

**BİLGİSAYAR DESTEKLİ TANI SİSTEMİ İLE AKCİĞER
NODÜLLERİNİN NİTELENDİRİLMESİ**

**CHARACTERIZATION OF LUNG NODULES WITH
COMPUTER AIDED DIAGNOSIS SYSTEM**

AYDIN KAYA

YRD. DOÇ. DR. AHMET BURAK CAN

Tez Danışmanı

Hacettepe Üniversitesi
Lisansüstü Eğitim – Öğretim ve Sınav Yönetmeliğinin
Bilgisayar Mühendisliği Anabilim Dalı İçin Öngördüğü
DOKTORA TEZİ
olarak hazırlanmıştır.

2015

AYDIN KAYA'nın hazırladığı “**Bilgisayar Destekli Tanı Sistemi İle Akciğer Nodüllerinin Nitelendirilmesi**” adlı bu çalışma jürimiz tarafından **BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI**'nda **DOKTORA TEZİ** olarak kabul edilmiştir.

Doç. Dr. Erdem Karabulut

Başkan

Yrd. Doç. Dr. Ahmet Burak Can

Danışman

Yrd. Doç. Dr. Nazlı İkizler Cinbiş

Üye

Yrd. Doç. Dr. Erhan Mengüşoğlu

Üye

Yrd. Doç. Dr. Mehmet Demirci

Üye

Bu tez Hacettepe Üniversitesi Fen Bilimleri Enstitüsü tarafından **DOKTORA TEZİ** olarak onaylanmıştır.

Prof. Dr. Fatma Sevin DÜZ
Fen Bilimleri Enstitüsü Müdürü

ETİK

Hacettepe Üniversitesi Fenbilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmasında

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahriyat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

02/07/2015

AYDIN KAYA

ÖZET

Bilgisayar Destekli Tanı Sistemi İle Akciğer Nodüllerinin Nitelendirilmesi

Aydın Kaya

Doktora Bilgisayar Mühendisliği

Tez Danışmanı: Yrd. Doç. Dr. Ahmet Burak Can

Haziran 2015

Akciğer kanseri, özellikle endüstrisi gelişmiş ülkelerde ölüm oranı en yüksek kanser türüdür. Tanısındaki en büyük sıkıntılar, görüntüleme verilerindeki fazlalık ve kansere sebep olan küçük boyutlu nodüllerin tespitinin ve değerlendirmesinin zorluğundan kaynaklanmaktadır. Bilgisayar destekli tanı sistemleri, bu temel sorunlarda radyologlara yardımcı olmak için geliştirilen uzman sistemlerdir.

Bu tez kapsamında küçük boyutlu akciğer nodüllerinin habislik derecesinin belirlenmesi için sınıflama yaklaşımları önerilmektedir. Çalışmada Amerikan Kanser Enstitüsü tarafından sunulan *Lung Image Database Consortium (LIDC)* veritabanı kullanılmıştır. LIDC veritabanı dört farklı enstitüden radyologlar tarafından değerlendirilmiş ve nodül habislik derecesi yanında radyografik tanımlayıcı bilgisini içeren bir veritabanıdır. Bu tez çalışmasında, radyografik tanımlayıcıların habislik derecesinin tahmininde faydalı olup olmadığının incelenmesi amaçlanmıştır. LIDC veritabanının dengesiz sınıf dağılımını ele almak için veri dengeleme yaklaşımları kullanılmıştır. Sınıflama yöntemleri temel olarak iki adımdan oluşmaktadır. İlk adımda alt düzey görüntü özniteliklerinden radyografik tanımlayıcılar tespit edilmekte, sonraki adımda ise radyografik tanımlayıcılar aracılığıyla habislik derecesi tahmin edilmektedir. Sınıflama adımlarında tekil sınıflayıcılar, takım sınıflayıcılar, bulanık mantık tabanlı ve kural tabanlı sınıflama yaklaşımları kullanılmıştır.

Çalışma sonuçları, sınıflama başarımı, duyarlılık ve seçicilik ölçümleri bağlamında literatürdeki önemli çalışmaların ve görüntü öznitelikleri ile eğitilmiş tekil sınıflayıcıların sonuçları ile karşılaştırılmıştır. Elde edilen sonuçlar radyografik tanımlayıcıların sınıflama başarımına olumlu katkı yaptığını göstermektedir. Sunulan yöntemlerin önemli bir bölümü, literatürdeki çalışmalar ile karşılaştırılabilir düzeyde başarılı sonuçlar üretmektedir.

Anahtar Kelimeler: akciğer nodülleri, radyografik tanımlayıcılar, sınıflama, LIDC, veri dengeleme, takım sınıflayıcılar, bulanık yöntemler.

ABSTRACT

Characterization Of Lung Nodules With Computer Aided Diagnosis System

Aydın Kaya

Doctor of Philosophy Department of Computer Engineering

Supervisor: Asst. Prof. Dr. Ahmet Burak Can

June 2015

Lung cancer is one of the leading causes of cancer related deaths worldwide, especially in industrially developed countries. Major problems in diagnosis are caused by high volume of visual data and difficulty of detection of solitary/small pulmonary nodules. Computer aided diagnosis systems are expert systems to assist radiologists on these issues.

In this thesis, classification approaches for predicting malignancy of solitary pulmonary nodules are presented. Publicly available Lung Image Database Consortium (LIDC) database which is proposed by USA National Cancer Institute is used in the study. LIDC contains malignancy and nodule characteristics evaluations of radiologists from four different institutions. The goal of this thesis is to examine the usefulness of radiographic descriptors in malignancy prediction. Dataset balancing approaches are used for addressing unbalanced class distribution of LIDC database. Classification methods basically consist of two phases. In the first phase, radiographic descriptors are determined from low level image features; and in the second step malignancy is predicted from these descriptors. Single classifiers, ensemble classifiers, fuzzy logic based and rule based methods are used in classification steps.

Results are compared with prominent studies in literature and single classifiers which are trained with image features in the context of classification accuracy, specificity and sensitivity measures. The obtained results indicate that radiographic descriptors contribute malignancy prediction. Moreover, the majority of the presented methods' results are successful and comparable with the methods in literature.

Keywords: solitary pulmonary nodules, nodule characteristics, classification, LIDC, data balancing, ensemble classifiers, fuzzy methods.

TEŞEKKÜR

Tez konusunun belirlenmesini sağlayan, tez metninin yazılmasına ve tez çalışmasının hazırlanmasına yardımcı olan ve bu süreçte desteğini eksik etmeyen Sayın Yrd. Doç. Dr. Ahmet Burak Can'a,

Tez izleme aşamasında fikirleriyle tezin gelişmesine ve tez metnini inceleyerek biçim ve içerik bakımından son halini almasına yardımcı olan Sayın Yrd. Doç. Dr. Nazlı İkizler Cinbiş'e ve Sayın Yrd. Doç. Dr. Erhan Mengüşođlu'na,

Tez metnini inceleyerek biçim ve içerik bakımından son halini almasına yardımcı olan Sayın Doç. Dr. Erdem Karabulut'a ve Sayın Yrd. Doç. Dr. Mehmet Demirci'ye,

Tez sürecinde maddi ve manevi desteklerini esirgemeyen değerli arkadaşlarım Ali Seydi Keçeli ve Hüseyin Temuçin'e,

Çalışma hayatım boyunca verdiği desteklerden dolayı Hacettepe Üniversitesi Bilgisayar Mühendisliği Bölümü çalışanlarına,

Hayatımın her aşamasında hep yanımda olan, sevgi ve desteklerini hiç eksik etmeyen çok sevgili annem ve babama,

ve

Tez sürecinde en büyük destekçim olan ve hayatımın sonraki aşamalarında da en büyük destekçim olacak olan; sevdiğim, meleğim ve nişanlım Özlem Ertan'a,

Canı gönülden teşekkür ederim.

İÇİNDEKİLER

Sayfa

ÖZET	i
ABSTRACT	iii
TEŞEKKÜR	iv
İÇİNDEKİLER	v
ŞEKİLLER	viii
ÇİZELGELER	x
SİMGELER VE KISALTMALAR	xiii
1. GİRİŞ	1
2. İLGİLİ ÇALIŞMALAR	6
2.1. Habislik Tahmini Üzerine Yapılan Çalışmalar	6
2.1.1. Tekil Sınıflayıcı Kullanılan Çalışmalar	6
2.1.2. Takım Sınıflayıcı Kullanılan Çalışmalar	11
2.2. Bilgi Erişim Sistemi Çalışmaları	12
2.3. Radyologlar Arası Anlaşma Üzerine Yapılmış Çalışmalar	13
2.4. Genel Değerlendirme	14
2.4.1. Veri Kümeleri	14
2.4.2. Ölçüm Metrikleri	15
2.4.3. Sınıflama Yaklaşımları	17
3. ARAÇ VE YÖNTEMLER	18
3.1. Görüntü Öznitelikleri	18
3.2. Öznitelik Kümelerinde Boyut İndirgeme	22
3.2.1. İleri ve Geri Adımlı Öznitelik Alt Küme Seçimi	23
3.2.2. Relieff Yöntemi ile Öznitelik Altküme Seçimi	26
3.3. Eksik Değerlerin Tamamlanması	27
3.4. Veri Kümelerinin Dengelenmesi	27
3.5. Sınıflama Yöntemleri	28
3.5.1. C4.5	28
3.5.2. Destek Vektör Makinaları (<i>Support Vector Machines - SVM</i>)	29
3.5.3. Random Forests (RF)	30

3.5.4. Naive Bayes.....	31
3.5.5. Adaboost.....	32
3.5.6. K-En Yakın Komşuluk (<i>K Nearest Neighbours - kNN</i>).....	33
3.5.7. Doğrusal Ayırım Analizi (<i>Linear Discriminant Analysis - LDA</i>).....	33
3.6. Sınıflama Algoritmasını Seçmek için Kullanılacak Yöntemler.....	34
3.6.1 Ortalama Sıralamalar (<i>Average Ranks (AR)</i>).....	34
3.6.2. Başarı Değer Oranları (<i>Success Rate Ratios (SRR)</i>).....	35
3.6.3. Önem Kazanımları (<i>Significant Wins (SW)</i>).....	35
3.7. Değerlendiriciler Arası Uyumun Tespiti.....	36
3.7.1. Krippendorff'un Alpha Katsayısı.....	36
3.8. İki Algoritma Arasındaki Farkın Test Edilmesi.....	37
3.9. Yöntem Başarımlarının Ölçülmesi.....	38
4. LİDC VERİ KÜMESİ VE ÇÖZÜMLEMESİ.....	39
4.1. LİDC Veritabanı.....	39
4.2. LİDC Okuyucu.....	43
4.3. Veri Dağılımı ve Çalışmada Kullanılacak Veri Kümelerinin Oluşturulması.....	43
4.4. Farklı Okuma Oturumları (<i>Reading Session</i>) Üzerinden Sınıflama.....	47
4.5. Kesin Referans ile Sınıflama Sonuçlarının Karşılaştırılması.....	55
5. İKİ KATLI TEKİL SINIFLAYICILAR İLE SINIFLAMA YÖNTEMİ.....	59
5.1. Veri Kümesi Hazırlama ve Önışleme Adımı.....	60
5.2. Sınıflama Aşaması.....	60
5.2. Görüntü Öznitelikleri ile İyileştirilmiş Sınıflama Aşaması.....	62
5.3. Sonuçlar.....	63
5.4. Tartışma.....	65
6. BULANIK YAKLAŞIMLAR İLE SINIFLAMA.....	66
6.1. Bulanık Küme Yaklaşımı ile Sınıflama.....	66
6.1.1. Sınıflama Aşaması.....	67
6.1.2. Sonuçlar.....	73
6.1.3. Tartışma.....	73
6.2. Nöro Bulanık Çıkarım ANFIS ile Sınıflama.....	74
6.2.1. Sınıflama Aşaması.....	75

6.2.2. Sonular	76
6.2.3. Tartışma	76
7. TAKIM SINIFLAYICILAR İLE SINIFLAMA	77
7.1. Takım Sınıflayıcılar Ve Kural Tabanlı Yaklaşım İle Sınıflama Yöntemi.....	77
7.1.1. Takım Sınıflayıcılar	78
7.1.2. Takım Sınıflama Aşaması	81
7.1.3. Sonular	84
7.1.4. Tartışma	87
7.2. Takım Sınıflayıcılar Ve Ağırlıklı Kurallar İle Sınıflama Yöntemi	88
7.2.1. Öznitelik Sayılarının Belirlenmesi	89
7.2.2. Takım Sınıflayıcılar	93
7.2.3. Ağırlıklı Kural Tabanlı Sistem	93
7.2.4. Sonular	95
7.2.5. Tartışma	99
7.3. Ağırlıklı Kural Tabanlı Sistemin Farklı Sınıflayıcılar İle Geliştirilmesi.....	100
7.3.1. Sonular	101
7.3.2. Tartışma	101
8. DENEY ÇIKTILARININ ANALİZİ VE TARTIŞMA.....	103
8.1. Sınıflama Sonularının Değerlendirilmesi	103
8.2. Yöntemler Arası Anlamlılık Değerlendirmesi	105
8.3. Literatürdeki Çalışmaların Karşılaştırılması	108
9. SONULAR.....	109
KAYNAKLAR.....	112
ÖZGEÇMİŞ.....	120

ŞEKİLLER

Sayfa

Şekil 2.1 ROC eğrisi gösterimi [44].	16
Şekil 3.1 (a) İstenen veri kümesi yapısı, (b) istenmeyen veri kümesi yapısı (<i>big p small n problem</i>) [59].	22
Şekil 3.2 Öznitelik alt küme seçimi için genel gösterim [60].	24
Şekil 3.3 Özyineli öznitelik seçim algoritması [61].	25
Şekil 3.4 Yeniden örnek kullanan özyineli öznitelik seçim algoritması [61].	26
Şekil 3.5 <i>Sublety</i> tanımlayıcısı için sınıf histogramı. (Dikey eksen örnek sayısıdır)	28
Şekil 3.6 SVM için iki boyutta ve iki öznitelikle ifade edilmiş, veri ve destek vektör gösterimi.	30
Şekil 3.7 kNN için örnek gösterim.	33
Şekil 4.1 Örnek LIDC tomografi görüntüleri.	40
Şekil 4.2 LIDC örneğinden bir CT kesiti. Aynı nodül üç farklı radyolog ratafindan değerlendirilmiştir.	40
Şekil 4.3 LIDC Okuyucu arayüzü	43
Şekil 4.4 (a-h) Radyografik tanımlayıcılarına ait veri kümelerinin histogramları.	46
Şekil 4.5. Yöntemin genel akış çizeneği.	48
Şekil 4.6 (a) RF algoritması ve tüm öznitelikler ile yapılmış sınıflama başarımlarının tüm okumalar ve tanımlayıcılar için gösterimi. (b) RF algoritması ve boyut indirgeme sonrası yapılmış sınıflama başarımlarının tüm okumalar ve tanımlayıcılar için gösterimi.	53
Şekil 4.7 a) SVM algoritması ve tüm öznitelikler ile yapılmış sınıflama başarımlarının tüm okumalar ve tanımlayıcılar için gösterimi. (b) SVM algoritması ve boyut indirgeme sonrası yapılmış sınıflama başarımlarının tüm okumalar ve tanımlayıcılar için gösterimi.	54
Şekil 4.8 SVM sınıflayıcı için R1-R4 ve kesin referans verikümelerinden elde edilen sınıflama başarımları sonuçları grafiği.	57
Şekil 5.1 Yöntemin genel akış çizeneği.	59
Şekil 5.2 Yöntem için çalıştırma örneği gösterimi.	62
Şekil 5.3 Yöntemin ikinci aşamasına ait gösterimin geliştirilmiş hali.	63
Şekil 5.4 Ek yöntem için çalıştırma örneği gösterimi.	63

Şekil 5.5 Habislik derecesi sınıflarının (c1-c5) farklı kesin referanslardaki dağılımı (1-4 arası radyolog anlaşması)	65
Şekil 6.1 Yöntemin genel akış çizeneği.....	66
Şekil 6.2 Bulanık küme ve ilişkilerin tıbbi teşhiste uygulanması yaklaşımlarının probleme uyarlanmış halinin gösterimi.	70
Şekil 6.3 Yöntem için çalıştırma örneği gösterimi.	72
Şekil 6.4 Yöntemin genel akış çizeneği.....	75
Şekil 7.1 Yöntemin genel akış çizeneği.....	77
Şekil 7.2 Takım sınıflayıcılar ile oylama algoritması.	79
Şekil 7.3 Kural tabanlı sistemin algoritması.....	83
Şekil 7.4 Yöntem için çalıştırma örneği gösterimi.	84
Şekil 7.5 Kural boyunun sınıflama sonucuna etkisini gösterir grafik.	86
Şekil 7.6 Yöntemin genel akış çizeneği.....	88
Şekil 7.7 Bir örnek için takım sınıflayıcı adımlarını gösteren algoritma.	92
Şekil 7.8 Yöntem için çalıştırma örneği gösterimi.	95
Şekil 7.9 Ağırlıklı kural tabanlı sistemin algoritması.	97
Şekil 7.10 Kural sayısının sınıflama başarımına etkisini gösteren grafik.	97

ÇİZELGELER

	<u>Sayfa</u>
Çizelge 3.1 Çıkarılan tüm öznelikleri kısa tanımları.....	20
Çizelge 4.1 Radyografik Tanımlayıcı tanımları ve dereceleri.....	42
Çizelge 4.2 Bir nodül üzerinde dört radyoloğun değerlendirmeleri. Her sütundaki, kalın ile işaretlenmiş örnekler ilgili tanımlayıcının veri kümesine eklenmiştir.	45
Çizelge 4.3 LIDC veritabanında paylaşılan nodüllere ait bazı temel istatistikler.	45
Çizelge 4.4 R1 okuması için Random Forest sınıflama algoritması ile elde edilen sonuçlar.	49
Çizelge 4.5 R1 okuması için SVM sınıflama algoritması ile elde edilen sonuçlar.	49
Çizelge 4.6 R2 okuması için Random Forest sınıflama algoritması ile elde edilen sonuçlar.	50
Çizelge 4.7 R2 okuması için SVM sınıflama algoritması ile elde edilen sonuçlar.	50
Çizelge 4.8 R3 okuması için Random Forest sınıflama algoritması ile elde edilen sonuçlar.	51
Çizelge 4.9 R3 okuması için SVM sınıflama algoritması ile elde edilen sonuçlar.	51
Çizelge 4.10 R4 okuması için Random Forest sınıflama algoritması ile elde edilen sonuçlar.....	52
Çizelge 4.11 R4 okuması için SVM sınıflama algoritması ile elde edilen sonuçlar.	52
Çizelge 4.12 <i>Calcification</i> tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.....	55
Çizelge 4.13 <i>Lobulation</i> tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.....	55
Çizelge 4.14 <i>Margin</i> tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.....	55
Çizelge 4.15 <i>Sphericity</i> tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.....	56
Çizelge 4.16 <i>Spiculation</i> tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.....	56
Çizelge 4.17 <i>Sublety</i> tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.....	56
Çizelge 4.18 <i>Texture</i> tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.....	56

Çizelge 4.19 SVM sınıflayıcı için R1-R4 ve kesin referans verikümelerinden elde edilen sınıflama başarımı sonuçları çizelgesi.....	58
Çizelge 5.1 Quick Hull algoritmasının her radyografik tanımlayıcı için en önemli 3-9 öznitelik için işlem süresi verilmiştir.....	61
Çizelge 5.2 Görüntü özniteliklerine göre ve radyografik tanımlayıcılara göre tekil sınıflayıcı sonuçları. En altta geliştirilen yöntemin sınıflama sonuçları verilmiştir.....	64
Çizelge 6.1 <i>Subtlety</i> tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.....	68
Çizelge 6.2 <i>Calcification</i> tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.....	68
Çizelge 6.3 <i>Sphericity</i> tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.....	68
Çizelge 6.4 <i>Margin</i> tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.....	69
Çizelge 6.5 <i>Lobulation</i> tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.....	69
Çizelge 6.6 <i>Spiculation</i> tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.....	69
Çizelge 6.7 <i>Texture</i> tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.....	70
Çizelge 6.8 Dört farklı üyelik işlevinden (Üİ) elde edilmiş sınıflama başarımı (SB) sonuçları.....	73
Çizelge 6.9 ANFIS tabanlı sınıflayıcı için sınıflama sonuçları.....	76
Çizelge 7.1 Bir örnek için her bir radyografik tanımlayıcıya ait oylama sonuç matrisi.....	81
Çizelge 7.2 Malignancy değeri ve diğer radyografik tanımlayıcıların olduğu örnek kural tablosu.....	81
Çizelge 7.3 Sublety için SVM ile sınıflama sonuçları.....	84
Çizelge 7.4 Sublety için en yüksek oylamayı kazanan sınıflar.....	85
Çizelge 7.5 Yöntemlerin sınıflama başarımı (5 sınıflı).....	86
Çizelge 7.6 Yöntemlerin sınıflama başarımı (3 sınıflı).....	87
Çizelge 7.7 (3-10, 20,30,40,50,60) öznitelik sayıları için AR ve SRR ölçümleri. Her satırdaki ölçümler için en önemli değerler kalın yazı tipi ile işaretlenmiştir.....	90
Çizelge 7.8 Quick Hull algoritmasının her radyografik tanımlayıcı için en önemli 3-9 öznitelik için işlem süresi verilmiştir.....	91

Çizelge 7.9 Nodül habislik derecesi için (<i>malignancy – Mal.</i>) örnek kural tablosu	93
Çizelge 7.10 Habislik kümesinden elde edilmiş, aynı nodule ait değerlendirme olan üç kuralın değerleri ve anlaşma katsayıları.	94
Çizelge 7.11 Yöntemlerin sınıflama başarımı (5 sınıflı veri için).	98
Çizelge 7.12 Yöntemlerin sınıflama başarımı (3 sınıflı veri için).	98
Çizelge 7.13 Kesim 7.2'deki yönteme ek olarak yapılan analiz sonuçları. Sadece örnek artırma bulunmaktadır.	101
Çizelge 7.14 Kesim 7.2'deki yönteme ek olarak yapılan analiz sonuçları. Örnek artırma ve azaltma bulunmaktadır.	101
Çizelge 8.1 Yöntemlerin sınıflama başarımları.	104
Çizelge 8.2 Beş sınıflı veri için, geliştirilen yöntemler ve tekil sınıflayıcılar arası istatistiksel anlamlılık testi sonuçları (p değerleri).	107
Çizelge 8.3 Üç sınıflı veri için, geliştirilen yöntemler ve tekil sınıflayıcılar arası istatistiksel anlamlılık testi sonuçları (p değerleri).	107

SİMGELER VE KISALTMALAR

CT	Computed Tomography
LIDC	Lung Image Database Consortium
CAD	Computer Aided Diagnosis
DICOM	Digital Imaging and Communications in Medicine
LDA	Doğrusal Ayırma Analizi
SVM	Destek Vektör Makinaları
RF	Random Forest
NB	Naive Bayes
kNN	k-En Yakın Komşuluk
AR	Average Ranks
SRR	Success Rate Ratios
SW	Significant Wins
EC	Takım Sınıflayıcı (Ensemble Classifier)
CH	Dışbükey Örtü (Convex Hull)
SB	Sınıflama Başarımı
Duy.	Duyarlılık
Seç.	Seçicilik
AUC	Eğri altındaki alan
EBNK	En Büyük Nodül Kesitinden
TNKO	Tüm Nodül Kesitlerinin Ortalamasından
SRND	Nodül Etrafındaki Alanlardan

1. GİRİŞ

Akciğer kanseri, özellikle endüstrisi gelişmiş ülkelerde hava kirliliğine bağlı olarak sıkça rastlanan ve büyük oranda ölümlere sebep olan bir kanser türüdür. Bu konuda yapılan araştırmalarda açıklanan verilere göre; Japonya’da erkeklerde bir numaralı ölüm sebebi olarak akciğer kanseri gösterilmiştir [1]. 2007 yılında Amerika Birleşik Devletleri’nde, 160.400 ölüm bu kanser türü sebebiyle olmuştur ve bu hastalığın tedavi maliyeti yıllık yaklaşık 10 milyar dolardır [2]. Dünya çapında her yıl 1.3 milyon yeni vakaya rastlanmaktadır [3].

Kansere sebep olan küçük boyutlu akciğer nodülleri (*small/solitary pulmonary nodules - SPN*) genellikle çapı 2-30 milimetre arasında değişen, sınırları çok kolay belirlenemeyen, küresel yapıya sahip ve erken aşamalarda tespit edilmesi bilgisayarlı tomografi (*computed tomography - CT*) veya X-ışını (*chest X-Ray*) taramalarında tesadüfen gerçekleşen nodüllerdir [4]. Yapılan nodül tanımları her ne kadar genelleştirilmiş olsa da, farklı nodül yapılarının tespit edilmesi ve değerlendirilmesi radyologlar için zor olmaktadır.

Tanı konusunda en önemli sorun, hastaların nodüllerinin iyi huylu (*benign*) ya da kötü huylu (*malignant*) olup olmadığının belirlenmesidir. Tomografi taramalarında nodüle rastlanan hastaların tanı aşaması, radyologların nodül şekli aracılığıyla belirledikleri (radyografik tanımlayıcılar ışığında) nodül yapısı olasılıklarının belirlenmesi ile başlar. Yüksek olasılıkla kötü huylu olduğu düşünülen nodüller takipten çıkarılarak tedavi süreci başlatılır. İyi huylu olduğu düşünülen nodüller ise zamanla uygulanan tomografi taramaları ile takibe alınır. Olasılıksal olarak iki yapı arasında tereddütte kalınır ise, biyopsi gibi ek testler uygulanır.

Radyografik tanımlayıcılar, radyologlar tarafından nodül değerlendirmesi için belirlenmiş, nodül şekline ve etrafındaki dokulara dayandırılarak çıkarılan öznel özelliklerdir. Bunlardan bazıları şöyledir [5]:

- kalsiyum yapılanması (*calcification*)
- nodül iç yapısı (*internal structure*)
- yuvarlaklık (*sphericity*)
- sınırların belirginliği (*margin*)
- nodülün farklı kesimlerindeki büyüme oranı (*lobulation*)
- sınırlardaki girinti-çıkıntı yapısı (*spiculation*)

- dokusu (*texture*)

Tıbbi görüntüleme ve tanısal radyolojide, bilgisayar destekli tanı (*Computer Aided Diagnosis- CAD*) ana araştırma konularından biri haline gelmiştir. CAD sistemlerinde bilgisayarın başarımı, doktorların başarımından daha iyi olmak ya da onlarla karşılaştırılmak durumunda değildir. Amaç doktora yardımcı olacak yöntemler geliştirip, tanının daha doğru olarak yapılabilmesini sağlamaktır [2], [6]. Birçok uluslararası firmanın bu konu üzerine geliştirdiği araçlar bulunmaktadır (ImageChecker[7], CADStream[8], Mevis LiverAnalyser[9], vs.).

Günümüzde kullanılan bilgisayarlı tomografi cihazları milimetre altı duyarlılık ve yüksek çözünürlük ile; hastanın nefesini tuttuğu anda bütün göğsün çekimini yapabilen sistemlerdir. Görüntülemedeki duyarlılık ve başarımın artması ile birlikte bu sistemlerin ürettiği veri miktarındaki artış radyologlar için de yük oluşturmaya başlamıştır [10].

Akciğer kanseri vakalarında, küçük boyuttaki akciğer nodüllerinin tespit edilmesi ve değerlendirilmesi radyologlar için zorlu bir süreçtir. Bu süreçte nodülün yerinin tespit edilmesi, ilgisiz verinin veri kümesinden çıkarılması, tespit edilen nodülün yapısı hakkında önerilerde bulunulması, farklı akciğer bileşenlerinin ayrılması gibi işlemler kullanıcılara büyük kolaylıklar sağlamaktadır [11]. Yapılan birçok CAD araştırmasında radyologların başarımı, sistemlerin kullanımı sonrası önemli oranda artmıştır. Hatta CAD sistemlerinin, tecrübesi fazla olan radyologlara yaptıkları katkının daha fazla olduğu gözlemlenmiştir [11].

Genel bir akciğer kanseri üzerine geliştirilmiş CAD sisteminde temel olarak üç bileşen bulunmaktadır [2], [11], [12].

1- Bölütleme:

CAD sisteminin genellikle ilk aşaması olarak bölütleme kesimi düşünülebilir. Akciğer nodüllerinin başarılı bir şekilde ayrılması için bazı anatomik nitelikler dikkate alınarak sırayla bazı bölütleme adımları tasarlanabilir. Eğer bölütleme işlemi el ile (radyoloğun kendisi tarafından) ya da yarı otomatik (yine radyolog tarafından kısmi işaretleme ile) yapılmıyorsa; sistem öncelikle akciğere ait bileşenleri otomatik olarak ayırmak durumundadır. Akciğer sınırlarının bulunması, soluma yollarının ayrıştırılması, nodüle etki edebilecek damarların ayrıştırılması ve nodülün birleşik olarak bulunabileceği akciğer zarının ayrılması bu işlemlerden bazılarıdır. Bu temel adıma ek olarak, sistemin görüntüyü

bölütlemesinin ardından tespit edilen nodül adaylarının elenmesi gibi bir adım da gerekli görülebilir.

2- Öznitelik Çıkarımı:

Doğru bir biçimde bölütlendiğine emin olunan nodülün iyi ya da kötü huylu olduğuna karar vermek için, nodülleri ifade edecek tanımlayıcılara (özniteliklere) ihtiyaç duyulur. Nodülün iki boyutlu ya da üç boyutlu yapısından, çevre bileşenler ile etkileşimini değerlendirerek birçok öznitelik değerlendirmeye alınabilir. Bazı CAD sistemlerinde hastanın yaşı, cinsiyeti, sigara geçmişi, aile geçmişi gibi klinik bilgiler de sınıflama işlemi için öznitelik olarak kullanılabilir [13].

3- Sınıflama:

Nodüller ve öznitelikleri elde edildikten sonra, nodülün iyi ya da kötü huylu olması ya da sebep olabileceği hastalıkların tespiti için çeşitli sınıflama yaklaşımları kullanılabilir. Öznitelik kümeleri anlamlı ya da anlamsız birçok veriyi barındırabileceği için, en anlamlı alt kümenin seçimini yapmak akciğer CAD sistemlerinde sıkça kullanılan bir yaklaşımdır.

Sınıflama adımına ek olarak yanlış pozitif (*false positive*) ya da yanlış negatif (*false negative*) sonuçların elenmesi aşaması da gerekli görülmektedir. Çünkü verilecek bir karar sonrası yapılacak işlemlerin hastaya rahatsızlık verebilecek olmasından dolayı böyle bir aşamaya ihtiyaç vardır.

Tez çalışması kapsamında, akciğer nodüllerinin habislik derecesinin tespiti için bir bilgisayar destekli tanı yöntemi sunulmuştur. Bu yöntemde, yukarıda verilen CAD sistemi parçalarından öznitelik çıkarımı ve sınıflama aşamaları üzerine çalışılmıştır. Bu amaçla, Amerika Birleşik Devletleri'nde birkaç devlet kuruluşunun birlikte çalışması ile oluşturulan Lung Image Database Consortium (LIDC) veritabanı [5] tarafından sağlanan veriler kullanılmıştır. LIDC veritabanı, 1010 hasta üzerinde ve yaklaşık 2600 nodül için dört farklı radyolog tarafından yapılmış değerlendirmeleri içermektedir. Değerlendirmeler içinde nodül radyografik tanımlayıcıları, nodülün habislik derecesi ve radyologlar tarafından sağlanan nodül alan bilgileri (*region of interest*) bulunmaktadır.

LIDC veritabanı, Zinovev ve diğerlerinin [14] de belirttiği üzere yapısı nedeniyle farklı bilgisayar destekli tanı yöntemlerinin geliştirilmesine uygun bir veri kümesidir. LIDC, 2009 yılında pilot sürümünün ve 2013 yılında tamamının yayınlandığı literatür için yeni sayılabilecek bir veri kaynağıdır. Literatürdeki birçok araştırmacı, veri edinmesi zor bir

alan olduđu için genellikle az sayıda nodül üzerinde çalışmıştır (Örn: [11], [15], [16]). LIDC arařtırmacıların veri bulma konusundaki sıkıntılarını gidermek için oluşturulmuş veritabanıdır. Literatürde de bu veritabanı üzerinde çalışmalar yeni yeni artmaktadır. Çalışmalar genellikle görüntü özniteliklerinden her bir radyografik tanımlayıcının tahmin edilmesi şeklindedir [14], [17]–[19]. Radyografik tanımlayıcıların sınıflamada alt düzey öznitelik olarak kullanıldığı [19], [20] ve radyologlar arasındaki değerlendirme farklılıklarını gidermek için radyografik tanımlayıcıların kullanıldığı çalışmalar [21], [22] da bulunmaktadır.

Bu tez çalışmasında sunulan bilgisayar destekli tanı yöntemlerinde, radyografik tanımlayıcılar nodülün habislik derecesinin tespitinde kullanılacaktır. Bu amaçla LIDC veritabanında sağlanan radyografik tanımlayıcı bilgileri kullanılmıştır. Görüntü öznitelikleri üzerinden radyografik tanımlayıcıların tahmin edileceđi ve bu çıktılar ile de habislik derecesinin tespit edileceđi iki aşamalı bir sınıflama yapısı kurulmuştur. Bu yapı üzerinden farklı sınıflama modelleri geliştirilmiş ve sonuçlar değerlendirilmiştir.

Literatürde yapılan çalışmalar göz önüne alındığında; radyografik tanımlayıcıların ara katman olarak kullanıldığı benzer çalışmalara sıklıkla rastlanmadığı görülmektedir. Radyografik tanımlayıcıların doğrudan sınıflamada kullanıldığı çalışmalara istinaden daha gerçekçi bir çözüm sunmaktadır. Tez kapsamında sunulan yöntemlerin kullanıcılara habislik derecesinin yanında radyografik tanımlayıcılar hakkında da ek bilgi vermesi açısından faydalı bir yaklaşım olarak düşünülmektedir. Bu yönüyle literatüre olumlu açıdan katkı sağlamaktadır.

Tez metni řu şekilde tasarlanmıştır:

2. Bölüm'de literatürde bulunan ilgili çalışmalara değinilmiştir. Bu çalışmalar içinde LIDC veritabanının kullanıldığı veya kullanılmadığı, farklı tanı sistemi parçaları üzerine sunulan, farklı sınıflama yöntemleri öneren arařtırmalar bulunmaktadır.

3.Bölüm'de, çalışmada kullanılan sınıflama, boyut indirgeme, öznitelik çıkarım yaklaşımlarından bahsedilmektedir.

4. Bölüm'de, LIDC veritabanı genel olarak tanıtılmıştır. LIDC veritabanından kesin referans (*ground truth/reference truth*) elde etme yöntemi hakkında bilgi verilmiştir. Ayrıca LIDC verisi üzerinde yapılmış temel sınıflama çözümlenmeleri de bulunmaktadır.

5-7. Bölümlerde, farklı iki katlı sınıflayıcı modelleri tanımlanarak deney sonuçları verilmiştir.

8-9. Bölüm'de, tez çalışmasından elde edilen sonuçlar incelenmiştir ve literatürdeki örnekleri arasında karşılaştırma yapılmıştır. Gelecekte yapılabilecek çalışmalar hakkında bilgi verilmiştir.

2. İLGİLİ ÇALIŞMALAR

Bu kesimde, alanda akciğer nodülleri üzerine çalışılmış olan bazı araştırmalara değinilecektir. Araştırmalar temel olarak beş başlıkta ele alınacaktır. Farklı başlıklar altında verilen çalışmaların, farklı parçalarında ortak kavramlar bulunabilir. Bu ayırım temelde tanı sistemlerinin kullanım amaçlarına göre yapılmıştır.

Alanda, burada değinilenler dışında akciğer nodülleri üzerine yapılan daha birçok çalışma vardır. Bu alanda yapılan farklı çalışmalara El-Baz ve diğerleri [23], Sluimer ve diğerleri [10], Suzuki [24], Abbod ve diğerleri [25] ve Mahfouf ve diğerleri [26] tarafından hazırlanan derlemelerden ulaşılabilir.

2.1. Habislik Tahmini Üzerine Yapılan Çalışmalar

2.1.1. Tekil Sınıflayıcı Kullanılan Çalışmalar

McNitt-Gray ve diğerleri [27], akciğer nodüllerini iyi huylu ya da kötü huylu olarak tanımlamaya çalışmıştır. Yöntem, dört ana kesime ayrılmıştır:

1- Görüntü elde etme: Veri kümesi, 31 hastaya ait (14 iyi huylu, 17 kötü huylu) yüksek çözünürlüklü tomografi görüntülerinden oluşmaktadır. Nodüllerin çapı 5 mm'den daha büyük, 3 cm'den daha küçüktür.

2- Nodül bölütleme: Mevcut bir görüntü işleme yazılımını genişleterek, tahmini nodül alanının işaretlendiği yarı-otomatik bir bölütleme yaklaşımı ile nodülün ayrılması işlemi gerçekleştirilmiştir.

3- Öznitelik çıkarımı: Bölütlenen nodül üzerinden şekil, boyut, doku ve yoğunluğa bağlı öznitelikler çıkarılmıştır. Bunlardan bazıları; en büyük/en küçük yarıçaplar (*max/min diameter*), alan (*area*), en-boy oranı (*aspect ratio*), düzensizlik (*irregularity*), yoğunluk (*compactness*), yuvarlaklık (*roundness*), dışbükeylik (*convexity*), katılık (*solidity*), şeklin merkezi ile sınırları arasında yapılan bazı ölçümler, fraktal çözümleme (*fractal analysis*); doku homojenliği (*homogeneity*), karmaşıklık (*complexity*), zıtlık (*contrast*), gibi şekil ve doku tabanlı özniteliklerdir.

4- Öznitelik seçme ve sınıflama: Öznitelik sayısı hayli fazla olduğu için hangi öznitelik kombinasyonunun daha iyi olduğunun belirlenmesi ve sonuca etki etmeyen özniteliklerin çıkarılması gerekmektedir. Öncelikle F istatistiği değerlendirilerek, her niteliğin tek başına sınıflamaya etkisi gözlemlenmiştir. Daha sonra en etkili öznitelik kümesinin seçilmesi için

doğrusal ayırım analizi kullanılmıştır. Genetik algoritma kullanılarak öznitelik kümesi belirleme yaklaşımı da gözden geçirilmiştir.

Minami ve diğerleri [28], zıtlığı artırılmış(*contrast enhanced*) tomografi görüntüleri üzerinde akciğer nodüllerini sınıflamışlardır. Biçim değiştirebilir yüzey modeli (*deformable surface model*) ve k-ortalamar kümeleme algoritması ile nodül bölütleme işlemi gerçekleştirilmiştir. Veri kümesi 68 nodülden - 28 iyi huylu, 40 kötü huylu - oluşmaktadır. Çıkarılan öznitelikler, şekil dizin değerleri (*shape index values*), eğrilik değerleri (*curvedness values*), uzaklık değerleri (*distance values*) bitişik bileşenlerin sayısı (*connected components*), boşluklar ve delikli yapılar gibi topolojik yapılardan elde edilmiştir. Nodüllerin kötü huylu olması potansiyelini araştırırken etrafındaki dokuların yapılarının da değerlendirilmesinin faydalı olacağını göstermişlerdir.

Gürcan ve diğerleri [15], tomografi görüntüleri üzerinde akciğer nodüllerinin tespiti için bir bilgisayar destekli tanı sistemi geliştirmişlerdir. Bölütleme aşamasında akciğer bölgeleri k-ortalamar algoritması ile ayrılmıştır. Her bir akciğer kesiti (*slice*) alt, üst ve orta kesim olarak sınıflandırılmıştır. Burada amaç farklı akciğer yapılarını (bronşlar, hava yolları, vs.) barındıran kesimleri ayırarak bölütleme işleminin başarımını arttırmaktır. Konum bilgisinin bölütleme başarımına etkisine dikkat çekilmiştir. Daha sonra bu kesimlerdeki akciğere ait bileşenler ağırlıklı k-ortalamar (*weighted k-means*) algoritması ile ayrılmıştır. İki ve üç boyutlu öznitelikler kullanılarak eğitilen kural tabanlı sınıflayıcılar ile yapı olarak birbirlerine benzeyebilen nodül ve kan damarları ayrıştırılmıştır. Doğrusal diskriminat çözümleme kullanılarak yanlış pozitif (*false positive*) oranı azaltılmıştır. Veri kümesi 34 hastadan edinilen, 63 nodüle sahip 1454 tomografi görüntüsü içermektedir (512x512, 12 bit). Çıkarılan öznitelikler üç boyutlu yüzey alanı, ortalama gri tonlama değeri; gri tonlama histogramının standart sapması, çarpıklık ve basıklık değerleridir. Ek olarak alan, çevre, yuvarlaklık, yoğunluk; en geniş ve en küçük eksen oranı; uydurulmuş elipsin dış merkezliliği (*fitted ellipse eccentricity*) hesaplanmıştır. İki boyutlu öznitelikler hesaplanırken, nesnelerin toplam çevresi ve alanlarının hızlı bulunması için “*bit quads*” algoritması kullanılmıştır. Gürcan ve diğerleri[15], çalışmanın sonunda, üç boyutlu özniteliklerin daha anlamlı olduğu sonucuna varmışlardır. Nodülü oluşturan tomografi kesimleri üst ve alt kesimlerde fazlaca küçüldüğü için bu kesimlerden çıkarılan iki boyutlu özniteliklerin güvenilir olmadığından bahsetmişlerdir.

Kawata ve diğerklerinin [16] geliřtirdiđi nodül sınıflama sisteminde özniteliklerin bir kısmı, üç boyutlu nodüllerin çevresindeki yapılar (damarlar, bronşlar, akciđer zarı) ile ilişkilerini ifade edecek şekilde çıkarılmıştır. Nodül iç yapısına ait öznitelikler, yoğunluk ve üç boyutlu eğrilikleri değerlendirilerek belirlenmiştir. Çok boyutlu öznitelik uzayından en iyi alt kümeyi seçmek için doğrusal ayırım analizi kullanılmıştır. Nodül bölütleme işlemi üç adımda gerçekleştirilmiştir: Ciđer bölgesinin çıkarılması, nodülün de bulunduğu ilgi alanının çıkarılması (*region of interest*), geometrik bir yaklaşım kullanarak nodülün bölütlenmesi. Sonuçlar ROC (*receiver operating characteristic*) yöntemi ile çözümlenmiş; sınıflama başarımı eğri altında kalan alanı ile değerlendirilmiştir. Veri kümesi 179'sı kötü huylu, 69'u iyi huylu 248 nodülden oluşmaktadır. Kötü huylu olanlar histolojik olarak da değerlendirilmiştir. İki yıl boyunca boyutu deđişmeyen nodüller iyi huylu olarak kabul edilmiştir.

Lo ve diğerkleri [29], tomografi üzerinde nodül bölütlemesi ve sınıflaması yapacak farklı yaklaşımlar geliřtirmişlerdir. Veri kümeleri 24'ü iyi huylu, 24'ü kötü huylu 48 akciđer nodülünden oluşmaktadır. Nodül bölütleme için kullanıcı tarafından nodül konumunun belirtilebileceđi gibi, CAD sistemi tarafından otomatik belirlenebilecek bir algoritma tasarlanmıştır. Bölütleme sırasında nodülün etrafındaki yapılar ile bağlantısı sonucu oluşabilecek hataları gidermek için bulanık kümeleme (*fuzzy clustering*) yöntemi kullanılmıştır. Öznitelik olarak; üç boyutlu hacim elde edildikten sonra en-boy oranı, yuvarlaklık, düzensizlik, genişlik, nodül yoğunluk dağılımı ve dışbükeylik kullanılmıştır. Ek olarak yoğunluk ve dokuya bađlı olarak, en büyük yoğunluk, ortalama, standart sapma, durađanlık, enerji, entropi ve fark entropisi kullanılmıştır. Sınıflama için yapay sinir ađları (*artificial neural network*) kullanılmıştır. Her aşamada yapay sinir ađını eğitmek için 47 nodül, deneme için de 1 adet nodül kullanılmıştır (*leave-one-out yaklaşımı*). Bu yolla sınıflama 48 adımda tamamlanmaktadır. ROC eğrisinin altında kalan alan değerlendirilerek sınıflama başarımı ölçülmüştür. Yöntemin çıktısı olarak iyi huylu (0) ya da kötü huylu (1) sonucu döndürölmektedir.

Xu ve diğerkleri [30], tomografi görüntülerinde tespit edilmiş akciđer nodüllerinin iyi ya da kötü huylu olarak ayrılmasında şekil, boyut, sınır ve yoğunluk deđerlerinin etkisini incelemişlerdir. Veri kümesi NELSON akciđer kanseri görüntüleme girişimine (NELSON *lung cancer screening trial* [31]) ait 405 katılımcıdan elde edilmiş, 469 belirsiz ya da kötü huylu olma potansiyeli olan nodülden oluşmaktadır. Bölütleme işlemi Lungcare yazılımı ile yapılmış, hatalı bölütlemelerde radyologlardan el ile bölütleme işlemi yapması

sağlanmıştır. Bu çalışmada nodüller; nodül şekline (yuvarlak/oval, poligonal ya da düzensiz) ve nodül sınırlarına göre (yuvarlak çıkıntılı, düzgün ya da düzensiz) sınıflandırılmıştır. Farklı nodül yapılarının kanser varlığına etkisi tek değişkenli ve çok değişkenli lojistik regresyon çözümlemesi yöntemiyle incelenmiştir. Bulgular, nodül yoğunluğu dışında diğer özelliklerin de kanser yapısını belirlemede etkisi olduğunu göstermiştir.

