

---

*Araştırma Makalesi / Research Article*

---

## **Mamografi görüntülerindeki anormalliklerin yerel ikili örüntü ve varyantları kullanılarak sınıflandırılması**

Volkan Müjdat TİRYAKİ\*

*Siirt Üniversitesi, Bilgisayar Mühendisliği Bölümü, Siirt  
(ORCID: 0000-0003-1824-5260)*

---

### **Öz**

Meme kanseri teşhisinde kullanılan mamografilerdeki anormalliklerin sınıflandırılması için makine öğrenme araştırmaları büyük önem arz etmektedir. Bu çalışmada Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM) görüntü tabanındaki kitlesel ve kalsifikasyonlu mamografi görüntüleri sınıflandırılmıştır. Veri setindeki görüntülerden Yerel İkili Örüntü(YİÖ), Yerel Türev Örüntü, Yerel Dörtlü Örüntü(YDÖ), ve Gürültüye Dirençli Yerel İkili Örüntü yöntemleri ile doku öznelikleri çıkarılmıştır. Öznelik çıkarım yöntemlerinden yerel çarpıklık örüntü tabanlı ayrıntılı histogram yöntemiyle de öznelik çıkarımı yapılmıştır. Daha sonra öznelik vektörleri doğrusal ve radyal tabanlı fonksiyon kernel destek vektör makineleri(DVM) ve yapay sinir ağları (YSA) kullanılarak sınıflandırılmıştır. Eğitim ve doğrulama verisi için 5-kez çapraz doğrulama yöntemi uygulanmıştır. En yüksek sınıflandırma performansı veren eşik seviyeleri ve pencere boyutları her bir öznelik çıkarım yöntemi için belirlenmiştir. Öznelik çıkarımı için gerekli olan süreler tablo halinde verilmiştir. Öznelik çıkarım yöntemi olarak farklı çap ve nokta sayısı ile hesaplanmış YİÖ vektörleri füzyonu ve sınıflandırıcı olarak 2 gizli katmanlı YSA kullanılması durumunda test verisi için %85.74 başarı oranı elde edilmiştir. Elde edilen başarı oranları literatürdeki makine öğrenmesi sonuçlarına göre yüksek ve derin öğrenme sonuçları ile kıyaslanabilir sonuçlardır.

**Anahtar kelimeler:** Dokusal analiz, görüntü sınıflandırma, meme kanseri, yapay sinir ağları.

---

## **Classification of abnormality in breast cancer mammography using local binary patterns and its variants**

### **Abstract**

Machine learning research for the classification of breast cancer abnormalities is very crucial. In this study, mass and calcification abnormalities of mammograms in the Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM) were classified. Local binary pattern (LBP), local derivative pattern, local tetra pattern (LTrP), and noise resistant local binary pattern features were extracted from the images in the data set. Histogram refinement methods of LBP and its variants based on local skewness pattern were also extracted. Feature vectors were then classified by using linear and radial basis function kernel support vector machines (SVM) and an artificial neural network (ANN). 5-fold cross-validation was conducted. The highest performance classification conditions including threshold level and the window size were reported for LBP, its variants, and the histogram refinement methods. Feature extraction time requirements for LBP and its variants were reported. Test accuracy of 85.74 was achieved when feature vectors were extracted at different radius and number of points and fused and classified by using 2 hidden-layer neural network. The test accuracies obtained in this study are better than the machine learning approaches in the literature and comparable with the deep learning results.

**Keywords:** Textural analysis, image classification, breast cancer, artificial neural networks.

---

---

\*Sorumlu yazar: [tiryakiv@siirt.edu.tr](mailto:tiryakiv@siirt.edu.tr)

