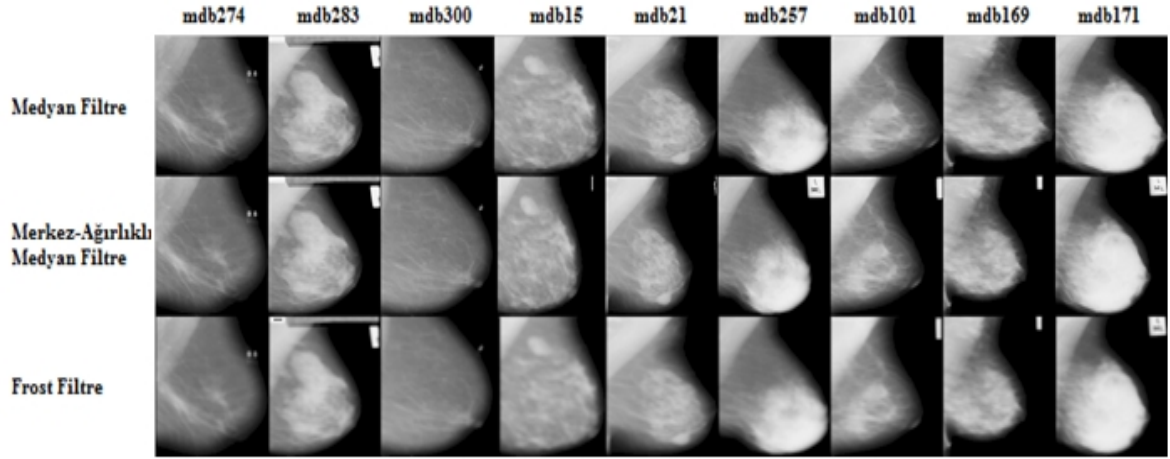
Tez çalışmasının ikinci bölümünde MIAS veritabanı kullanılmıştır. Bu bölümde ilk olarak, görüntüleme esnasında oluşan gürültülerin giderimi için imgelere ayrı ayrı medyan, CWM, Frost ve bilateral filtreler uygulanmıştır. Bu filtreler kullanılarak yapılan gürültü giderimi işlemi sonucunda elde edilen mamografi imgelerinden örnekler Şekil 9.12’de gösterilmiştir. Bilateral filtreleme sonucu düzgün sonuç alınamadığından bu filtre ile ön işlenmiş imgeler üzerinden çalışmaya devam edilmemiştir.

Gürültü giderimi aşamasından sonra imgelere morfolojik işlemler uygulanarak yapaylıkların giderimi sağlanmış, imgeler arka plandan arındırılmış ve pektoral kas bölgesi sol üst köşede kalacak şekilde sola dayalı olarak düzenlenmiştir. Bu işlemlerin medyan, CWM ve Frost filtreleme sonucu uygulanması ile elde edilen örnek mamografi imgeleri ve bu işlemler için imge başına ortalama çalışma süreleri sırasıyla Şekil 9.13 ve Çizelge 9.9’da verilmiştir.

Sonrasında, sola dayalı mamografi imgelerine akış şemaları verilen iki algoritma (Bkz. Şekil 8.9, Şekil 8.10) ayrı ayrı uygulanmış ve elde edilen bölgelere doğru oturtma yöntemi (line fitting) kullanılarak (Bkz. Şekil 8.12) pektoral kas bölgeleri bölütlenmiştir.



Şekil 9. 12. Gürültü giderimi sonucu elde edilen mamografi imgesi örnekleri

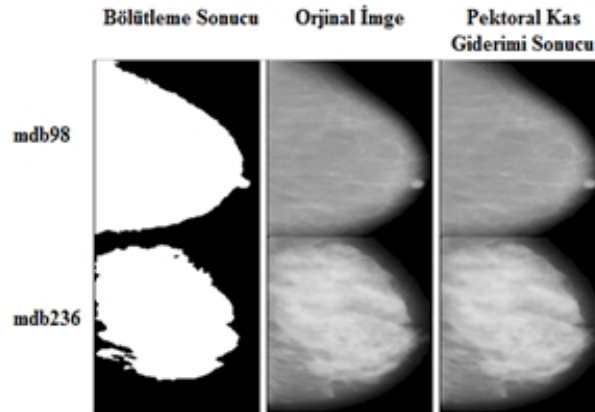


Şekil 9. 13. Gürültü ve yapaylık giderimi sağlanmış ve arka plandan arındırılmış mamografi imgesi örnekleri

Çizelge 9. 9. Gürültü giderimi için kullanılan filtrelerin parametre ve başarı oranları

Filtre Tipi	Filtreleme için İmge	Yapaylık Giderimi için	Başarı Oranı (%)
	Başına Ortalama Çalışma Süresi	İmge Başına Ortalama Çalışma Süresi	
Medyan	~16 ms	~31 ms	99,07
CWM	~1.2 sn	~31 ms	89,44
Frost	~6 sn	~31 ms	89,44

Pektoral kas bölütleme uygulaması sonucunda, Şekil 9.14'de örneklerinin olduğu sadece meme bölgesinin bölütlendiği ya da bölütleme işleminin gerçekleştirilemediği (tüm meme bölgesinin bölütlendiği) durumlarla karşılaşmıştır.



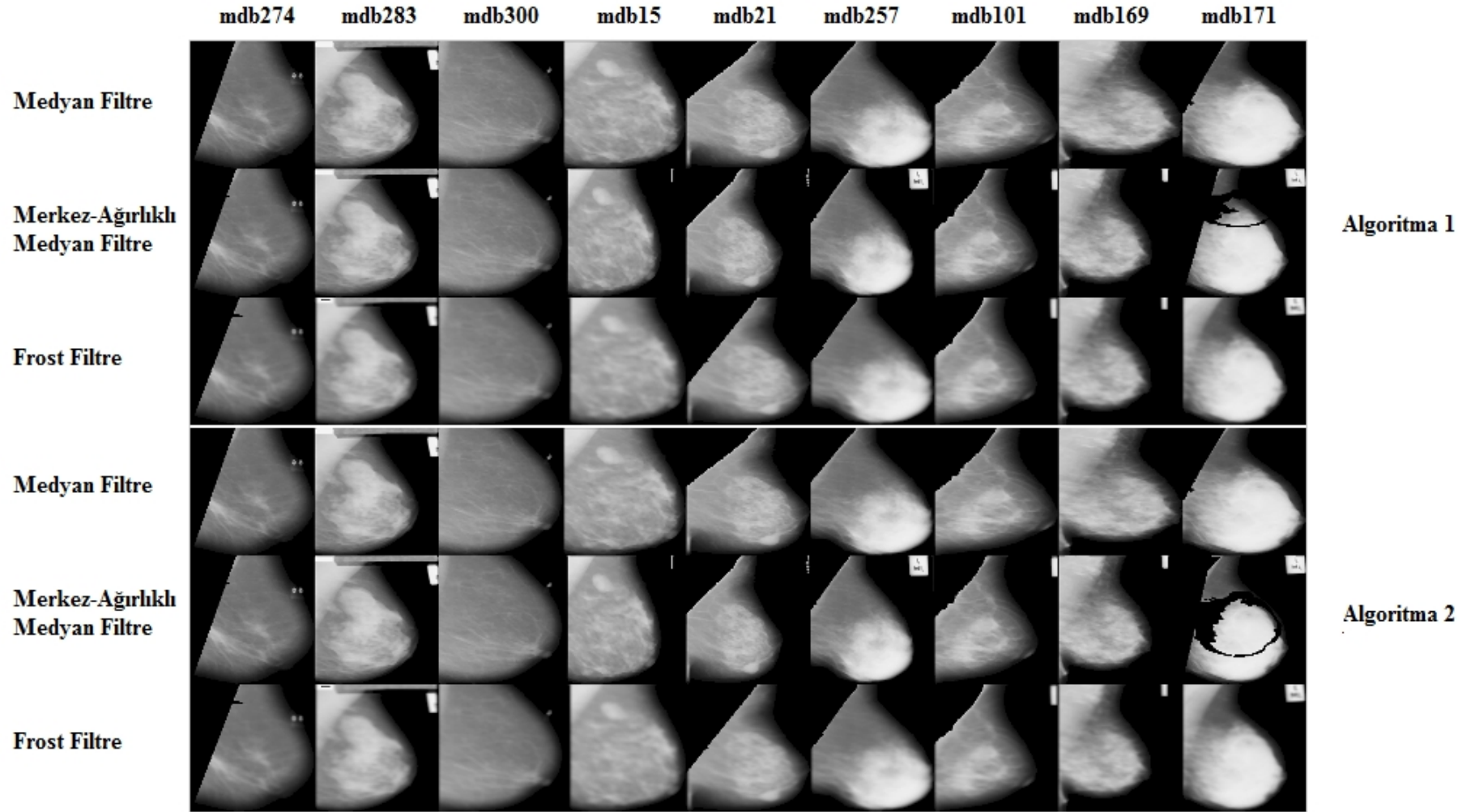
Şekil 9. 14. Bölütleme işlemi sonucunda pektoral kas bulundurmadığına karar verilen mamografi imgesi örnekleri

Yapılan incelemeler sonucunda, bu şekilde bölütlenmiş mamografi imgelerinde çoğunlukla pektoral kas bulunmadığı tespit edilmiştir ve bu imgelere pektoral kas giderimi amaçlı hiçbir algoritma uygulanmamıştır.

İki farklı bölge büyütme algoritması kullanılarak gerçekleştirilen pektoral kas bölütleme çalışmasının başarı oranları Çizelge 9.10'da örnek sonuçlar ise Şekil 9.15'te verilmiştir.

Çizelge 9. 10. Bölge büyütme yöntemi sonucu elde edilen başarı oranları

	Algoritma Tipi	
	Algoritma 1	Algoritma 2
Medyan	%52,48	%50,93
CWM	%55,59	%54,04
Frost	%39,13	%37,58



Şekil 9. 15. Bölge büyütme yöntemi sonucu elde edilen mamografi imgeleri

Şekil 9.15'te verilen mdb300 imgesinden de görüldüğü üzere, pektoral kas bölgesindeki piksel yeğliliklerinin göğüs dokusundaki piksel yeğliliklerine benzer olduğu durumlarda oldukça başarısız sonuçlar elde edilmektedir. Bu durumun üstesinden gelebilmek için ön işlenmiş mamografi imgelerine ilk olarak histogram eşitleme uygulanarak zıtlıklar vurgulanmış ve sonrasında yukarıda bahsedilen süreç gerçekleştirilmiştir. Bu durumda elde edilen sonuçlar Çizelge 9.11'de verilmiştir.

