

Pap Smear Test Görüntülerinde Hücre Çekirdeklerinin Bölütlenmesi

Cell Nuclei Segmentation in Pap Smear Test Images

Aslı Kale, Selim Aksoy

Sevgen Önder

Bilgisayar Mühendisliği Bölümü,
Bilkent Üniversitesi,
Bilkent, 06800, Ankara

{akale, saksoy}@cs.bilkent.edu.tr

Patoloji Bölümü,
Hacettepe Üniversitesi,
Sıhhiye, 06100, Ankara

sonder@hacettepe.edu.tr

Özetçe

Erken tanı ile önlenebilir bir kanser olan serviks kanserinin hücrelerde oluşturduğu displazi pap smear testi ile taranabilmektedir. Testin yaygınlaşması ve güvenilirliğinin artması için bilgisayar destekli tanılayıcı bir sistemin geliştirilmesi yardımcı olabilir. Hücrelerdeki displazinin otomatik olarak belirlenmesinde çekirdek ve sitoplazma bölütlenmesi en önemli aşamalardandır. Bu çalışmada, serviks hücre görüntülerindeki çekirdeklerin bölütlenmesi amaçlanmaktadır. Öncelikle, hücre çekirdekleri üzerinde işaretler matematiksel biçimcilik yöntemleri kullanılarak hesaplanmıştır. Bulunan işaretler temel alınarak işaret esaslı havza bölütlenmesi ve balon snake modeli çekirdeklerin çevrilerini elde etmek üzere serviks hücresi görüntülerinden oluşan veri kümesi üzerinde uygulanmıştır. Veri kümesi displazi derecesine göre ayrılmış 6 sınıftan oluşmaktadır. Elde edilen sonuçlar bağıl uzaklık hatası kullanılarak karşılaştırılmış ve yöntemlerin başarımları ve eksiklikleri tartışılmıştır.

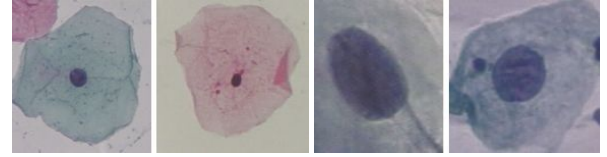
Abstract

Cervical cancer is a preventable disease and the dysplasia it causes can be scanned by using a pap smear test. It can be beneficial to develop a computer-assisted diagnosis system to make the pap smear test robust and widespread. The most fundamental part of such a system is the segmentation of nuclei and cytoplasm in cervical cell images. The aim of this study is to segment the nuclei in such images. First, markers on the nuclei are found by using mathematical morphology operations. Based on the obtained markers, marker-based watershed segmentation and balloon snake model are applied to find the nuclei contours in a data set consisting of cervical cell images. The data set is composed of six classes ranging according to the dysplasia degree of the cells. The results are evaluated according to the relative distance error measure, and the strengths and weakness of the methods are discussed.

1. Giriş

Dünya genelinde her yıl 500.000 kadına serviks kanseri tanısı konulmakta ve 250.000 kadın bu kanser nedeniyle ölmektedir.

Bu çalışma TÜBİTAK KARIYER 104E074 numaralı proje tarafından desteklenmiştir.



(a)

(b)

Şekil 1: Hematoksilen-Eozin ile boyanmış skuamöz hücre örnekleri. (a) Normal hücreler. (b) Olağandışı hücreler.

Histopatolojik bir inceleme olan pap smear testi ile serviks kanserinin öncü değişiklikleri taranarak belirlenebilir. Serviks kanseri erken tanı ile önlenebilir bir kanser olduğu için serviks tarama çalışmalarının düzenli olarak uygulandığı ülkelerde invaziv serviks kanseri büyük oranda azalmıştır [1].

Serviks hücrelerinde meydana gelen kanser öncüsü değişiklikler displazi olarak adlandırılmaktadır. Hücre çekirdeğinin ve sitoplazmasının büyüklüğü, dokusu, şekli ve birbirlerine göre alanlarının oranı displazi ve kanser tanısı konulurken etkili olan faktörlerdendir. Şekil 1'de Hematoksilen-Eozin ile boyanmış normal ve olağandışı skuamöz hücre örnekleri gösterilmiştir.

