

## MAMOGRAFİ NORMALİZASYONUNUN SINIFLANDIRMA PERFORMANSI ÜZERİNDEKİ ETKİLERİ

*UMAS 2017'de sunulmuş ve genişletilmiş bildiridir.*

Ayşe AYDIN YURDUSEV<sup>1</sup> Canan ORAL<sup>2</sup> Mahmut Hekim<sup>2</sup>

<sup>1</sup> Amasya Üniversitesi, Teknoloji Fak., Elektrik-Elektronik Müh., 05100, Amasya, TÜRKİYE

<sup>2</sup> Amasya Üniversitesi, Teknoloji Fak., Elektrik-Elektronik Müh., 05100, Amasya, TÜRKİYE

<sup>3</sup> Gaziosmanpaşa Üniversitesi, Mühendislik ve Doğa Bilimleri Fak., Elektrik-Elektronik  
Mühendisliği, 60150, Tokat, TÜRKİYE

ayse.yurdusev@amasya.edu.tr

**Özet-** Bilgisayar destekli teşhis için dijital mamogramların belirli ön-işlemlerden geçirilmesi gerekmektedir. Ön-işleme mamografi imgelerinde gürültü temizleme ve yapay olguları kaldırma amacıyla yapılmaktadır. Ön-işlemeden geçen mamogram imgeleri çeşitli tekniklerle ve çeşitli sınıflandırıcılarla anomali bulguları için işlenmektedir. Bu çalışmanın amacı uygulanan çeşitli ön-işleme adımlarının, dokusal özellikler kullanılarak sınıflandırma sonuçlarına etkisini incelemektir.

Çalışmada gerçekleştirilen ön-işleme algoritmasında MIAS (The Mammographic Image Analysis Society) veri tabanına ait mamogramlar ortanca süzgeç ile küçük gürültülerden, eşikleme yöntemleri ve morfolojik işlemler ile yapay gürültülerden ve pektoral kasta temizlenmiştir. Temizlenen imgeler 512×256 piksel boyutlarına getirilmiş, son olarak ise kontrast sınırlı adaptif histogram eşitleme (CLAHE) yöntemi ile normalizasyon sağlanmıştır.

Sunulan çalışmada mamogram görüntüleri 4 farklı gruba ayrılmıştır:

Grup 1: özgün MIAS veri tabanından alınan, ön-işleme tabii tutulmamış mamogram görüntüleri

Grup 2: gürültülerden ve pektoral kasta temizlenmiş mamogram görüntüleri

Grup 3: temizlenmiş ve boyutu 256x512'ye sabitlenmiş mamogram görüntüleri

Grup 4: temizlenmiş, boyutlandırılmış ve kontrast eşitlenmiş mamogram görüntüleri

Her bir grup mamogram görüntülerinden yaygın olarak kullanılan birinci dereceden özellikler hesaplanarak DVM (Destek Vektör Makineleri) sınıflandırıcıya uygulanmıştır. Sınıflandırma sonuçlarına göre ön-işleme adımlarından gürültülerin ve pektoral kasta temizlenmesi en iyi sonucu vermiştir.

**Anahtar Kelimeler-** Dijital Mamografi, Pektoral Kasta, Dokusal Özellikler, Destek Vektör Makineleri, MIAS Veri Tabanı

## THE EFFECTS OF MAMMOGRAPHY NORMALIZATION ON CLASSIFICATION

**Abstract-** Digital mammograms should be preprocessed for computer-aided diagnosis. The aims of preprocessing are denoising and eliminating of artifacts. Mammograms are computed by different techniques and classifiers, after preprocessing step. The purpose of this study is the evaluation of the classification rates with first order textural features by different preprocessing steps.

In the study digital mammograms are taken from MIAS database. The algorithm of the preprocessing step of this study includes noise clearance by median filter and artifact noise and pectoral muscle elimination by threshold techniques and morphological operations. Denoised images are normalized by a size of 512x256 pixels. Then, contrast-limited adaptive histogram equalization (CLAHE) is applied.

