



Mamogramların Sınıflandırılmasında Dokusal Özelliklerin Etkileri

Canan ORAL*

Amasya Üniversitesi, Teknoloji Fakültesi, Elektrik Elektronik Mühendisliği Bölümü, Amasya
canan.oral@amasya.edu.tr ORCID: 0000-0001-9819-3979, Tel: (358) 260 00 67 (1472)

Ayşe AYDIN YURDUSEV

Amasya Üniversitesi, Teknoloji Fakültesi, Elektrik Elektronik Mühendisliği Bölümü, Amasya
ayse.yurdusev@amasya.edu.tr ORCID: 0000-0001-6516-3976

Erhan BERGİL

Amasya Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Elektronik ve Otomasyon Bölümü, Amasya
erhan.bergil@amasya.edu.tr ORCID: 0000-0021-6526-1661

Geliş 09.03.2018, Kabul Tarihi: 30.03.2018

Öz

Dokusal özellikler mamografi görüntülerini sınıflandırmada yaygın olarak kullanılmaktadır. Bu çalışma, mamografi görüntülerini sınıflandırmada birinci derece ve ikinci derece dokusal özelliklerin etkisini incelemeyi amaçlamaktadır. Bu amaçla öncelikle bilgisayar destekli teşhis (BDT) sistemi tasarlanmıştır. Bilgisayar destekli teşhis sistemi temel olarak, teşhis doğruluğunu artırmayı ve desteklemeyi amaçlamaktadır. Tasarlanan sistem ön işleme, özellik çıkarımı ve sınıflandırma olarak üç adımdan oluşmaktadır. Ön işleme adımında öncelikle gürültü yok edilmiştir. Sonra pektoral kas temizlenmiş ve görüntü iyileştirme gerçekleştirilmiştir. Özellik çıkarımı aşamasında iyileştirilmiş mamogram görüntülerinden birinci derece ve ikinci derece dokusal özellikler hesaplanmıştır. Son aşamada, k en yakın komşuluk (k -EYK) algoritması, destek vektör makineleri (DVM) ve Lineer Diskriminant Analizi (LDA) yöntemleriyle mamografi görüntüleri normal ve anormal olarak sınıflandırılmıştır. Sınıflandırıcı açısından bakıldığında, DVM yönteminin k -EYK ve LDA sınıflandırıcılarına göre daha iyi bir performans sergilediği görülmektedir. LDA için %77.63, k -EYK için %82.89 ve DVM için %85.53 doğruluk ile mamografi sınıflandırılması gerçekleştirilmiştir. Her bir özellik grubu ve sınıflandırıcı için elde edilen sınıflandırma sonuçları karşılaştırılmıştır. İkinci derece dokusal özelliklerin sınıflandırma başarımını artırdığı gözlemlenmiştir.

Anahtar Kelimeler: Mamografi, dokusal özellikler, bilgisayar destekli teşhis, sınıflandırma

* Yazışmaların yapılacağı yazar:

Giriş

Meme kanseri tüm dünyada kadınlarda en yaygın görülen ve ölümlere neden olan kanser türüdür. Meme kanserinin erken tespiti ve tedavisi iyileşme oranını artırmada önemli rol oynamaktadır. Mamografi meme kanserinin erken tespitinde en etkili yöntemlerden biridir. Bilgisayar destekli teşhis (BDT) memedeki anormallikleri tespit etmek için bilgisayarlaştırılmış algoritmaları kullanmaktadır. BDT sistemi yanlış negatif teşhisleri azaltmak için radyoloji uzmanına ikinci görüş sağlamaktadır (Sahiner vd.,1996; Chen vd., 2011).

Literatürde meme kanserinin bilgisayar destekli tespiti için birçok çalışma mevcuttur. Shainer vd. (1996), dokusal özellikleri ve konvolusyonel sinir ağları kullanarak mamogramları normal ve anormal olarak sınıflandırmışlardır. Wei vd. (1995), mamogramlarda normal meme dokusundan kütleleri ayırt etmek için uzaysal gri seviye bağlılık matrisinden çıkarılan özelliklere dayalı çoklu çözünürlüklü doku analizinin uygulanabilirliğini incelemiştir. Chen vd. (2011), sonogram görüntülerinde meme lezyonları için dokusal özellikleri kullanan bilgisayar destekli teşhis sistemi önermişlerdir. Çalışmalarında dokusal özelliklerin teşhis uygulamalarında etkili sonuçlar verdiğini gözlemlemişlerdir. Zheng vd. (2015), gri seviye histogram özellikleri, eş oluşum matrisi özellikleri, akış uzunluğu doku özelliklerinden yararlanarak mamogramlarda parankimal doku analizi gerçekleştirmiştir. Severoğlu (2016) gri seviye eş oluşum matrisleri özellikleri ile istatistiksel özelliklerin sınıflandırma başarımına olan etkilerini incelemiştir.

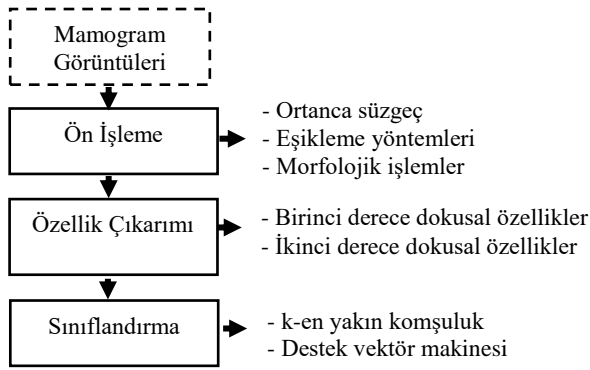
Mamografilerdeki dokusal değişimler mamografideki anormallikleri belirlemek için yararlı bilgiler içermektedir (Sahiner vd.,1996; Chen vd.,2011).