Çizelge 9. 11. Histogram eşitleme akabinde bölge büyütme yöntemi sonucu elde edilen başarı oranları

Filtre Tipi	Algoritma Tipi	
	Algoritma 1	Algoritma 2
Medyan	%63,04	%62,42
CWM	%62,42	%63,04
Frost	%64,91	%64,29

Bölge büyütme yöntemi ardından T_1 eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları Çizelge 9.12'de ve histogram eşitleme akabinde bölge büyütme yöntemi uygulanması ardından T_1 eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları Çizelge 9.13'te verilmiştir.

Çizelge 9. 12. Bölge büyütme yöntemi ardından T_1 eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları

Filtre Tipi	Algoritma Tipi	
	Algoritma 1	Algoritma 2
Medyan	%67,70	%63,35
CWM	%63,98	%58,70
Frost	%58,39	%57,14

Bölge büyütme yöntemi ardından T_2 eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları Çizelge 9.14'te ve histogram eşitleme akabinde bölge büyütme yöntemi uygulanması ardından T_2 eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları Çizelge 9.15'te verilmiştir.

Çizelge 9. 13. Histogram eşitleme akabinde bölge büyütme yöntemi uygulanması ardından T_1 eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları

Filtre Tipi	Algoritma Tipi	
	Algoritma 1	Algoritma 2
Medyan	%70,19	%71,43
CWM	%69,25	%68,94
Frost	%74,78	%73,91

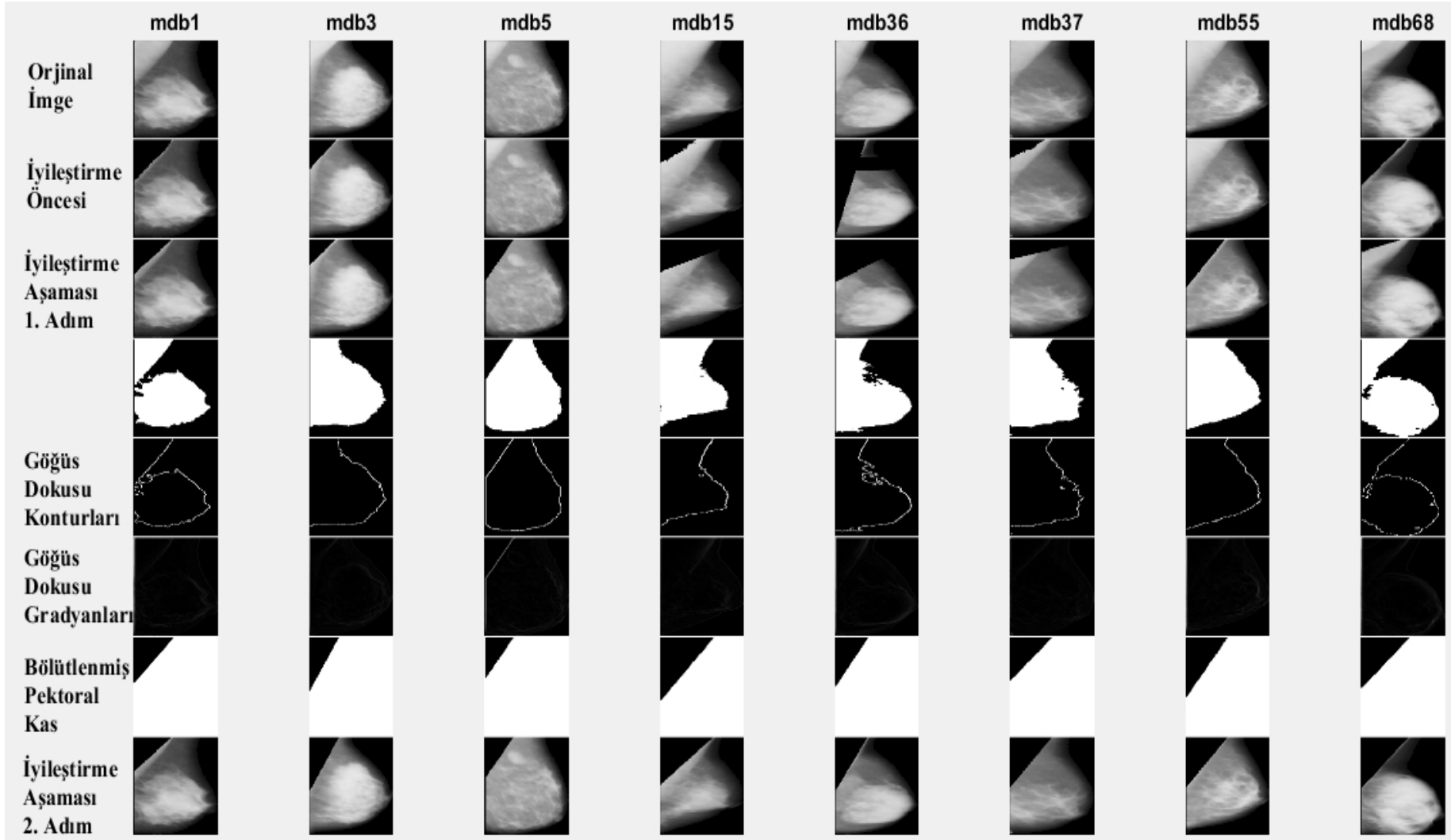
Çizelge 9. 14. Bölge büyütme yöntemi ardından T_2 eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları

Filtre Tipi	Algoritma Tipi	
	Algoritma 1	Algoritma 2
Medyan	%51,55	%59,94
CWM	%58,70	%59,32
Frost	%51,55	%51,55

Çizelge 9. 15. Histogram eşitleme akabinde bölge büyütme yöntemi uygulanması ardından T_2 eşik değeri kullanılarak iyileştirme yapıldığında elde edilen başarı oranları

Filtre Tipi	Algoritma Tipi	
	Algoritma 1	Algoritma 2
Medyan	%74,84	%71,12
CWM	%74,53	%74,22
Frost	%72,05	%71,12

İyileştirme aşamasının ikinci adımı, ilk adımda en yüksek başarı oranının (%74,84) sağlandığı, medyan filtreleme uygulamış imgelere histogram eşitleme akabinde bölge büyütme yöntemi uygulanması ardından T_2 eşik değeri kullanılarak gerçekleştirilen iyileştirme sonucu elde edilen imgelere uygulanmıştır. Bu durumda MIAS veritabanında bulunan 322 imgenin 304'ünden pektoral kaslar kabul edilebilir düzeyde kesilerek %94,40 oranında başarı sağlanmıştır. İyileştirme aşamasının 2. adımı sonucunda elde edilen pektoral kaslardan arındırılmış mamografi imge örnekleri ve uygulama aşamaları Şekil 9.16'da gösterilmiştir.



Şekil 9. 16. Örnek mamografi imgelerinde pektoral kas bölütleme ve iyileştirme adımları

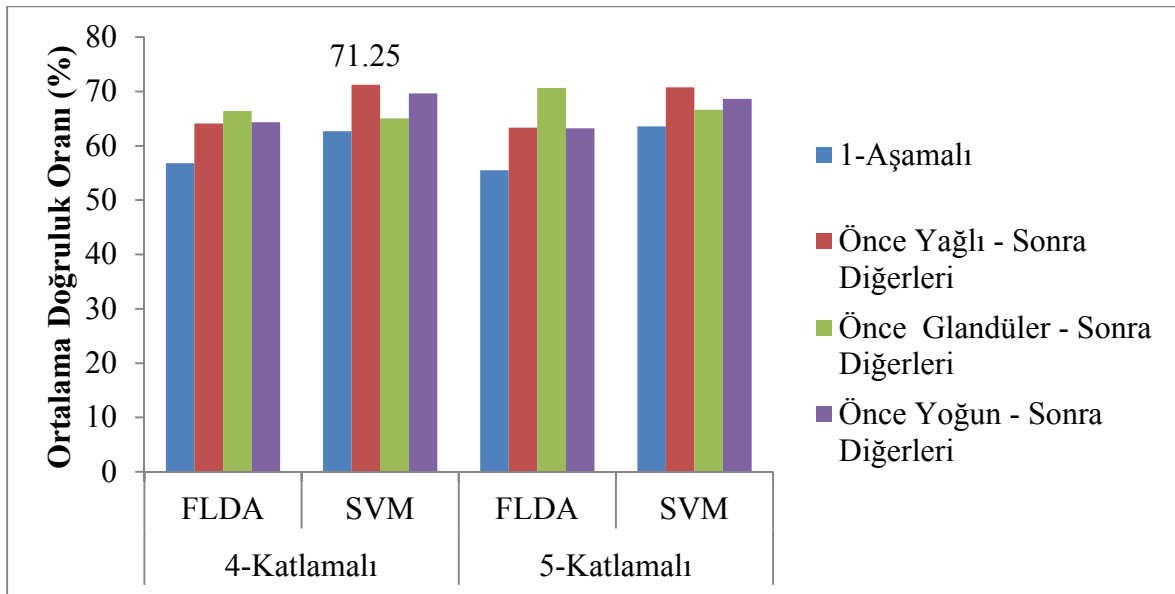
Mamografi imgelerinin pektoral kaslardan arındırılması çalışmasının performansı *The Area Normalized Error* kriterine göre değerlendirilmiştir. Bu kriterlere göre değerlendirme sonuçları ve sonuçların literatürde yer alan bazı çalışmalar ile karşılaştırılması Çizelge 9.16'da verilmiştir.