Geliş Tarihi: 24.04.2019, Kabul Tarihi: 11.07.2019

## 1. Giriş

Meme kanseri dünyada ve Türkiye’de kadınlarda en sık görülen kanser türüdür ve hastalığın görülme sıklığı gittikçe artmaktadır [1]. Meme kanserinin tedavisinde erken teşhis önemlidir. Meme kanserinin görülme oranı yaşlandıkça artmaktadır ve 40 yaş sonrası kadınların düzenli meme muayenesi ve mamografi yaptırmaları tavsiye edilmektedir. Mamografi en yaygın meme tarama teknolojilerinden birisidir. Meme taraması için ultrason, manyetik rezonans görüntüleme (MRI), X-ışın görüntüleme, moleküler meme görüntüleme ve dijital meme tomosentezi (DBT) yöntemleri mevcuttur. Mamografi görüntüleri memenin düşük dozlu X-ışınlarına maruz kalması ile elde edilir [2] ve meme anormalliklerinin taranması için en güvenli yöntemdir [3]. Mamografi, meme kanseri teşhisinde “altın standart” olarak kabul edilmektedir [4].

Meme kanseri teşhisi ile ilgili olarak bilgisayar destekli teşhis (Computer Aided Diagnosis (CADx)) ve bilgisayar destekli anormallik bulunması (Computer Aided Detection (CADE)) farklı kavramlardır. Bilgisayar destekli meme kanseri teşhisi dünyada henüz klinik olarak onaylanmamıştır fakat bilgisayar destekli anormallik bulunması ile ilgili araştırmalar yapılmaktadır. Bu çalışmada güncel bir veri seti olan Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM) içindeki mamografi görüntüleri kullanılarak kitle ve kalsifikasyon türü anormallikleri sınıflandırabilen bir makine öğrenme sistemi geliştirilmiştir [5-7].

Bilgisayar destekli anormallik bulunması bir mamografi sınıflandırma problemi olarak ele alınmamalıdır, çünkü mamografi veri setlerinde mamografi piksel sayıları 1 milyon ile 13.8 milyon arasında değişmekte olup anormalliğin bulunduğu bölge genellikle 200x200 piksel gibi küçük bir bölge içindedir. Bu oran mamografi içindeki toplam alanın %4 ile %0.29’unu oluşturmaktadır ve günümüzde makine öğrenme yöntemlerinin bu kadar küçük bir alan içindeki anormalliği sınıflandırmada başarılı olması mümkün gözükmemektedir [2]. Bunun yerine anormalliğin olduğu bölgeler makine öğrenme aşamasında kullanılmaktadır. Mammographic Image Analysis Society (MIAS) [8] ve CBIS-DDSM gibi veri setlerinde hem mamografiler hem de anormalliğin bulunduğu görüntü parçaları veya koordinatları verilmektedir.

Mamografi görüntülerinden kanserli olanların tespiti radyolog ve doktorlar tarafından yapılır. Bu işlem için bilgisayar destekli bir sistemin kullanılması fikri ilk olarak 1972 yılında Ackerman ve Gose tarafından ortaya atılmıştır [9]. Bu araştırmacılar çalışmalarında bir radyoloğun tanıma oranına yakın bir performansa sahip uzman sistem tasarlamışlardır. Kanserli dokunun otomatik tespiti dokusal analiz ile ilgilidir. 1970’li yıllardan günümüze kadar yeni dokusal öznitelik tanımları ve sınıflandırma teknikleri kullanılarak farklı yöntemler uygulanmıştır. Kutluk ve Günsel meme kanseri tanısı için SIFT (scale invariant feature transform) öznitelik çıkarımı ve LVQ (learning vector quantization) sınıflandırıcı kullanarak meme dokusu yoğunluk sınıflandırıcısı geliştirmişlerdir [10]. Başka bir çalışmada mamografi sınıflandırılması için MIAS veri setinde ayrık kosinüs dönüşümü ve ayrık dalgacık dönüşümü füzyonu kullanılmıştır [11]. Spanhol ve diğerleri DeCaf derin öznitelik çıkarımı kullanarak meme kanseri histopatolojik görüntü sınıflandırması gerçekleştirmişlerdir [12]. Başka bir çalışmada ise gri seviyeli eş oluşum matrisleri ve destek vektör makinesi (DVM) yöntemi ile mamografilerden otomatik doku sınıflandırması yapılmıştır [13].