In proposed study, mammogram images divided into 4 groups.

Group 1: Original MIAS mammograms, without any preprocessing operation

Group 2: Mammograms that are cleaned noises and pectoral muscles

Group 3: The dimension normalized mammograms at 512x256 pixels

Group 4: CLAHE applied mammograms

For each group mammograms commonly used features are extracted and SVM classifier are used. According to classification results, the best classification rate is implemented by noise and pectoral muscle are eliminated groups (Group 2).

**Key Words-** Digital mammography, Pectoral Muscle, Textural Features, Support Vector Machines, MIAS database

## 1. GİRİŞ (INTRODUCTION)

Meme kanseri kadınlarda en çok ölümlü sonuçlanan ve kadın ve erkeklerde, akciğer kanserinden sonra ikinci en çok ölümlü sonuçlanan kanser türüdür [1][2]. Dünya çapında her yıl yaklaşık 170 bin kadında meme kanseri görülmektedir ve kanserden dolayı ölümlerin %17'sinin sebebi meme kanseridir [3]. Bu sebeple meme kanseri teşhis ve tedavisi hayati önem taşımaktadır. Meme kanseri teşhisi için çeşitli görüntüleme teknikleri kullanılarak kanserli hücrenin tespiti yapılmaktadır. Görüntüleme teknikleri arasında mamografi, ya da mamogram, tarama ve teşhis için bilinen en kullanışlı yöntemdir [1,4–6]. Bilgisayar destekli teşhis (BDT) ise dijital mamografi üzerinden memedeki anormal dokuları tespit edebilen bir bilgisayar algoritmasıdır. BDT sistemleri mamogram okunmasında uzmanlara yardımcı olmakta, ikinci bir görüş sunmaktadır [7,8]. Böylelikle mamografi üzerinden yapılan teşhislerde yanlış pozitif/negatif oranının düşürülmesi hedeflenmektedir.

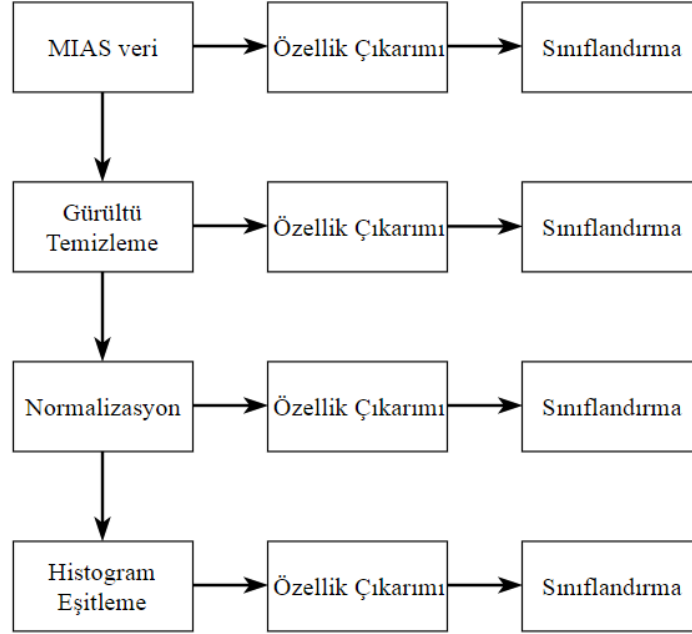
Literatürde BDT için kullanılan geniş bir algoritma yelpazesi bulunmaktadır. Bunlardan bir tanesi de dijital mamogram üzerinde dokusal özelliklerin hesaplanarak sınıflandırma yapılmasıdır [7–9]. Sahiner ve ark. uzamsal bölgede dokusal özellikleri kullanarak konvolüsyonel sinir ağları ile mamogram kesitlerini normal ve kütleli olarak sınıflandırmışlardır [7]. Chen ve ark. ise meme sonogramlarını dokusal özellikler kullanarak sınıflandırmış, kütleli memelerde kütleyle ilişkili kanser ve kanser değil sınıflandırması yapmışlardır [8]. Jaleel ve ark. ise mamogram kesitlerinde dalgacık dönüşümü üzerinden dokusal özellikleri hesaplayarak radyal taban fonksiyonlu yapay sinir ağları kullanarak sınıflandırma yapmışlardır [9].