Wolf ve diğerleri [12], farklı kaynaklardan elde edilmiş ve farklı parametreler ile çekilmiş tomografi görüntüleri üzerinde işlem yapabilecek bir tanı sisteminin gerekliliği konusuna değinmişlerdir. Veri kümesi bir Kuzey Amerika ve üç Avrupa kaynaklı siteden elde edilmiş 185 yüksek çözünürlüklü tomografi görüntüsünden oluşmaktadır. Tasarladıkları tanı sistemi beş bölümden oluşmaktadır: 1) akciğerin bölütlenmesi, 2) nodülleri tespit edilmesi, 3) nodül olarak tespit edilen kesimlerden özniteliklerin çıkarılması, 4) yanlış negatif durumların azaltılması ve 5) bulguların kullanıcıya sunulması. Bölütleme işlemi, hacim üzerinde Hounsfield birimi üzerinden eşik değeri belirlenmesi ve çeşitli morfolojik işlemlerin uygulaması ile gerçekleştirilmiştir. Nodül tespiti için akciğer sınırlarının iyi bölütlenmesi gerekmektedir. Nodül kesimlerinin belirlenmesi için şüpheli konumların listesi elde edilmektedir. Bu nodül adayı kesimlerinden elde edilen iki/üç boyutlu öznitelikler ve konum bilgisi kullanılarak tanımlayıcılar oluşturulmaktadır. Elde edilen çok sayıda niteliğin içinden anlamlı alt kümenin saptanması için açgözlü (*greedy*) algoritma yöntemi kullanılmıştır. Bu çalışmada, SVM sınıflama yönteminin dengesiz dağılmış veri kümeleri için sorun oluşturduğundan bahsedilmiştir.

Way [2], tomografi görüntülerinden bulunan nodülleri iyi huylu ve kötü huylu olarak sınıflamak için bilgisayar destekli tanı sistemi geliştirmiştir. Sistem, nodülü, kendisini çevreleyen yapısal arkaplandan otomatik olarak bölütleyip; çıkarılan hacimsel bölgeden öznitelikleri elde ederek sınıflamaktadır. Veri kümesi, 44 kötü huylu ve 52 iyi huylu, toplam 96 nodülden oluşan; 58 hastaya ait tomografi görüntülerinden elde edilmiştir. Kötü huylu olanların 25 tanesi biopsi ile kesinleştirilmiş, kalanları PET taramaları veya metastas yapılarından dolayı vücudun farklı bölgelerindeki kanser sebebi ile bu sınıfa dahil edilmiştir. İyi huylu nodüllerde ise 10 tanesi biopsi ile kesinleştirilmiş, diğerleri 2 yıl boyunca süren tomografi taramalarında yapılarını bozmadıkları için bu sınıfa dahil edilmişlerdir. Bu veri kümesine ek olarak LIDC veri tabanına ait 23 nodül de bölütleme başarımlarını denetlemek için kullanılmıştır.

Bölütleme iki aşamada gerçekleştirilmiştir. Nodülün ilk sınırları k-ortalamar algoritması ile tespit edildikten sonra üç boyutlu etkin çevre modeli (*3D active contour model*) ile düzeltilmesi yapılmıştır. Çıkarılan morfolojik öznitelikler, hacim, yüzey alanı, çevre, en büyük yarıçap ve nodül içindeki en büyük/en küçük CT değeri (HU) kullanılarak oluşturulmuştur. Ek olarak nodülün gri-seviye yoğunluğundan elde edilen histogramın; ortalama, varyans, çarpıklık ve basıklık değerleri özniteliklere katılmıştır. Nodülün sınırları çevresindeki yapıların dokuları da önemli özniteliklerin çıkarılmasını sağlamaktadır. Öznitelik kümesi üzerinden, en önemli alt küme seçimini yapmak için doğrusal ayırım çözümlemesi (LDA) kullanılmıştır. Bir örneği dışarıda bırakarak (*leave-one-out, round robin*), sınıflayıcı eğitilmiştir. Sınıflayıcı başarımı ROC eğrisi altında kalan alanın ölçümü ile değerlendirilmiştir.

Kawata ve diğerleri [32] bu çalışmalarında diğerinden farklı olarak melez sınıflama yöntemi kullanmışlardır. Tomografi görüntülerinden elde edilen yoğunluk bilgisi ve üç boyutlu eğrilerden, öznitelikler türetilmiştir. Sınıflama adımında gözetimsiz k-ortalamar kümeleme yöntemi ve gözetimli doğrusal ayırım sınıflayıcı kullanılmıştır. Çok boyutlu öznitelik kümesinden anlamlı alt küme seçimi sonrası, sınıflama başarımı ROC eğrisinin altında kalan alan değerlendirilerek ölçülmüştür.

Aoyama ve diğerleri [33] akciğer nodüllerinin kötü huylu olabirlik ölçümünü yapan bir sistem tasarlamışlardır. Veri kümeleri 76 akciğer kanseri bulunan 147 tomografi kesiti ve 413 iyi huylu nodül bulunan 576 kesitten oluşmaktadır. Nodül bölütleme işleminde öncelikle nodül konumu bir radyolog tarafından işaretlenmektedir. Daha sonra dinamik programlama yaklaşımı ile nodül ayırma işlemi tamamlanmaktadır. Olabirlik ölçümü doğrusal ayırım analizi ile yapılmaktadır. Bu işlemde kullanılmak üzere 43 adet öznitelik tanımlanmıştır. Bunlardan 41'i nodül kesitinin iç ve dışından çıkarılan görüntü öznitelikleridir. Bazıları, kenar eğim yönelimleri (*edge gradient orientation*), asıl görüntü ve kenar eğimleri hesaplanmış görüntü üzerinde gri-seviye dağılımlarına bağlı bilgiler, dairesellik derecesi, elipslik derecesi, düzensizlik derecesi, nodülün iç ve dış kesimindeki yapılara bağlı histogramlardan elde edilen öznitelikler olarak verilebilir. Bunlara ek olarak klinik özniteliklerden yaş ve cinsiyet de kullanılmıştır. Bu çalışmada sadece iki boyutlu özniteliklerle ilgilenilmiştir. Nodülün her kesitinden çıkarılanlar değerlendirmeye alınarak bilgi kaybı engellenmeye çalışılmıştır.

Hanamiya ve diğeri [13], “*ekstrapulmoner malign neoplazm*” bulunan hastalarda, akciğeri nodüllerinin bulunma sıklığını belirleme için bir çalışma yapmışlardır. Veri kümesi 2mm kesit kalınlığı ile çekilmiş 308 hastaya ait tomografi görüntülerinden oluşmaktadır. Nodüller ve boyutları, büyüklükleri, akciğeri zarına olan uzaklıkları, on yıldan fazla tecrübesi olan üç radyolog tarafından işaretlenmiştir. Radyologlar tarafından çıkarılan öznitelikler Fisher’in kesin testi kullanılarak çözümlenmiştir. Nodülün habislik oranı çok değişkenli lojistik regresyon modeli kullanılarak tespit edilmiştir. Bu modelde nodül sayısı, yaş, cinsiyet, sigara içme geçmişi, nodül boyu ve nodül-akciğeri zarı arasındaki uzaklık öznitelik olarak kullanılmıştır. Bulgular nodül boyu ve akciğeri zarına uzaklığın habislik oranında etkili olduğunu göstermiştir.

2.1.2. Takım Sınıflayıcı Kullanılan Çalışmalar

Lee ve diğeri [20], radyografik tanımlayıcıları LIDC veri kümesi dışında kullanan nadir çalışmalardan birini yapmışlardır. Kendi veri kümelerini, LIDC ile ortak olan ve olmayan birkaç radyografik tanımlayıcı ile (cavitation, calcification, margin, vs) oluşturarak, radyografik tanımlayıcıları alt düzey özniteliklerle birlikte kullanarak habislik derecesini tahmin etmeye çalışmışlardır. Yöntemleri iki düzeyden oluşmaktadır. İlk adımda rastgele altalan (*random subspace*) ve genetik programlama kullanarak önemli olan özniteliklerin bulunmakta, ikinci adımda ise bu öznitelikler ve takım sınıflayıcılarla sınıflama yapılmaktadır.

Zinovev ve diğeri [17] tarafından yapılan çalışma, LIDC veritabanından elde edilen ve radyografik tanımlayıcılar üzerinde yapılan ilk çalışmalardandır. Araştırmacılar bu çalışmalarında, LIDC veritabanının 2009 yılında pilot sürümden önce yayınlanan 207 vaka ve 914 nodül örneğinin bulunduğu sürümünü kullanmışlardır. Etkin öğrenme (*active learning*) ve takım sınıflayıcılar ile görüntü özniteliklerini kullanarak radyografik tanımlayıcıları tahmin etmeye çalışmışlardır. Kesin referans olarak ortanca oylamayı (*median voting*) kullanmışlardır.

Zinovev ve diğeri [14], diğeri bir çalışmalarında *belief decision tree* tekil sınıflayıcı ve takım sınıflayıcı türevlerini kullanarak tekil/takım sınıflayıcı başarımlarını LIDC verisi üzerinde karşılaştırmışlardır. Bu çalışmada da görüntü özniteliklerini girdi olarak kullanmış ve radyografik tanımlayıcıları tahmin etmeye çalışmışlardır. Sonuçlar, takım sınıflayıcıların tekil sınıflayıcılara göre daha yüksek başarımlar elde ettiğini göstermiştir.

Vinay ve diğeri [34] farklı tekil sınıflayıcı türlerinin, LIDC veritabanından elde edilmiş kümeler üzerindeki başarımlarını incelemişlerdir. Bu çalışma için 2009 yılında yayınlanan 399 vakadan oluşan pilot sürümü kullanılmıştır. Çalışmada görüntü öznitelikleri ile radyografik tanımlayıcılar birleştirilerek habislik derecesinin tahmini yapılmaya çalışılmıştır. Radyografik tanımlayıcılar, görüntü öznitelikleri gibi alt düzey öznitelik olarak düşünülmüştür. Bu özniteliklerin ayrı ayrı sınıflama başarımına etkisi incelenmemiştir. Ayrıca kesin referans belirleme yaklaşımı tam olarak belli değildir. Her bir örnek radyologlar arası bir anlaşma beklenmeden kullanılmış olabilir.

Vinay ve diğeri [18] önceki çalışmalarına ek olarak takım sınıflayıcıların başarımlarını aynı sürümdeki LIDC veritabanı üzerinde denemişlerdir. Önceki yöntem ile aynı veri kümesi ve öznitelik grubunu kullanmışlardır. Takım sınıflayıcıların tekil sınıflayıcılardan başarılı sonuçlar ürettiğini ileri sürmüşlerdir.

Vinay ve diğeri [19], bu çalışmalarında ise takım sınıflayıcıları farklı yapılarda bir araya getirerek bir deney ortamı geliştirmişlerdir. Önceki yöntemleri ile aynı veri kümesi ve öznitelik grubunu kullanmışlardır.

2.2. Bilgi Erişim Sistemi Çalışmaları

Kawata ve diğeri [35], üç boyutlu akciğer tomografi veritabanında benzer görüntülerden faydalanarak, yeni bir vaka üzerinde tanı koymayı kolaylaştırmayı amaçlamışlardır. Bu amaçla veri tabanından benzer yapıda olan nodülleri getiren bir sistem tasarlamışlardır. Veri kümesi, “*National Cancer Center Hospital East and Tochigi Cancer Center*” tarafından sağlanan, 179 kötü huylu ve 69 iyi huylu toplam 248 nodülden oluşmaktadır. Nodül bölütleme işlemi üç adımda gerçekleştirilmiştir: Ciğer bölgesinin çıkarılması, nodülün de bulunduğu ilgi alanının çıkarılması (*region of interest*), geometrik bir yaklaşım kullanarak nodülün bölütlenmesi. Öznitelikler nodülün çevre yapılar ile olan ilişkisi ve üç boyutlu yapısından Thirion [36] tarafından önerilen yaklaşım ile elde edilmiştir. Nodüller arasındaki benzerlik ölçümü, iki nodülün ortak histogramlarının korelasyon katsayısı değerlendirilerek yapılmıştır.

Li ve diğeri [37], radyologlara yardımcı olan bilgisayar destekli tanı sistemlerinin başarımını arttırmak için bir tasarım geliştirmişlerdir. İçerik tabanlı görüntü erişim sistemleri (*Content-based image retrieval*), bu çalışma ile benzer bir araştırma alanına girmektedir. Mamogramlar üzerinde yaptıkları araştırmalarda, yapıları bilinmeyen kitleleri teşhis etmek için benzer görüntüleri bulmaya yarayacak, psikofiziksel benzerlik ölçüleri

bulmaya çalışmışlardır. Veri kümesi olarak, 76 kötü huylu nodüle ait ve 413 iyi huylu nodüle ait 489 ilgi alanı (*region of interest*) kullanılmıştır. Nodül birden çok tomografi kesiminde tespit edildiyse, bu kesitler içinde en geniş nodül alanına sahip olan değerlendirmeye alınmıştır. Nodül bölütlemesi için bölge büyütme (*region growing*) ve devingen programlama (*dynamic programming*) yaklaşımları kullanılmıştır. Öznitelik olarak etkin yarıçap, dairesellik derecesi ve zıtlık bilgisi kullanılmıştır. Bu öznitelikler radyologlar tarafından temel olarak ele alındığı için bunlar seçilmiştir. Veriler daha sonra 0-1 aralığında normalleştirilmiştir. Bu çalışmada en iyi bölütleme yaklaşımının radyolog tarafından yapıldığına değinilmiştir. Bölütleme sonucu nodülün öznitelikleri üzerinde önemli bir etki yapmaktadır.

Kawata ve diğerleri [38], iyi huylu ve kötü huylu akciğer nodüllerinin ayrımı için nodüllerin hacimsel gelişimini değerlendirmek için bir yöntem sunmuşlardır. Yöntem dört ana adımdan oluşmaktadır:

- 1) İki ardıl üç boyutlu tomografi görüntüsünün esnek olmayan çakıştırması (*rigid registration*).
- 2) İki ardıl ilgi alanı (ROI) görüntüsünün ilgin çakıştırması (*affine registration*)
- 3) Yerel hacimsel ilgi alanlarının (*local volumetric ROI*) esnek çakıştırması (*non-rigid registration*).
- 4) Ardıl zamana bağlı görüntülerin yerel değiştirme alanlarının çözümlenmesi.

Çalışma sonucunda deneysel çalışmalar lezyonların hacimsel ve çevresel yapılarının gelişimini değerlendirmek için önemli bir yöntem geliştirildiğini göstermiştir.

Jabon ve diğerleri [39], görüntü öznitelikleri ve radyografik tanımlayıcılar kullanarak içerik ve anlamsal tabanlı görüntü erişim sistemi tanımlamışlardır. Sistem, girdi olarak verilen CT örneklerine benzer örnekleri, Öklid ve kosinüs benzerlikleri kullanarak kullanıcıya döndürmektedir. Bu çalışma LIDC veritabanı üzerinde geliştirilmiştir.

2.3. Radyologlar Arası Anlaşma Üzerine Yapılmış Çalışmalar

Matsuki ve diğerleri [40], yapay sinir ağların nodül yapısının tespitinde kullanılabilirliğini ve bu yöntemin radyologların değerlendirme başarımı üzerindeki etkisini değerlendirmişlerdir. Veri kümesi altı yıl içinde çekilmiş tomografi görüntülerinden çıkarılmış; çapları 3 cm'den küçük, 99'u iyi huylu, 56'sı kötü huylu olmak üzere 155 nodülden oluşmaktadır. Öznitelik olarak yedi klinik veri ve çalışmaya katılan radyologlar

tarafından öznel olarak derecelendirilmiş 16 radyolojik bulgu kullanılmıştır. Klinik veriler yaş, cinsiyet, sigara içme geçmişi, belli başlı habislik, aile geçmişinde yatan habislik bilgisi, kilo kaybı ve belirtilerin şiddetidir. Radyolojik bulgular üç kesimde ele alınmıştır: 1) Nodülle alakalı öznelilikler: boyut şekil, içbükeylik, sınıf tanımı, sınırlardaki dalgalanması ve sivrileşme, hava bronkogramı, hava boşluğu, buzlu-cam opaklığı ve kalsitleşme. 2) İkincil anormallikler: Akciğer zarı girintileri, uydu lezyonlar, atardamar ilişkisi, toplardamar ilişkisi. 3) Ek anormallikler: Amfizematöz değişimler ve lenfdenopati. Bu öznelilikler tecrübesi 10 yılın üzerinde olan üç radyolog tarafından 0-10 arasında derecelendirilmiştir. Sınıflandırma için round-robin yöntemi (birini dışarıda bırakma yaklaşımı) kullanılarak eğitilmiş yapay sinir ağları kullanılmıştır. Gözlemci testi (*Observer test*) 25 iyi huylu ve 25 kötü huylu nodül seçilerek, 12 radyolog eşliğinde yapılmıştır.

Li ve diğerleri [22] literatürde çalışmamıza en çok benzeyen yapıyı kurmuşlardır. İki katlı sınıflayıcılarında, ilk aşamada 3 adet radyografik tanımlayıcıyı tahmin edip, ikinci aşamada habislik derecesine ulaşmaya çalışmışlardır. Amaçları radyologlar arasındaki değerlendirme çeşitliliğini azaltarak sistemin nesneliliğini artırmaktır. Sonuçları Kappa istatistiği [41] kullanarak değerlendirmişlerdir.

Horsthemke ve diğerleri [21], [42] de çalışmalarında, Li ve diğerlerinin [22] çalışmasında olduğu gibi, radyologlar arasındaki değerlendirme özneliliğini azaltmaya yönelik çalışma yapmışlardır. Nodüllerin sınırlarını ve radyologların değerlendirmelerini birleştirerek ortak karar elde etmeye çalışmışlardır. Sonuçlarını Kappa istatistiğine göre değerlendirmişlerdir.

2.4. Genel Değerlendirme

2.4.1 Veri Kümeleri

Alanda yapılan çalışmalar incelendiğinde, özellikle 2010 yılından önce yapılan çalışmaların çoğunlukla iyi değerlendirilmiş ve yüksek sayıda örnek bulunan ortak bir veritabanı kullanmadığı görülmüştür.

NELSON Trial[31]: NELSON veritabanı Belçika ve Hollanda'da bulunan hastalardan elde edilmiş verilerle Hollanda Sağlık Bakanlığı'nın onayıyla oluşturulmuş veritabanıdır. Yaşları 50-70 arasında erkek/kadın katılımcılar ile oluşturulmuştur. Ağ tabanlı katılım ve değerlendirmelerin yapıldığı geliştirme ortamına sahiptir. Hastalar üzerinde 1., 2., 4., 6. yıllarda klinik ölçümler ve CT ölçümleri tekrarlanmıştır. Bölütleme işlemi yarı otomatik olarak LungCare yazılımı ile yapılmıştır. Tamamlanamayan bölütlemeler için iki radyoloğun müdahalesi ile bölütleme yapılmıştır. Bölütleme için son karar üçüncü

radyolog tarafından verilmiştir. Nodülün yüzeyi ve şekline dair değerlendirmeler de veritabanına eklenmiştir.

LIDC[5]: *Lung Image Database Consortium(LIDC)* veritabanı, Amerika Birleşik Devletleri Ulusal Kanser Araştırmaları Enstitüsü'nün talebi ile oluşturulmuştur. Tamamlanması 12 yıl süren bu veritabanı son 5-6 yılda yapılan çalışmalarda sıklıkla kullanılmaya başlanmıştır. Yaklaşık 1000 hastaya ait akciğer nodül görüntü ve değerlendirmesini barındırmaktadır. Ayrıntılı bilgi Kesim 4'te verilmiştir.

LIDC veritabanı ile büyük boyutlu ve çoklu değerlendirmenin olduğu verilerin kullanımı artmıştır. LIDC, değerlendirici sayısı ve radyografik tanımlayıcılar ile habislik tahmin çalışmalarında kullanılması kolay bir veritabanı iken; NELSON Trial veritabanı zamana bağlı ölçümlerin yapılması için daha uygun bir veritabanıdır.

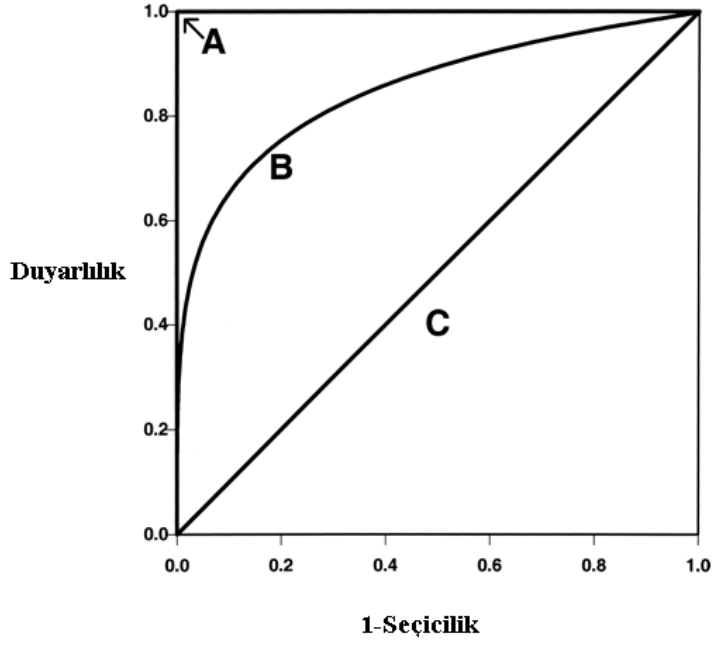
2.4.2. Ölçüm Metrikleri

Tekil sınıflayıcıların ve iki sınıflı habislik verisinin kullanıldığı çalışmalarda ROC eğrisinin altında kalan alan ve F-ölçümünün (*F-measure*) çoğunlukla kullanıldığını görmekteyiz.

ROC, y ekseninde doğru tanıma oranı (duyarlılık) ve x ekseninde yanlış tanıma oranının (1-seçicilik) farklı eşiklerde grafiğe dökülmüş halidir [43].

Şekil 2.1'de ROC eğrisi için örnek gösterim verilmiştir [44]. Şekildeki A testi ulaşılabilecek en iyi sonuçtur. Yani tüm örneklemin duyarlılığı ve seçiciliği %100'dür. C testi ise ayırtecdiliği olmayan en kötü sonuca sahiptir. A testinde eğri altında kalan alan 1 iken, C testinde bu alan 0.5'tir. Bir yöntemin tahmin etmede başarımı değerlendirmeye alınacaksa en azından C eğrisinin üstünde bir sonuç üretmesi beklenmelidir [45].

F-ölçümü, ROC ölçümüne alternatif olarak görülen bir ölçümdür [46]. Kesinlik (*precision, kes..*) ve geri çağırma (*recall, gç..*) üzerinden hesaplanır. Kesinlik, tüm doğru sınıflanan örneklere (*true positive + false positive*), doğru tanımlanan (*true positive*) örneklerin sayısının oranıdır. Geri çağırma ise doğru tanımlanan örneklerin, doğru ya da hatalı tanımlanan (*true positive + false negative*) örnek sayısına bölümüdür. F-ölçümü eşitlik (2.1)'deki ifadeyle bulunur.



Şekil 2.1 ROC eğrisi gösterimi [44].

$$F \text{ ölçümü} = \frac{2 \times \text{kes.} \times \text{gç.}}{\text{kes.} + \text{gç.}} \quad (2.1)$$

Takım sınıflayıcı çalışmaları sınıflama başarımı sıklıkla kullanılan bir metriktir. Duyarlılık ve seçicilik tıbbi araştırmalarda dikkate alınan diğer önemli ölçüm metrikleridir. Bu ölçümler hakkında ayrıntılı bilgi Kesim 3.9'da verilmiştir.

Radyologlar arası anlaşmanın ölçüldüğü çalışmalarda Cohen'in Kappa Katsayısı [41] sıklıkla kullanılmaktadır. Kappa katsayısı, kategorik ölçümlerde, değerlendiriciler arası anlaşmanın oranını bulmaktadır.

$P(a)$ değerlendiriciler arası ilişkili gözlenen anlaşma, $P(b)$ anlaşmanın şans eseri ortaya çıkma değeri olmak üzere Cohen'in Kappa Katsayısı (k) eşitlik (2.2)'deki ifade ile bulunur.

$$k = \frac{P(a) - P(b)}{1 - P(b)} \quad (2.2)$$

Kappa ve diğer benzer katsayılar da çıkan sonuç 1'e yakınsadıkça anlaşma oranı artmakta, 0'a yakınsadıkça bu oran azalmaktadır.

Krippendorff'un Alpha Katsayısı [47], Scott'ın Pi Katsayısı [48] da bu amaçla kullanılabilir ölçümlerdir.

2.4.3. Sınıflama Yaklaşımları

Çalışmalar genel olarak tam bir bilgisayarlı tanı destek sisteminin, bölütleme dahil, tüm parçalarını geliştirmeye odaklıdır. Sınıflama aşamalarında çoğunlukla LDA yaklaşımı kullanıldığı dikkat çekmektedir.

Klinik bilgi bazı çalışmaların sınıflama aşamasında görüntü öznelikleri ile birlikte kullanılmıştır. Özellikle, nodülün gelişiminin zaman içerisinde takip edildiği çalışmalarda bu bilgiden faydalanılmıştır.

LIDC veritabanının kullanılması ile birlikte radyografik tanımlayıcıların sınıflanması, radyologlar arası anlaşmanın değerlendirilmesi yönünde araştırmalar yapılmaya başlanmıştır. Takım sınıflayıcıların kullanıldığı çalışmalarda da bu veritabanı ile birlikte artış olmuştur. Ayrıca, bilgi erişim sistemlerinde sadece görüntü nitelikleri değil, radyografik tanımlayıcı bilgisi de kullanılabilir.

Tez kapsamında önerilen yöntemlerle 2015 yılına kadar yapılmış alandaki bilgimiz dahilindeki çalışmalar karşılaştırıldığında;

- Önerilen yöntemlerin radyografik tanımlayıcılar üzerinde yapılmış sayılı çalışmalardan olduğunu,
- Li ve diğerlerinin [22] çalışması dikkate alındığında, radyografik tanımlayıcıları iki aşamalı sınıflamada kullanan çalışma bulunmadığını,
- Sınıflama başarımı bağlamında, radyografik tanımlayıcıların habislik derecesinin tahmininde kullanıldığı çalışma olmadığını,

görmekteyiz.

Önerilen yöntemler, alandaki takım sınıflayıcılar grubuna dahil edilebilir. Takım sınıflayıcılar radyografik tanımlayıcılar düzeyinde tanımlanmıştır. Habislik derecesi tahmininde ise bulanık yaklaşımlar, kural tabanlı ve tekil sınıflayıcı tabanlı yaklaşımlar sunulmuştur. Alandaki takım sınıflayıcılar genelde görüntü nitelikleri üzerinde habislik tahmini aşamasında kullanılmaktadır [14], [17]–[19], [34]. Uygulama farklılığı olsa da, çalışmalarımızda olan takım sınıflama adımında boyut indirgeme işleminin kullanıldığı benzer çalışma da bulunmaktadır [20].

3. ARAÇ VE YÖNTEMLER

Bu kesimde kullanılacak araç ve yöntemlerden bahsedilecektir. Sırasıyla çalışmalarda kullanılan sınıflama yöntemleri, öznitelik ağırlıklandırma ve seçme yöntemleri, görüntü öznitelikleri, sınıflama algoritması seçimi için kullanılacak yöntemler anlatılacaktır.

3.1. Görüntü Öznitelikleri

Taramalar sonucunda belirlenen akciğer nodülleri bir ya da daha fazla görüntü kesitinde bulunabilir. Bunlar üzerinden çıkarılan öznitelikler 2 boyutlu (2B), 2.5 boyutlu (2.5B) ve 3 boyutlu (3B) olarak üçe ayrılır. 2B öznitelikler, nodülün en geniş olduğu kesitten çıkarılır. 2.5B öznitelikler ise tüm kesitlerden çıkarılır ve ortalaması değerlendirmeye alınır. 3B öznitelikler ise nodülün 3 boyut yapısının çıkarılması sonucu elde edilir. Çizelge 3.1’de tez çalışmasında kullanılan görüntü özniteliklerinin kısa tanımları verilmiştir.

Gri seviye eşdizimlilik matrisi (*Gray Level Co-Occurrence Matrix- GLCM*) üzerinden elde edilen öznitelikler [49], [50]:

Zıtlık (*Contrast – varianca, inertia*): Tüm görüntü üzerindeki pikseller ve komşulukları üzerinde zıtlığın ölçülmesi ile elde edilir (3.1).

$$\sum_{i,j} |i - j|^2 p(i, j) \quad (3.1)$$

Korelasyon (*Correlation*): Bir pikselin komşuluğu üzerinde nasıl korelasyona sahip olduğu ölçüsüdür (3.12).

$$\sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)p(i, j)}{\sigma_i \sigma_j} \quad (3.2)$$

Enerji (*Energy – uniformity, angular second moment*) : i ve j matris indislerini göstermek üzere, GLCM üzerinde tüm elemanların karelerinin toplamını döndürür (3.3).

$$\sum_{i,j} p(i, j)^2 \quad (3.3)$$

Homojenlik (*Homogeneity*): i ve j matris indislerini göstermek üzere, GLCM üzerindeki elemanların dağılımının, GLCM diagonal elemanlarına yakınlığını ölçer (3.4).

$$\sum_{i,j} \frac{p(i, j)}{1 + |i - j|} \quad (3.4)$$

Entropy: Entropi, görüntünün karmaşıklığını ifade eden ölçüm değeridir. Düz bir görüntüde entropi sıfırdır. Görüntüde karmaşıklık arttıkça entropi de artar. Görüntünün

sıkıştırılabilirliğinin belirlenmesi için de kullanılan bir ölçüdür (i entropi değerlerini göstermektedir, 3.5).

$$-\sum_{i=1}^n p_i \log_2 p_i \quad (3.5)$$

Zernike Momentleri [51]: Zernike momentleri, birim disk üzerinde dikey bir küme oluşturan karmaşık polinom kümeleridir. Zernike momentleri dönüşe karşı duyarsız özneliklerdir. Fakat ötelenmeye (*translation*) karşı duyarlıdırlar. Bu yüzden çalışma esnasında nodüller 128x128 genişliğinde bir kare alana, ağırlık merkezleri taşınarak hesaplama yapılmıştır.

Çizelge 3.1 Çıkarılan tüm öznelikleri kısa tanımları.

Öznelik	Kısa Tanım
max_eccentricity	EBNK* için çıkarılan "eccentricity" değeri.
max_solidity	EBNK için çıkarılan "solidity" değeri.
max_circularity	EBNK için çıkarılan "circularity" değeri.
max_aspectratio	EBNK için çıkarılan "aspectratio" değeri.
max_areaboundingbox	EBNK için çıkarılan sarmalayan kutu alan değeri.
max_std_deviation	EBNK için çıkarılan standart sapma değeri.
max_glcmm_contrast	EBNK için GLCM'den elde edilen zıtlık değeri.
max_glcmm_correlation	EBNK için GLCM'den elde edilen korelasyon değeri.
max_glcmm_energy	EBNK için GLCM'den elde edilen enerji değeri.
max_homogeneity	EBNK için GLCM'den elde edilen homojenlik değeri.
max_e	EBNK için GLCM'den elde edilen entropi değeri.
tmp_eccentricity	Tüm nodül kesitlerinin ortalamasından (TNKO) elde edilen "eccentricity" değeri
tmp_solidity	TNKO elde edilen "solidity" değeri
tmp_circularity	TNKO elde edilen "circularity" değeri
tmp_aspectratio	TNKO elde edilen "aspectratio" değeri
tmp_areaboundingbox	TNKO elde edilen sarmalayan kutu alan değeri.
tmp_std_deviation	TNKO elde edilen standart sapma değeri.
tmp_glcmm_contrast	TNKO elde edilen zıtlık değeri.
tmp_glcmm_correlation	TNKO elde edilen korelasyon değeri.
tmp_glcmm_energy	TNKO elde edilen değeri.
tmp_homogeneity	TNKO elde edilen homojenlik değeri.
tmp_e	TNKO elde edilen entropi değeri.
zm_nX_mY	m=X ve n=Y değerleri için Zernike moment değerleri. Çeşitli x,y değerleri için 64 öznelik seçilmiştir. Bu öznelikler, en büyük nodül kesitinden çıkarılmıştır.
srnd_max_glcmm_contrast	SRND ve EBNK elde edilen zıtlık değeri
srnd_max_glcmm_correlation	SRND ve EBNK elde edilen korelasyon değeri
srnd_max_glcmm_energy	SRND ve EBNK elde edilen enerji değeri
srnd_max_homogeneity	SRND ve EBNK elde edilen homojenlik değeri
srnd_max_e	SRND ve EBNK elde edilen entropi değeri
srnd_max_std_deviation	SRND ve EBNK elde edilen standart sapma değeri
srnd_tmp_glcmm_contrast	SRND ve TNKO elde edilen zıtlık değeri
srnd_tmp_glcmm_correlation	SRND ve TNKO elde edilen korelasyon değeri
srnd_tmp_glcmm_energy	SRND ve TNKO elde edilen enerji değeri
srnd_tmp_homogeneity	SRND ve TNKO elde edilen homojenlik değeri
srnd_tmp_e	SRND ve TNKO elde edilen entropi değeri
srnd_tmp_std_deviation	SRND ve TNKO elde edilen standart sapma değeri
tmp_glcmm_autoc	TNKO elde edilen "auto correlation" değeri[52]
tmp_glcmm_cprom	TNKO elde edilen "cluster prominence" değeri[52]
tmp_glcmm_cshad	TNKO elde edilen "cluster shade" değeri[52]
tmp_glcmm_denth	TNKO elde edilen "difference entropy" değeri[49]
tmp_glcmm_dissi	TNKO elde edilen "dissimilarity" değeri[52]
tmp_glcmm_dvarh	TNKO elde edilen "difference variance" değeri[49]
tmp_glcmm_idmnc	TNKO elde edilen "inv. dif. m. norm" değeri[53]
tmp_glcmm_indnc	TNKO elde edilen "inv. dif. norm" değeri[53]
tmp_glcmm_inf1h	TNKO elde edilen "inf. m. of corr. 1" değeri[49]
tmp_glcmm_inf2h	TNKO elde edilen "inf. m. of corr. 2" değeri[49]
tmp_glcmm_maxpr	TNKO elde edilen "max. probability" değeri[52]

tmp_glc_m_savgh	TNKO elde edilen "sum average" değeri[49]
tmp_glc_m_senth	TNKO elde edilen "sum entropy" değeri[49]
tmp_glc_m_sosvh	TNKO elde edilen "sum of squares" değeri[49]
tmp_glc_m_svarh	TNKO elde edilen "sum variance" değeri[49]
srnd_tmp_glc_m_autoc	SRND ve TNKO elde edilen "auto correlation" değeri[52]
srnd_tmp_glc_m_cprom	SRND ve TNKO elde edilen "cluster prominence" değeri[52]
srnd_tmp_glc_m_cshad	SRND ve TNKO elde edilen "cluster shade" değeri[52]
srnd_tmp_glc_m_denth	SRND ve TNKO elde edilen "difference entropy" değeri[49]
srnd_tmp_glc_m_dissi	SRND ve TNKO elde edilen "dissimilarity" değeri[52]
srnd_tmp_glc_m_dvarh	SRND ve TNKO elde edilen "difference variance" değeri[49]
srnd_tmp_glc_m_idmnc	SRND ve TNKO elde edilen "inv. dif. m. norm" değeri[53]
srnd_tmp_glc_m_indnc	SRND ve TNKO elde edilen "inv. dif. norm" değeri[53]
srnd_tmp_glc_m_inf1h	SRND ve TNKO elde edilen "inf. m. of corr. 1" değeri[49]
srnd_tmp_glc_m_inf2h	SRND ve TNKO elde edilen "inf. m. of corr. 2" değeri[49]
srnd_tmp_glc_m_maxpr	SRND ve TNKO elde edilen "max. probability" değeri[52]
srnd_tmp_glc_m_savgh	SRND ve TNKO elde edilen "sum average" değeri[49]
srnd_tmp_glc_m_senth	SRND ve TNKO elde edilen "sum entropy" değeri[49]
srnd_tmp_glc_m_sosvh	SRND ve TNKO elde edilen "sum of squares" değeri[49]
srnd_tmp_glc_m_svarh	SRND ve TNKO elde edilen "sum variance" değeri[49]
max_glc_m_autoc	EBNK elde edilen "auto correlation" değeri[52]
max_glc_m_cprom	EBNK elde edilen "cluster prominence" değeri[52]
max_glc_m_cshad	EBNK elde edilen "cluster shade" değeri[52]
max_glc_m_denth	EBNK elde edilen "difference entropy" değeri[49]
max_glc_m_dissi	EBNK elde edilen "dissimilarity" değeri[52]
max_glc_m_dvarh	EBNK elde edilen "difference variance" değeri[49]
max_glc_m_idmnc	EBNK elde edilen "inv. dif. m. norm" değeri[53]
max_glc_m_indnc	EBNK elde edilen "inv. dif. norm" değeri[53]
max_glc_m_inf1h	EBNK elde edilen "inf. m. of corr. 1" değeri[49]
max_glc_m_inf2h	EBNK elde edilen "inf. m. of corr. 2" değeri[49]
max_glc_m_maxpr	EBNK elde edilen "max. probability" değeri[52]
max_glc_m_savgh	EBNK elde edilen "sum average" değeri[49]
max_glc_m_senth	EBNK elde edilen "sum entropy" değeri[49]
max_glc_m_sosvh	EBNK elde edilen "sum of squares" değeri[49]
max_glc_m_svarh	EBNK elde edilen "sum variance" değeri[49]
srnd_max_glc_m_autoc	SRND ve EBNK elde edilen "auto correlation" değeri[52]
srnd_max_glc_m_cprom	SRND ve EBNK elde edilen "cluster prominence" değeri[52]
srnd_max_glc_m_cshad	SRND ve EBNK elde edilen "cluster shade" değeri[52]
srnd_max_glc_m_denth	SRND ve EBNK elde edilen "difference entropy" değeri[49]
srnd_max_glc_m_dissi	SRND ve EBNK elde edilen "dissimilarity" değeri[52]
srnd_max_glc_m_dvarh	SRND ve EBNK elde edilen "difference variance" değeri[49]
srnd_max_glc_m_idmnc	SRND ve EBNK elde edilen "inv. dif. m. norm" değeri[53]
srnd_max_glc_m_indnc	SRND ve EBNK elde edilen "inv. dif. norm" değeri[53]
srnd_max_glc_m_inf1h	SRND ve EBNK elde edilen "inf. m. of corr. 1" değeri[49]
srnd_max_glc_m_inf2h	SRND ve EBNK elde edilen "inf. m. of corr. 2" değeri[49]
srnd_max_glc_m_maxpr	SRND ve EBNK elde edilen "max. probability" değeri[52]
srnd_max_glc_m_savgh	SRND ve EBNK elde edilen "sum average" değeri[49]
srnd_max_glc_m_senth	SRND ve EBNK elde edilen "sum entropy" değeri[49]
srnd_max_glc_m_sosvh	SRND ve EBNK elde edilen "sum of squares" değeri[49]
srnd_max_glc_m_svarh	SRND ve EBNK elde edilen "sum variance" değeri[49]

*EBNK: En büyük nodül kesiti, TNKO: Tüm nodül kesitlerinin ortalamasından, SRND: Nodül etrafındaki alanlardan.

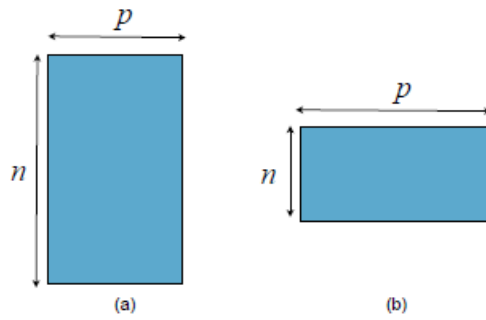
3.2. Öznitelik Kümelerinde Boyut İndirgeme

Sınıflama, kümeleme veya regresyon gibi makine öğrenimi yöntemlerinde kullanılan veri kümeleri genelde yüksek boyutlu ve veriyi tam anlamıyla ifade edip etmediği kesinleşmemiş özniteliklerden oluşmaktadır. Bu öznitelikler üzerinde boyut indirgeme, sadece hesaplama etkinliği ve karmaşıklığın düşürülmesi değil, analizin tutarlılığının artırılması için de kullanılmalıdır [54].

Özellikle biyoinformatik ve çoklu ortam veri analizlerinde büyük boyutlarda öznitelik sayılarına ulaşılabilmektedir. İstatistikçilere göre tanımlayıcı değişkenlerin örnek sayısından fazla olması bir problem olarak görülmektedir (*Big p small n problem*, Şekil 3.1). Teoride, bir veri kümesi üzerinde çıkarım yapılacaksa örnek sayısı, öznitelik sayısına oranla üssel olarak artmalıdır [55].

Birçok makine öğrenimi yöntemi yüksek boyutlu verilere karşı duyarlıdır (*curse of dimensionality*) [56]. Hız ve analiz başarımı yüksek boyutlu veriler nedeniyle düşmektedir.

Yüksek boyutlu veri kümeleri yukarıda belirtilen bazı sorunları ortaya çıkarsa da bazı fırsatlar da sunmaktadırlar [57]. Veri kümesinin ifade edip etmediği bilinmeyen bazı öznitelikler değerlendirilerek yeni teorik gelişmeler kaydedilebilir [58]. Her ne kadar yeni geliştirilen yöntemler büyük hesaplama karmaşıklığına sahip olsalar ve yüksek boyutları ele alabilseler de boyut indirgemek veri modellenmeden önce birçok uygulamada değerlendirilebilecek bir süreçtir.



Şekil 3.1 (a) İstenen veri kümesi yapısı, (b) istenmeyen veri kümesi yapısı (*big p small n problem*) [59].

Neden boyut indirgeme kullanılmalıdır [59]?

- Makine öğrenimi algoritmalarında karmaşıklık veri kümesinin boyu ve özniteliklerin boyutu ile artmaktadır. Hesaplama ve bellek ihtiyacını azaltmak, oluşturulacak sonuç modelin daha basit şekilde ifade edilmesini sağlamak için gereklidir.
- Veriyi ifade eden özniteliklerin azaltılması, daha sonraki adımlarda bu özniteliklerin hesaplanması maliyetini azaltır.
- Basit modeller, küçük veri kümelerinde daha güvenilir sonuçlar vermektedir.
- Veri daha az öznitelikle ifade edildiğinde, öznitelikler ve veri arasında ilişki kurmak ve bundan bilgi sahibi olmak daha kolay olacaktır.
- Veri boyutunun azalması çözümlenme ve aykırı değerlerin bulunması aşamasında görselliğin daha kolay oluşturulmasını sağlamaktadır.

Öznitelik boyut indirgemedede kullanılan iki genel yaklaşım vardır. Bunlar öznitelik alt küme seçimi (*feature subset selection, feature selection*) ve öznitelik çıkarımıdır (*feature extraction*).

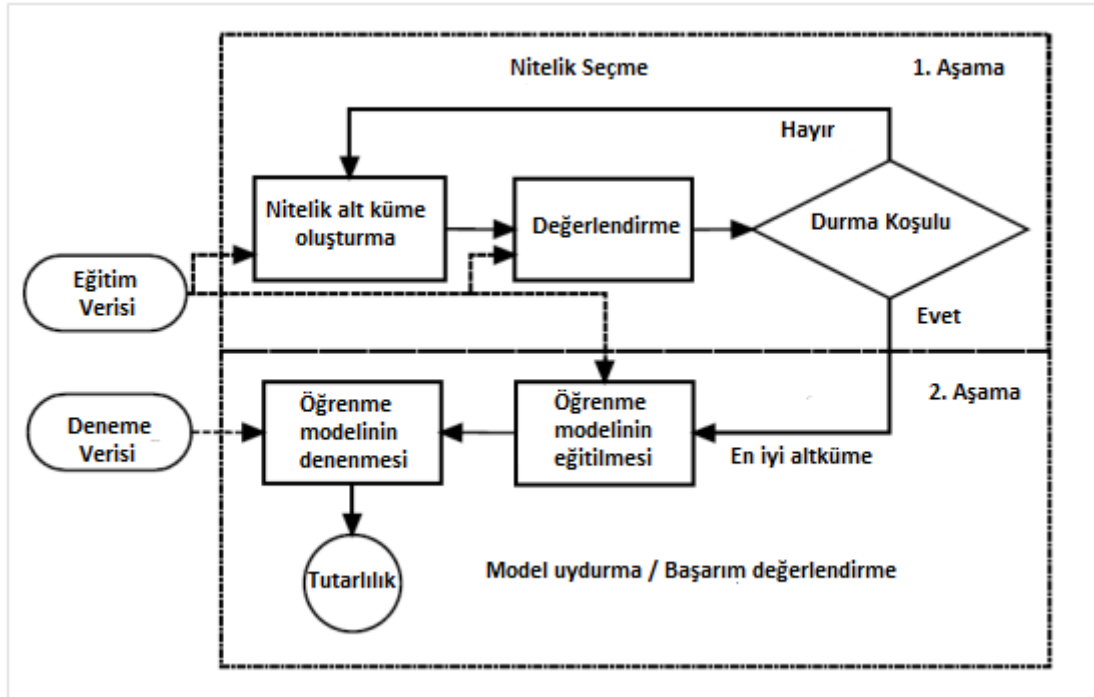
3.2.1. İleri ve Geri Adımlı Öznitelik Alt Küme Seçimi

Öznitelik alt küme seçimi, verideki örnekleri tanımlayan birçok öznitelik içinden bilgi kazanımını (*information gain*) veya doğruluk (*accuracy*) değerlerini en yüksek yapacak alt kümenin belirlenmesi işlemidir (Şekil 3.2).