Pap smear testi insan gözüyle uygulanan bir yöntem olduğu için gözlemci tutarsızlığı ve her bir test için sarf edilmesi gereken çaba gibi dezavantajlar içermektedir. Testin ülke genelinde düzenli olarak uygulanması durumunda eldeki yüklü miktarda verinin zamanında işlenmesi problemi ortaya çıkacaktır. Testin makine destekli hale getirilmesi maliyeti azaltırken başarılı bir algoritma test sonuçlarının güvenilirliğini arttıracaktır. Böylelikle test ülke genelinde düzenli olarak uygulanabilir hale gelecektir. Bilgisayar destekli tanılayıcı bir sistemin geliştirilebilmesi için hücre çekirdeğinin ve sitoplazma çevritinin belirlenmesi en önemli aşamalardandır.

Wu ve arkadaşları [2] hücrelerin bölütlenmesi için en iyi parametrik bir yöntem geliştirmiştir ancak bu yöntem kullanılırken hücrelerin şeklini, büyüklüğünü ve çevresine göre yeğinliğini önceden bilmek önemlidir. Yang-Mao ve arkadaşları [3] sitoplazma ve çekirdek kenarlarını pekiştirerek bölütleme yapan yöntemlerini serviks hücre görüntüleri üzerinde uygulamış ve elde edilen sonuçları karşılaştırmak üzere bağıl uzaklık hatası (relative distance error) ölçüsünü tanımlamışlardır. Norup [5] sitoplazma ve çekirdek bölütlenmesi için ticari CHAMP sayısal görüntü yazılımını

kullanmıştır. Ancak, bu yöntem ile elde edilen sonuçlar olağandışı hücreler için tatmin edici değildir [3].

Yapılan çalışma, pap smear test görüntülerinin otomatik analizinin ilk aşaması olarak hücre çekirdeklerini bölütleme problemi üzerinde yoğunlaşmaktadır. Çekirdekler üzerinde elde edilen önplan işaretleri temel alınarak işaret esaslı biçimbilimsel havza (watershed) bölütleme algoritması ve balon dış kuvvetini kullanan aktif çevrit modeli (balloon snake) [6] çekirdek bölütlemesi için kullanılmıştır. İki yöntem serviks hücre görüntülerinden oluşan veri kümesi üzerinde uygulanmış ve elde edilen bölütleme sonuçları bağıl uzaklık hatası kullanılarak karşılaştırılmıştır. Farklı bölütleme sonuçlarını içeren örneklerle her iki yöntemin zayıf özellikleri tartışılmıştır. Bildiri, çekirdek bölütlemesi için yapılan çalışmaya gelecekte eklenebilecek hususların tartışılması ile sona ermektedir.

2. Hücre Çekirdeklerinin Bölütlenmesi

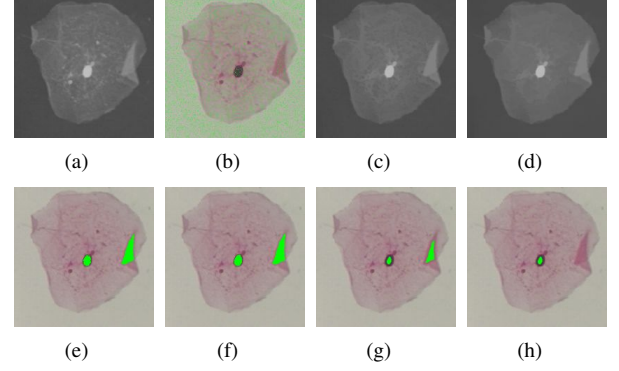
Hücre çekirdeklerini bölütlemeyi amaçlayan çalışmada çekirdekler üzerinde elde edilen işaretler kullanılarak işaret esaslı havza bölütlemesi ve aktif çevrit modeli uygulanmıştır.

Gri tonlu pap smear test görüntülerinde hücre çekirdeklerinin yeğniliğinin sitoplazma ve arka plana göre daha düşük olduğu gözlemlenmektedir. Bölütlenmek istenen çekirdek alanlarının yeğniliğini diğer görüntü alanlarına göre daha yüksek hale getirmek amacı ile gri tonlu görüntülerin tümleyeni alınmış ve bölütlemedeki bütün işlemler elde edilen bu görüntüler üzerinde uygulanmıştır.