Bu çalışmada MIAS veri tabanından alınan mamogram görüntülerinin 4 farklı şekilde önışlemeden geçirilmesinden sonra dokusal özellikleri kullanılarak sınıflandırma yapılmıştır. Böylelikle dokusal özellikler ile sınıflandırmada en başarılı önışleme adımı bulunmuştur.

## 2. YÖNTEM (METHOD)

Bu çalışmada MIAS (Mammographic Image Analysis Society Mini Mammographic Database) veri tabanına ait dijital mamografi görüntüleri kullanılmıştır [10]. Veri tabanında kayıtlı mamogramlar MLO (the mediolateral oblique) tekniğine göre çekilmiştir. Her bir mamogram 50µm piksel çözünürlüğündedir ve 1024x1024 piksel boyutundadır; toplamda 322 mamografi

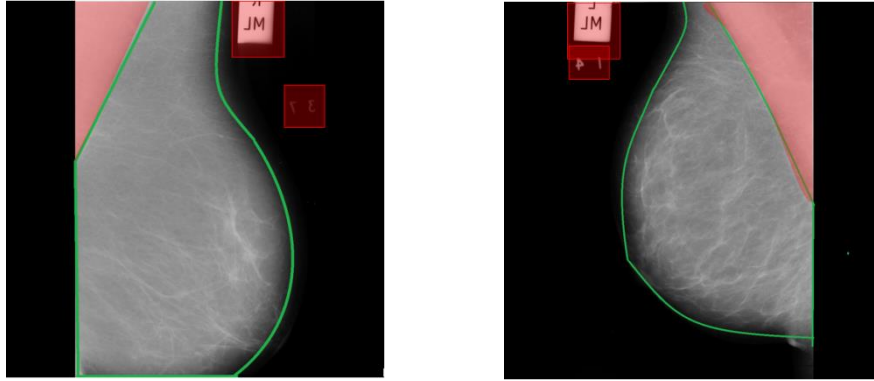
görüntüsü bulunmaktadır. [10]. Mamogramlar içerdiği anomalilere ve iyi huylu/kötü huylu olmalarına göre etiketlenmiştir. Veri tabanından alınan mamogramlar direkt olarak, pektoral kas ve gürültü temizlenmiş olarak, boyutlanmış olarak ve kontrast eşitlenmiş olarak 4 farklı durumda özellik çıkarımı yapılarak sınıflandırılmıştır. Çalışmada tasarlanan sistemin blok diyagramı şekil 1’de verilmiştir.



Şekil 1 Tasarlanan Sistem (Proposed Algorithm)

## 2.1. Ön İşleme Adımları (Preprocessing Steps)

Öncelikle çalışmada MIAS veri tabanından alınan mamogramlar hiçbir işlemten geçirilmeden kullanılmıştır. Bu mamografi görüntüleri grup 1 bölümünü oluşturmaktadır. MLO tipinde çekilen mamogramlarda yanlış bilgi verebilecek olan etiket ve pektoral kas gibi bölümler temizlenmelidir. Çünkü, normal bir mamogramda kist/kitleli bölüm memeden daha beyaz tonlarda temsil edilmektedir ve pektoral kas ve etiketler de meme dokusundan daha parlak olduğu için sağlıklı bir memede dahi anomali varmış gibi bir sonuca sebep olabilmektedir [11–13]. Şekil 2 ‘de bu bölümler kırmızı ile gösterilmiştir. Etiketlerin temizlenmesi için görüntünün siyah-beyaz renklendirilmesiyle gerçekleştirilen morfolojik işlemlerden yararlanılmış, sonrasında ise mamogram sadece meme dokusu ele alacak şekilde kırılmıştır..