Literatürde farklı özellik gruplarının

kombinasyonları kullanılarak daha yüksek doğrulukta sınıflandırma işlemleri gerçekleştirilmiştir. Bu çalışmada, yaygın olarak kullanılan birinci derece dokusal özellikler ile ikinci derece dokusal özelliklerin mamogram görüntülerinin sınıflandırılması üzerindeki etkileri incelenmiştir. Bu amaçla, sınıflandırma işlemi için sadece dokusal özellik çıkarım yöntemleri kullanılmıştır. Öncelikle, mamogram görüntüleri üzerinde ön işleme yapılarak görüntülerin ilgili bölümleri alınmıştır. Elde edilen her bir mamogram görüntüsünden birinci derece ve ikinci derece dokusal özellikler hesaplanarak özellik kümeleri oluşturulmuştur. Oluşturulan özellik kümeleri yağın olarak kullanılan k-En Yakın Komşuluk (k-EYK) algoritması, Destek Vektör Makineleri (DVM) ve Lineer Diskriminant Analizi (LDA) kullanılarak sınıflandırma gerçekleştirilmiştir. Birinci derece ve ikinci derece dokusal özelliklerin sınıflandırma başarımına olan etkileri ve sınıflandırıcıların performansları karşılaştırmıştır.

Materyal ve Yöntem

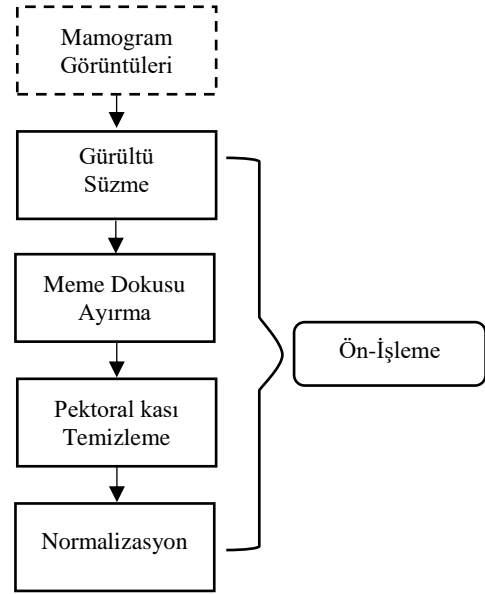
Bu çalışmada mamogram görüntüleri MIAS (Mammographic Image Analysis Society Mini Mammographic Database) veri tabanından alınmıştır. Veri tabanındaki görüntüler 50µm x 50µm piksel çözünürlüğüne sahiptir. Her görüntü 1024x1024 piksel boyutunda ve her piksel 8 bit ile temsil edilmektedir. Veri tabanı 208'i normal, 63'ü iyi huylu ve 51'i kötü huylu olmak üzere 322 görüntü içermektedir. (Suckling vd., 1994). Bu çalışmada birinci ve ikinci derece dokusal özelliklerin mamogramların sınıflandırılması üzerindeki etkilerini incelemek için bir bilgisayar destekli teşhis sistemi (BDT) tasarlanmıştır. Tasarlanan sistem ön işleme, özellik çıkarımı ve sınıflandırma olarak üç aşamadan oluşmaktadır. Tasarlanan BDT sistemin blok diyagramı Şekil 1'de verilmiştir.



Şekil 1. Tasarlanan BDT sistemi blok diyagramı

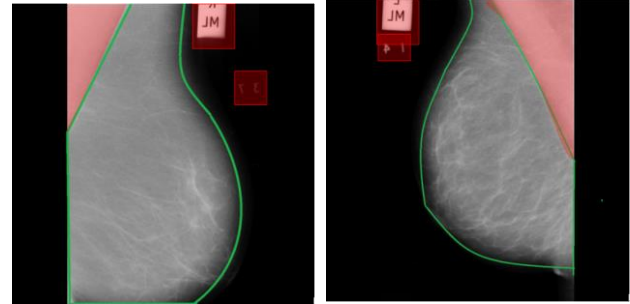
Ön işleme

Sayısal görüntülerde ön işleme gürültüyü bastırmak ve görüntü kontrastını iyileştirmek için gerçekleştirilmektedir (Li vd., 2013; Sreedevi ve Sherly, 2015; Gonzalez ve Woods, 2014). Bir mamogram imgesinde yumuşak meme dokusu, meme derisi, pektoral kas dokusu ve etiketler bulunmaktadır. Buradaki her bilginin bir amacı olmasına karşın tümör, kalsifikasyon gibi yumuşak doku lezyonları incelendiğinden dolayı bu bilgilerin temizlenmesi gerekmektedir (Saltanat vd., 2010; Ganesan vd., 2013; Sreedevi ve Sherly, 2015; Pak vd., 2015). Etiketler, ışık patlamaları, mamografi cihazından kaynaklı gürültüler ve pektoral kas alanı yüksek yoğunluk değerine sahip oldukları için ve BDT sistemlerinde yanlış sonuçlara neden olabileceklerinden dolayı mamografi görüntülerinden temizlenmeleri gerekmektedir (Saltanat vd., 2010; Ganesan vd., 2013; Pak vd., 2015). Örneğin sağlıklı bir memede hiç lezyon yokken pektoral kas yüksek yoğunluk içerdiği için tümör varmış gibi algılanabilmektedir. Literatürde pektoral kasın temizlenmesi ile ilgili parlaklık tabanlı, Hough dönüşümlü, dalgacık dönüşümlü gibi farklı yöntemlerle yapılmış çalışmalar mevcuttur (Saltanat vd., 2010; Ganesan vd., 2013; Pak vd., 2015). Bu çalışmada ise ağırlıklı olarak morfolojik işlemler kullanılmıştır. Ön işleme aşamaları Şekil 2’de gösterilmiştir.