2.1. Önplan İşaretlerinin Bulunması

Hücre çekirdeklerinin bölütlenmesinde ilk aşama her bir çekirdeğe ait işaretlerin bulunması olarak belirlenmiştir. Elipitik bir şekle sahip olan hücre çekirdekleri üzerinde önplan alanları bulunurken şekil ve büyüklük bilgisinin kullanımına olanak veren matematiksel biçimbilim yöntemleri kullanılmıştır. Geri çatılma ile açma (opening by reconstruction) ve geri çatılma ile kapama (closing by reconstruction) işlemleri disk şeklinde yapısal öğeler kullanılarak sırasıyla uygulanmıştır. Bu işlemleri uygulamakla ön ve arka planda yer alan en büyük gri tonları düzleştirerek her bir çekirdek üzerinde tek sayıda düz bir ton elde etmek amaçlanmıştır. Kullanılan disk yapısal öğelerinin yarıçapları seçilirken en küçük hücre çekirdeğinin yarıçapını geçmemesine dikkat edilmiştir. Şekil 2'de verilen bir görüntüye ait önplan işaretlerinin elde edilmesi sırasında uygulanan işlemlerin sonuçları gösterilmiştir. Geri çatılma ile açma ve kapama işlemlerinin görüntüdeki en büyük değerlere olan etkisi Şekil 2(b) ve 2(e)'de yeşil alanlar ile gösterilen bölgelere bakılarak gözlemlenebilir.

Biçimbilimsel işlemlerle değiştirilmiş görüntüde yer alan en büyük değerler önplan işaretleri olarak belirlenmiştir. Elde edilen işaretler çekirdek kenarlarına değebileceği için işaretlere Şekil 2(f) ve 2(g)'deki gibi kapama (closing) ve kemirme (erosion) işlemleri uygulanarak alanları daraltılmıştır. Çekirdekler üzerinde bulunanlara göre çok büyük ve çok küçük alanlı işaretleri süzgeçlemek amacıyla işaretlerin alanlarına alt ve üst sınır getirilmiştir. Şekil 2(h) çekirdek dışında yer alan geniş alanlı işaretin süzgeçlenmesinin ardından elde edilen önplan işaretlerini göstermektedir.



Şekil 2: Önplan işaretlerinin bulunması. (a) Gri tonlu görüntü. (b) Görüntü üzerindeki en büyük değerler. (c) Geri çatılma ile açma. (d) Geri çatılma ile kapama. (e) (d) üzerindeki en büyük değerler. (f) Kapama işlemi. (g) Kemirme işlemi. (h) Süzgeçleme işlemi.

2.2. İşaret Esaslı Havza Bölütleme Yöntemi

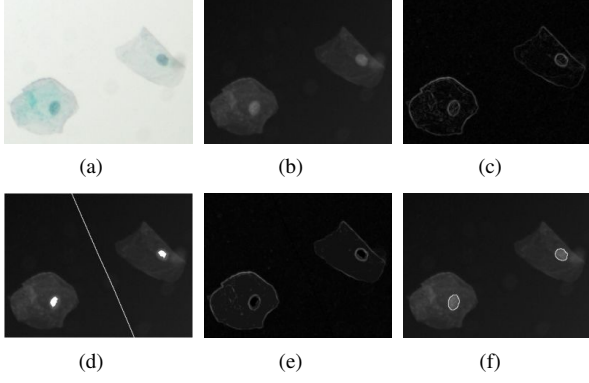
Havza, etkili bir biçimbilimsel bölütleme yöntemi olarak kullanılmaktadır. Topoğrafya alanından gelmekte olan su ayırımı hattı anlamındaki havza kavramı, gri tonlu bir görüntüyü yükseltelerin piksel değerleri olarak belirlendiği topoğrafik bir yüzey olarak değerlendirilmektedir. Böylece yüksek gradyanlı görüntü kenarları su ayırımı hattını, düşük gradyanlı alanlar ise su havzalarını (catchment basins) temsil etmektedir. Üzerlerine gelen suyu aynı en küçük değere boşaltan pikseller su havzalarını meydana getirirken bu havzaları ayıran sınırdaki piksellerin birleşmesiyle su ayırımı hattı oluşmaktadır.

