

Mamografide Şüpheli Kitle Adayı Bölgelerin Belirlenmesi

Burçin KURT^a, Vasif V. NABIYEV^b, Kemal TURHAN^a

^a Biyoistatistik ve Tıp Bilişimi AD, Karadeniz Teknik Üniversitesi, Trabzon

^b Bilgisayar Mühendisliği AD, Karadeniz Teknik Üniversitesi, Trabzon

Determination of Suspicious Mass Candidate Regions on Mammography

Abstract: The segmentation of suspicious regions, which will be used for the detection of related abnormalities in computer-aided diagnosis (CAD) of breast cancer, is an important step. These regions of the image that are meaningful for us are candidate regions which may include mass. In the study, segmentation of suspicious regions which is a process of breast cancer CAD system has been implemented. For this, Havrda&Charvat entropy based segmentation method has been used which is one of non-Shannon entropy measures. Thus, suspicious mass candidate (which may include mass) regions have been segmented from the tissue background in the mammogram. Successful results have been obtained on international MIAS (Mammographic Image Analysis Society) mammogram database used in the study.

Key Words: CAD, breast cancer, suspicious regions, segmentation, mass candidate regions

Özet: Meme kanserinin bilgisayar destekli tanısında, ilgili anormalliklerin tespiti için kullanılacak olan şüpheli bölgelerin bölütlenmesi oldukça önemli bir aşamadır. Bu bölgeler, görüntünün bizim için anlamlı olan yani kitle içerebilecek olan aday bölgeleridir. Çalışmada bilgisayar destekli meme kanseri tanı sisteminin bir aşaması olan şüpheli bölgelerin bölütlenmesi gerçekleştirilmiştir. Bunun için non-Shannon entropi ölçülerinden Havrda&Charvat entropi tabanlı bölütleme yöntemi kullanılmıştır. Böylece, mamografi görüntüsündeki şüpheli kitle adayı(kitle içerebilecek) bölgeler doku arka planından ayrılmıştır. Çalışmada kullanılan uluslararası MIAS (Mammographic Image Analysis Society) mamografi veritabanı üzerinde başarılı sonuçlar elde edilmiştir.

Anahtar Kelimeler: CAD, meme kanseri, şüpheli bölgeler, bölütleme, kitle aday bölgeler

1. Giriş

Meme kanseri, akciğer kanserinden sonra dünyada görülme sıklığı en yüksek olan kanser türüdür. Her 8 kadından birinin hayatının belirli bir zamanında meme kanserine yakalanacağı bildirilmektedir. Meme kanseri, yayılmadan önce, erken tespit edilirse, hasta %96 yaşam şansına sahiptir. Meme kanserine karşı en iyi koruyucu yöntem erken

teşhistir [1].

CAD (Bilgisayar destekli tanı) sistemleri, tıbbi görüntülerin yorumlanmasında doktorlara yardımcı olarak kullanılan sistemlerdir. Dijital radyolojik görüntülerde(mamografi, MRI, ultrason vb.), matematiksel ve istatistiksel analiz yöntemlerini kullanarak radyolog için aşikar olmayan alanları(hedef tanıları veya patolojileri) belirler. CAD, ikincil ve yardımcı ek bir tanı sistemidir, son karar radyoloğa aittir.

Son zamanlarda, CAD sistemleri Amerika Birleşik Devletlerinde tarama merkezleri ve hastanelerde mamogramlar üzerinden meme kanseri tespiti için yapılan rutin klinik çalışmalarının bir parçası haline gelmiştir. Bu da CAD sistemlerinin farklı görüntüleme yöntemlerinden elde edilmiş medikal görüntülerdeki birçok farklı anormalliklerin tespit ve teşhisinde yaygın olarak uygulanmaya başlanacağını göstermektedir[2]. CAD ile radyoloğun yeteneği ve bilgisayarın gücü birleştirilerek kazanç sağlanır. Böylece, meme kanseri erken tespit oranı artırılabilir.

Mamografide kitle aday şüpheli bölgelerin belirlenmesi, geliştirilmesi hedeflenen meme kanseri CAD sisteminin bir aşamasıdır. Önceki çalışmalarımızda, mamografi görüntülerinin iyileştirilmesi, meme bölgesinin bölütlenmesi ve meme sınırının belirlenmesi aşamalarını gerçekleştirmiştik. Yapılan çalışmada da, şüpheli bölgelerin bölütlenmesi ele alınmıştır.