Çizelge 9. 16. Mamografi imgelerinde pektoral kas tespiti değerlendirme sonuçları

Değerlendirme Kriterleri	Önerilen	Özvektör Bölütleme (Abdellatif vd., 2012)	Hough Dönüşümü (Camilus vd., 2011)	Gabor Dalgacıkları (Camilus vd., 2011)	Adaptif Piramid (Camilus vd., 2011)	Minimum Tarama Ağacı (Camilus vd., 2011)	Graph-Cut Based Merging (Camilus vd., 2011)	Watershed Dönüşümü (Camilus vd., 2011)
FP_m	0,0415	0,0074	0,0198	0,0058	0,0371	0,0255	0,0064	0,0085
FN_m	0,0328	0,0412	0,2519	0,0577	0,0595	0,1168	0,0558	0,0488
FP < 0.05 & FN < 0.05	322	51	10	45	50	40	43	46
min(FP , FN) < 0.05 & 0.05 < max(FP , FN) < 0.10	0	18	0	0	18	20	19	25
min(FP , FN) < 0.05 & max(FP , FN) > 0.10	0	10	0	0	11	18	22	13
0.05 < FP < 0.10 & 0.05 < FN < 0.10	0	1	8	22	0	0	0	0
0.05 < min(FP , FN) < 0.10 & max(FP , FN) > 0.10	0	—	0	0	0	1	0	0
FP > 0.10 & FN > 0.10	0	0	66	17	5	3	0	0
Hausdorff Uzaklığı	15,97	—	7,08	3,84	—	—	—	3,85

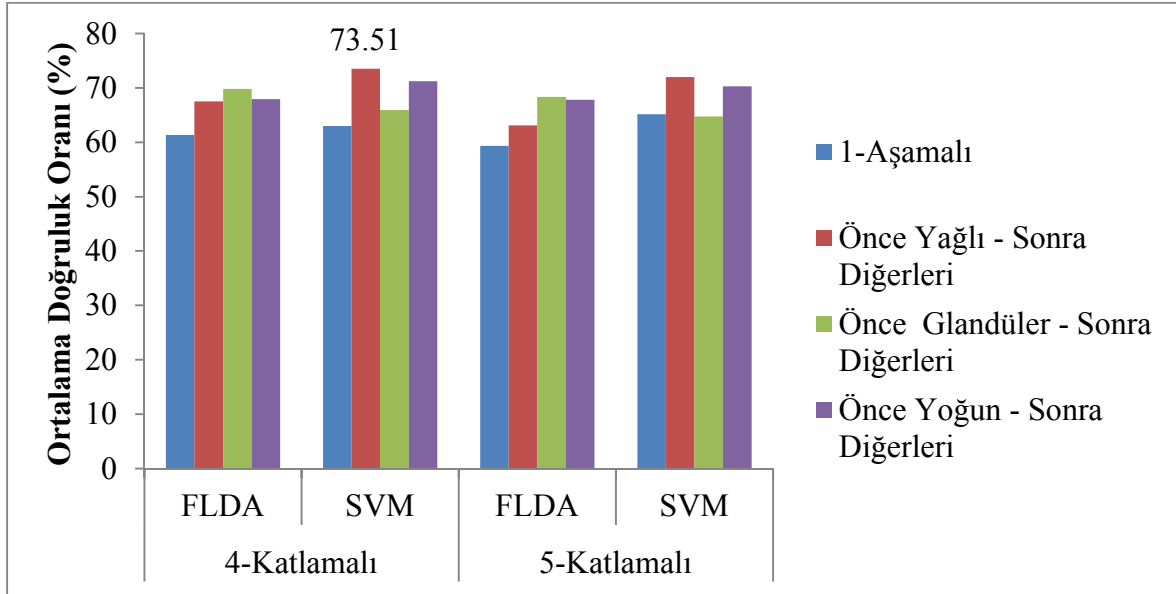
Çizelge 9.16’da verilen sonuçlar incelendiğinde, elde ettiğimiz imge başına ortalama yanlış-pozitif sayısının diğer çalışmalara göre daha yüksek oluşu referansa göre pektoral kas olmayan pikselleri algoritmamızın daha fazla kas olarak tespit ettiğini göstermektedir. İmge başına ortalama yanlış-negatif sayısının daha düşük oluşu ise referansa göre pektoral kas olan piksellerin algoritmamız tarafından daha az kaçırıldığını göstermektedir. Bu durumda, çalışmalarda kullanılan imge sayısı da düşünüldüğünde, elde ettiğimiz sonuçların başarısını göstermektedir.

Pektoral kas tespiti aşamasından sonra imgelerin meme doku tipi tespiti gerçekleştirilmiştir. Meme doku tipi tespiti çalışması, MIAS veritabanında yer alan 322 tane imgeden pektoral kasları kabul edilebilir düzeyde kesilmiş 304 imge üzerinde gerçekleştirilmiştir. Bu imgelerin 103 tanesi yağlı, 98 tanesi yağlı-glandüler ve geriye kalan 103 tanesi de yoğun-glandüler doku tipine sahiptir. İmgelerden doku öznitelikleri (Bkz. Çizelge 6.2) alt bölüm 8.2.3’de anlatıldığı gibi yedi farklı öznitelik grubu şeklinde çıkartılmıştır. Bu öznitelikler grupları FLDA ve SVM sınıflandırıcıları kullanılarak 4-ve 5-katlamalı çapraz-doğrulama tekniği ile 1-aşamalı ve 2-aşamalı süreçlerde sınıflandırılmıştır. Birinci öznitelik grubu kullanılarak yapılan meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları Şekil 9.17’de gösterilmiştir.



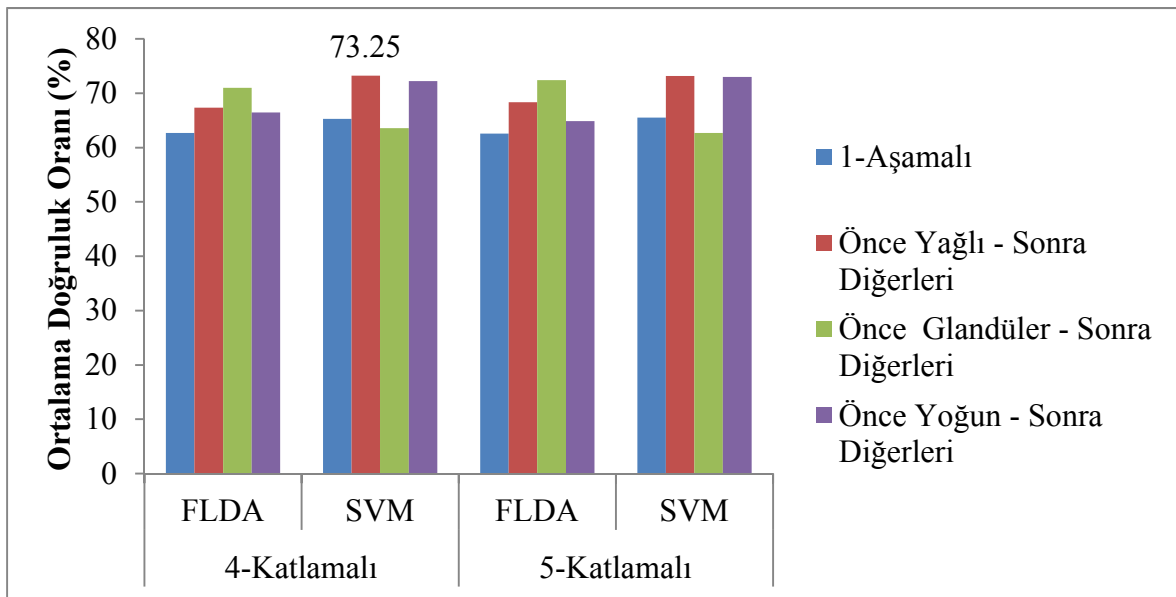
Şekil 9. 17. Birinci öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları

İkinci öznitelik grubu kullanılarak yapılan meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları Şekil 9.18’de gösterilmiştir.



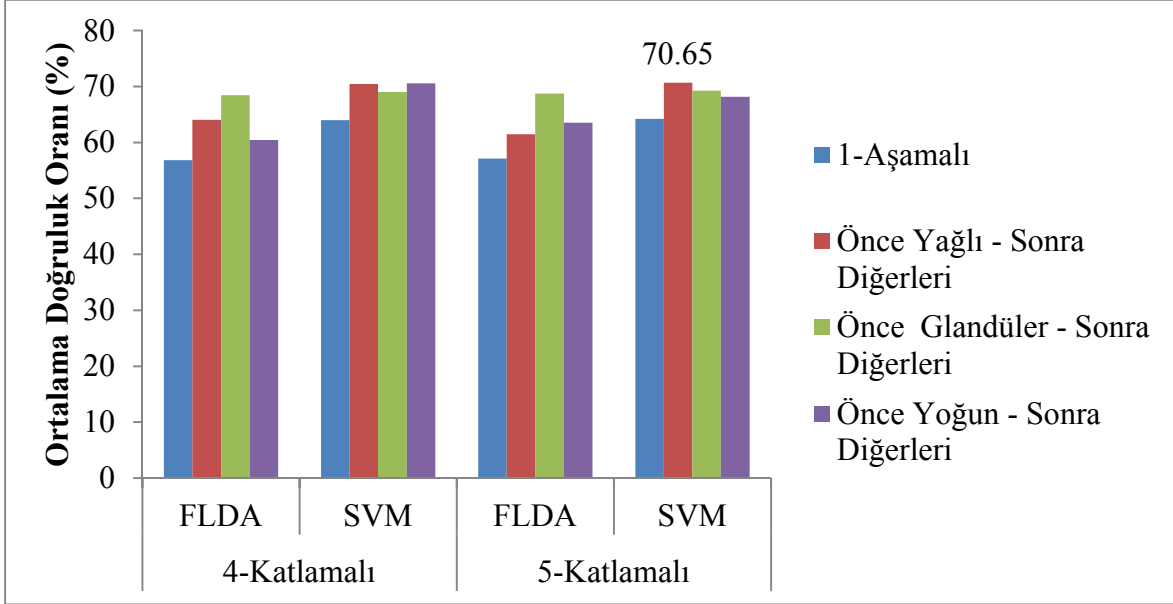
Şekil 9. 18. İkinci öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları

Üçüncü öznitelik grubu kullanılarak yapılan meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları Şekil 9.19’da gösterilmiştir.