Havza, oluşturduğu çevritlerin kapalı, bitişik ve görüntü kenarları üzerinde yer almasıyla avantajlı bir bölütleme yöntemi olmasına rağmen genellikle aşırı bölütleme (over-segmentation) problemini beraberinde getirmektedir. Oluşan çevritler bütün önemli nesne kenarlarını içerdiklerinden yapılması gereken işlem istenmeyen kenarların süzgeçlenmesi olarak belirlenebilir. İşaret esaslı biçimbilimsel havza bölütleme (marker-based watershed segmentation) istenmeyen kenarların süzgeçlenmesine yönelik geliştirilmiş yöntemlerden biridir.

İşaret esaslı havza bölütlemesi, görüntü gradyanını işaretlerle belirlenen yerlerde en küçük değere sahip olacak şekilde biçimbilimsel işlemler uygulayarak değiştirdikten sonra havza bölütlemesi yapmaktadır. Yöntemde işaretleri bölütlenmek istenen nesnelere hesaplamak gerekmektedir. Yapılan çalışmada çekirdekler üzerinde bulunan önplan işaretleri ve çekirdekler dışındaki alanda hesaplanan arka plan işareti kullanılarak işaret esaslı havza bölütlemesi serviks hücre görüntülerinde uygulanmıştır. Arka plan işareti görüntüyü çevreleyen sınır pikselleriyle önplan işaretlerinin uzaklık dönüşümünün (distance transform) havza bölütlemesinden elde edilen havza çizgisinin birleşiminden oluşmaktadır. Böylece görüntü Şekil 3'te verilen örnekte olduğu gibi çekirdekler ve çekirdekler dışındaki alan olmak üzere iki çeşit alana bölütlenebilmektedir.

2.3. Aktif Çevrit Modeli

Snake adıyla bilinmekte olan aktif çevrit modeli (active contour model) Kass ve arkadaşları tarafından önerilmiştir [7].



Şekil 3: İşaret esaslı havza bölütlemesi. (a) Renkli görüntü. (b) Gri tonlu görüntü. (c) Görüntü gradyanı. (d) Önplan ve arka plan işaretleri. (e) Değiştirilmiş görüntü gradyanı. (f) Bölütleme sonucu.

Snake şekline ve görüntü üzerindeki konumuna göre tanımlanan bir enerjiye sahiptir. Modelde enerjisi en küçük olan snake eğrisini bulmak amaçlanmaktadır. Kısıtlayıcı kuvvetlere göre biçimini değiştiren snake, birbirine bağlı kontrol noktalarından, $v(s) = [x(s), y(s)]$, $s \in [0, 1]$, oluşmaktadır. Çıkarılmak istenen görüntü kenarlarını snake ile iyi bir şekilde ifade edebilmek için kontrol noktalarının birbirlerine çok yakın ya da çok uzak olmaksızın eşit aralıklarla yerleştirilmesi gerekmektedir.

Snake, verilen bir ilk konumdan başlayarak enerjisi en küçükleyen kenarlar üzerinde sabitlenene değin şekil değiştirmektedir. Modelde snake toplam enerjisi

$$E_{snake} = \frac{1}{2} \int_s [\alpha(s)|v_s|^2 + \beta(s)|v_{ss}|^2] + \gamma E_{dış}(v(s)) ds \quad (1)$$

olarak tanımlanmaktadır [7]. $\alpha(s)$ tarafından kontrol edilen birinci dereceden türev terimi $|v_s|$ snake eğrisinin bir zar gibi hareket etmesini sağlarken $\beta(s)$ tarafından kontrol edilen ikinci dereceden türev terimi $|v_{ss}|$ eğrinin ince bir levha gibi hareket etmesine neden olur. Dış enerji terimi $E_{dış}$ 'in toplam enerjiye olan katkısı γ ile kontrol edilmektedir. E_{snake} 'in sabitlendiği eğriyi bulmak amacıyla Euler-Lagrange türevsel denklemi kullanıldığında