Mamogram görüntüleri üzerinden kitle tespiti ile ilgili W.B. Sampaio [3] ve arkadaşları tarafından yapılan çalışmada şüpheli bölgelerin bölütlenmesi için hücreler sinir ağları kullanılmıştır. Burada, Textudil ve Blur şablonlarından yararlanılmıştır. Bölütleme algoritmasınının 623 kitleden 526'sını bölütlediğinden ve başarısının %90.85 olduğundan bahsedilmektedir. Fatima[4] ve arkadaşları tarafından yapılan çalışmada ise kitle tespitinde ilgili alanların çıkarılmasında maksima eşikleme kullanılmıştır. Ancak, burada maksima eşiklemede tüm veritabanı incelenerek ortalama sabit bir eşik değeri seçilerek işlemler gerçekleştirilmiştir.

Otomatik kitle belirlenmesi için Hussein[5] ve arkadaşları tarafından yapılan çalışmada, morfolojik erozyon ve yeniden yapılanma yöntemlerinin hibridlenmesiyle en iyi 10 kitle aday bölge belirlenmiştir. Ardından, alan minimizasyonuna dayalı bir bölütleme algoritması ile kitle aday bölgeler çıkarılmıştır. Bir diğer çalışmada, Ravi[9] tarafından, meme kanseri bilgisayar destekli tanı için etkili şüpheli bölge bölütlenmesi algoritması ile ilgili yapılan tez çalışmasında, şüpheli bölgelerin bölütlenmesi için sağlam bulanık c-ortalama(robust fuzzy c-means-FCM) ve uzaysal bulanık c-ortalama (spatial fuzzy c-means-SFCM) kümeleme algoritmaları kullanılmıştır.

Burada, şüpheli kitle aday bölgelerin çıkarılmasında etkili istatistiksel yöntemlerden olan Shannon ve non-Shannon entropi ölçüleri [6] incelenmiş ve Havrda&Charvat entropi tabanlı bölütleme gerçekleştirilmiştir. Çalışmada, 322 dijital mamogram içeren Mamografik Görüntü Analiz Derneği (MIAS) veritabanı kullanılmıştır[7].

2. Gereç ve Yöntem

Entropi, bir bilgi teorisi kavramıdır ve bilgi miktarını ölçmek için kullanılır[8]. Entropi, bilgi kaynağının olasılıklı davranışı olarak da tanımlanmaktadır. Bu tanıma göre, rastgele bir A olayının P(A) olasılığı ile oluşması için aşağıdaki bilgi birimlerini içermesi gerektiğinden bahsedilmektedir[8].

$$I(A) = \log[1/P(A)] = -\log[P(A)] \quad (1)$$

$I(A)$ miktarı, A olayının öz bilgisi olarak adlandırılmaktadır. Olayın öz bilgi miktarı, olasılığı ile ters orantılıdır. Bilgi teorisindeki entropinin temel içeriği bir sinyal ya da rastgele bir olaydaki rastgeleliğin ne kadar olduğuyla ilgilidir. Buna bakmanın alternatif yolu, sinyal ile ne kadar bilgi taşındığı hakkında konuşularak yapılabilir. Entropi, rastgelelik ölçütüdür.

$p_1, p_2, \dots, p_i, \dots, p_k$ ayrık kaynağın olasılık dağılımı olsun. Bu nedenle $0 \leq p_i \leq 1$ ve $\sum_{i=1}^k p_i = 1$, burada k durumların toplam sayısıdır. Bir ayrık kaynağın entropisi genellikle olasılık dağılımından elde edilir. Çeşitli entropi ölçütleri[6] tanımlanabilir.

Shannon Entropisi

Mamografi; meme arka planı, doku arka planı ve şüpheli bölge olmak üzere üç farklı nesnenin birleşimidir. Meme arka planı teşhis için bir bilgi sağlamaz. Bu nedenle, mamogram analiz çalışmalarında göz ardı edilebilir. Yaklaşık olarak, bir mamogramın üçte birinden fazlası meme arka planıdır. Bu da, meme dokularının ortalama gri seviye değerini etkileyecektir[8]. Bu nedenle, meme dokularının ortalama gri seviye değeri, meme arka plan pikselleri çıkarılarak aşağıdaki gibi tanımlanabilir[9].

$$k = \frac{1}{n} \sum_{i,j \in R_i} \sum_{j \in R} f(i, j) \quad (2)$$

Burada M, N mamogramın boyutları; R , gri seviye değerleri >100 olan bölge; n , bu bölgedeki piksellerin sayısı; $f(i, j)$ ise, i ve j koordinatlarındaki gri seviye değeridir.

