

## AKCİĞER BİLGİSAYARLI TOMOGRAFİ GÖRÜNTÜLERİNDE GÖRÜNTÜ İŞLEME UYGULAMALARI İLE TÜMÖRLERİNİN TESPİT EDİLMESİ

Berat ÖZDET\* 

Semra İÇER\*\* 

Alınma: 03.06.2021; düzeltme: 08.02.2022; kabul: 23.02.2022

**Öz:** Türkiye’de bir yılda tespit edilen kanser vakalarının büyük çoğunluğu olan akciğer kanseri toplam vaka sayısının yaklaşık %20’sini oluşturmaktadır. En çok ölüm oranını oluşturan akciğer kanseri günümüzde Türkiye ve dünya için önemli bir sağlık sorunu durumundadır. Bu sorunun en önemli kaynağı erken tanısında tedavisi çok daha mümkün olan birçok vakanın erken teşhis edilememesidir. Bu çalışmada Bilgisayarlı Tomografi (BT) görüntüleri kullanılarak tümörlerin ve nodüllerin tespit edilmesi, görüntülerden çıkarılan özelliklerin farklı sınıflandırma algoritmaları ile sınıflandırılması amaçlanmıştır.

Kullanılan görüntüler DICOM formatında olup RIDER-Lung CT veri setine ait 26 görüntü üzerinde çalışılmıştır. Tümör bölgesi farklı akciğer segmentasyon yöntemleri kullanılarak elde edilmiş, tümöre ait pek çok özellik hesaplanmıştır. Hesaplanan özelliklerden istatistiksel olarak anlamlı ( $p<0,05$ ) olanları sınıflandırma için kullanılmıştır. Anlamlı özellikler Karar ağaçları (Decision Trees) algoritmaları, Destek Vektörü Makinesi (SVM), Yakın Komşuluk Sınıflandırması (KNN) sınıflandırıcı algoritmaları ve Diskriminant analizi ile sınıflandırılarak sonuçlar karşılaştırılmıştır. Bu algoritmaların doğruluk oranları karar ağaçları %97, SVM %96,6, KNN %93,6, Diskriminant analizi %97 olarak sonuç vermiştir. Yöntemler hassasiyet ve duyarlılık olarak karşılaştırıldığında ise her iki nicelik Kuadratik SVM ve Diskriminant analizinde % 95 üstüdür. Bu karşılaştırmalar sonucunda yöntemlerin yüksek başarı oranları ile umut verici olarak gelecek çalışmalarda kullanılabileceği görülmüştür.

**Anahtar Kelimeler:** Akciğer kanseri, Akciğer segmentasyonu, Görüntü işleme, Makinalı öğrenme

### Determination Of Tumors By Image Processing Applications In Lung Computerized Tomography Images

**Abstract:** Lung cancer, which is the majority of cancer cases detected in a year in Turkey, constitutes approximately 20% of the total number of cases. Lung cancer, which constitutes the highest mortality rate, is an important health problem for Turkey and the world today. The most important source of this problem is that many cases that are much more possible to treat in early diagnosis cannot be diagnosed early. In this study, it is aimed to detect tumors and nodules using Computed Tomography (CT) images, and to classify features extracted from images with different classification algorithms. The images used are in DICOM format and 26 images of the RIDER-Lung CT data set were studied. The tumor region was obtained using different lung segmentation methods, and many features of the tumor were calculated. Statistically significant ( $p<0.05$ ) calculated features were used for classification. Significant features were classified by Decision Trees algorithms, Support Vector Machine (SVM), Close Neighborhood Classification (KNN) classifier algorithms and Discriminant Analysis and the results were compared. The accuracy rates of these algorithms were 97% for decision trees, 96.6% for SVM, 93.6% for KNN, and

\* Mazaka Yazılım, Erciyes Teknopark Tekno 5, 38039, Melikgazi, Kayseri, [brt\\_ozdt@hotmail.com](mailto:brt_ozdt@hotmail.com)

\*\* Erciyes Üniv. Mühendislik Fak. Biyomedikal Müh. Böl., 38039, Melikgazi, Kayseri, [ksemra@erciyes.edu.tr](mailto:ksemra@erciyes.edu.tr)  
İletişim Yazarı: Semra İÇER ( [ksemra@erciyes.edu.tr](mailto:ksemra@erciyes.edu.tr) )

97% for Discriminant analysis. When the methods are compared in terms of sensitivity and sensitivity, both quantities are above 95% in Quadratic SVM and Discriminate analysis. As a result of these comparisons, it has been seen that the methods can be used in future studies with high success rates.

**Keywords:** Lung cancer, Lung segmentation, Image processing, Machine learning

## 1. GİRİŞ

Günümüzde akciğer kanseri insan hayatını tehdit eden en büyük kanser hastalıklarının başında gelmektedir. Erken tespiti durumunda tedavi yöntemleri çok daha etkili olmaktadır ve bu hastalığa bağlı olan birçok hastalığın da önüne geçilmiş olup insan hayatı ve hastalığın tedavisinde çok önemli bir rol oynayacaktır. Akciğer kanseri temelde küçük hücre dışı ve küçük hücreli olmak üzere iki çeşittir ve bunların alt tipleri mevcuttur.

Akciğer tümörü tespitinde yaygın olarak BT incelemesi kullanılmaktadır. Akciğer BT görüntüleri üzerinde görüntü işleme ile tümör tespiti yaparken karşılaşılan en büyük zorluk akciğerin segmentasyonu aşamasında tümör dokularının akciğer dış dokusuna ve toraks boşluğuna çok yakın olması durumlarıdır. Bu tür durumlarda İlgi Alanı (Region Of Interest) denilen bölgeyi çıkartırken veri kaybı yaşanmaktadır dolayısıyla tümörün tespitinin yapılmasını hem zor hale getirdiği gibi yanlış tespit yapılması durumunu da ortaya çıkarmaktadır. Çeşitli çalışmalarda akciğer segmentasyonu otsu eşikleme metodu ile yapılmaktadır (Adams ve diğ. ,2018). Ancak tek başına bu metot tam anlamıyla istenen bir ilgi alanı vermemektedir çünkü nodüller dış dokuya çok yakın olduğunda bu nodül görüntüleri ilgi alanının dışında kalmaktadır. Çeşitli çalışmalarda bu sorun alansal morfolojik görüntü işleme işlemleri olan açma kapama işlemleri ile çözülmeye çalışılsa da her görüntü üzerinde bu işlemlerin yapılması ne kadar büyüklükte bir açma kapama yapısı kullanılacağı veya bu açma kapama işlemlerinin ne kadar tekrar halinde yapılacağı görüntüden görüntüye göre değişeceği için otomatik veya yarı-otomatik bir sistem için bu çözümler çok elverişli olmamaktadır (Adams ve diğ. ,2018). Eşikleme yöntemini çeşitli optimizasyon algoritmalarıyla birleştirerek akciğer segmentasyonu işlemleri geleneksel eşikleme işlemine göre çok daha yüksek başarı ile akciğer segmentasyonu yapabilmektedir (Widodo ve diğ. ,2020). Bu yöntemlerin yanında bir akciğer maskesi veri seti kullanarak yapay sinir ağları ile segmentasyon işlemi yapılan çalışmalar da bulunmaktadır (Widodo ve diğ. ,2020). Bu tür çalışmalarda istenmeyen dokuların da segmentasyon maskesi içerisine alınması durumu ile karşı karşıya kalınmaktadır. Yöntem farklılıklarının yanında kullanılacak görüntü formatı da segmentasyon işlemi için önemli bir parametredir. JPEG, TIFF vb. formatlarda olan görüntülerde eşikleme işlemleri veya eşik değerlerinin hesaplanması için ekstra bir işlem gücü gerekmekte ve hesaplanan değerler görüntüden görüntüye değişebilmektedir. Ancak DICOM görüntüleri üzerinden bir segmentasyon işlemi yapmak daha standart olup, işlem gücünü de iyileştirmektedir. DICOM görüntüleri Hounsfield skalasına göre oluşturulmaktadır çeşitli çalışmalarda akciğer loblarının Hounsfield skalasındaki değerleri uzmanlarca bulunarak direkt DICOM üzerinden 400 HU akciğer loblarının eşiği olarak belirlenmiştir (Elsayed ve diğ. ,2013).