Şekil 2. Ön işleme aşamaları

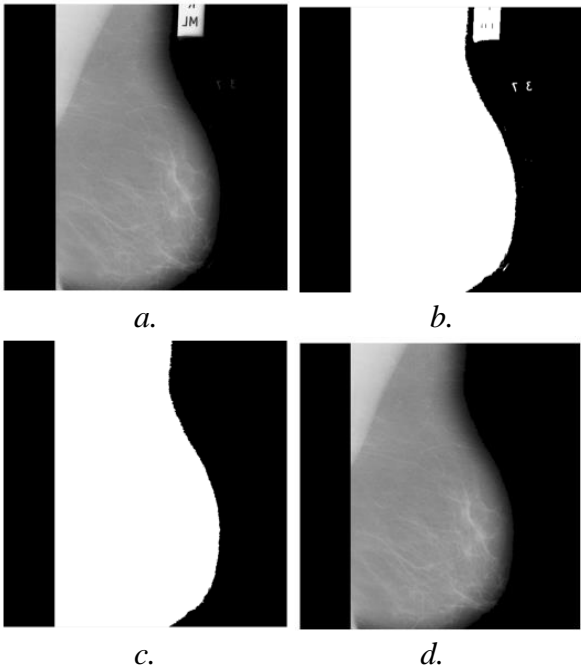
Şekil 3’te bir mamografi imgesinden yukarıda bahsedilen temizlenmesi gereken alanlar kırmızı ile gösterilmiştir. Gerçekleştirilen ön işlem sonucunda kırmızı alanlar temizlenerek sadece meme dokusu elde edilmiştir.



Şekil 3. Mamografi segmentasyon alanları

Gürültü süzme aşamasında, veri tabanından alınan mamografi imgesini küçük gürültülerden arındırmak için 3×3 boyutunda bir orta süzgeç kullanılmıştır. Orta süzgeç, süzgeç matrisi boyutunda gruplarla çalışarak, bir pikselin komşuları ile olan ilişkisine bakarak gruptaki değerlerden orta sırada olanı merkez piksele atayan bir filtredir (Gonzalez ve Woods, 2014). Küçük gürültülerden temizlenen imgenin bir sonraki aşamada yönlendirme etiketi, düşük yoğunluk değerli etiket, tarama ve kayıt gürültülerinden temizlenmesi gerekmektedir (Mustra vd., 2006). Bunun için sunulan ön işleme eşikleme ve morfolojik işlemler kullanılmıştır. Şekil 4.a’daki imge

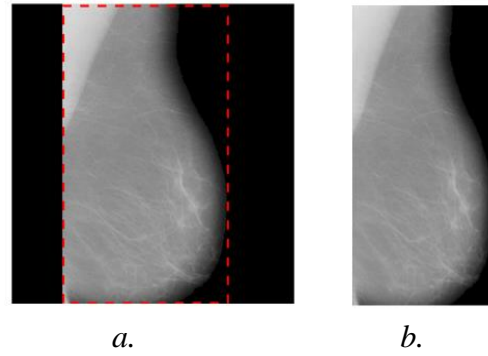
düşük bir eşik değeri kullanılarak siyah beyaz hale dönüştürülmüştür (Şekil 4.b). Siyah beyaz imgede sadece meme dokusunu ele almak için etiket ve rakamlar silinmiştir. Bunun için morfolojik işlemler yapılmış, en büyük siyah-beyaz alan bulunmuş ve bu alandan küçük alanların silinmesi gerçekleştirilmiştir. (Şekil 4.c) Son aşamada ise bulunan siyah-beyaz görüntü maske olarak kullanılmış, orijinal (Şekil 4.a) görüntü ile çarpılarak mamografide sadece bir bütün olarak meme dokusu elde edilmiştir (Şekil 4.d).



Şekil 4. a. Etiket ve yapay gürültülü mamogram görüntüsü b. Eşikleme yöntemi ile mamogramın siyah-beyaza dönüştürülmesi c. Etiketlerin ve yapay gürültülerin temizlenmesi d. Temizlenen mamogram görüntüsü.

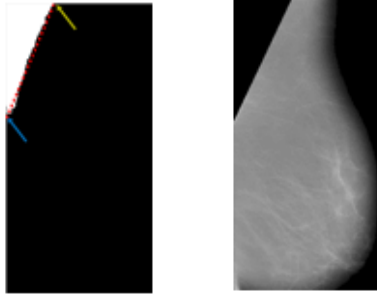
Etiketlerden temizlenen mamogram görüntüsü için bir sonraki önemli adım pektoral kasın temizlenmesidir. Pektoral kas mamogram görüntülerinde memenin sağ ya da sol üst köşesinde yer almaktadır. Fakat MIAS veri tabanındaki mamografi görüntüleri 1024x1024 piksel boyutunda eşitlendiği ve mamogram bu boyutta ortalandığı için pektoral kasın imgenin köşesinde olup olmayacağı bilinmemektedir, çünkü her bir mamogramda bilinmeyen bir pozisyonlandırma ve yerleştirme yapılmıştır. Bu

sebepten dolayı mamogram sadece meme dokusunu ele alacak yani arka plan alanları kesecek şekilde yeniden boyutlanmalıdır (Şekil 5.a- kırmızı çizgili alan). Şekil 5.b haline getirilen imge üst sağ veya sol köşelerde pektoral kası aramak için hazır hale getirilmiştir. Gerekli işlemleri hızlandırmak ve bir standarda kavuşturmak amacı ile görüntü 512x256 piksel boyutlarına indirgenmiştir. Yeniden boyutlandırılan mamografi imgesinde pektoral kas köşede aranabilir.



