



NORMAL VE KANSERLİ HÜCRE ÇEKİRDEKLERİNİN DALGACIK DÖNÜŞÜM YÖNTEMİ İLE SINIFLANDIRILMASI

*(CLASSIFICATION OF NORMAL AND CANCEROUS NUCLEI BY WAVELET
TRANSFORM)*

Metehan MAKİNACI*

ÖZET/ABSTRACT

Bu çalışmada, mikroskop, kamera ve sayısallaştırıcı kart kullanılarak elde edilmiş olan prostat hücre çekirdek imgeleri, durağan dalgacık dönüşüm yöntemi kullanılarak modellenmiştir. İşlemlerde Daubechies, Coiflet, çift dikgen ve simetrik dalgacıkların değişik tipleri kullanılmıştır. Öznitelik vektörleri, her kanal için enerji, entropi ve ortalama sapma hesaplanarak oluşturulmuştur. Modelin başarımını ölçmek için, doğrusal ayırtaç sınıflandırıcı kullanılmıştır. Çapraz sağlama yöntemi ile elde edilen sınıflandırma sonuçlarına bakıldığında, çift dikgen dalgacık dönüşümünden elde edilen ortalama sapma özniteliklerinin en iyi sonucu verdiği görülmüştür.

Nucleus images of prostate cells, acquired using a microscope, camera, and a digitizing board, are mathematically modeled by stationary wavelet transform. Daubechies, Coiflet, biorthogonal, and symmetric spline wavelets with different orders are used for the transforms. Feature vectors are calculated from the energy, entropy, and mean deviation representations of each channel output. The performance of the wavelet signatures is measured by using linear discriminant classifier. From the cross-validated classification results, it is demonstrated that mean deviation signatures calculated from the biorthogonal wavelet transform gave the best result.

ANAHTAR KELİMELER/KEYWORDS

Öznitelik çıkarma, Doku analizi, Doku sınıflandırma, Dalgacık dönüşümü
Feature extraction, Texture analysis, Texture classification, Wavelet transform

1. GİRİŞ

Patoloji ve sitolojide, incelenen organ ve hücre imgeleri yoğun doku (*texture*) bilgisi içerir. Uzmanlar, mikroskop ile yapılan incelemelerde, dokusal ve şekilsel değişimleri analiz ederek hastalık ile ilgili tanı koymaktadır. Parça kesitlerinin mikroskop altında incelenmesi çok dikkat ve zaman gerektiren bir iştir (Robbins ve Kumar, 1987; Schenck ve Planding, 1998). Ayrıca, bazı vakaların analizinde uzmanların tanıları arasında %30-40 farklılık olabilmektedir. Bu aşamada, iyi tanımlanmış nesnel ölçütler kullanılarak gerçekleştirilen sayısal imge analizi, tanı başarımlarının artmasında uzmana yardımcı olmaktadır. Yöntemin bir diğer faydası ise analizin otomatik hale gelmesi ile sonuçların kısa sürede elde edilmesidir. Otomasyona bağlı olarak, patoloğun günümüzde giderek artan iş yükünün hafifletilmesi, başarımlarını olumlu yönde etkileyecektir.

Bilgisayar destekli tanı çalışmaları birçok alanı kapsamaktadır. Bunlardan bazıları bağırsak, deri, prostat, meme, serviks ve karaciğer gibi biyolojik dokular ile gerçekleştirilmiştir (Esgiar vd., 1998; Gao vd., 1998; Dalton, 1992; Teague vd., 1997; Erp vd., 1997; Mojsilovic vd., 1998). Smir örneklerinin otomatik olarak tekrar kontrolünü yapan bazı sistemler tarafından değerlendirilmiştir (Grohs vd., 1997; Colgan vd., 1997; Linder, 1995). İmge analizinin sitolojideki uygulamaları ile ilgili genel bir değerlendirme Bahr vd. (1992)'nin çalışmasında verilmiştir.