Segmentasyon işlemlerinden sonra elde edilen görüntü üzerinde sınıflandırıcılar için bir aday çıkarımı işlemleri yapılmaktadır. Aday çıkarımında kullanılan çeşitli metotlar arasında eşikleme yöntemi yine en sık kullanılan algoritmalarıdır. Akciğer içerisinde aday nodülleri çıkartma işlemi için k-ortalamlar metodu da çok sıklıkla kullanılan bir metottur (Aniketbombale ve Patil ,2017). Eşikleme ve k-ortalamlar metodu arasında performans farkı çok bulunmamaktadır ancak bu konudaki çalışmalarda otsu yönteminin daha iyi sonuçlar verdiği görülmüştür (Eset ve diğ. ,2015). Bunun yanında k-ortalamlar metodu işlem süresini ve gerekli işlem gücünü arttırmaktadır. Segmente edilmiş akciğer üzerinden aday nodüllerin parlaklığı akciğer dokusuna göre çok daha yüksek olduğu için eşikleme işlemi sonuç almak için kullanılabilecek en uygun ve basit metottur (Eset ve diğ. ,2015). Bunun yanı sıra bu adaylar

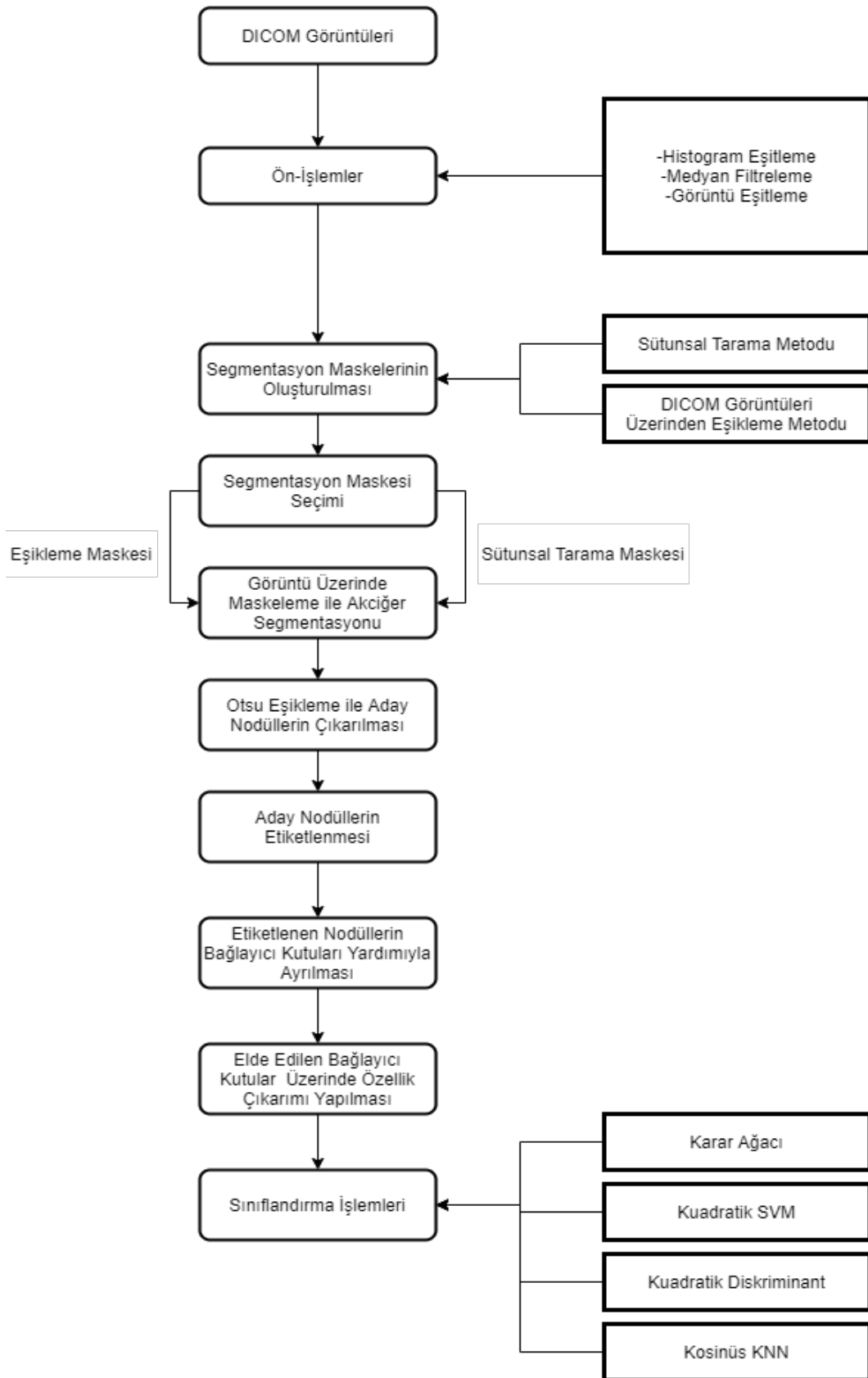
içerisinde istenmeyen ya da kontrol edilmeyecek tipte nodüller bulunabilir. Bu yüzden elde edilen adaylar arasından alanı belli bir seviyenin üzerinde veya altında bulunan dokuların adaylıktan çıkarma işlemi de çok sık kullanılan bir işlemdir.

Bu çalışmada önerilen tümör tespit metotları ile akciğer tümörlerinin tespiti yapılarak bunun otomatik bir tespit yöntemi geliştirmekte kullanılması amaçlanmıştır. Tümör tespitini yapmakta olan sınıflandırma algoritmalarında en büyük sorunlardan birisi olan segmentasyon işlemlerinde istenilmeyen veri kayıplarının yaşanmasıdır. Çalışmamızda kullandığımız sütunsal tarama ile akciğer segmentasyonu algoritması kenar bölgelerde bulunan tümör hücrelerini de segmentasyon alanına almamızı sağlamaktadır. Başarılı bir şekilde segmentasyonu yapılmış tümörün çeşitli özellikleri çıkarılarak bu özelliklerden istatistiksel anlamlı olanlara makine öğrenmesi yöntemlerinin uygulanması hedeflenmektedir. Bunun sonucunda yaygın görülen bir problemin basit ve kullanışlı bir çözümüne dayalı sonuçlar ortaya konularak hastalığın teşhis ve tedavi süreçlerini iyileştirmeye katkıda bulunulması amaçlanmıştır.

## 2. MATERYAL ve METOT

Bilgisayar destekli görüntü işleme ile tümörlerin tespiti için bu çalışmada küçük hücreli dışı akciğer kanseri tipindeki vakaların tespiti üzerine çalışılmıştır. RIDER Lung CT veri setindeki 26 adet ayrı hastanın BT görüntüleri üzerinden tümör tespiti yapılmaktadır. Çalışmada DICOM formatındaki görüntüler kullanılmış, görüntüler üzerinde ön-işlemler olarak adlandıracağımız bazı işlemler yapılarak görüntü kalitesinin işlenmeye uygun hale getirilmesi sağlanmıştır. Ön-işlem uygulanan görüntüler iki farklı segmentasyon işlemlerine tabii tutulmuştur. Segmentasyon işlemi tamamlanan görüntülere yapılan segmentasyona uygun olan maske seçimi akciğer lobları elde edilmiştir. Elde edilen akciğer loblarındaki nodül ve tümörler için özellik çıkarımı yapılarak çıkarılan özellikler sınıflandırma algoritmaları ile sınıflandırılmıştır.