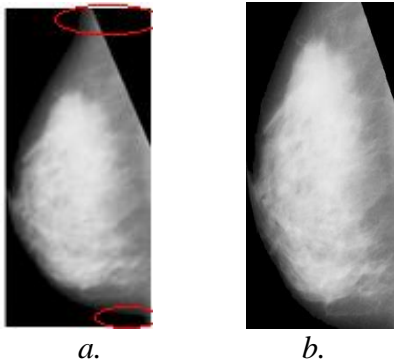
Şekil 5. a. Mamografi arka planının kırılması gereken alan b. Kırılmış meme alanı.

512x256 boyutlu imgede en yüksek yağlık değerli alanlar eşikleme yöntemi ile bulunmuş, eşik değerinden küçük pikseller en düşük değere, eşik değeri ve yukarısındaki değerleri ise en yüksek değere atanmıştır. Böylelikle görüntünün ikili sistem bir eşleniği oluşturulmuş ve morfolojik işlemler için hazırlanmıştır. En yüksek yağlık içeren alanlardan pektoral kasa ait olmayan kitle gibi başka yapılar var ise doldurma, aşındırma, yok etme gibi morfolojik işlemlerle elenmiştir. Sonuçta pektoral kas tespit edilmiştir (Şekil 6.a). Pektoral kasın kapladığı uç noktalarındaki piksel koordinatları $I(x_1, y_1)$ ve $I(x_2, y_2)$ şeklinde tanımlanmıştır. Bu koordinat noktaları kullanılarak pektoral kasın eğik kenarı doğru denklemi ile düz olarak varsayılmıştır. Düz bir üçgen şekilde elde edilen pektoral kas görüntüsü ile temizlenen 512x256 boyutundaki görüntü maskelenerek pektoral kas temizlenmiş mamogram görüntüsü oluşturulmuştur (Şekil 6.b).



a. b.
Şekil 6. a. Pektoral kasın koordinatları b. Pektoral kasın temizlenmesi

Pektoral kasın temizlenmesiyle oluşan mamogram görüntüsünde meme altında veya meme üstünde ilgisiz alanlar, boşluklar oluştuğunda (Şekil 7.a) bu alanlar yok edilerek meme görüntüsü elde edilmiştir. Son olarak elde edilen görüntüler tekrar 512x256 boyutuna getirilerek ön işlem aşaması tamamlanmıştır (Şekil 7.b). Ön işleme aşamasında tüm mamografiler birbirleriyle benzer biçime getirilerek özellik çıkarımı ve sınıflandırma için hazır hale getirilmiştir.



a. b.
Şekil 7. Temizlenen meme alanı b. Yeniden boyutlandırma

Özellik Çıkarımı

Dokusal özellikler mamogramları sınıflandırmak için yaygın olarak kullanılmaktadır. Bu çalışmada iyileştirilmiş mamogram görüntülerinden yaygın olarak kullanılan 6 birinci derece dokusal özellik ve 4 ikinci derece dokusal özellik hesaplanmıştır.

Birinci derece dokusal özellikler

Birinci derece dokusal özellikler, doğrudan mamogram görüntülerinden hesaplanan,

görüntünün gri seviye histogramının istatistiksel özelliklerine dayalıdır. Bu çalışmada iyileştirilmiş mamogramlardan ortalama, standart sapma, yamukluk, basıklık, homojenlik ve entropi özellikleri çıkarılmıştır.

Bu özellikler;

$$\text{Ortalama;} \quad m = \sum_{i=0}^{L-1} ip_i \quad (1)$$

$$\text{Standart sapma;} \quad \sigma = \sqrt{\sum_{i=0}^{L-1} (i-m)^2 p_i} \quad (2)$$

$$\text{Yamukluk;} \quad \frac{1}{\sigma^3} \sum_{i=0}^{L-1} (i-m)^3 p_i \quad (3)$$

$$\text{Basıklık;} \quad \frac{1}{\sigma^4} \sum_{i=0}^{L-1} (i-m)^4 p_i \quad (4)$$

$$\text{Homojenlik;} \quad \sum_{i=1}^{L-1} p_i^2 \quad (5)$$

$$\text{Entropi;} \quad - \sum_{i=0}^{L-1} p_i \log_2 p_i \quad (6)$$

olarak ifade edilmektedir.

Yukarıda verilen eşitliklerde, N toplam piksel sayısı ve L görüntüdeki gri seviyelerin sayısı olmak üzere $\{h_i, i=0,1,\dots,L-1\}$ görüntünün gri seviye histogramını, $p_i=h_i/N$, $\{p_i, i=0,1,2,\dots,L-1\}$ görüntünün normalize edilmiş histogram kümesini ifade etmektedir (Haralick vd., 1973; Gonzalez ve Woods, 2014).

İkinci derece dokusal özellikler

İkinci derece doku hesaplamalar, asıl görüntüden oluşturulan yardımcı matrisler üzerinden gerçekleştirilmektedir. Bu amaçla iyileştirilmiş mamogram görüntülerinden gri seviye eş-oluşum matrisi hesaplanmıştır. Gri seviye eş-oluşum matrisi (GSEOM) görüntüyü oluşturan piksellerin ikili komşuluğa göre gri seviyelerinin kombinasyonlarının hangi sıklıkta olduğunun tablolandığı iki boyutlu matristir. Bu matrisi oluşturan değerler üzerinden dokuya özgü hesaplamalar yapılmaktadır (Haralick vd., 1973; Chen vd., 2011).