Doku analizi için bu güne kadar bir çok yöntem kullanılmıştır. Bu yöntemler, son 30 sene içinde, dokunun yapısal, istatistiksel veya yapısal-istatistiksel olma durumuna göre geliştirilmiştir. Bu yöntemlerden bazıları, Fourier güç spektrumu, birinci ve ikinci derece gri seviye istatistikleridir (Weszka vd., 1976; Haralick 1979). Bu çalışmalardan yola çıkılarak geliştirilen fraktal boyut, Gabor süzgeçleri, ve dört evreli süzgeçler hep insanın görme mekanizmasını modellemeyi amaçlamıştır (Du Buf vd., 1990). Yapılan araştırmalar sonucunda, insan görme sisteminde (*striate visual cortex*'de) bulunan nöral algılama alanlarının belli yönelim ve uzaysal frekansa duyarlı, birbirinden bağımsız birçok kanal gibi modellenilebildiği görülmüştür. Doku görüntüsünün insan görme sistemi tarafından çok ölçekli olarak işlenmesi, benzer özelliklere sahip matematiksel modellerin doku analizinde kullanılmasının temel sebebinin oluşturmaktadır (Bovik vd., 1990, Van de Wouwer vd., 1999). Aynı şekilde, dalgacık dönüşümü de bu özelliklere sahip olduğu için çalışmada modelleme amacı ile kullanılmıştır.

Bu çalışmada, ilk olarak Dokuz Eylül Üniversitesi, Tıp Fakültesi, Patoloji Ana Bilim Dalı'nda bulunan prostat kesitleri mikroskop, kamera ve sayısallaştırıcı kart kullanılarak imge olarak bilgisayara aktarılmış ve hücre çekirdekleri uzman tarafından etiketlenmiştir. Durağan dalgacık dönüşümü uygulanmış imgelerin öznitelik vektörleri, enerji, entropi ve ortalama sapma hesaplanarak oluşturulmuştur. Başarımları ölçmek için doğrusal ayırtaç sınıflandırıcı kullanılmıştır.

2. MATERYAL VE YÖNTEM

2.1. Histoloji

Bu çalışmada, hastalardan alınan prostat parçalarından elde edilmiş kesitler kullanılmıştır. Patoloji laboratuvarında uzmanlar tarafından gerçekleştirilen hazırlık işlemleri şu şekilde özetlenebilir: Parçalar önce %10'luk formalin çözeltisinde 24-48 saat bekletilerek sabitlenmiştir. Rutin biyolojik doku incelemesinden sonra elde edilen parafin bloklar, 5-µm kalınlığında kesilerek hematoksilin-eosin ile boyanmıştır.

2.2. İmge Kayıt Sistemi

Prostat doku kesitlerinin görüntüsü Nikon Labophot-2 mikroskop kullanılarak x100 büyütülmüştür. Mikroskoba bağlı renkli kamera (Sony DXC-107AP) ile elde edilen analog görüntüler, bilgisayarda bulunan sayısallaştırıcı kart ve yazılım aracılığıyla 768x576 piksel boyutunda 16 milyon renk olarak kayıt edilmiştir.

Patoloji uzmanının etiketlediği 84 normal ve 86 kanserli hücre çekirdek görüntüsünden 32x32 piksel boyutunda bloklar alınmıştır. Seçilen çekirdeklerin net olmasına, çok koyu renk olmamasına, görsel bozukluklar içermemesine, ayrık olmasına ve bloğu tamamen kapsamasına dikkat edilmiştir. Seçilen çekirdek imgeleri gri tona çevrilmiş ve öznetelik çıkarma işlemi için kayıt edilmiştir.