Tanım 1:

$X = \{x_i \mid i = 1 \dots N\}$ şeklinde verilen bir öznitelik kümesinde, amaç işlevi $J(Y)$ 'yi en yüksek kılacak ve $M < N$ olacak şekilde bir Y_M altkümesi seçilmelidir (3.6).

$$Y_M = \{x_{i_1}, x_{i_2}, \dots, x_{i_M}\} = \arg \max_{M, i_M} J \{x_i \mid i = 1 \dots N\} \quad (3.6)$$



Şekil 3.2 Öznitelik alt küme seçimi için genel gösterim [60].

Alt küme seçimi için ileri (*forward*) ve geri (*backward*) adımlı yaklaşımlar kullanılabilir. İleri adımlı yaklaşımlarda, öncelikle seçilecek küme boştur. Daha sonra öznitelikler bu kümeye teker teker eklenerek kazanım hesaplanır. Yeni bir öznitelik eklendiğinde kazanımın artmadığı (ya da hatanın azalmadığı) koşullarda işlem sonlandırılır (Tanım 2). Geri adımlı yaklaşımlarda ise tam tersine ilk önce tüm öznitelikler seçili başlanır. Her adımda öznitelikler çıkarılıp; çıkan öznitelik hatayı yüksek oranda arttırdığı durumlarda işlem sonlandırılır (Tanım 3). Şekil 3.3’de genelleştirilmiş algoritma [60] verilmiştir.

Tanım 2 [59]:

$X = \{ x_i \mid i = 1 \dots N \}$ şeklinde verilen bir öznitelik kümesinde, ilk adımda öznitelik kümesi boştur. $F = \{ \}$

Tüm olası x_i için model eğitilir ve $E(F \cup x_i)$ hesaplanır.

$$j = \arg \min_i E (F \cup x_i) \quad (3.7)$$

olacak şekilde x_j aşağıdaki koşulu sağlarsa kümeye eklenir.

$$E(F \cup x_j) < E(F) \quad (3.8)$$

Tanım 3 [59]:

$X = \{ x_i \mid i = 1 \dots N \}$ şeklinde verilen bir öznitelik kümesinde, ilk adımda öznitelik kümesi boştur. $F = \{ \}$

Tüm olası x_i için model eğitilir ve $E(F \cup x_i)$ hesaplanır.

$$j = \arg \min_i E(F - x_i) \quad (3.9)$$

olacak şekilde x_j aşağıdaki koşulu sağlarsa kümeden çıkarılır.

$$E(F - x_j) < E(F) \quad (3.10)$$

1. Veri kümesi üzerinde tüm kestiricileri (*predictors*) kullanarak modeli eğit
2. Model başarımını hesapla
3. Değişken önemliliğini hesapla (*variable importance*) ya da değişkenleri derecelendir (*ranking*)
4. Her alt küme boyutu için $S_i, i = 1 \dots S$
 - 4.1. En önemli S_i değişkeni tut.
 - 4.2. (tercihe bağlı) Veriyi ön işle
 - 4.3. Veri kümesi üzerinde S_i kümesindeki kestiricileri kullanarak modeli eğit
 - 4.4. Model başarımını hesapla
 - 4.5. (tercihe bağlı) Her kestirici için dereceleri yeniden hesapla.
5. S_i üzerindeki başarımlarını hesapla
6. Yeterli kestirici sayısını hesapla
7. Her kestiricinin nihai derecelerini hesapla
8. En iyileştirilmiş S_i üzerinde son modeli oluştur.

Şekil 3.3 Özyineli öznitelik seçim algoritması [61].

Öznitelik alt kümesi seçilirken mevcut olan tek veri kümesi üzerinde seçim yapmak hatalara neden olabilmektedir. Çünkü daha sonra genişleyecek ya da farklı örnekler gelecek bir veri kümesinde daha farklı alt kümeler anlamlı çıkabilmektedir. Bu yüzden algoritmalar üzerinde *bootstrap* ya da çapraz geçişleme (*cross validation*) gibi genelleme yaklaşımları ile daha tutarlı sonuçlar elde etmek mümkündür. Şekil 3.4'de, Şekil 3.3 için yeniden örnek alma ile genelleştirme eklenmiş hali verilmiştir.

1. Her bir yeniden örnek (*resampling*) adımı için
 - 1.1. Yeniden örnek ile veriyi eğitim ve deneme kümesi olarak ayır.
 - 1.2. Eğitim veri kümesine göre tüm kestiricileri kullanarak modeli eğit.
 - 1.3. Test veri kümesi üzerinde kestirim yap.
 - 1.4. Değişken önemliliğini hesapla (variable importance) ya da değişkenleri derecelendir
 - 1.5. Her alt küme boyutu için $S_i, i = 1 \dots S$
 - 1.5.1. En önemli S_i değişkeni tut.
 - 1.5.2. (tercihe bağlı) Veriyi ön işle
 - 1.5.3. Veri kümesi üzerinde S_i kümesindeki kestiricileri kullanarak modeli eğit
 - 1.5.4. Deneme verilerini kullanarak kestirim yap.
 - 1.5.5. (tercihe bağlı) Her kestirici için dereceleri yeniden hesapla.
2. S_i üzerindeki başarımlar profilini deneme veri kümesini kullanarak hesapla.
3. Yeterli kestirici sayısını hesapla
4. Her kestiricinin nihai derecelerini hesapla
5. En iyileştirilmiş S_i üzerinde son modeli, tüm veri kümesini kullanarak oluştur.

Şekil 3.4 Yeniden örnek kullanan özyineli öznelik seçim algoritması [61].

3.2.2. Relieff Yöntemi ile Öznelik Altküme Seçimi

Tez çalışmasında, öznelik alt küme seçim yöntemi olarak Relieff [62] yöntemi kullanılmıştır. Bu yöntem Kira ve Rendell tarafından sunulan Relief [63] yönteminin adımlı olarak geliştirilmiş halidir.

Relief yönteminde, bir verikümesini tanımlayan n öznelik 0 ile 1 arasında ağırlıklandırılacaktır (W). Algoritmanın her adımında veri kümesinden rastgele bir örnek seçilir. Ve belirli bir komşuluk sayısında (k), bu örneğe en yakın örnekler nitelik vektörleri (x) arasındaki Öklid uzaklığı ölçülerek seçilir. Bu örnekler içindeki, seçilen örnek ile aynı sınıfa ait olanlar yakın uyuşan (*near hit - nh*), farklı sınıfa ait olanlar yakın uyuşmaz (*near miss - nm*) olmak üzere, i . ağırlık eşitlik (3.11)'deki ifadeyle güncellenir.

$$W_i = W_i - (x_i - nh_i)^2 + (x_i - nm_i)^2 \quad (3.11)$$

İşlemin sonunda ağırlık vektörü W , adım sayısına bölünür. Tüm niteliklerin ağırlıklarının normleştirilmiş haline bu vektördeki değerlerden ulaşılır. Öznelik seçimi için kullanıcının belirlediği eşik değerleri kullanılabilir.

Relief yöntemi gürültüye duyarlı, doğrusal zaman gerektiren, kesikli ya da sürekli veri üzerinde çalışabilen bir yaklaşımdır. Eksik ve az veri üzerinde yetersiz kalan bu algoritmanın geliştirilmiş sürümlerinde bu sorunlar giderilmeye çalışılmıştır.

Relieff yöntemi çoklu sınıflı verilerde ve eksik değerlerin olduğu verilerde çalışabilmektedir. Yöntem, özniteliklere ağırlıklar atayarak, en önemliden en az önemliye doğru sıralamaktadır. Bu yöntemde eksik olan özniteliklerin tamamlanması için de bir yaklaşım geliştirilmiştir. Tez kapsamında farklı bir eksik değer tamamlama yaklaşımı kullanıldığı için bu özellik dikkate alınmamıştır.

3.3. Eksik Değerlerin Tamamlanması

Her bir veri kümesinde nodül boyutunun küçüklüğüne bağlı olarak bazı örneklerinin öznitelik değerleri eksiktir. Tez çalışması bağlamında, bu eksik değerleri (*missing values*) tamamlamak için k en yakın komşuluk yakıştırma (*k-nearest neighbour imputation*) [64] yöntemi kullanılmıştır. Bu yöntem, örneğin eksik olmayan değerlerini kullanarak en yakın k komşuluğunu bulur. Eksik olan veri, komşuluklarda eksik olmayan değerlerin ortalaması kullanılarak tamamlanır. *Expectation maximization, regression, single value decomposition* gibi yaklaşımlar da bu problemi çözmek için kullanılabilir.

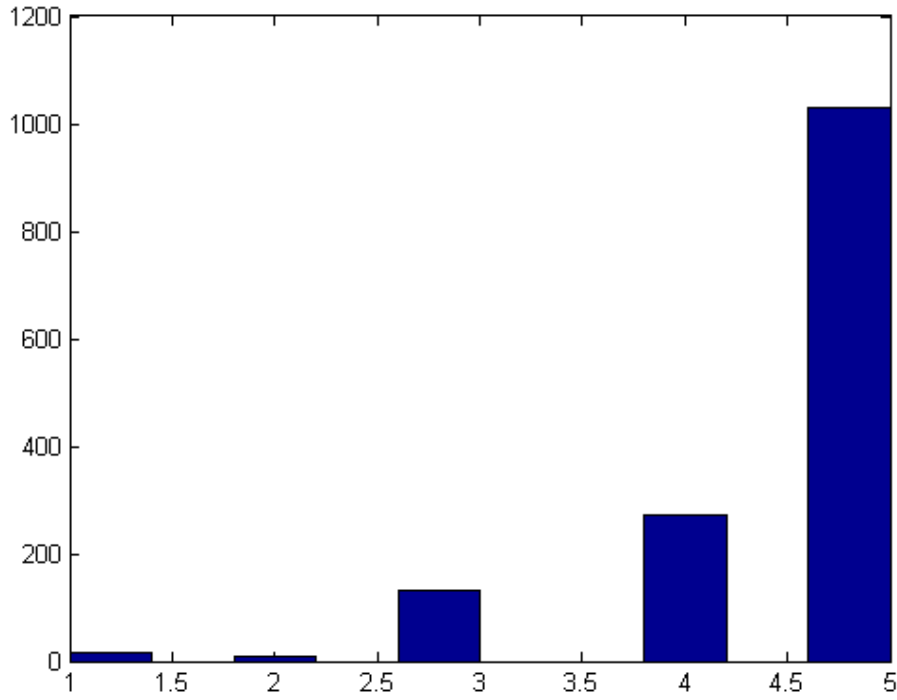
3.4. Veri Kümelerinin Dengelenmesi

Birçok radyografik tanımlayıcının veri kümesi yüksek oranda dengesiz dağılmıştır. Şekil 3.5’de “*sublety*” tanımlayıcısı için sınıf histogramı verilmiştir. Bu tanımlayıcıda nodüllerin %70’e yakın oranı 5 sınıfına aittir. Benzer durum birçok tanımlayıcı için geçerlidir. Az örneğe sahip sınıfın daha az ifade edilmesi sorununu çözmek için veri dengeleme yöntemleri kullanılmıştır.

Verinin örnek arttırımı için *Synthetic Minority Over-sampling Technique (SMOTE)* [65] yöntemi kullanılmıştır. SMOTE, aynı örnekleri çoğaltmak yerine, mevcut örnekleri kullanarak yapay veriler üretmektedir. İhtiyaç duyulan örnek arttırımına bağlı olarak, azınlıkta olan sınıfın seçilen bir örneği k komşuluğundaki örnekleri kullanılarak çoğaltılır. Örnek ve komşuları seçildikten sonra, öznitelik vektörleri arasındaki fark hesaplanır. Bulunan fark 0 ile 1 arasında seçilen rastgele bir değer ile çarpılır ve örneğin değerleriyle toplanır. Böylece yeni örnek, seçilen örnek ve komşuları arasında bir düzlemde yer alır [65]. SMOTE yöntemi, basit ve etkili bir yöntemdir ve örnek tekrar etme gibi aşırı uyum (*over fitting*) sorunlarına yol açmamaktadır.

Aşırı örneğe sahip sınıflarda ise, sınıfa özel öznitelik uzayında tanımlanan dışbükey örtü yöntemi ile veri indirgenir. Dışbükey örtü üzerinde öznitelikleri bulunan örnekler seçilir ve geri kalanlar atılır. Eşitlik (3.12)'de, CH , n boyutlu bir örneklem uzayında noktalar kümesi olmak üzere; N noktaların sayısı (p_1, \dots, p_N) ve λ_j , p_j 'nin ağırlığıdır. CH noktalarına sahip tüm dışbükey kümelerin kesişimine dışbükey örtü denir [66].

$$CH \equiv \{\sum_{j=1}^N \lambda_j p_j : \lambda_j \geq 0 \forall j, \sum_{j=1}^N \lambda_j = 1\} \quad (3.12)$$



Şekil 3.5 *Sublety* tanımlayıcısı için sınıf histogramı. (Dikey eksen örnek sayısıdır)

3.5. Sınıflama Yöntemleri

Geliştirilen yöntemler ile karşılaştırmada kullanılacak sınıflayıcılara ait bilgiler bu kesimde verilecektir. Her başlık altında kullanılan parametreler verilmiştir. Parametrelerin seçilmesi için Orange [67] yazılımının parametre en iyileştirme yaklaşımları kullanılmıştır.

3.5.1. C4.5

C4.5 [68] sınıflama algoritması ID3 [69] ve CLS [70] gibi algoritmalarla aynı kökenden gelmektedir. C4.5 karar ağacı olarak ifade edilebilen sınıflandırıcılar üretir. Aynı zamanda daha anlaşılır kural kümeleri şeklinde gösterilebilen sınıflandırıcılar da oluşturabilir [71].

Karar ağaçları, bir örneğin kök düğümünden başlayarak her bir düğümde test edildiği, yaprak düğümlerine ulaştığında ise sınıf değerleri ile sınıflama sonucuna ulaşılan yapılardır.

Bir S örneklem verildiğinde, C4.5 ilk olarak böl ve fethet (*divide and conquer*) yaklaşımı kullanarak bir karar ağacı üretir:

- Eğer S'deki tüm örnekler aynı sınıfa aitse ya da S küçükse, düğüm yaprak olarak belirlenir ve S kümesinde en sık rastlanan sınıfla etiketlenir.
- Aksi durumda; iki ya da daha fazla çıktısı olan tek bir öznelik seçilir. Bu öznelik ağacın kökü gibi düşünülür ve her bir çıktı için test yapılır. Test sonucu en çok kazanımın olduğu öznelik ve çıktılar ağacın düğümü olarak eklenir.

Düğümde kullanılacak niteliğin belirlenmesi için farklı testler uygulanabilir. Bu testlere Gini index, bilgi kazanımı oranı (*information gain ratio*) örnek verilebilir.

Karar ağacını temel alan algoritmalarda aşırı büyüyen ağaçları düzenlemek ve verinin bu bağlamda aşırı uyuma (*overfitting*) neden olmaması için budama (*pruning*) yaklaşımları da uygulanmaktadır. Budama yapraklardan köke doğru uygulanır. En basit budama yöntemi olarak; yapraklardan başlayarak her karar düğümü, düğümün altında yer alan en popüler sınıf değeri ile değiştirilir. Eğer sınıflama başarımında değişiklik yoksa bu işleme devam edilir.

3.5.2. Destek Vektör Makinaları (*Support Vector Machines - SVM*)

SVM [72] sağlam teorik altyapısı ile makine öğrenimi algoritmaları içinde güvenilir ve tutarlı sonuçlar üreten, az sayıda örnek ile eğitilebilen, verinin boyutuna duyarsız olan bir algoritmadır [71].

Verinin doğrusal ayrılabilirliği durumları daha iyi ifade etmek için aşağıdaki örnek verilebilir [59]. İki sınıflı ve her biri -1 ve +1 olarak işaretlendiği varsayalım. Örneklem $X = \{x^t, r^t\}$, eğer $x^t \in C_1$ ise $r^t = +1$, eğer $x^t \in C_2$ ise $r^t = -1$ olsun. En iyi ayrılmayı sağlayan hiper düzlem w ve w_0 aşağıdaki gibi ifade edilebilir (3.13 ve 3.14).

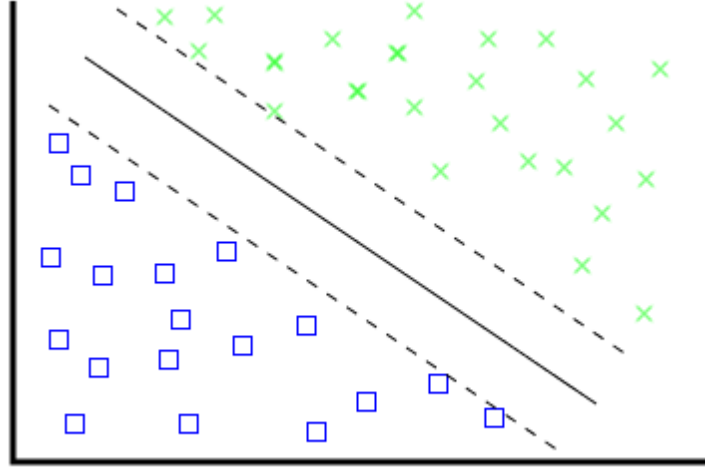
$$r^t = +1 \text{ için } w^t x^t + w_0 \geq +1 \quad (3.13)$$

$$r^t = -1 \text{ için } w^t x^t + w_0 \leq -1 \quad (3.14)$$

Yukarıdaki eşitlikler aşağıdaki gibi tekrar yazılabilir (3.15).

$$r^t (w^t x^t + w_0 \geq +1) \quad (3.15)$$

Hiperdüzlem hesaplandıktan sonra Şekil 3.6’de de görüleceği üzere iki sınıf tam ortadan geçen düzlem ile doğrudan ayrılabilir. Fakat daha iyi bir genelleştirme sağlanması açısından hiperdüzlem ile sınıfların hiperdüzleme olan mesafelerini (*margin*) en iyi şekilde ifade etmek gereklidir.



Şekil 3.6 SVM için iki boyutta ve iki öznelikle ifade edilmiş, veri ve destek vektör gösterimi.

SVM başlangıçta doğrusal olarak ayrılabilen veri kümelerinde çalışabilir olarak tasarlanmışsa da, çekirdek işlevleri (*kernel functions*) ile aralarında doğrusal ilişkiler olmayan veriler üzerinde de uygulanabilmektedir. RBF-SVM (*radial basis function-svm*), Polynomial-SVM bu çekirdeklere örnek verilebilir. Bu özelliği ile birçok algoritmaya göre daha esnek uygulanabilirlik sağlamaktadır.

Tez kapsamında karşılaştırma için kullanılan tekil sınıflayıcılarda çekirdek işlevi doğrusal, maliyet değeri 0.10, hata payı parametresi 0.001 seçilmiştir.

3.5.3. Random Forests (RF)

Karar ağaçları (*decision trees*) sınıflama için sıkça kullanılan yaklaşımlardır. RF algoritması ise bünyesinde rastgele birçok karar ağacı oluşturan ve bunlar içinden en uygununu belirleyen bir takım (*ensemble*) algoritmadır. Bu “rastgelelik” hem öznelik seçme, hem de deneme/ eğitim veri kümelerini tanımlama aşamasında uygulandığı için ayrı bir genelleştirme (*generalization*) yaklaşımına gerek duymaz [73].

RF’de oluşturulan her bir karar ağacı, orijinal veri kümesinden elde edilmiş farklı örneklem (*bootstrap sample*) üzerine oluşturulur. Veri kümesinin bir kısmı deneme için

ayrılır ve her bir ağacın oluşturulması esnasında eğitim setine dahil edilmez. Genelde eğitim verisi örneklemin %66'sından, deneme verisi ise %34'ünden oluşmaktadır.

Karar ağaçlarının oluşturulması aşamasında, ağacın her bir düğümü belirlenirken; tamamı K kadar olan öznelik içinden rastgele seçilen k ($k < K$) öznelik kullanılır. Düğümde kullanılacak öznelikler *gini importance* denen bir ölçüm ile belirlenir. Değişken/Öznelik önemliliği (*variable importance*) değerlendirmesi yine bu aşamada elde edilen hesaplamalarla yapılır.

Bir örnek öznelikleri ile RF ile eğitilmiş sisteme verildiğinde, üretilen tüm karar ağaçlarından sınıflamaya sokulur. Sonuç kesikli veri için çoğunluk oylama, sürekli veri için ortalama ya da ağırlıklı ortalama ile belirlenebilir.

RF algoritmasının sahip olduğu bazı özellikler [73]:

- Birçok sınıflama algoritmasının aksine mikroarray veriler gibi büyük boyutlu ve az örnek olan veriler üzerinde güvenilir şekilde sonuç üretebilir.
- Büyük veri kümelerinde etkin olarak çalışır.
- Hangi değişkenlerin sınıflama için önemli olduğuna dair bilgi verir.
- Eksik veriyi tahmin etmek için etkili bir yöntemi vardır.
- Nüfusun dengesiz dağıldığı veri kümelerinde, hatayı azaltmak için yöntemler barındırmaktadır.
- Oluşturulan “ormanlar” başka veriler üzerinde kullanılmak için saklanabilir.
- Sahip olduğu özellikler genişletilerek etiketlenmemiş veri üzerinde, kümeleme ve aykırı değer tespitinde kullanılabilir.

Tez kapsamında karşılaştırma için kullanılan sınıflayıcılarda ağaç sayısı parametresi 100 seçilmiştir.

3.5.4. Naive Bayes

Naive Bayes algoritması, Bayes Teoremine bağlı olarak (3.16), koşullu olasılıklara dayanarak hesaplanan; karmaşık yapısı olmayan, tekrarlı parametre kestirim düzeneklerine ihtiyaç duymayan bir algoritmadır [71].

$$p(i|x) = \frac{p(x,i)}{p(x)} = \frac{p(x|i)p(i)}{\sum_{i=1}^C p(x|i)p(i)} \quad (3.16)$$

$i = 0,1$ şeklinde etiketlenmiş iki sınıflı bir veri olsun. Eğitim kümesindeki örneklerin başlangıç sınıf üyeliklerini kullanarak bir değer (*score*) bulunsun ve bu değerden büyük olanları 1 sınıfıyla, küçük olanları 0 sınıfıyla ilişkilendirilsin. Daha sonra belirlenen bir t eşik değeri bu değerler karşılaştırılarak sınıflama işlemi yapılır.

Naive Bayes’de bütün öznitelikler koşullu olarak birbirinden bağımsız varsayılır (3.17: x öznitelik vektörü, i sınıf etiketi, N öznitelik sayısı). Çoğu durumda böyle olmadığı halde sonuç model sınıflandırmayı başarılı yapmaktadır.

$$p(x|i) = \prod_n^N p(x_n|i) \quad (3.17)$$

3.5.5. Adaboost

Bütünleşik/takım öğrenme (*Ensemble learning*) bir problemi çözmek için birden fazla öğrenme yöntemini bir araya getiren yaklaşımlardır [73], [74]. Bütünleşik yöntemlerin genelleştirme yöntemleri çoğunlukla tekil yöntemlerinkinden daha iyi olmaktadır.

Freund ve Schapire [75] tarafından sunulan Adaboost algoritması en önemli bütünleşik algoritmalarından birisidir. Kolay gerçekleştirimi ve sağlam teorik altyapısı ile doğru tahmin yapan ve başarılı uygulamaları olan bir yaklaşımdır.

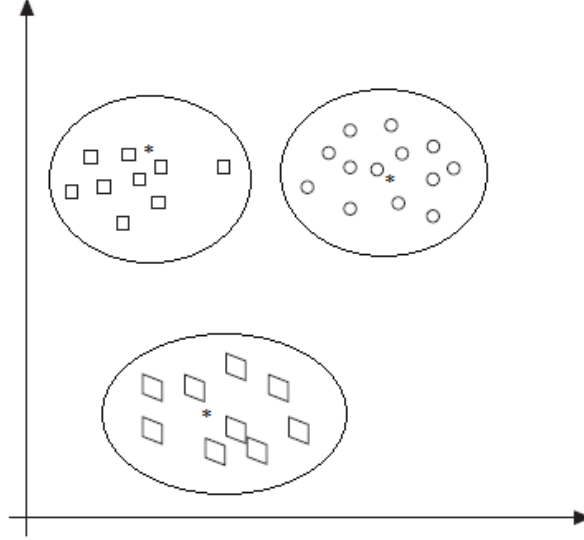
X örneklem uzayı ve $Y = \{-1, +1\}$ sınıf etiketleri kümesi olsun. Eğitim kümesi $\{(x_1, y_1), \dots, (x_m, y_m)\}$ ve t 'nci öğrenme adımında ağırlıkların (*weight*) dağılımları D_t olarak ifade edilsin. Adaboost algoritması öncelikle tüm eğitim örneklerine, (x_i, y_i) ($i \in \{1, \dots, m\}$), eşit ağırlık atar. Temel öğrenme algoritması; eğitim kümesi ve D_t 'den zayıf öğrenici (*weak learner*) $h_t : X \rightarrow Y$ elde eder. Bu aşamadan sonra hatalı sınıflanan örneklerin ağırlıkları artırılır ve D_{t+1} hesaplanır. Bu işlemler T adım kadar ağırlıklar güncellenerek devam eder. Sonuç model, T adet zayıf öğrenicinin aldığı ağırlıklı oy çoğunluğu değerlendirilerek oluşturulur.

Çok sınıflı problemleri çözmek için Ada-Boost.M1 algoritması sunulmuştur [75]. Aynı problemin çözümü için önerilen başka bir sürümü de AdaBoost.MH algoritmasıdır [76].

Tez kapsamında karşılaştırma için kullanılan Adaboost sınıflayıcı, Naive Bayes algoritması üzerinden artırma (*boosting*) yöntemi kullanılarak oluşturulmuştur.

3.5.6. K-En Yakın Komşuluk (*K Nearest Neighbours - kNN*)

kNN örneği, eğitim kümesinde en yakın k komşusunun sınıf değerleri ele alınarak sınıflar. Bu sınıflamanın sonucunu etkileyen üç temel kavram vardır: eğitim kümesindeki verilerin etiketleri, k değeri ve özniteliklerin yakınlığının belirlendiği uzaklık ölçü yöntemi.



Şekil 3.7 kNN için örnek gösterim.

k değerinin küçük seçilmesi sınıflama aşamasında verinin gürültüden (*noise*) etkilenmesine sebep olabilir. Büyük seçilirse örneğin ait olduğu sınıf dışından da birçok komşu değerlendirmeye katılabilir. Uzaklık ölçümü ise yapılan sınıflandırma türüne göre belirlenmesi gerekebilir. Örneğin belgeler sınıflanırken Euclidean uzaklığı kullanmak yerine *cosine* ölçümü kullanmak daha doğru olabilir [71], [77]. Özniteliklerin değer olarak birbirinden çok farklı olduğu (örn: boy ölçüsü, maaş) durumlarda uzaklığın hesaplanması esnasında büyük veriler baskın olabilir. Bu sorunun çözülmesi için normalleştirme yaklaşımları (örn: z-score) kullanılabilir.

Şekil 3.7’de iki boyutlu öznitelik uzayında (x-y) dağılan örneklem kümesinin, k değeri 3 olan kNN işlemi sonrası kümelenmiş hali gösterilmiştir. Her bir örnek, ilgili kümelerin merkezlerine uzaklıklarına göre kümeye üyeliklerine karar verilmektedir.

Tez kapsamında karşılaştırma için kullanılan sınıflayıcılarda komşuluk değeri 5, uzaklık ölçüm türü Öklid seçilmiştir.

3.5.7. Doğrusal Ayrım Analizi (*Linear Discriminant Analysis - LDA*)

Doğrusal ayrım analizi Fisher tarafından sunulmuş bir yöntemdir [78]. LDA, sınıfları en iyi ayıran doğrusal öznitelik kombinasyonlarını arama üzerine geliştirilmiştir. Bu özelliği ile

aynı zamanda boyut indirgeme yaklaşımı olarak da kullanılmaktadır. LDA, sınıflar arasında ayırım bilgisini kaybetmeyecek en uygun şekilde boyut indirgemeyi amaçlar.

Bu sınıflama yönteminde, bağımlı değişken (Y) grup/sınıf, bağımsız değişkenler (X) ise bu sınıfları tanımlayan öznitelikleri tanımlar. Bağımlı değişken her zaman kategori, bağımsız değişkenler ise sayısal veri olabilir.

Fisher'in önerisine göre, ortalamalar arasındaki farkın en yüksek değere yaklaştırılması sınıf içi dağınıklığı normalleştirerek sağlanır. Bu dağınıklık varyans olarak da ifade edilebilir. Fisher'in doğrusal ayırımı, kriter işlevini en yüksek yapan doğrusal $w^T x$ işlevi olarak tanımlanır. Sonuç olarak aynı sınıftan örneklerin birbirine en yakın olduğu, farklı sınıf ortalamalarının ise en uzak olduğu muhtemel yansıma bulunmaya çalışılır [79].

Tez kapsamında karşılaştırma için kullanılan sınıflayıcılarda karşılaştırma parametresi olarak "doğrusal sınırlar" ve gama değeri 0 olarak kullanılmıştır.

3.6. Sınıflama Algoritmasını Seçmek için Kullanılacak Yöntemler

Literatürde tanımlı birçok sınıflama algoritması bulunmaktadır ve bu algoritmaların temelleri istatistik, makine öğrenimi ve sinir ağları gibi farklı alanlardan gelmektedir [80]. *Bagging*, *boosting* gibi yaklaşımların bu algoritmalarla bütünleştirilmesiyle sayıları daha da artmaktadır. Bu yüzden hangi algoritmanın kullanılacağını belirlemek önemli bir problemdir [81].

Brachman [82]'a göre algoritmanın seçimi analizcinin tecrübesine ve araştırma alanına bağlı, deneysel bir süreçtir. Algoritmalar farklı veri kümeleri üzerinde, hatta aynı veri kümesinin farklı altkümeleri üzerinde farklı başarımlar elde edebilir. Buna rağmen algoritmaların değerlendirilmesi ve başarımı en yüksek olanın belirlenmesi mümkündür.

Brazdil ve Soares [81]'in çalışmalarında önerdikleri değerlendirme yöntemleri ele alınıp, algoritma seçiminde dayanak olarak kullanılmıştır. Bu değerlendirme yöntemleri Ortalama Sıralamalar (*Average Ranks (AR)*), Başarı Değer Oranları (*Success Rate Ratios (SRR)*) ve Önem Kazanımları (*Significant Wins (SW)*)'dır.

3.6.1 Ortalama Sıralamalar (*Average Ranks (AR)*)

Bu ölçüm Friedman'ın M istatistiğinden yola çıkılarak geliştirilmiş basit bir yaklaşımdır [83]. Algoritmalar, her bir veri kümesi için, ölçülen hata oranına (*error rate*) göre sıralanır ve buna göre derecelendirilir. r_j^i , j algoritmasının i veri kümesi üzerindeki derecesidir $-n$

toplam veri kümesi sayısı olmak üzere her algoritma için ortalama dereceler aşağıdaki eşitlik ile hesaplanır (3.18).

$$r_j = (\sum_i r_j^i)/n \quad (3.18)$$

Son değerlendirme ortalaması alınmış dereceler üzerinden yapılır.

3.6.2. Başarı Değer Oranları (*Success Rate Ratios (SRR)*)

Bu derecelendirme ölçümünde algoritmaların ikili olarak, karşılıklı başarımları değerlendirilir. Her bir j,k algoritma ikilisi için; ER_j^i , algoritma j'nin i veri kümesi üzerinde hesaplanan hata oranı olmak üzere;

$$SRR_{j,k}^i = (1 - ER_j^i) / (1 - ER_k^i) \quad (3.19)$$

(3.19) eşitliği hesaplanır. Bu durumda veri kümesinde daha başarılı sonuç üreten algoritma 2, algoritma 1'in üzerinde değer alacaktır.

Daha sonra iki başarı oranlarının ortalama değerleri hesaplanır. n veri kümesi sayısı, j,k algoritma ikilisi olmak üzere;

$$SRR_{j,k} = (\sum_i SRR_{j,k}^i)/n \quad (3.20)$$

(3.19) eşitliği hesaplanır. Bu değer j algoritmasının k algoritması üzerindeki genel avantaj/dezavantajı olarak adlandırılabilir.

Son olarak algoritma j'nin genel başarı değer oranı, m algoritması sayısı olmak üzere (3.20) şeklinde hesaplanır.

$$SRR_j = (\sum_k SRR_{j,k}) / (m - 1) \quad (3.21)$$

3.6.3. Önem Kazanımları (*Significant Wins (SW)*)

Bu yöntemde derecelendirme, iki yöntemin karşılıklı ikili hipotez test sonuçlarına dayanarak yapılır. Öncelikle iki algoritma arasındaki başarımlar farklılıklarının önemi test edilir. Bu işlem tüm veri kümeleri için yapılır (Brazdil ve Soares önemlilik oranı %5 olan ikili t testi kullanmıştır). Amaç algoritmalar arasındaki farkı bulurken, bu farkın önemini de değerlendirmeye alarak güvenilir sonuç elde etmektir.

$ER_j^i \ll ER_k^i$ olacak şekilde, eğer i veri kümesi üzerinde, algoritma j algoritma k'dan dikkate değer bir biçimde daha iyiye; j algoritması k üzerinde kazanmış olur. Bu kazanım

değerleri bir kazanım tablosunda bütünleştirilebilir. $W_{j,k}^i$, algoritma j 'nin k üzerindeki kazanımını göstermek üzere, (3.22)'de ifade eden değerleri alır.

$$W_{j,k}^i = \begin{cases} 1, & \text{Eğer } ER_j^i \ll ER_k^i \\ -1, & \text{Eğer } ER_k^i \ll ER_j^i \\ 0, & \text{Diğer durumlarda} \end{cases} \quad (3.22)$$

Tanıma göre $W_{j,k}^i = -W_{k,j}^i$ olmaktadır.

Daha sonra algoritmaların karşılıklı kazanımlarının olasılığı olan $pw_{j,k}$ hesaplanır. Bu değere j algoritmasının k 'dan önemli derecede başarılı olduğu veri kümesi sayısı toplam veri kümesi sayısı olan n 'ye bölünür.

Son olarak algoritmaların genel önem kazanım olasılıkları, m algoritma sayısı olacak şekilde (3.23) hesaplanır.

$$pw_j = (\sum_k pw_{j,k}) / (m - 1) \quad (3.23)$$

3.7. Değerlendiriciler Arası Uyumun Tespiti

LIDC veritabanı gibi çoklu değerlendiricinin olduğu verilerde, farklı değerlendiriciler arasında uyumun olup olmadığının ölçümünü yapmak için bazı yöntemler kullanılmaktadır. Bunlardan en bilinenlerine Krippendorff'un Alpha katsayısı [47], Cohen'in Kappa katsayısı [41], Scott'un Pi katsayısı [48] örnek olarak verilebilir. Bu tez kapsamında LIDC verisi üzerinde tüm değerlendiricilerin eşit düzeyde tecrübeye sahip olduğu varsayımı yapılmıştır. Bu varsayım literatürdeki tüm çalışmalarda yapılmaktadır. LIDC'de tüm örnekleri dört farklı radyolog değerlendirmiştir, fakat her örnek aynı radyolog kümesi ile değerlendirilmemiştir (buna dair bir bilgi veritabanında bulunmamaktadır –doktorlara ait bilgi anonim olarak işaretlenmiştir). Çalışmada farklı okuma oturumlarından yakın başarımlar elde edip etmediğimizin ölçümünde Krippendorff'un Alpha katsayısı kullanılmıştır.

3.7.1. Krippendorff'un Alpha Katsayısı

Krippendorff alfa (α), gözlemciler veya değerlendiriciler arasındaki uyumun ölçümü için geliştirilmiş bir güvenilirlik katsayısıdır [47]. Literatürde en yaygın olarak kullanılan değerlendiriciler-arası güvenilirlik yöntemidir [84]. Özellikle ikiden fazla gözlemcinin varlığında güvenilirliğin zorlaştığına dair bir inanç mevcuttur; ancak Krippendorff alfa ikili karşılaştırmalar yaptığından, gözlemci sayısından bağımsız bir ölçüm sağlamak ve α 'nın iyi bir güvenilirlik ölçümünün tüm özelliklerini karşılayan standart olarak en uygun ölçüm

olduğu ileri sürülmektedir [85]. α , içerik analizinden ortaya çıkmakla beraber; daha geniş olarak uygulanabilir [47]. α 'nın genel formu (3.24)'de verilmiştir.

$$\alpha = 1 - \frac{D_o}{D_e} \quad (3.24)$$

D_o , gözlenen uyumsuzluk, D_e ise birimlerin özelliklerinden değil şans eseri ortaya çıkan; yani beklenen uyumsuzluk olarak ifade edilebilir. Uyum ölçütü, gözlenen uyumsuzluğun, beklenen uyumsuzluğa bölünmesi ile elde edilir. $D_o=0$ ve dolayısıyla $\alpha=1$ olduğunda gözlemciler arası uyum mükemmeldir. Ancak gözlemciler şans eseri uyumlu ise, $D_o=1$ ve $\alpha=0$ olur. Bu durumda güvenilirlik söz konusu değildir; yani kategoriler ya da ölçek puanları istatistiksel olarak ilişkili değildir [47], [85]. α 'nın aralığı ise şöyledir: $1 \geq \alpha \geq 0$.

Krippendorff Alfa, araştırmacılara aynı güvenilirlik standardı ile çeşitli verileri değerlendirme imkânı sağlar. α ; sadece iki değil daha fazla sayıda gözlemcinin varlığında, herhangi bir sayıda kategori, ölçek değeri veya ölçümün varlığında, her tür ölçek düzeyinde (sınıflayıcı, sıralı, aralıklı veya oranlı), tamamlanmamış veya kayıp verinin varlığında ve hem büyük hem de küçük örneklerde uygulanabilir [47].

3.8. İki Algoritma Arasındaki Farkın Test Edilmesi

İki algoritma arasındaki hata oranı arasındaki anlamlılığın test edilmesi için Snedecor ve Cochran [86] tarafından basit bir yöntem önerilmiştir [87].

A algoritması tarafından yanlış sınıflanan örnek olasılığı p_A , B algoritması tarafından yanlış sınıflanan örnek olasılığı p_B olsun (3.25, 3.26) (n_{00} : Her iki algoritma tarafından yanlış sınıflanan örnek sayısı, n_{01} : Sadece A algoritması tarafından yanlış sınıflanan örnek sayısı, n_{10} : Sadece B algoritması tarafından yanlış sınıflanan örnek sayısı, n : toplam örnek sayısı).

$$p_A = (n_{00} + n_{01})/n \quad (3.25)$$

$$p_B = (n_{00} + n_{10})/n \quad (3.26)$$

p iki hata olasılığının ortalaması olmak üzere (3.27), yokluk hipotezi altında, standart hata se eşitlik (3.28) gösterildiği şekilde hesaplanır.

$$p = (p_A + p_B)/2 \quad (3.27)$$

$$se = \sqrt{\frac{2p(1-p)}{n}} \quad (3.28)$$

Çözümleme sonucunda, yaklaşık olarak standart normal dağılıma sahip z istatistiği elde edilir (3.29). Eğer $|z| > Z_{0.975} = 1.96$ ise yokluk hipotezi reddedilir ($p < 0.05$).

$$z = \frac{p_A - p_B}{\sqrt{\frac{2p(1-p)}{n}}} \quad (3.29)$$

3.9. Yöntem Başarımlarının Ölçülmesi

Tez çalışmaları kapsamında önerilen yöntemlerin ve karşılaştırılan sınıflayıcıların başarımlarının ölçülmesinde sınıflama başarımları (*Classification accuracy*, SB), duyarlılık (*sensitivity*, Duy.) ve seçicilik (*specificity*, Seç.) ölçütleri kullanılmıştır.

Sınıflama başarımları, sınıflayıcıların bir örneklem içindeki doğru sınıfladığı örnek sayısının tüm örnek sayısına bölüldüğü ölçümdür. TP, doğru tanımlanmış örnek sayısı; TN, doğru reddedilmiş örnek sayısı; FP, hatalı tanımlanmış örnek sayısı; FN, yanlış reddedilmiş örnek sayısı olmak üzere sınıflama başarımları Eşitlik (3.30)'daki gibi ifade edilir.

$$SB = (TP + TN) / (TP + TN + FP + FN) \quad (3.30)$$

Duyarlılık, doğru tanımlanmış örnek oranını belirtmektedir (*true positive rate*). Eşitlik (3.31)'deki gibi ifade edilir.

$$Duy. = (TP) / (TP + FN) \quad (3.31)$$

Seçicilik, doğru reddedilmiş örnek oranını belirtmektedir (*true negative rate*). Eşitlik (3.32)'deki gibi ifade edilir.

$$Seç. = (TN) / (FP + TN) \quad (3.32)$$

4. LIDC VERİ KÜMESİ VE ÇÖZÜMLEMESİ

4.1. LIDC Veritabanı

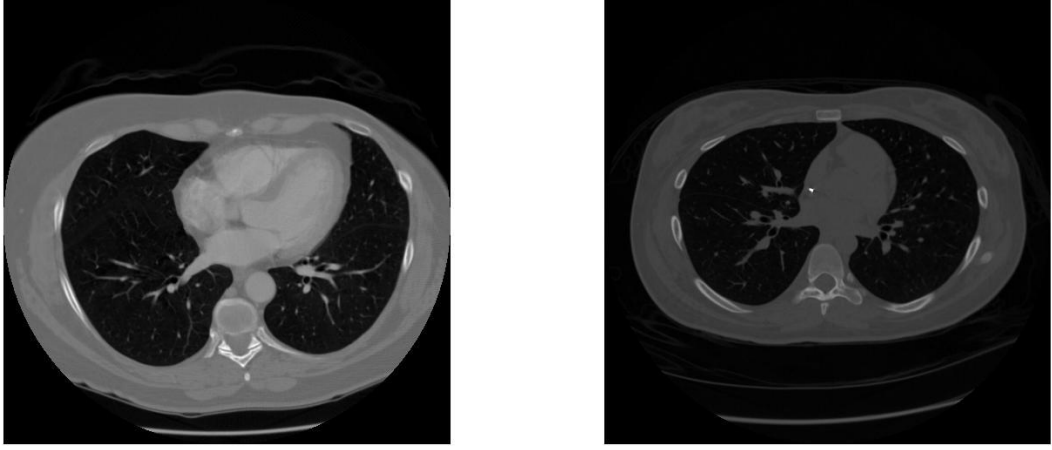
Bilgisayarlı tomografi sistemlerinin gelişmesinin, bu görüntüleme tekniğinin sıkça uygulanması üzerinde çok önemli etkileri olmuştur. Hastanın hareketlerinden daha az etkilenmesini sağlayacak, daha ayrıntılı çekim için sık aralıklarla çekim yapacak sistemler görüntü kalitesini arttırdığı gibi; önemli miktarda verinin oluşmasına da sebep olmuşlardır. Bu gelişmeler ile radyologların üzerine düşen yük de artmaya başlamıştır.

Bilgisayar destekli tanı (*computer aided diagnosis - CAD*) sistemleri, birçok tıp alanında olduğu gibi bu alanda da radyologların işleri kolaylaştırmak için yaygın kullanılmaktadır. Geliştirilen sistemler için en büyük sıkıntı, doğru şekilde etiketlenmiş ve genelleştirme için kullanılabilir verinin elde edilmesidir.

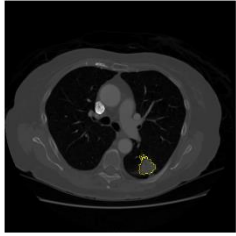
2001 yılının Nisan ayında, National Cancer Institute akciğer CT görüntülerinden oluşan, İnternet üzerinden ulaşılabilecek bir veri ambarı kurulması için “*Lung Image Database Resource for Imaging Research*” başlığı altında bir talep oluşturulmuştur. Bu amaçla, Amerika Birleşik Devletlerinde etkinlik gösteren beş akademik enstitünün çabası ile (Cornell University, University of California, University of Chicago, University of Iowa, University of Michigan), bilgisayar destekli tanı sistemlerinin geliştirilmesi için kullanılacak, üzerinde ortak görüş sağlanmış bir görüntü veri tabanı olan *Lung Image Database Consortium*'u (LIDC) geliştirmiştir.

Görüntüler toplanmadan önce, bilimsel olarak güvenilir bir veritabanı oluşturmak için çalışmalar yapılması gerekli olmuştur. Tomografi görüntülerinin veritabanına uygunluğunun denetlenmesi, nodüllerin değerlendirilmesi için uygun tanımların yapılması, veritabanının genişlemesi için uygun modelin seçilmesi, kullanıcılara yol gösterecek istatistiksel bir çatının oluşturulması gibi hazırlık aşamalarından sonra verilerin toplanıp değerlendirilmesi aşamasına geçilmiştir.

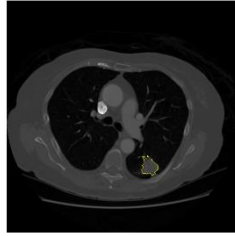
Toplanılan DICOM verilerinde akciğer görüntülerinin yanı sıra; tomografi cihazının teknik parametreleri, modeli ve yeniden yapılandırma algoritması; yaş, cinsiyet, sigara geçmişi ve ölçümü yapılmış herhangi klinik tanı bilgisi de XML yapısında eklenmiştir.