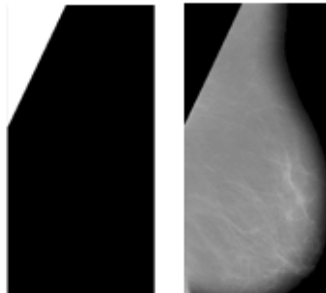
Şekil 2. Sol ve sağ MLO görüntülerinde temizlenmesi gereken alanlar (areas to clean on left and right MLO mammograms)

Şekil 3'te görüldüğü gibi etiketlerden temizlenen mamogramda bir sonraki aşama pektoral kasın temizlenmesidir. Bunun için pektoral kasın meme dokusundan daha yoğun yeğinlik değerleri içerdiği yani daha beyaz olduğu için eşikleme yöntemi kullanılmıştır. Bulunan pektoral kasın kenarı doğru olarak kabul edilmiş ve memeden çıkarılmıştır.



Şekil 3. Etiketlerin temizlenmesi (Cleaning Labels)

Şekil 4'te pektoral kasın bulunması ve pektoral kasta temizlenen meme dokusu görülmektedir. Buna ek olarak pektoral kasın temizlenmesi ve eşiklemeden sonra, bazı mamogramlarda çok küçük alt ve üst dokular açığa çıkmıştır. Bunların kırılması da bu önışlemin bir parçasıdır (Şekil 5).



Şekil 4. Pektoral kasın temizlenmesi (Eliminating pectoral muscle)

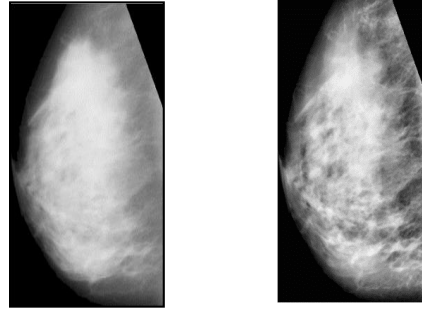
Pektoral kasta ve etiketlerden temizlenmiş mamogramlar grup 2'yi oluşturmaktadır



Şekil 5. Meme alt ve üstünde kalan boşluklar (gaps under and over the breast )

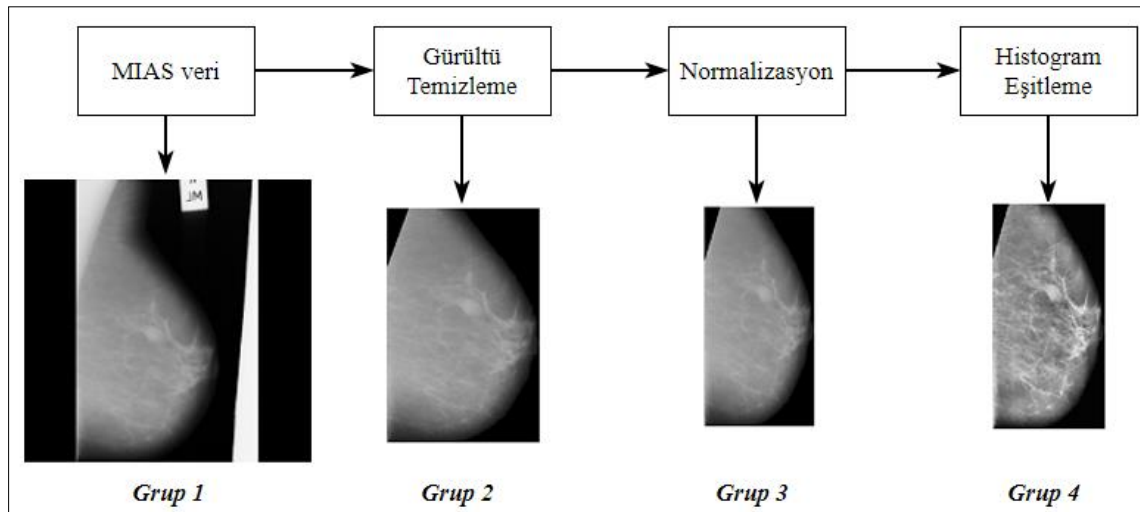
Pektoral kastan temizlenen mamogramlar 256x512 piksel boyutlarına sabitlenmiştir. Boyut normalizasyonu yapılan mamogramlar grup 3'ü oluşturmaktadır.