2.3. Dalgacık Dönüşümü

2 boyutlu ayrık dalgacık dönüşümü, ayrıştırıcı süzgeç gurubu ile imgeyi işleyerek aşağıdaki gibi gerçekleştirilir.

$$L_n(b_i, b_j) = [H_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \quad (1)$$

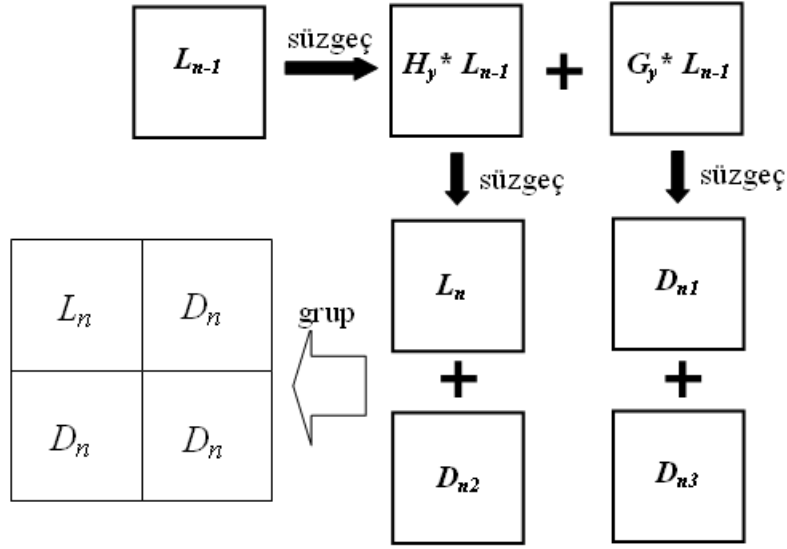
$$D_{n1}(b_i, b_j) = [H_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \quad (2)$$

$$D_{n2}(b_i, b_j) = [G_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \quad (3)$$

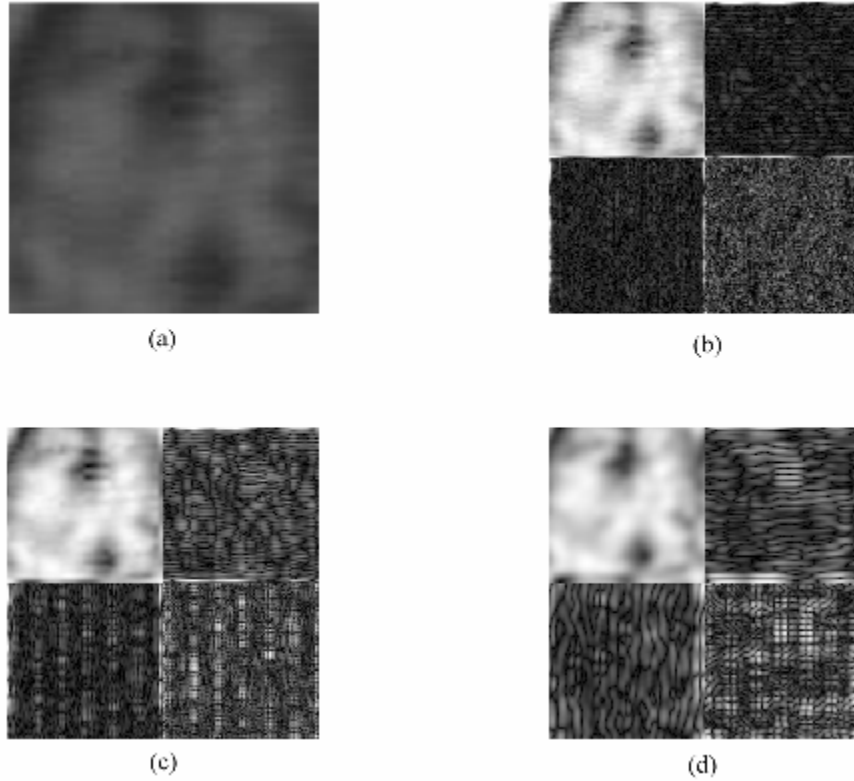
$$D_{n3}(b_i, b_j) = [G_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(b_i, b_j) \quad (4)$$