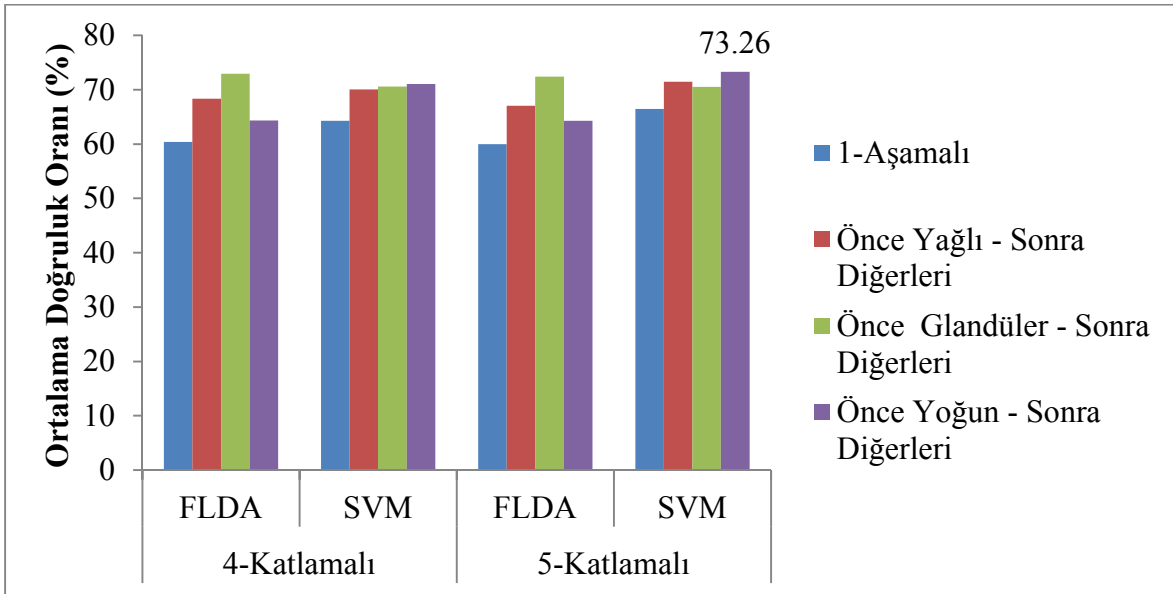
Şekil 9. 19. Üçüncü öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları

Dördüncü öznitelik grubu kullanılarak yapılan meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları Şekil 9.20’de gösterilmiştir.



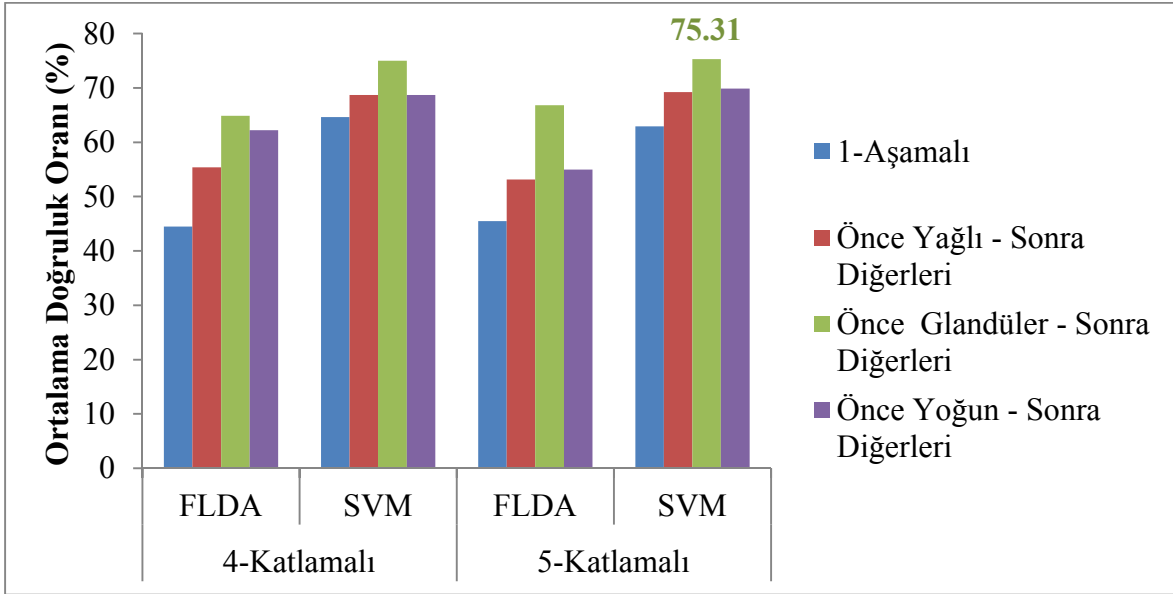
Şekil 9. 20. Dördüncü öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları

Beşinci öznitelik grubu kullanılarak yapılan meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları Şekil 9.21’de gösterilmiştir.



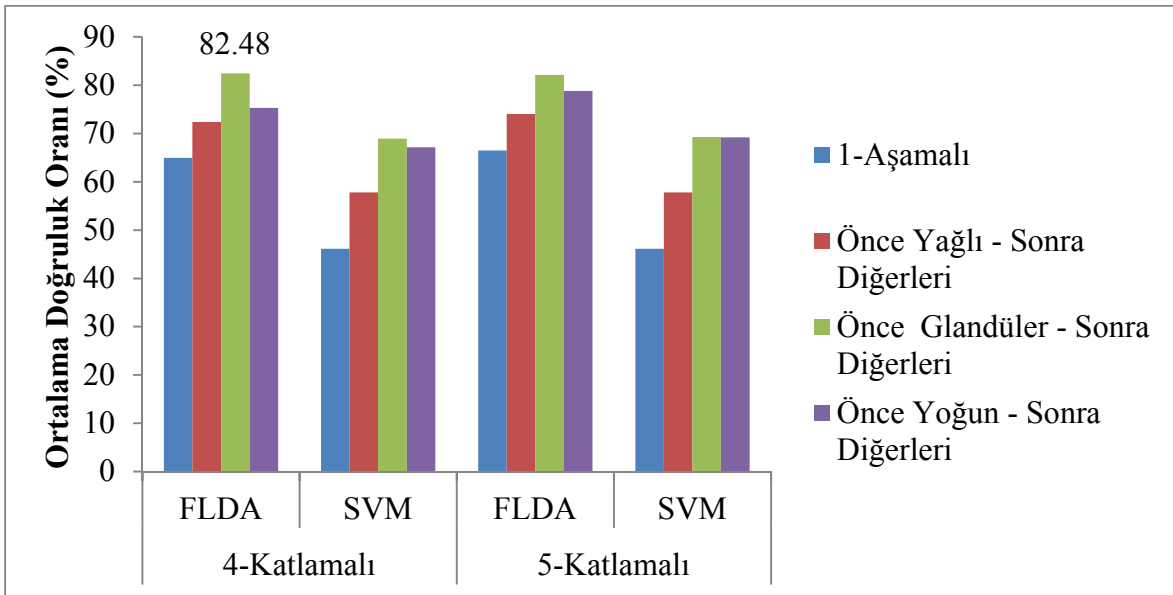
Şekil 9. 21. Beşinci öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları

Altıncı öznitelik grubu kullanılarak yapılan meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları Şekil 9.22’de gösterilmiştir.



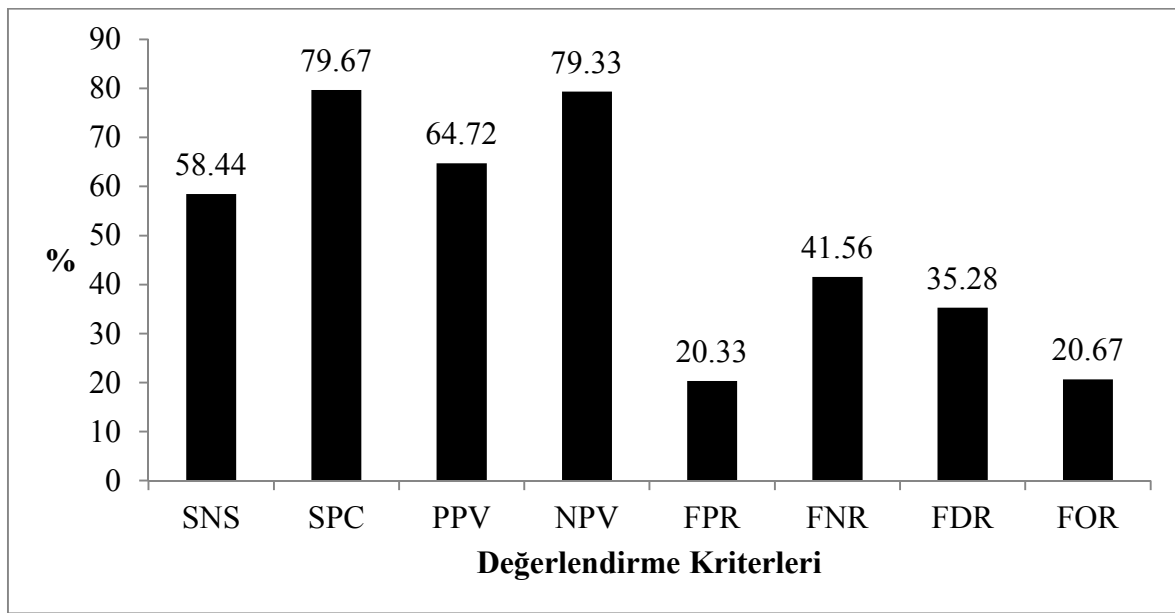
Şekil 9. 22. Altıncı öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları

Yedinci öznitelik grubu kullanılarak yapılan meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları Şekil 9.23’te gösterilmiştir.



Şekil 9. 23. Yedinci öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmalarından elde edilen ortalama doğruluk oranları

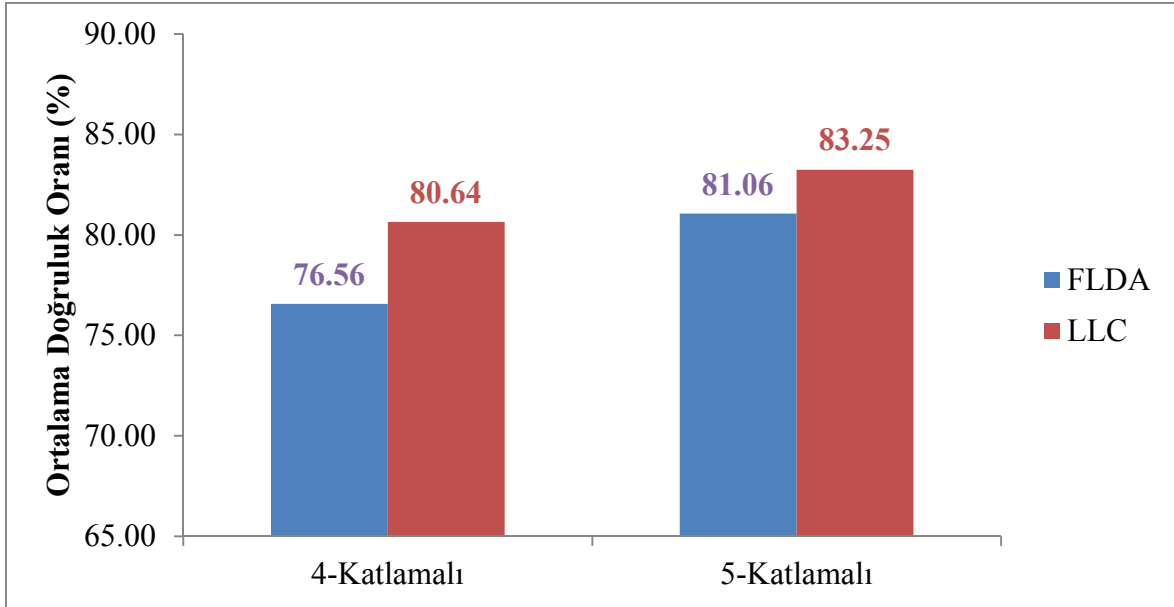
Yukarıdaki çizelgeler incelendiğinde meme doku tipi tespiti en yüksek %82,48'lik doğruluk oranı ile gerçekleştirildiği görülmektedir. Bu oran, yedinci öznitelik grubunun SVM sınıflandırıcısı ile “Önce Glandüler – Sonra Diğerleri” sınıflandırma sürecinde 5-katlamalı çapraz doğrulama tekniği kullanılarak sınıflandırıldığında sağlanmıştır (Işıklı Esener vd., 2016). Yedinci öznitelik grubunun belirtilen süreçte sınıflandırmasına ilişkin performans değerlendirmesi Şekil 9.24'te verilmiştir.