$p_1, p_2, \dots, p_i, \dots, p_{k-1}, p_k, p_{k+1}, \dots, p_L$ mamogramdaki gri seviye değerlerinin olasılık dağılımı olsun. Burada, p_k mamogramın normalize histogramıdır.

$$p_k = h_k / (M \times N)$$

ve h_k , mamogramın gri seviye histogramıdır. Bu dağılımdan, üç tane olasılık dağılımı çıkarabiliriz; birincisi meme arka planı için, ikincisi doku arka planı için ve üçüncüsü de şüpheli bölge için. Bu olasılık dağılımları aşağıdaki gibi ifade edilebilir:

Meme arka planına ait olasılık dağılımı:

$$\frac{p_1}{p_{k-1}}, \frac{p_2}{p_{k-1}}, \dots, \frac{p_{k-1}}{p_{k-1}} \quad (3)$$

Doku arka planına ait olasılık dağılımı:

$$\frac{p_k}{p_t - p_{k-1}}, \frac{p_{k+1}}{p_t - p_{k-1}}, \dots, \frac{p_t}{p_t - p_{k-1}} \quad (4)$$

Şüpheli bölgeye ait olasılık dağılımı:

$$\frac{p_{t+1}}{1-p_t}, \frac{p_{t+2}}{1-p_t}, \dots, \frac{p_L}{1-p_t} \quad (5)$$

Burada, $p_{k-1} = \sum_{i=1}^{k-1} p_i$, $p_t = \sum_{j=1}^t p_j$, t eşik değerini ve L mamogramdaki maksimum gri seviye değerini göstermektedir.

Shannon entropisi eşikleme yaklaşımına dayanmaktadır. Shannon entropisine göre doku arka planı piksellerinin entropisi ve şüpheli bölge piksellerinin entropisi aşağıdaki şekilde ifade edilebilir:

Doku arka planı pikselleri için Shannon entropisi:

$$S_{TB}(t) = - \sum_{i=k}^t \frac{p_i}{p_t - p_{k-1}} \log_e \left(\frac{p_i}{p_t - p_{k-1}} \right) \quad (6)$$

Şüpheli bölge pikselleri için Shannon entropisi:

$$S_{SR}(t) = - \sum_{i=t+1}^L \frac{p_i}{1-p_t} \log_e \left(\frac{p_i}{1-p_t} \right) \quad (7)$$

Shannon entropisi, eşikleme yaklaşımına dayalı, $S_{TB}(t) + S_{SR}(t)$ değerini maksimize ederek optimum eşik değerini t_{opt}^s , elde eder[6].

$$t_{opt}^s = \text{Arg}_{t=k}^L \max [S_{TB}(t) + S_{SR}(t)] \quad (8)$$

Bu optimum eşik değeri şüpheli bölgeyi, doku arka planından ayırmaktadır.

Non-Shannon Entropi Ölçütleri

a. Renyi Entropisi

Benzer şekilde Renyi entropisi de eşikleme yaklaşımına dayanmaktadır ve Renyi entropisine göre doku arka planı pikselleri ile şüpheli bölge piksellerine ait entropiler aşağıdaki gibi ifade edilmektedir:

Doku arka planı pikselleri için Renyi entropisi:

$$R_{TB}(t) = \frac{1}{1-\alpha} \log_e \left[\sum_{i=k}^t \left(\frac{p_i}{p_t - p_{k-1}} \right)^\alpha \right] \quad (9)$$

Şüpheli bölge pikselleri için Renyi entropisi:

$$R_{SR}(t) = \frac{1}{1-\alpha} \log_e \left[\sum_{i=t+1}^L \left(\frac{p_i}{1-p_t} \right)^\alpha \right] \quad (10)$$

Burada $\alpha \neq 1$, $\alpha > 0$

Renyi entropisi, eşikleme yaklaşımına dayalı, $R_{TB}(t) + R_{SR}(t)$ değerini maksimize ederek optimum eşik değerini t_{opt}^R , elde eder[6].