Şekil 4.1 Örnek LIDC tomografi görüntüleri.



Reading Session 1	
<i>Sublety:</i>	5
<i>Internal Str:</i>	1
<i>Calcification:</i>	6
<i>Sphericity:</i>	3
<i>Margin:</i>	3
<i>Lobulation:</i>	3
<i>Spiculation:</i>	4
<i>Texture:</i>	5
<i>Malignancy:</i>	5



Reading Session 2	
<i>Sublety:</i>	5
<i>Internal Str:</i>	1
<i>Calcification:</i>	6
<i>Sphericity:</i>	3
<i>Margin:</i>	4
<i>Lobulation:</i>	4
<i>Spiculation:</i>	5
<i>Texture:</i>	5
<i>Malignancy:</i>	5



Reading Session 3	
<i>Sublety:</i>	5
<i>Internal Str:</i>	1
<i>Calcification:</i>	6
<i>Sphericity:</i>	5
<i>Margin:</i>	4
<i>Lobulation:</i>	1
<i>Spiculation:</i>	5
<i>Texture:</i>	4
<i>Malignancy:</i>	4

Şekil 4.2 LIDC örneğinden bir CT kesiti. Aynı nodül üç farklı radyolog ratafından değerlendirilmiştir.

Görüntü kalitesi tanımlanması zor bir özneliktir. Hasta hareketinden ya da tarayıcıdan kaynaklı hatalı görüntüler oluşabilmektedir. Bu veri tabanında da benzer kusurlu görüntüler bulunsa da, bunlar veritabanından çıkarılmamıştır; fakat görüntü kalitesi hakkında bilgi vermek amacıyla “*marginal*” ya da “*unacceptable*” bilgisi ile kaydedilmiştir. Veri kümesine literatürde tanımlanmış nodül çapı olan 30 mm.’nin üstünde nodül tanımlanmamıştır. En küçük yarıçap olarak ise 3 mm. sınırı koyulmuştur.

Görüntülerin değerlendirilmesi için dört farklı enstitüden farklı radyologlar görevlendirilmiştir. Verileri değerlendirme aşaması “*blinded*” (her radyolog ilk aşamada tek başına değerlendirmiştir) ve “*unblinded*” (her radyolog diğer radyologların değerlendirmelerini de göz önüne alarak değerlendirmiştir) olarak ikiye ayrılmıştır. İki aşamada da aynı adımlar gerçekleşir: bir radyolog bir taramadaki tüm nodülleri olabildiğince tespit edip radyolojik yapısı hakkında bilgi vermelidir. Boyutu 3mm’den küçük olduğu halde nodül olabileceğine dair kuşku duyulduğunda bu bölge konumu işaretlenmektedir. “*Blinded*” okuma aşaması tamamlandığında, bütün radyologların bilgisi bir araya getirilip, her bir radyologa dağıtılmıştır. Böylece her bir radyolog, diğer radyologların değerlendirmesini de göz önünde bulundurarak kararlarını düzenleyebilmiştir. Hasta ve tomografi çekimlerine ait bilgilerin bulunduğu XML kütüğünde sadece “*unblinded*” okuma aşaması eklenmiştir.

Veri tabanı farklı CAD yapısı tanımlayan araştırmacılar için kaynak olabilecek özneliktir. Eğer araştırmacı sıralı görüntüler üzerinde nodül gelişimini denetleyen bir çalışma yapacaksa, aynı hastaya ait, belli bir süre aralıkla çekilmiş birden fazla tomografi taramasını kullanabilecektir. Eğer tek bir tarama üzerinde çalışma yapacaksa mevcut patolojik bilgiyi kullanabilecektir.

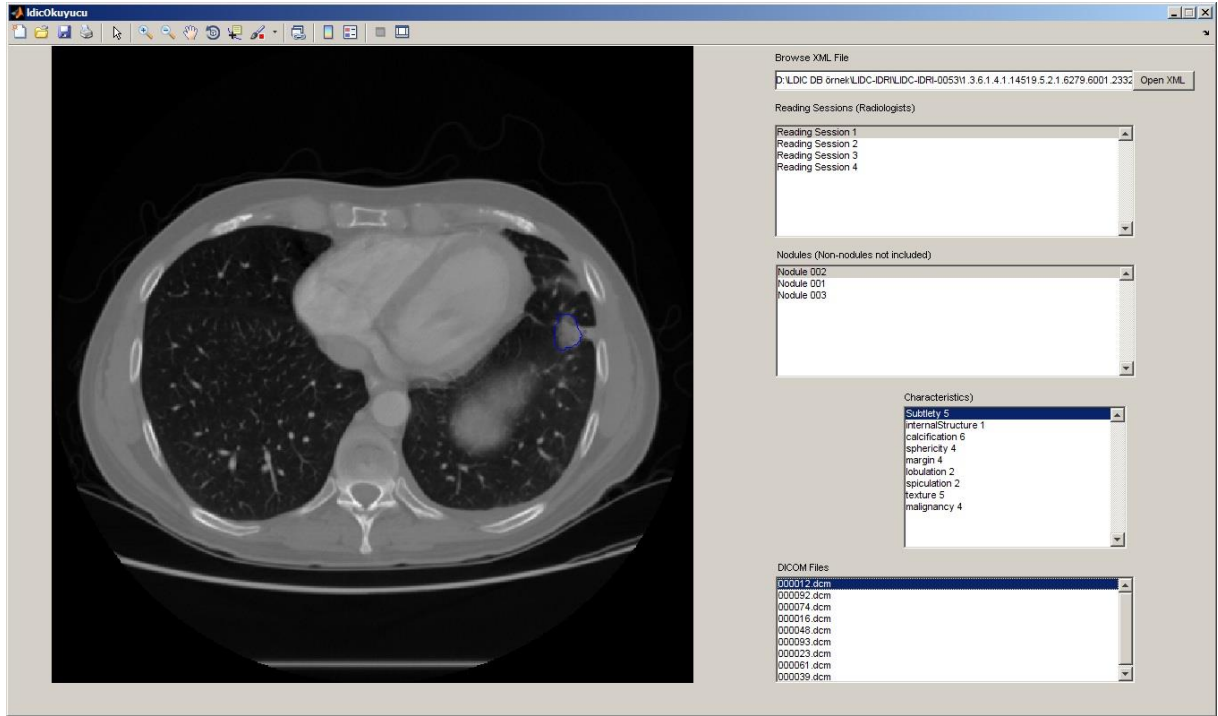
Toplam 1010 hasta için tarama sonucu veritabanının son sürümünde bulunmaktadır.

Çizelge 4.1 Radyografik Tanımlayıcı tanımları ve dereceleri.

Radyografik Tanımlayıcı	Tanım	Dereceler
Calcification	Nodüldeki kireçlenme yapılanması.	1. <i>Popcorn</i> 2. <i>Laminated</i> 3. <i>Solid</i> 4. <i>Non-central</i> 5. <i>Central</i> 6. <i>Absent</i>
Internal Structure	Nodülün beklenen iç yapısı.	1. <i>Soft Tissue</i> 2. <i>Fluid</i> 3. <i>Fat</i> 4. <i>Air</i>
Lobulation	Lobular şeklin sınır ile belirginliğinin olup olmaması.	1. <i>Marked</i> 2. . 3. . 4. . 5. <i>None</i>
Malignancy	Nodülün habislik oranı.	1. <i>Highly Unlikely</i> 2. <i>Moderately Unlikely</i> 3. <i>Indeterminate</i> 4. <i>Moderately Suspicious</i> 5. <i>Highly Suspicious</i>
Margin	Sınırların belirginliği.	1. <i>Poorly Defined</i> 2. . 3. . 4. . 5. <i>Sharp</i>
Sphericity	Yuvarlaklık bağlanmında nodülün boyutsal şekli.	1. <i>Linear</i> 2. . 3. <i>Ovoid</i> 4. . 5. <i>Round</i>
Spiculation	Nodüldeki sınırlarındaki sivri yapılaşmanın derecesi.	1. <i>Marked</i> 2. . 3. . 4. . 5. <i>None</i>
Sublety	Nodül ve civarındaki yapılar arasındaki zıtlık.	1. <i>Extremely Subtle</i> 2. <i>Moderately Subtle</i> 3. <i>Fairly Subtle</i> 4. <i>Moderately Obvious</i> 5. <i>Obvious</i>
Texture	Nodülün iç yoğunluğu.	1. <i>Non-Solid</i> 2. . 3. <i>Part Solid</i> 4. . 5. <i>Solid</i>

4.2. LIDC Okuyucu

Bu kesimde LIDC veri tabanı [10] bilgilerini okumak ve göstermek için tasarlanan araç anlatılacaktır. Araç, MATLAB ortamı üzerinde geliştirilmiştir.



Şekil 4.3 LIDC Okuyucu arayüzü

Şekil 4.3’de arayüzü verilen aracın kullanımı ve bileşenleri hakkında bilgi:

- İncelenecek hastaya ait XML kütüğü seçimi yapılır.
- Seçim yapıldıktan sonra ilk listeye dört radyoloğa ait oturum numaraları listelenir.
- Oturum seçildikten sonra ikinci listede radyolog tarafından bu hastada tespit edilmiş nodüllerin kimlik numaraları listelenir.
- Bu listeden nodül seçildikten sonra eğer nodül sınırları belirlenmiş bir nodül ise 3. listeye nodül yapısı hakkında radyologun girdiği değerlendirme sonuçları gelir. 4. listede ise bu nodülün tespit edildiği DICOM görüntü kütüklerinin listesi eklenir.
- Listedeki DICOM kütüğü seçildikten sonra ekranın sol tarafında ayrılan kesimde ilgili görüntü gösterilir ve nodülün etrafını değerlendiren radyologun belirttiği koordinatlar ile çizilir.

4.3. Veri Dağılımı ve Çalışmada Kullanılacak Veri Kümelerinin Oluşturulması

LIDC, farklı enstitülerden doktorların aynı hasta kümesi üzerinde yaptıkları çalışmalar üzerinde yaptıkları değerlendirme sonuçlarından oluşmaktadır. LIDC veritabanında radyologların yaptıkları değerlendirmelerde anlaşmaları beklenmediği için ve biyopsi

sonuçları olmadığı için, bir nodül için yapılmış değerlendirme sonuçları uyumsuzluk içerebilir. Bu nedenle kesin referans (*ground truth*) ya da altın oran (*golden ratio*) denilen veri doğrudan elde edilememektedir. Buna rağmen LIDC veri ambarı, aynı hastanın aynı nodülü üzerinde birden fazla değerlendirmenin yapılmış olması Zinovev ve diğerlerinin [17] belirttiği üzere bazı varsayımların yapılması ve farklı yöntemlerin uygulanması imkanlarını sağlamaktadır.

Çalışmalarda kullanılacak kesin referansın elde edilebilmesi için bazı varsayımların yapılması gerekmektedir. Bu çalışmada çoğunluk oylama yöntemi kullanarak, kesin referans olarak kullanılacak veri elde edilmiştir. Çoğunluk oylama yönteminde tüm değerlendiricilerin konu üzerinde aynı uzmanlık derecesi olduğu varsayılır. İkili olarak işaretlenmiş bir veride, R değerlendirici için genel ifadesi Eşitlik 4.1'deki gibi verilebilir. y_i , i . örnek için sınıf etiketi ve y_i^j , j . değerlendiricinin i . örnek için değerlendirme sonucudur.

$$y_i = \begin{cases} 1, & \left(\frac{1}{R}\right) \sum_{j=1}^R y_i^j > 0.5 \\ 0, & \left(\frac{1}{R}\right) \sum_{j=1}^R y_i^j < 0.5 \end{cases} \quad (4.1)$$

LIDC'nin wiki sayfasından [88] edinilen bilgiler ışığında bir nodül üzerinde, bir radyografik tanımlayıcı için en az üç doktor tarafından anlaşma sağlanmışsa (3/3, 3/4, 4/4) bu bilginin kesin referans olduğu kabul edilip veri kümelerine eklenmiştir.

Çoğunluk oylama ile her bir radyografik tanımlayıcı için ayrı veri kümeleri oluşturulmuştur. Bir nodül örneğini veri kümelerine eklemek için en az üç radyolog tarafından anlaşmaya varılması beklenmiştir. Bu durumda elimizde her bir tanımlayıcı için farklı sayıda örnek sayısına sahip veri kümeleri bulunmaktadır. Çizelge 4.2'de, dört ayrı radyolog tarafından değerlendirilmiş bir nodülün sonuçları verilmiştir. Eğer tanımlayıcı üzerinde bir anlaşmaya varılmışsa, örnek veri kümesine dahil edilmiştir. Örneğin, çizelgedeki örneklerin hepsi *sphericity* ve *lobulation* kümeleri hariç, tüm verikümelerine eklenmiştir. Bu yöntemle oluşturulan veri kümelerine ilişkin dağılım histogramları Şekil 4.4'de verilmiştir. Radyografik tanımlayıcıların histogramlarında görüldüğü gibi bazı tanımlayıcıların sınıfları aşırı dengesiz dağılmıştır.

Çizelge 4.2 Bir nodül üzerinde dört radyoloğun değerlendirmeleri. Her sütundaki, kalın ile işaretlenmiş örnekler ilgili tanımlayıcının veri kümesine eklenmiştir.

	Subtlety	Calcification	Sphericity	Margin	Lobulation	Spiculation	Texture	Malignancy
Rad. A	5	6	3	3	3	5	5	5
Rad. B	5	6	3	4	4	5	5	5
Rad. C	5	6	5	4	1	5	4	4
Rad. D	4	3	4	4	3	5	5	5

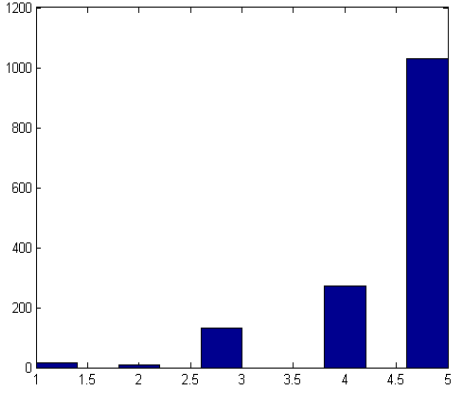
LIDC nodül veri kümesinin genel tanımı S_{LIDC} Eşitlik 4.2'de verilmiştir [89]. Eşitlikte, x_i , i . görüntü niteliği, m toplam öznelik sayısı, y_j , j . radyografik tanımlayıcının derecesi (sınıf değeri), c radyografik tanımlayıcı sayısı ve n örnek sayısıdır. Oluşturulan radyografik tanımlayıcı verikümelerinin histogramları Şekil 4.4'de gösterilmiştir.

$$S_{LIDC} = \{x_1, \dots, x_m, y_1, \dots, y_c\}^{1..n} \in \mathbb{R}^m \quad (4.2)$$

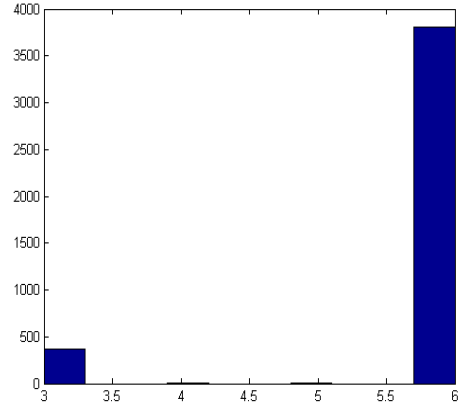
Veri kümesinde, 1010 vaka için 2635 nodül eşleşmesi (ortak değerlendirmesi) bulunmaktadır. Çoğunluk oylama sonrasında habislik veri kümesinde 438 biricik nodül eşleşmesi kalmıştır (derece 1:135, derece 2:29, derece 3:174, derece 4:32, derece 5:68). Tekrarlı değerlendirmelerle örnek sayısı toplamda 1304'tür (derece 1:445, derece 2:79, derece 3:490, derece 4:87, derece 5:203). Nodüllerin boyutları için bazı temel istatistikler Çizelge 4.3'te verilmiştir. İstatistikler LIDC wiki sayfasındaki bilgilerden elde edilmiştir [88].

Çizelge 4.3 LIDC veritabanında paylaşılan nodüllere ait bazı temel istatistikler.

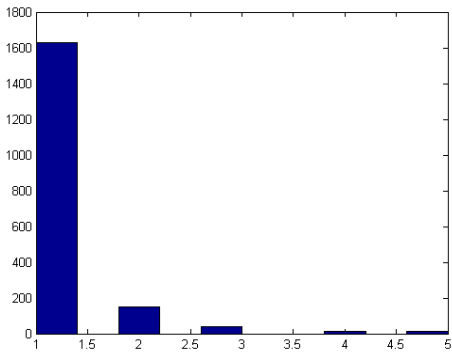
İstatistik	Sayı	Birim
Vaka sayısı – (aynı hastaya ait birden fazla vaka olabilir.)	1010	adet
Radyografik tanımlayıcı sayısı	9	adet
Ortalama nodül hacmi	618,90	mm3
Nodül hacmi standart sapma	3650,65	
En büyük nodül hacmi	27067,00	mm3
En küçük nodül hacmi	4,39	mm3



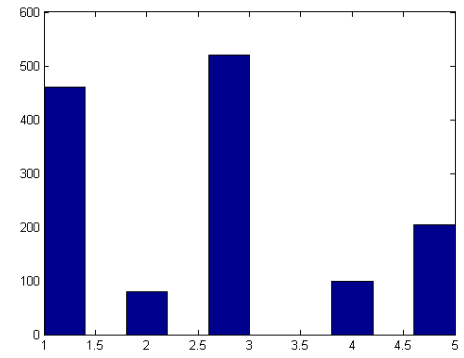
(a) *Subtlety*



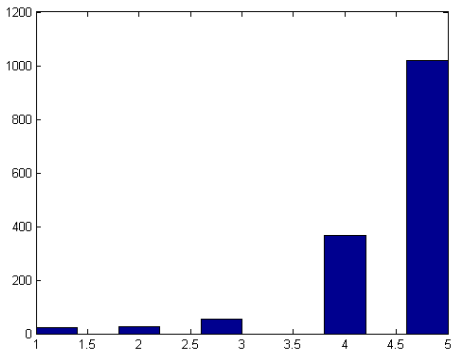
(b) *Calcification*



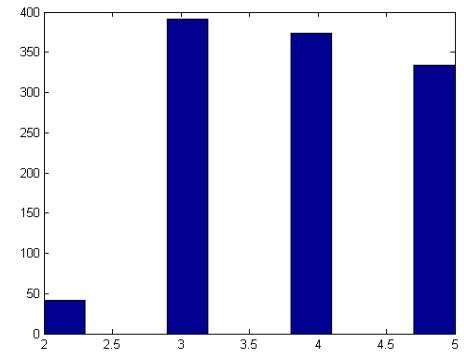
(c) *Lobulation*



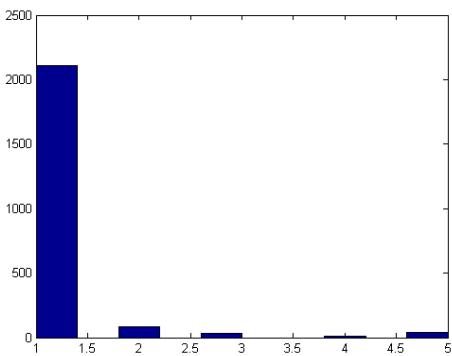
(d) *Malignancy*



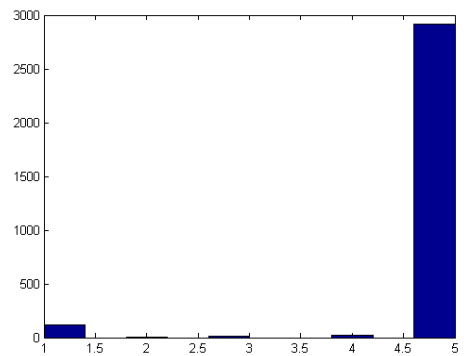
(e) *Sphericity*



(f) *Spiculation*



(g) *Texture*



(h) *Margin*

Şekil 4.4 (a-h) Radyografik tanımlayıcılarına ait veri kümelerinin histogramları.

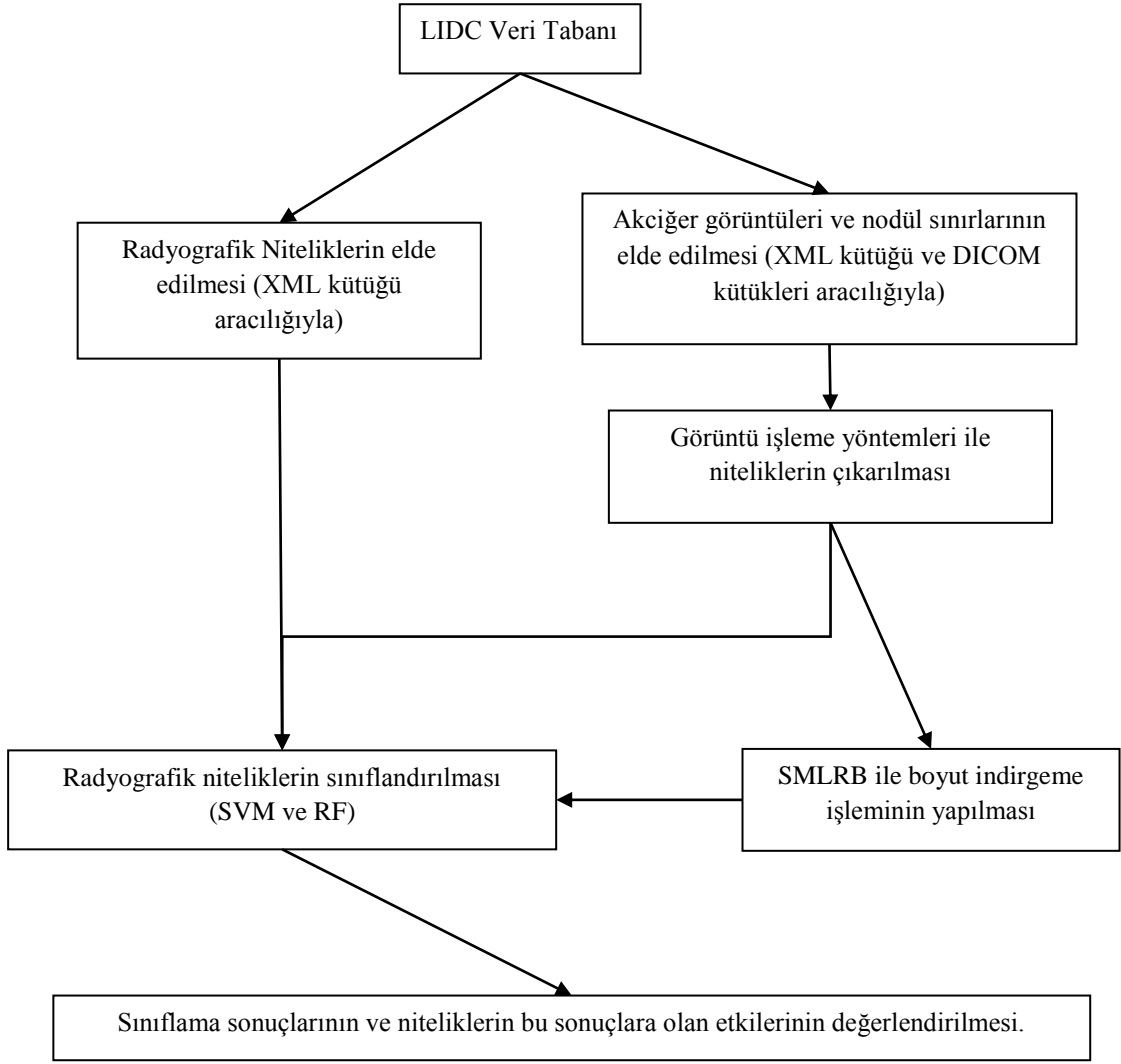
4.4. Farklı Okuma Oturumları (*Reading Session*) Üzerinden Sınıflama

Tez kapsamında LIDC veritabanından sağlanan radyolog değerlendirmelerinin her birini aynı tecrübeye ve aynı derecede bilgi sahibi olarak varsayılmıştır. Literatürdeki çalışmalar da bu şekilde varsayımlar yapmışlardır / yapmak durumunda kalmışlardır. Bu durumun geçerliliğini test etmek için aşağıdaki deney uygulanmıştır.

Bu kesimde, her bir radyografik tanımlayıcı ayrı ayrı sınıflandırılmıştır. Veride her bir vaka için dört okuma oturumu (*reading session-R1,R2,R3,R4*) sonuçları gruplanarak sınıflama başarımları incelenecektir. Buradaki amaç veride farklı ensitülerden farklı okuyucuların değerlendirmeleri arasında benzerlik ya da ilişkinin olup olmadığının incelenmesi ve rastgele seçim sayılabilecek bu işlemin sonuçlarının tutarlılığının denetlenmesidir.

Deney için radyolog değerlendirmeleri R1-R2-R3-R4 başlıkları altında, LIDC veri tabanında sağlanan xml kütüğündeki sırasıyla (rastgele olarak ifade edilebilir) gruplanmıştır. R1-R4 altında gruplanan değerlendirmelerin hepsinin aynı enstitü ya da aynı radyolog tarafından yapıldığı kesin değildir.

Bu deneyde Çizelge 3.1’de verilen özniteliklerden ilk 54 tanesi kullanılmıştır. Sınıflama için SVM ve RF yöntemleri kullanılmıştır. Boyut indirgeme yöntemi olarak SMLRB (*Sparse Multinomial Logistic Regression via Bayesian L1 Regularisation*) [90] yöntemi uygulanmıştır. Genelleştirme yöntemi olarak 10 parçalı çapraz geçerlilik (*10 fold cross validation*) kullanılmıştır. Deneyde farklı grupların başarımları yanında, boyut indirgeme sonucunda başarımın nasıl değiştiğine dair bilgi alınması da amaçlanmıştır. Yöntemin genel akış çizeneği Şekil 4.5’te verilmiştir.



Şekil 4.5. Yöntemin genel akış çizeneği.

Çizelge 4.4 - Çizelge 4.11’da, R1-R4 için alınan deney sonuçları gösterilmiştir. Şekil 4.6 ve Şekil 4.7’te sınıflama başarımlarının grafiksel gösterimi verilmiştir. RF ve SVM yöntemleri ayrı ayrı değerlendirildiğinde, boyut indirgeme işleminin sınıflama başarımını olumsuz yönde etkilemediği görülmüştür. R1-R4 arasındaki sınıflama sonuçlarına bakıldığında, tüm radyografik tanımlayıcılar arasında yüksek oranda farklılık bulunmadığı gözlemlenmiştir.

Çizelge 4.4 R1 okuması için Random Forest sınıflama algoritması ile elde edilen sonuçlar.

Radyografik Tanımlayıcı	SB	AUC
	<i>Tüm Öznitelikler ile</i>	
Calcification	0.9300	0.8636
Lobulation	0.6294	0.6974
Margin	0.5074	0.7007
Sphericity	0.4205	0.6831
Spiculation	0.7152	0.7526
Sublety	0.4158	0.7082
Texture	0.7762	0.7478
	<i>Boyut indirgeme ile</i>	
Calcification	0.9266	0.8618
Lobulation	0.6361	0.5778
Margin	0.4892	0.6910
Sphericity	0.3751	0.5912
Spiculation	0.7130	0.5977
Sublety	0.3717	0.6211
Texture	0.7875	0.7677

Çizelge 4.5 R1 okuması için SVM sınıflama algoritması ile elde edilen sonuçlar.

Radyografik Tanımlayıcı	SB	AUC
	<i>Tüm Öznitelikler ile</i>	
Calcification	0.9266	0.8682
Lobulation	0.6351	0.7175
Margin	0.5288	0.7386
Sphericity	0.4137	0.6834
Spiculation	0.7300	0.7528
Sublety	0.4159	0.7028
Texture	0.7932	0.7606
	<i>Boyut indirgeme ile</i>	
Calcification	0.9175	0.8685
Lobulation	0.6373	0.5925
Margin	0.4847	0.6899
Sphericity	0.3673	0.5654
Spiculation	0.7153	0.5961
Sublety	0.3944	0.6329
Texture	0.7932	0.7956

Çizelge 4.6 R2 okuması için Random Forest sınıflama algoritması ile elde edilen sonuçlar.

Radyografik Tanımlayıcı	SB	AUC
	<i>Tüm Öznitelikler ile</i>	
Calcification	0.9490	0.9310
Lobulation	0.4895	0.6685
Margin	0.4425	0.7370
Sphericity	0.4665	0.7222
Spiculation	0.6004	0.6840
Sublety	0.4036	0.7212
Texture	0.7912	0.8045
	<i>Boyut indirgeme ile</i>	
Calcification	0.9520	0.9277
Lobulation	0.5115	0.5935
Margin	0.4475	0.7428
Sphericity	0.3477	0.5294
Spiculation	0.6174	0.6095
Sublety	0.4036	0.6790
Texture	0.8012	0.8351

Çizelge 4.7 R2 okuması için SVM sınıflama algoritması ile elde edilen sonuçlar.

Radyografik Tanımlayıcı	SB	AUC
	<i>Tüm Öznitelikler ile</i>	
Calcification	0.9560	0.9442
Lobulation	0.5225	0.6723
Margin	0.4375	0.7563
Sphericity	0.4695	0.7325
Spiculation	0.6134	0.6799
Sublety	0.3986	0.7336
Texture	0.7952	0.8312
	<i>Boyut indirgeme ile</i>	
Calcification	0.9461	0.9371
Lobulation	0.5145	0.5952
Margin	0.4725	0.7669
Sphericity	0.3536	0.5385
Spiculation	0.6174	0.6146
Sublety	0.4116	0.7155
Texture	0.8042	0.8377

Çizelge 4.8 R3 okuması için Random Forest sınıflama algoritması ile elde edilen sonuçlar.

Radyografik Tanımlayıcı	SB	AUC
	<i>Tüm Öznitelikler ile</i>	
Calcification	0.9432	0.9272
Lobulation	0.6045	0.7187
Margin	0.4300	0.7188
Sphericity	0.4200	0.6667
Spiculation	0.6927	0.7453
Sublety	0.4341	0.7657
Texture	0.6937	0.7281
	<i>Boyut indirgeme ile</i>	
Calcification	0.9493	0.9227
Lobulation	0.6146	0.5928
Margin	0.4472	0.7397
Sphericity	0.4301	0.6637
Spiculation	0.6887	0.6020
Sublety	0.3540	0.6707
Texture	0.7140	0.7483

Çizelge 4.9 R3 okuması için SVM sınıflama algoritması ile elde edilen sonuçlar.

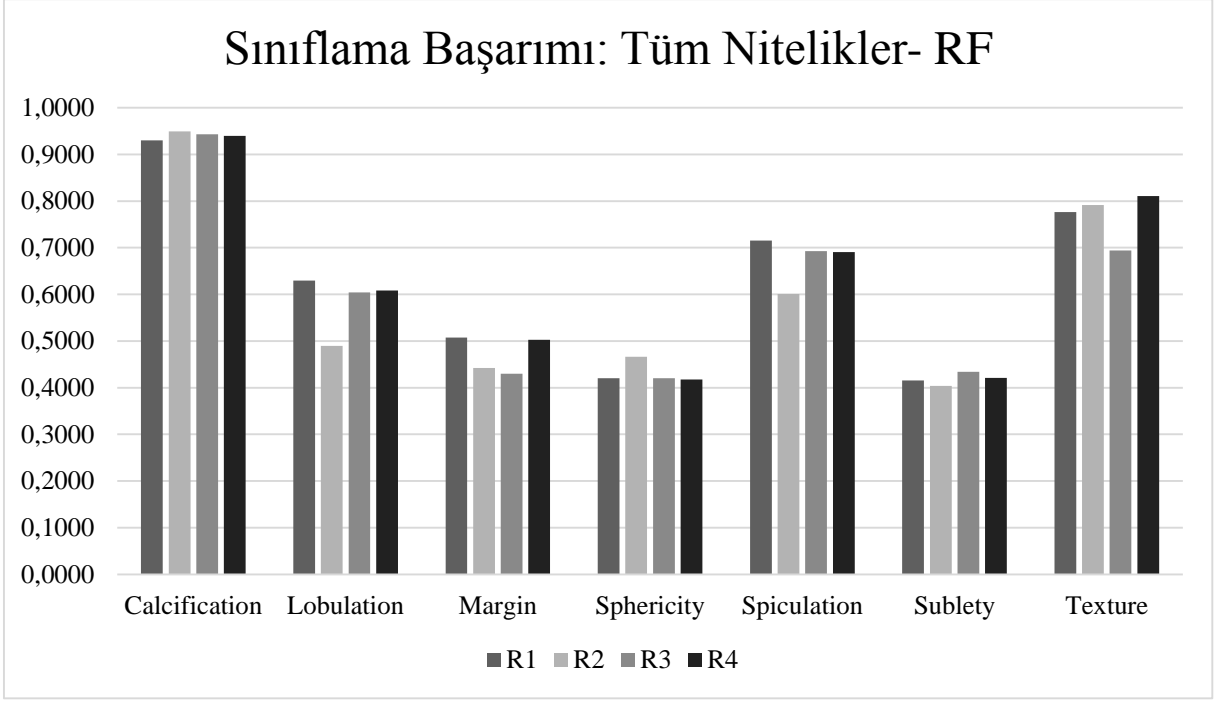
Radyografik Tanımlayıcı	SB	AUC
	<i>Tüm Öznitelikler ile</i>	
Calcification	0.9514	0.9472
Lobulation	0.6247	0.7459
Margin	0.4543	0.7491
Sphericity	0.4616	0.7140
Spiculation	0.6998	0.7440
Sublety	0.4310	0.7625
Texture	0.7129	0.7555
	<i>Boyut indirgeme ile</i>	
Calcification	0.9422	0.9454
Lobulation	0.6126	0.5813
Margin	0.4189	0.6852
Sphericity	0.4575	0.7160
Spiculation	0.6937	0.5779
Sublety	0.3763	0.6991
Texture	0.7241	0.7417

Çizelge 4.10 R4 okuması için Random Forest sınıflama algoritması ile elde edilen sonuçlar.

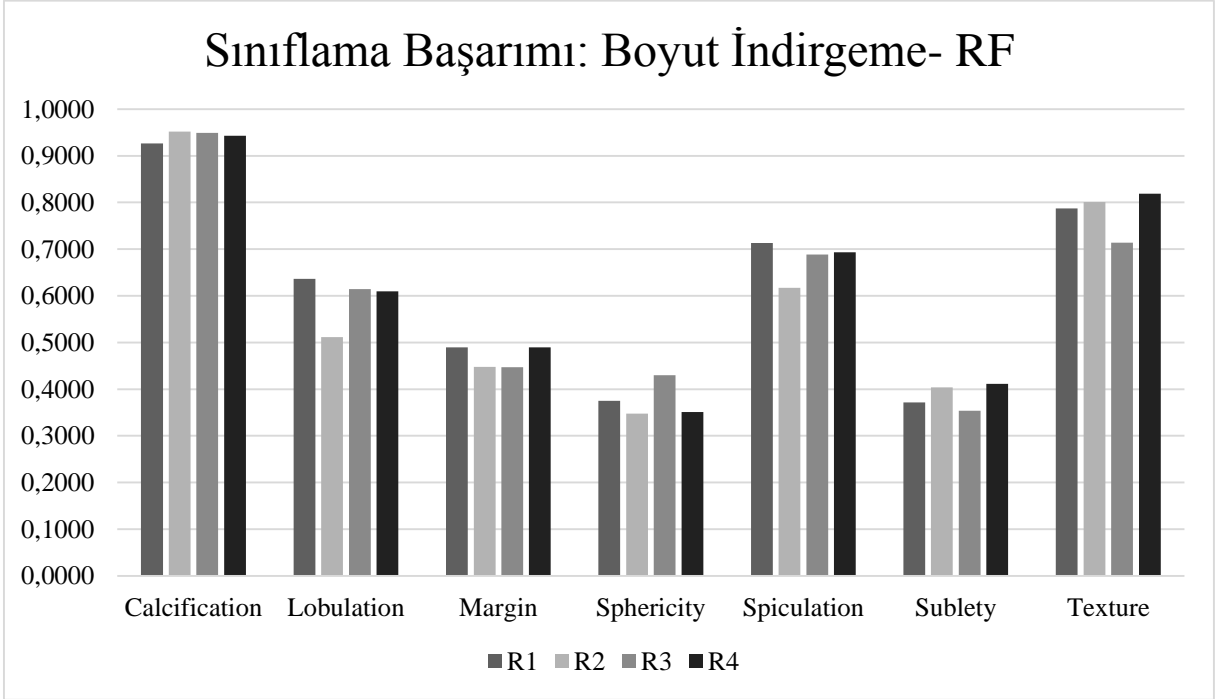
Radyografik Tanımlayıcı	SB	AUC
	<i>Tüm Öznitelikler ile</i>	
Calcification	0.9395	0.8800
Lobulation	0.6081	0.7043
Margin	0.5023	0.7188
Sphericity	0.4174	0.6606
Spiculation	0.6907	0.6719
Sublety	0.4209	0.7436
Texture	0.8105	0.8050
	<i>Boyut indirgeme ile</i>	
Calcification	0.9430	0.8963
Lobulation	0.6093	0.5977
Margin	0.4895	0.7251
Sphericity	0.3512	0.5706
Spiculation	0.6930	0.6284
Sublety	0.4116	0.7230
Texture	0.8186	0.8157

Çizelge 4.11 R4 okuması için SVM sınıflama algoritması ile elde edilen sonuçlar.

Radyografik Tanımlayıcı	SB	AUC
	<i>Tüm Öznitelikler ile</i>	
Calcification	0.9442	0.9057
Lobulation	0.6233	0.6707
Margin	0.5291	0.7592
Sphericity	0.3919	0.6726
Spiculation	0.6942	0.6679
Sublety	0.4500	0.7611
Texture	0.8198	0.8295
	<i>Boyut indirgeme ile</i>	
Calcification	0.9395	0.9128
Lobulation	0.6186	0.5711
Margin	0.5047	0.7488
Sphericity	0.3523	0.5723
Spiculation	0.6919	0.5882
Sublety	0.4174	0.7231
Texture	0.8302	0.8133

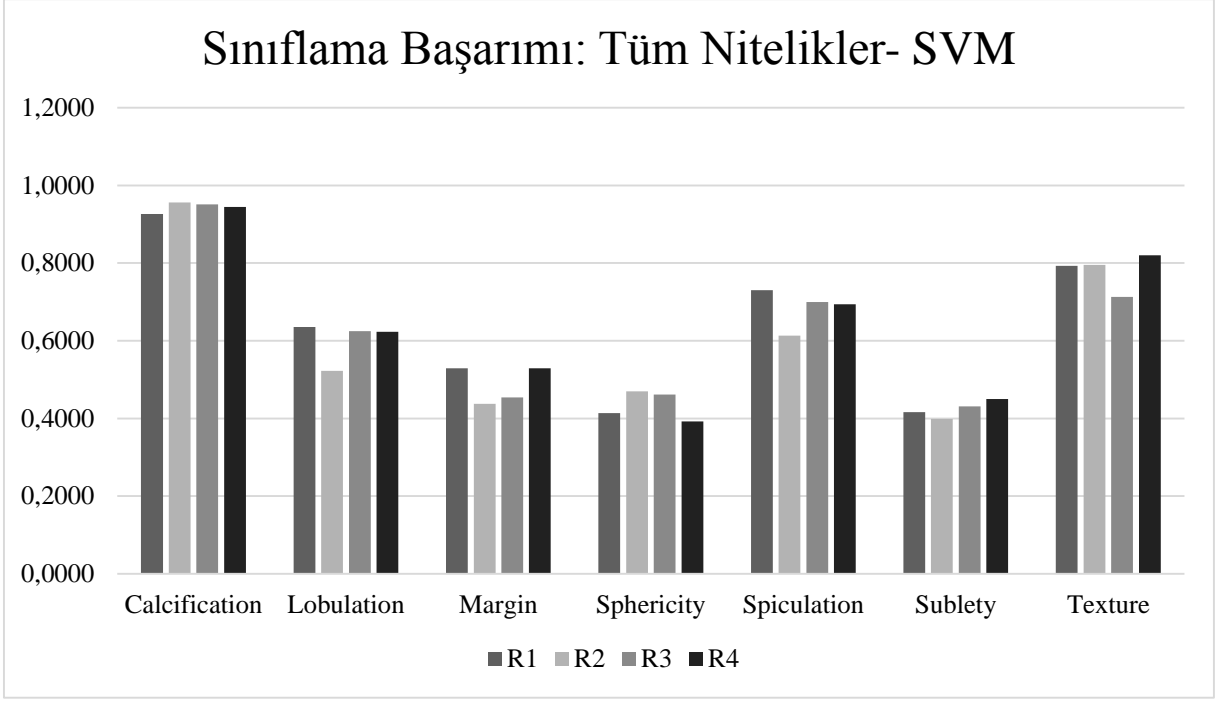


(a)

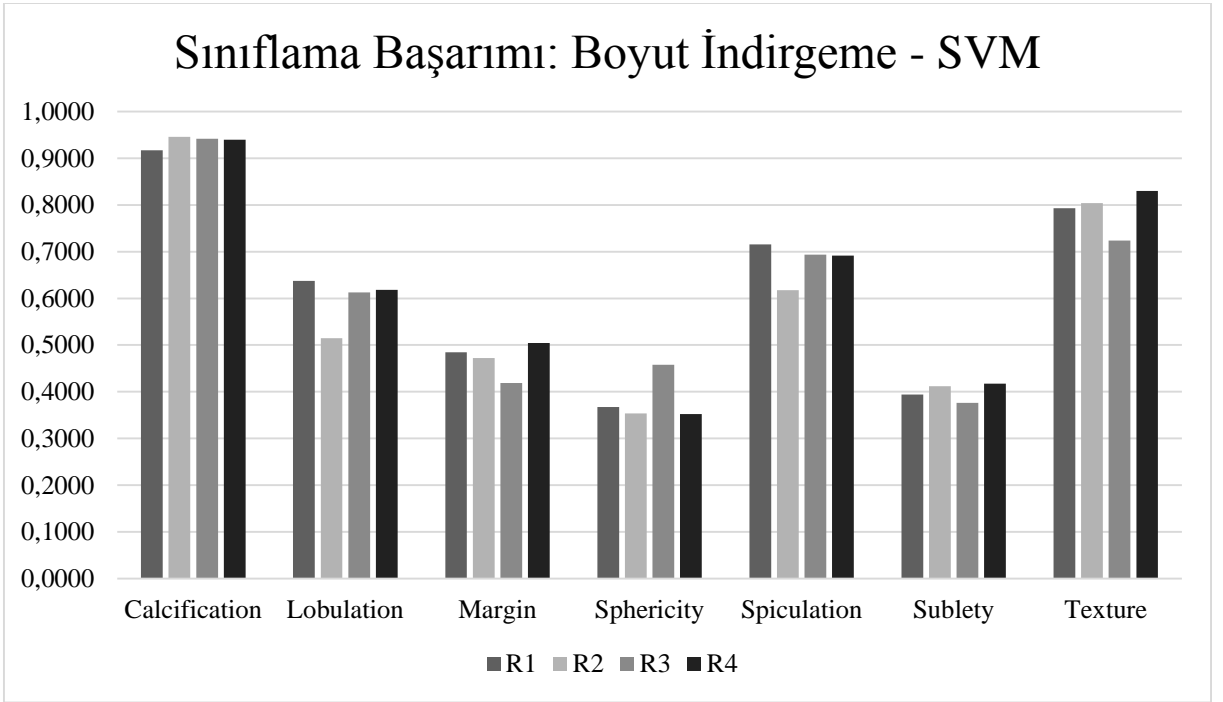


(b)

Şekil 4.6 (a) RF algoritması ve tüm öznelikler ile yapılmış sınıflama başarımlarının tüm okumalar ve tanımlayıcılar için gösterimi. (b) RF algoritması ve boyut indirgeme sonrası yapılmış sınıflama başarımlarının tüm okumalar ve tanımlayıcılar için gösterimi.



(a)



(b)

Şekil 4.7 a) SVM algoritması ve tüm öznelikler ile yapılmış sınıflama başarımlarının tüm okumalar ve tanımlayıcılar için gösterimi. (b) SVM algoritması ve boyut indirgeme sonrası yapılmış sınıflama başarımlarının tüm okumalar ve tanımlayıcılar için gösterimi.

4.5. Kesin Referans ile Sınıflama Sonuçlarının Karşılaştırılması

Bu kesimde Kesim 5.1’de oluşturulan kesin referans verisi ile radyografik tanımlayıcıları sınıflama amacıyla kullanılacak 5 farklı tekil sınıflayıcının sonuçları verilecektir. Ayrıca bu veri ile Kesim 5.2’de ölçümleri verilen R1, R2, R3, R4 kümeleri ile karşılaştırma yapılacaktır.

Çizelge 4.12 *Calcification* tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.

Yöntem	SB	AUC
SVM	0.9797	0.9851
kNN	0.9775	0.9629
Naive Bayes	N/A	0.9715
Random Forest	0.9816	0.9572
Adaboost	N/A	0.5005

Çizelge 4.13 *Lobulation* tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.