Burada, * konvolüsyon işleci, $\downarrow_{2,1}(\downarrow_{1,2})$ satır (sütun) boyunca seyrek örnekleme ve $L_0 = I(x, y)$ başlangıç imgesidir. H ve G sırası ile alçak geçiren ve yüksek geçiren süzgeçtir. Eşitlik 1’de alçak geçiren süzgeçlerin kullanılması ile elde edilen L_n , başlangıç imgesinin n . ölçekteki düşük çözünürlükteki imgesidir. D_{ni} ler n . ölçekteki ayrıntı bilgilerini içeren imgelerdir ve süzgeçlerin belli yönlerde başlangıç imgesine uygulanması sonucu elde edilir (Van de Wouwer vd., 1999). Dalgacık dönüşümünün bir görüntü bloğuna uygulanması Şekil 1’de verilmiştir.

Bu çalışmada, hücre çekirdeklerinden seçilen blokların boyutu küçük (32x32 piksel) olduğu için, durağan dalgacık dönüşümü kullanılmıştır. Bu yöntemde seyrek örnekleme yapılmadan filtreler uygulanmaktadır. Şekil 2’de, örnek bir çekirdek görüntü bloğuna, 3 seviyeye kadar uygulanmış dalgacık ayrıştırması sonuçları verilmiştir.



Şekil 1. Dalgacık dönüşümünün bir görüntü bloğuna uygulanması



Şekil 2. (a) orijinal görüntü, dalgacık ayrıştırma sonuçları: (b) 1. seviye, (c) 2. seviye, (d) 3. seviye

2.4. Öznitelik Hesaplama

Dalgacık dönüşümü ile elde edilen farklı ölçekteki imgelerin (katsayıların) öznitelik vektörleri 3 farklı yöntem ile hesaplanmıştır (Eşitlik 5, Eşitlik 6 ve Eşitlik 7). Belli bir ölçekteki enerji dağılım parametresi, doku analizinde kullanılan önemli bilgiler içerir. Kanal çıkışının enerjisi eşitlik 5 ile hesaplanır. N ve M sırası ile piksel cinsinden blok yüksekliği ve genişliğidir. Eşitlik 6'da verilen ortalama sapma, enerji parametresine benzer şekilde, bir

kanaldaki dalgacık katsayılarının saçılma miktarının ölçütüdür ve bu iki parametre birbiri ile ilintilidir (Van de Wouwer vd., 1999). Entropi ise, kanalın bilgi miktarını veren bir ölçüttür ve öznel olarak doku analizinde kullanılmaktadır (Chen ve Kundu, 1994; Laine ve Fan, 1993). Bu hesaplamalarda üçüncü ölçeğe kadar olan ayrıntı imgeleri kullanıldığı için, her çekirdeğe ait öznel vektörü (enerji, ortalama sapma ve entropi vektörünün her biri) toplam 9 eleman içermektedir.