$$t_{opt}^R = \text{Arg}_{t=k}^L \max [R_{TB}(t) + R_{SR}(t)] \quad (11)$$

b. Havrda&Charvat Entropisi

Renyi entropisinden sonra Havrda ve Charvat entropisi de eşikleme tekniğine dayanmaktadır. Buna göre, doku arka planı pikselleri ve şüpheli bölge pikselleri için Havrda ve Charvat entropisi de aşağıdaki şekilde tanımlanmaktadır:

Doku arka planı pikselleri için Havrda ve Charvat entropisi:

$$HC_{TB}(t) = \frac{1}{1-\alpha} \left[\sum_{i=k}^t \left(\frac{p_i}{p_i - p_{k-1}} \right)^\alpha - 1 \right] \quad (12)$$

Şüpheli bölge pikselleri için Renyi entropisi:

$$HC_{SR}(t) = \frac{1}{1-\alpha} \left[\sum_{i=t+1}^L \left(\frac{p_i}{1-p_t} \right)^\alpha - 1 \right] \quad (13)$$

Havrda ve Charvat entropisi, yaygın olmayan bir özelliğe sahiptir. Böylece eşikleme yaklaşımı tabanlı Havrda ve Charvat entropisi,

$HC_{TB}(t) + HC_{SR}(t) + (1-\alpha)HC_{TB}(t)HC_{SR}(t)$ değerini maksimize ederek optimum eşik değerini t_{opt}^{HC} , elde eder[6].

$$t_{opt}^{HC} = \text{Arg}_{t=k}^L \max [HC_{TB}(t) + HC_{SR}(t) + (1-\alpha)HC_{TB}(t)HC_{SR}(t)] \quad (14)$$

c. Kapur Entropisi

Son olarak da, yine eşikleme yaklaşımına dayalı Kapur entropisinden bahsedeceğiz. Kapur entropisinin tanımına göre doku arka planı ve şüpheli bölge piksellerinin olasılık dağılımlarına ait entropiler aşağıdaki gibidir:

Doku arka planı pikselleri için olasılık dağılımına ait Kapur entropisi:

$$K_{TB}(t) = \frac{1}{\beta - \alpha} \log_e \left[\frac{\sum_{i=k}^t \left(\frac{p_i}{p_i - p_{k-1}} \right)^\alpha}{\sum_{i=k}^t \left(\frac{p_i}{p_i - p_{k-1}} \right)^\beta} \right] \quad (15)$$

Şüpheli bölge pikselleri için Kapur entropisi:

$$K_{SR}(t) = \frac{1}{\beta - \alpha} \log_e \left[\frac{\sum_{i=t+1}^L \left(\frac{p_i}{p_i - p_{k-1}} \right)^\alpha}{\sum_{i=t+1}^L \left(\frac{p_i}{p_i - p_{k-1}} \right)^\beta} \right] \quad (16)$$

Burada $\alpha \neq \beta$, $\alpha > 0$, $\beta > 0$

Kapur entropisi, eşikleme yaklaşımına dayalı, $K_{TB}(t) + K_{SR}(t)$ değerini maksimize ederek optimum eşik değerini t_{opt}^K , elde eder[6].

$$t_{opt}^R = \text{Arg}_{t=k}^L \max [K_{TB}(t) + K_{SR}(t)] \quad (17)$$

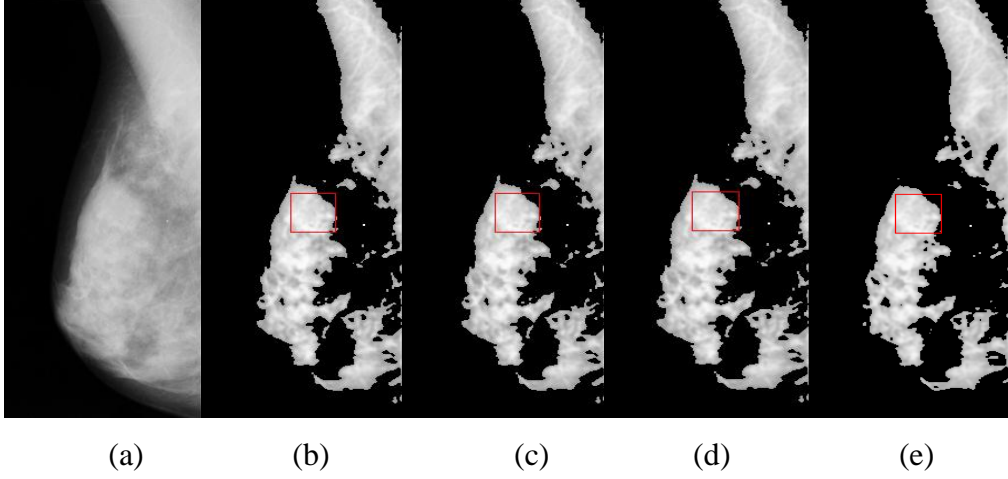
Optimum eşik değeri, t_{opt}^* , $* \in \{S, R, HC, K\}$, şüpheli bölgeyi doku arka planından ayırır.