Yöntem	SB	AUC
SVM	0.8867	0.6847
kNN	0.8401	0.7211
Naive Bayes	0.7214	0.8369
Random Forest	0.8819	0.8274
Adaboost	0.8092	0.7188

Çizelge 4.14 *Margin* tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.

Yöntem	SB	AUC
SVM	0.6827	0.7148
kNN	0.6492	0.7170
Naive Bayes	0.6279	0.7568
Random Forest	0.6868	0.7276
Adaboost	0.6279	0.6705

Çizelge 4.15 *Sphericity* tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.

Yöntem	SB	AUC
SVM	0.6333	0.8603
kNN	0.5561	0.7873
Naive Bayes	0.6105	0.8620
Random Forest	0.6026	0.8529
Adaboost	0.6105	0.7895

Çizelge 4.16 *Spiculation* tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.

Yöntem	SB	AUC
SVM	0.9333	0.7747
kNN	0.9140	0.7744
Naive Bayes	N/A	0.8632
Random Forest	0.9315	0.8323
Adaboost	N/A	0.5530

Çizelge 4.17 *Sublety* tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.

Yöntem	SB	AUC
SVM	0.7858	0.9246
kNN	0.7583	0.8739
Naive Bayes	0.6615	0.8934
Random Forest	0.7741	0.9145
Adaboost	0.7096	0.8526

Çizelge 4.18 *Texture* tanımlayıcısı için sınıflama yöntemlerinin sınıflama başarımları ve eğri altında kalan alan değerleri.

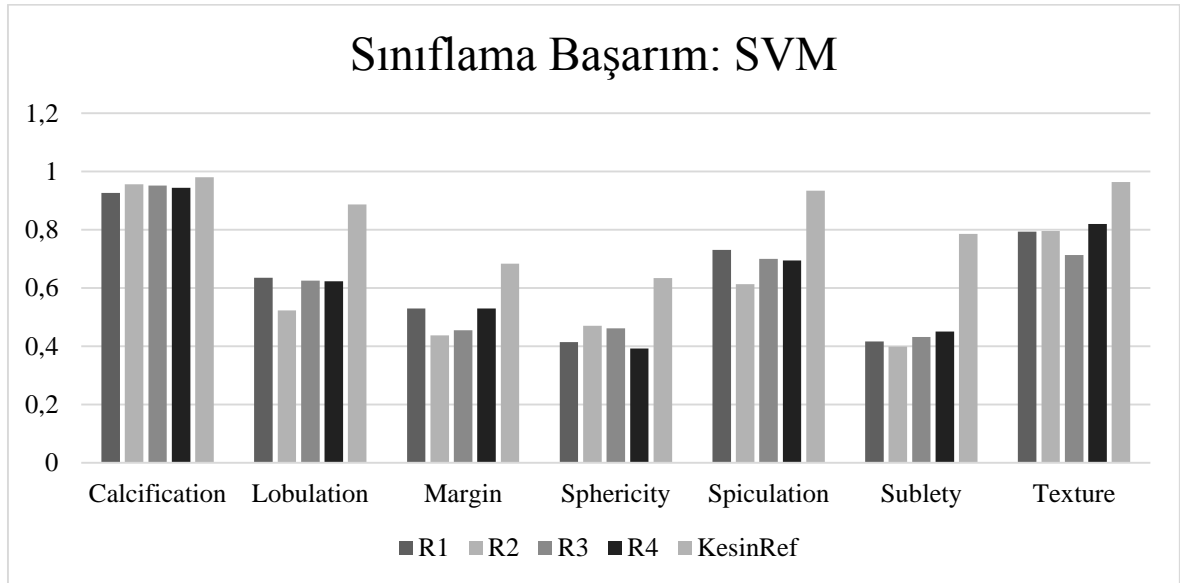
Yöntem	SB	AUC
SVM	0.9636	0.9285
kNN	0.9600	0.8594
Naive Bayes	N/A	0.8905
Random Forest	0.9613	0.8833
Adaboost	N/A	0.5674

Çizelge 4.12 - Çizelge 4.18'de her bir radyografik tanımlayıcı için elde edilen sonuçlar gösterilmiştir. Önceki kesimde R1-R4 için yapılan sınıflamada kullanılan SVM ve RF yöntemleri kesin referans için de en başarılı sonuçları üretmiştir. Şekil 4.8'de ve Çizelge 4.19'de en başarılı sınıflayıcı olan SVM'den elde edilen sınıflama başarımları sonuçları gösterilmiştir.

Ayrıca, Kesim 3.7.1'de tanımı verilen Krippendorff'un Alpha katsayısı kullanılarak R1-R4 için Çizelge 4.19'de verilen değerler üzerinde uyum değerlendirme yapılmıştır. Değerlendirme sonucunda, $\alpha = 0.9514$ çıkmıştır. Tanıma göre α değeri en yüksek 1 en düşük 0 olmaktadır. Rastgele örneklem sayılabilecek R1-R4 başarımlarının birbirine yakın olması, α değerinin 1'e yakın çıkması ve çoğunluk oylama ile elde edilen kesin referans üzerinde yapılan çalışmanın daha yüksek başarımlı olması;

- Tüm değerlendiricilerin aynı düzeyde tecrübeye sahip olarak değerlendirilmesi,
- Çoğunluğun ortak kararının üzerinde daha başarılı sonuçların alınabileceği,

Varsayımlarını doğrular öznitelikte görünmektedir. Bundan sonraki kesimlerde, bu sonuçları dikkate alarak, çoğunluğun ortak kararından elde edilen bilgi üzerinde yöntemler geliştirilmeye çalışılacaktır.



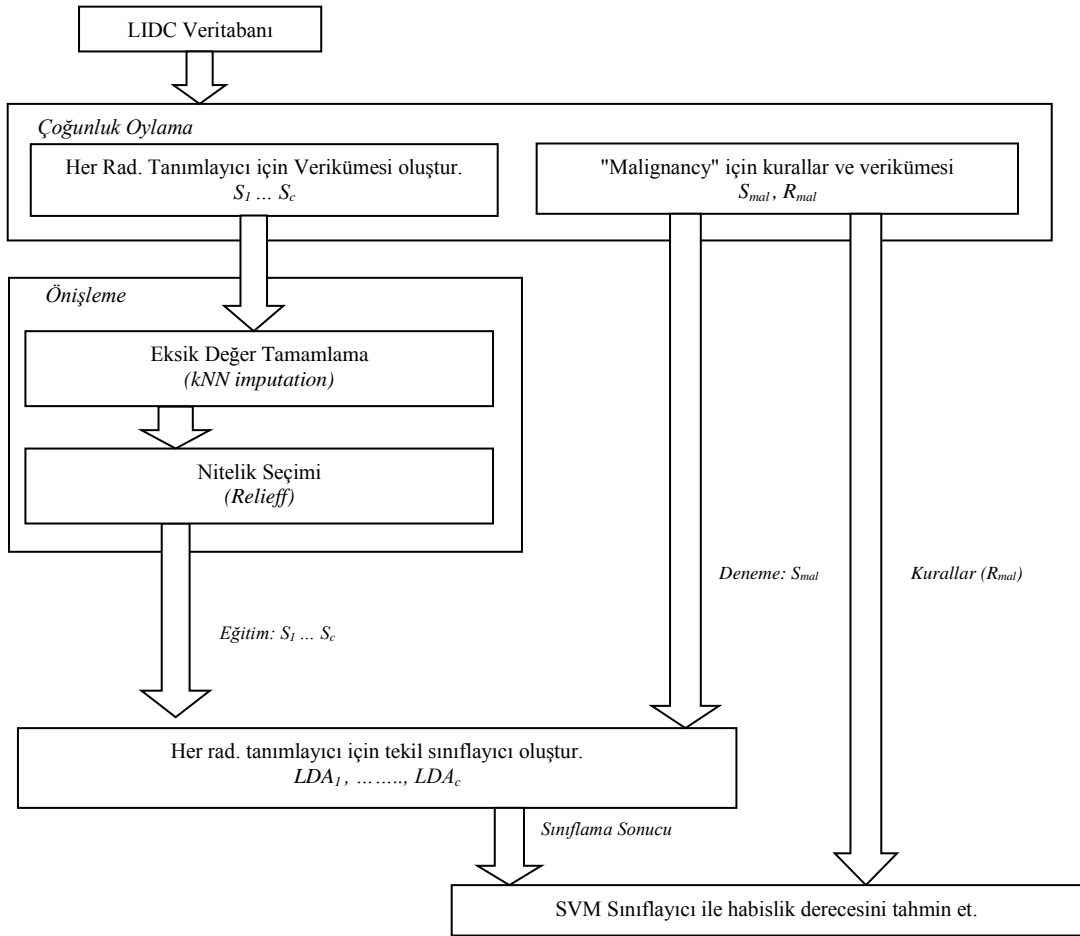
Şekil 4.8 SVM sınıflayıcı için R1-R4 ve kesin referans verikümelerinden elde edilen sınıflama başarımları sonuçları grafiği.

Çizelge 4.19 SVM sınıflayıcı için R1-R4 ve kesin referans verikümelerinden elde edilen sınıflama başarımı sonuçları çizelgesi.

	R1	R2	R3	R4	KesinRef
Calcification	0,9266	0,9560	0,9514	0,9442	0,9797
Lobulation	0,6351	0,5225	0,6247	0,6233	0,8867
Margin	0,5288	0,4375	0,4543	0,5291	0,6827
Sphericity	0,4137	0,4695	0,4616	0,3919	0,6333
Spiculation	0,7300	0,6134	0,6998	0,6942	0,9333
Sublety	0,4159	0,3986	0,4310	0,4500	0,7858
Texture	0,7932	0,7952	0,7129	0,8198	0,9636

5. İKİ KATLI TEKİL SINIFLAYICILAR İLE SINIFLAMA YÖNTEMİ

Bu çalışmada, LIDC veri tabanı kullanarak radyografik tanımlayıcılar üzerinden nodülün habisliğini tespit etmek için doğrusal ayırma analizi (LDA) ve destek vektör makinaları (SVM) tabanlı iki katlı sınıflayıcı oluşturulmuştur. Bu deney düzeneği ile görüntü üzerinden çıkarılmış öznelikler sonrası radyologların anlamaya vardığı tanımlayıcı verilerini kullanarak habislik derecesi tahmin edilmiştir.



Şekil 5.1 Yöntemin genel akış çizeneği.

Şekil 5.1'de gösterilen yöntem üç ana kesimde ele alınabilir. Önişleme kesimi; LIDC veri tabanından verilerin elde edilmesi, radyografik tanımlayıcılar için veri kümelerinin oluşturulması, boyut indirgemenin yapılması ve veri kümesi dengeleme yaklaşımlarının uygulanması işlemlerinin yapıldığı kesimdir. Tekil sınıflayıcılar kesiminde ise her bir radyografik tanımlayıcı için ayrı LDA sınıflayıcılar tanımlanmıştır. Son kesimde ise, tekil

LDA sınıflayıcılardan gelen sonuçlar SVM sınıflayıcıya iletilerek habislik derecesi (*malignancy*) tahmin edilmiştir.

5.1. Veri Kümesi Hazırlama ve Önleme Adımı

Her bir radyografik tanımlayıcının veri kümesi en az üç radyologun anlaştığı örneklerin hiçbir değişiklik yapılmadan eklenmesi ile oluşturulmuştur. Her bir veri kümesinde nodül boyutunun küçüklüğüne bağlı olarak bazı örneklerin nitelik değerleri eksiktir. Bu eksik değerleri tamamlamak için *k-nearest neighbour imputation* [64] yöntemi kullanılmıştır. Veri kümelerinin dengelenmesi için SMOTE [65] ve dışbükey örtü tabanlı örneklem azaltma yaklaşımı kullanılmıştır.

Bu yöntemler hakkında ayrıntılı bilgi Kesim 3'te verilmiştir.

5.2. Sınıflama Aşaması

Her radyografik tanımlayıcı için ayrı veri kümesi oluşturduktan sonra, sınıflayıcı modelleri geliştirilmeden, öznelik seçme (*feature selection*) ile boyut indirgeme işlemi uygulanmıştır. Bu işlem için Relieff [62] yöntemi kullanılmıştır ve en etkin altı öznelik seçilmiştir. Seçilen öznelik sayısı isteğe bağlı olarak ya da başarımlar gereksinimlerine göre değiştirilebilir. Çalışma kapsamında öncelikle yöntemin zaman gereksinimi gözönüne alınarak bu değer belirlenmiştir.

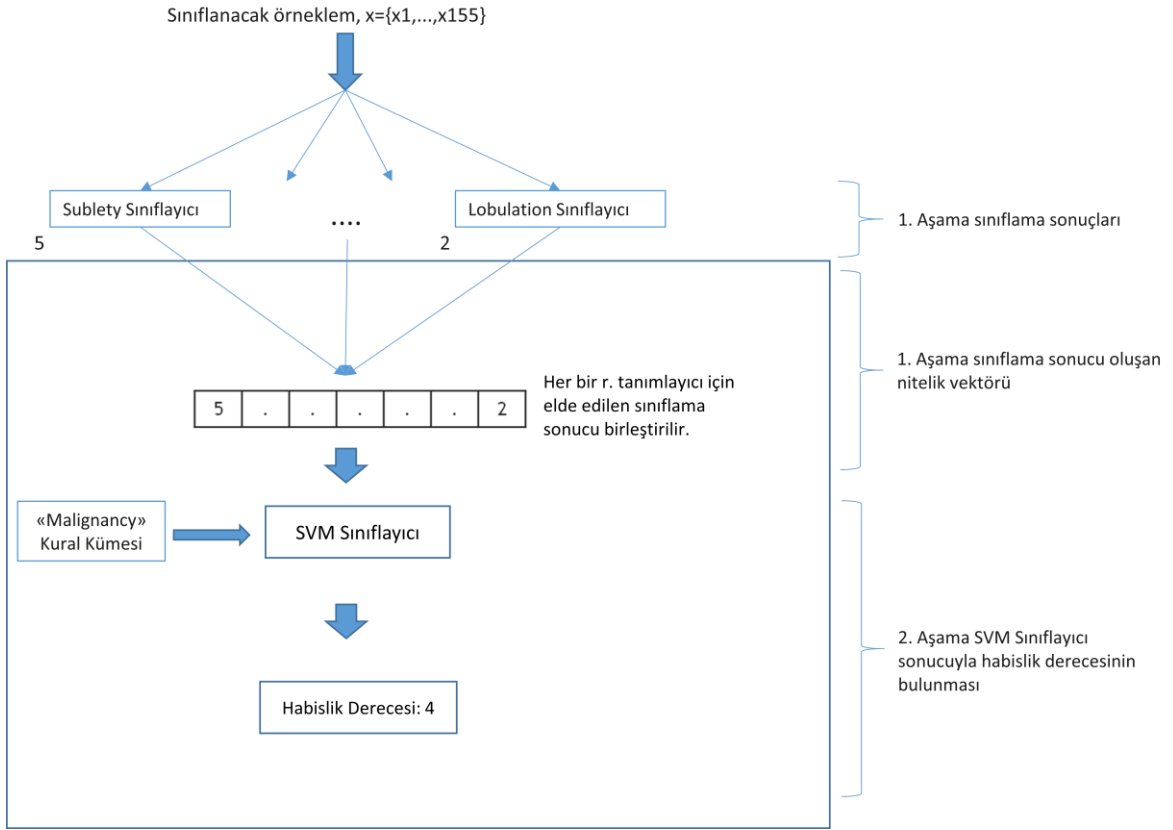
Dış bükey örtü yöntemi, zaman gereksinimi yüksek olan bir yöntemdir. Karmaşıklığı örnek sayısı ve verikümesi boyutu ile artmaktadır. Bu yöntem için Barber ve diğerlerinin önerdiği *Quick Hull* [91] algoritması kullanılmıştır. *Quick Hull* algoritması yüksek boyutlu verileri, benzer amaçlı algoritmalarından daha etkin olarak hesaplamaktadır, fakat yine de dokuz boyut üzerindeki verilerde kayda değer bir işlem süresi almaktadır. İstenen, en makul hesaplama süresinde bilgi kaybını en aza indirecek öznelik sayısına erişmektir. Çizelge 5.1'de görüleceği üzere, altı boyuttan sonra algoritmanın hesaplama süresi üssel olarak artmaktadır. Bu sebeple, Relieff algoritması ile belirlenen en önemli altı öznelik, dış bükey örtüyü hesaplarırken kullanılmıştır.

Çizelge 5.1 Quick Hull algoritmasının her radyografik tanımlayıcı için en önemli 3-9 öznitelik için işlem süresi verilmiştir.

	3	4	5	6	7	8	9
Calcification	0,10	0,04	0,05	0,59	5,40	56,60	787,49
Lobulation	0,80	0,77	0,76	1,04	3,71	27,78	452,09
Margin	0,19	0,19	0,20	0,29	2,64	19,21	95,56
Sphericity	0,07	0,08	0,13	0,33	1,44	8,90	50,62
Spiculation	0,38	0,38	0,45	0,72	3,26	22,09	291,25
Sublety	0,53	0,53	0,55	0,78	4,20	57,76	315,04
Texture	0,53	0,53	0,55	0,88	6,49	50,80	760,84

Katlı sınıflayıcının ilk katmanında her bir radyografik tanımlayıcı için elde edilen eğitim veri kümesi ile ayrı ayrı LDA sınıflayıcıları oluşturulmuştur. Her bir sınıflayıcı, bir örneğe ilişkin görüntü özniteliklerine karşılık verilen radyografik tanımlayıcı değerlendirmeleri kullanılarak eğitilmektedir. Her model farklı veri kümesi üzerinde olduğu gibi, öznitelik seçme ile de farklı öznitelik kümelerine sahiptir. Deneme kümesinden alınan bir örnek her bir sınıflayıcıya yollanır ve elde edilen sınıflama sonuçları ikinci katman için deneme verisi olarak tutulur.

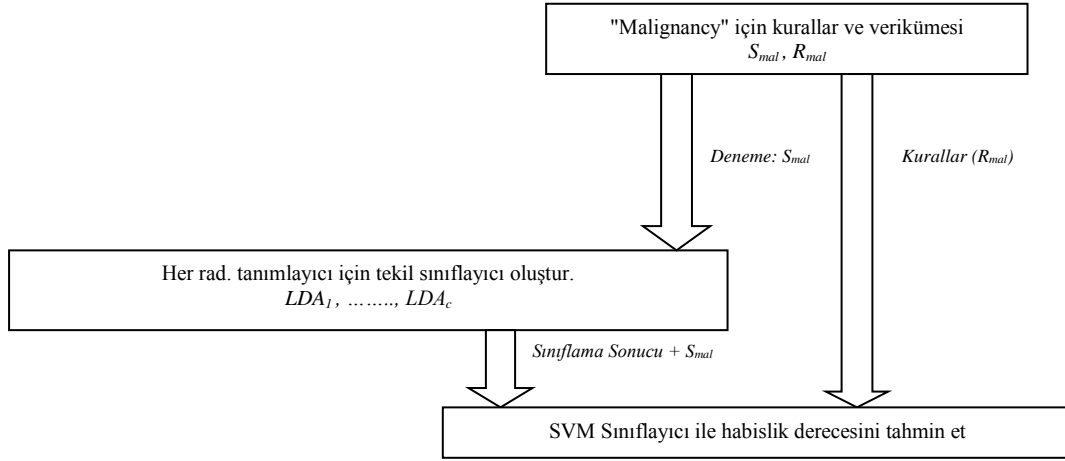
İkinci katmanda ise habislik (*malignancy*) tanımlayıcısı için eğitilmiş bir SVM sınıflayıcısı kullanılır. Bu sınıflayıcı bir örneğe ilişkin LIDC veritabanında bulunan radyografik tanımlayıcı değerlerine karşılık verilen habislik değeri kullanılarak eğitilmiştir. İlk katmandan elde edilen çıktı bu sınıflayıcıya verilir ve nodülün habislik derecesine ait çıktı elde edilir. Bu işlem *malignancy* veri kümesindeki her bir eleman için yapılır ve yanlılığı önlemek için diğer veri kümelerinde aynı nodüle ait bir bilgi varsa sınıflayıcılar eğitilmeden önce veri kümelerinden çıkarılır. Şekil 5.2'de sınıflama için örnek gösterim bulunmaktadır.



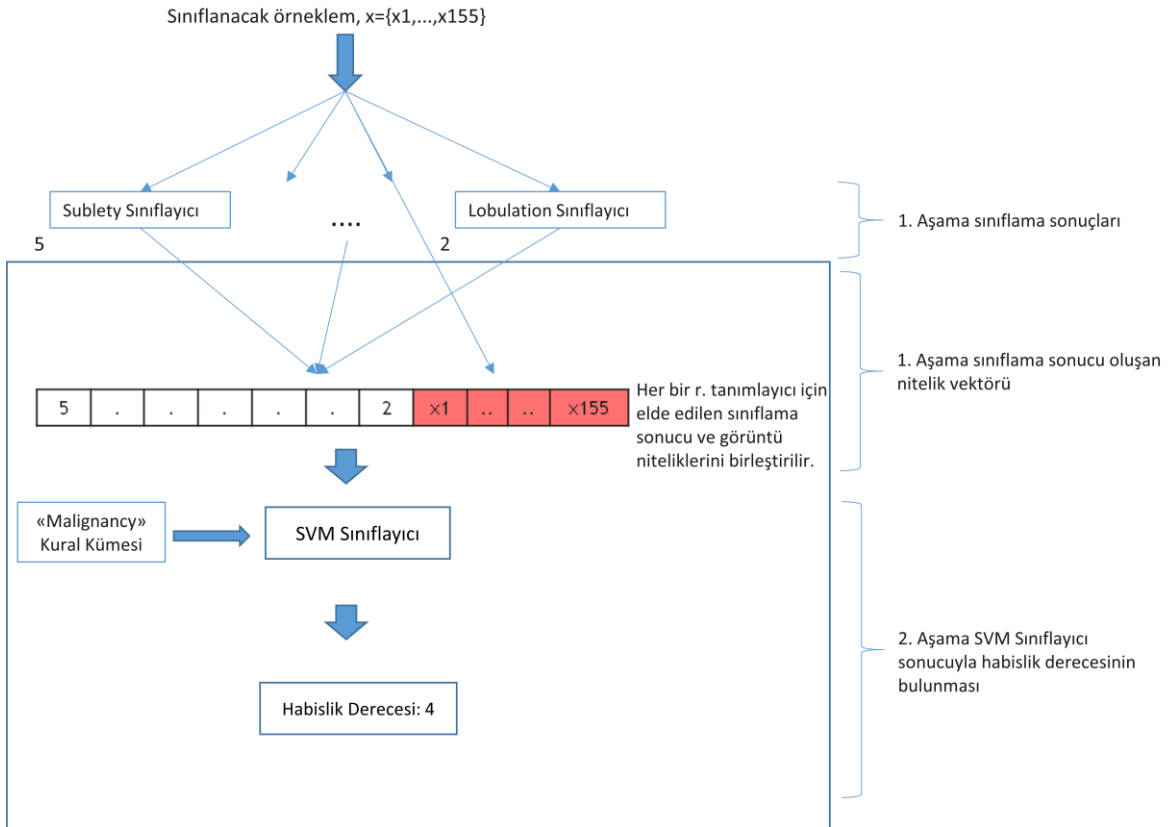
Şekil 5.2 Yöntem için çalıştırma örneği gösterimi.

5.2. Görüntü Öznitelikleri ile İyileştirilmiş Sınıflama Aşaması

Bu kesimde önerilen sınıflama yönteminin ikinci aşamasında, radyografik tanımlayıcılardan elde edilen sınıflama sonuçlarına ek olarak, birinci aşama sınıflayıcılara gönderilen görüntü öznitelikleri de eklenmiştir. Şekil 5.3'de, Şekil 5.1'de verilen akış çizeneğinde ikinci aşamada gösterilen sınıflama adımının geliştirilmiş hali verilmiştir. Şekil 5.4'de ise, sınıflama için örnek gösterim bulunmaktadır.



Şekil 5.3 Yöntemin ikinci aşamasına ait gösterimin geliştirilmiş hali.



Şekil 5.4 Ek yöntem için çalıştırma örneği gösterimi.

5.3. Sonuçlar

Bu çalışmada her bir radyografik tanımlayıcı için ayrı veri kümeleri ve doğrusal ayırım sınıflayıcıları oluşturulmuştur. Habislik tanımlayıcısı için elde edilen veri kümesi ise bu sınıflayıcılara girdi olarak verilip radyografik tanımlayıcı dereceleri alınmıştır. Bu sınıflayıcılardan çıkan veri ise SVM sınıflayıcısına girdi olarak verilmiştir. SVM

sınıflayıcısının eğitim kümesi olarak LIDC veri tabanından elde edilmiş, habislik üzerinde anlaşmaya varılmış radyolog değerlendirmeleri kullanılmıştır.

Genelleştirme yöntemi olarak bir örneği dışarıda bırakma (*leave one out*) kullanılmıştır. Tüm deneme kümesi elemanları için deney tekrarlanmıştır. Yanlılığı azaltmak için her adımda deneme örneğindeki nodül birinci aşama eğitim veri kümelerinde mevcutsa bu kümelerden o deneme adımı için atılmıştır. Habislik derecelerindeki belirsizlik sebebiyle, bu dereceleri (1,2) iyi huylu, (3) belirsiz , (4,5) kötü huylu olarak gruplanarak deney tekrarlanmıştır.

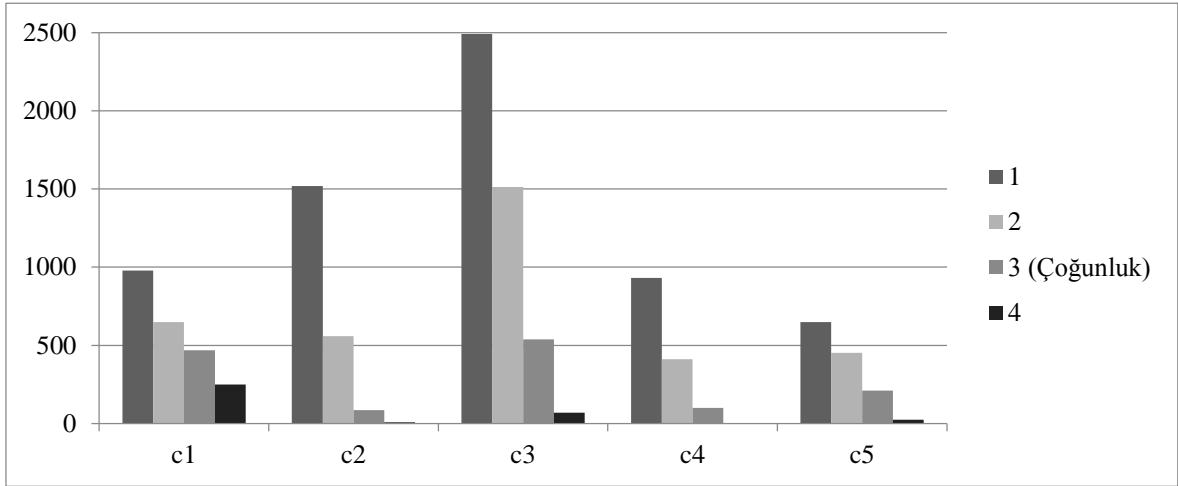
Elde edilen sonuçlar farklı özellikteki tekil sınıflayıcılar (Naive Bayes, Adaboost, kNN, SVM, Random Forest) ile karşılaştırılmıştır.

Çizelge 5.2 Görüntü özneliklerine göre ve radyografik tanımlayıcılara göre tekil sınıflayıcı sonuçları. En altta geliştirilen yöntemin sınıflama sonuçları verilmiştir.

Yöntem	5 Sınıf SB	5 Sınıf Duy.	5 Sınıf Seç.	3 Sınıf SB	3 Sınıf Duy.	3 Sınıf Seç.
LDA	0.6610	0.3832	0,8886	0.8200	0.8001	0.9006
Adaboost	0.7254	0.6737	0,9311	0.8135	0.8265	0.9055
Naive Bayes	0.7250	0,6700	0,9325	0.8100	0.8250	0.9060
kNN	0.7665	0.6314	0.9350	0.7959	0.8003	0.8930
SVM	0.7828	0.5384	0.9345	0.8005	0.8108	0.9021
RF	0.8040	0.5797	0.9421	0.8151	0.8161	0.9035
Yöntem GN	0.7791	0,5171	0.9285	0.8006	0.7863	0.8945
Yöntem GN+IK	0.8159	0.5665	0.9384	0.8221	0.8219	0.9096

Çizelge 5.2'den görüleceği üzere doğrudan öznelikler ve radyolog tanımlayıcıları tarafından elde edilen sınıflama sonuçları, katlı sınıflayıcı ile birbirine yakındır.

Şekil 5.5'da habislik derecesi sınıflarının (c1-c5) farklı kesin referans kabullerindeki örnek sayısı dağılımı gösterilmiştir (en az 1-4 radyoloğun anlaşığı nodül örneklerinin sayısı). Tez çalışmasında çoğunluk oylama kullanılmıştır ve grafikte en az 3 radyoloğun anlaşığı (Çoğunluk) etiketi ile gösterilmiştir. Radyologlar arasında beklenen anlaşma oranı arttıkça her sınıfa düşen örnek sayısının azalması beklenen bir durumdur. Fakat ara değerlendirme (2 - moderately unlikely ve 4 - moderately suspicious) sınıflarındaki azalma oranı diğer sınıflara göre çok daha fazladır. Bu durum, radyologların daha az belirsiz sınıflar (1 - highly unlikely , 3 - indeterminate , 5 – highly suspicious) üzerinde anlaşmaya meyilli olduğunu göstermektedir. Bundan yola çıkarak 3 sınıflı habislik veri kümesi üzerinde yapılan deneyler, beklendiği gibi daha başarılı sonuçlar üretmiştir.



Şekil 5.5 Habislik derecesi sınıflarının (c1-c5) farklı kesin referanslardaki dağılımı (1-4 arası radyolog anlaşması)

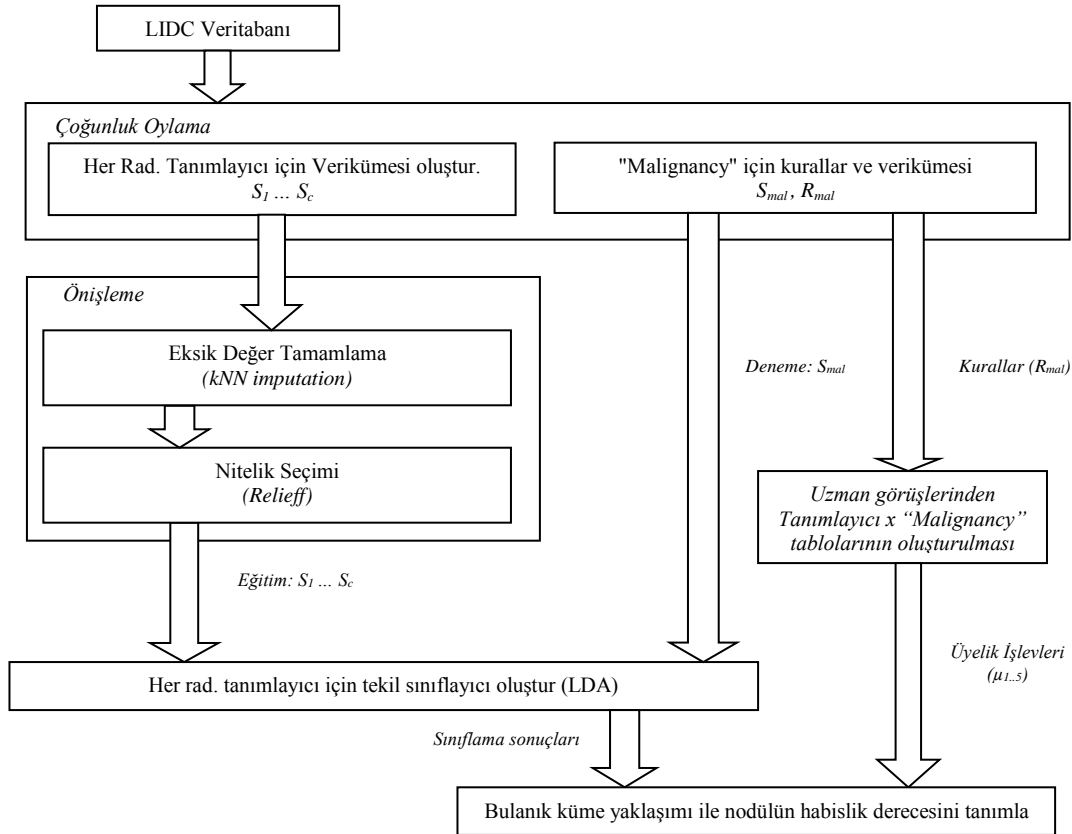
5.4. Tartışma

Bu çalışmada alt düzey öznitelikler kullanılarak radyografik tanımlayıcılar üzerinden habislik derecesinin tahmin edilmesi için bir çalışma yapılmıştır. Ek çalışmada ise birinci aşamada kullanılan görüntü öznitelikleri iki aşamaya da girdi olarak eklenmiştir. Sınıflama başarımı denenen tüm tekil sınıflayıcıların üzerinde olmuştur. Elde edilen sonuçlar, radyografik tanımlayıcıların bilgisi ile sınıflamanın mümkün olduğunu, ayrıca bu şekilde bir sınıflamanın tekil sınıflama yöntemlerine göre daha başarılı sonuçlar elde edebileceğini göstermiştir.

6. BULANIK YAKLAŞIMLAR İLE SINIFLAMA

6.1. Bulanık Küme Yaklaşımı ile Sınıflama

Bu çalışmada akciğer nodüllerinin habislik derecesinin belirlenmesi için bulanık küme yaklaşımları kullanılmıştır. LIDC veritabanı, çoklu değerlendirici tarafından öznel değerlendirmelere sahip olması nedeniyle, bulanık yaklaşımların kullanılmasına uygun bir veri tabanıdır [92]. Önceki çalışmalara benzer olarak ilk adımda tekil sınıflayıcılardan görüntü öznitelikleri kullanılarak radyografik tanımlayıcıların sınıf değerleri elde edilmiştir. Daha sonra LIDC veritabanındaki uzman değerlendirmeleri ile nodülün habislik derecesi üzerinden yapılan çoklu oylama sonuçlarından üyelik işlevleri hesaplanmıştır. Radyografik tanımlayıcı değerlerinin üyelik işlevlerine verilmesi sonrasında, her bir nodülün hangi habislik derecesine, ne kadar oranda ait olduğu sonucuna ulaşılmıştır. Bu çalışmada farklı üyelik işlevleri kullanılarak ölçümler yapılmıştır. Bu yöntemin ön işleme ve veri kümesi oluşturma aşamaları Kesim 5 'te verilen yöntem ile aynıdır. Yöntemin genel çizeneği Şekil 6.1'de verilmiştir.



Şekil 6.1 Yöntemin genel akış çizeneği.

6.1.1. Sınıflama Aşaması

Bu yöntemde de, önceki yöntemlerde kullandığı üzere, radyografik tanımlayıcılardan elde edilen sınıflama bilgisi nodülün habislik derecesini tespit etmek için kullanılacaktır. Sınıflama yöntemi ilk adımda görüntü özniteliklerinden radyografik tanımlayıcı sonuçlarını elde etme; daha sonra ise *malignancy* veri kümesinde radyologların değerlendirmeleri sonucu elde edilen uzman sonuçlarından faydalanarak oluşturulan matrisler aracılığıyla belirlenen üyelik işlevleri ile habislik derecesi üyeliklerinin bulunmasından oluşmaktadır.

6.1.1.1. Tekil sınıflayıcılar ile radyografik tanımlayıcıların elde edilmesi

Kesim 5’de tanımı verilen *tekil sınıflayıcılar ile iki aşamalı sınıflama yönteminde* uygulanan adımlardan ilki kullanılmıştır. Bu adımda, çoğunluk oylama ile her radyografik tanımlayıcı için ayrı veri kümesi oluşturduktan sonra, sınıflayıcı modelleri geliştirilmeden, öznitelik seçme ile boyut indirgeme işlemi uygulanmıştır. Bu işlem için Relieff [62] yöntemi kullanılmıştır ve en etkin altı öznitelik seçilmiştir.

Bu katmanda her bir radyografik tanımlayıcı için elde edilen eğitim veri kümesi ile ayrı ayrı doğrusal ayırım sınıflayıcıları (*LDA*) oluşturulmuştur. Her model farklı veri kümesi üzerinde olduğu gibi, öznitelik seçme ile de farklı öznitelik kümelerine sahiptir.

6.1.1.2. Bulanık küme yaklaşımı ile habislik derecesinin tespiti

Bulanık küme teorisinde, üyelik işlevleri tanımlanırken uzman bilgisinden faydalanılır [93]. Tanımladığımız sistemde üyelik işlevleri radyologların habislik (*malignancy*) derecesi üzerinde anlaştıkları, radyografik tanımlayıcı kümesindeki bilgilerden elde edilmiştir.

Öncelikle her bir tanımlayıcının derecelerinin, habislik derecelerinde olasılıksal dağılımları hesaplanmıştır. Yedi tanımlayıcı için 6x5 boyutunda birer matris oluşturulmuştur [94].

Çizelge 6.1 *Subtlety* tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.

		Malignancy				
		1	2	3	4	5
Subtlety	1	0,1622	0,0541	0,7838	0,0000	0,0000
	2	0,1531	0,1531	0,6735	0,0204	0,0000
	3	0,1897	0,1437	0,6207	0,0345	0,0115
	4	0,2919	0,0676	0,4973	0,1189	0,0243
	5	0,4528	0,0192	0,1648	0,0560	0,3072
	6	0,0000	0,0000	0,0000	0,0000	0,0000

Çizelge 6.2 *Calcification* tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.

		Malignancy				
		1	2	3	4	5
Calcification	1	0,0000	0,0000	1,0000	0,0000	0,0000
	2	1,0000	0,0000	0,0000	0,0000	0,0000
	3	0,9975	0,0000	0,0025	0,0000	0,0000
	4	0,4000	0,2000	0,4000	0,0000	0,0000
	5	0,9333	0,0333	0,0000	0,0000	0,0333
	6	0,0139	0,0891	0,5625	0,1007	0,2338

Çizelge 6.3 *Sphericity* tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.

		Malignancy				
		1	2	3	4	5
Sphericity	1	0,5000	0,5000	0,0000	0,0000	0,0000
	2	0,2222	0,1111	0,2963	0,0864	0,2840
	3	0,2436	0,0703	0,4005	0,0679	0,2178
	4	0,3056	0,0382	0,4315	0,0742	0,1506
	5	0,5330	0,0630	0,2951	0,0516	0,0573
	6	0,0000	0,0000	0,0000	0,0000	0,0000

Çizelge 6.4 *Margin* tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.

		Malignancy				
		1	2	3	4	5
Margin	1	0,0270	0,0000	0,4054	0,2703	0,2973
	2	0,0253	0,0886	0,4557	0,1266	0,3038
	3	0,0185	0,0556	0,4506	0,0741	0,4012
	4	0,1045	0,0734	0,5141	0,0960	0,2119
	5	0,5982	0,0551	0,2738	0,0313	0,0417
	6	0,0000	0,0000	0,0000	0,0000	0,0000

Çizelge 6.5 *Lobulation* tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.

		Malignancy				
		1	2	3	4	5
Lobulation	1	0,4592	0,0815	0,3920	0,0401	0,0272
	2	0,1877	0,0345	0,4713	0,1149	0,1916
	3	0,1888	0,0350	0,2867	0,0909	0,3986
	4	0,0964	0,0241	0,2771	0,0964	0,5060
	5	0,1364	0,0000	0,0000	0,1136	0,7500
	6	0,0000	0,0000	0,0000	0,0000	0,0000

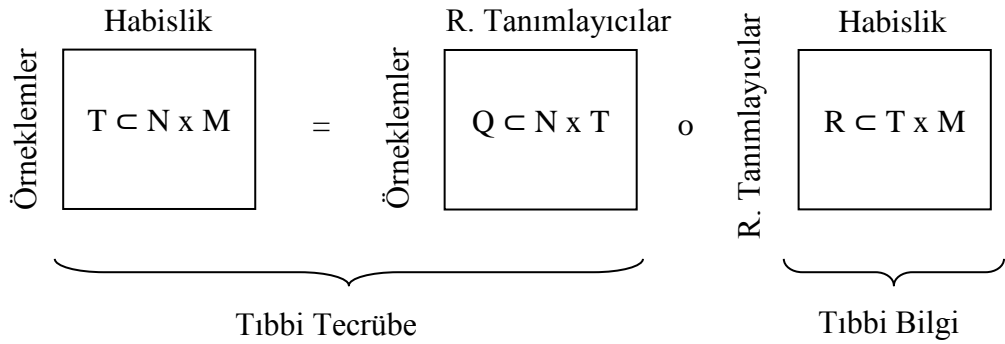
Çizelge 6.6 *Spiculation* tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.

		Malignancy				
		1	2	3	4	5
Spiculation	1	0,4530	0,0817	0,3937	0,0391	0,0324
	2	0,0867	0,0204	0,5306	0,1224	0,2398
	3	0,1226	0,0094	0,2925	0,0943	0,4811
	4	0,0769	0,0192	0,0577	0,1923	0,6538
	5	0,1071	0,0000	0,0000	0,1429	0,7500
	6	0,0000	0,0000	0,0000	0,0000	0,0000

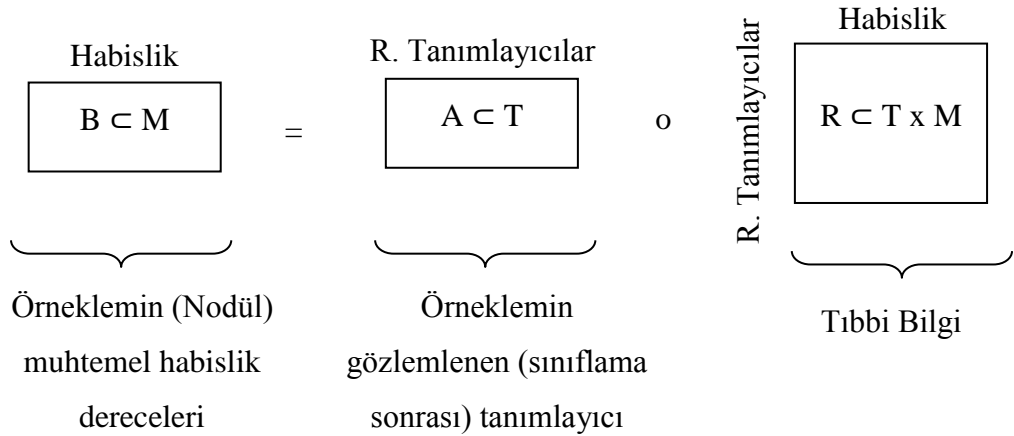
Çizelge 6.7 *Texture* tanımlayıcısı dereceleri ile habislik derecesi arasındaki dağılım matrisi.

		Malignancy				
		1	2	3	4	5
Texture	1	0,0000	0,1000	0,6667	0,1167	0,1167
	2	0,0000	0,1579	0,5263	0,2632	0,0526
	3	0,0233	0,0930	0,5814	0,1395	0,1628
	4	0,0148	0,0889	0,5630	0,1481	0,1852
	5	0,4222	0,0516	0,3238	0,0468	0,1557
	6	0,0000	0,0000	0,0000	0,0000	0,0000

Tıbbi Bilginin Derlenmesi



Teşhis Süreci



Şekil 6.2 Bulanık küme ve ilişkilerin tıbbi teşhiste uygulanması yaklaşımlarının probleme uyarlanmış halinin gösterimi.

Şekil 6.2’de Sanchez [95] tarafından önerilen ve Klir ve Yuan tarafından sadeleştirilen [93] bulanık küme ve ilişkilerin tıbbi teşhiste uygulanması yaklaşımlarının tezde belirlenen probleme uyarlanmış hali gösterilmiştir. Kısaca, Q ve T bilinen örneklerden elde edilmiş,

habislik derecesi ve radyografik tanımlayıcı sonuçlarıdır. R ise radyologlar tarafından verilen ölçümlerden elde edilmiş tanımlayıcı ve habislik derecesi arasındaki ilişkiyi ifade eden birikimli tıbbi tecrübeyi ifade eder (N : Örnekler (nodüller), M : Habislik derecesi (*malignancy*), T : Radyografik tanımlayıcılar). R ilişkisini, onaylanmış teşhis bilgisinden ziyade (biyopsi vs. yoluyla), teşhis hipotezi olarak düşünmek daha doğru olacaktır. Bu bilgi her ne kadar uzmanlardan elde edilmiş olsa da, kesin anlaşmaya varmaları beklenmeyen, öznel değerlendirmelerdir.

Tanımlanan sistemde, radyografik tanımlayıcılar ile habislik derecesi arasındaki ilişki önceki kesimde anlatılan habislik verikümesinden elde edilmiştir. Çizelge 6.1 - Çizelge 6.7’de her bir tanımlayıcı derecesinin, habislik dereceleri üzerinde bireysel üyelik puanları bu kural kümesinden elde edilerek verilmiştir.

T kümesindeki c adet radyografik tanımlayıcı elemanı t_1, \dots, t_c ile, LIDC veri kümesinde habislik derecesi olan *malignancy*’nin sözel değişkenleri (*linguistic variables*) m_1 (*highly unlikely*), m_2 (*moderately unlikely*), m_3 (*indeterminate*), m_4 (*moderately suspicious*), m_5 (*highly suspicious*) ile ifade edilmiştir. Radyografik tanımlayıcıların üyelik puanlarının olduğu matrisler P_1, \dots, P_c ve nodülün habislik derecesi ile aralarındaki korelasyon analizi sonucu elde edilen katsayı $\alpha_1, \dots, \alpha_c$ ’dir. Her bir m derecesi üzerinden, uzman görüşlerine dayanarak dört farklı kombinasyonla üyelik işlevi tanımlanmıştır. μ_{m_n}, m_n ($n=1-5$) habislik dereceleri için tanımlanmış üyelik işlevidir.