Bu çalışmada BT görüntüleri üzerinde tümör dokularının tespitini yapmak için 5 ana bölümden oluşan işlemler görüntüler üzerinde uygulanmıştır. Yapılan tüm işlemlerin akış diyagramı Şekil 1’de verilmiştir.



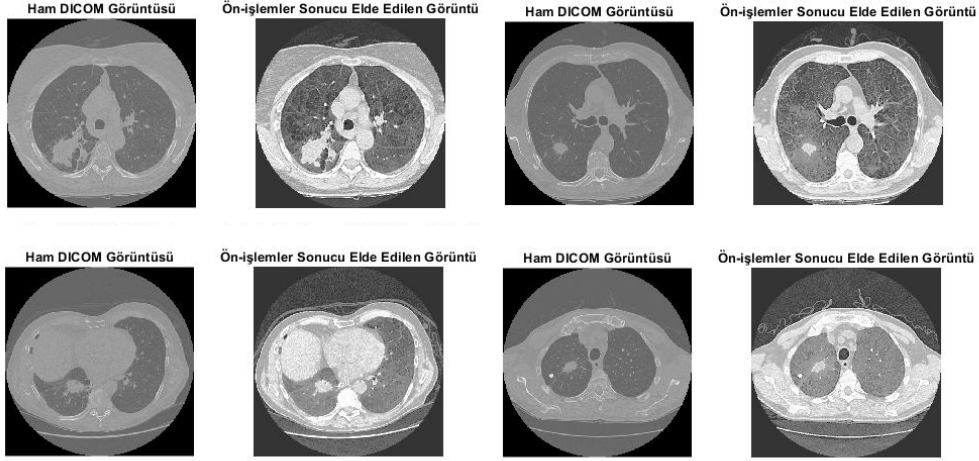
**Şekil 1:**  
*Tümör Tespit Uygulaması Akış Diyagramı*

## 2.1. Görüntü Veri Seti

Çalışmada The Cancer Imaging Archive (TCIA) kurumundan alınan RIDER Lung-CT veri setindeki 32 hasta görüntüsünden 26 tanesi kullanılmıştır ((Zhao ve diğ. ,2015), (Zhao ve diğ. ,2009), (Clark ve diğ. 2013)). Çekim kalitesi çok düşük olan, içerisinde aşırı parlama ve hasta kaynaklı çekim hataları olan görüntüler dahil edilmemiştir. Görüntüler 512x512 piksel boyutunda DICOM formatında okutulmuş ve ilk segmentasyon işlemi yapıldıktan sonra gri tonlu JPEG formatına çevrilip işlemler uygulanmıştır.

## 2.2. Ön-işlem Safhası

DICOM formatında okutulan veri seti öncelikle segmentasyon işlemlerinden birine tabii tutulmuştur. Segmentasyon işleminden sonra gri formata çevrilmiştir çünkü görüntü işleme uygulamaları gri formatlı double türünde veriler ile yapılmaktadır. Gri formata çevrilen veriler üzerinde ilk olarak medyan filtre ile iyileştirme işlemleri uygulanmıştır. Medyan filtre uygulanan görüntüler histogram eşitleme algoritmasıyla daha kullanılabilir ve elverişli görüntülere dönüştürülmüştür. Histogram eşitleme işlemi sonrasında görüntü kontrast iyileştirme işlemine tabii tutulmuştur. Görüntü eşitleme işlemi ön-işlemlerin son aşamasıdır. Ön-işlemlerin uygulanması sonrası elde edilen görüntüler Şekil 2'deki gibidir.

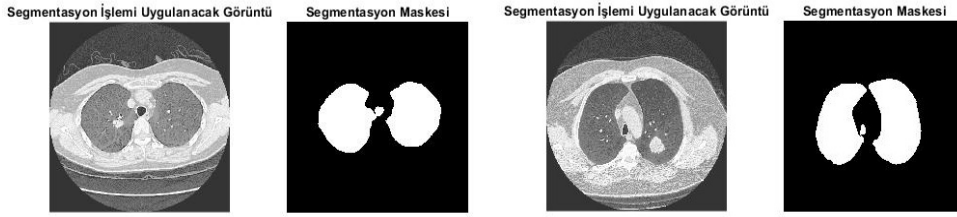


**Şekil 2:**

*Ön-işlemler Öncesi ve Sonrası Elde Edilen Görüntüler*

## 2.3. Akciğer Segmentasyonu

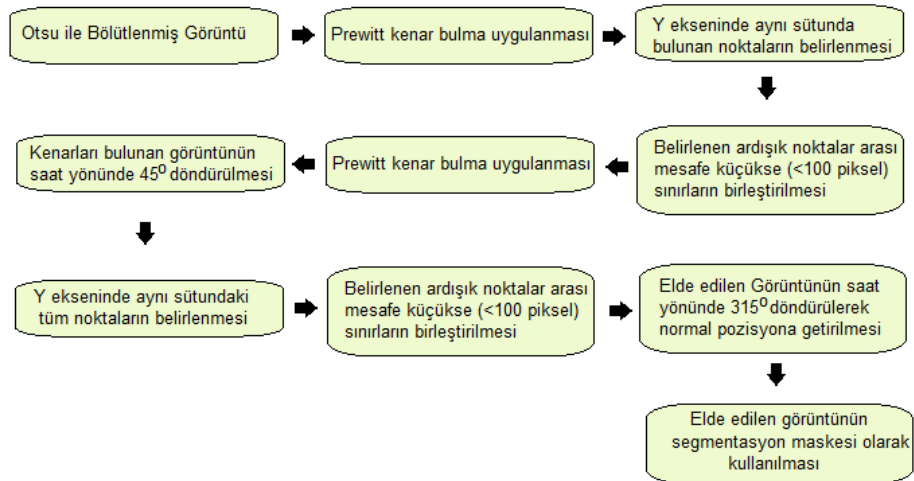
Segmentasyon için iki farklı yöntem kullanılmaktadır. Bunlardan birincisi geleneksel olan eşikleme yöntemidir. Eşikleme yöntemi DICOM görüntüleri üzerinden Hounsfield skalasına göre yapılmıştır. Bu skalada akciğer loblarının boşluğu 400 hounsfield biriminin altında kalırken daha yoğun bölümleri oluşturan kemik doku, akciğer dış dokusu ve toraks borusunun BT çekimlerinde 400 hounsfield biriminin üzerinde olduğu çeşitli çalışmalarda gösterilmiştir. Eşikleme değeri için kullanılan değer 400 hounsfield biriminden düşük olan değerleri çıkarmaya yöneliktir. Bu eşik değerinin aşağısındaki değerleri lojik 1, üzerindeki değerleri lojik 0 yapacak bir eşikleme işlemi uygulanması sonucu elde edilen akciğer segmentasyon maskeleri kullanım için son derece elverişli olmakla birlikte uygulama ve işlem gücü anlamında da diğer eşikleme algoritmaları ile aynı sonuçları verdiği gözlenmiştir. Eşikleme sonrasında akciğer alanını temizlemek için bir alan doldurma algoritması uygulanır. Bu işlemler sonucu elde edilen segmentasyon maskesi örnekleri Şekil 3'te verilmiştir.