Yerel ikili örüntü (YİÖ) yöntemi ilk olarak 2002 yılında Ojala ve diğerleri tarafından ileri sürülmüştür [14]. İlerleyen yıllarda bu özneliğin farklı türevleri tasarlanmıştır ve biyomedikal işaret ve görüntü işleme alanlarında yaygın olarak kullanılmıştır [15-17]. Mamografi görüntülerinin sınıflandırılması için YİÖ yöntemi bazı makalelerde ele alınmıştır [18-19]. Meme kanseri teşhisinde IRMA veri setinde YİÖ yöntemlerinden yerel konfigürasyon örüntüsü kullanılmıştır [20]. Bir başka çalışmada BreakHis veri setindeki görüntüler düşük seviyeli YİÖ kullanılarak sınıflandırılmıştır [21]. Bu çalışmada mamografilerdeki anormallik sınıflandırması için veri setleri içinde en büyüğü ve en yenisi olan CBIS-DDSM veri seti kullanılmıştır. Bu veri setindeki mamografilerden YİÖ, yerel türev örüntü (YTÖ), yerel dördü örüntü (YDÖ) ve Gürültüye Dirençli Yerel İkili Örüntü (GDYİÖ) yöntemleri kullanılarak öznitelik çıkarımı yapılmıştır. Mamografi başına öznitelik çıkarım süreleri tablo halinde verilmiştir. Kitle ve kalsifikasyon ikili sınıflandırma problemi için en yüksek tanıma oranı veren öznitelik parametreleri ve sınıflandırıcı özellikleri belirlenmiştir. Sınıflandırıcı olarak doğrusal ve radyal tabanlı fonksiyon kernel DVM ile yapay sinir ağı (YSA) performansları karşılaştırılmıştır. Öznitelik çıkarım yöntemi olarak YİÖ farklı çap ve nokta sayısı ile elde edilmiş öznitelik vektörleri birleştirilmesi (füzyon) ve sınıflandırıcı olarak YSA kullanılması durumunda değerlendirilmesinde %85.54 ve test

verisinde %85.74'lük doğruluk oranına ulaşılmıştır. Bu değerler CBIS-DDSM veri seti kullanılarak bugüne kadar makine öğrenmesi yöntemi ile elde edilmiş en yüksek doğruluk oranlarıdır. Ayrıca önerilen makine öğrenmesi yaklaşımı ile derin öğrenme sonuçları ile kıyaslanabilir sonuçlar elde edilmiştir.

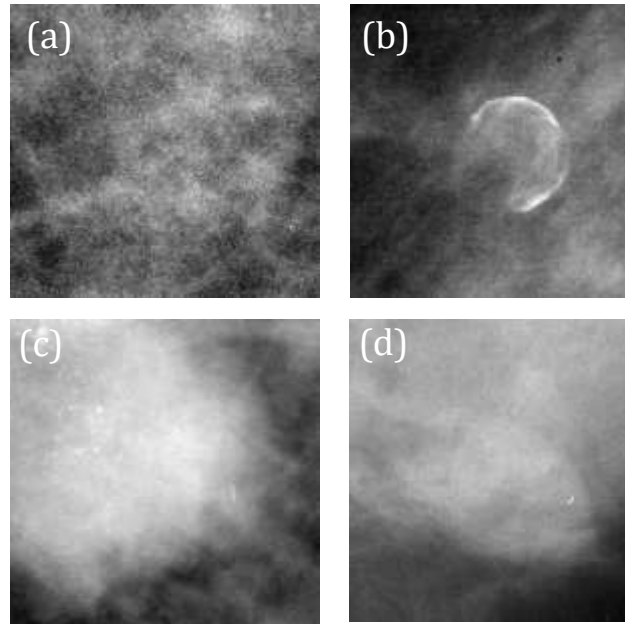
## 2. Materyal ve Metot

### 2.1. Veri Toplama

Mamografi taramasında iki ana anormallik türü kalsifikasyon ve kitledir. Kalsifikasyonlu mamografilerde parlak benek olarak gözükken küçük kalsiyum birikintileri bulunmaktadır. Kitleli mamografilerde ise birden fazla izdüşümü üzerinde yer kaplayan lezyonlar bulunmaktadır [22]. Bu çalışmada mamografilerde karşılaşılan kalsifikasyon ve kitleli mamografi sınıflandırılması problemi ele alınmıştır.