GSEOM ikinci derece olasılık yoğunluk fonksiyonudur ve $p(i,j|d, \theta)$ olarak ifade edilir. $p(i,j|d, \theta)$ uzaklık d ve açı θ iken $p(i,j)$ i gri seviyesinden j gri seviyesine geçiş olasılığını

gösterir. İyileştirilmiş mamogram görüntüsünden $d=1$ ve $\theta=0^\circ, 45^\circ, 90^\circ, 135^\circ$ yönelme açıları için dört farklı gri seviye eş oluşum matrisi (GSEOM) hesaplanmıştır. Haralick vd. (1973) GSEOM için 14 özellik önermiştir. Bu çalışmada hesaplanan her bir eş oluşum matrisi için literatürde en yaygın kullanılan karşıtlık, korelasyon, enerji ve homojenlik özellikleri hesaplanmıştır. İşlem karmaşıklığını azaltmak ve özelliklerin açıdan bağımsız olmasını sağlamak amacıyla her 4 farklı matris için hesaplanan özelliklerin ortalaması alınmıştır. Hesaplanan özellikler;

$$\text{Karşıtlık; } \sum_{i,j} |i-j|^2 p_{ij} \quad (7)$$

$$\text{Korelasyon; } \frac{\sum_{i,j} (ij) p_{ij} - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (8)$$

$$\text{Enerji; } \sum_{i,j} p_{ij}^2 \quad (9)$$

$$\text{Homojenlik; } \sum_{i,j} \frac{p_{ij}}{1+|i-j|} \quad (10)$$

olarak ifade edilmektedir. μ_x, μ_y, σ_x ve σ_y değerleri p_{ij} 'nin satır ve sütunlarının ortalaması ve standart sapmasıdır (Haralick vd., 1973; Chen vd., 2011).

Sınıflandırma

Sınıflandırma adımında mamogramları normal ve anormal olarak sınıflandırmak için en yakın komşuluk yöntemi, destek vektör makineleri ve lineer diskriminant analizi kullanılmıştır.

a) k-En yakın Komşuluk Yöntemi (k-EYK)

k-en yakın komşuluk yöntemi (k-EYK); verinin eğitim kümesindeki örneklerle arasındaki benzerliğe bakılarak sınıflandırma gerçekleştiren örnek tabanlı bir sınıflandırıcıdır. Sınıflandırıcı bir verinin hangi sınıfa ait olacağına karar verirken eğitim kümesinden ilgili örneğe en yakın k örnek belirler. Sınıf etiketine en yakın komşuların sınıf etiketlerinin çoğunluk oylamasına göre karar verilir. Çoğunluk oylamasına katılacak olan komşu sayısı (k) sınıflandırıcı performansını önemli ölçüde etkilemektedir (Mitchell, 1997; Han vd., 2006). Bu çalışmada k=3 seçilmiştir.

b) Destek Vektör Makinaları (DVM)

Destek vektör makineleri (DVM); istatistiksel öğrenme temeline dayalı bir denetimli öğrenme yöntemidir. Geleneksel öğrenme yöntemlerine göre oldukça avantajlara sahiptir. 1960'lı yılların sonlarında Vapnik tarafından teorik temelleri atılmıştır. 1995 yılında ise yine Vapnik tarafından ilk defa bir sınıflandırma probleminin çözümünde kullanılmıştır (Vapnik, 1998; Zheng ve Xue, 2009). DVM, yapısal risk azaltmaya dayalı veri bölme tabanlı bir öğrenme yöntemi olarak bilinmektedir. DVM'lerin genelleme hataları, verileri ayırdığı sınır genişliği ile ilişkilidir. DVM, sınıflandırma işlevi (hiper-düzlem) yardımıyla veri uzayını bölerek bilinmeyen verilerin sınıf etiketlerini tanımlamaktadır. Bu amaçla DVM, veriyi bölmek için en uygun ayırıcı hiper-düzlemi kullanmaktadır. DVM, el yazısını tanıma, görüntü sınıflandırma, ses tanıma ve biyolojik bilgi işleme gibi birçok uygulama alanına sahiptir (El-Naqa vd., 2002; Zheng ve Xue, 2009; Abe, 2010).

b) Lineer Diskriminant Analizi (LDA)

Doğrusal diskriminant analizi (LDA), iki ya da daha fazla sınıfa ait örnekleri ayırmak için doğrusal bir sınıflandırıcı olarak kullanılabilen özelliklerin doğrusal birleşimini bulmaya çalışan bir yöntemdir. Bir başka ifadeyle LDA, sınıfları birbirinden en iyi ayırt eden uzaya ait vektörleri bulmaya çalışır (Panahi vd., 2011; Junoh and Mansor, 2012). Bu yöntem; farklı sınıflara ait veri noktalarını birbirinden uzaklaştıran ve aynı sınıfta bulunan veri noktalarını birbirine yakınlaştıran, mevcut verilerin bileşkesi olan yeni bir değişken üretmektedir. İlk tanımlanan sınıfların yeni değişkene göre aralarındaki farklılıkları maksimumlaştırmayı amaçlar (Fielding, 2007; Panahi vd., 2011; Fraz vd., 2013).

Uygulama ve Başarımlar

Bu çalışmada birinci ve ikinci derece dokusal özelliklerin, mamogramların normal ve anormal olarak sınıflandırılmasında başarımlar etkileri incelenmiştir.

Ön-işleme adımından sonra elde edilen mamogram görüntülerinden hesaplanan birinci ve ikinci derece dokusal özellikler kullanılarak üç ayrı özellik kümesi oluşturulmuştur. Birinci özellik kümesi, 6 birinci derece, ikinci özellik kümesi 4 ikinci derece, üçüncü özellik kümesi ise 6 birinci ve 4 ikinci derece dokusal özellik olmak üzere toplam 10 özellikten oluşmaktadır.

Çalışmada, 227 mamogram görüntüsü kullanılmıştır. Eğitim seti ve test seti çapraz geçerlilik yöntemi ile belirlenmiştir. Eğitim ve test verileri sayısı Tablo 1’de verilmiştir.