**Şekil 3:**

*Eşikleme Yöntemi Sonucu Elde Edilen Segmentasyon Maskeleri*

Segmentasyon işleminde ikinci olarak kullandığımız yöntem ise algoritması tarafımızca geliştirilen akciğer çeperine çok yakın tümörleri segmentasyon maskesi içerisine almamızı sağlayan ve bu husustaki kayıpları önleyebilen önerdiğimiz yöntemdir. Bu yöntemde ilk olarak segmentasyon aşamasında elde ettiğimiz maske üzerinde prewitt kenar bulma algoritması uygulanır. Daha sonra elde edilen görüntü üzerinde sütunsal olarak tarama yapan ve aynı sütun üzerinde bulunan sınır noktalarının birleştirilmesi temeline dayanan bir algoritmadır. Yöntemde bu sınır noktaları koordinatları arasında 100 piksellik farktan az olan noktaların aralarındaki 0 değerlerini 1 yapılarak maske oluşturulmuştur. Bu işlemden sonra görüntüye tekrar kenar sezme algoritması uygulanıp saat yönünün tersine 45 derece döndürülme işlemi yapılarak sütun tarama işlemi tekrar uygulanmaktadır. Diyagonal olarak da gerçekleştirilen sütun tarama işlemi yapıldıktan sonra 315 derece döndürülerek görüntünün eski pozisyonuna gelmesi sağlanmış olur. Bu algoritma akciğer dış dokusuna ve toraks borusuna çok yakın tümörleri segmentasyon maskesi içerisine almak için büyük oranda başarı göstermektedir. Sütunsal tarama algoritmasının akış diyagramı Şekil 4’te verilmektedir. Bu yöntem sonucu oluşan segmentasyon maskeleri Şekil 5’te verilmiştir.

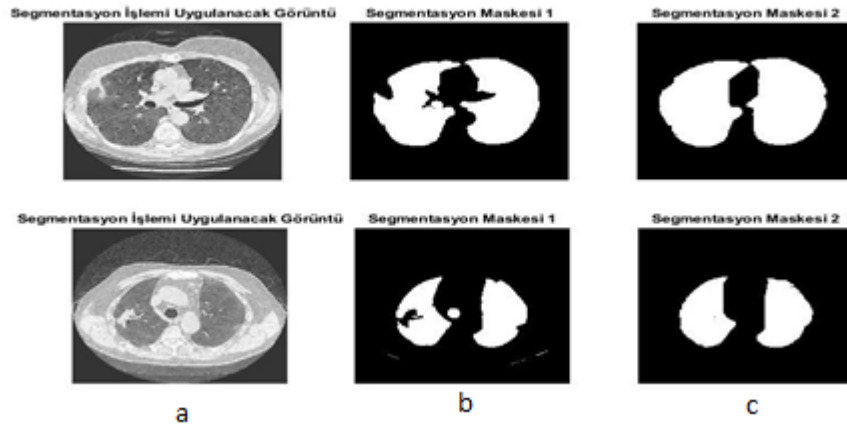


**Şekil 4:**

*Sütunsal Tarama Algoritması Akış Diyagramı*

Çalışmada kullanılan görüntüler üzerinde uygulanan segmentasyon yöntemlerinin kullanılan görüntüler üzerinde etkileri izlendiğinde klasik eşikleme yönteminde 26 görüntü üzerinden 8 tanesi klasik eşikleme metodu ile segmente edilmiş akciğer görüntüleri ile geri kalan 18 tanesi de sütunsal tarama yöntemi ile elde edilen segmente edilmiş akciğer görüntüleri kullanılarak maskeleme işlemleri yapılmıştır.

Bu iki veri seti üzerindeki başarı oranları karşılaştırıldığında klasik eşikleme metodu ile 8 görüntüde, sütunsal tarama yöntemi ile 21 görüntüde ilerideki adımlarda kullanılmak üzere seçilen başarılı bir akciğer segmentasyonu işlemi yapılmıştır. Bu iki yöntemin kullanılan veri setinde başarılı akciğer segmentasyonu yapma oranları klasik eşikleme için %30,77 olup sütunsal tarama metodu için %80,77'dir.



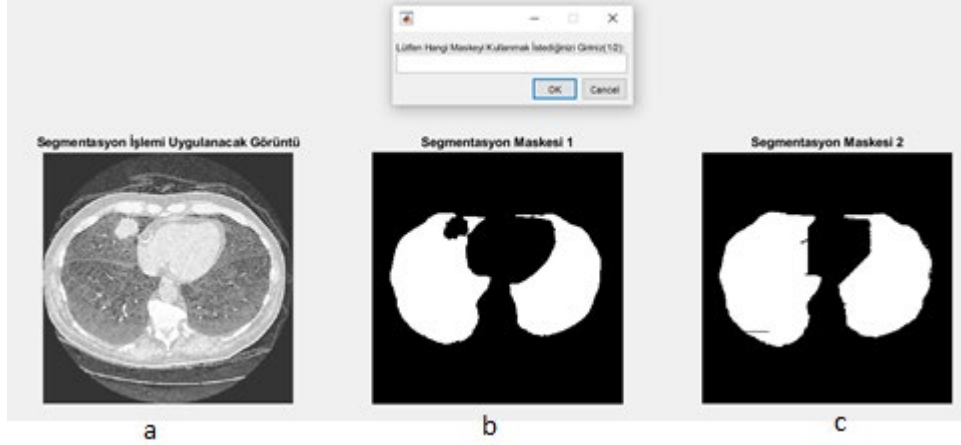
**Şekil 5:**

*a-Segmentasyon İşlemi Uygulanacak Görüntü*

*b-Klasik Eşikleme Metodu İle Elde Edilen Segmentasyon Maskesi*

*c-Sütunsal Tarama Yöntemi ile Elde Edilen Segmentasyon Maskesi*

Sistem yarı otomatik olarak öngörülmüştür, uygulamada kullanılacak segmentasyon maskeleri kullanıcı tarafından seçilecek şekilde uygulanmıştır. Bu seçim ekranı Şekil 6'daki gibidir. Burada kullanıcıya Şekil 6-b'de gösterildiği gibi klasik eşikleme metodu ve Şekil 6-c'de gösterildiği gibi sütunsal tarama metodu ile oluşturulan akciğer segmentasyon maskelerinden hangisinin kullanılacağını seçmesi için bir seçim ekranı sunulmaktadır.



**Şekil 6:**

*a-Segmentasyon İşlemi Uygulanacak Görüntü*

*b-Otsu Eşikleme Yöntemi ile Elde Edilmiş Segmentasyon Maskesi*

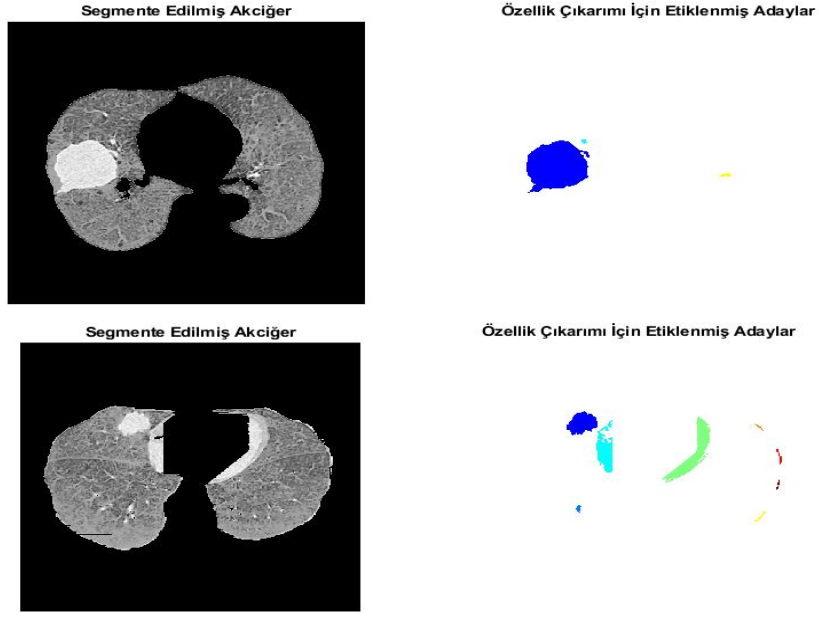
*c-Çalışmada Önerilen Sütunsal Tarama Yöntemi ile Elde Edilmiş Segmentasyon Maskesi*