Makine öğrenme sistemlerinin tasarlanabilmesi için veri seti gereklidir. Mamografi görüntüleri için standart değerlendirme verisi bulunmamaktadır. Bundan dolayı birçok bilgisayar destekli teşhis algoritmaları özel veri setlerinde uygulanmaktadır. Bu durum geliştirilen yöntemlerin karşılaştırılmasını zorlaştırmaktadır. Bu çalışmada makine öğrenme sisteminin eğitim, doğrulama ve test aşamaları için mamografi veri setleri içindeki en büyük ve herkesin erişimine açık olan CBIS-DDSM veri seti kullanılmıştır [5-7]. Bu veri seti 2017 yılında yayınlanmış olup içinde sadece kitle ve kalsifikasyon türünde anormallik içeren mamografiler bulunmaktadır. Veri setindeki mamografilerin bit derinliği 16-bittir, toplam mamografi sayısı toplam 3568'dir ve mamografiler Digital Imaging and Communications in Medicine (DICOM) formatındadır. Veri setinin toplam büyüklüğü 160 GB'tır. Veri setinde 4000x3600 pikselden yaklaşık olarak 200x200 piksele kadar değişen mamografi ve mamografi parçaları bulunmaktadır. Veri setindeki mamografiler bilgisayara indirildikten sonra veri seti ile birlikte gelen csv uzantılı dosya yardımı ile okunmuştur.

CBIS-DDSM veri setinde sadece anormallik içeren mamografiler bulunmaktadır. Bu çalışmada veri setinde 450,000 pikselden küçük olan görüntüler makine öğrenme aşamasına katılmıştır. 1502 kalsifikasyonlu ve 1644 kitleli olmak üzere toplam 3146 görüntü kullanılmıştır. Görüntülerin %15'i test ve %85'i eğitim ve doğrulama verisi olarak kullanılmıştır. Ele alınan makine öğrenme yöntemi öğreticili öğrenmedir. Şekil 1'de kitle ve kalsifikasyon anormallikleri içeren örnek mamografiler verilmiştir.



**Şekil 1.** CBIS-DDSM veri setinden bazı mamografi görüntüleri a) Nokta şeklinde pleomorfik kalsifikasyon b) Yumurta kabuğu şeklinde kalsifikasyon c) Spiküle düzensiz kitle d) Lobüle çevrenmiş kitle

## 2.2. Öznitelik Seçimi ve Öznitelik Çıkarımı

Mamografilerdeki piksel sayısı 40000 ile 450000 arasında değişmektedir. Bu görüntülerden uygun olan özniteliklerin seçilmesi önemlidir aksi halde çok boyutluluk problemi ile karşılaşılır. YİÖ özniteliği gri seviyeli resimler için tanımlanmıştır [14],[23] ve yerel komşuluğa bağlı bir özniteliktir.  $(x_c, y_c)$  merkezli  $R$  yarıçaplı bir çember üzerinde  $P$  üyesi olan ve  $YİÖ_{P,R}(x_c, y_c)$  şeklinde gösterilen bir YİÖ operatörü denklem 1 ve 2’de verilmiştir.

$$YİÖ_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} S(I(x_p, y_p) - I(x_c, y_c)) \times 2^p \quad (1)$$

$$S(I(x_p, y_p) - I(x_c, y_c)) = \begin{cases} 1 & I(x_p, y_p) - I(x_c, y_c) \geq 0 \\ 0 & \text{aksi halde} \end{cases} \quad (2)$$