1. Tekil sınıflayıcılardan gelen sınıf değerine göre, her bir radyografik tanımlayıcı tablosundan elde edilen üyelik puanlarının ortalaması ile üyelik işlevlerinin tanımlanması (6.1).

$$\mu_{m_n}(t_1, \dots, t_c) = \frac{\sum_{j=1}^c P_j(t_j, n)}{c} \quad (6.1)$$

2. Tekil sınıflayıcılardan gelen sınıf değerine göre, her bir radyografik tanımlayıcı tablosundan elde edilen üyelik puanlarının en büyüğü ile üyelik işlevlerinin tanımlanması (6.2).

$$\mu_{m_n}(t_1, \dots, t_c) = \max_{j=1..c} [P_j(t_j, n)] \quad (6.2)$$

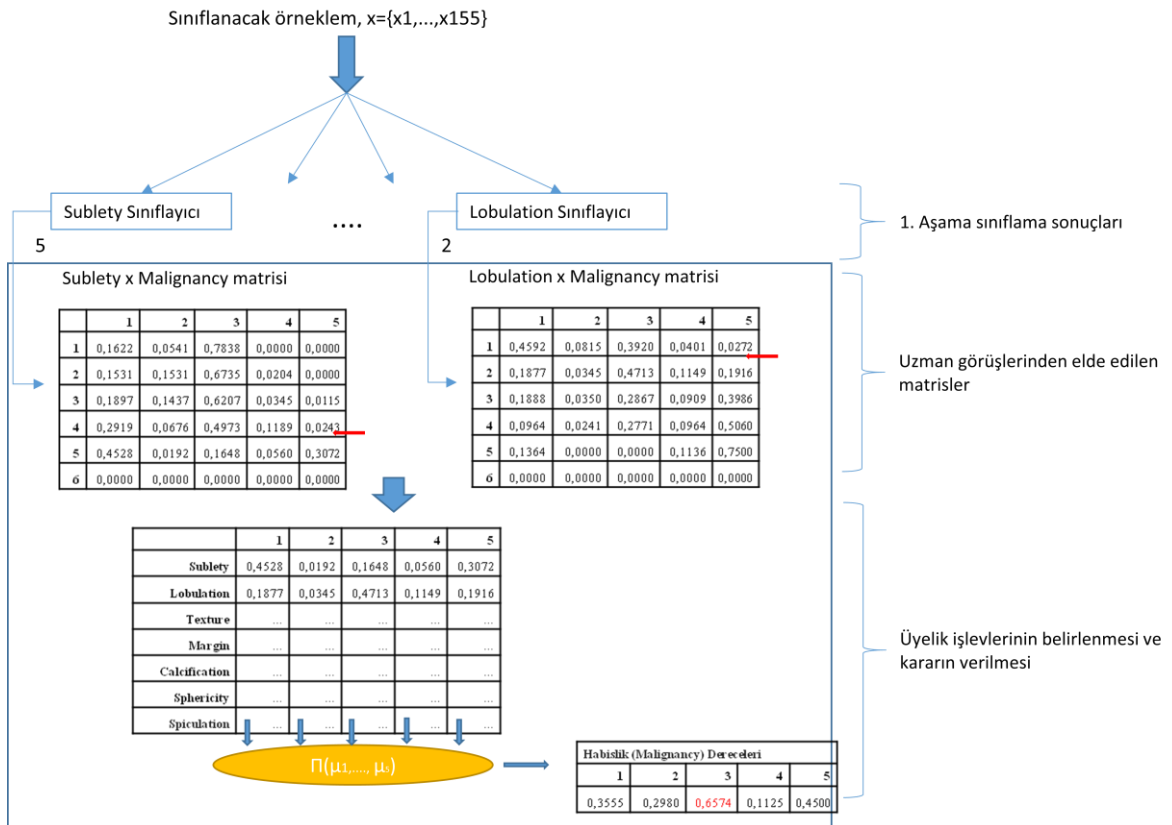
3. Tekil sınıflayıcılardan gelen sınıf değerine göre, her bir radyografik tanımlayıcı tablosundan elde edilen üyelik puanlarının ortalaması ve korelasyon katsayısı ile üyelik işlevlerinin tanımlanması (6.3).

$$\mu_{m_n}(t_1, \dots, t_c) = \frac{\sum_{j=1}^c P_j(t_j, n) \exp(\alpha_j)}{c} \quad (6.3)$$

4. Tekil sınıflayıcılardan gelen sınıf değerine göre, her bir radyografik tanımlayıcı tablosundan elde edilen üyelik puanlarının en büyüğü ve korelasyon katsayısı ile üyelik işlevlerinin tanımlanması (6.4).

$$\mu_{m_n}(t_1, \dots, t_c) = \max_{j=1..c} [P_j(t_j, n) \exp(\alpha_j)] \quad (6.4)$$

Şekil 6.3'de yöntem için çalıştırma örneği verilmiştir.



Şekil 6.3 Yöntem için çalıştırma örneği gösterimi.

6.1.2. Sonuçlar

Yöntemde verilen dört ayrı üyelik işlevi sonrası, bir nodülün her bir habislik derecesinde sahip olduğu üyelik dereceleri elde edilmektedir. Bulanıklığın ortadan kaldırılması aşamasında (durulaştırma, *defuzzification*), bu üyelik derecelerinden en büyük olanı ve/veya en büyük iki değeri sistemin çıktısı olarak kabul edilip sonuçlar verilecektir.

Çizelge 6.8 Dört farklı üyelik işlevinden (Üİ) elde edilmiş sınıflama başarımı (SB) sonuçları.

	Tek Sonuç SB	İki Sonuç SB	Gruplu Veri SB
Üİ-1	0.6488	0.8282	0.6552
Üİ-2	0.6235	0.8206	0.6457
Üİ-3	0.7531	0.8083	0.7615
Üİ-4	0.6549	0.8451	0.6549

Çizelge 6.8’de dört farklı üyelik işlevinden elde edilen sonuçlar verilmiştir. En yüksek sınıflama başarımı kalın yazı ile gösterilmiştir (Üİ-1: Üyelik puanlarının ortalaması ile hesaplanan üyelik işlevi. Üİ-2: Üyelik puanlarının ek yükseği ile hesaplanan üyelik işlevi. Üİ-3: : Üyelik puanlarının ortalaması ve korelasyon katsayısı ile hesaplanan üyelik işlevi. Üİ-4: Üyelik puanlarının ek yükseği ve korelasyon katsayısı ile hesaplanan üyelik işlevi.) Tek Sonuç, sistemin çıktılarında en büyük üyelik derecesinin doğru kabul edildiği durumdur. İki Sonuç, sistemin çıktılarında en büyük iki üyelik derecesinin doğru kabul edildiği durumdur. Gruplu Veri ise, önceki kesimlerde de deneye dahil edilmiş; *Malignancy* derecelerinin (1,2) iyi huylu, (3) belirsiz , (4,5) kötü huylu olarak gruplanmış halidir (Burada gruplanmış üyelik işlevlerinin aynı sınıfa ait sonuç ürettiği varsayılmıştır).

6.1.3. Tartışma

Bu kesimde bulanık küme ve mantık yaklaşımı kullanılarak temel düzeyde dört adet üyelik işlevi tanımıyla sınıflama işlemi yapılmıştır. İncelenen bulanık mantık yöntemleri, diğer yöntemlere göre daha yalın yapıda hazırlanmıştır.

Tıbbi veri, bu çalışmalarda kullandığımız LIDC veri kümesi de dahil olmak üzere, bulanık yaklaşımlarda kullanılmaya en uygun veri türüdür [92]. Çoklu değerlendirici, değerlendiricilerin farklı düzeyde tecrübeye ve bilgiye sahip olması ve hastalık karar süreçlerinin ve biyolojik sistemlerin karmaşıklığı bunun sebeplerinden bazılarıdır [25].

Pereira ve diğerleri [96] çalışmalarında, otomatikleştirilmiş mühendislik süreçlerinde bulanıklığın ortadan kaldırılması aşamasında, bulanık sonuçlardan kesin bir sonuç

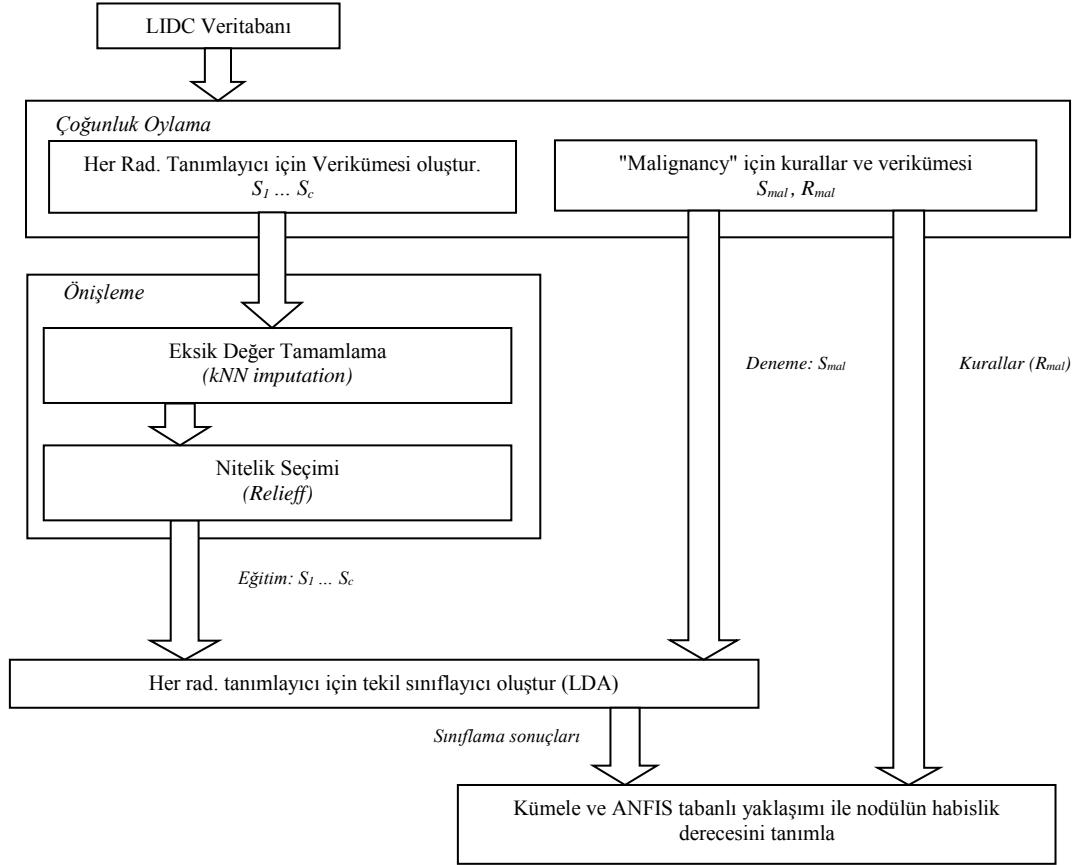
çıkarmanın kaçınılmaz olduğunu vurgulamışlardır. Hguyen ve Walker'ın [97] belirttiği gibi, bir tıbbi sistem, nihai kararı verecek olan doktorun görevini yapmaktan ziyade, tamamlayıcı/yardımcı bilgi sağlaması gerekir. Dolayısıyla, bulanık sistemlerden kesin (*crisp*) sonuçlar elde etmektense, ihtimalleri dereceleriyle birlikte sunmak daha faydalı olacaktır.

Üyelik işlevleri hakkında, ortalama alınarak oluşturulan üyelik işlevlerinde her bir radyografik tanımlayıcının üyelik puanı hesaba katılarak değerlendirme yapıldığı için daha yüksek başarımlar elde edildiğini düşünülebilir (Korelasyon katsayısının hesaba katılmadığı Üİ-1 ve Üİ-2 sırasıyla 0.6488, 0.6235 başarımlar elde edilmiştir.). Korelasyon katsayısının hesaba katılması ile iki grup arasındaki başarımlar artmıştır (Üİ-1, 0.6488'den, Üİ-3'de 0.7531'e; Üİ-2, 0.6235'den, Üİ-4'te 0.6549'a yükselmiştir).

6.2. Nöro Bulanık Çıkarım ANFIS ile Sınıflama

Bu çalışmada akciğer nodüllerinin habislik derecesinin belirlenmesi için uyarlanabilir nöro bulanık çıkarım (*adaptive neuro fuzzy inference system - ANFIS* [98]) yöntemi kullanılmıştır. Kesim 6.1'de, LIDC veri kümesi, çoklu değerlendirici tarafından öznel değerlendirmelere sahip olması nedeniyle, bulanık yaklaşımların kullanılmasına uygun bir veri tabanı olduğu ifadesi kullanılmıştır. ANFIS, bir işlev/parametre kestirim yaklaşımıdır. Sınıflamada kullanıldığında, ürettiği çıktılar kesikli olmadığı için, sonuçlar yukarı ya da aşağı yuvarlanarak örneğin ait olduğu sınıf bulunmaktadır [99], [100]. Çıkan sonucu yuvarlayarak sınıflama, çıktılarının yuvarlanması durumunda 0 ya da en üst sınıftan daha büyük değerlerde sınıflama sonucuna neden olacağından çok kabul edilebilir bir yaklaşım değildir. Çetişli'nin [99], [100] sunduğu ANFIS ve kümeleme tabanlı yöntem ile bu sorun aşılabilir, nodülün habislik derecesinin tahmini yapılmıştır.

Yöntemin ilk adımında tekil sınıflayıcılardan görüntü öznitelikleri kullanılarak radyografik tanımlayıcıların sınıf değerleri elde edilmiştir. İkinci adımda habislik kural kümesi üzerinden kümeleme ve ANFIS tabanlı sistem eğitilerek habislik derecesi tahmini yapılmıştır. Bu yöntemin ön işleme ve veri kümesi oluşturma aşamaları Kesim 5 'te verilen yöntem ile aynıdır. Yöntemin genel çizeneği Şekil 6.4'de verilmiştir.



Şekil 6.4 Yöntemin genel akış çizeneği.

6.2.1. Sınıflama Aşaması

Bu yöntemde de, önceki yöntemlerde kullandığı üzere, radyografik tanımlayıcılardan elde edilen sınıflama bilgisi nodülün habislik derecesini tespit etmek için kullanılacaktır. Sınıflama yöntemi ilk adımda görüntü özneliklerinden radyografik tanımlayıcı sonuçlarını elde etme; daha sonra ise habislik veri kümesinde radyologların değerlendirmeleri sonucu elde edilen kural kümesi ile tahmin etme şeklindedir.

6.2.1.1. Tekil Sınıflayıcılar İle Radyografik Tanımlayıcıların Elde Edilmesi

Kesim 5’da tanımı verilen *tekil sınıflayıcılar ile iki aşamalı sınıflama yönteminde* uygulanan adımlardan ilki kullanılmıştır. Bu adımda, çoğunluk oylama ile her radyografik tanımlayıcı için ayrı veri kümesi oluşturduktan sonra, sınıflayıcı modelleri geliştirilmeden öznelik seçme ile boyut indirgeme işlemi uygulanmıştır. Bu işlem için Relieff [62] yöntemi kullanılmıştır ve en etkin altı öznelik seçilmiştir.

Bu katmanda her bir radyografik tanımlayıcı için elde edilen eğitim veri kümesi ile ayrı ayrı doğrusal ayırım sınıflayıcıları (*LDA*) oluşturulmuştur. Her model farklı veri kümesi üzerinde olduğu gibi, öznelik seçme ile de farklı öznelik kümelerine sahiptir.

6.2.1.2. ANFIS ile Sınıflama

Yöntemin girişinde anlatıldığı üzere, ANFIS yönteminin doğrudan kesikli sınıf değerleri üzerinde kullanılması çok kabul edilebilir değildir. Çetişli'nin [99], [100] sunduğu ANFIS ve kümeleme tabanlı yöntem ile bu sorun aşularak, habislik derecesi tespit edilmiştir.

Çetişli, farklı bir uyarlanır nöro-bulanık sınıflayıcı önermiştir. Öncelikle bulanık kuralları oluşturmak için k-ortalamlar algoritması kullanılmıştır. Kullanıcı başlangıçta küme sayısını parametre olarak geçmektedir. Basit türetilebilir ifadeler olduğu için, bulanık küme tanımlarında Gauss üyelik işlevleri kullanılmıştır. Yöntem, Jang'ın [98] sunduğu ANFIS yöntemini temel almaktadır. Farklılıklar ise, kural ağırlıkları ve parametre en iyileştirmesidir. *Scaled conjugate gradient (SCG)* algoritması ile doğrusal olmayan parametrelerin en iyi değerleri bulunmaktadır. Bu yöntem büyük ölçekli problemlere de uyarlanabilir bir yaklaşımdır [101].

6.2.2. Sonuçlar

Çizelge 6.9'de sınıflamaya ait sınıflama sonuçları verilmiştir. Gruplu veri; *Malignancy* derecelerinin (1,2) iyi huylu, (3) belirsiz , (4,5) kötü huylu olarak gruplanmış halidir. Önceki deneylerde de gözlemlendiği üzere, sınıflama başarıımı gruplu veride artış göstermektedir.

Çizelge 6.9 ANFIS tabanlı sınıflayıcı için sınıflama sonuçları

	SB	Duy.	Seç.
5 Sınıflı Veri Deneyi	0.7607	0.5053	0.9250
3 Sınıflı Veri Deneyi	0.7715	0.7438	0.8787

6.2.3. Tartışma

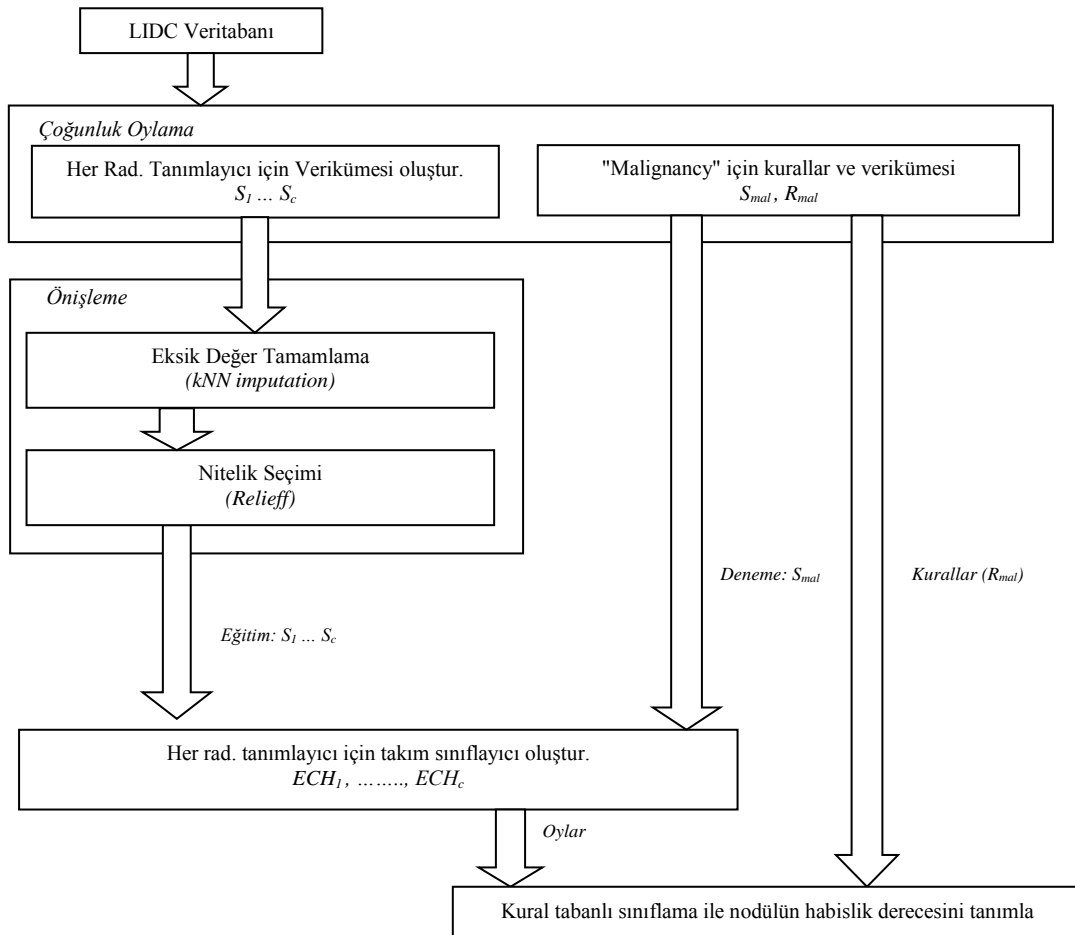
Önceki kesimde bulanık kümeler ile yapılan çalışma, ANFIS tabanlı sınıflayıcı ile genişletilmiştir. Önceki bulanık küme yöntemi bu yönteme göre daha basit bir yapıdadır. ANFIS ise uyarlanır bir yöntem olması nedeniyle kendi içinde bir en iyileştirme yaklaşımı ile daha iyi sınıflama sonuçları üretmiştir.

7. TAKIM SINIFLAYICILAR İLE SINIFLAMA

7.1. Takım Sınıflayıcılar Ve Kural Tabanlı Yaklaşım İle Sınıflama Yöntemi

Bu kesimde nodülün habislik derecesini, radyografik tanımlayıcılar üzerinde kurulmuş takım sınıflayıcıların (*ensemble classifiers - EC*) üyelik işlevleri olarak kullanıldığı bir kural tabanlı sistem sunulacaktır. Radyografik tanımlayıcılar üzerinden elde edilen oylama değerlerine takım sınıflayıcıların sonuçlarına ek olarak dışbükey örtü (*convex hull -CH*) yaklaşımından elde edilen sonuçlar da eklenmiştir. Bu yöntemlerin çıktıları habislik derecesinin belirlenmesi için temel olarak kullanılmıştır.

Elde edilen sonuçlar nodülün habislik derecesinin diğer radyografik tanımlayıcılara bağlı olarak tahmin edilebileceğini göstermektedir. Bu yöntemin önışleme ve veri kümesi oluşturma aşamaları Kesim 5 'te verilen yöntem ile aynıdır. Yöntem için tanımlanan genel akış çizeneği Şekil 7.1'de verilmiştir.



Şekil 7.1 Yöntemin genel akış çizeneği.

Yöntemi Şekil 7.1’de gösteildiği üzere üç ana kesimde ele alabiliriz. Önışleme kesimi; LIDC veritabanından verilerin elde edilmesi, radyografik tanımlayıcılar için veri kümelerinin oluşturulması, boyut indirgemenin yapılması ve veri kümesi dengeleme yaklaşımlarının uygulanması işlemlerinin yapıldığı kesimdir. Takım sınıflayıcılar kesimi; her bir radyografik tanımlayıcı ve onların sınıfları için ayrı SVM tabanlı sınıflayıcıların oluşturulması, her ayrı kümeyle ait dışbükey örtü noktalarının elde edilmesi ve bu noktaları temel alan ayrıştırıcı işlevlerin oluşturulması işlemlerini içeren kesimdir. Kural tabanlı sistem kesimi ise; takım sınıflayıcılardan elde edilen oylama bilgileri ve önışlemede oluşturulan kuralları kullanarak sisteme verilen örneğin habislik derecesinin tespit edildiği kesimdir.

7.1.1. Takım Sınıflayıcılar

Veri kümesi ön işleme adımından sonra, Şekil 7.2’de verilen algoritma ile takım sınıflayıcılar oluşturulmuştur. Her bir radyografik tanımlayıcının her bir derecesi için ayrı bir sınıflayıcı eğitilmiştir. Bu sebeple her bir radyografik tanımlayıcı derecesi için de ayrı eğitim kümeleri hazırlanmıştır. Örneğin “*sublety*” tanımlayıcısı için 5 ayrı veri kümesi oluşturulmuştur. Her bir dereceye karşılık, kalan diğer derecelerin hepsinin ayrı bir sınıfta olduğu düşünülmüştür. Böylece verilen “*sublety*” örneğinde her bir derece için 2 sınıflı sınıflayıcılar kullanılarak 5 adet sınıflayıcı eğitilmiştir.

Girdi:

t : örneğin nitelik vektörü.

S : rad. tanımlayıcı için veri kümesi n örnek, m öznelik, $S = \{x_1, \dots, x_m, y\}^{1..n}$

c : rad. tanımlayıcı için en büyük sınıf değeri.

f_S : rad. tanımlayıcı için nitelik sayısı

f_{CH} : dışbükey örtü için nitelik sayısı

Başla:

Döngü $i=1$ 'den c 'ye

Sınıf i için veri kümesi oluştur (i vs diğerleri) $S_i = \{x_1, \dots, x_m, y_i = \{0,1\}\}^{1..n}$

Nitelik seçimi $F_i = \text{relieff}(S_i)$

$$\text{Veri dengele, } S'_i = \begin{cases} \text{arttır, } & \text{size}(S_i(y=1)) < \frac{\text{size}(S)}{2c} \\ \text{azalt} & , \text{size}(S_i(y=1)) > \frac{3\text{size}(S)}{2c} \\ S_i(y=1), & \text{diğer} \end{cases}$$

Taban sınıflayıcıyı eğit $T_i = \text{TabanSınıflayıcı}(S'_i, F_i (1 \dots f_S))$

S'_i nin dışbükey örtü noktalarını bul, $H_i = \text{DışbükeyÖrtü}(S'_i(y=1), F_i (1 \dots f_{CH}))$

Döngü bitir

Oylama vektörü v ilkle, $\forall i: v(i) = 0$

Döngü $i=1$ 'den c 'ye

Eğer sınıfla $(t, T_i) = 1$ ise $v(i) = v(i) + 1$

Eğer dışbükeyörtüiçindemi(H_i, t) = 1 ise $v(i) = v(i) + 1$

Döngü bitir

Bitir**Çıktı:**

v : oylama vektörü

Şekil 7.2 Takım sınıflayıcılar ile oylama algoritması.

Her bir tanımlayıcı derecesi için öznitelik seçme işlemi uygulanmıştır. Öznitelik seçme işlemi olarak Relieff [62] yöntemi kullanılmış, bu yöntem sonucu ağırlıklandırılan tüm öznitelikler içinde en önemli altı tanesi seçilmiştir. Öznitelik sayısı isteğe bağlı olarak ya da yöntemlerin başarımları maliyetleri düşünülerek değiştirilebilir. Bu çalışmada dışbükey örtü hesaplamaları da göz önüne alınarak, en az bilgi kaybı olacak ve en iyi hesaplama başarımları elde edilebilecek değer olarak altı belirlenmiştir. İki sınıflı veri kümesi üzerinde yapılan bu öznitelik seçme yaklaşımının özellikle dengesiz veri kümesi dağılımı sonucu az ifade edilen sınıfların diğer sınıflardan ayrımını daha belirgin hale getirmesi için uygulanmıştır.

Öznitelik seçiminden sonra, dengeleme işlemi uygulanmış veri kümesi ile her radyografik tanımlayıcısının derecesi için 2-sınıflı SVM [72] sınıflayıcıları eğitilmiştir (doğrusal çekirdek işlevi (*linear kernel function*) Matlab® varsayılan parametreleriyle kullanılmıştır. Sınıflama aşaması tamamlanamadığında ise 1000 adımlı ikilenik maliyet işlevi (*quadratic cost function*) kullanılmıştır). Deneme örnekleri her tanımlayıcıya ait sınıflayıcılara ayrı ayrı verilmiştir. Eğer örnek bu sınıflayıcılardan birine aitse, ilgili derecenin oyu bir artırılmıştır. Bu çalışmada her sınıflayıcının aynı ağırlığa sahip olduğu varsayılmıştır.

SVM sınıflayıcılarından sonra yine her bir radyografik tanımlayıcının her bir derecesi için dışbükey noktalar kümesi belirlenmiştir. Eğer daha önceden indirgenmiş bir derece ise bu işlem boyut indirgemede yapıldığı için, işlem tekrarlanmamıştır. Kodell ve diğerleri [102] ve Nalbantov ve diğerleri [103] dışbükey örtü kullanımının sınıflama başarımlarına katkısı olduğunu göstermişlerdir. Deneme örneklerinin, ilgili derecenin dışbükey örtüsü içinde olup olmadığı denetlenmiştir. Eğer bu bölgenin içinde ise oylama sonucu bir artırılmaktadır, eğer değilse oylamaya etki etmemektedir. Bir örnek için takım sınıflayıcı ve dışbükey yönteminden elde edilen sonuç oylama matrisi Çizelge 7.1’de gösterilmiştir.

Çizelge 7.1 Bir örnek için her bir radyografik tanımlayıcıya ait oylama sonuç matrisi.

Oylama Matrisi	Sınıf değerleri (1-6)					
	(1)	(2)	(3)	(4)	(5)	(6)
Subtlety	1	1	0	0	2	-
Calcification	0	0	1	0	1	2
Sphericity	0	1	1	0	0	-
Margin	1	1	1	1	0	-
Lobulation	1	1	1	1	1	-
Spiculation	0	1	1	1	1	-
Texture	1	2	1	1	2	-

7.1.2. Takım Sınıflama Aşaması

Oylama aşamasından sonra, nodülün habislik derecesini (*malignancy*) belirlemek için kural tabanlı sistem kurulmuştur. Radyologların derecelendirmelerini kullanarak habislik için bir kural kümesi oluşturulmuştur. Kural kümesinin birkaç satırından oluşan örnek satırlar Çizelge 7.2’de gösterilmiştir. Kural kümesinde, en az üç radyoloğun habislik tanımlayıcısı üzerinde anlaştığı değer ve diğer radyografik tanımlayıcıların değerleri bulunmaktadır. Diğer tanımlayıcılar üzerinde bir anlaşma beklenmemiştir.

Çizelge 7.2 Malignancy değeri ve diğer radyografik tanımlayıcıların olduğu örnek kural tablosu.

Sub.	Cal.	Sph.	Mar.	Lob.	Spi.	Tex.	Mal.
5	3	4	5	2	1	5	1
5	3	5	5	1	1	5	1
5	6	4	2	4	5	5	5
.
.
.
4	6	4	5	1	1	5	3
4	6	5	5	1	1	5	3
4	6	4	4	2	2	5	4

Ek olarak, kural tabanlı sınıflama sistemine girdi olarak verilecek deneme kümesi olarak kullanmak üzere, kural kümesinin her satırına ait nodüller için alt düzey görüntü öznelikleri de hesaplanmıştır. Bu öznelikler takım sınıflayıcılara girdi olarak verilmiştir. Böylece bu örnekler için oylama matrisleri hesaplanmıştır. Daha sonra bu oylama matrisini

kullanan kural tabanlı sistem (7.1)'i kullanarak her kural satırı için bir değer hesaplamaktadır. Yöntemin adımları Şekil 7.3'deki algorithmada verilmiştir.

$$P(j,) = \{ \sum_{i=1}^7 V(i, R(j, i)) , mr(j) \} \quad (7.1)$$

Eşitlik (7.2)'de V her bir satırı bir radyografik tanımlayıcı ait 7×6 'lık oylama matrisidir (Çizelge 7.1'de örnek matris verilmiştir). R , r satıra sahip $r \times 7$ boyutta kural tablosudur. j , R içindeki hesaplanan sıradaki kural sayısıdır. $P(j)$ eşitlik (7.2)'den hesaplanan değeri ve o j 'nci kuralın habislik derecesini (mr) tutan değerdir. Her bir kural için eşitlik (7.2)'deki değer hesaplandıktan sonra P dizisi azalan değerde sıralanmıştır ve ilk N kuralda en çok rastlanan habislik derecesi sistemin sonucu olarak kabul edilmiştir. N isteğe bağlı herhangi bir değer olabilir. Bu yöntem için, en iyileştirme çalışmalarında bu değer 50 olarak bulunmuştur. Şekil 7.4'de yöntem için bir çalıştırma örneği gösterilmiştir.

Girdi:

t : habislik derecesi bulunacak örneğin öznitelik vektörü

tm : örneğin gerçek habislik derecesi

N : habislik tahmini için kural sayısı

R : kural matrisi, (her r satırda 7 tanımlayıcı)

mr : her kural için habislik derecesi listesi

f : nitelik sayısı

Başla:

Oylama matrisi V 'yi ilkle

Döngü $i = 1$ 'den 7 'ye (her tanımlayıcı için)

Her takım sınıflayıcıdan oy al, $V(i) = \text{takımSınıflayıcı}(t, S_i, c_i, f)$

Döngü bitir

Döngü $i=1$ 'den r 'ye

Her kural için puan hesapla, $P(i, 1) = \sum_{j=1}^7 V(i, R(i, j))$

Gerçek habislik derecesini ekle, $P(i, 2) = mr(j)$

Döngü bitir

P 'yi sırala ($P(,1)$ üzerinden)

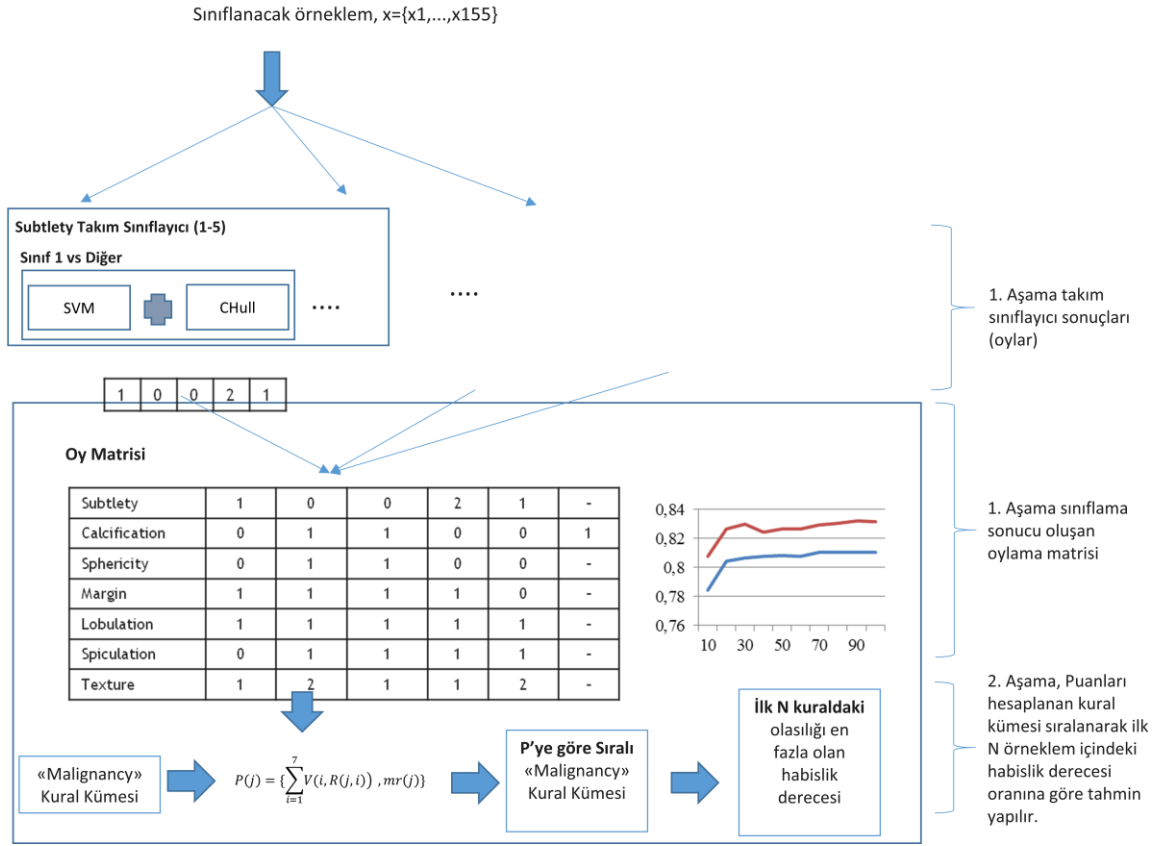
P içindeki ilk N kural içinde en çok tekrarlı değeri (fm) bul

Eğer $tm = fm$ ise $out = \text{doğru}$ değilse $out = \text{yanlış}$

Bitir**Çıktı:**

out : sınıflama sonucu

Şekil 7.3 Kural tabanlı sistemin algoritması



Şekil 7.4 Yöntem için çalıştırma örneği gösterimi.

7.1.3. Sonuçlar

7.1.3.1. Takım Sınıflayıcı Sonuçları

Yüksek oranda dengesiz dağılmış veri kümelerinde, düşük oranda ifade edilmiş sınıflar tekli sınıflayıcılarda başarısız sonuçlar vermektedir. Örnek olarak *subtlety* radyografik tanımlayıcısı için örnek SVM sınıflayıcı sonuçları Çizelge 7.3'te verilmiştir. 1 ve 2 derecelerinin veri kümesindeki ağırlığı %1'in altındadır. Baskın olan 5 derecesi içinse bu oran %70'in üzerindedir. Tablo 4'den görüleceği üzere düşük oranlı sınıfın hatalı sınıflama oranı fazladır.

Çizelge 7.3 Subtlety için SVM ile sınıflama sonuçları.

	1	2	3	4	5	Örnek sayısı	Örnek oranı
1	0	0	9	1	5	15	<0,01
2	0	0	5	4	0	9	<0,01
3	0	0	60	64	7	131	~0,09
4	1	0	32	146	92	271	~0,19
5	1	0	9	75	945	1030	~0,71

Bu yöntemde oylama mekanizması olarak kullanılan takım sınıflayıcılar, tekil sınıflayıcıların aksine sonuç olarak her bir dereceye ait değerlerin bulunduğu vektörler döndürmektedir. Bu oylama sistemi düşük oranlı sınıfların ifade edilebilirliğini artırmaktadır. Çizelge 7.4'te, sublety tanımlayıcısı için en yüksek oylama sonuçları gösterilmiştir. Tablodaki diagonal eksende kalın yazıyla verilmiş değerler ilgili sınıfın en yüksek aldığı oylama sayısıdır. Diğer değerler ise farklı dereceye ait sınıflayıcılardan en yüksek alınan oylama sayısıdır. Bu durum belirsizliğe neden olmaktadır. Çünkü takım sınıflayıcılardan elde edilen sonuçlarda bir örnek bir radyografik tanımlayıcı için birden fazla derecede en yüksek oylama sonucuna ulaşabilir. Örneğin Çizelge 7.1'de sublety takım sınıflayıcısı 5 derecesi için en yüksek sonuçla olan 2'yi döndürmüştür. Fakat *lobulation* takım sınıflayıcısı her derece için 1 sonucunu vermiştir. Bu sınıflama aşaması için yüksek oranda belirsizlik taşıyan bir durumdur. Bu radyografik tanımlayıcının (bu örnek için) habislik derecesini belirlemede ayırt edici bir yanı yoktur. Fakat *sublety*, *calcification* ve *texture* tanımlayıcıları daha ayırt edici sonuçlara sahiptir.

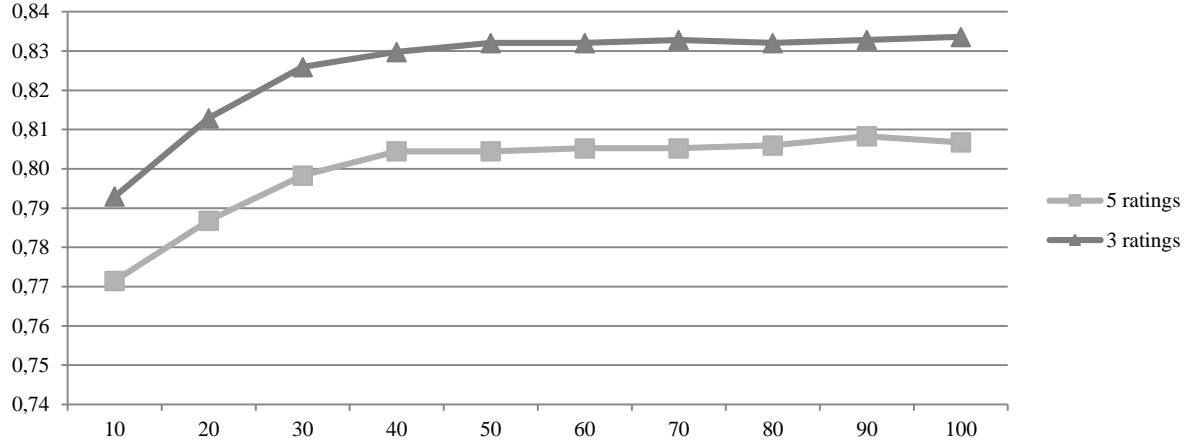
Çizelge 7.4 Sublety için en yüksek oylamayı kazanan sınıflar

	1	2	3	4	5	Örnek sayısı	Örnek oranı
1	14	10	2	5	0	15	<0,01
2	6	9	1	1	0	9	<0,01
3	42	38	53	46	1	131	~0,09
4	111	115	87	179	22	271	~0,19
5	395	383	16	125	793	1030	~0,71

7.1.3.2. Kural Tabanlı Sınıflama Sonuçları

Yöntemin başarımının ölçülmesi için bir örneği dışarıda bırakma genelleştirme yaklaşımı uygulanmıştır. Habislik veri kümesinden bir örneklem deneme verisi olarak ayrılmış ve bu örneğe ait kural satırı ise kural kümesinden çıkarılmıştır. Geri kalan örneklerin kuralları sistem için kural tablosunda kullanılmıştır. Her bir örneğin (*nodül*) hasta kodu (*case id*) ve nodül kodu (*nodule id*) bulunmaktadır. Deneme örneği olarak kullanılacak nodülle aynı kodlara sahip olan nodüller (her radyografik tanımlayıcıda farklı sayıda nodül bulunmaktadır, aynı nodül birden fazla kümede mevcut olabilir) radyografik tanımlayıcıların eğitim kümesinden de çıkarılmıştır. Bu işlem habislik kümesindeki her bir örnek, bir kez deneme örneği olana kadar her biri için tekrarlanmıştır.

Şekil 7.2’teki algoritma için, en uygun kural boyunu (N) farklı kural uzunluklarında hesaplamalar tekrarlanarak bulunmuştur. Habislik kümesindeki derecelerin örnek sayıları da göz önünde bulundurularak kural boyutu 10 ila 100 arasında denetlenmiştir. Şekil 7.5’te görüleceği üzere en uygun kural sayısına 50 ila 60 arasında erişilmiştir.



Şekil 7.5 Kural boyunun sınıflama sonucuna etkisini gösterir grafik.

Son olarak, sunulan yöntem ile tekil sınıflayıcılar sınıflama başarımı, duyarlılık ve seçicilik metrikleri ile karşılaştırılmıştır. Habislik tanımlayıcısının 5 derecesi vardır: Yüksek olasılıkla iyi huylu (*highly unlikely-1-*), orta olasılıkla iyi huylu (*moderately unlikely-2-*), belirsiz (*indeterminate-3-*), orta olasılıkla kötü huylu (*moderately suspicious-4-*), yüksek olasılıkla kötü huylu (*highly suspicious-5-*). Sınıflama başarımları 5 sınıflı habislik kural kümesi için Çizelge 7.5’te gösterilmiştir.

Çizelge 7.5 Yöntemlerin sınıflama başarımı (5 sınıflı).

Yöntem	5 Sınıf SB	5 Sınıf Duy.	5 Sınıf Seç.
LDA	0.6610	0.3832	0,8886
Adaboost	0.7254	0.6737	0,9311
Naive Bayes	0.7250	0,6700	0,9325
kNN	0.7665	0.6314	0.9350
SVM	0.7828	0.5384	0.9345
RF	0.8040	0.5797	0.9421
Önerilen Yöntem	0.8044	0.5526	0.9403

Habislik derecelerindeki belirsizlik sebebiyle, bu dereceleri (1,2) iyi huylu, (3) belirsiz , (4,5) kötü huylu olarak gruplanarak deney tekrarlanmıştır. Çizelge 7.6’da görüleceği üzere gruplamadan sonra tüm yöntemlerin sınıflama başarımı artmıştır.

Çizelge 7.6 Yöntemlerin sınıflama başarımı (3 sınıflı).