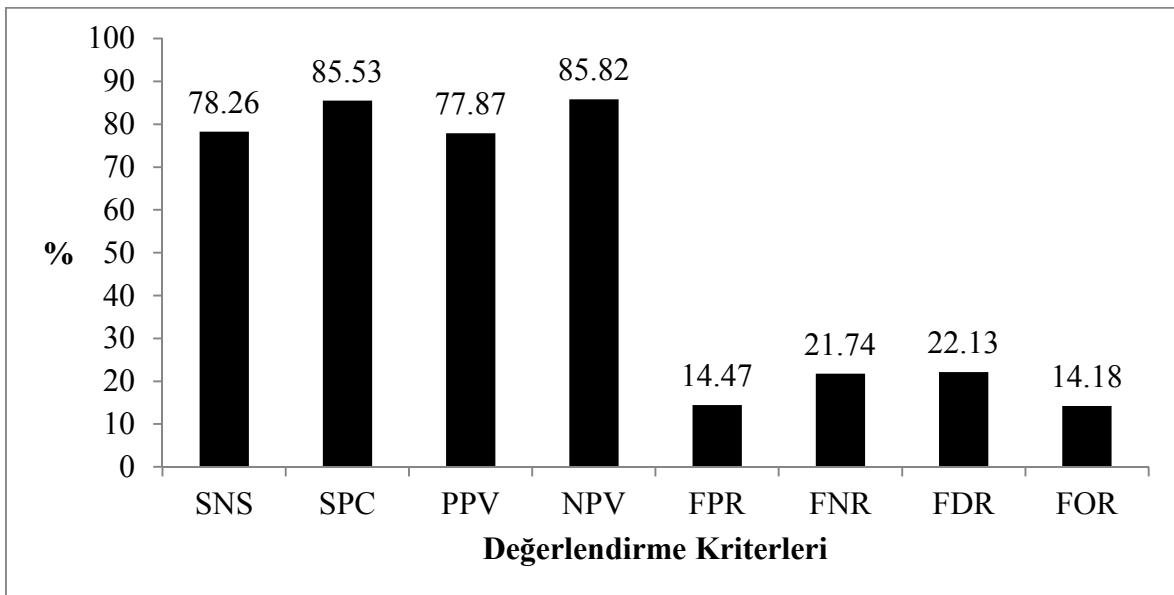
Şekil 9. 24. Yedinci öznitelik grubu kullanılarak gerçekleştirilen meme doku tipi tespiti çalışmasının performans değerlendirmesi

Tez çalışmasının ikinci döneminde, meme doku tipi tespitinin ardından alt bölüm 8.2.4'te ayrıntılı olarak anlatıldığı şekilde ROI tespiti gerçekleştirilmiştir. 52 normal, 40 iyi huylu kanser ve 18 kötü huylu kanser içeren toplamda 110 tane seçilmiş ROI'den alt bölüm 6.5'de anlatılan 11-boyutlu mamografik öznitelikler çıkartılmıştır. Bu öznitelikler, çok büyük değerlerin öznitelik vektör uzayını daraltmasını önlemek amacı ile 1-10 arası değerlere indekslenmiştir. İndekslenmiş öznitelik vektörlerinin FLDA ve LLC sınıflandırıcıları kullanılarak 4- ve 5- katlamalı çapraz doğrulama tekniği ile sınıflandırılması sonucu gerçekleştirilen meme kanseri teşhis çalışmasından elde edilen sonuçlar Şekil 9.25'te verilmiştir.

Çalışmada en yüksek doğruluk oranını (%83,25) sağlayan LLC sınıflandırıcısının performans değerlendirmesi Şekil 9.26'da verilmiştir.



Şekil 9. 25. MIAS veritabanı meme kanseri teşhisi çalışmasından elde edilen sonuçlar



Şekil 9. 26. LLC sınıflandırıcısı ile MIAS veritabanı meme kanseri teşhisi çalışmasının performans değerlendirilmesi

Çalışmada LLC sınıflandırıcısı kullanılarak oluşturulan toplam karmaşıklık matrisi Çizelge 9.17’de verilmiştir. Çizelge incelendiğinde, kötü huylu kanser sınıfının başarılı bir şekilde sınıflandırıldığı gözlenmiştir. Bu durum, sınıflandırma çalışmasının güvenilirliğini ortaya koymaktadır. İyi huylu kanser sınıfına ait yanlış-pozitif ve yanlış-negatif sonuçların ise normal sınıfına ait olduğu görülmektedir. Bu durum, çalışmada kullanılan mamografik özneliklerin kötü huylu sınıfı diğerlerinden ayırt eden öznelikler olduğunu

belirtmektedir. Sonuç olarak, iyi huylu kanser sınıfı ile normal sınıfları arasında ayırt edici özneliklerin bulunması ile sınıflandırma başarısının yükseltilebileceği düşünülmektedir.

Çizelge 9. 17. LLC sınıflandırıcısı ile MIAS veritabanı meme kanseri teşhisi çalışmasına ait karmaşıklık matrisi

		Sınıflandırma Sonucu		
		Normal	İyi Huylu	Kötü Huylu
Referans Sınıflar	Normal	46	9	0
	İyi Huylu	11	25	4
	Kötü Huylu	1	0	19

10. SONUÇ VE ÖNERİLER

Kadın ölüm nedenlerinde ön sıralarda yer alan meme kanserinin erken teşhis ile ölüm oranı azaltılabilmektedir (Jemal vd., 2011). Tez çalışması kapsamında incelenen mamografi, meme kanserinin erken tespit ve teşhisinde kullanılan en yaygın yöntemdir. Buna rağmen, meme dokusundaki sağlıklı bölgeler ile kanserli bölgelerin X-ışını geçirgenliği çok farklı olmadığı için mamografi görüntüleri oldukça az bilgi içermektedir (Ergin ve Kılınç, 2014). Meme doku yoğunluğu arttıkça mamogramlardan elde edilen bilgi daha da azalmakta ve bu durum radyoloji uzmanları açısından tespit ve teşhis sürecini zorlaştırmaktadır. Mamogramların analiz süresini kısaltmak, radyologların mamogramları analizine katkı sağlamak ve insan hatasının teşhisteki rolünü azaltmak maksatlarıyla CAD sistemlerin geliştirilmesi önem arz etmektedir.

Bu tez çalışmasında, mamografi imgelerinden meme doku tespiti, meme kanseri tespit ve teşhisi yapabilen bir CAD sistemi önerilmiştir. Önerilen CAD sistemi literatürde kabul görmüş IRMA ve MIAS veritabanları kullanılarak gerçekleştirilmiştir.

Tez çalışmasının ilk bölümünde, IRMA veritabanı kullanılarak meme kanseri teşhisi yapabilen bir CAD sistemi önerilmiştir. Önerilen CAD sisteminde ilk olarak, imgelere histogram eşitleme ve NLM filtreleme işlemleri uygulanarak kontrastı artırılmış ve gürültü giderimi sağlanmış imgeler elde edilmiştir. Öznitelik çıkarımı aşamasında, ön işlenmiş imgelere Daubechies1 dalgacığı ile birinci seviyede iki boyutlu dalgacık dönüşümü uygulanmıştır. Dönüşüm sonucu elde edilen her bir alt banda uygulanan LCP dönüşümü ile imgelere ait 81×1 -boyutlu vektörler elde edilmiştir. Elde edilen vektörlerin 81. boyutundaki değerler diğer boyutlardaki değerlerden çok büyük olduğu için 81. boyutlar atılmış ve her bir alt banttan elde edilen vektörler $\{1-1-1-0\}$, $\{1,4-1-1-0\}$ ve $\{1,8-1-1-0\}$ katsayıları ile ağırlıklandırılarak toplanmıştır. Bu öznitelik vektörlerine 12 tane istatistiksel ve 16 tane frekans-düzlemi öznitelik eklenerek 108-boyutlu LCP-tabanlı öznitelik vektörleri elde edilmiştir. Meme kanseri teşhisi, 108-boyutlu LCP-tabanlı öznitelik vektörlerinin FLDA, LDC, LLC, SVM, k -NN ($k=5$), Naïve Bayes, karar ağacı ve rastgele orman sınıflandırıcıları kullanılarak sınıflandırılması ile gerçekleştirilmiştir.

Teşhis çalışması öncelikle veritabanının yalnızca yağlı doku tipine sahip imgeleri üzerinde 3-sınıflı olarak gerçekleştirilmiştir. Bu çalışmada $\{1,4-1-1-0\}$ katsayıları kullanılarak elde edilen 108-boyutlu LCP-tabanlı öznelik vektörlerin %92,23 doğruluk oranı ile en iyi performansı sağladığı tespit edilmiştir. Çalışmada, ayrıca, 12-boyutlu istatistiksel, 16-boyutlu frekans-düzlemi, 80-boyutlu LCP öznelikleri ve bu üç öznelik grubundan elde edilen 108-boyutlu LCP-tabanlı öznelik vektörlerinin performansları karşılaştırılmış, ilk üç öznelik grubunun ayırt ediciliğinin beraber kullanıldıklarında daha da arttığı gözlemlenmiştir.

İstatistiksel, frekans-düzlemi ve LCP öznelik vektörleri literatürde mevcut iken, bu üç vektörün birleştirilmesi ile elde edilen ve veriyi ayırt ediciliği daha yüksek olan 108-boyutlu LCP-tabanlı öznelik vektörlerinin oluşturulması çalışmanın özgünlüğünü göstermektedir.