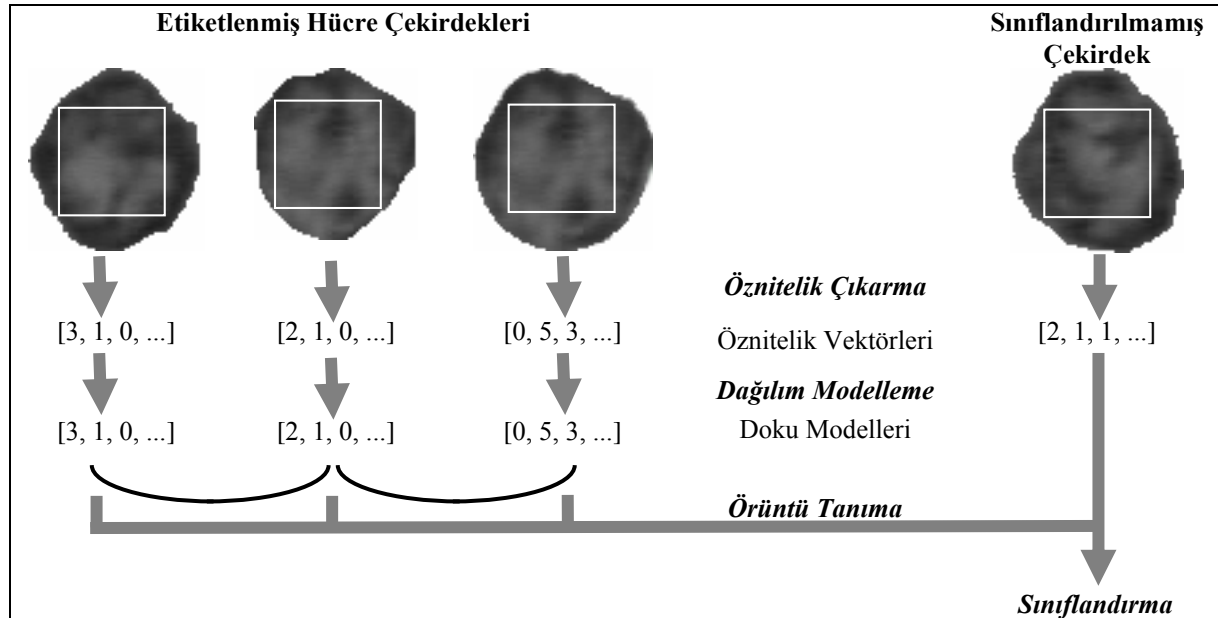
$$\text{Enerji} \quad E_{ni} = \frac{1}{NM} \sum_j^{N-1} \sum_k^{M-1} D_{ni}(b_j, b_k)^2 \quad (5)$$

$$\text{Ortalama Sapma} \quad OS_{ni} = \frac{1}{NM} \sum_j^{N-1} \sum_k^{M-1} |D_{ni}(b_j, b_k)| \quad (6)$$

$$\text{Entropi} \quad ENT_{ni} = - \sum_j^{N-1} \sum_k^{M-1} |D_{ni}(b_j, b_k)|^2 \log |D_{ni}(b_j, b_k)|^2 \quad (7)$$

2.5. Sınıflandırma

Sınıflandırma işleminde 84 normal ve 86 kanserli hücre çekirdek imgesinin öznel vektörleri kullanılmıştır. Enerji, ortalama sapma ve entropi öznelikleri ayrı ayrı kullanılarak doğrusal ayırtaç sınıflandırıcılar oluşturulmuştur. Sınıflandırıcılar oluşturulurken birini dışarıda bırak (çapraz sağlama) yöntemi uygulanmıştır. Bu yöntemde, sınıflandırıcıyı hesaplamak için tek bir öznel vektörü hariç tüm vektörler kullanılır (Şekil 3).



Şekil 3. Sınıflandırma sistemi

Başarımı ölçmek için de ayrılan vektör kullanılır. Bu şekilde, her defasında bir öznel vektörü dışarıda bırakılarak yaratılan sınıflandırıcıların başarımlarından genel başarımları hesaplanır.