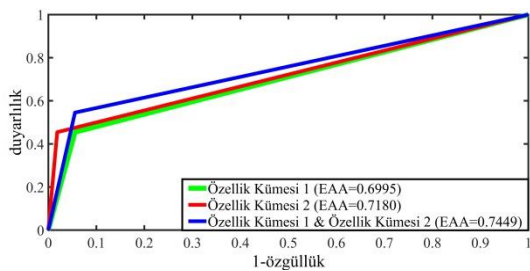
Tablo 1. Eğitim ve test verileri

	Normal	Anormal
Eğitim verisi	107	44
Test verisi	54	22
Toplam	151	66

Sınıflandırma aşamasında; oluşturulan üç ayrı özellik kümesi ayrı ayrı olarak k-EYK, DVM, ve LDA sınıflandırıcılara uygulanarak sınıflandırma gerçekleştirilmiştir ve sonuçları değerlendirilmiştir. Önerilen sistemin performansı işlem karakteristik eğrileri (İKE) ile analiz edilmiştir. Bu çalışmada k=3 seçildiğinde en yüksek sınıflandırma başarımı elde edilmiştir. k-EYK sınıflandırma doğrulukları Tablo 2’de verilmiştir. k-EYK için İKE ve eğri altındaki alan (EAA) değerleri Şekil 8’de verilmiştir.

Tablo 2. k-EYK sınıflandırma doğrulukları

Özellik Kümesi	Doğruluk (%)
Özellik kümesi 1	80.26
Özellik kümesi 2	82.89
Özellik kümesi 1 & Özellik kümesi 2	82.89



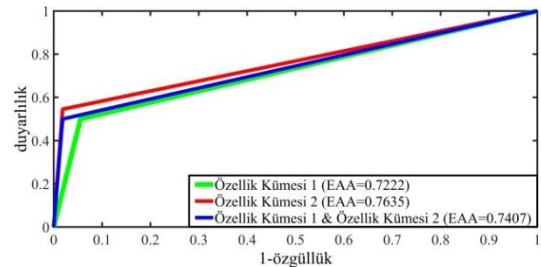
Şekil 8. k-EYK için İKE ve eğri altındaki alan (EAA) değerleri

k-EYK sınıflandırıcısı ile en yüksek sınıflandırma doğruluğuna (% 82.89), Özellik Kümesi 2 ve iki özellik kümesinin birlikte kullanıldığı durumlarda ulaşılmıştır. Bu iki özellik kümesi için sınıflandırma doğruluğu eşit olmasına rağmen, EAA değerleri dikkate alındığında iki özellik kümesinin birlikte kullanıldığı durum için daha yüksek değer elde edilmiştir. Bunun nedeni iki özellik kümesinin birlikte kullanıldığı sınıflandırma işleminin duyarlılığının daha yüksek olmasıdır. En düşük sınıflandırma doğruluğu ve en küçük EAA değeri Birinci dereceden dokusal özellikler için elde edilmiştir.

Bu çalışmada DVM için çekirdek olarak, radyal tabanlı çekirdek seçilmiştir. DVM sınıflandırma doğrulukları Tablo 3’de verilmiştir. DVM için İKE ve eğri altındaki alan (EAA) değerleri Şekil 9’da verilmiştir.

Tablo 3. DVM sınıflandırma doğrulukları

Özellik Kümesi	Doğruluk (%)
Özellik kümesi 1	81.58
Özellik kümesi 2	85.53
Özellik kümesi 1 & Özellik kümesi 2	84.21



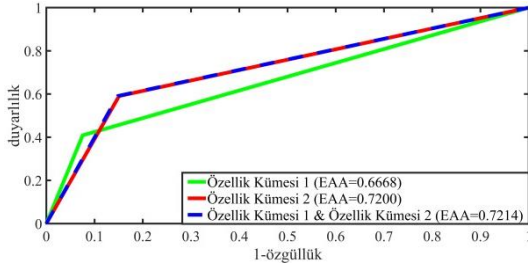
Şekil 9. DVM için İKE ve eğri altındaki alan (EAA) değerleri

DVM sınıflandırma sonuçlarına göre, en yüksek sınıflandırma doğruluğuna (% 85.53) ve EAA değerine (0.7635), Özellik Kümesi 2 (ikinci derece dokusal özellikler) ile ulaşılmıştır. En düşük sınıflandırma doğruluğu ve en küçük EAA değeri Birinci dereceden dokusal özellikler için elde edilmiştir.

LDA sınıflandırma doğrulukları Tablo 4’de verilmiştir. LDA için İKE ve eğri altındaki alan (EAA) değerleri Şekil 10’da verilmiştir.

Tablo 4. LDA sınıflandırma doğrulukları

Özellik Kümesi	Doğruluk (%)
Özellik kümesi 1	77.33
Özellik kümesi 2	77.33
Özellik kümesi 1 & Özellik kümesi 2	77.63

**Şekil 10.** LDA için İKE ve eğri altındaki alan (EAA) değerleri

LDA sınıflandırıcısı ile en yüksek sınıflandırma doğruluğuna (% 77.63) ve EAA değerine (0.7214) iki özellik kümesinin birlikte kullanıldığı durumda ulaşılmıştır. Birinci ve ikinci derece dokusal özellikler için eşit sınıflandırma doğruluğu elde edilmiştir (%77.33). Ancak anormal mamogram görüntüsü duyarlılığı ikinci derece dokusal özelliklerde daha yüksek olduğu için EAA değeri (0.72) birinci derece dokusal özelliğe (0.6668) göre daha yüksek olduğu görülmektedir.