#### **2.4. Özellik Çıkarımı**

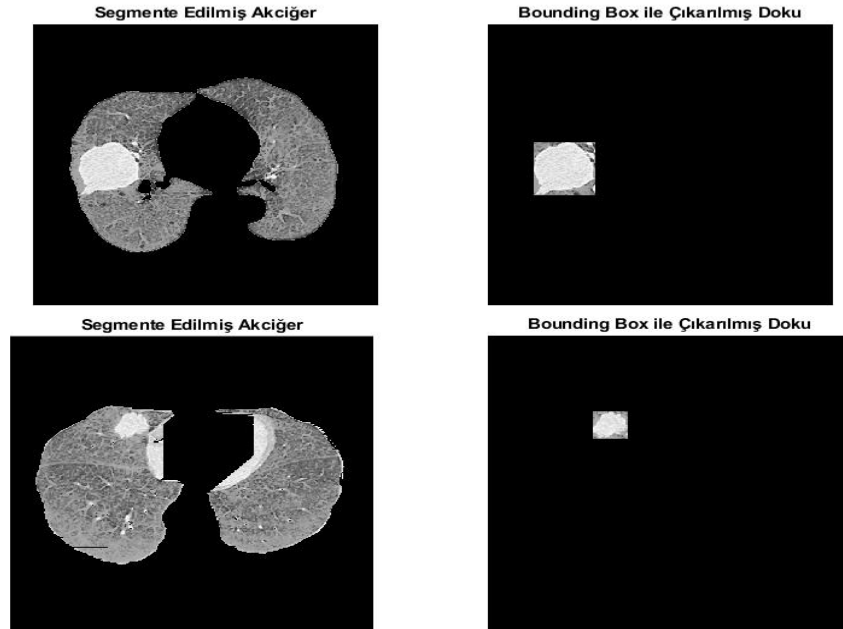
Özellik çıkarımı aşamasında sınıflandırma işlemleri için gerekli olan öznelikler elde edilir. Makine öğrenmesi ve sınıflandırma araçları görüntü işleme uygulamalarında direk olarak görüntü üzerinden değil görüntünün içerdiği özellikler üzerinden çalışmaktadır.

Özellik çıkarma işlemleri için özellikleri çıkarılacak parçalar segmente edilmiş akciğer görüntüleri üzerinden otsu eşikleme metodu ile seçilmiştir. Eşikleme sonucu oluşan görüntü Şekil 7'deki gibidir. Elde edilen bu görüntü üzerindeki her bir parça görüntü etiketleme yöntemi ile otomatik olarak etiketlenmiştir. Etiketleme işlemi sonrasında her görüntü içindeki her bir parça için ayrı ayrı sınırlayıcı kutu haritaları çıkartılmıştır. Bu sınırlayıcı kutu haritaları lojik formatta olup tümör ve nodüllerin görüntülerinin çıkarılması için birer maske olarak kullanılmıştır. Bu maskeleri kullanarak elde edilen tümör görüntüleri Şekil 8'de gösterilmektedir. Tümör etrafındaki küçük bölgelere veya tümöre benzer yoğunluk gösteren diğer akciğer bölgeleri nodül olarak değerlendirilmiştir. Özellik çıkarımı işlemleri bu şekiller üzerinde yapılmaktadır.





**Şekil 7:**  
*Segmente Edilmiş Akciğer Görüntüleri Üzerinden Etiketlenmiş Bölgelerin Çıkarımı*



**Şekil 8:**  
*Adayların Bounding Box ile Özellik Çıkarımı Ayrılması*

Doku analizi görüntüdeki gri renk seviyesindeki komşulukların mekansal dağılımı ile karakterize edilen tıbbi görüntülerde de oldukça sık kullanılan bir özellik çıkarma metodudur. Doku analizinde en çok kullanılan matris Gri Seviye Eş Oluşum Matrisi (GLCM) olup bu matris her bir komşu ilişki için gri renk seviyesinin beraber görülme sayısını yöne göre düzenleyen bir matristir(Bino vs diğ. ,2012).

Gri düzeyi  $i$  ve  $j$  olan iki piksel, bu durumda konumsal yönün  $\theta$ , uzaklığın  $d$  olma olasılığı  $P(i,j,d,\theta)$  olarak gösterilir. Genellikle  $\theta$  4 değere sahiptir:  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  ve  $135^\circ$ .  $M(k,l)=i$  ve  $M(m,n)=j$  piksel çiftinin bu dört yönde görünme olasılığı aşağıdaki gibidir.

$$P(i,j,d,0^\circ)=\#\{(k,l),(m,n) \mid k-m=0,|l-n|=d\}$$

$$P(i,j,d,45^\circ)=\#\{(k,l),(m,n) \mid k-m=d, l-n=d\}$$

$$P(i,j,d,90^\circ)=\#\{(k,l),(m,n) \mid |k-m|=d, l-n=0\}$$

$$P(i,j,d,135^\circ)=\#\{(k,l),(m,n) \mid k-m=-d, l-n=-d\}$$

Bu çalışmada GLCM matrisi  $0^\circ$  ye göre elde edilmiş ve matrisin aşağıda belirtilen dört özelliği kullanılmıştır. Her bir özelliğin hesap formülleri özellikler içerisinde verilmiştir.

Kontrast: Gri düzeydeki eş oluşum matrisindeki yerel değişikliklerin ölçüsünü verir.

$$Kontrast = \sum_{i,j} |i - j|^2 p(i,j) \quad (1)$$

Korelasyon: Görüntü üzerindeki ilgili pikselin gri seviye değeri ile belirlenen renk arasındaki lineer bağımlılığın ölçüsünü verir.

$$Korelasyon = \sum_{i,j} \left( \frac{(i - \mu_i)(j - \mu_j)p(i,j)}{\sigma_i \sigma_j} \right) \quad (2)$$

Enerji: Açısal ikinci momentum olarak da bilinen bu özellik GLCM matrisindeki elemanların karelerinin toplamı olarak hesaplanır.

$$Enerji = \sum_{i,j} p(i,j)^2 \quad (3)$$

Homojenite: Görüntünün farklı bölgelerindeki benzerliklerin bir ölçüsüdür. Verilen denklemlerde,  $i$  ve  $j$  matrisin satır ve sütun indislerini,  $p$  belirttiği GLCM elemanını ve  $\sigma$  standart sapmayı  $\mu$  ise ortalamayı gösterir.

$$Homojenite = \sum_{i,j} \frac{p(i,j)}{1 + |i - j|} \quad (4)$$

Bu çalışmada bu özelliklerin yanı sıra etiketlenmiş görüntü parçalarının alanları, konveks alanları, sağlamlık oranları(solidity), dış merkezlilik oranları(eccentricity), entropileri, standart sapmaları, ortalamaları, çarpıklık oranları(skewness), basıklık oranları(kurtosis) gibi özellikler de hesaplanmıştır. Ayrıca dalgacık dönüşümü kullanılarak görüntünün dalgacık dönüşümü sonrasındaki ortalama piksel yoğunluğu ve standart sapma özellikleri elde edilmiştir. Bu çalışmada Haar-Dalgacık dönüşümü metodu kullanılmıştır. Haar-Dalgacık dönüşümü görüntü matrisini Haar fonksiyonu ile ayrıştırarak yeniden birleştirme işlemi ile görüntü içerisindeki

özellikleri diyagonal ayrışım yaptırarak özelliklerin karakteristiğini güçlendirmektedir ( Piotr ve Agnieszka, 2004). Belirtilen tüm bu özelliklerin t-test ile anlamlılık düzeyleri istatistiksel olarak hesaplanarak çalışmada sadece anlamlı özellikler kullanılmıştır. Anlamlılık seviyesi için  $p < 0,05$  olmak üzere elde edilen özellikler ve bunlara ait örnek değerler Tablo 1’de verilmiştir.