3. SONUÇLAR

Öznitelik vektörlerini hesaplamak için farklı derecelere sahip dalgacık dönüşümleri kullanılmıştır. Bu dalgacık dereceleri: Daubechies için 1-10, Coiflet için 1-5, simetrik dalgacık (symlet) için 1-8 ve çift dikgen için 1.1, 1.3, 1.5, 2.2, 2.4, 2.6, 2.8, 3.1, 3.3, 3.5, 3.7, 3.9, 4.4, 5.5, 6.8'dir. Çift dikgen dalgacığın a.b şeklinde ifade edilen derecesinde a yeniden oluşturma, b ise ayrıştırma süzgecinin derecesini vermektedir. Her dalgacık türü ve derecesi ile elde edilen öznitelik vektörlerinin ayırım gücü, doğrusal ayırtaç sınıflandırıcı kullanılarak ölçülmüştür. Enerji, entropi ve ortalama sapma özniteliklerinden çapraz sağlama yöntemi ile elde edilen genel başarımlar oranları Şekil 4'de gösterilmektedir. Ortalama sapma öznitelikleri, tüm dalgacık türlerinde en iyi başarımlar oranlarını vermiştir ve bu öznitelik vektörleri ile elde edilen en iyi başarımlar oranları Daubechies, Coiflet, simetrik ve çift dikgen için sırası ile %80.6, %81.8, %81.2 ve % 82.9 olarak hesaplanmıştır. Başarımlar oranlarına dalgacık türü açısından bakılacak olursa, en yüksek değerler çift dikgen ile elde edilmiştir. Bu değerler ise, ortalama sapma, enerji ve entropi için sırası ile %82.9, %75.9 ve %66.5'dir. Çift dikgen dalgacık dönüşümünün sağladığı özniteliklerin diğerlerine göre çok az daha başarılı sonuçlar vermesi, temelde bu dönüşümün ayrıştırma için ayrı, yeniden oluşturma için ayrı süzgeçler kullanması ile açıklanabilir. Bu sayede, sadece analiz için gerekli olan (osilasyon, sıfır momenti gibi) özelliklerin ayrıştırma süzgecinde yoğunlaştırılması daha iyi sonuçlar elde edilmesini sağlamıştır (Misiti vd., 1996).

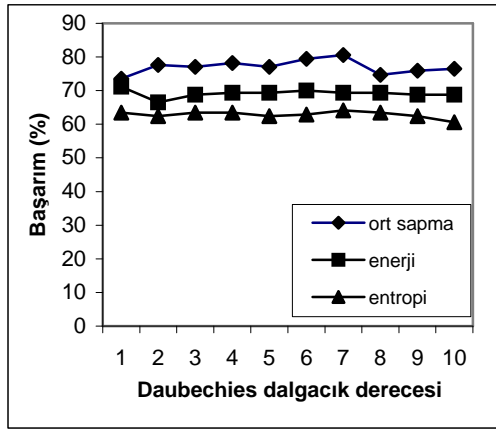
4. TARTIŞMA

Bu çalışmada, normal ve kanserli hücre çekirdeklerinin sınıflandırılması probleminin çözümü için, dalgacık dönüşümü tabanlı, çoklu çözünürlüklü bir yaklaşım sunulmuştur. Kullanılan dalgacık parametrelerinin çoklu çözünürlükteki etkin modelleme yeteneği deneysel sonuçlar ile gösterilmiştir. Genelde, dalgacık dönüşüm derecesinin artması ile başarımlar oranının arttığı görülmüştür. Bu duruma aykırı olarak, Daubechies'in 8., 9. ve 10. derece dönüşümleri 7. dereceden daha düşük başarımlar oranları vermiştir. Birden fazla ölçekte yapılan dönüşümler, hücre çekirdeklerinde bulunan hem yerel hem de genel özellikleri modellemede başarılı olmuştur. Ortalama sapma ve enerji özniteliklerinin birbiri ile oldukça ilintili olmasına rağmen ilk yöntemin daha başarılı sonuç vermesinin nedeni, D_{ni} 'lerin mutlak değerlerinin alınması sonucunda oluşan sınıf için uzaklıklarının daha az olması ve buna bağlı olarak normal ve kanserli sınıflarının birbiri içine daha az girmesidir.