Son olarak boyutu sabitlenen mamogramların CLAHE (kontrast sınırlı adaptif histogram eşitleme) yöntemi ile kontrast eşitlenmesi sağlanarak grup 4 mamogramları oluşturulmuştur (Şekil 6).



Şekil 6. Normal mamogram ve kontrast eşitlenmiş mamogram (Original contrast and CLAHE applied mammograms)

Şekil 7'de MIAS veri tabanından alınan 012 numaralı mamogram görüntüsünün 4 grup için hazırlık aşması görülmektedir.



Şekil 7. Örnek mamogram grupları (Example of mammography groups)

## 2.2. Dokusal Özellik Çıkarımı (Textural Feature Extraction)

Her bir mamografi görüntüsünün 4 grup için kendi piksel değerleri üzerinden ortalama, standart sapma, entropi, homojenlik, yamukluk ve basıklık değerleri hesaplanarak özellik çıkarımı yapılmıştır. Bu özellikler:

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

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

$$\text{Entropi;} \quad - \sum_{i=0}^{L-1} p_i \log_2 p_i \quad (3)$$

$$\text{Homojenlik;} \quad \sum_{i=1}^{L-1} p_i^2 \quad (4)$$

$$\text{Yamukluk;} \quad \frac{1}{\sigma^3} \sum_{i=0}^{L-1} (i - m)^3 p_i \quad (5)$$

$$\text{Basıklık;} \quad \frac{1}{\sigma^4} \sum_{i=0}^{L-1} (i - m)^4 p_i \quad (6)$$

olarak gösterilmektedir. Eşitliklerde  $p_i$  görüntü matrisinin normalize edilmiş histogramıdır ve formülü  $h_i/L$ 'dir ( $L$  görüntü matrisindeki gri seviye sayısı ve  $h_i$  gri seviye histogramıdır) [12–14].

## 2.3. Sınıflandırma (Classification)

Öznitelik çıkarımının ardından DVM sınıflandırıcısı kullanılarak mamogramların normal/anormal olarak sınıflandırılması yapılmıştır. DVM destek vektörlerini kullanarak veriyi sınıflandırır ve oldukça yaygın kullanılan bir algoritmadır [15,16]. DVM öğrenmesi için seçilen 161 normal ve 66 anormal mamografi imgesi 2 gruba ayrılmıştır. Test ve eğitim verileri kfold algoritması kullanılarak 2/3'ü eğitim verisi ve 1/3'ü test verisi olacak şekilde seçilmiştir. DVM çekirdeği radyal tabanlı olarak belirlenmiştir.

## 3.BULGULAR (FINDINGS)

İncelenen veri kümesine ait özellikler Tablo 1'de verilmiştir. Veri tabanından alınan mamografi görüntüleri aynı eğitim ve test gruplarına bağlı kalınarak DVM sınıflandırıcı ile sınıflandırılmış ve sınıflandırma sonuçları Tablo 2'de verilmiştir.