**Tablo 1. Tümör Doku ve Normal Doku Özellikleri**

Özellik	Tümör	Normal
Korelasyon	0,914	0,679
Enerji	0,113	0,331
Alan	4294	45
Entropi	4,78	3,60
Kurtosis	1,710	2,356
Skewness	-0,0600	-0,7969
Solidity	0,593	0,978
Dalgacık Standart Sapması	0,420	0,375
Dalgacık Ortalaması	1,281	1,535

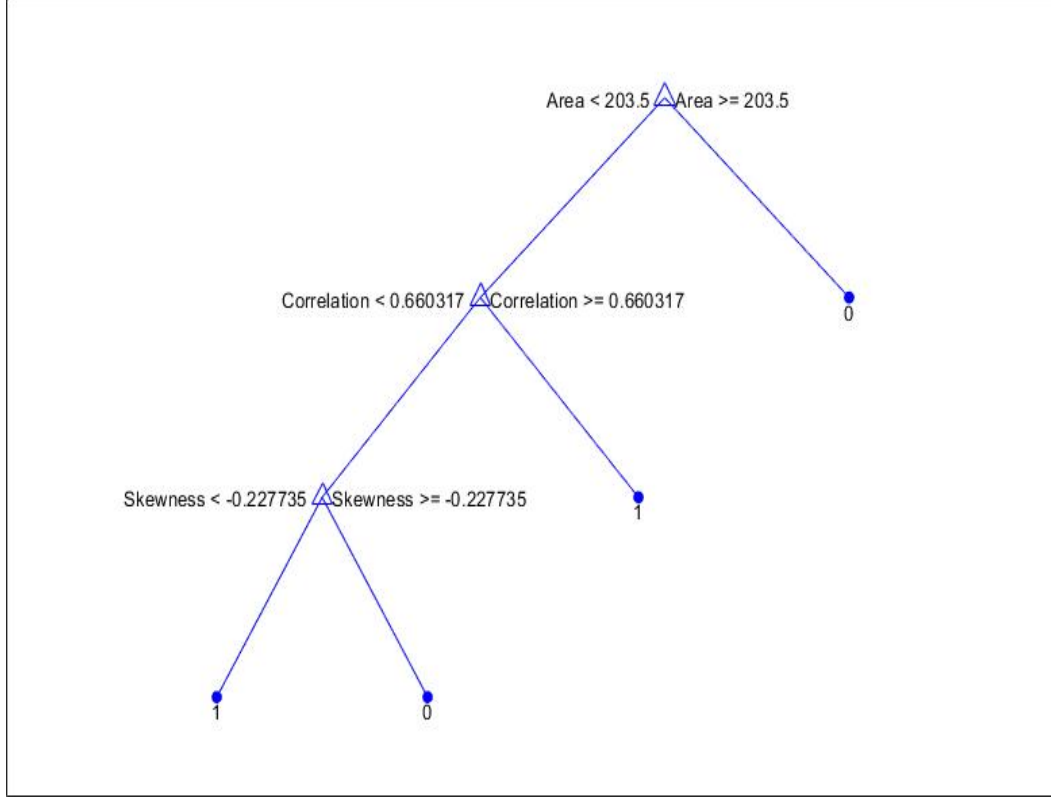
Özellik çıkarımı sonrasında veri setinde radyologlar tarafından belirlenen tümör dokularının ve diğer nodüllerin ayrıştırılması sonucu sınıflandırma aşaması için veri seti oluşturulmuştur. Oluşturulan normal veri seti içerisinde segmentasyon sonucu elde edilen görüntüler üzerinde bir takım dış dokuların segmentasyon alanlarına girdiği tespit edilip bu dokuların bulunduğu etiketler normal set içerisinde alanları 500 pikselden fazla olanların çıkartılması işlemine tabii tutularak segmentasyon içerisinde olmaması gereken parçaların elenmesi sonucu istenilen eğitim setleri elde edilmiştir. 26 görüntü üzerinde yapılan etiketleme işlemleri sonucunda 26 adet tümör dokusu ve 177 adet nodül tespiti yapılmıştır. Tümör dokusu ve tümör olmayan nodüllerin farklı şekilde etiketledikten sonra eğitim için kullanılacak veri kümesi elde edilmiştir.

## 2.5. Sınıflandırma

Elde edilen özneliklerin sınıflandırılması için SVM, KNN, karar ağacı, diskriminant sınıflandırıcı algoritmaları olmak üzere 4 farklı makine öğrenme algoritması uygulanıp sonuçları karşılaştırılmıştır. Sınıflandırma işlemi Matlab Learner App’ ta yapılmış olup kullanılan yöntemler, kısa açıklamaları ve uygulanan parametre değerleri aşağıda verilmiştir.

*Karar Ağaçları;* Karar ağaçları veri madenciliğinde verileri sınıflandırmak için kullanılan kuvvetli bir sınıflandırma yöntemidir. Bu çalışmada CART algoritması ile oluşturulan karar ağacı kullanılmıştır. İkili karar ağaçları, bir özellik uzayının iki alt uzaya tekrarlanan bölümlerinden oluşur ve sınıflarla ilişkili uç düğümlerin sonuçlarının seçilmesi ile sınıflandırma yapar. CART algoritmasında doğru sınıflandırmanın bilindiği bir eğitim verisi örneğini kullanarak bir ikili karar ağacının tanımlanmasını ve oluşturulmasını içerir (Helio ve Robin, 2003). Ayrım kriteri olarak gini çeşitlilik indeksi kullanılmıştır. Karar ağacı modeli Şekil 9’da

gösterilmektedir. Gösterilen sınıflardan 0 etiketi tümörlü dokuları, 1 etiketi diğer nodülleri göstermektedir.



**Şekil 9:**  
Karar Ağacı Modeli

*Destek Vektör Makinesi (SVM) sınıflandırıcısı;* Yaygın kullanılan güçlü bir sınıflandırma algoritmasıdır. SVM, doğrusal SVM ve doğrusal olmayan SVM olmak üzere iki gruba ayrılabilir. Bu iki grup, ayrı ayrı hiper düzlem ve porsiyon tuzağı kullanarak bilgileri izole etmek için alternatif bir stratejiye sahiptir. Doğrusal SVM verileri 2 boyutta birbirinden ayırırken doğrusal olmayan SVM ise kernel ayarlamaları ile verileri daha yüksek boyutlarda ayırabilmektedir (Muhamad ve diğ. ,2020). Bu çalışmada lineer olmayan Karesel (Quadratic) SVM algoritması ile sınıflandırma işlemleri yapılmıştır. SVM için kernel fonksiyonu Karesel (Quadratic) olarak seçilerek kutu kısıtlama seviyesi 1 ve kernel skalası 2 olarak belirlenmiştir.