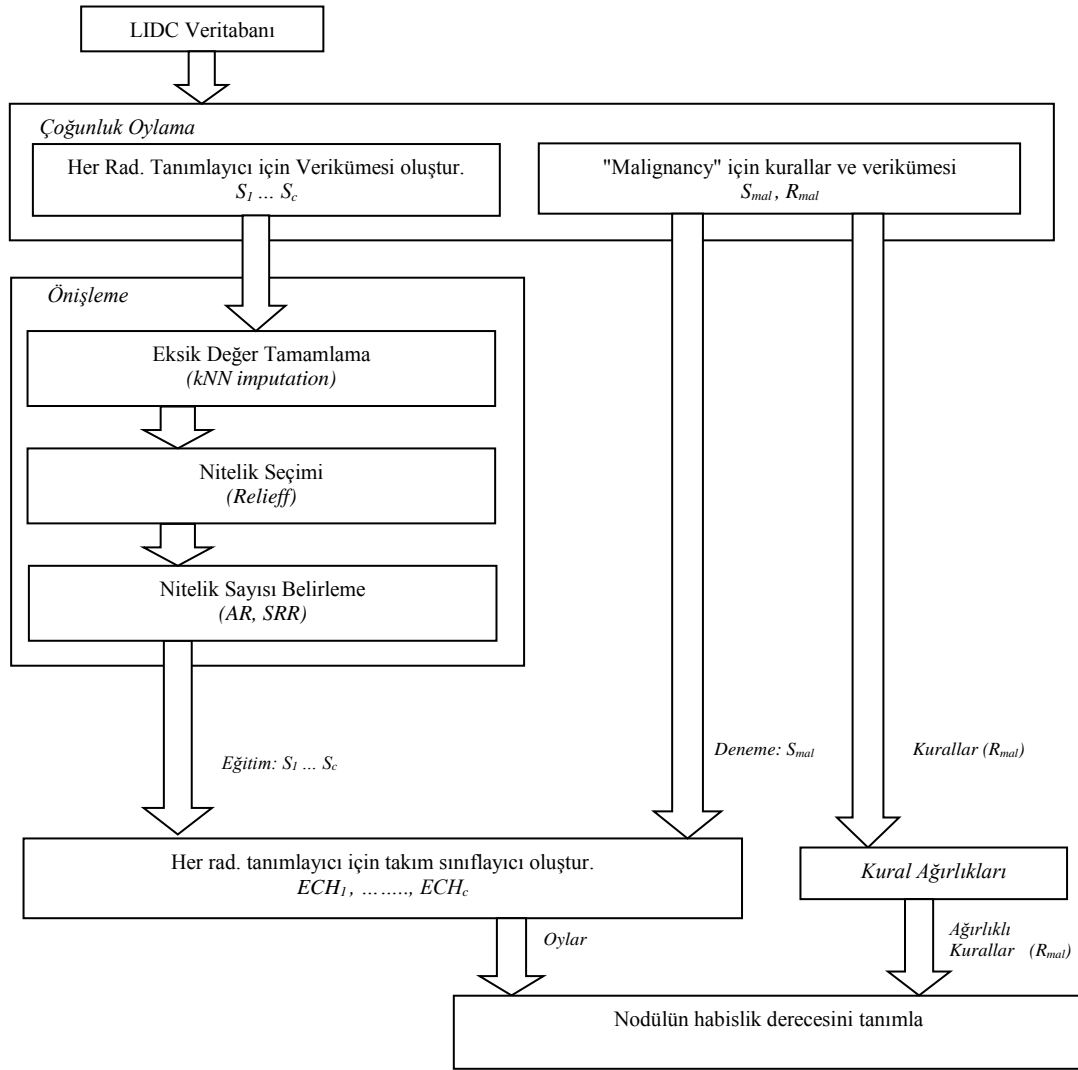
Yöntem	3 Sınıf SB	3 Sınıf Duy.	3 Sınıf Seç.
LDA	0.8200	0.8001	0.9006
Adaboost	0.8135	0.8265	0.9055
Naive Bayes	0.8100	0.8250	0.9060
kNN	0.7959	0.8003	0.8930
SVM	0.8005	0.8108	0.9021
RF	0.8151	0.8161	0.9035
Önerilen Yöntem	0.8328	0.8234	0.9118

7.1.4. Tartışma

Bu çalışmada küçük boyutlu akciğer nodüllerinin habislik derecesinin tespiti için kural tabanlı bir sınıflama yaklaşımı geliştirilmiştir. SVM sınıflayıcısı ve dışbükey örtü noktaları kullanılarak her radyografik tanımlayıcı için takım sınıflayıcılar oluşturulmuş, bu sınıflayıcıların sonuçları kural tabanlı sınıflama sistemi için girdi olarak kullanılmıştır. Kural tabanlı sınıflama sistemine ek girdi olarak LIDC veri tabanından radyologların habislik üzerinde ortak kararları ile elde edilmiş radyografik özniteliklerden oluşan kurallar kümesi de verilmiştir. Her radyografik tanımlayıcı için ayrı veri kümeleri oluşturulmuştur. Bu veri kümelerinin birçoğu dengesiz dağılmıştır. Bu dengesiz dağılım sorunu veri kümesi dengeleme yaklaşımı ve sınıfa özel boyut indirgeme ile çözülmeye çalışılmıştır. Alınan sonuçlar ışığında takım sınıflayıcılar ile kural tabanlı sınıflama sistemi üzerine oluşturduğumuz yöntemin tekil sınıflayıcılardan daha iyi sonuç verdiği gözlemlenmiştir. Habislik değerini gruplamak sınıflama başarımlarını tüm sınıflayıcılar nezdinde arttırmıştır. Gruplama yapıldığında, LDA tüm sınıflayıcılar içinde en çok artış göstereni olmasına rağmen, sunulan yöntem en iyi sınıflama sonucunu vermiştir.

7.2. Takım Sınıflayıcılar Ve Ağırlıklı Kurallar İle Sınıflama Yöntemi

Bu çalışmada, akciğer nodüllerinin habislik derecesinin belirlenmesi için ağırlıklı kural tabanlı bir sınıflama yöntemi sunulmuştur. Önceki yönteme ek olarak her bir tanımlayıcı için, ortalama dereceler (*average ranks*) ve başarı oran orantıları (*success rate ratios*) kullanılarak farklı sayıda ve farklı türde öznelik kümeleri oluşturulmuştur. Takım sınıflayıcıların çıktıları, ağırlıklı kural tabanlı sınıflayıcıya girdi olarak verilmiştir. Nodülün habislik derecesi ve radyografik tanımlayıcılar arasındaki doğrusal ilişki korelasyon çözümlemesi kullanılarak belirlenmiştir. Ayrıca, radyologların aynı nodüller üzerindeki değerlendirmeleri de çözümlenerek, her bir kural içinde uzmanlar arasındaki anlaşmanın derecesi tespit edilmiştir. Sunulan yöntemin genel çizeneği Şekil 7.6'de verilmiştir.



Şekil 7.6 Yöntemin genel akış çizeneği.

7.2.1. Öznitelik Sayılarının Belirlenmesi

Seçilecek öznitelik sayıları isteğe bağlı olarak veya başarı ve maliyet ihtiyaçlarına göre de belirlenebilir. Seçilecek özniteliklerin, özellikle az ifade edilmiş sınıfların ayırt edilmesinde faydalı olacağı öngörülmektedir. Bu işlem sınıflama işlemi ve dışbükey örtü işleminin hesaplama maliyetlerini, bilgi kaybı olmaksızın, azaltmaktadır.

Her bir radyografik tanımlayıcı veri kümesinde, öncelikle özniteliklerin önemlerine göre derecelendirilmesi işlemi, Relieff [62] yöntemi kullanılarak yapılmıştır. Relieff yöntemi, özniteliklere ağırlıklar atayarak, en önemliden en az önemliye doğru sıralamaktadır. Daha sonra Brazdil ve Sores'in [81] ölçüm yöntemleri kullanılarak uygun öznitelik kümeleri saptanmıştır. Bunlar, ortalama dereceler (*average ranks-AR*) ve başarı oran orantıları (*success rate ratios-SRR*) ve önemli kazanım (*significant wins-SW*) yöntemleridir. Bu çalışmada, sadece SRR ve AR yöntemleri kullanılmıştır. Bu yöntemler, hangi sınıflama yönteminin daha başarılı olduğunu tespit etmekte kullanılmaktadır. Bu çalışma kapsamında, SVM algoritması için hangi öznitelik sayısının daha anlamlı olduğunu bulmakta kullanılmıştır.

AR yönteminde, sınıflama algoritmaları ölçülen hata oranlarına göre sıralanır. En iyi algoritma, en düşük ortalama hata oranına sahip olan algoritmadır. r_j^i , deneme kümesi i üzerinde öznitelik küme sayısı j 'nin oranı ve n rastgele oluşturulan deneme kümesi sayısıdır. Ortalama dereceler, eşitlik (7.2)'deki gibi hesaplanmaktadır.

$$r_j = (\sum_i r_j^i)/n \quad (7.2)$$

SRR yönteminde, ikili algoritmalar arasındaki başarı oranı ölçülmektedir. ER_j^i , deneme kümesi i üzerinde öznitelik küme sayısı j 'nin hata oranı olmak üzere; öznitelik küme sayısı j 'nin, öznitelik sayısı k 'ya deneme kümesi i 'deki avantajı $SRR_{j,k}^i$, eşitlik (7.3)'deki gibi hesaplanır.

$$SRR_{j,k}^i = (1 - ER_j^i) / (1 - ER_k^i) \quad (7.3)$$

Avantajlar hesaplandıktan sonra, öznitelik sayıları j ve k için ikili ortalama başarı oran orantıları (*pairwise mean success rate ratio*) $SRR_{j,k}$ eşitlik (7.4)'teki gibi hesaplanır. n rastgele deneme kümelerinin sayısıdır. Bu oran ikililer için genel avantaj/dezavantajın değerlendirilmesi için kullanılmaktadır.

$$SRR_{j,k} = (\sum_i SRR_{j,k}^i)/n \quad (7.4)$$

Son olarak, genel ortalama başarı oran orantıları SRR_j her bir öznitelik sayısı j için eşitlik (7.5)'teki gibi hesaplanır. m denenen farklı öznitelik küme sayısıdır.

$$SRR_j = (\sum_k SRR_{j,k}) / (m - 1) \quad (7.5)$$

Yukarıda tanımlanan yöntemler farklı öznitelik sayılarıyla oluşturulmuş aynı sınıflama yöntemi üzerinde uygulanmıştır (3,4,5,6,7,8,9,10,20,40,60). Her bir radyografik tanımlayıcı için, veri kümesinin %10'u deneme verisi, kalanı ise eğitim verisi olarak ayrılmıştır. Eğitim ve deneme verisi rastgele seçim ile (*bootstrap*) seçilmiştir ve bu seçim 50 kere tekrarlanmıştır. Her bir veri kümesi, her sınıflayıcıda kullanılmıştır. Çizelge 7.7'de, her radyografik tanımlayıcının her bir öznitelik sayısı için elde edilmiş ortalama AR ve SRR değerleri verilmiştir. En anlamlı öznitelik sayısı kalın yazı tipi ile işaretlenmiştir.

Çizelge 7.7 (3-10, 20,30,40,50,60) öznitelik sayıları için AR ve SRR ölçümleri. Her satırdaki ölçümler için en önemli değerler kalın yazı tipi ile işaretlenmiştir.

Rad. Tan.		3	4	5	6	7	8	9	10	40	60
Calcification	AR	92,65	92,50	87,10	75,85	69,40	58,90	56,35	132,55	145,90	129,25
	SRR	1,0821	1,0830	1,0851	1,0898	1,0919	1,0947	1,0972	1,0672	1,0630	1,0676
Lobulation	AR	77,80	68,35	33,70	33,10	18,10	146,80	115,75	126,70	179,65	179,35
	SRR	1,1535	1,1564	1,2151	1,2172	1,2421	1,0082	1,0596	1,0495	0,9913	0,9916
Margin	AR	128,40	119,20	109,80	105,90	156,40	154,60	148,45	109,80	104,50	89,65
	SRR	1,0643	1,0680	1,0735	1,0749	1,0543	1,0549	1,0578	1,0735	1,0761	1,0825
Sphericity	AR	14,05	15,40	18,55	109,80	91,30	92,65	141,40	150,40	142,00	142,30
	SRR	1,8496	1,8380	1,8287	1,0322	1,0486	1,0501	0,9569	0,9424	0,9625	0,9627
Spiculation	AR	167,40	163,30	163,30	29,65	36,10	31,15	31,45	30,25	119,95	163,30
	SRR	1,0270	1,0290	1,0290	1,1686	1,1651	1,1674	1,1671	1,1680	1,0426	1,0290
Sublety	AR	12,40	4,90	80,35	62,65	106,5	96,25	75,25	67,15	192,85	187,75
	SRR	1,5351	1,5503	1,1909	1,2138	1,1268	1,1465	1,1770	1,1880	0,8726	0,8775
Texture	AR	40,90	39,55	7,15	28,45	47,65	122,50	104,20	80,65	201,85	194,95
	SRR	1,2643	1,2655	1,3357	1,2846	1,2499	1,0490	1,0781	1,1180	0,9020	0,9245

Sınıflayıcılar için en uygun öznitelik sayısını bulduktan sonra, takım sınıflayıcı ve veri dengemeleme aşamalarında kullanılacak dış bükey örtü yöntemi için de en uygun öznitelik sayısı tespit edilmelidir. Dış bükey örtü yöntemi, zaman gereksinimi yüksek olan bir yöntemdir. Karmaşıklığı örnek sayısı ve verikümesi boyutu ile artmaktadır. Bu yöntem için Barber ve diğerlerinin önerdiği *Quick Hull* [91] algoritması kullanılmıştır. *Quick Hull* algoritması yüksek boyutlu verileri, benzer amaçlı algoritmalarından daha etkin olarak hesaplamaktadır, fakat yine de dokuz boyut üzerindeki verilerde kayda değer bir işlem süresi almaktadır. İstenen, en makul hesaplama süresinde bilgi kaybını en aza indirecek

öznitelik sayısına erişmektir. Çizelge 7.8'te görüleceği üzere, altı boyuttan sonra algoritmanın hesaplama süresi üssel olarak artmaktadır. Bu sebeple, Relieff algoritması ile belirlenen en önemli altı öznitelik, dış bükey örtüyü hesaplarken kullanılmıştır.

Çizelge 7.8 Quick Hull algoritmasının her radyografik tanımlayıcı için en önemli 3-9 öznitelik için işlem süresi verilmiştir.

	3	4	5	6	7	8	9
Calcification	0,10	0,04	0,05	0,59	5,40	56,60	787,49
Lobulation	0,80	0,77	0,76	1,04	3,71	27,78	452,09
Margin	0,19	0,19	0,20	0,29	2,64	19,21	95,56
Sphericity	0,07	0,08	0,13	0,33	1,44	8,90	50,62
Spiculation	0,38	0,38	0,45	0,72	3,26	22,09	291,25
Sublety	0,53	0,53	0,55	0,78	4,20	57,76	315,04
Texture	0,53	0,53	0,55	0,88	6,49	50,80	760,84

Girdi:

t : örneğin nitelik vektörü.

S : rad. tanımlayıcı için veri kümesi n örnek, m öznelik, $S = \{x_1, \dots, x_m, y\}^{1..n}$

c : rad. tanımlayıcı için en büyük sınıf değeri.

f_S : AR ve SRR yöntemlerinden elde edilen nitelik sayısı

f_{CH} : dışbükey örtü için nitelik sayısı

Başla:

Döngü $i=1$ 'den c 'ye

Sınıf i için veri kümesi oluştur (i vs diğerleri) $S_i = \{x_1, \dots, x_m, y_i = \{0,1\}\}^{1..n}$

Nitelik seçimi $F_i = \text{relieff}(S_i)$

$$\text{Veri dengele, } S'_i = \begin{cases} \text{arttır, } & \text{size}(S_i(y=1)) < \frac{\text{size}(S)}{2c} \\ \text{azalt, } & \text{size}(S_i(y=1)) > \frac{3\text{size}(S)}{2c} \\ S_i(y=1), & \text{diğer} \end{cases}$$

Taban sınıflayıcıyı eğit $T_i = \text{TabanSınıflayıcı}(S'_i, F_i (1 \dots f_S))$

S'_i nin dışbükey örtü noktalarını bul, $H_i = \text{DışbükeyÖrtü}(S'_i(y=1), F_i (1 \dots f_{CH}))$

Döngü bitir

Oylama vektörü v ilkle, $\forall i: v(i) = 0$

Döngü $i=1$ 'den c 'ye

Eğer sınıfla $(t, T_i) = 1$ ise $v(i) = v(i) + 1$

Eğer dışbükeyörtüiçindemi $(H_i, t) = 1$ ise $v(i) = v(i) + 1$

Döngü bitir

Bitir**Çıktı:**

v : oylama vektörü

Şekil 7.7 Bir örnek için takım sınıflayıcı adımlarını gösteren algoritma.

7.2.2. Takım Sınıflayıcılar

Veri kümesi ön işleme adımlarından sonra, takım sınıflayıcılar Şekil 7.7'de gösterilen algoritma ile oluşturulmuştur.

7.2.3. Ağırlıklı Kural Tabanlı Sistem

Oylama aşamasından sonra, ağırlık kural tabanlı sistem ile nodülün habislik derecesi tahmin edilmiştir. Bu sistem, habislik veri kümesinden elde edilen kuralları kullanarak eğitilmektedir. Radyologların değerlendirmeleri kullanılarak, nodülün habislik derecesi için bir kural kümesi oluşturulmuştur. Çizelge 7.9'te, kural kümesinden örnek satırlar verilmiştir. Kural kümesine, LIDC verisinden, habislik değeri üzerinde en az üç radyoloğun anlaşmış olduğu verilerin radyografik tanımlayıcıları çıkarılarak kural kümesine eklenmiştir.

Çizelge 7.9 Nodül habislik derecesi için (*malignancy – Mal.*) örnek kural tablosu

Sub.	Cal.	Sph.	Mar.	Lob.	Spi.	Tex.	Mal.
5	3	4	5	2	1	5	1
5	3	5	5	1	1	5	1
5	6	4	2	4	5	5	1
.
.
.
4	6	4	5	1	1	5	3
4	6	5	5	1	1	5	3
4	6	4	4	2	2	5	3

Nodül habislik (*Malignancy*) veri kümesi, diğer radyografik tanımlayıcılar gibi habislik derecesi üzerinden çoğunluk oylama ile oluşturulmuştur, fakat diğer tanımlayıcı değerlerindeki de anlaşmaya varılmış olması beklenmemiştir. Diğer tanımlayıcılar üzerinde anlaşma beklenmediği için, her tanımlayıcının önemini belirleyecek katsayılar tanımlanmıştır. Her bir kuralda, her bir tanımlayıcı için farklı katsayılar atanmıştır. Örneğin, Çizelge 7.10'da aynı nodüle ait üç kural (her biri ayrı radyoloğun değerlendirmesi) gösterilmiştir. $rule_{i-1}$, $rule_i$, ve $rule_{i+1}$ üç ayrı radyoloğun değerlendirmesidir ve anlaşmaya vardıkları habislik değeri 5'tir. İki tanesi *calcification* tanımlayıcısında anlaşmıştır ve anlaşma katsayıları $2/3$ 'tür; diğer radyoloğun ise $1/3$ 'tür. Tüm radyologların anlaştığı *texture/subtlety* değerlerinde bu katsayı $3/3$ 'tür, vs.

Çizelge 7.10 Habislik kümesinden elde edilmiş, aynı nodule ait değerlendirme olan üç kuralın değerleri ve anlaşma katsayıları.

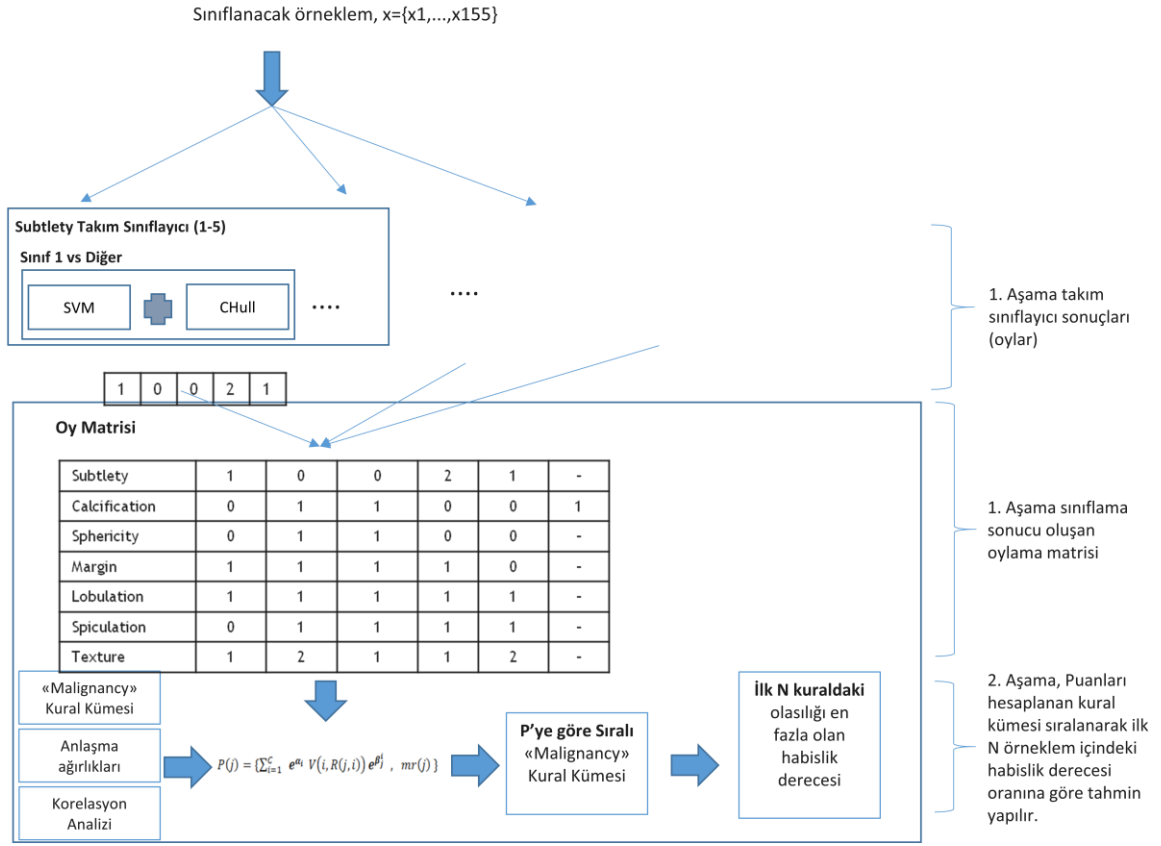
R. Tanımlayıcı	<i>rule_{i-1}</i>		<i>rule_i</i>		<i>rule_{i+1}</i>	
	değer	Anlaşma k.	değer	Anlaşma k.	değer	Anlaşma k.
<i>calcification</i>	6	2/3	6	2/3	3	1/3
<i>sphericity</i>	3	2/3	4	1/3	3	2/3
<i>lobulation</i>	5	1/3	1	2/3	1	2/3
<i>texture</i>	5	3/3	5	3/3	5	3/3
<i>margin</i>	5	1/3	3	2/3	3	2/3
<i>subtlety</i>	2	3/3	2	3/3	2	3/3
<i>spiculation</i>	1	1/3	3	1/3	4	1/3
<i>malignancy</i>	5	-	5	-	5	-

Ek olarak, habislik derecesi ve diğer radyografik tanımlayıcılar arasındaki ilişkiyi ölçmek için korelasyon analizi yapılmıştır. Bu analiz sonucu elde edilen korelasyon katsayıları, tanımlayıcılar arasındaki doğrusal ilişkiyi ölçmektedir. Büyük oranda korelasyona sahip tanımlayıcılar hesaplamalarda baskın gelebilir. Bu yüzden katsayıların etkileri Eşitlik (7.6)'da belirtildiği üzere, üstel/üssel işlev (*exponential function*) kullanılarak yumuşatılmıştır.

$$P(j) = \{ \sum_{i=1}^C e^{\alpha_i} V(i, R(j, i)) e^{\beta_j^i}, mr(j) \} \quad (7.6)$$

Eşitlik (10.6)'da C , radyografik tanımlayıcıların sayısıdır (*malignancy* hariç), V , $C \times 6$ boyutlarında oylama matrisidir ve bu matriste her satır ilgili tanımlayıcı için elde edilen oylama sonuçlarını tutar. R , $r \times C$ boyutlarında kural matrisi ve r kural sayısıdır. j , herhangi bir kuralın indisidir. β_j^i , j . kuraldaki i . tanımlayıcı anlaşma katsayısıdır. $P(j)$, iki değer tutar: $P(j,1)$ j . kuralın ağırlıklı oy toplamı ve $P(j,2)$ j . kuralın habislik derecesi (mr). Daha sonra P 1. sütununa göre azalan sırada sıralanır. Örneğin habislik derecesi olasılığı, ilk N kuraldaki habislik derecelerinin bulunma oranına göre hesaplanır.

Şekil 7.9'teki algoritmada, bir nodül örneği için ağırlıklı kural tabanlı sınıflama yaklaşımını göstermektedir. Algoritma ilk döngüde, Şekil 7.7'deki yöntemden her bir radyografik tanımlayıcı için elde edilmiş oyları alarak oy matrisini oluşturur. İkinci döngüde ise, Eşitlik (10.5) kullanılarak her bir kural satırı için ağırlıklı oy toplamları (P) hesaplanır. P sıralandıktan sonra ilk N eleman içindeki habislik derecesi olasılıkları hesaplanır. Sonuçta, en yüksek olasılığa sahip habislik derecesi tahmin edilen derece olarak verilir. Şekil 7.8'de yöntem için çalıştırma örneği gösterilmiştir.



Şekil 7.8 Yöntem için çalıştırma örneği gösterimi.

7.2.4. Sonuçlar

7.2.4.1. Takım Sınıflayıcı Sonuçları

Takım sınıflayıcı sonuçları, önceki yöntemde ayrıntılı olarak incelenmiştir. Yöntem 2. adımıyla farklılık gösterdiği için bu kesimde sonuçlara değinilmeyecektir.

Girdi:

t : habislik derecesi bulunacak örneğin öznitelik vektörü

C : rad. tanımlayıcı sayısı

tm : örneğin gerçek habislik derecesi

N : habislik tahmini için kural sayısı

R : kural matrisi, (her r satırda 7 tanımlayıcı)

mr : her kural için habislik derecesi listesi

α : korelasyon katsayısı

β : anlaşma oranı

f_S : AR ve SRR yöntemlerinden elde edilen nitelik sayısı

f_{CH} : dışbükey örtü için nitelik sayısı

Başla:

Oylama matrisi V 'yi ilkle

Döngü $i = 1$ 'den C 'ye (her tanımlayıcı için)

S_i : i . tanımlayıcı için veri kümesi

c_i : i . tanımlayıcı için en büyük sınıf değeri

Her takım sınıflayıcıdan oy al, $V(i) = \text{takımSınıflayıcı}(t, S_i, c_i, f_S, f_{CH})$

Döngü bitir

Döngü $j=1$ 'den r 'ye (her kural için)

Her kural için ağırlıklı puanı hesapla, $P(j, 1) = \sum_{i=1}^C e^{\alpha_i} V(i, R(j, i)) e^{\beta_j}$

Her kural için gerçek habislik derecesini ekle, $P(j, 2) = mr(j)$

Döngü bitir

P' 'yi sırala, $P' = \text{sırala}(P, 1)$

Döngü $i = 1$ 'den 5'e (her habislik derecesi için)

i . derece için olasılık değerini bul, $p(i) = (P'$ içinde ilk N kural içinde i say) / N

Döngü bitir

En yüksek olasılıklı dereceyi bul, $m = \{i: \max_{i=1, \dots, 5} p(i)\}$

Bitir

Çıktı:

p : örnein habislik derece olasılığı.

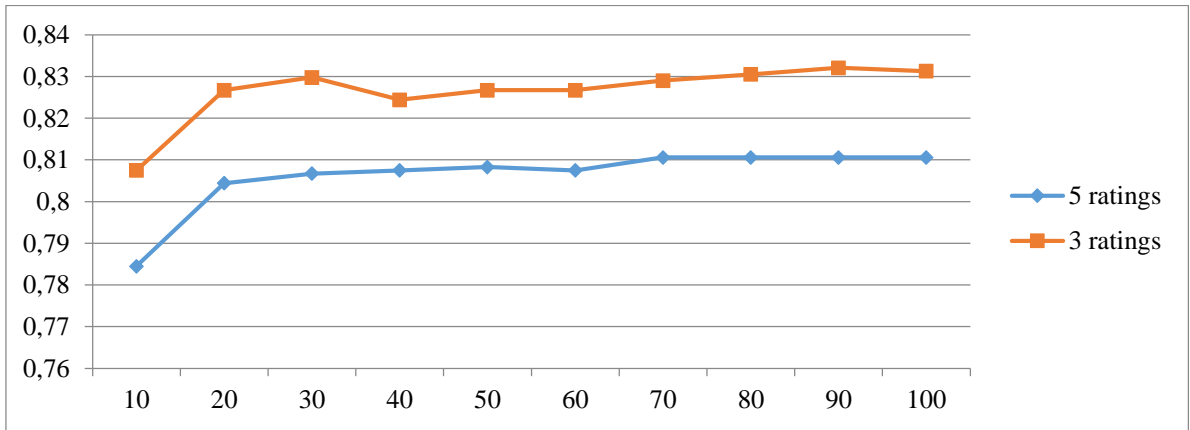
m : en yüksek olasılıklı habislik derecesi.

Şekil 7.9 Ağırlıklı kural tabanlı sistemin algoritması.

7.2.4.2. Ağırlıklı Kural Tabanlı Sınıflayıcı Sonuçları

Yöntemlerin değerlendirilmesi için tek örnek bırakarak çapraz geçerleme kullanılmıştır. *Malignancy* veri kümesinden bir örnek seçilmiş ve geçerleme verisi olarak kullanılmıştır. Bu örneğe ait kural satırı da kural kümesinden çıkarılmıştır. Kalan örneklere ait kurallar ise eğitim verisi olarak kullanılmıştır. Her örneğin bir vaka kodu (*case id*) ve nodül kodu (*nodule id*) vardır. Seçilen örnek eğer diğer radyografik tanımlayıcı veri kümelerinde de mevcutsa (kodlar aracılığıyla bakılır) bu veri kümelerinden de çıkarılır. Bu işlem, tüm örnekler bir kez geçerleme örneği olana kadar tekrarlanır.

Şekil 7.9'deki algoritma için en uygun (*optimum*) N değeri, ölçümler farklı kural boyutları için tekrarlayarak bulunur. Ölçek olarak 10 ile 100 arası seçilmiştir. Şekil 7.10'de verilen grafiğe göre en uygun kural sayısına 70 civarında ulaşılmaktadır – en az kural ile en yüksek sınıflama başarımı.



Şekil 7.10 Kural sayısının sınıflama başarımına etkisini gösteren grafik.

Ağırlıklı kural tabanlı sistemin başarımı ve farklı türde tekil sınıflayıcıların başarımı karşılaştırılmıştır: örnek tabanlı *k-nearest neighbor* (*kNN*), takım sınıflayıcı tabanlı *Adaboost*, işlev tabanlı *support vector machine* (*SVM*), *linear discriminant classifier* (*LDA*) ve *naive bayes* sınıflayıcıları. Bu sınıflayıcılar *malignancy* veri kümesindeki görüntü

öznitelikleri kullanılarak eğitilmiştir. Sınıfların denenmesi için tek örnek bırakmalı çapraz geçerleme yöntemi kullanılmıştır.

Nodül habislik veri kümesi *malignancy* beş sınıftan oluşmaktadır: oldukça düşük (*highly unlikely -1-*), orta ihtimal düşük (*moderately unlikely -2-*), belirsiz (*indeterminate -3-*), orta şüpheli (*moderately suspicious -4-*), oldukça şüpheli (*highly suspicious -5-*). 5 sınıflı *malignancy* veri kümesi için sınıflama başarımları Çizelge 7.11’da verilmiştir. Önerilen sınıflayıcı bu veri kümesinde %81.10 sınıflama başarımına ulaşmıştır. Diğer sınıflayıcılar arasında %80.40 ile RF en yüksek başarıma sahiptir. LDA ise %66.10 ile en düşük başarıma ulaşmıştır.

Çizelge 7.11 Yöntemlerin sınıflama başarımı (5 sınıflı veri için).

Yöntem	5 Sınıf SB	5 Sınıf Duy.	5 Sınıf Seç.
LDA	0.6610	0.3832	0.8886
Adaboost	0.7254	0.6737	0.9311
Naive Bayes	0.7250	0.6700	0.9325
kNN	0.7665	0.6314	0.9350
SVM	0.7828	0.5384	0.9345
RF	0.8040	0.5797	0.9421
Önerilen Yöntem	0.8110	0.5520	0.9410

Habislik veri kümesindeki sınıf değerleri (1,2)’yi düşük, (3)’ü belirsiz ve (4,5)’i ise şüpheli olarak gruplanarak, deney tekrar edilmiştir. Çizelge 7.12’da görüleceği üzere, tüm yöntemlerin sınıflama başarımı bu veri üzerinde daha yüksek çıkmıştır. Yöntem %83.20 sınıflama başarımı elde etmiştir. LDA ise önceki veri kümesine göre %82.00 sınıflama başarımı ile sonucu en çok yükselen yöntemdir. *Adaboost* ve *Naive Bayes* sınıflayıcıları da %81.35 ve %81.00 ile yüksek oranda gelişim göstermiştir. Habislik derecelerini gruplandırmak genel olarak sınıflama başarımını artırmıştır.

Çizelge 7.12 Yöntemlerin sınıflama başarımı (3 sınıflı veri için).

Yöntem	3 Sınıf SB	3 Sınıf Duy.	3 Sınıf Seç.
LDA	0.8200	0.8001	0.9006
Adaboost	0.8135	0.8265	0.9055
Naive Bayes	0.8100	0.8250	0.9060
kNN	0.7959	0.8003	0.8930
SVM	0.8005	0.8108	0.9021
RF	0.8151	0.8161	0.9035
Önerilen Yöntem	0.8320	0.8160	0.9100

7.2.5. Tartışma

Bu çalışmada Kesim 9'da verilen takım sınıflayıcılar ile sınıflama yönteminin ikinci adımı geliştirilmiştir. Her radyografik tanımlayıcı için SVM ve dışbükey örtü tabanlı takım sınıflayıcılar oluşturulmuştur. Takım sınıflayıcı çıktıları (oylar), kural tabanlı sistem için girdi olarak kullanılmıştır. Kural tabanlı sistemde kurallar, radyologların anlaşmaları ve radyografik tanımlayıcılar ve habislik derecesi arasındaki doğrusal ilişki dikkate alınarak ağırlıklandırılmıştır. Ek olarak, AR ve SRR yöntemleri ile her bir radyografik tanımlayıcı için farklı sayıda öznitelik küme boyu belirlenmiştir. Önceki yöntemle göre sınıflama başarımlarında artış gözlemlenmektedir.

7.3. Ağırlıklı Kural Tabanlı Sistemin Farklı Sınıflayıcılar İle Geliştirilmesi

Bu kesimde, "Takım Sınıflayıcılar ve Ağırlıklı Kurallar ile Sınıflama Yöntemi" başlığında önerilen yöntemde farklı taban sınıflayıcılar kullanarak yapılan ek analizler sunulacaktır. Önceki yöntemde, ilk sınıflama aşamasında taban sınıflayıcı olarak SVM kullanılmıştır. Taban sınıflayıcıya ek olarak dışbükey örtü yöntemi ile de oylama sonucu alınmıştır. Veri kümesi dengeleme yönteminde de dışbükey örtü ve SMOTE yönteminden faydalanılmıştır.

Ek analizlerde taban sınıflayıcı olarak RF, NB, Adaboost ve kNN sınıflayıcıları denenmiştir. Dışbükey örtü yönteminin (CH) hem sınıflamada hem de dengelemede olduğu ve olmadığı durumlar da değerlendirmeye alınmıştır. Tekil sınıflayıcılarla karşılaştırmada ise tekil RF yöntemi sonuçlara eklenmiştir.

Yapılan işlemler madde halinde verildiğinde:

- Kesim 7.2'de verilen yöntemde birinci sınıflama aşamasındaki takım sınıflayıcılar veri kümesi dengeleme ve oylama adımlarından oluşmaktadır.
- Veri kümesi dengeleme veri artırma (az olan sınıfın örneklerini SMOTE yöntemi ile çoğaltma) ve veri azaltma (çok olan sınıfın örneklerini dışbükey örtü yöntemi ile azaltma) adımlarından oluşmaktadır.
- Oylamada, tekil sınıflayıcılardan gelen (2-sınıflı SVM taban sınıflayıcısı ile) sınıflama sonucu ve örneğin dışbükey örtünün içinde olup olmasına göre belirlenen sonuç değerlendirilmektedir.
- Ek analizlerde dışbükey örtü (CH) aşamasının sınıflama başarısına etkisi incelenmektedir.
 - Veri dengeleme sadece veri arttırmanın olduğu ve arttırma/azaltmanın olduğu durumlar.
 - Oylamada sadece taban sınıflayıcının ve taban sınıflayıcı / dışbükey örtünün birlikte olduğu durumlar.
 - Taban sınıflayıcı için sadece SVM değil, ek olarak LDA, Adaboost, RF ve kNN yöntemlerinin de değerlendirmeye alındığı durumlar.

Çizelge 7.13'de sadece örnek arttırmanın olduğu ve oylamada her iki durumun ele alındığı sınıflama sonuçları verilmiştir. Çizelge 7.14'de ise dengelemede ve oylamada her iki durumun da ele alındığı sınıflama sonuçları verilmiştir.

Çizelge 7.13 Kesim 7.2'deki yöntem ek olarak yapılan analiz sonuçları. Sadece örnek artırma bulunmaktadır.

	5 sınıflı			3 Sınıflı		
	SB	Duy.	Seç.	SB	Duy.	Seç.
SVM	0,8067	0,5575	0,9420	0,8336	0,8259	0,9117
SVM + CH	0,8052	0,5539	0,9409	0,8275	0,8152	0,9081
RF	0,8252	0,5586	0,9474	0,8489	0,8311	0,9209
RF + CH	0,8144	0,5362	0,9419	0,8267	0,7907	0,9079
LDA	0,7853	0,5450	0,9349	0,8090	0,8082	0,9001
LDA + CH	0,7730	0,5269	0,9292	0,7845	0,7696	0,8860
Adaboost	0,7914	0,5433	0,9360	0,8098	0,8000	0,8999
Adaboost + CH	0,7776	0,5250	0,9308	0,7891	0,7685	0,8887
kNN	0,7224	0,4479	0,9128	0,7278	0,6701	0,8560
kNN + CH	0,7316	0,4609	0,9161	0,7385	0,6871	0,8620

Çizelge 7.14 Kesim 7.2'deki yöntem ek olarak yapılan analiz sonuçları. Örnek artırma ve azaltma bulunmaktadır.

	5 sınıflı			3 Sınıflı		
	SB	Duy.	Seç.	SB	Duy.	Seç.
SVM	0,8010	0,5450	0,9390	0,8210	0,8040	0,9050
SVM + CH	0,8110	0,5520	0,9410	0,8320	0,8160	0,9100
RF	0,8121	0,5421	0,9417	0,8275	0,7966	0,9078
RF + CH	0,7922	0,5145	0,9351	0,8060	0,7648	0,8964
LDA	0,7899	0,5433	0,9360	0,8129	0,8068	0,9018
LDA + CH	0,7669	0,5141	0,9268	0,7784	0,7558	0,8822
Adaboost	0,7868	0,5403	0,9352	0,8090	0,8012	0,8995
Adaboost + CH	0,7669	0,5142	0,9273	0,7784	0,7546	0,8821
kNN	0,7293	0,4715	0,9172	0,7370	0,6932	0,8627
kNN + CH	0,7278	0,4518	0,9142	0,7316	0,6726	0,8573

7.3.1. Sonuçlar

5 sınıflı ve 3 sınıflı verinin kullanıldığı deney sonuçları incelendiğinde RF tabanlı yöntemin %82.52 ve %84.89 sınıflama başarımı ile en başarılı yöntem olduğu görülmektedir. Duyarlılık ve seçicilik ölçümlerinde de bu durum aynıdır. LDA, kNN ve Adaboost tabanlı yöntemler daha başarısız sonuçlar üretmiştir.

7.3.2. Tartışma

Bu kesimdeki sonuçlar incelendiğinde, dışbükey örtü yaklaşımının kNN ve SVM yöntemlerine katkıda bulunduğu halde, diğer taban sınıflayıcılarda sınıflama sonuçlarına

olumsuz etki ettiđi grlmŖtr. DıŖbkey rt yaklaŖımının kullanılıp kullanılmayacađının seilecek taban sınıflayıcıya gre belirlenmesi daha dođru olacaktır. En iyi baŖarımı veren RF tabanlı sınıflayıcı sonuları dikkate alındıđında, yntemin zaman karmaŖıklıđını artıran dıŖbkey rt yaklaŖımı yntemden ıkarılabilir. İkinci en baŖarılı sonuca sahip SVM taban sınıflayıcısında ise olumlu katkısı grlmektedir. Ayrıca sonular Nalbantov ve diđerlerinin [103], dıŖbkey rtnn SVM sınıflayıcının baŖarısını arttırdıđı savını desteklemektedir.

8. DENEY ÇIKTILARININ ANALİZİ VE TARTIŞMA

Deneyle habislik derecesinin gruplandığı 3 sınıflı ve gruplanmadığı 5 sınıflı kümeler üzerinde gerçekleştirilmiştir. Genelleştirme yöntemi olarak tek örneği dışarda bırakarak çapraz geçerlilik yöntemi kullanılmıştır. Geliştirilen yöntemler doğrudan görüntü öznitelikleri üzerinde eğitilmiş farklı çeşitlerdeki tekil sınıflayıcılar ile karşılaştırılmıştır. Çizelge 8.1’de yöntemler için sınıflama başarımı, duyarlılık ve seçicilik ölçümleri verilmiştir. Bu çizelgeye literatürde LIDC veritabanı üzerinde yapılan habislik derecesi tahmini çalışmalarından edinilen sonuçlar da eklenmiştir.

8.1. Sınıflama Sonuçlarının Değerlendirilmesi

5 sınıflı habislik veri kümesi üzerinde yapılan deneylerde; Çizelge 8.1’deki verilere göre, iki katlı sınıflayıcı ve görüntü özniteliklerinin birlikte kullanıldığı yöntem (İK+GN), SVM tabanlı takım sınıflayıcı (AKT SVM+CH) ve RF tabanlı takım sınıflayıcı (AKT RF) sırasıyla %81.59, %81.10 ve %82.52 sınıflama başarımı ile en yüksek ilk üç yöntemdir. LDA ve Adaboost tabanlı takım sınıflayıcılar (AKT LDA ve Adaboost) %78.53 ve %79.14 sınıflama başarımı ile birçok tekil sınıflayıcıdan daha başarılı sonuç üretmelerine rağmen RF tekil sınıflayıcısından daha düşük başarıma ulaşmışlardır. RF tekil sınıflayıcısı ise %80.40 sınıflama başarımı ile en yüksek sonuca ulaşan tekil sınıflayıcıdır.

Duyarlılık ölçümü dikkate alındığında tekil NB, kNN ve Adaboost %67.00, %63.14 ve %67.37 ile en yüksek sonuca ulaşmıştır. Sunulan yöntemlerden İK+GN, KT SVM+CH ve AKT RF %56.65, %55.26 ve %55.86 ile en yüksek duyarlılık sonucuna sahiptirler. Bu ölçümde, tekil NB, kNN, Adaboost ve RF çalışmadaki yöntemlerden daha iyi başarıma ulaşmışlardır.

Seçicilik ölçümünde ise AKT RF, AKT RF+CH ve KT SVM+CH sırasıyla %94.74, %94.19 ve %94.03 ile sunulan yöntemler içinde en yüksek sonuca ulaşmışlardır. RF tekil sınıflayıcısı %94.21 ile bu yöntemlere en yakın tekil sınıflayıcıdır. Sunulan yöntemler ortalama olarak tekil sınıflayıcılardan daha başarılı olsalar da, genel olarak bu ölçümde yüksek farklar bulunmamaktadır.

Çizelge 8.1 Yöntemlerin sınıflama başarımları.

Yöntem	5 Sınıf SB	5 Sınıf Duy.	5 Sınıf Seç.	3 Sınıf SB	3 Sınıf Duy.	3 Sınıf Seç.
LDA	0.6610	0.3832	0,8886	0.8200	0.8001	0.9006
Adaboost	0.7254	0.6737	0,9311	0.8135	0.8265	0.9055
Naive Bayes	0.7250	0,6700	0,9325	0.8100	0.8250	0.9060
kNN	0.7665	0.6314	0.9350	0.7959	0.8003	0.8930
SVM	0.7828	0.5384	0.9345	0.8005	0.8108	0.9021
RF	0.8040	0.5797	0.9421	0.8151	0.8161	0.9035
Zinovev v.d. (2009)	0.7251 ± 0.18	-	-	-	-	-
Zinovev v.d. (2011)	0.5900 ± 0.04	-	-	-	-	-
Vinay v.d. (2014)	0.7828 ± 0.06	-	-	-	-	-
İK	0.7791	0,5171	0.9285	0.8006	0.7863	0.8945
İK + GN	0.8159	0.5665	0.9384	0.8221	0.8219	0.9096
BK Üİİ	0.7531	0.4904	0.9218	0.7615	0.7249	0.8725
ANFIS	0.7607	0.5053	0.9250	0.7715	0.7438	0.8787
KT SVM + CH	0.8044	0.5526	0.9403	0.8328	0.8234	0.9118
AKT SVM + CH	0.8110	0.5520	0.9410	0.8320	0.8160	0.9100
AKT SVM	0.8010	0.5450	0.9390	0.8210	0.8040	0.9050
AKT RF	0.8252	0.5586	0.9474	0.8489	0.8311	0.9209
AKT RF + CH	0.8144	0.5362	0.9419	0.8267	0.7907	0.9079
AKT LDA	0.7853	0.5450	0.9349	0.8090	0.8082	0.9001
AKT LDA + CH	0.7730	0.5269	0.9292	0.7845	0.7696	0.8860
AKT Adaboost	0.7914	0.5433	0.9360	0.8098	0.8000	0.8999
AKT Adaboost + CH	0.7776	0.5250	0.9308	0.7891	0.7685	0.8887
AKT kNN	0.7224	0.4479	0.9128	0.7278	0.6701	0.8560
AKT kNN + CH	0.7316	0.4609	0.9161	0.7385	0.6871	0.8620

*İK: İki katlı tekil, ATK: Ağırlıklı kural tabanlı, KT: Kural tabanlı, BK: Bulanık küme CH: Dışbükey örtü (*convexhull*), GN: Görüntü öznitelikleri

Genel olarak tüm yöntemler bazında duyarlılık ölçümü, seçicilik ölçümünden daha düşüktür. Bu durum yöntemlerin negatif örnekleri tespit etmekte, olumlu örnekleri tespit etmektен daha başarılı olduğu gözlemlenmiştir.