IRMA veritabanı kullanılarak meme kanseri teşhis çalışması, veritabanındaki tüm imgeler ele alınarak aynı öznelik vektörleri ile 12-sınıflı olarak da gerçekleştirilmiştir. 12-sınıflı sınıflandırma çalışması için 1- , 2- ve 3-aşamalı sınıflandırma süreçleri önerilmiştir. 1-aşamalı sınıflandırma sürecinde, mamografi imgeleri meme doku tipi dikkate alınmaksızın direk olarak sağlık durumlarına göre sınıflandırılmıştır. 2-aşamalı süreçte ilk olarak meme doku tipi sınıflandırması gerçekleştirilmiş, sonrasında her bir doku tipi kendi içinde sağlık durumlarına göre sınıflandırılmıştır. 3-aşamalı sınıflandırma sürecinde ise meme doku tiplerine göre sınıflandırılan imgeler öncelikle *sağlıklı – kanserli* olarak sınıflandırılmış ve üçüncü aşamada, kanserli bulunan imgelerin iyi huylu ya da kötü huylu kanserli sınıflandırması gerçekleştirilmiştir. 1-aşamalı sınıflandırma sürecinde en yüksek %87,51 doğruluk oranı elde edilirken, bu oran, 2-aşamalı süreç izlenince %88,79'a ve 3-aşamalı süreç ile de %93,52'ye yükselmiştir. Ayrıca, tüm süreçlerden elde edilen sonuçlar incelendiğinde, 3-aşamalı sınıflandırma sürecinin standart sapmasının diğer süreçlere göre daha düşük çıktığı gözlemlenmiştir. Bu durum, 3-aşamalı sınıflandırma sürecinin katlamalarda kullanılan imgelerden bağımsız çalıştığını göstermekte ve güvenilirliğini vurgulamaktadır.

Tez çalışmasının ikinci bölümünde ise MIAS veritabanı kullanılarak meme doku tipi tespiti, meme kanseri tespit ve kanseri teşhisi yapabilen bir CAD sistemi önerilmiştir.

Önerilen CAD sisteminde ilk olarak, imgelere medyan filtreleme uygulanarak gürültü giderimi, morfolojik işlemler sonucu yapaylık giderimi sağlanmış ve arka plandan arındırılmış mamografi imgeleri pektoral kas bölgesi sol üst köşede kalacak şekilde düzenlenmiştir. Sonrasında, pektoral kas tespiti için, ön işlenmiş mamografi imgelerine histogram eşitleme ardından iki farklı algoritma kullanılarak bölge büyütme yöntemi uygulanmış ve elde edilen bölgelere doğru oturtma gerçekleştirilerek başlangıç bölütleri elde edilmiştir. İmge gradyan bilgisinin de kullanıldığı iyileştirme işlemleri sonucunda, mamografi imgelerinde pektoral kas tespiti gerçekleştirilmiştir. Çalışmadan elde edilen sonuçlar literatürde yer alan çalışmalar ile karşılaştırıldığında, önerilen yöntem ile yanlış negatiflik oranının daha düşük olduğu, yanlış pozitiflik oranının daha yüksek çıkmasına rağmen kullanılan imge sayısı düşünüldüğünde daha başarılı sonuçlar elde edildiği gözlemlenmiştir.

Bölge büyütme ve doğru oturtma yöntemleri literatürde pektoral kas tespiti için kullanılsa da bu yöntemlerin birlikte kullanılması çalışmanın özgünlüğünü göstermektedir.

Pektoral kas tespitinin ardından gerçekleştirilen meme doku tipi tespiti çalışması için GLCM matrislerinden elde edilen doku öznelikleri, FLDA ve SVM sınıflandırıcıları ile 1- ve 2-aşamalı sınıflandırma süreçleri izlenerek sınıflandırılmıştır. 1-aşamalı meme doku tipi sınıflandırma sürecinde mamografi imgeleri doku tiplerine göre *yağlı*, *glandüler* ve *yoğun* olmak üzere üç sınıfa ayrılmıştır. 2-aşamalı meme doku tipi tespiti ise “*önce yağlı – sonra diğerleri*”, “*önce glandüler – sonra diğerleri*” ve “*önce yoğun – sonra diğerleri*” olmak üzere üç farklı süreç izlenerek gerçekleştirilmiştir. Çalışmada en yüksek başarı 108-boyutlu LCP-tabanlı öznelik vektörünün (Bkz. Çizelge 8.3) MIAS veritabanı için oluşturulması ve sırasıyla $\theta = 0^{\circ}, 45^{\circ}, 90^{\circ}$ ve 135° yönleri, $d = \{1, 2, 3, 4\}$ uzaklık parametreleri kullanılarak oluşturulan dört GLCM matrisinden çıkartılan 22-boyutlu doku özneliklerin ortalamalarının alınması ile hesaplanan 22-boyutlu ortalama vektörlerinin alt alta dizilmesi ile elde edilen 88-boyutlu özneliklerin birlikte kullanıldığı öznelik grubunun, FLDA sınıflandırıcısı ile “*Önce Glandüler – Sonra Diğerleri*” sınıflandırma sürecinde sınıflandırıldığında sağlanmıştır.

Meme doku tipi tespiti ardından Chan-Vese aktif kontur modelleme yöntemi kullanılarak ROI tespiti gerçekleştirilmiştir. Meme kanseri teşhisi için, ROI'lerden 11-

boyutlu mamografik öznitelikler çıkartılmış ve 1-10 arası değerlere indekslenmiştir. İndekslenen özniteliklerin FLDA ve LLC sınıflandırılması sonucu gerçekleştirilen meme kanseri teşhisi çalışmasında %83,25 doğruluk oranı elde edilmiştir. Çalışmada, kötü huylu kanser sınıfının başarılı bir şekilde sınıflandırıldığı gözlenmiştir. Bu durum, sınıflandırma çalışmasının güvenilirliğini ortaya koymaktadır. İyi huylu kanser sınıfının ise normal sınıftan tam olarak ayrılamadığı tespit edilmiştir. Bu durum, çalışmada kullanılan mamografik özniteliklerin kötü huylu sınıfı diğerlerinden ayırt eden öznitelikler olduğunu belirtmektedir.

Tez çalışması değerlendirildiğinde, ilerleyen dönemlerde, iyi huylu kanser ile normal sınıfları arasında ayırt edici mamografik özniteliklerin bulunması ile sınıflandırma başarısının yükseltilebileceği ve meme doku tipi tespitinin ardından her bir doku tipine ayrı ROI tespiti ve öznitelik çıkarımı uygulandığında daha başarılı sonuçlar elde edilebileceği düşünülmektedir.

Gerçekleştirilen tez çalışması ile özgün bir öznitelik çıkartım ve pektoral kas tespiti yöntemi önerilerek literatüre katkıda bulunulmuştur. Ayrıca, tez çalışmasından elde edilen sonuçların, önerilen sistemin radyologlar tarafından etkin bir CAD sistemi olarak kullanılabilirlikte olduğunu gösterdiği düşünülmektedir.

KAYNAKLAR DİZİNİ

- Abdellatif, H., Taha, T. E., Zahran, O. F., Al-Nauimy, W., Abd El-Samie, F. E., 2012, Automatic pectoral muscle boundary detection in mammograms using eigenvectors segmentation, 29thNational Radio Science Conference (NRSC), 633-640.
- Agrawal, P., Vatsa, M., Singh, R., 2014, Saliency based mass detection from screening mammograms, *Signal Processing*, 99, 29-47.
- 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.
- Anitha, J., Peter, J. D., Pandian, I. A. S., 2017, A dual stage adaptive thresholding (DuSAT) for automatic mass detection in mammograms, *Computer Methods and Programs in Biomedicine*, 138, 93–104.
- Beura, S., Majhi, B., Dash, R., 2015, Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer, *Neurocomputing*, 154, 1–14.
- Biswas, S. K., Mukherjee, D. P., 2011, Recognizing architectural distortion in mammogram: A multiscale texture modeling approach with GMM, *IEEE Transactions on Biomedical Engineering*, 58,7, 2023-2030.
- Breiman, L., 2001, Random Forests, *Machine Learning*, 45, 5-32.
- Bria, A., Karssemeijer, N., Tortorella, F., 2014, Learning from unbalanced data: A cascade-based approach for detecting clustered microcalcifications, *Medical Image Analysis*, 18, 241-252.
- Camilus, K. S., Govindan, V. K., Sathidevi, P. S., 2010, Computer-aided identification of the pectoral muscle in digitized mammograms, *Journal of Digital Imaging*, 23,5, 562-580.
- Camilus, K. S., Govindan, V. K., Sathidevi, P. S., 2011, Pectoral muscle identification in mammograms, *Journal of Applied Clinical Medical Physics*, 12,3, 215-230.
- Casti, P., Mencattini, A., Salmeri, M., Ancona, A., Mangieri, F., Rangayyan, R. M., 2013, Measures of radial correlation and trend for classification of breast masses in mammograms, 35th Annual International Conference of Engineering in Medicine and Biology Society (EMBC), 6490-6493.
- Casti, P., Mencattini, A., Salmeri, M., Ancona, A., Mangeri, F., Pepe, M.L., Rangayyand, R.M., 2016, Contour-independent detection and classification of mammographic lesions”, *Biomedical Signal Processing and Control*, 25, 165–177.

KAYNAKLAR DİZİNİ (devam)

- Chakraborty, J., Mukhopadhyay, S., 2012, Automatic detection of pectoral muscle using average gradient and shape based feature, *Journal of Digital Imaging*, 25, 387-399.
- Chapelle O., Haffner P., Vapnik V. N., 1999, Support vector machines for histogram-based image classification, *IEEE Transaction on Neural Networks*, 10,5, 1055-1064.
- Chen, J., Huang, H., Tian, S., Qu, Y., 2009, Feature selection for text classification with Naive Bayes, *Expert Systems with Applications*, 36, 5432-5435.
- Choi, J. Y., Kim, D. H., Plataniotis, K. N., Ro, Y. M., 2016, Classifier ensemble generation and selection with multiple feature representations for classification applications in computer-aided detection and diagnosis on mammography, *Expert Systems With Applications*, 46, 106–121.
- Chockri, F., Farida, M. H., 2017, Mammographic mass classification according to Bi-RADS lexicon, 11, 3, 189-198.
- 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.
- David, R., Arnau, O., Joan, M., Marta, P., Joan, E., 2005, Breast segmentation with pectoral muscle suppression on digital mammograms, *Lecture Notes in Computer Science*, 153-158.
- Deserno, T. M., de Oliveira, J. E. E., Araujo, A. A., 2011, Towards computer-aided diagnostics of screening mammography using content-based image retrieval, *Proc IEEE Sibgrapi*, 211-219.
- Dik, E. C., Aktaş, A. Z., 2014, Hausdorff uzaklığı kullanılarak hitit çiviyazısı işaretlerinin tanınması, *TBD 31. Ulusal Bilişim Kurultayı*, 1-5.
- Domingues, I., J. Cardoso, S., Amaral, I., Moreira, I., Passarinho, P., Comba, J. S., Correia, R., Cardoso, M. J., 2010, Pectoral muscle detection in mammograms based on the shortest path with endpoints learnt by SVMs, *Engineering in Medicine and Biology Society*, 3158-3161.
- Eltoukhy, M. M., Faye, I., Samir, B. B., 2012, A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multi resolution representation, *Computers in Biology and Medicine*, 42, 123–128.
- Ergin, S., Kılınç, O., 2014, A new feature extraction framework based on wavelets for breast cancer diagnosis, *Computers in Biology and Medicine*, 51, 171-182.