Sonuçlar ve Tartışma

Bu çalışmada birinci ve ikinci derece dokusal özelliklerin mamogramların sınıflandırılması üzerindeki etkilerini incelemek için bir bilgisayar destekli teşhis sistemi (BDT) tasarlanmıştır. Tasarlanan sistem ön işleme, özellik çıkarımı ve sınıflandırma olarak üç aşamadan oluşmaktadır. Ön işleme adımı; ortanca süzgeç ile küçük gürültüler süzölmüştür. Küçük gürültülerden temizlenen görüntüde bulunan etiketler ve yüksek yoğunluk değerine sahip pektoral kas morfolojik işlemler ile temizlenmiştir. Normalizasyon işlemi ile tüm mamogram görüntüleri birbirleriyle benzer biçime getirilerek özellik çıkarımı için hazır hale getirilmiştir.

Ön işlem sonucunda alınan görüntülerden birinci derece ve ikinci derece dokusal özellikler

hesaplanmıştır. Mamogramları normal ve anormal olarak sınıflandırmak için k-EYK, DVM ve LDA yöntemleri kullanılmıştır. k-EYK, DVM ve LDA sınıflandırma sonuçları Tablo 5’de verilmiştir.

Tablo 5. k-EYK ve DVM sınıflandırma doğrulukları

Özellik Kümesi	k-EYK Doğruluk (%)	DVM Doğruluk (%)	LDA Doğruluk (%)
Özellik Kümesi 1	80.26	81.58	77.33
Özellik Kümesi 2	82.89	85.53	77.33
Özellik Kümesi 1 & Özellik kümesi 2	82.89	84.21	77.63

Analiz sonuçları değerlendirildiğinde mamogramları sınıflandırmada, DVM yöntemi k-EYK ve LDA yöntemlerine göre daha başarılıdır. Her üç sınıflandırıcı sonuçları dikkate alındığında, ikinci derece dokusal özelliklerin birinci derece dokusal özelliklere göre daha etkili olduğu görülmektedir.

Literatürdeki çalışmalarda genellikle birinci ve ikinci derece dokusal özellikler birlikte kullanılarak özellik kümeleri oluşturulmuştur. Aynı zamanda birçok çalışmada dokusal özelliklere ek olarak farklı özellik grupları (dalgalı tabanlı özellikler, spektral özellikler vb.) birlikte değerlendirilmiştir. Bu nedenle birinci ve ikinci derece dokusal özelliklerin sınıflandırma başarımlarını ayrı ayrı değerlendirme imkanı bulunmamaktadır.

Severoğlu (2016), MIAS veri tabanı için 4 birinci derece ve 4 GSEOM özelliği kullanarak sınıflandırma gerçekleştirdiği çalışmada; birinci derece özellik grubu için DVM ile % 63, GSEOM özellikleri için DVM ile %65 sınıflandırma başarımları elde etmiştir. Özellik gruplarının sınıflandırma başarımları üzerindeki etkileri dikkate alındığında sonuçlar, bu çalışmada elde edilen sonuçları benzerlik göstermektedir. İkinci derece dokusal özelliklerin sınıflandırma başarımlarını artırdığı gözlenmektedir. Bu çalışmada mamogram sınıflandırma işlemi daha yüksek doğrulukla (birinci derece özellik grubu için DVM ile %81.58, ikinci derece özellik grubu için DVM ile %85.53) gerçekleştirilmiştir. Daha yüksek

başarımın elde edilmesinde ön işlem aşamasında; diğer çalışmadan farklı olarak gürültü süzme, pektoral kas temizleme ve normalizasyon işlemlerinin uygulanmasının etkili olduğu sonucu çıkarılmıştır.

Sonuç olarak, dokusal özelliklerin teşhis uygulamaları için kullanışlı ve etkili olduğu tespit edilmiştir. Ancak özellik seçimi veya boyut azaltımı gerçekleştirilecek uygulamalarda ikinci derece dokusal özelliklere öncelik verilmesinin daha uygun olacağı görülmektedir.

Teşekkür

Bu çalışma Amasya Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü tarafından desteklenmiştir (Proje Numarası: FMB-BAP 16-0181).

Kaynaklar

- Abe, S., (2010). *Support Vector Machines for Pattern Classification*, 486, Springer.
- Chen, D. R., Huang, Y. L., ve Lin, S. H., (2011). Computer-aided diagnosis with textural features for breast lesions in sonograms. *Computerized Medical Imaging and Graphics*, **35**,3, 220-226.
- El-Naqa, I., Yang, Y., Wernick, M. N., Galatsanos, N. P., ve Nishikawa, R. M., (2002). A support vector machine approach for detection of microcalcifications. *IEEE Transactions on Medical Imaging*, **21**, 12, 1552-1563.
- Fraz, M. M., Remagnino, P., Hoppe, A., & Barman, S. A. (2013, January). Retinal image analysis aimed at extraction of vascular structure using linear discriminant classifier. International Conference on Computer Medical Applications (ICCM), 1-6, Sousse.
- Fielding, A. (2007). *Cluster and classification techniques for the biosciences*, 259, Cambridge University Press, , New York.
- Ganesan, K., Acharya, U. R., Chua, K. C., Min, L. C., ve Abraham, K. T., (2013). Pectoral muscle segmentation: a review. *Computer Methods and Programs in Biomedicine*, **110**,1, 48-57.
- Gonzalez, R. C. ve Woods, R. E., (Z. Telatar, H. Tora, H. Arı, and A. Kalaycıoğlu), *Sayısal Görüntü İşleme*. Ankara: Palme Yayıncılık, 2014.
- Han, J., Kamber, M. Ve Pei, J., (2006). *Data Mining*. 703, Morgan Kaufmann, USA.
- Haralick, R. M., & Shanmugam, K. ve Dinstein, I., (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-3**, 6, 610-621.
- Junoh, A. K., & Mansor, M. N. (2012). Safety system based on linear discriminant analysis. International Symposium on Instrumentation & Measurement, Sensor Network and Automation (IMSNA), 32-34, Sanya.
- Li, Y., Chen, H., Yang, Y., ve Yang, N., (2013). Pectoral muscle segmentation in mammograms based on homogenous texture and intensity deviation. *Pattern Recognition*, **46**,3, 681-691.
- Mitchell, T. M., (1997). *Machine learning*, 432, McGraw-Hill Education, USA.
- Musta, M., Grgic, M., ve Rangayyan, R. M., (2016). Review of recent advances in segmentation of the breast boundary and the pectoral muscle in mammograms. *Medical & Biological Engineering & Computing*, **54**,7, 1003-1024.
- Pak, F., Kanan, H. R., ve Alikhassi, A., (2015). Breast cancer detection and classification in digital mammography based on Non-Subsampled Contourlet Transform (NSCT) and Super Resolution. *Computer Methods and Programs in Biomedicine*, **122**,2, 89-107.
- Panahi, N., Shayesteh, M. G., Mihandoost, S., & Varghahan, B. Z. (2011). Recognition of different datasets using PCA, LDA, and various classifiers. 5th International Conference on Application of Information and Communication Technologies (AICT), 1-5, Baku.
- Sahiner, B., Chan, H. P., Petrick, N., Wei, D., Helvie, M. A., Adler, D. D., ve Goodsitt, M. M., (1996). Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *IEEE Transactions on Medical Imaging*, **15**,5, 598-610.
- Saltanat, N., Hossain, M. A., ve Alam, M. S., (2010). An efficient pixel value based mapping scheme to delineate pectoral muscle from mammograms. IEEE Fifth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA), 1510-1517, Changsha.
- Severoğlu, N. (2016). Mammogram images classification using Gray Level Co-occurrence Matrices. 24th Signal Processing and Communication Application Conference (SIU), 1781-1784, Zonguldak.