1 ve 2 nolu denklemlerde  $I(x,y)$   $(x,y)$  koordinatındaki pikselin parlaklığını gösterir. YİÖ özniteliği çıkarımında merkezi piksel ve bunun komşu pikselleri arasındaki türevlerin karşılaştırması kodlanır. Daha sonra  $0^\circ, 45^\circ, 90^\circ$  ve  $135^\circ$  olmak üzere dört farklı yönde hesaplanmış türevler birleştirilir [24]. YDÖ yönteminde  $0^\circ$  ve  $90^\circ$  yönlerinde yüksek dereceden türev yönleri elde edilir ve türevlerin büyüklükleri karşılaştırma amacı ile dört farklı seviyeye kodlanır [25]. GDYİÖ yönteminde ise pikseller arası küçük seviye farklılıklarına karşı duyarlılığı azaltan bir gürültü azaltma yöntemi kullanılmıştır [26]. Ayrıntılı histogram yönteminde YİÖ öznitelik vektörü ile yerel çarpıklık örüntüsü vektörü birleştirilir ve ikili özdeğer eşleme yöntemi uygulanır [27]. Çarpıklık bir istatistiksel dağılımın ortalaması etrafındaki asimetrisinin bir ölçütüdür ve bu değer dağılıma göre pozitif veya negatif olabilir. Ayrıntılı histogram yönteminin içerik tabanlı görüntü erişimi (CBIR) problemlerinde performans artırdığı ortaya konmuştur [27] ve bu çalışmada da kullanılmıştır. YİÖ ve varyantları ile elde edilmiş öznitelik vektör uzunlukları Tablo 1’de gösterilmiştir. Kullanılan öznitelik çıkarım yöntemlerinin hesaplama maliyetini ortaya koyabilmek için mamografilerden mevcut donanım sisteminde ortalama olarak YİÖ ve varyantlarından mamografi başına öznitelik çıkarım süreleri de Tablo 1’de verilmiştir.

**Tablo 1.** YİÖ, varyantları ve ayrıntılı histogram öznitelik vektör uzunlukları ve çıkarım süreleri (sn.)

Örüntü adı	Orijinal		Ayrıntılı histogram	
	Uzunluk	Süre	Uzunluk	Süre
YİÖ	59	0.16	236	0.33
YTÖ	236	0.26	590	0.43
YDÖ	708	0.53	1534	0.71
GDYİÖ	59	10.81	236	11.04

CBIS-DDSM veri seti içindeki mamografiler YİÖ, YTÖ, YDÖ ve GDYİÖ ve ayrıntılı histogram yöntemleri ile öznitelik vektörleri çıkarılmıştır. Sınıflandırma işlemi öncesinde veri noktalarının ortalaması sıfır ve standart sapması bir olacak şekilde ölçeklenmiştir. Sınıflandırıcı olarak doğrusal ve radyal tabanlı fonksiyon (RBF) kernel DVM ve YSA kullanılmıştır [28]-[30]. Eğitim-doğrulama ve test verileri için sırası ile %85 ve %15 oranında görüntüler kullanılmıştır. Doğruluk tahmini için 5-kez çapraz doğrulama yöntemi uygulanmıştır. Her bir alt bölmede verinin %80’i eğitim ve kalanı test için kullanılmıştır. Doğrulama verisi doğruluk oranı beş sınıflandırma döngüsünün ortalaması alınarak hesaplanmıştır. Eğitim ve doğrulama esnasında test verisi kullanılmamıştır. Başarı ölçütü olarak doğruluk oranı kullanılmıştır. Doğruluk oranı formülü 3 nolu denklemde verilmiştir.

$$\text{Doğruluk} = \frac{\text{Doğru sınıflandırılan görüntü sayısı}}{\text{Toplam görüntü sayısı}} \quad (3)$$

CBIS-DDSM veri seti için YİÖ ve varyantları kullanılarak elde edilen doğrulama verisi DVM doğrusal ve radyal tabanlı fonksiyon kernel sınıflandırmada elde edilen en yüksek doğruluk oranları hesaplanmıştır. Ayrıntılı histogram yöntemi için pencere boyutu 9, 13, 17, 21 ve 25 piksel olması ile eşik seviyesinin 1.0025, 1.0045, 1.05 ve 1.25 olması durumlarında gerçekleştirilmiştir.