KAYNAKLAR DİZİNİ (devam)

- Ferrari, R. J., Rangayyan, R. M., Desautels, J. E. L., Borges, R. A., Frere, A. F., 2004, Automatic identification of the pectoral muscle in mammograms, *IEEE Transactions on Medical Imaging*, 23,2, 232-245.
- Fisher, R.A., 1936, The use of multiple measurements in taxonomic problems, *Annals of Eugenics*, 7, 179-188.
- Fix, E., Hodges, J., 1952, Discriminatory analysis, nonparametric discrimination: Consistency properties, Report 4 on the Project 21-49-004, US Air Force School of Aviation Medicine, Randolph Field.
- Ganesan, K., Acharya, U. R., Chua, K. C., Min, L. C., Abraham, K. T., 2013, Pectoral muscle segmentation: A review, *Computer Methods and Programs in Biomedicine*, 110, 48-57.
- Ganesan, K., Acharya, U. R., Chua, C. K., Lim, C. M., Abraham, K. T., 2014, One-class classification of mammograms using trace transform functions, *IEEE Transactions on Instrumentation and Measurement*, 63,2, 304-311.
- Guliatto, D., Rangayyan, R. M., Carvalho, J. D., Santiago, S. A., 2008, Polygonal modelling of contours of breast tumors with the preservation of spicules, *IEEE Transactions on Biomedical Engineering*, 55,1, 14-20.
- Guo, Y., Dong, M., Z., Xiaoli, G., Wang, K., Luo, C., Ma, Y., Zhang, J., 2016, A new method of detecting micro-calcification clusters in mammograms using contourlet transform and non-linking simplified PCNN, *Computer methods and programs in biomedicine*, 130, 31-45.
- Hachama, M., Desolneux, A., Richard, F. J. P., 2012, Bayesian technique for image classifying registration, *IEEE Transactions on Image Processing*, 21,9, 4080-4091.
- Haider, W., Sharif, M., Raza, M., 2011, Achieving accuracy in early stage tumor identification systems based on image segmentation and 3d structure analysis, *Computer Engineering and Intelligent Systems*, 2, 6, 96-103.
- Haralick, R. M., Shanmugam, K., Dinstein, I., 1973, Textural features of image classification, *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3, 6, 1973.
- Işıklı Esener, İ., Ergin, S., Yüksel, T., 2015, A new ensemble of features for breast cancer diagnosis, *IEEE 38th International Conference on Information and Communication Technology, Electronics and Microelectronics (MIPRO'15)*, 1415-1420.

KAYNAKLAR DİZİNİ (devam)

- Işıklı Esener, İ., Ergin, S., Yüksel, 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.
- Işıklı Esener, İ., Ergin, S., Yüksel, T., 2017,, A new feature ensemble with a multi-stage classification scheme for breast cancer diagnosis, *Journal of Healthcare Engineering*, 2017, 1-15.
- Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., Forman, D., 2011, Global cancer statistics”, *CA: A Cancer Journal for Clinicians*, 61, 2, 69-90.
- Karahaliou, A. N., Boniatis, I. S., Skiadopoulos, S. G., Sakellaropoulos, F. N., Arikidis, N. S., Likaki, E. A., Panayiotakis, G. S., Costaridou, L. I., 2008, Breast cancer diagnosis: Analyzing texture of tissue surrounding microcalcificatons, *IEEE Transactions on Information Technology in Biomedicine*, 12,1, 731-738.
- Karssemeijer, N., 1998, Automated classification of parenchymal patterns in mammograms, *Physics in Medicine and Biology*, 43,2, 365-378.
- Kavzoğlu, T., Çölkesen, İ., 2010, Karar Ağaçları ile Uydu Görüntülerinin Sınıflandırılması: Kocaeli Örneği, *Harita Teknolojileri Elektronik Dergisi*, 2,1, 36-45.
- Keleş, A., Keleş, A., Yavuz, U., Expert system based on neuro-fuzzy rules for diagnosis breast cancer, 2011, *Expert Systems with Applications*, 38, 5719–5726.
- Kekre, H. B., Sarode, T., Gharge, S., 2009, Tumor detection in mammography images using vector quantization technique, *International Journal of Intelligent Information Technology Application*, 2, 5, 237-242.
- Kekre, H. B., Sarode, T., Gharge, S., Raut, K., 2010, Detection of cancer using vector quantization for segmentation, *International Journal of Computer Applications*, 4,9, 14-19.
- Krishnan, M. M. R., Banerjee, S., Chakraborty, C., Chakraborty, C., Ray, A. K., 2010, Statistical analysis of mammographic features and its classification using support vector machine, *Expert Systems with Applications*, 37, 470–478.
- Kwok, S. M., Chandrasekhar, R., Attikiouzel, Y., 2001, Automatic pectoral muscle segmentation on mammograms by straight line estimation and cliff detection, 7. *Australian and New Zealand Intelligent Information Systems Conference*, 67-72.
- Kwok, S. M., Chandrasekhar, R., Attikiouzel, Y., Rickard, M. T., 2004, Automatic pectoral muscle segmentation on mediolateral oblique view mammograms, *IEEE Transactions on Medical Imaging*, 23,9, 1129-1140.

KAYNAKLAR DİZİNİ (devam)

- Lado, M. J., Cadarso-Suárez, C., Roca-Pardiñas, J., Tahoces, P. G., 2008, Categorical variables, interactions and generalized additive models. Applications in computer-aided diagnosis systems, *Computers in Biology and Medicine*, 38, 475-483.
- Li, J. B., 2012, Mammographic image based breast tissue classification with kernel self-optimized fisher discriminant for breast cancer diagnosis, *Journal of Medical Systems*, 36, 2235-2244.
- Li, J. B., Wang, Y.-H., Chu, S.-C., Roddick, J. F., 2014, Kernel self-optimization learning for kernel-based feature extraction and recognition, *Information Sciences*, 257, 70-80.
- Li, X.-Z., Williams, S., Bottema, M. J., 2013, Background intensity independent texture features for assessing breast cancer risk in screening mammograms, *Pattern Recognition Letters*, 34, 1053-1062.
- Li, X.-Z., Williams, S., Bottema, M. J., 2014, Texture and region dependent breast cancer risk assessment from screening mammograms, *Pattern Recognition Letters*, 36, 117-124.
- Liu, C. C., Tsai, C. Y., Liu, J., Yu, C. Y., Yu, S. S., 2012, A pectoral muscle segmentation algorithm for digital mammograms using Otsu thresholding and multiple regression analysis, *Computers & Mathematics with Applications*, 64,5, 1100-1107.
- Liu, L., Wang, J., He, K., 2010, Breast density classification using histogram moments of multiple resolution histograms, 3. *International Conference on Biomedical Engineering and Informatics (BMEI)*, 146-149.
- Ma, W. Y., Manjunath, B. S., 2000, Edge flow: A technique for boundary detection and image segmentation, *IEEE Transactions on Image Processing*, 9, 1375-1388.
- Maitra, I. K., Nag, S., Bandyopadhyay, S. K., 2011, Identification of abnormal masses in digital mammography images, *International Journal of Computer Graphics*, 2,1, 17-30.
- Maitra, I K., Nag, S., Bandyopadhyay, S. K., 2012, Technique for preprocessing digital mammogram, *Computer Methods and Programs in Biomedicine*, 107, 175-188.
- Malar, E., Kandaswamy, A., Chakravarthy, D., Dharan, A. G., 2012, A novel approach for detection and classification of mammographic microcalcifications using wavelet analysis and extreme learning machine, *Computers in Biology and Medicine*, 42, 898-905.

KAYNAKLAR DİZİNİ (devam)

- Menon D. P. S., Phadke, A. C., 2013, Classification of circular and lobulated masses in mammograms using fractal analysis, International Conference on Circuits, Power and Computing Technologies (ICCPCT-2013), 1219-1221.
- Muramatsu, C., Hara, T., Endo, T., Fujita, H., 2016, Breast mass classification on mammograms using radial localternary patterns, Computers in Biology and Medicine 72, 43–53.
- Mustra, M., Bozek, J., Grgic, M., 2009, Breast border extraction and pectoral muscle detection using wavelet decomposition, IEEE International Conference on Computer as a Tool (EUROCON), 1426-1433.
- Mustra, M., Grgic, M., 2013, Robust automatic breast and pectoral muscle segmentation from scanned mammograms, Signal Processing, 93, 2817-2827.
- Moayedi, F., Azimifar, Z., Boostani, R., Katebi, S., 2010, Contourlet-based mammography mass classification using the SVM family, Computers in Biology and Medicine, 40, 373-383.
- Nagi, J., Kareem, S. A., Nagi, F., Ahmed, S. K., 2010, Automated breast profile segmentation for ROI detection using digital mammograms”, IEEE EMBS Conference on Biomedical Engineering & Sciences (IECBES), 87-92.
- Otsu, N., 1979, A threshold selection method from gray-level histograms, IEEE Transactions on Systems, Man, and Cybernetics, 9, 1, 62-66.
- 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.
- Pal, N. R., Bhowmick, B., Patel, S. K., Pal, S., Das, J., 2008, A multi-stage neural network aided system for detection of microcalcifications in digitized mammograms, Neurocomputing, 71, 2625-2634.
- Papadopoulos, A., Fotiadis, D.I., Costaridou, L., 2008, Improvement of microcalcification cluster detection in mammography utilizing image enhancement techniques, Computers in Biology and Medicine, 38, 1045-1055.
- Pereira, D.C., Ramos, R.P., Nascimento, M.Z., 2014, Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm, Computer Methods and Programs in Biomedicine, 114,1, 88-101.
- Petroudi, S., Constantinou, I., Tziakouri, C., Pattichis, M., Pattichis, C., 2013, Investigation of AM-FM methods for mammographic breast density classification, 13th IEEE International Conference on Bioinformatics and Bioengineering (BIBE-2013).