DeneySEL sonuçlara bakıldığında, non-Shannon (Renyi, Havrda&Charvat, Kapur) entropi ölçülerinin, mamogram görüntülerinde şüpheli bölgeyi doku arka planından ayırmada daha tatmin edici sonuçlar verdiğini ortaya koymaktadır[6]. Non-Shannon entropi ölçülerinin, Shannon entropi ölçüsüne bir başka üstünlüğü de değerleri ayarlama da kullanılabilir α , β gibi parametreler içermesidir. Bu parametreler, aynı sınıf mamogram görüntüleri için görüntü işleme zincirinde ayar parametreleri olarak önemli bir rol oynayabilirler.

Ayrıca, Havrda&Charvat entropisinin non-Shannon ölçüleri arasında yaygın olmayan bir özelliğe sahip olduğu görülmektedir. Yani, Havrda&Charvat entropi tabanlı bölütleme yaklaşımı, mamogram görüntülerinde eklemeli olmayan içeriğin varlığını yakalama yeteneğine sahiptir[6]. Bu nedenlerle çalışmada, non-Shannon entropi ölçülerinden Havrda&Charvat entropi tabanlı bölütleme gerçekleştirilmiştir. Böylece, mamogram görüntüsündeki şüpheli kitle aday bölgeler bölütlenmiştir.

3. Bulgular

HC_{TB} , HC_{SR} ve t_{opt} değerlerinin hesaplanmasında kullanılan α değeri çalışmada[6] en iyi sonucu 0,2 alındığında vermektedir. Biz, çalışmamızda farklı α değerlerini denedik ve aşağıdaki sonuçları elde ettik. (Kırmızı kare çerçeve, o görüntüye ait anormalliğin yerini göstermektedir, programın tespiti değildir, merkez x,y koordinatları verilerek bölütleme sonucu şüpheli bölgelerde kapsadığını gösterme amaçlı çizilmiştir.)



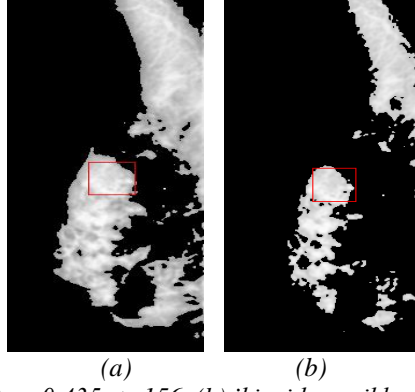
Şekil 1. (a) Orijinal görüntü - MIAS mdb017, (b) $\alpha=0,2$; $t=156$, (c) $\alpha=0,4$; $t=156$, (d) $\alpha=0,6$; $t=156$, (e) $\alpha=0,8$; $t=168$

Şekil 1.'de görüldüğü gibi mdb017 mamogram görüntüsü için $\alpha=0,2$; 0,4 ve 0,6 değerleri için optimum eşik değeri aynı çıkarken $\alpha=0,8$ değeri için farklı bir optimum eşik değeri elde edilmektedir. Ancak, farklı görüntüler için farklı sonuçlar elde edilebilmektedir. Buna göre, biz çalışmamızda MIAS veritabanı üzerinde yaptığımız deneysel sonuçları inceleyerek α değerini sabit almaktansa her görüntü için dinamik bir biçimde değişecek şekilde aşağıdaki formülü geliştirerek hesapladık.