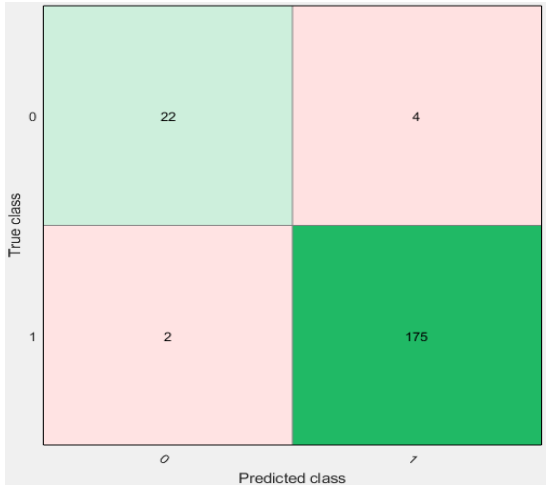
*Diskriminant Sınıflandırıcı;* Çok özellikli sınıflandırma metotları arasında son derece etkili bir sınıflandırma algoritmasıdır. Bir model veya örnek, bir vektör veya m boyutlu uzayda (R<sup>m</sup>) model uzayı olarak adlandırılan bir noktayı temsil eden bir dizi m özelliği ile temsil edilir. Örüntü sınıflandırma sürecinin amacı, bilinmeyen bir desene bir sınıf etiketi atamak için etiketlenmiş desenleri kullanarak bir modeli eğitmektir. Sınıf etiketleri, her sınıf için diskriminant fonksiyonlarını hesaplamak için kullanılan etiketli modellerin sınıflarını veya kategorilerini temsil eder. Diskriminant fonksiyonları daha sonra her sınıf için karar sınırlarını ve karar bölgelerini belirlemek için kullanılır. İki tip diskriminant sınıflandırıcı tipi vardır. Bunlar lineer ve Karesel diskriminant sınıflandırıcılardır. Karesel diskriminant sınıflandırıcıda sınır çizgileri Karesel çizgilerken lineer diskriminant sınıflandırıcıda bu çizgiler doğrusaldır

(Tharwat, 2016). Bu çalışmada Karesel (Quadratic) diskriminant sınıflandırıcı algoritması kullanılmıştır. Diskriminant sınıflandırıcıda tam kovaryans kullanılmıştır.

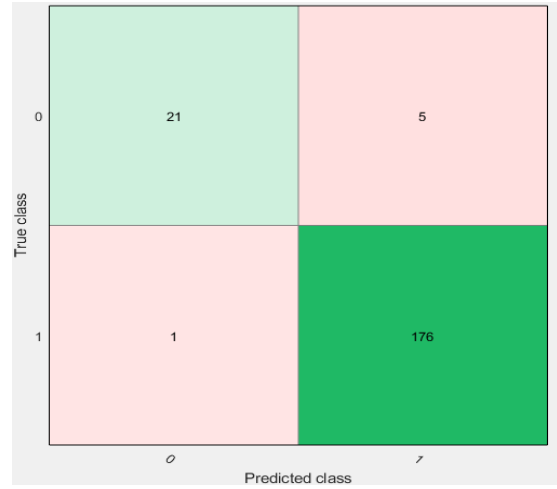
*K-En Yakın Komşu (KNN) sınıflandırıcı*; KNN algoritması veri madenciliğinde oldukça fazla kullanılan bir yöntemdir. K-En Yakın Komşular (kNN), basit ancak birçok durumda etkili olan parametrik olmayan bir sınıflandırma yöntemidir. Bir  $t$  veri kaydının sınıflandırılması için, en yakın  $k$  komşusu alınır ve bu,  $t$ 'nin bir komşuluğunu oluşturur. Komşuluktaki veri kayıtları arasında çoğunluk oylaması, genellikle mesafeye dayalı ağırlıklandırma dikkate alınarak veya dikkate alınmadan  $t$  için sınıflandırmaya karar vermek için kullanılır (Gongde ve diğ. , 2003). Bu çalışmada en verimsiz sonuçlar KNN sınıflandırma yöntemi ile elde edilmiştir.  $K$  en yakın komşu algoritmasında uzaklık metriği olarak kosinüs, uzaklık ağırlığı eşit ve komşu sayısı 10 olarak ayarlanmıştır.

### 3. SONUÇLAR ve TARTIŞMA

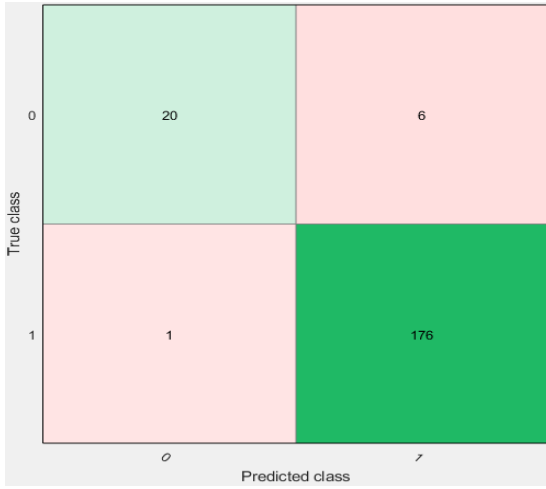
Uygulanan sınıflandırma yöntemleri sonucunda elde edilen doğruluk tabloları Şekil 10, 11, 12 ve 13'te verilmiştir. Doğruluk tablolarında görüldüğü gibi en iyi sonuç veren sınıflandırma algoritması karar ağacı algoritmasıdır. Bu sonuçlar üzerinden hesaplanan veriler Tablo 2'de karşılaştırılmıştır. Tabloda görüldüğü gibi anlamlı özelliklerin sınıflandırılması ile yüksek doğruluk oranları elde edilmiş, en başarılı sonuçları kuadratik diskriminat ve karar ağaçları yöntemleri vermiştir. İki yöntem arasında ise doğru pozitif oranı karar ağaçlarında küçük bir yükseklik gösterirken hassasiyet ise kuadratik diskriminat sınıflandırıcıda küçük miktarda artmıştır.



**Şekil 10:**  
Karar Ağacı Sınıflandırma Sonucu

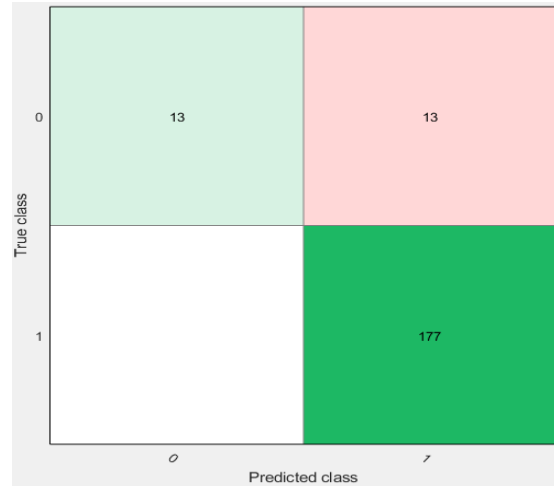


**Şekil 11:**  
Karesel Diskriminant Sınıflandırma Sonucu



Şekil 12:

Karesel SVM Sınıflandırma Sonucu



Şekil 13:

KNN Sınıflandırma Sonucu

Tablo 2. Sınıflandırma Sonuçları

	Kuadratik SVM	Kuadratik Diskriminant	Karar Ağacı	Cosine KNN
<b>Doğru Pozitif Oranı (%)</b>	77	81	85	50
<b>Yanlış Negatif oranı (%)</b>	1	1	1	0
<b>Doğru Negatif Oranı (%)</b>	99	99	99	100
<b>Yanlış Pozitif Oranı (%)</b>	23	19	15	50
<b>Doğruluk (%)</b>	96,6	97	97	93,6
<b>Hassasiyet (%)</b>	95	95	92	100
<b>Duyarlılık(%)</b>	96,7	97,2	97,8	93,2

Yukarıdaki tabloda verilen performans oranları aşağıda belirtildiği şekilde konfüzyon matrislerinden hesaplanmıştır.