$$\alpha(s)v_{ss} - \beta(s)v_{ssss} - \nabla E_{dış} = 0 \quad (2)$$

elde edilmektedir. (2) numaralı denklem kuvvet denge sistemine

$$F_{iç} + F_{dış} = 0 \quad (3)$$

$$F_{iç} = \alpha(s)v_{ss} - \beta(s)v_{ssss} \quad (4)$$

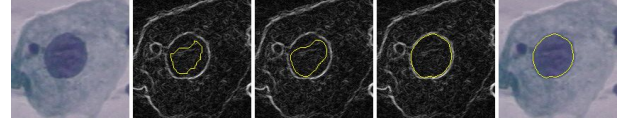
$$F_{dış} = -\nabla E_{dış} \quad (5)$$

denklemleri ile dönüştürülebilir. Bu kuvvet denge sisteminin son çözümü

$$\mathbf{v}_t = [A + \tau \mathbf{I}]^{-1} [\tau \mathbf{v}_{t-1} - \mathbf{F}(\mathbf{v}_{t-1})] \quad (6)$$

olarak elde edilir [7]. Burada, t dürüm sayısını, τ bir dürümdeki adım sayısını, \mathbf{v} snake eğrisini, A kuvvet denge sistemi çözümlenirken oluşturulan matrisi, \mathbf{F} ise görüntü gradyanını temsil etmektedir.

Kass ve arkadaşları tarafından önerilen model [7] dar yakalama aralığına sahip olduğundan kontrol noktalarının ilk konumlarını elde edilmek istenen çevrite çok yakın seçmek gerekmektedir. Dar yakalama aralığı problemini çözmek amacı



Şekil 4: Balon snake modelinin uygulanması. Soldan sağa: renkli görüntü, snake eğrisinin ilk hali, 5 dürüm sonraki hali, 75 dürüm sonraki hali, snake eğrisinin son hali. Sarı eğri snake eğrisini göstermektedir.

ile Cohen [6] tarafından balon gibi davranan aktif çevrit modeli geliştirilmiştir. Balon snake orijinal modelden farklı olarak çevritin normal boyunca şişme ve sönmeye sebep olan bir dış kuvvete sahiptir. Snake bu dış kuvvet sayesinde düşük gradyanlı kenarlar yerine belirgin kenarlar üzerinde son bulabilmektedir.

Balon modelinde toplam kuvvet fonksiyonu

$$F_{snake} = (F_{iç} + F_{dış}) + F_{basınç} \quad (7)$$

$$F_{dış} = -k_2 \frac{\nabla E_{görüntü}}{\|\nabla E_{görüntü}\|} \quad (8)$$

şeklinde ifade edilmiştir. Verilen bir $v(s)$ noktasındaki dış basınç kuvveti

$$F_{basınç} = k_1 n(s) \quad (9)$$

olarak modellenmiştir. Burada $n(s)$, o noktadaki eğrinin normal yönündeki birim vektörü temsil etmektedir. Snake, k_1 'in pozitif değerleri için şişerken negatif değerleri için sönmektedir. k_1 'in mutlak değeri arttıkça snake daha hızlı şekil değiştirmektedir. Balon snake F_{snake} kuvvetini en küçüklemektedir. Balon snake için son çözüm

$$\mathbf{v}_t = [A + \tau \mathbf{I}]^{-1} [\tau \mathbf{v}_{t-1} - \mathbf{F}(\mathbf{v}_{t-1}) + \mathbf{F}_{basınç}] \quad (10)$$