KAYNAKLAR DİZİNİ (devam)

- Radovic, M., Djokovic, M., Peulic, A., Filipovic, N., 2013, Application of data mining algorithms for mammogram classification, 13th IEEE International Conference on Bioinformatics and Bioengineering.
- Ramos, R. P., Nascimento, M. Z., Pereira, D. C., 2012, Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms, *Expert Systems with Applications*, 39, 11036-11047.
- Rangayyan, R. M., El-Faramawy, N. M., Desautels, J. E. L., Alim, O. A., 1997, Measures of acutance and shape for classification of breast tumors, *IEEE Transactions on Medical Imaging*, 16, 6, 799-810.
- Roshan, D. Y., Harada, K., 2007, A connected component labeling algorithm for grayscale images and application of the algorithm on mammograms, *ACM Symposium on Applied Computing*, 146-152.
- Safavian, S.R., Landgrebe, D., 1991, A survey of decision tree classifier methodology, *IEEE Transactions on Systems, Man, and Cybernetics*, 21, 660-674.
- Saltanat, M., Hossain, M. A., Alam, M. S., 2010, An efficient pixel value based mapping scheme to delineate pectoral muscle from mammograms, 5. *IEEE International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA)*, 1510-1517.
- Savitha, R., Suresh, S., Sundararajan, N., 2013, Projection-based fast learning fully complex-valued relaxation neural network”, *IEEE Transactions on Neural Networks and Learning Systems*, 24,4, 529-541.
- 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, 1999.
- Subashini, T. S., Ramalingam, V., Palanivel, S., 2010, Pectoral muscle removal and detection of masses in digital mammogram using CCL, *International Journal of Computer Applications*, 1,6, 71-76.
- Suckling, J. et al., 1994, The mammographic image analysis society digital mammogram database, *Exerpta Medica International Congress Series*, 1069, 375-378.
- Sultana, A., Ciuc, M., Strungaru, R., 2010, Detection of pectoral muscle in mammograms using mean-shift segmentation approach, 8. *International Conference on Communications (COMM)*, 165-168.

KAYNAKLAR DİZİNİ (devam)

- Swiderski, B., Osowski, S., Kureka, J., Kruka, M., Lugowska, I., Rutkowski, P., Barhoumi, W., 2017, Novel methods of image description and ensemble of classifiers in application to mammogram analysis, *Expert Systems With Applications*, 81, 67–78.
- Talha, M., Sulong, G. B., 2012, Preprocessing and pectoral muscle separation from breast mammograms, *International Journal of the Physical Sciences*, 7,3, 471-477.
- Thangavel, K., Karnan, M., 2005, Computer aided diagnosis in digital mammograms: Detection of microcalcifications by meta heuristic algorithms, *International Journal on Artificial Intelligence and Machine Learning*, 5, 29-40.
- Tzikopoulos, S. D., Mavroforakis, M. E., Georgiou, H. V., Dimitropoulos, N., Theodoridis, S., 2011, A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry, *Computer Methods and Programs in Biomedicine*, 102, 47-63.
- Vaidehi, K., Subashini, T. S., 2013, Automatic identification and elimination of pectoral muscle in digital mammograms, *International Journal of Computer Applications*, 75,14, 15-18.
- Vallez, N., Bueno, G., Deniz, O., Dorado, J., Seoane, J. A., Pazos, A., Pastor, C., 2013, Breast density classification to reduce false positives in CADE systems, *Computer Methods and Programs in Biomedicine*, 113,2, 569-584.
- Verma, B., McLeod, P., Klevansky, A., 2009, A novel soft cluster neural network for the classification of suspicious areas in digital mammograms, *Pattern Recognition*, 42, 1845—1852.
- a, Webb, A.R., 2002, Linear discriminant analysis. in: *Statistical Pattern Recognition*, John Wiley & Sons, New York, p. 158-159.
- b, Webb, A.R., 2002, Linear discriminant analysis. in: *Statistical Pattern Recognition*, John Wiley & Sons, New York, p. 123-124.
- Weidong, X., Lihua, L., Wei, L. A., 2007, Novel pectoral muscle segmentation algorithm based on polyline fitting and elastic thread approaching, 1. *International Conference on Bioinformatics and Biomedical Engineering (ICBBE)*, 837-840.
- Yam, M., Brady, M., Highnam, R., Behrenbruch, C., English, R., Kita, Y., 2001, Three-dimensional reconstruction of microcalcification clusters from two mammographic views, *IEEE Transactions on Medical Imaging*, 20,6, 479-489.
- Zhang, Z., Lu, J., Yip, J., 2008, Computer aided mammography, *School of Computing and Engineering Researchers' Conference*, University of Huddersfield, 125-130.

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı: İdil IŞIKLI ESENER

Doğum Yeri / Tarihi: Ankara / 01.11.1986

Vatandaşlığı: Türkiye Cumhuriyeti

İletişim

E-posta Adresi: idil.isikli (at) gmail.com

Öğrenim Durumu

Lisans: Elektrik – Elektronik Mühendisliği Bölümü, Eskişehir Osmangazi Üniversitesi, Eskişehir, 2010

Yüksek Lisans: Elektrik – Elektronik Mühendisliği Anabilim Dalı, Bilecik Şeyh Edebali Üniversitesi, Bilecik, 2012

Tez Başlığı: Akıllı sistemler kullanılarak güç sistemlerinde yük tahmini analizi ve uygulaması

Tez Danışmanı: Prof. Dr. Mehmet KURBAN, Yrd. Doç. Dr. Tolga YÜKSEL

Doktora: Telekomünikasyon / Sinyal İşleme Bilim Dalı, Elektrik – Elektronik Mühendisliği Anabilim Dalı, Eskişehir Osmangazi Üniversitesi, Eskişehir, devam ediyor

Tez Başlığı: Göğüs Kanseri İçin Şüpheli Bölgelerin Mamogram İmgeleri Üzerinde Belirlenmesi ve Göğüs Kanseri Türünün Sınıflandırılması

Tez Danışmanı: Doç. Dr. Semih ERGİN, Yrd. Doç. Dr. Tolga YÜKSEL

SCI / SCI-E Dergi Yayınları

1. Işıklı Esener İ., T. Yüksel ve M. Kurban, Short-Term Load Forecasting Without Meteorological Data Using AI Based Structures, Turkish Journal of Electrical Engineering and Computer Sciences, 23: 370-380, 2015.
2. Özkan K., S. Ergin, Ş. Işık ve İ. Işıklı Esener, A New Classification Scheme of Plastic Wastes Based Upon Recycling Labels, Waste Management, 35: 29-35, 2015.

ÖZGEÇMİŞ (devam)

3. Işıklı Esener İ., S. Ergin ve T. Yüksel, A New Feature Ensemble with a Multi-Stage Classification Scheme for Breast Cancer Diagnosis, Journal of Healthcare Engineering, vol. 2017, Article ID 3895164, 15 pages, 2017. doi: 10.1155/2017/3895164.

Diğer Dergi Yayınları

1. Işıklı Esener İ., S. Ergin ve T. Yüksel, A Genuine GLCM-based Feature Extraction for Breast Tissue Classification on Mammograms, International Journal of Intelligent Systems and Applications in Engineering (IJISAE), 4: 124-129, 2016.

Uluslararası Konferans Yayınları

1. Işıklı İ., E. Açıkcalp, H. Yamık ve M. Kurban, Biyodizelin Dizel Santrallerde Kullanım Analizi, The 6th International Advanced Technologies Symposium (IATS'11), s. 20-24, Elazığ, Türkiye, 2011.
2. Bozdemir M., İ. Işıklı ve M. Kurban, Hidromobillerde Kullanılan Motorlar ve Sürücü Devreleri, The 6th International Advanced Technologies Symposium (IATS'11), s. 282-286, Elazığ, Türkiye, 2011.
3. Işıklı Esener İ., T. Yüksel ve M. Kurban, Artificial Intelligence based Hybrid Structures for Short-Term Load Forecasting without Temperature Data, IEEE 11th International Conference on Machine Learning and Applications (ICMLA'12), s. 457-462, Boca Raton, Florida, ABD, 2012.
4. Işıklı Esener İ., T. Yüksel ve M. Kurban, Short-Term Load Forecasting without Weather Data: A Comparison Study, The 12th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS'12), İstanbul, Türkiye, 2012.
5. Işıklı Esener İ., T. Yüksel ve M. Kurban, Wavelet Transform in Power System Analysis, The 2th International Eurasian Conference on Mathematical Sciences and Applications (IECMSA'13), Saraybosna, Bosna-Hersek, 2013.

ÖZGEÇMİŞ (devam)

6. Işıklı Esener İ., S. Ergin ve T. Yüksel, A New Ensemble of Features for Breast Cancer Diagnosis, IEEE 38th International Conference on Information and Communication Technology, Electronics and Microelectronics (MIPRO'15), s. 1415-1420, Opatija, Hırvatistan, 2015.

Ulusal Konferans Yayınları

1. Işıklı Esener İ., T. Yüksel ve M. Kurban, Sıcaklık Verisi Olmadan Kısa Dönem Yük Tahmini için Yapay Zeka Tabanlı Melez Yapılar, 2012 Elektrik – Elektronik ve Bilgisayar Mühendisliği Sempozyumu (ELECO'12), s. 84-88, Bursa, 2012.
2. Balcı H., İ. Işıklı Esener ve M. Kurban, Regresyon Analizi Kullanılarak Kısa Dönem Yük Tahmini, 2012 Elektrik – Elektronik ve Bilgisayar Mühendisliği Sempozyumu (ELECO'12), s. 796-801, Bursa, 2012.
3. Işıklı Esener İ., S. Ergin ve T. Yüksel, Göğüs Kanseri Teşhisinde Bir Öznitelik Seçim Analizi, IEEE Tıp Teknolojileri Ulusal Kongresi (TIPTEKNO'15), s. 412-415, Bodrum, Muğla, 2015.

Projeler

1. “Mamogram Görüntülerinde Meme Kanseri için Şüpheli Bölgelerin İncelenerek Meme Kanserinin Otomatik Teşhisi”, Bilimsel Araştırma Projesi, Proje No. 2015-698, Eskişehir Osmangazi Üniversitesi, Araştırmacı, 2015-2017.