- Sreedevi, S., ve Sherly, E., (2015). A novel approach for removal of pectoral muscles in digital mammogram. *Procedia Computer Science*, **46**, 1724-1731.
- Suckling, J., Parker, J., Dance, D.R., Astley, S., Hutt, I., Boggis, C. R. M., Ricketts, I., Stamatakis, E., Cerneaz, N., Kok, S.L., Taylor, P., Betal, D., ve Savage, J., (1994). The mammographic image analysis society digital mammogram database. *Exerpta Medica. International Congress Series* 1069, 375-378.
- Vapnik, V. N., (1998), *Statistical Learning Theory*, 736, Wiley-Interscience Publication.
- Wei, D., Chan, H.P., Helvie, M.A., Sahiner, B., Petrick, N., Adler, D.D., ve G. M.M., (1995). Classification of mass and normal breast tissue on digital mammograms: multiresolution texture analysis. *Medical Physics*, **22**, 1501-1513.
- Zheng, J. ve Xue, N., (2009) *Advances in Pattern Recognition*, 365, Springer.
- Zheng, Y., Keller, B. M., Ray, S., Wang, Y., Conant, E. F., Gee, J. C., ve Kontos, D., (2015). Parenchymal texture analysis in digital mammography: a fully automated pipeline for breast cancer risk assessment. *Medical physics*, **42**, 7, 4149-4160.

The Effects Of Textural Features On Mammograms Classification

Extended abstract

Textural features are widely used to classify mammographic images. The aim of this paper is to investigate the first order and second order textural features' effects in mammograms classification. For this motivation, firstly computer-aided diagnosis (CAD) system is designed. Computer-aided diagnosis system basically is to purpose enhance and support the diagnose accuracy. The designed system consists of preprocessing, feature extraction and classification steps. In this study, MIAS database is used.

In preprocessing step, firstly noise and label are removed. A (3x3) median filter is used to eliminate small noises. After smoothing images, the radiopaque artifact noises such as labels, numbers, words and flashing lights are removed by morphological operators. Then pectoral muscle removal and image enhancement are implemented. Pectoral muscle is segmented and removed by threshold techniques and black and white morphological operations. Mammographic images are normalized to make them ready for feature extraction.

In feature extraction step, first order and second order textural features are calculated from enhanced mammographic images. 6 first order textural features are mean, standard deviation, skewness, kurtosis, uniformity, and entropy. 4 features extracted for calculated each gray level co-occurrence matrix (GLCM). We averaged these features. These features are contrast, correlation, energy, and homogeneity.

In this paper; three features sets are generated. First feature set contains first order textural features. Second feature set contains second order textural features. Third feature set contains both first order and second order textural features.

In the last step, mammographic images are classified into normal and abnormal by k- nearest neighbor (k-NN), support vector machines (SVM), and Linear Discriminant Analysis (LDA) methods. Classification is performed for all possible combinations of generated feature sets separately.

The classification results are evaluated and compared.

For k-NN classifier the Feature set 2 (second order textural features) yields better results than the Feature set 1 (first order textural features). The highest classification accuracy (82.89%) with the k-NN classifier was achieved when Feature Set 2 and two feature sets were used together. The lowest classification accuracy and the smallest AUC value were obtained for first order textural features.

According to SVM classification results, the highest classification accuracy (85.53%) and AUC value (0.7635) were achieved with Feature Set 2 (second order textural features). The lowest classification accuracy and the smallest AUC value were obtained for first order textural features.

LDA has reached best classification performance (77.63%) and AUC value (0.7214) for both first order and second order textural features. Equal classification accuracy (77.33%) was obtained for first and second order textural features. However second order textural feature set is more sensitive than first order textural feature set for abnormal mammographic image. Second order textural features have higher AUC value than first order textural features.

LDA achieves 77.63%, k-NN achieves 82.89%, and SVM achieves 85.53%. According to classification results, SVM performs better than k-NN and LDA. Experimental results demonstrate that the proposed CAD system obtained best diagnostic performance with second order textural features. Second order textural features outperformed first order textural features to classify mammographic images into normal and abnormal.

Simulation results indicate that textural features are useful for the diagnostic application.

Keywords: Mammography, textural features, computer-aided diagnosis, classification