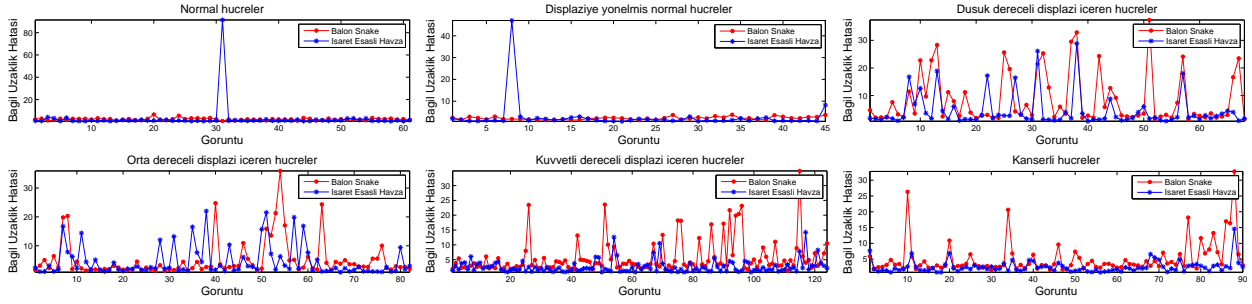
olarak bulunur.

Yapılan çalışmada hücre çekirdeklerinin çevritlerini elde etmek amacıyla balon snake modeli kullanılmıştır. Snake çekirdekler üzerinde bulunan önplan işaretlerinin dış çevresinden başlatılarak belirgin kenarlarda sabitlenene kadar şişerek biçim değiştirmektedir. Çekirdek çevritini elde etmek için uygun parametrelerin seçilmesi önemlidir. Şekil 4'te örnek bir görüntü için aktif çevrit modeliyle elde edilen bölütleme sonuçları gösterilmiştir.

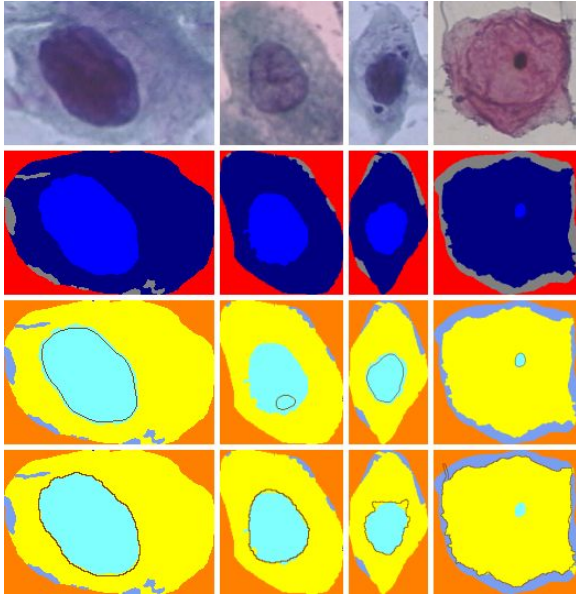
3. Deneysel Sonuçlar

İşaret esaslı havza bölütlemesi ve balon dış kuvvetini kullanan aktif çevrit modeli yapılan deneylerle hücre çekirdeğinin bölütlenmesinde gösterdikleri başarıma göre karşılaştırılmıştır. Her iki yöntem <http://fuzzy.iau.dtu.dk/download/smear2005/> adresinden elde edilen veri kümesi üzerinde uygulanmıştır. Veri kümesi displazi derecesine göre 6 sınıfa ayrılmış serviks hücreleri görüntüleri ve bu görüntülerin uzmanlarca belirlenmiş bölütleme sonuçlarından meydana gelmektedir.

Şekil 5'te işaret esaslı havza bölütlemesi ve balon snake modelinin uygulanmasıyla elde edilen örnek bölütleme sonuçları gösterilmiştir. İlk örnekte iki yöntemle birbirine benzer sonuçlar elde edilmiştir. Snake modeli için belirlenen parametrelerin uygun olmamasından dolayı ikinci örnekte snake eğrisi çekirdek kenarlarına ulaşmadan yerel enküçükte sabitlenmiştir. İşaret esaslı havza bölütlemesi şekil bilgisinden yararlanmadığı için üçüncü örnekte çekirdek kenarlarındaki



Şekil 6: Farklı sınıflar için bölütleme sonuçları. Yukarıdan aşağıya: normal hücreler, displaziye yönelmiş normal hücreler, düşük dereceli displazi içeren hücreler, orta dereceli displazi içeren hücreler, kuvvetli derecede displazi içeren hücreler, kanserli hücreler.