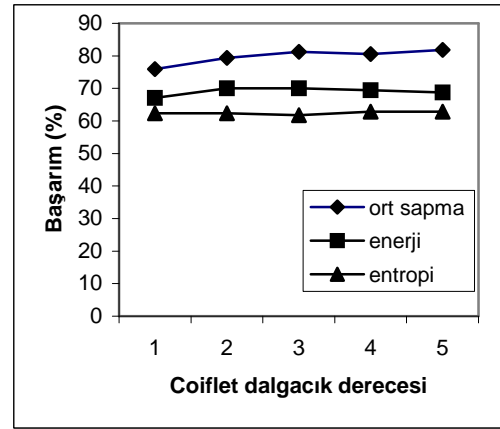
Başarımlar sonuçlarına bakıldığında, birden fazla ölçekte bulunan, istatistiksel düzenliliğe sahip bilginin dalgacık dönüşümü ile modellenilebildiği görülmektedir. Yöntemin tıbbi tanı sürecinde kullanılması ile, özellikle anormal prostat hücre çekirdeklerinin belirlenmesinde patoloğa önemli ek bilgiler sağlanabilecektir.

Çalışmanın daha sonraki genişleme konuları aşağıda özetlenmektedir:

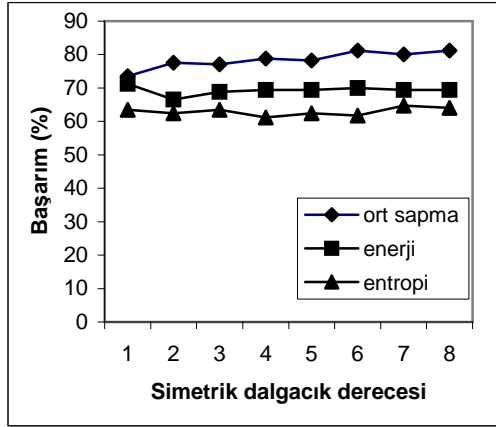
- § Özellikle dokusal özniteliklerin hesaplanmasında kullanılan ve zaman alan fonksiyonlar, optimizasyon ve paralel çalışma algoritmaları ile daha kısa sürede, gerçek zamanlı olarak gerçekleştirilebilir.
- § Farklı öznitelik çıkarma ve sınıflandırma algoritmaları kullanılarak daha başarılı sonuçlar elde edilebilir.
- § İmge analiz yazılımı, mikroskop-kamera sistemi ile birleştirilerek gerçek zamanlı çalışan, otomatik bir sistem elde edilebilir.



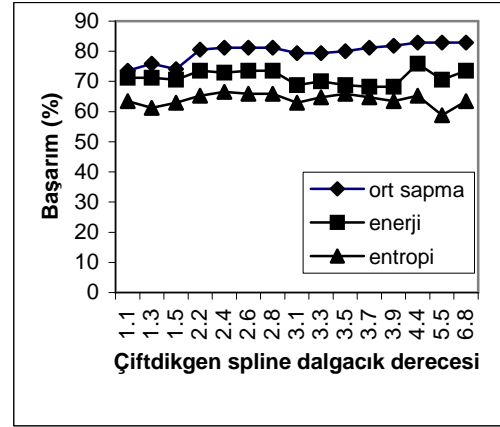
(a)



(b)



(c)



(d)

Şekil 4. Farklı dalgacık dönüşümü ile elde edilen öznelitliklerin sınıflandırma başarımları

TEŞEKKÜR

Görüntülerin sağlanmasında ve etiketlenmesinde yaptığı katkılardan dolayı DEÜ Tıp Fakültesi, Patoloji A.B.D.'dan Prof. Dr. Kutsal Yörükoğlu'na teşekkür ederim.