A: Gerçekte hasta olup, tanı testi sonucuna göre de hasta olarak belirlenen olgulardır (Doğru pozitif, DP)

B: Gerçekte sağlam oldukları halde testin hatalı olarak hasta dediği olgulardır (Yanlış pozitif, YP)

C: Gerçekte hasta olup, tanı testi sonucuna göre sağlam olarak belirlenen olgulardır (Yanlış negatif, YN)

D: Gerçek tanı sonucuna uygun olarak testinde sağlam dediği gerçek negatif olgulardır (Doğru negatif, DN)

Doğru Pozitif Oranı =  $A / (A+B)$ , Doğru Negatif Oranı =  $D / (C+D)$

Yanlış Pozitif Oranı =  $B / (B+D)$ , Yanlış Negatif Oranı =  $C / (A + C)$

Hassasiyet =  $A / (A+C)$ , Duyarlılık =  $D / (B + D)$ , Doğruluk =  $(A+D) / (A+B+C+D)$

Bu çalışmada Akciğer BT görüntüleri kullanılarak akciğer tümörünün segmentasyonu akciğerde bulunan başka nodül ve lezyonlardan ayrımı ve sınıflandırması basit ve kullanışlı bir

yöntemle gerçekleştirilmiştir. Çalışmada elde edilen sonuçlar değerlendirildiğinde Akciğer BT görüntüleri üzerinde tümör tespiti yapılırken en kritik adımların başında segmentasyon aşaması gelmektedir. Bu aşamada özellik çıkarmak ve aday dokuları çıkarmak için ihtiyaç duyulan ilgi alanı (ROI) ne kadar iyi şekilde elde edilirse sınıflandırma sonucu da aynı oranda artacaktır. İlgi alanını daha iyi elde etmek için aktif görünüş modelleri ve optimizasyon algoritmaları birlikte kullanılabilir (Widodo ve diğ. ,2017). Yarı otomatik bir uygulama olan bu çalışmada elde edilen model farklı veri setlerinde de kullanılabilir ve geliştirilebilir. Daha büyük veri setlerinin kullanılması elde edilen doğruluk değerlerinin artmasına katkıda bulunacaktır. Bunların yanında çıkarılan özellikler temel bileşen analizi, faktör analizi metotlarıyla etkinlik seviyelerine göre gruplanabilir. Özelliklerin sayısının ve türünün artırılması mümkün olmakla birlikte çok fazla özellik işlem yükünü artırarak karmaşıklığa yol açabilecektir.

Elde edilen sınıflandırma sonuçlarına bakıldığında tümör dokularının tespiti yüksek başarı oranlarıyla test edilmiştir. Birçok otomatik tümör tespiti çalışmasına kıyasla son derece iyi başarı oranları basit ve temel adımlarla elde edilmiştir (Adams ve diğ. ,2018). Yapay sinir ağları, derin öğrenme metotları ve daha farklı sınıflandırma yöntemleri ilerideki çalışmalara dahil edilerek tam otomatik yüksek doğrulukla bir tümör tespit yöntemi üzerine çalışmalar geliştirilebilecektir.

## ÇIKAR ÇATIŞMASI

Yazarlar, bilinen herhangi bir çıkar çatışması veya herhangi bir kurum/kuruluş ya da kişi ile ortak çıkar bulunmadığını onaylamaktadırlar.

## YAZAR KATKISI

Yazarlardan Semra İÇER çalışmanın kavramsal ve tasarım süreçlerinin belirlenmesi ve yönetimi, Berat ÖZDET ise veri toplama, veri analizi, ve yorumlama konularında öncelikli olarak rol almış olmakla birlikte tüm aşamalarda yazarlar koordineli olarak çalışarak eşit katkı vermişlerdir.

## KAYNAKLAR

1. Adams, T., Dörpinghaus, J., Jacobs, M., Steinhage, V., (2018) Automated lung tumor detection and diagnosis in ct scans using texture feature analysis and svm. *Communication Papers of the 2018 Federated Conference on Computer Science and Information Systems, POZNAN*, 17, 13–20. doi: 10.15439/2018F176
2. Aniketbombale ,C.G.Patil , (2017) Segmentation of Lung Nodule in CT Data Using K-Mean Clustering, *International Journal of Electrical, Electronics and Data Communication (IJEEDC)*, 5(2), 36-39. doi: IJEEDC-IRAJ-DOIONLINE-6985
3. Vadakkenveetil, B. S., Unnikrishnan, A., Balakrishnan, K., (2012) Grey Level Co-Occurrence Matrices: Generalisation And Some New Features. *International Journal of Computer Science, Engineering and Information Technology (IJCEIT)*, 2(2), 151-157. doi: 10.5121/ijceit.2012.2213
4. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., & Prior, F. (2013) The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *Journal of Digital Imaging*, 26(6), 1045–1057. doi:10.1007/s10278-013-9622-7
5. Elsayed, O., Mahar, K.M., Kholief, M., Khater, H., (2015) Automatic detection of the pulmonary nodules from CT images, *2015 SAI Intelligent Systems Conference (IntelliSys)*, London, UK, 742-746. doi:10.1109/IntelliSys.2015.7361223

6. Eset, K., İçer, S., Karaçavuş, S., Yılmaz, B., Kayaaltı, Ö., Ayyıldız, O., Kaya, E., (2015) Comparison of lung tumor segmentation methods on pet images. *TipTekno-15*, Bodrum, Turkey. 77-80. doi: 10.1109/TIPTEKNO.2015.7374569
7. Guo, G., Wang, H., Bell, D.A., Bi, Y., Greer K.,(2003) KNN Model-Based Approach in Classification, *Lecture Notes in Computer Science*, 2888:986-996. doi:10.1007/978-3-540-39964-3\_62
8. Bittencourt, H. R. and Clarke, R. T., (2003) Use of classification and regression trees (CART) to classify remotely-sensed digital images, *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings*, 6, 3751-3753. doi: 10.1109/IGARSS.2003.1295258.
9. Raif, M., Ismail, N., Nor Azah M. A., Mohd H. F. R., & Tajuddin, S. N., Taib, M. N., (2020) Quadratic tuned kernel parameter in Non-linear support vector machine (SVM) for agarwood oil compounds quality classification. *Indonesian Journal of Electrical Engineering and Computer Science*. 17(3), 1371-76. doi: 10.11591/ijeecs.v17.i3.pp1371-1376.
10. Porwik, P., Lisowska A., (2004) The Haar-Wavelet Transform in Digital Image Processing: Its Status And Achievements, *Machine Graphics & Vision*, 13(1-2),79-98
11. Haralick, R.M., Shanmugam, K., Dinstein, I.,(1973) Textural Features for Image Classification, *IEEE Transactions On Systems Man And Cybernetics*. SMC-3 (6), 610-621.
12. Tharwat, A. (2016). Linear vs. Quadratic Discriminant Analysis Classifier: A Tutorial, *International Journal of Applied Pattern Recognition* , 3(2)2, 145–180. doi: 10.1504/IJAPR.2016.079050
13. Widodo, S., Rohmah, N.R., Handaga, B., (2017) Classification of lung nodules and arteries in computed tomography scan image using principle component analysis, *2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering, Yogyakarta*. 153-158. doi:10.1109/ICITISEE.2017.8285485
14. Widodo, S., Rosyid, I., Faizuddin, M., Roslan, R.B., (2020) Improved accuracy in detection of lung cancer using self organizing map, *Journal of Critical Reviews 2020*, 7(14), 685-689. doi:10.31838/jcr.07.14.121
15. Zhao, Binsheng, Schwartz, Lawrence Kris, Mark (2015). Data From RIDER\_Lung CT. *The Cancer Imaging Archive*. doi: 10.7937/K9/TCIA.2015.U1X8A5NR
16. Zhao, B., James, L. P., Moskowitz, C. S., Guo, P., Ginsberg, M. S., Lefkowitz, R. A., Qin, Y., Riely, G. J., Kris, M. G., & Schwartz, L. H. (2009). Evaluating variability in tumor measurements from same-day repeat CT scans of patients with non-small cell lung cancer. *Radiology*, 252(1), 263–272. <https://doi.org/10.1148/radiol.2522081593>