

Mamografi İmgelerinden HOG Öznitelikleri Çıkartılarak Hastaların Kansere Seviyelerinin Belirlenmesi

Murat YOLDAŞ

Eskişehir Osmangazi Üniversitesi
Elektrik-Elektronik Mühendisliği Bölümü
muratyoldas44@gmail.com

Mehmet DİRLİKLİ

Eskişehir Osmangazi Üniversitesi
Elektrik-Elektronik Mühendisliği Bölümü
mehmetdirlikli@yahoo.com

Semih ERGİN

Eskişehir Osmangazi Üniversitesi
Elektrik-Elektronik Mühendisliği Bölümü
sergin@ogu.edu.tr

M. Oğuz ŞAKAR

Eskişehir Osmangazi Üniversitesi
Elektrik-Elektronik Mühendisliği Bölümü
mehmetoguz89@yahoo.com.tr

Onur KILINÇ

Eskişehir Osmangazi Üniversitesi
Elektrik-Elektronik Mühendisliği Bölümü
onur_kilinc@anadolu.edu.tr

Özet

Dünya geneli açısından bakıldığında, meme kanserinin kadınlarda en sık görülen kanser türü olduğu ifade edilmektedir. Son yıllarda her yaştaki bayanların korkulu rüyası haline gelen meme kanseri, daha çok 40 yaş üstü bayanlarda görülmektedir. Meme kanserinin gerek hastanın gerekse de yakınlarının üzerinde fiziksel, sosyal ve ruhsal etkileri bulunmaktadır. Meme kanserinin insan hayatı açısından taşıdığı riskler göz önüne alındığında, belirli aralıklarla mamografilerin çekilmesi gerekmektedir. Bu mamografi imgelerinin uzmanlar tarafından incelenmesi istenilen sonuçları tam olarak vermediğinden dolayı, sayısal ortamda işlenip incelenmesi daha faydalı olacaktır. Bu çalışmada, öncelikle, mamografi imgelerinden “Histogram of Oriented Gradient (HOG)” yöntemiyle öznitelikler çıkartılmış ve hesaplanan bu öznitelikler k-En Yakın Komşuluk (k- Nearest Neighbourhood) sınıflandırma yönteminde kullanılarak, bilgisayar ortamında bireylerin kanser seviyelerinin isabetli bir şekilde tespiti hedeflenmiştir. Elde edilen tanıma oranları incelendiğinde, oldukça karmaşık bir örüntü tanıma problemi olan meme kanseri teşhisine dikkate değer bir çözüm getirilmiştir.

Anahtar kelimeler: Meme kanseri, Mamogram, HOG Öznitelik Çıkartımı, İmge Sınıflandırma.

1. Giriş

Kanser çağımızın en tehlikeli ve ölümcül hastalıklarından birisidir. Bu ölümcül hastalıkta, erken tanı, teşhis ve tedavi hayatı önem taşımaktadır. Meme kanseri ise, özellikle kadınlarda en sık görülen kanser tipi olmasıyla beraber nadiren (bayanlara nazaran 1/150 oranında) erkeklerde de görülmektedir [1, 2]. Kanser, hücrelerde çevresel ve kalıtsal nedenlerden dolayı DNA'nın hasar alması sonucu, hücrelerin kontrolsüz veya anormal bir şekilde büyümesi ve çoğalmasıdır. Dünyada genelinde meme kanseri bayanlarda, kansere bağlı ölümlerde ilk sırada yer almaktadır. Erken tanıda eski dönemlere nazaran fark edilebilir bir gelişme gözlenmiştir. Erken teşhiste hekim muayenesi, meme ultrasonografisi ve iki taraflı mamografi ile memelerin kansere karşı etkin bir şekilde incelenmesi mümkündür. Bu yöntemlerden mamografi, erken teşhiste en etkin yöntem olarak bilinmektedir [3]. Mamografi düşük dozda X-ray ışını kullanan bir görüntüleme metodudur. Mamografi etkinliği kanıtlanmış bir tarama yöntemidir ve taramalar

sayesinde meme kanserinden ölüm oranlarının düştüğü ortaya konmuştur [4, 5]. Literatürde yer alan bilgiler incelendiğinde; dünyada her üç dakikada bir kadına meme kanseri tanısı konulduğu ifade edilmektedir [6, 7]. Mamografilerden kanser hastalığının hangi aşamada olduğu ve kişilerin ne tür bir tümörle karşı karşıya oldukları belirlenebilir. Fakat mamografi kullanılarak yapılan doktor tanımlarında hâlâ birçok hatalar yapılmaktadır. Uzmanlar tarafından mamografilere bakılarak meme kanseri teşhisi konulan hastalarda, bu teşhislerin çoğunun yanlış olduğu yürütülen biyopsi sayesinde belirlenmiştir [8]. İlk olarak Nico Karssemeijer dijital mamogram imgelerinden istatistiksel metotla öznitelik çıkarımı uygulamıştır. Bu metot temelde Bayes sınıflayıcı yöntemlerinden esinlenerek gerçekleştirilmiştir. HOG tanımlayıcılar ilk olarak Shashua ve Dalal tarafından önerilmiştir. Nesne ve örüntü tanıma için yaygın olarak kullanılmaya başlanan bu basit ve etkili yöntem, birçok konuda yüksek oranda başarıyla çalışabilmektedir.