$$\alpha = 1 - \left(\frac{k}{255} \right) \quad (18)$$

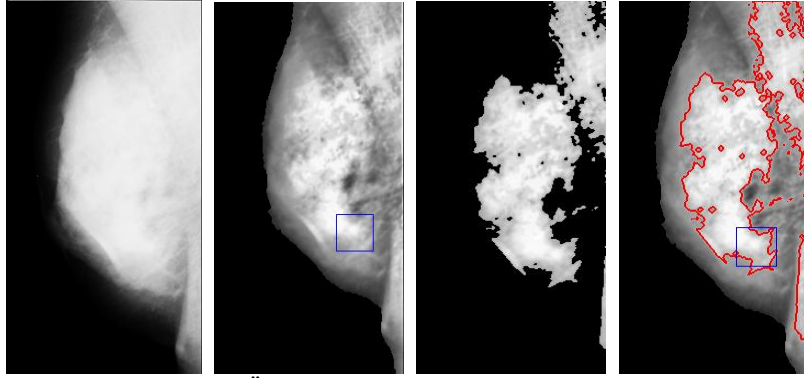
Her görüntü için dinamik olarak hesaplanan k (*formül (2)*) değerinin en yüksek gri seviye değerine bölerek normalize ettik ve bunu maksimum değer olan 1'den çıkararak α değerini hesapladık. Buna göre, elde ettiğimiz bölütleme sonucu Şekil 2'deki gibidir.

Daha sonra şüpheli bölgeleri biraz daha filtrelemek için Havrda&Charvat entropi tabanlı bölütleme sonucu elde edilen şüpheli bölgelerin ortalama gri seviye değeri alınarak ikinci bir eşikleme işlemi uygulanmıştır.

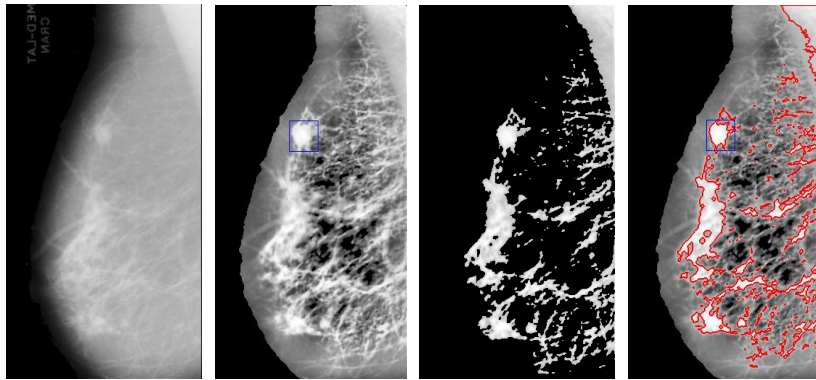


Şekil 2. (a) $\alpha=0,435$; $t=156$, (b) ikinci kez eşiklenmiş görüntü

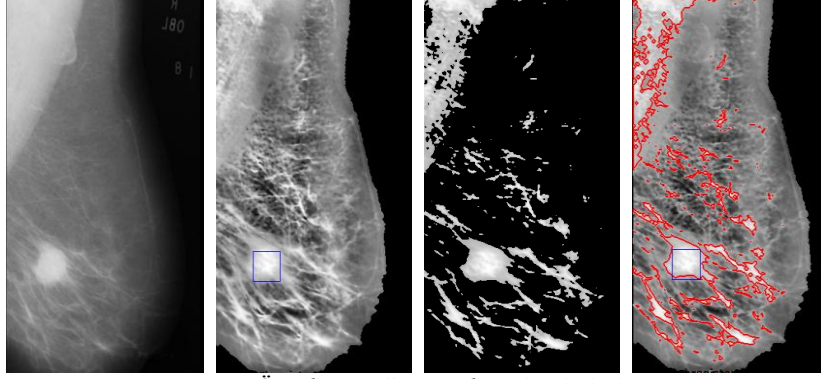
Şüpheli kitle adayı bölgelerinin bölütlenmesine ilişkin MIAS mamografi görüntü veritabanına ait bazı örneklerin sonuçları aşağıda gösterilmiştir.