Şekil 5: İşaret esaslı havza bölütlemesi ve balon snake yöntemlerinin örnek görüntüler için verdiği sonuçları. İlk satır orjinal görüntüleri, ikinci satır doğru bölütleme sonuçlarını, üçüncü satır snake ile elde edilen sonuçları ve son satır havza yöntemi ile elde edilen sonuçları göstermektedir.

koyu renkli alanlar da çekirdek sınırları içerisine alınmıştır. Dördüncü örnekte olduğu gibi çekirdek kenarlarının gradyanı sitoplazma kenarlarının gradyanından daha düşük olduğunda işaret esaslı havza bölütlemesi sitoplazma kenarlarını bölütleme çizgisi olarak belirlemektedir.

Şekil 6'da veri kümesinde yer alan 6 sınıfa ait görüntülerin bölütlenmesiyle elde edilen sonuçların hata grafikleri gösterilmiştir. Hata ölçüsü olarak bağıl uzaklık hatası [3] kullanılmıştır. Şekil 6'da farklı sınıflara ait grafiklerde gösterilen yüksek hata değerleri Şekil 5'te örneklendirilmiş nedenlerden kaynaklanmaktadır. Buna göre bölütleme hatalarında etkili olan faktörler balon snake modelinde snake eğrisinin başlangıç konumu ve snake parametreleri iken işaret esaslı havza bölütlemesinde işaretlerin yeri ve alanlarıdır.

4. Sonuçlar

Serviks hücrelerinin displazi derecesini belirleyebilmek için çekirdek ve sitoplazma çevritlerinin çok iyi hesaplanması ge-

reklemektedir. Sunulan çalışmada şekil ve büyüklük bilgisini kullanmaya olanak veren matematiksel biçimbilim işlemleri ile önce çekirdekler üzerinde işaretler hesaplanmış ardından çekirdek çevritlerini bulmak için işaret esaslı havza bölütlemesi ve balon snake modeli denenmiştir. Yapılan deneylerle iki yöntem bağıl uzaklık hatası kullanılarak karşılaştırılmıştır. İşaret esaslı havza bölütlemesi çekirdek kenarlarında yüksek gradyana sahip görüntüler için diğer yöntemlere göre daha iyi sonuç verirken, balon snake modeli uygun parametreler belirlendiğinde düşük gradyanlı çekirdek kenarı ve çekirdek çevresindeki gürültüye karşı daha gürbüz çalışmaktadır. Gelecek çalışma olarak görüntülere gürültü süzgeçlemesinin ardından çekirdek bölütlemesi uygulanabilir ve çekirdekler elipse yakın bir şekle sahip olduğundan balon snake modeline elips şekil ön bilgisi eklenebilir. Ayrıca, çekirdek ve sitoplazma bölütleri elde edildikten sonra bu bölgeleri doku, şekil ve görelî büyüklük gibi özniteliklerle modelleyen, bu modellere göre görüntüler arasında benzerlikler bulan, ve görüntüleri içerdikleri displazi derecesine göre sıralayan bir görüntü erişim sistemi geliştirilmesi planlanmaktadır.

5. Kaynakça

- [1] N. C. Institute, "Health Report Fiscal Years 2005-2006," 2007. [Online]. Available: <http://women.cancer.gov/planning/>
- [2] H. Wu, J. Gil, and J. Barba, "Optimal segmentation of cell images," in *Vision, Image and Signal Processing, IEE Proceedings-*, vol. 145, no. 1, 1998, pp. 50-56.
- [3] S. Yang-Mao, Y. Chan, and Y. Chu, "Edge Enhancement Nucleus and Cytoplasm Contour Detector of Cervical Smear Images," vol. 38, no. 2, pp. 353-366, 2008.
- [4] E. Martin, J. Jantzen, A. Oersted-DTU, and B. Bjerregaard, "Pap-Smear Classification," *Master's Thesis, Technical University of Denmark, Oersted-DTU, Automation*, 2003.
- [5] J. Norup, "Classification of pap-smear data by transductive neuro-fuzzy methods," *Master's Thesis, Technical University of Denmark: Oersted-DTU, Automation*, 2005.
- [6] L. Cohen, "On active contour models and balloons," *CV-GIP: Image Understanding*, vol. 53, no. 2, pp. 211-218, 1991.
- [7] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321-331, 1988.