Bu çalışmada, öncelikle, mamografi imgelerinden “Histogram of Oriented Gradient (HOG)” yöntemiyle öznitelikler çıkartılmıştır. Çalışmanın ikinci aşamasında ise, hesaplanan öznitelikler k-En Yakın Komşuluk (k-Nearest Neighbourhood) sınıflandırma algoritması kullanılarak, bilgisayar ortamında bireylerin kanser seviyelerinin isabetli bir şekilde tespiti hedeflenmiştir. Yürütülen bütün deneysel çalışmalar Matlab ortamında yapılmıştır. Bildirinin geri kalan kısmında, önce, mamografi imgelerine bir ön işlem olarak uygulanan “Non Local Means” filtreleme yöntemi anlatılmış olup; sırasıyla, öznitelik çıkartım ve sınıflandırma metotları anlatılmıştır. Deneysel çalışmaların sonuçları verildikten sonra, en son olarak, çalışmadan elde edilen sonuçlar yazılmıştır.

2. Filtreleme

Çalışmada gözlemlendiğimiz sonuçların daha etkin ve geçerli olmasını sağlamak amacıyla elimizde bulunan mamografi imgeleri filtreleme aşamasından geçirilmiştir. Filtreleme sayesinde, görüntü üzerinde netleştirme, belirli ayrıntıları ortaya çıkarma, görüntüyü yumuşatma, kenar keskinleştirme veya kenar bulma gibi işlemler gerçekleşir [9]. Mamografi çekilirken ortamda bulunan birçok sebepten (gürültü, nem, elektromanyetik dalgalar vb.) dolayı görüntüde istenmeyen parazitler oluşur. Bu problemi gidermek için mamografi imgeleri, Non Local Means (NLM) tipi filtreden geçirilmiştir. Bu filtre görüntüyü istenilen seviyeye netleştirirken herhangi bir piksel kaybına yol açmamaktadır.

3. Öznitelik Çıkartımı

Bu çalışmanın öznitelik çıkartım aşamasında, bir nesne tanımlama algoritması da olan Histogram of Oriented Gradient (HOG) yöntemi kullanılmıştır. Bu yöntem, imge üzerine yerleştirilen bir pencerenin birbiri üzerine binen bloklara ayrılması ve her blok içindeki piksellerin genliklerinin bir yönelim histogramında toplanmasıyla öznitelik verilerinin elde edilmesidir [10]. Bu çalışmada blok boyutunun etkisini daha iyi gözlemleyebilmek için, 6x6, 8x8, 16x16 şeklinde 3 ayrı blok boyutunda HOG öznitelikleri çıkartılmıştır.

4. Sınıflandırma

Kanserli dokular, benign (iyi huylu) ve malignant (kötü huylu) tümörler olmak üzere iki şekilde isimlendirilirler. Kötü huylu tümörler, iyi huylu tümörlerin aksine başka dokulara ve organlara sızma ve yayılma özelliği göstermektedir. İyi huylu tümörler dokudan alındığı zaman tekrardan oluşma ihtimalleri yokken; kötü huylu tümörlerin dokudan alınsa bile, bireyin yaşamı boyunca tekrar kendilerini yenileme olasılıkları vardır [11]. Bundan dolayı kanser vakalarında tümörlerin çeşidinin belirlenerek tedaviye başlanması gerekmektedir. İki yönlü mamografi tetkikinde, her bir meme için 2 imge bulunmaktadır. Çalışma veri kümemizde 50 hastaya ait 200 mamografi imgesi bulunmaktadır. Bu imgelerden, 67 adedi normal, 112 adedi iyi huylu ve 14 adedi kötü huylu mamografi imgesidir. Deneyisel sonuçların kıyaslanabilmesi için, her veri sınıfından 60 adet mamografi imgesi kullanılmasına karar verilmiştir. Eksik olan 46 adet kötü huylu mamografi imgesi ise, veri kümesinde bulunan 14 adet kötü huylu mamografi imgesinin 12'li alt kümelerinin içerisinde oluşan imgelerin ortalamasından elde edilen sanal imgeler ile tamamlanmıştır.