Örnek 1: mdb179 – kötü huylu kitle



Örnek 2: mdb023 – kötü huylu kitle



Örnek 3: mdb028 – kötü huylu kitle

Şekil 3. (a) Orijinal görüntü, (b) iyileştirilmiş ve meme bölgesi bölütlenmiş görüntü (c) şüpheli bölgeler bölütlenmiş görüntü (d) şüpheli bölgeler kırmızı ile çizilmiş görüntü

Şekil 3'te görüldüğü gibi şüpheli kitle aday bölgeler çıkarılmıştır ve veritabanında koordinatları verilen kitle bölgeleri de elde edilen şüpheli bölgeler arasındadır. Bununla birlikte koltuk altı bölgesi de yüksek gri seviye değerinden dolayı şüpheli bölgeler arasındadır. Bir sonraki aşamada, elde edilen şüpheli bölgeler arasından kitle bölgesi tespit edilecektir.

4. Tartışma

Önerilen yöntemde kullanılan α parametresi 322 mamografi görüntüsü taranarak, hiçbir kitle bölgesini kaçırmayacak şekilde formül (18) geliştirilerek hesaplanmıştır. Koltuk altı bölgesi de bir radyolog uzmana sorularak kitle içerebileceğinden bırakılmıştır.

5. Sonuç

Sonuç olarak, şüpheli kitle aday bölgelerinin çıkarılması için önerilen algoritma başarılı sonuçlar vermektedir. Ayrıca, algoritmadaki α parametresi geliştirilen formül (18) ile her görüntü için dinamik hale getirilmiştir. Bu da, yöntem katılan oldukça önemli bir yeniliktir. Bununla birlikte, geliştirilen sistemin en önemli avantajı hiçbir kitle bölgesini kaçırmamasıdır.

Şüpheli kitle aday bölgelerin çıkarılması, otomatik kitle tespiti için ön bir filtrelemedir ve böylece gereksiz yerlerin taranması engellenmiştir. Bu da kitle tespit başarısını arttıracaktır.

Yapılan çalışma, yürütülen SANTEZ projemizin bir aşamasıdır, bundan sonraki çalışmamızda elde edilen şüpheli bölgeler üzerinden, otomatik olarak kitle bölgesinin tespiti hedeflenmektedir.

6. Teşekkür

Yapılan çalışma, T.C. Bilim, Sanayi ve Teknoloji Bakanlığı ile AKGÜN Bilgisayar Programları ve Hizmetler Sanayi Tic. Ltd. Şti. tarafından desteklenen ve geliştirilmekte

olan “DICOM Uyumlu Meme Kanseri Karar Destek Sistemi” isimli SANTEZ projemizin bir aşamasıdır.

7. Kaynakça

- [1] Meme kanseri, http://tr.wikipedia.org/wiki/Meme_kanseri#cite_note-1 Last accessed: 12th September 2012.
- [2] Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, vol. 31, 2007; p. 198–211.
- [3] Wener BS, Edgar MD, Aristófanés CS, Anselmo CP, and Gattas M. Detection of masses in mammogram images using CNN, geostatistic functions and SVM. *Computers in Biology and Medicine*. vol. 41 (8), 2011; p. 653–664.
- [4] Eddaoudi F, Regragui F, Mahmoudi A, and Lamouri N. Masses detection using svm classifier based on textures analysis. *Applied Mathematical Sciences*. vol. 5, no. 8, 2011; p. 367 -379.
- [5] Samma H, Lim CP, and Samma A. A Computer-aided detection system for automatic mammography mass identification. *ICONIP 2010, Part II, LNCS 6444, Springer-Verlag Berlin Heidelberg*. 2010; p. 226–233.
- [6] Pharwaha APS, and Khehra BS. Shannon and Non-Shannon Measures of Entropy for Statistical Texture Feature Extraction in Digitized Mammograms. *Proceedings of the World Congress on Engineering and Computer Science, WCECS 2009*, vol. II. San Francisco, USA, 2009.
- [7] University of Essec, Mamographic Image Analysis Society. <http://peipa.essex.ac.uk/ipa/pix/mias/> Last accessed: 27th August 2012.
- [8] Khehra BS, and Pharwaha APS. Digital mammogram segmentation using non-shannon measures of entropy. *Proceedings of the World Congress on Engineering, WCE 2011*. vol. II. London, U.K., 2011.
- [9] Mohanalini J, Kalra PK, and Kumar N. Fast communication: An automatic method to enhance microcalcifications using Normalized Tsallis entropy. *Journal Signal Processing*. vol.90, 2010; p. 952-958.

8. Sorumlu Yazarın Adresi

Burçin KURT, Karadeniz Teknik Üniversitesi, Tıp Fak., Tıp Bilişimi AD. burcinkurt@gmail.com.