KAYNAKLAR

- Bahr G.F., Bartels P.H., Dytch H.E., Koss L.G., Wied G.L. (1992): "Image Analysis and Its Applications to Cytology", p. 1572-1605, Diagnostic Cytology and its Histopathologic Bases. (4th ed.), Koss, L.G.(ed) , Lippincott Williams & Wilkins Publishers.
- Bovik A.C., Clark M., Geisler W.S. (1990): "Multichannel Texture Analysis using Localized Spatial Filters", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 12, p. 55-73.
- Chen J.L., Kundu A. (1994): "Rotation and Gray Scale Transform Invariant Texture Identification using Wavelet Decomposition and Hidden Markov Model", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 16, p. 208-214.
- Colgan T.J., Bon N., Lee J.S.J., Patten S.F.Jr. (1997): "AutoPap 300 QC System Scoring of Cervical Smears without Epithelial Cell Abnormalities", Acta Cytol. Vol. 41, p. 45-49.

- Dalton III L.W. (1992): "Computer-Based Image Analysis of Prostate Cancer: Comments with Emphasis on Use of Commercially Available Systems", *Human Pathology*, Vol.23, p. 280-286.
- Du Buf J. M.H., Kardan M., Spann M. (1990): "Texture Feature Performance for Image Segmentation", *Pattern Recognition*, Vol. 23, No. 3/4 , p.291-309.
- Erp A.J.M.V.A.-V., Hof-Grootenboer B.E.V., Brugal G., Vooijs G.P. (1997): "Identifying Cytologic Characteristics and Grading Endocervical Columnar Cell Abnormalities: A Study Aided by High-definition Television", *Acta Cytol.*, Vol. 41, p. 1659-1670.
- Esgiar A.N., Naguib R.N.G., Sharif B.S., Bennett M.K., Murray A. (1998): "Microscopic Image Analysis for Quantitative Measurement and Feature Identification of Normal and Cancerous Colonic Mucosa", *IEEE Tran. Inf. Tech. In Biomedicine*, Vol. 2, no. 3, p. 197-203.
- Gao J., Zhang J., Fleming M.G., Pollak I., Coggnetta A.B. (1998): "Segmentation of Dermatoscopic Images by Stabilised Inverse Diffusion Equations", *Proc. Int. Conf. on Image Processing*, Vol. 3, p. 823 -827.
- Haralick R.M. (1979): "Statistical and Structural Approaches to Texture", *Proc. Of the IEEE*, Vol. 67, no. 5, p. 786-804.
- Laine A., Fan J. (1993): "Texture Classification by Wavelet Packet Signatures", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 15, pp 1186-1191.
- Linder J. (1995): "Evaluation of PAPNET System for Rescreening of Negative Cervical Smears" *Diag. Cytopathology*, Vol. 13, no. 1, p. 31-36.
- Misiti M., Misiti Y., Oppenheim G., Poggi J.M. (1996): "Wavelet Toolbox for Use with Matlab®", The MathWorks Inc.
- Mojsilovic A., Popovic M., Markovic S., Krstic M. (1998): "Characterization of Visually Similar Diffuse Diseases from B-scan Liver Images using Nonseparable Wavelet Transform", *IEEE Trans.. Medical Imaging*, Vol. 17, p.541-549.
- Robbins S.L., Kumar V. (1987): "Basic Pathology", Saunders Co.
- Schenck U., Planding W. (1998): "Quantitation of Visual Screening Technique in Cytology", *Proc. Image Analysis in Medicine, II. National Symposium*, p. 7-14.
- Teague M.W., Wolberg W.H., Street W.N., Mangasarian O.L., Lambremont S., Page D.L. (1997): "Indeterminate Fine-needle Aspiration of the Breast. Image Analysis-assisted Diagnosis", *Cancer* Vol. 81, p. 129-135.
- Van de Wouwer G., Scheunders P., Van Dyck D. (1999): "Statistical Texture Characterization from Discrete Wavelet Representation", *IEEE Trans. on Image Processing*, Vol. 8, pp. 592-598.
- Weszka J. S., Dyer C. R., Rozenfeld A. (1976): "A Comparative Study of Texture Measures for Terrain Classification", *IEEE Trans. on Systems, Man, and Cybenetics*, Vol. SMC-6, no. 6, p. 269-285.