5. Deneyisel Çalışmalar

Yürütülen deneyisel çalışmalarda, öncelikle normal, iyi huylu ve kötü huylu kanser dokularına ait 3 farklı veri sınıfındaki toplam 180 imge ayrı ayrı NLM filtresinden geçirilerek, imgelerin üzerindeki istenmeyen gürültüler ortadan kaldırılmıştır. Temizlenen bu imgelerden, 6x6, 8x8, 16x16 şeklinde 3 ayrı blok boyutunda HOG öznitelikleri çıkartılmıştır. Elde edilen öznitelik matrisleri, sırasıyla, 14535x58, 8192x58, 2048x58 boyutlarındadır. Bu öznitelik matrislerinin her birinin sütunlarında bulunan 14535, 8192, 2048 boyutlu vektörlerden, enerji, ortalama, standart sapma, maksimum, minimum, sivruluk (kurtosis) ve yamukluk (skewness) değerleri hesaplanmıştır. Böylece, her bir ayrı blok boyutundan 7x58'lik matrisler bulunmuştur. Bu matrislerin sütunları alt alta getirilerek 406x1 boyutlu vektörler elde edilmiştir. Daha sonra, bu vektörler, k-En Yakın Komşuluk (k-Nearest Neighbourhood) sınıflandırma algoritmasına tabi tutulmuştur. Yapılan tüm sınıflandırma çalışmalarında k-En Yakın Komşuluk yöntemi k parametresi 5 olarak seçilmiştir. Sınıflandırma algoritmasının eğitim aşamasında her bir veri sınıfına ait 45, test aşamasında ise 15 adet imge kullanılmıştır. Test aşamasında kullanılan imgeler, tamamıyla rastgele seçilmiştir. Elde edilen tüm

tanıma oranları ayrıntılı bir biçimde Tablo 1' de verilmiştir. Tanıma oranları, normal, iyi huylu ve kötü huylu kanser veri sınıflarının ortalama tanıma oranlarıdır. Bu sayede, üç sınıflı karmaşık bir örüntü tanıma problemine çözüm getirilmeye çalışılmıştır.

Tablo 1: Meme Kanseri Seviyesi Ortalama Tanıma Oranları

HOG Öznitelikleri Hesaplanırken Kullanılan Blok Boyutu	Tanıma Oranları (%)
6x6	55.6
8x8	64.4
16x16	68.9

6. Sonuç

Bu çalışmada, olası yanlış teşhislere karşın, etkin bir öznitelik çıkarımı metodu olan HOG ve örüntü sınıflandırma algoritması kullanılarak hastaların kanser seviyeleri belirlenmeye çalışılmıştır. Mamografi imgelerinden bireylerin, hastalığın hangi aşamasında olduğu ve ne tür bir tümörle karşı karşıya oldukları saptanmış oldu. Buna ek olarak biyopsiye olan gereksinimi azalttığı için meme kanseri taramalarındaki maliyeti düşürmesi beklenmektedir. Dolayısıyla meme kanseri taramalarının daha kısa zaman dilimleri içinde tekrarlanması yaygınlaştıracaktır. Deneyisel çalışmaların test-eğitim aşamasında elde edilen sonuçlar incelendiğinde, en iyi tanıma oranına 16x16 boyutlarındaki HOG öznitelik vektörleriyle ulaşılmıştır. Ayrıca bu incelemeler sonucunda, HOG öznitelik vektörlerinin blok boyutları küçüldükçe tanıma oranının da düştüğü gözlemlenmiştir.

7. Kaynaklar

- (1) Hossfeld, DK. Manual of Clinical Oncology, 5th ed. Springer- Verlag, UICC, 1992.
- (2) Topuz, E. Aydinler, A. Dincer, M. Meme Kanseri, Nobel Tip Kitabevi, 2003.
- (3) Baldi, V. Cicalese, A. Coppola, G. Scarano, M. Gatto, E. Results of Mammographic Screening, 2003.
- (4) Humprey, LL. Helfand, M. Chang, BK. Woolf, SH. "Breast cancer screening: a summary of the evidence for the U.S Preventive Services Task Force", 2009.
- (5) Kopans, D.B. Feig, S.A. The Canadian National Breast Cancer Screening Study, 1993.
- (6) Howlader, N. Noone, AM. Krapcho, M. et. al. (eds). SEER Cancer Statistics Review, 1975-2009
- (7) National Cancer Institute. Bethesda. 2012.
- (8) American Cancer Society Guidelines for The Early Detection of Cancer, 2001.
- (9) Kevrann, C. Boulanger, J. Coupe, P. Bayesian Non-local Means Filter, Image Redundancy and Adaptive Dictionaries for Noise Removal, 2007.
- (10) Barman, H. Gösta, HG. and Haglund, L. Feature Extraction for Computer-Aided Analysis of Mammograms, 1993.
- (11) Wooster, R. Neuhausen, SL. Mangion, J. Quirk, Y. Ford, D. Nyungen, K. Seal, S. Tran, T. Averill, D. Localization of breast cancer susceptibility, 1994.