Üç sınıflı veri kümesi üzerinde yapılan deneylerde;

Amaç birbirine yakın olan değerlendirmelerin (örn; 5 – *highly suspicious* / 4 – *suspicious* ve 1 – *highly unlikely* / 2 – *unlikely*) gruplandırılarak değerlendirmeler arasındaki saçılmanın azaltılmasıdır.

Çizelge 8.1’deki verilere göre, AKT RF, KT SVM+CH ve AKT SVM+CH yöntemleri sırasıyla %84.89, %83.28 ve %83.20 sınıflama başarımlarıyla tüm yöntemler içinde en

yüksek sonuca ulaşmışlardır. Tekil sınıflayıcılarda ise LDA, %82.00 ile 5 sınıflı deneyin aksine en yüksek sınıflama sonucunu elde etmiştir. İki deney arasında %16 ile en çok gelişme gösteren yöntem tekil LDA'dır. RF ve Adaboost, %81.51 ve %81.35 ile diğer yüksek sonuca ulaşan tekil sınıflayıcılardır.

Duyarlılık ölçümü dikkate alındığında AKT RF, KT SVM+CH ve İK+GN sırasıyla %83.11, %82.34 ve %82.19 ile sunulan yöntemler arasında en yüksek sonuçlara ulaşmışlardır. Tekil sınıflayıcılarda duyarlılık sonucu ise Adaboost, NB ve RF sırasıyla %82.65, %82.50 ve %81.61'dir. 5 sınıflı deneyin aksine sunulan yöntem ve tekil sınıflayıcılar arasında fark düşüktür. K11 RF yöntemi tüm yöntemler içinde en yüksek duyarlılık değerine sahiptir.

Seçicilik ölçümünde ise AKT RF, KT SVM+CH ve AKT SVM+CH sırasıyla %92.09, %91.18 ve %91.00 ile tüm yöntemler içinde en yüksek sonuca ulaşmışlardır. Tekil yöntemler içinde ise NB, Adaboost ve RF %90.60, %90.55 ve %90.35 ile en yüksek seçicilik sonuçlarını elde etmişlerdir.

Genel olarak seçicilik ölçümünde de yöntemler arasında sınıflama başarımları kadar belirgin farklılık yoktur. Bu durum yöntemlerin olumsuz örnekleri ve olumlu örnekleri tespit etmede farklılık arz etmediklerini göstermektedir.

Beş sınıflı ve üç sınıflı deneyler karşılaştırıldığında, geliştirilen ve karşılaştırılan tüm yöntemler dikkate alındığında;

- Sınıflama başarımları ortalama olarak %4 artış göstermiştir.
- Duyarlılık değeri ortalama olarak %25 artış göstermiştir.
- Seçicilik değeri ortalama olarak %4 düşüş olmuştur.

Ağırlıklı kural tabanlı yöntemde kullanılan taban sınıflayıcılar ve karşılığı olan tekil sınıflayıcılar dikkate alındığında ise, kNN taban sınıflayıcı yöntem hariç tüm yöntemlerde 5 sınıflı deneyde sınıflama başarımları artmıştır. 3 sınıflı deneyde ise kNN ve LDA taban sınıflayıcı yöntemler hariç diğerlerinde başarımları artmıştır.

8.2. Yöntemler Arası Anlamlılık Değerlendirmesi

Sunulan yöntemler ve tekil sınıflayıcılar arasında, sınıflama başarımları açısından istatistiksel olarak anlamlı fark olup olmadığının test edilmesi için Kesim 3.8'de verilen istatistiksel test uygulanmıştır. 5 sınıflı ve 3 sınıflı veri kümeleri üzerinde yapılan deneyler sonrasında sınıflama hataları üzerinden elde edilen test sonuçları Çizelge 8.2 ve Çizelge

8.3'de gösterilmiştir. Yokluk hipotezi iki algoritma arasında anlamlı farklılık bulunmamasıdır. Çizelgelerde verilen p değerlerinin 0.05'in altında olması, yokluk hipotezinin reddedilmesi demektir ve sunulan yöntem ile tekil sınıflayıcı arasında anlamlı farklılık olduğunu göstermektedir. Bu çizelgeler incelenirken Çizelge 8.1'de dikkate alınmalıdır. İki yöntem arasında anlamlı farklılık olduğu halde, sunulan yöntemin sınıflama başarımı karşılaştırılan yöntemden daha düşük ise, bu anlamlılık olumsuz yönde oluşmuştur.

Çizelge 8.2'de 5 sınıflı veri için verilen istatistiksel anlamlılık sonuçlarına göre; sunulan yöntemler içinde en yüksek başarıma ulaşmış olan AKT RF yönteminin sınıflama başarımı, tekil RF yöntemi hariç ($p = 0.164$) diğer tüm tekil sınıflayıcılara göre ($p < 0.05$) anlamlıdır. En başarılı ikinci yöntem olan İK+GN ise aynı şekilde tekil RF hariç ($p = 0.436$) diğer tüm sınıflayıcılara göre anlamlıdır. Tüm yöntemler en düşük başarıma sahip tekil LDA'ya göre anlamlıdır ($p < 0.05$). AKT kNN, AKT kNN+CH, BK Üİ3 ve ANFIS ile en başarılı olan tekil RF arasında anlamlılık varsa da sınıflama başarımı daha düşük olduğu için bu anlamlılık olumsuzdur.

Çizelge 8.3'de 3 sınıflı veri için verilen istatistiksel anlamlılık sonuçlarına göre; sunulan yöntemler içinde en yüksek başarıma ulaşmış olan AKT RF yönteminin sınıflama başarımı ile diğer tüm tekil yöntemler arasında anlamlı farklılık vardır ($p < 0.05$). AKT kNN, AKT kNN+CH, BK Üİ3 yöntemleri ile diğer tekil yöntemlerin başarımı arasında istatistiksel fark olsa da ($p < 0.05$) sınıflama başarımı daha düşük olduğu için bu anlamlılık olumsuzdur.

Çizelge 8.2 Beş sınıflı veri için, geliştirilen yöntemler ve tekil sınıflayıcılar arası istatistiksel anlamlılık testi sonuçları (p değerleri).

Yöntem	LDA	Adaboost	Naive Bayes	kNN	SVM	RF
İK	0,000	0,001	0,001	0,443	0,819	0,118
İK + GN	0,000	0,000	0,000	0,002	0,035	0,439
BK Üİİ	0,000	0,107	0,102	0,423	0,072	0,002
ANFIS	0,000	0,039	0,037	0,727	0,179	0,007
KT SVM + CH	0,000	0,000	0,000	0,018	0,173	0,979
AKT SVM	0,000	0,000	0,000	0,032	0,252	0,847
AKT SVM + CH	0,000	0,000	0,000	0,005	0,073	0,650
AKT RF	0,000	0,000	0,000	0,000	0,006	0,164
AKT RF + CH	0,000	0,000	0,000	0,003	0,044	0,499
AKT LDA	0,000	0,000	0,000	0,250	0,877	0,237
AKT LDA + CH	0,000	0,005	0,005	0,693	0,547	0,053
AKT Adaboost	0,000	0,000	0,000	0,125	0,592	0,423
AKT Adaboost + CH	0,000	0,002	0,002	0,499	0,748	0,097
AKT kNN	0,001	0,864	0,882	0,010	0,000	0,000
AKT kNN + CH	0,000	0,722	0,705	0,040	0,002	0,000

Çizelge 8.3 Üç sınıflı veri için, geliştirilen yöntemler ve tekil sınıflayıcılar arası istatistiksel anlamlılık testi sonuçları (p değerleri).

Yöntem	LDA	Adaboost	Naive Bayes	kNN	SVM	RF
İK	0,206	0,404	0,544	0,765	0,995	0,347
İK + GN	0,889	0,569	0,425	0,089	0,159	0,643
BK Üİİ	0,000	0,001	0,003	0,034	0,016	0,001
ANFIS	0,002	0,008	0,016	0,130	0,071	0,006
KT SVM + CH	0,388	0,196	0,129	0,015	0,033	0,235
AKT SVM	0,947	0,620	0,469	0,103	0,181	0,696
AKT SVM + CH	0,419	0,216	0,143	0,018	0,038	0,258
AKT RF	0,047	0,016	0,008	0,000	0,001	0,021
AKT RF + CH	0,654	0,380	0,269	0,044	0,086	0,440
AKT LDA	0,470	0,769	0,948	0,401	0,584	0,690
AKT LDA + CH	0,023	0,065	0,105	0,475	0,314	0,051
AKT Adaboost	0,502	0,809	0,990	0,372	0,549	0,729
AKT Adaboost + CH	0,047	0,118	0,183	0,669	0,471	0,096
AKT kNN	0,000	0,000	0,000	0,000	0,000	0,000
AKT kNN + CH	0,000	0,000	0,000	0,001	0,000	0,000

8.3. Literatürdeki Çalışmaların Karşılaştırılması

LIDC veritabanı, araştırmacılara açık bir veritabanı olmasına karşın, varsayımların çeşitliliği sebebiyle bu veritabanı üzerinde yapılmış literatürdeki çalışmaları karşılaştırmak zordur. Zinovev ve diğerlerinin [14], [17] ve Vinay ve diğerlerinin [18], [19], [34] çalışmaları bu alanda öne çıkmaktadır. Bu çalışmalardan Zinovev ve diğerlerinin çalışmaları doğrudan görüntü özneliklerini habislik derecesinin tahmininde kullanırken, Vinay ve diğerleri radyografik tanımlayıcıları Lee ve diğerleri gibi [20] alt düzey öznelik gibi kullanmışlardır. Zinovev ve diğerleri [17] çalışmalarında ortalama oylama (*median voting*) ve LIDC veritabanının 2009 yılının ilk sürümlerinden birini (207 hasta ve 917 nodül) kullanmışlardır. Çeşitli genelleştirme ve sınıflama yöntemleri sonucu ortalama olarak 0.7251 ± 0.18 sınıflama başarımı elde etmişlerdir. Diğer bir çalışmalarında ise 0.5900 ± 0.04 sınıflama başarımı elde etmişlerdir. Bu makalelerinde, etkin öğrenme ve takım sınıflayıcıların tekil sınıflayıcılara karşı olan başarımlarını incelemişlerdir. Vinay ve diğerleri ise [18], [19], [34], çalışmalarında LIDC'nin 2009 yılı pilot sürümünü (399 hasta) kullanmışlardır. İlk önce farklı türden tekil sınıflayıcıların başarımlarını incelemişlerdir [34] ve sonraki çalışmalarında farklı yapıda takım sınıflayıcılar sunarak başarımlarını karşılaştırmışlardır [18], [19]. Ortalama 0.7828 ± 0.06 sınıflama başarımı elde etmişlerdir. Vinay ve diğerleri [19] çalışmalarında radyografik tanımlayıcıları, Lee ve diğerlerinin [20] çalışmalarında olduğu gibi, doğrudan görüntü öznelikleriyle birlikte alt düzey öznelik gibi kullanmıştır. Lee ve diğerleri LIDC'den farklı bir veritabanı kullanmıştır. Vinay ve diğerleri ise LIDC veritabanındaki 8 radyografik tanımlayıcıyı da verikümelerine eklemiştir, fakat çalışmalarında radyografik tanımlayıcıların olmadığı durumları değerlendirmeye almamışlardır. Gerçek dünya problemi bağlamında düşünülürse, radyologlardan radyografik tanımlayıcıları görüntü öznelikleri gibi (nodül alanlarının kendileri tarafından ya da otomatik/yarı otomatik bölütleme işleminden belirlendiğini kabul edersek) kullanmalarını beklemek gerçekçi olmayacaktır. Çünkü radyografik tanımlayıcıların tespit edilmesi, nodülün habislik derecesinin tespiti kadar zaman alacak bir süreçtir. Radyolog bu süre için habislik derecesini tahmin etmeyi tercih edecektir.

Bu çalışmalar arasında genelleştirme yöntemleri, veri kümeleri ve kesin referans varsayımları farklılıklarından [104] dolayı doğrudan karşılaştırma yapmak zordur. Bu karşılaştırmayı yapmak zor olmasına karşın elde ettiğimiz sonuçlar, alandaki çalışmalar ile karşılaştırılabilir düzeydedir.

9. SONUÇLAR

Bu tez kapsamında, literatürdeki çalışmalardan farklı olarak radyografik tanımlayıcılar kullanılarak nodülün habislik derecesi tespit edilmeye çalışılmıştır. Bu amaçla görüntü öznitelikleri üzerinden radyografik tanımlayıcıların tahmin edileceği ve bu çıktılar ile de habislik derecesinin tespit edileceği iki aşamalı bir sınıflama yapısı üzerinde tekil sınıflayıcılar, takım sınıflayıcılar, bulanık mantık tabanlı ve kural tabanlı yöntemler değerlendirmeye alınmıştır. Veri kümesi olarak Amerikan Kanser Araştırmaları Enstitüsü'nün yayınladığı LIDC veritabanı kullanılmıştır. LIDC veritabanının en büyük sorunlarından birisi kesin referansın bulunmaması ve dengesiz veri dağılımıdır. Kesin referansın belirlenmesi için çoğunluk oylama yöntemi kullanılmıştır. Dengesiz verilerin ele alınmasında ise doğrudan ve dolaylı yaklaşımlar kullanılmıştır. Doğrudan yaklaşımlar, veri kümesi dengeleme yöntemleridir. Dolaylı yaklaşımlar ise takım sınıflayıcıların ve sınıf tabanlı boyut indirgemenin kullanılmasıdır.

Tez çalışmasının ilk aşamalarında, LIDC veritabanındaki verinin analizi ve kesin referans kabul edilebilecek verinin elde edilmesi ile ilgili çalışılmıştır. Daha sonra ilk olarak iki katlı tekil sınıflayıcılar ile habislik tespiti üzerine gidilmiştir. Tekil sınıflayıcılara göre bir miktar daha iyi sonuca veren bu yaklaşımda, ilk katmandan elde edilen radyografik tanımlayıcı değerlerine görüntü özniteliklerinin eklenmesinin sonucu iyileştirdiği gözlenmiştir. LIDC veritabanı, çoklu değerlendirici tarafından öznel değerlendirmelere sahip olması nedeniyle, bulanık yaklaşımların kullanılmasına uygun bir veri tabanıdır [92]. Bu nedenle, sonraki aşamalarda bulanık küme ve nöro-bulanık yaklaşımlar kullanılmış fakat iki katlı tekil sınıflayıcıya göre daha düşük başarımın elde edildiği için alternatif yöntem arayışlarına girilmiştir. Takım sınıflayıcıların, literatürde birçok çalışmada dengesiz veri kümelerini ele almak amacıyla kullanılan bir araç olmasından yola çıkılarak [105]–[108], sonraki kesimlerdeki çalışmalar takım sınıflayıcılar ile genişletilmiştir. Bu amaçla ilk katmanda takım sınıflayıcılar ve dışbükey örtü yaklaşımı (*convex hull*) kullanılarak, radyografik tanımlayıcılar için oylar üretilmiştir. İkinci katmanda, radyolog değerlendirmelerinden elde edilen kurallarla oluşturulan kural tabanlı bir yaklaşım kullanılarak bu oylar değerlendirilmiş ve habislik derecelendirmesi yapılmıştır. Daha sonraki çalışmalarda, bu oylama yaklaşımı, radyolog değerlendirmelerinin birbirleri arasındaki tutarlılığının ve radyografik tanımlayıcılar arasındaki doğrusal bağıntının değerlendirilmesi ile genişletilmiş ve sınıflama başarımı artırılmıştır. En son olarak da, farklı takım sınıflayıcılar kullanarak sınıflama başarımı artırılmıştır.

Takım sınıflayıcılar ile yapılan çalışma değerlendirmeye alındığında; sonuçlar dışbükey örtü yaklaşımının, SVM ve kNN taban sınıflayıcı yöntemlerde başarıyı arttırdığını, diğer yöntemlerde ise düşürdüğünü göstermektedir. Bu durum Nalbantov ve diğerlerinin [103], dışbükey örtünün SVM sınıflayıcısının başarıyı arttırdığı savını desteklemektedir. Dışbükey örtü yaklaşımı, bu algoritmalar dışında, RF, Adaboost ve LDA takım sınıflayıcılarında başarıyı olumsuz yönde etkilemektedir. Çalışmalarda ayrıca, habislik derecelendirmenin gruplandırma yaklaşımı ile yapılması üzerine de çalışılmıştır. Genel olarak habislik derecelerinin gruplandırılması, radyologlarca üzerinde çok anlaşılmasayan ara derecelerin (2 - *moderately unlikely* ve 4 - *moderately suspicious*) temsil oranını artırdığı için sınıflama başarıyı artırdığı gözlenmiştir.

Çalışmalarda, deneysel sonuçlardan radyografik tanımlayıcıların nodüllerin habislik tahmininde sınıflama sonuçlarına katkıda bulunduğu gözlemlenmiştir. İki katlı sınıflayıcı yaklaşımında, tekil sınıflayıcılar, takım sınıflayıcılar, bulanık mantık tabanlı ve kural tabanlı yöntemler kullanılmıştır. Tekil sınıflayıcılar kullanılarak oluşturulan katlı sınıflayıcı ise görüntü öznitelikleriyle birleştirildiğinde, sınıflama başarılarında dikkate değer bir artış gözlemlenmiştir. Tez kapsamında kullanılan bulanık küme yaklaşımları diğer yöntemlere göre daha az işlem karmaşıklığına sahip yöntemlerdir. Sınıflama başarıları daha düşük olsa da elde edilen sonuçlar yöntemlerin geliştirilebilir olduğunu göstermektedir. Takım sınıflayıcılar ve kural tabanlı yaklaşım genel olarak karşılaştırılan yöntemler içinde sınıflama başarıları en yüksek olanlardır. Radyolog değerlendirmelerinin birbirleri arasındaki tutarlılığının ve radyografik tanımlayıcılar arasındaki doğrusal bağıntının dikkate alınmasının sınıflama başarısına olumlu katkıları olduğu gözlenmiştir. Yapılan istatistiksel anlamlılık testlerinde de, sınıflama başarıları en yüksek olan RF tabanlı takım sınıflayıcısının istatistiksel olarak da diğer yöntemlerden üstün olduğu gözlenmiştir.

Gelecekte bu alanda yapılabilecek çalışmalar:

- Öznitelikler bağlamında: Öznitelik kümeleri ek görüntü nitelikleri ile genişletilebilir. Çalışmada 2 ve 2.5 boyutlu öznitelikler dikkate alınmıştır. Üç boyutlu öznitelikler de çalışmaya dahil edilebilir. Momentler üzerinde farklı parametreler değerlendirmeye alınabilir.
- Boyut indirgeme bağlamında: Kullanılan Relief yöntemine ek olarak farklı yapıda boyut indirgeme yöntemleri ile çalışma genişletilebilir. Bu yöntemler ile

özniteliklerin azaltılması hız başarımına olumlu katkıda bulunmaktadır. Koşut işlem yaklaşımı tanımlanan probleme uyarlanarak nitelik sayısının hız başarımına olumsuz etkisi azaltılabilir.

- Kesin referans bağlamında: Tez çalışmasında çoğunluk oylama yaklaşımı ile kesin referanslar belirlenmiştir. Ortanca (*median*) oylama gibi farklı oylama yaklaşımları da değerlendirmeye alınabilir. Horsthemke ve diğerlerinin [42] yaklaşımında olduğu gibi olasılık haritaları kullanılarak radyologların değerlendirmeleri bir araya getirilebilir. Radyologlar arası anlaşma bağlamında karşılaştırmalar yapılan çalışmalar [21], [22], [42] yönünde geliştirmeler yapılabilir.
- Bulanık yaklaşımlar bağlamında: Tez çalışmasında önerilen bulanık yöntemlere ek olarak farklı bulanık yaklaşımlar da uyarlanabilir. LIDC verilerinin bu yaklaşım yönünde değerlendirilerek daha uygun hale getirilmesi sağlanabilir.
- Tez kapsamında bir bilgisayar destekli tanı sisteminin parçası olarak nitelendirilebilecek bölütleme aşamasına değinilmemiştir. LIDC üzerinde denenecek bölütleme yaklaşımları ile önerilen yaklaşımlar birleştirilerek tanı sistemi bir bütün olarak ortaya koyulabilir.

KAYNAKLAR

- [1] Kaneko, M., Eguchi, K., Ohmatsu, H., Kakinuma, R., Naruke, T., Suemasu, K., Moriyama, N., “Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography.,” *Radiology*, vol. 201, pp. 798–802, **1996**.
- [2] Way, T. W., “Computer Aided Diagnosis of Pulmonary Nodules in Thoracic Computed Tomography,” **2008**.
- [3] “W.H.O Cancer.” Bağlantı: <http://www.who.int/cancer/en/>. [Erişim: 29-01-2012].
- [4] Austin, J. H., Müller, N. L., Friedman, P. J., Hansell, D. M., Naidich, D. P., Remy-Jardin, M., Webb, W. R., Zerhouni, E. A., “Glossary of terms for CT of the lungs: recommendations of the Nomenclature Committee of the Fleischner Society.,” *Radiology*, vol. 200, no. 2, pp. 327–31, **1996**.
- [5] Armato, S. G., McLennan, G., McNitt-Gray, M. F., Meyer, C. R., Yankelevitz, D., Aberle, D. R., Henschke, C. I., Hoffman, E. A., Kazerooni, E. A., MacMahon, H., Reeves, A. P., Croft, B. Y., Clarke, L. P., “Lung image database consortium: developing a resource for the medical imaging research community.,” *Radiology*, vol. 232, pp. 739–748, **2004**.
- [6] Doi, K., “Computer-aided diagnosis in medical imaging: historical review, current status and future potential.,” *Comput. Med. Imaging Graph.*, vol. 31, no. 4–5, pp. 198–211.
- [7] “ImageChecker.” Bağlantı: <http://www.pamf.org/radiology/services/mammography/imagechecker.html>. [Erişim: 01-06-2015].
- [8] “CADStream.” Bağlantı: <http://www.merge.com/Solutions/Radiology/Merge-CADstream.aspx>. [Erişim: 01-06-2015].
- [9] “Mevis Liver Analyser.” Bağlantı: <http://www.mevis.fraunhofer.de/index.php?id=548&L=1>. [Erişim: 01-06-2015].
- [10] Sluimer, I., Schilham, A., Prokop, M., Van Ginneken, B., “Computer analysis of computed tomography scans of the lung: A survey,” *IEEE Trans. Med. Imaging*, vol. 25, pp. 385–405, **2006**.
- [11] Shiraishi, J., Abe, H., Engelmann, R., Aoyama, M., MacMahon, H., Doi, K., “Computer-aided diagnosis to distinguish benign from malignant solitary pulmonary nodules on radiographs: ROC analysis of radiologists’ performance--initial experience.,” *Radiology*, vol. 227, pp. 469–474, **2003**.
- [12] Wolf, M., Krishnan, A., Salganicoff, M., Bi, J., Dunder, M., Fung, G., Stoeckel, J., Periaswamy, S., Shen, H., Herzog, P., Naidich, D. P., “CAD performance analysis for pulmonary nodule detection on thin-slice MDCT scans,” *Int. Congr. Ser.*, vol. 1281, pp. 1104–1108, **2005**.

- [13] Hanamiya, M., Aoki, T., Yamashita, Y., Kawanami, S., Korogi, Y., “Frequency and significance of pulmonary nodules on thin-section CT in patients with extrapulmonary malignant neoplasms,” *Eur. J. Radiol.*, vol. 81, pp. 152–157, **2012**.
- [14] Zinovev, D., Furst, J., Raicu, D., “Building an Ensemble of Probabilistic Classifiers for Lung Nodule Interpretation,” *2011 10th Int. Conf. Mach. Learn. Appl. Work.*, vol. 2, pp. 155–161, **2011**.
- [15] Gurcan, M. N., Sahiner, B., Petrick, N., Chan, H.-P., Kazerooni, E. A., Cascade, P. N., Hadjiiski, L., “Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system,” *Med. Phys.*, vol. 29, pp. 2552–2558, **2002**.
- [16] Kawata, Y., Niki, N., Ohmatsu, H., Kusumoto, M., Kakinuma, R., Mori, K., Nishiyama, H., Eguchi, K., Kaneko, M., Moriyama, N., “Computerized analysis of 3-D pulmonary nodule images in surrounding and internal structure feature spaces,” *Proc. 2001 Int. Conf. Image Process. (Cat. No.01CH37205)*, vol. 2, **2001**.
- [17] Zinovev, D., Raicu, D., Furst, J., Armato III, S. G., “Predicting Radiological Panel Opinions Using a Panel of Machine Learning Classifiers,” *Algorithms*, vol. 2, pp. 1473–1502, **2009**.
- [18] Vinay, K., Rao, A., Hemanthakumar, G., “Combining Ensemble of Classifiers Using Voting-Based Rule to Predict Radiological Ratings for Lung Nodule Malignancy,” *Emerg. Res. Electron.*, vol. 248, pp. 443–451, **2014**.
- [19] Vinay, K., Rao, A., Kumar, G., “Classifiers in Context: Prediction of Radiological Characteristic Ratings for Lung Nodule Malignancy,” **2013**.
- [20] Lee, M. C., Boroczky, L., Sungur-Stasik, K., Cann, A. D., Borczuk, A. C., Kawut, S. M., Powell, C. A., “Computer-aided diagnosis of pulmonary nodules using a two-step approach for feature selection and classifier ensemble construction,” *Artif. Intell. Med.*, vol. 50, pp. 43–53, **2010**.
- [21] Horsthemke, W., Varutbangkul, E., Raicu, D., Furst, J., “Predictive data mining for lung nodule interpretation,” in *Proceedings - IEEE International Conference on Data Mining, ICDM, 2007*, pp. 157–162.
- [22] Li, G., Kim, H., Tan, J. K., Ishikawa, S., Hirano, Y., Kido, S., Tachibana, R., “Semantic characteristics prediction of pulmonary nodule using Artificial Neural Networks,” in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2013*, pp. 5465–5468.
- [23] El-Baz, A., Beache, G. M., Gimel’Farb, G., Suzuki, K., Okada, K., Elnakib, A., Soliman, A., Abdollahi, B., “Computer-aided diagnosis systems for lung cancer: Challenges and methodologies,” *International Journal of Biomedical Imaging*, vol. 2013, p. 46, **2013**.
- [24] Suzuki, K., “Machine Learning in Computer-Aided Diagnosis of the Thorax and Colon in CT: A Survey,” *IEICE Trans. Inf. Syst.*, **2013**.

- [25] Abbod, M. F., von Keyserlingk, D. G., Linkens, D. A., Mahfouf, M., “Survey of utilisation of fuzzy technology in Medicine and Healthcare,” *Fuzzy Sets and Systems*, vol. 120. pp. 331–349, **2001**.
- [26] Mahfouf, M., Abbod, M. F., Linkens, D. A., “A survey of fuzzy logic monitoring and control utilisation in medicine,” *Artif. Intell. Med.*, vol. 21, pp. 27–42, **2001**.
- [27] McNitt-Gray, M. F., Hart, E. M., Wyckoff, N., Sayre, J. W., Goldin, J. G., Aberle, D. R., “A pattern classification approach to characterizing solitary pulmonary nodules imaged on high resolution CT: preliminary results,” *Med. Phys.*, vol. 26, pp. 880–888, **1999**.
- [28] Minami, K., Kawata, Y., Niki, N., Mori, K., Ohmatsu, H., Kakinuma, R., Eguchi, K., Kusumoto, M., Kaneko, M., Moriyama, N., “Computerized characterization of contrast enhancement patterns for classifying pulmonary nodules,” *Proc. 2001 Int. Conf. Image Process. (Cat. No.01CH37205)*, vol. 2, **2001**.
- [29] Lo, S.-C. B., Hsu, L.-Y., Freedman, M. T., Lure, Y. M. F., Zhao, H., “Classification of lung nodules in diagnostic CT: an approach based on 3D vascular features, nodule density distribution, and shape features,” in *Medical Imaging 2003*, **2003**, pp. 183–189.
- [30] Xu, D. M., van Klaveren, R. J., de Bock, G. H., Leusveld, A., Zhao, Y., Wang, Y., Vliegenthart, R., de Koning, H. J., Scholten, E. T., Verschakelen, J., Prokop, M., Oudkerk, M., “Limited value of shape, margin and CT density in the discrimination between benign and malignant screen detected solid pulmonary nodules of the NELSON trial,” *Eur. J. Radiol.*, vol. 68, pp. 347–352, **2008**.
- [31] “NELSON Trial.” Bağlantı: <http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=636>. [Erişim: 29-01-2012].
- [32] Kawata, Y., Niki, N., Ohmatsu, H., Kusumoto, M., Kakinuma, R., Mori, K., Nishiyama, H., Eguchi, K., Kaneko, M., Moriyama, N., “Computer-Aided Diagnosis of Pulmonary Nodules Using Three-Dimensional Thoracic CT Images,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2001 SE - 242*, vol. 2208, W. Niessen and M. Viergever, Eds. Springer Berlin Heidelberg, **2001**, pp. 1393–1394.
- [33] Aoyama, M., Li, Q., Katsuragawa, S., Li, F., Sone, S., Doi, K., “Computerized scheme for determination of the likelihood measure of malignancy for pulmonary nodules on low-dose CT images,” *Med. Phys.*, vol. 30, pp. 387–394, **2003**.
- [34] Vinay, K., “Prediction of Lung Nodule Characteristic Rating using Best Classifier Model,” *Int. J. Comput. Appl.*, vol. 56, no. 18, **2012**.
- [35] Kawata, Y., Niki, N., Ohmatsu, H., Kusumoto, M., Kakinuma, R., Mori, K., Nishiyama, H., Eguchi, K., Kaneko, M., Moriyama, N., “Searching similar images for classification of pulmonary nodules in three-dimensional CT images,” in

- Proceedings IEEE International Symposium on Biomedical Imaging*, **2002**, pp. 189–192.
- [36] Thirion, J., Gourdon, A., “Computing the differential characteristics of iso-intensity surfaces,” *Comput. Vis. image Underst.*, **1995**.
- [37] Li, Q., Li, F., Shiraishi, J., Katsuragawa, S., Sone, S., Doi, K., “Investigation of new psychophysical measures for evaluation of similar images on thoracic computed tomography for distinction between benign and malignant nodules,” *Med. Phys.*, vol. 30, pp. 2584–2593, **2003**.
- [38] Kawata, Y., Niki, N., Ohmatsu, H., Kusumoto, M., Kakinuma, R., Mori, K., Nishiyama, H., Eguchi, K., Kaneko, M., Moriyama, N., “Analysis of Pulmonary Nodule Evolutions Using a Sequence of Three-Dimensional Thoracic CT Images,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2001 SE - 13*, vol. 2208, W. Niessen and M. Viergever, Eds. Springer Berlin Heidelberg, **2001**, pp. 103–110.
- [39] Jabon, S. A., Raicu, D. S., Furst, J. D., “Content-based versus Semantic-based Retrieval: A LIDC Case Study,” in *Medical Imaging 2009: Image Perception, Observer Performance, and Technology Assessment.*, **2009**, vol. 47803.
- [40] Matsuki, Y., Nakamura, K., Watanabe, H., Aoki, T., Nakata, H., Katsuragawa, S., Doi, K., “Usefulness of an artificial neural network for differentiating benign from malignant pulmonary nodules on high-resolution CT,” *Am. J. Roentgenol.*, vol. 178, pp. 657–663, **2002**.
- [41] Cohen, J., “Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit,” *Psychol. Bull.*, **1968**.
- [42] Horsthemke, W., “Predicting LIDC diagnostic characteristics by combining spatial and diagnostic opinions,” in *Proc. SPIE 7624, Medical Imaging 2010: Computer-Aided Diagnosis*, **2010**.
- [43] Metz, C. E., “Basic principles of ROC analysis,” *Semin. Nucl. Med.*, vol. 8, no. 4, pp. 283–98, **1978**.
- [44] Zou, K. H., O’Malley, A. J., Mauri, L., “Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models,” *Circulation*, vol. 115, no. 5, pp. 654–7, **2007**.
- [45] KILIÇ, S., “ROC Analysis in Clinical Decision Making,” *J. Mood Disord.*, vol. 3, no. 3, pp. 135–40, **2013**.
- [46] Powers, D., “Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation,” **2011**.
- [47] Krippendorff, K., “Computing Krippendorff’s alpha reliability,” *Dep. Pap.*, p. 43, **2007**.

- [48] Scott, W., “Reliability of content analysis: The case of nominal scale coding,” *Public Opin. Q.*, **1955**.
- [49] Haralick, R. M., Shanmugam, K., Dinstein, I., “Textural Features for Image Classification,” *Syst. Man Cybern. IEEE Trans.*, vol. 3, pp. 610–621, **1973**.
- [50] Haralick, R. M., Shapiro, L. G., *Computer and Robot Vision*, vol. 1. **1992**.
- [51] Zernike, F., “Diffraction theory of the cut procedure and its improved form, the phase contrast method,” *Physica*, **1934**.
- [52] Soh, L. K., Tsatsoulis, C., “Texture analysis of sar sea ice imagery using gray level co-occurrence matrices,” *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 2 I, pp. 780–795, **1999**.
- [53] Clausi, D. A., “An analysis of co-occurrence texture statistics as a function of grey level quantization,” *Can. J. Remote Sens.*, vol. 28, no. 1, pp. 45–62, **2002**.
- [54] Cunningham, P., “Dimension Reduction,” **2007**.
- [55] West, M., “Bayesian factor regression models in the ‘large p, small n’ paradigm,” in *Bayesian Statistics 7 - Proceedings of the Seventh Valencia International Meeting*, **2003**, pp. 723–732.
- [56] Bellman, R., Kalaba, R., “On adaptive control processes,” *IRE Trans. Autom. Control*, vol. 4, **1959**.
- [57] Donoho, D., “High-dimensional data analysis: The curses and blessings of dimensionality,” *AMS Math Challenges Lect.*, pp. 1–33, **2000**.
- [58] Fodor, I. K., “A survey of dimension reduction techniques,” *Cent. Appl. Sci. Comput. Lawrence Livermore Natl. Lab.*, vol. 9, pp. 1–18, **2002**.
- [59] Alpaydm, E., *Introduction to Machine Learning*, 2nd Ed. The MIT Press, **2009**.
- [60] Zhao, Z., Morstatter, F., Sharma, S., Alelyani, S., Anand, A., Liu, H., “Advancing Feature Selection Research,” *ASU Featur. Sel. Repos. Arizona State Univ.*, pp. 1 – 28, **2010**.
- [61] Kuhn, M., “Building Predictive Models in R Using the caret Package,” *J. Stat. Softw.*, vol. 28, pp. 1–26, **2008**.
- [62] Kononenko, I., Šimec, E., Robnik-Šikonja, M., “Overcoming the myopia of inductive learning algorithms with RELIEFF,” *Appl. Intell.*, vol. 7, pp. 39–55, **1997**.
- [63] Kira, K., Rendell, L. A., “The Feature Selection Problem: Traditional Methods and a New Algorithm,” in *Proceedings of the 9th International Conference on Machine Learning*, **1992**, pp. 129–134.

- [64] Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., Altman, R. B., “Missing value estimation methods for DNA microarrays,” *Bioinformatics*, vol. 17, pp. 520–525, **2001**.
- [65] Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W. P., “SMOTE: Synthetic minority over-sampling technique,” *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, **2002**.
- [66] “Wolfram Mathworld.” Bağlantı: <http://mathworld.wolfram.com/>. [Erişim: 01-11-2014].
- [67] “Orange.” Bağlantı: <http://orange.biolab.si/>. [Erişim: 10-01-2015].
- [68] Quinlan, J. R., *C4.5: Programs for Machine Learning*, vol. 1. **1993**.
- [69] Quinlan, J. R., “Discovering Rules by Induction from large collections of examples,” in *Expert Systems in the micro-electronic age*, **1979**, pp. 168–201.
- [70] Hunt, E., Marin, J., Stone, P., “Experiments in induction. 1966,” **1986**.
- [71] Wu, X., Kumar, V., Quinlan, J., Ghosh, J., “Top 10 algorithms in data mining,” *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, **2008**.
- [72] Cortes, C., Vapnik, V., “Support-vector networks,” *Mach. Learn.*, vol. 20, pp. 273–297, **1995**.
- [73] Breiman, L., “Random forests,” *Mach. Learn.*, vol. 45, pp. 5–32, **2001**.
- [74] Dietterich, T. G., “Machine-learning research: Four current directions,” *AI Mag.*, vol. 18, pp. 97–136, **1997**.
- [75] Freund, Y., Schapire, R. E., “A Decision-theoretic Generalization of On-line Learning and an Application to Boosting,” *J. Comput. Syst. Sci.*, vol. 55, pp. 119–139, **1997**.
- [76] Schapire, R. E., Singer, Y., “Improved boosting algorithms using confidence-rated predictions,” *Mach. Learn.*, vol. 37, pp. 297–336, **1999**.
- [77] Fix, E., Hodges, J., “Discriminatory analysis. Nonparametric discrimination: Consistency properties,” ... *Stat. Rev. Int. Stat.*, vol. 57, pp. 238–247, **1989**.
- [78] Fisher, R., “The Use of Multiple Measurements in Taxonomic Problems,” *Ann. Eugen.*, vol. 7, pp. 179–188, **1936**.
- [79] Gutierrez-Osuna, R., “Linear Discriminant Analysis.” Bağlantı: http://research.cs.tamu.edu/prism/lectures/pr/pr_110.pdf. [Erişim: 28-04-2014].
- [80] Michie, E. D., Spiegelhalter, D. J., Taylor, C. C., “Machine Learning , Neural and Statistical Classification,” *Technometrics*, vol. 37, p. 459, **1994**.

- [81] Brazdil, P., Soares, C., “A Comparison of Ranking Methods for Classification Algorithm Selection,” in *Machine Learning: Proceedings of the 11th European Conference on Machine Learning {ECML2000}*, **2000**, pp. 63–74.
- [82] Brachman, R., Anand, T., *The process of knowledge discovery in databases*. American Association for Artificial Intelligence, **1996**.
- [83] Neave, H., Worthington, P., “Distribution-free tests,” **1988**.
- [84] Gwet, K. L., “On The Krippendorff’s Alpha Coefficient,” *Manuscr. Submitt. Publ.*, pp. 1–16, **2011**.
- [85] Hayes, A. F., Krippendorff, K., “Answering the Call for a Standard Reliability Measure for Coding Data,” *Communication Methods and Measures*, vol. 1. pp. 77–89, **2007**.
- [86] Snedecor, G. W., Cochran, W. G., *Statistical Methods*. Ames, Iowa: Blackwell Publishing Professional, **1989**.
- [87] Dietterich, T., “Approximate statistical tests for comparing supervised classification learning algorithms,” *Neural Comput.*, **1998**.
- [88] “The Cancer Imaging Archive Wiki Page.” Bağlantı: <https://wiki.cancerimagingarchive.net/display/Public/Wiki>. [Erişim: 01-04-2014].
- [89] Raykar, V. C., Yu, S., Zhao, L. H., Valadez, G. H., Florin, C., Bogoni, L., Moy, L., “Learning From Crowds,” *J. Mach. Learn. Res.*, vol. 11, pp. 1297–1322, **2010**.
- [90] Cawley, G., “Sparse multinomial logistic regression via bayesian l1 regularisation,” *Adv. Neural Inf. Process. Syst.*, vol. 19, p. 209, **2007**.
- [91] Barber, C. B., Dobkin, D. P., Huhdanpaa, H., “The quickhull algorithm for convex hulls,” *ACM Transactions on Mathematical Software*, vol. 22. pp. 469–483, **1996**.
- [92] Sikchi, S. S., Sikchi, S., Ali, M. S., “Fuzzy Expert Systems (FES) for Medical Diagnosis,” *Int. J. Comput. Appl.*, vol. 63, no. 11, **2013**.
- [93] Klir, G. J., Yuan, B., *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. **1995**.
- [94] Adlassnig, K.-P., “Fuzzy Set Theory in Medical Diagnosis,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 16. pp. 260–265, **1986**.
- [95] Sanchez, E., “Inverses of fuzzy relations. Application to possibility distributions and medical diagnosis,” *Fuzzy sets Syst.*, **1979**.
- [96] Pereira, J. C., Tonelli, P. A., Barros, L. C., Ortega, N. R. S., “Defuzzification in Medical Diagnosis,” in *Advances in logic, artificial intelligence and robotics*, **2002**, pp. 202–207.
- [97] Nguyen, H. T., Walker, E. A., *A first course in fuzzy logic*. CRC press, **2005**.

- [98] Jang, J., “ANFIS: adaptive-network-based fuzzy inference system,” *Syst. Man Cybern.*, vol. 23, no. 3, **1993**.
- [99] Cetisli, B., “The effect of linguistic hedges on feature selection: Part 2,” *Expert Syst. Appl.*, **2010**.
- [100] Cetisli, B., “Development of an adaptive neuro-fuzzy classifier using linguistic hedges: Part 1,” *Expert Syst. Appl.*, **2010**.
- [101] Cetişli, B., Barkana, A., “Speeding up the scaled conjugate gradient algorithm and its application in neuro-fuzzy classifier training,” *Soft Comput.*, **2010**.
- [102] Kodell, R. L., Zhang, C., Siegel, E. R., Nagarajan, R., “Selective voting in convex-hull ensembles improves classification accuracy,” *Artificial Intelligence in Medicine*, vol. 54. pp. 171–179, **2012**.
- [103] Nalbantov, G. I., Groenen, P. J. F., Bioch, J. C., “Nearest Convex Hull Classification,” pp. 1–9, **2006**.
- [104] Ochs, R., Kim, H., Angel, E., “Forming a reference standard from LIDC data: impact of reader agreement on reported CAD performance,” in *Medical Imaging. International Society for Optics and Photonics*, **2007**, p. 65142A–65142A.
- [105] Galar, M., Fernandez, A., Barrenechea, E., Bustince, H., Herrera, F., “A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches,” *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 42. pp. 463–484, **2012**.
- [106] Nikulin, V., McLachlan, G., Ng, S., “Ensemble approach for the classification of imbalanced data,” *AI 2009 Adv. Artif. Intell.*, pp. 291–300, **2009**.
- [107] Guo, H., Viktor, H., “Learning from imbalanced data sets with boosting and data generation: the DataBoost-IM approach,” *ACM SIGKDD Explor. Newsl.*, **2004**.
- [108] Yang, P., Liu, W., Zhou, B., *Ensemble-based wrapper methods for feature selection and class imbalance learning*, vol. 7818. **2013**.

ÖZGEÇMİŞ

Kimlik Bilgileri

Adı Soyadı : Aydın Kaya
Doğum Yeri : Ankara
Medeni Hali : Bekar
E-posta : aydinkaya83@gmail.com
Adresi : Kocatepe Cad. 199. Sok 1/10 Eryaman, Ankara

Eğitim

Lise : Mehmet Emin Resulzade Anadolu Lisesi
Lisans : Hacettepe Üniversitesi Bilgisayar Mühendisliği
Yüksek Lisans : Hacettepe Üniversitesi Bilgisayar Mühendisliği
Doktora : Hacettepe Üniversitesi Bilgisayar Mühendisliği

Yabancı Dil ve Düzeyi

İngilizce- KPDS 93 (2006-Güz)

İş Deneyimi

Akanus Bilişim : Yazılım Mühendisi, 2006 - 2007
Hacettepe Üniversitesi : Araştırma Görevlisi, 2007 - (Devam ediyor)

Deneyim Alanları

Makine öğrenmesi, Görüntü işleme, Veri madenciliği, Yazılım Mühendisliği

Tezden Üretilmiş Projeler ve Bütçeleri

Tezden Üretilmiş Yayınlar

- Aydın Kaya, Ahmet Burak Can, A Weighted Rule Based Method for Predicting Malignancy of Pulmonary Nodules by Nodule Characteristics, Journal of Biomedical Informatics, 2015 (Kabul edildi, DOI: 10.1016/j.jbi.2015.05.011).
- Aydın Kaya, Ahmet Burak Can, Malignancy Prediction By Using Characteristic-Based Fuzzy Sets: A Preliminary Study, IEEE International Conference on Fuzzy Systems, 2015 (Kabul edildi).
- Aydın Kaya, Ahmet Burak Can, eFis: A Fuzzy Inference Method for Predicting Malignancy of Small Pulmonary Nodules, 11th International Conference on Image Analysis and Recognition (Lecture Notes in Computer Science 8815), 2014.

- Aydın Kaya, Ahmet Burak Can, Determining Malignancy of Small Pulmonary Nodules by Nodule Characteristics, 6th European Conference of the International Federation for Medical and Biological Engineering, 2014.
- Aydın Kaya, Ahmet Burak Can, Akciğer Nodüllerinin Nitelendirilmesi, IEEE 21. Sinyal İşleme ve İletişim Uygulamaları (SIU) Kurultayı, 2013.

Tezden Üretilmiş Tebliğ ve/veya Poster sunumu ile Katıldığı Toplantılar

- IEEE 21. Sinyal İşleme ve İletişim Uygulamaları Kurultayı,2013.
- 6th European Conference of the International Federation for Medical and Biological Engineering, 2014.