Tablo 1. Eğitim ve Test Verileri

	<b>Eğitim Verisi</b>	<b>Test Verisi</b>	<b>Toplam</b>
<b>Normal</b>	107	54	161
<b>Anormal</b>	44	22	66
<b>Toplam</b>	151	76	228

Sınıflandırma sonuçlarının incelenmesinde karışıklık matrisleri kullanılmıştır. Buna göre her bir grup için sınıflandırma sonucunda doğru pozitif (TP- True Positive), yanlış pozitif (FP- False Positive), yanlış negatif (FN-False Negative) ve doğru negatif (TN- True Negative) etiketlenmiş

veriler kullanılarak doğruluk, özgünlük ve hassasiyet değerleri hesaplanmıştır. Sonuçlara göre mamografi imgelerinde etiketlerin ve pektoral kasın temizlenmesi sınıflandırma sonuçlarında %10 oranında doğruluk, %17 özgünlük artışı sağlarken %9 oranında hassasiyet düşümü gerçekleştirmiştir. Benzer şekilde boyutlandırılmış ve kontrast eşitlenmiş, grup 3 ve 4, mamogramlar için doğruluk ve özgünlük değerlerini yükseltirken hassasiyet oranını düşürmüştür.

Bu sonuçlara bağlı kalarak, mamogramlarda önışleme yapılmasının yanlış pozitif değerlendirmesini düşürdüğü görülmüştür. Bu sebeple mamogramlarda önışleme yapmak doğruluğu artıracaktır. Önışleme adımlarından sadece pektoral kası ve etiketleri temizleme dört grup arasından en başarılı sonuçları sağlamıştır. Temizlenen mamogram görüntülerinin boyutlarının eşitlenmesi de özgün mamogram sonuçlarına göre daha iyi bir sonuç vermektedir. Diğer taraftan boyut eşitlemesi yapılması işlem süresini kısalttığı ve işleme hızını artırdığı için ileri algoritmalar için kullanılabilir.

<b>Grup</b>	<b>Doğruluk</b>	<b>Özgünlük</b>	<b>Hassasiyet</b>
	$\frac{TP + TN}{TP + FP + FN + TN}$	$\frac{TN}{TN + FP}$	$\frac{TP}{TP + FN}$
<b>Grup 1</b>	0,6184	0,59	0,68
<b>Grup 2</b>	0,7105	0,76	0,59
<b>Grup 3</b>	0,6842	0,72	0,59
<b>Grup 4</b>	0,6447	0,72	0,45

Tablo 2. Sınıflandırma sonuçları

#### 4. SONUÇ VE TARTIŞMA (CONCLUSION AND DISCUSSION)

MIAS veri tabanında mamogramlar 1024x1024 piksel boyutunda sabit tutulsa da her bir mamogramın tümör ya da kalsifikasyon gibi gerçek bilgi içerdiği meme dokusu alanı birbirinden farklıdır. Kimi mamogramlarda yağ dokusu baskın bir meme varken pektoral kasın ve etiketlerin varlığı BDT sistemini yanlış yönlendirmeye yol açmaktadır. Bunun için mamogramların önışleme tabii tutulması elzemdir.

Çalışmada 4 farklı grupta mamogramlar sınıflandırılmış ve sınıflandırma sonuçları ile karşılaştırma yapılmıştır. Bu sonuçlara göre, en başarılı geliştirmeyi, mamogramlardaki gürültülerin ve pektoral kasın temizlendiği grup (grup 2) sağlamıştır. Orijinal mamogramlarda sınıflandırma %61.8 oranında doğruluğa sahipken, bu oran grup 2 mamogramlarda %71'e çıkmıştır. Ayrıca, işlem süresi bakımından tüm mamogramların aynı boyuta getirilmesi programlama aşamasında kolaylık sağlayacağı için grup 3 mamogramları, gürültü ve pektoral kastan temizlendikten sonra 512x256 boyutuna getirilen mamogramlar, doğruluk oranını %68.4'e çıkarmışlardır. Bu sebeple grup 3 önışleme adımı da uygulanacak programa göre tercih edilebilir.

Gelecek çalışmalarda bu algoritma farklı veri tabanları ile de çalıştırılmalı ve sonuçları karşılaştırılmalıdır. Tüm meme dokusu kullanılarak gerçekleştirilen BDT sistemleri için grup 2 ve grup 3 algoritmaları önerilmektedir.

## 5. KAYNAKLAR (REFERENCES)

- [1] Aslan Avdan A., (2013). Duktal Karsinoma İn Situ'da BI-RADS Tanımlayıcıları İle Moleküler Prognostik Faktörler Arasındaki İlişki. *Gazi Üniversitesi Tıp*.
- [2] Gülsün M, Başaran Demirkazık F, Köksal A, Arıyürek M., (2002). Meme mikrokalsifikasyonlarının BI-RADS kriterlerine göre değerlendirilmesi ve yorumcular arasındaki uyumun araştırılması. *Tanısal ve Girişimsel Radyoloji* ;8:358–63.
- [3] Özşen Ö., (2004). *Mamogramlar Üzerinde Uygulanan Görüntü İşleme Tekniklerinin İncelenmesi*.
- [4] Memiş A., (2002). Meme Radyolojisi.
- [5] Akbay C.,(2015). Application Of Image Enhancement Algorithms To Improve The Visibility And Classification Of Microcalcifications In Mammograms. *Middle East Technical University*.
- [6] Redman A, Lowes S, Leaver A., (2015). *Imaging techniques in breast cancer*. Surg (United Kingdom) 34:8–18.
- [7] Sahiner B, Chan HP, Petrick N, Wei D, Helvie MA, Adler DD, ve ark. (1996). *Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images*. IEEE Trans Med Imaging;15:598–610. doi:10.1109/42.538937.
- [8] Chen DR, Huang YL, Lin SH., (2011). *Computer-aided diagnosis with textural features for breast lesions in sonograms*. Comput Med Imaging Graph 35:220–6. doi:10.1016/j.compmedimag.2010.11.003.
- [9] Jaleel JA, Salim S, Archana S., (2014). *Textural features based computer aided diagnostic system for mammogram mass classification*. 2014 Int Conf Control Instrumentation, Commun Comput Technol ICCICCT 2014 2014:806–11. doi:10.1109/ICCICCT.2014.6993069.
- [10] Suckling J, Hutt I, Boggis CRM, Astley S, Betal D, Cerneaz N, ve ark.(1994). *The Mammographic Image Analysis Society Digital Mammogram Database*. Int. Congr. Ser., International Congress Series, 375–8.
- [11] Saltanat N, Hossain MA, Alam MS.,(2010). *An Efficient Pixel Value based Mapping Scheme to Delineate Pectoral Muscle from Mammograms*. IEEE 2010:1510–7.
- [12] Ganesan K, Acharya UR, Chua KC, Min LC, Abraham KT. (2013). *Pectoral muscle segmentation: A review*. Comput Methods Programs Biomed 110:48–57. doi:10.1016/j.cmpb.2012.10.020.
- [13] Pak F, Kanan HR, Alikhassi A.,(2015). *Breast cancer detection and classification in digital mammography based on Non-Subsampled Contourlet Transform (NSCT) and Super Resolution*. Comput Methods Programs Biomed 122:89–107. doi:10.1016/j.cmpb.2015.06.009.
- [14] Gonzalez RC, Woods RE, Telatar Z, Tora H, Arı H, Kalaycıoğlu A., (2014). *Sayısal Görüntü İşleme*. Ankara: Palme Yayıncılık.
- [15] Mohamed WA, Alolfe MA, Kadah YM., (2008). *Microcalcifications enhancement in digital mammograms using fractal modeling*. 2008 Cairo Int Biomed Eng Conf CIBEC 2008:1–5. doi:10.1109/CIBEC.2008.4786034.
- [16] Wang D, Shi L, Heng PA.,(2009). *Automatic detection of breast cancers in mammograms using structured support vector machines*. Neurocomputing 2009;72:3296–302. doi:10.1016/j.neucom.2009.02.015.