### 3. Bulgular ve Tartışma

YİÖ ve varyantları ile öznelik çıkarımı neticesinde en yüksek doğruluk oranı YDÖ öznelik çıkarımı yönteminde pencere boyutu 21 ve eşik seviyesi 1.0045 olduğunda elde edilmiştir. Doğrulama verisi için doğruluk oranları Tablo 2’de gösterilmiştir.

**Tablo 2.** Doğrulama verisi için YİÖ ve varyantları öznelik çıkarımı ve doğrusal ile radyal tabanlı fonksiyon kernel DVM sınıflandırma sonuçları (% doğruluk oranları)

Örüntü	Orijinal		Ayrıntılı histogram	
	Doğrusal	Radyal tabanlı	Doğrusal	Radyal tabanlı
YİÖ	49.96	65.54	56.42	69.58
YTÖ	48.76	65.96	57.36	67.98
YDÖ	55.30	66.07	<b>72.16</b>	63.27
GDYİÖ	51.86	65.62	57.28	69.77

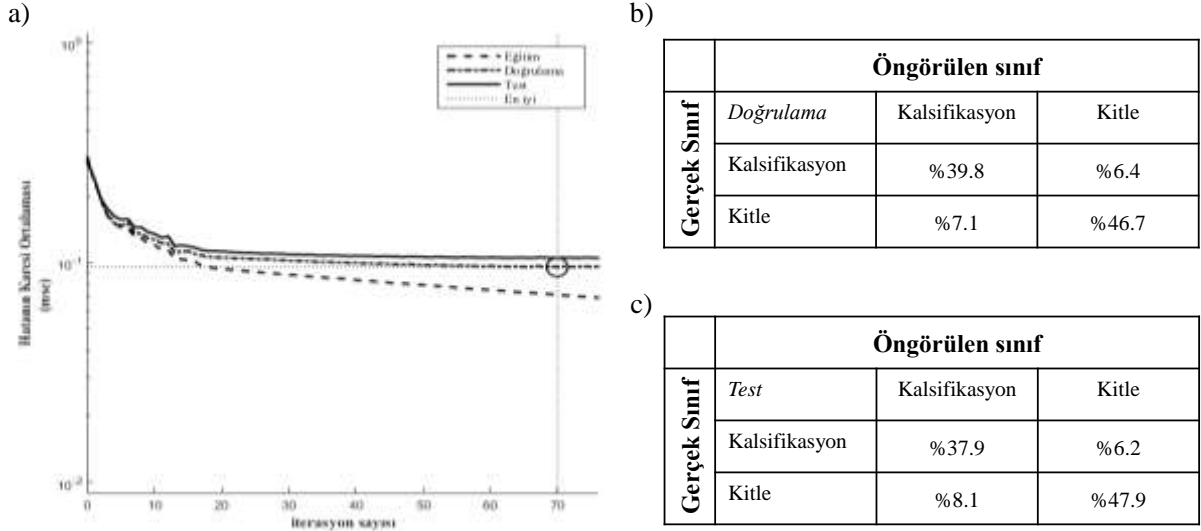
DVM ile elde edilen sonuçlara bakıldığında en yüksek doğruluk oranının ayrıntılı histogram YDÖ öznelik çıkarımı ve sınıflandırıcı olarak doğrusal kernel DVM kullanıldığında elde edildiği görülmüştür. YDÖ öznelik matrisi alınarak YSA kullanılarak da sınıflandırılmıştır. Bu işlemin amacı doğruluk oranını daha da artırmadığını görmektedir. YSA ile doğrulama veri setinde %73.55 doğruluk oranı elde edilmiştir. Bu aşamadan sonra sınıflandırma işlemi için YSA kullanılmıştır.

Doğruluk oranının artırılabilmesi için öznelik füzyonu işleme alınmıştır. Kalsifikasyon görüntülerinde farklı büyüklükte parlak benekler bulunduğu için bu beneklerin büyüklüğü ile orantılı öznelik çıkarımı gerekmektedir. Bu amaçla, sınıflandırma doğruluğunun artırılabilmesi için YİÖ’nün farklı çap (1-27 piksel) ve nokta sayısından (8-28 nokta) elde edilmiş öznelik vektörleri birleştirilmiştir. YİÖ’nün döndürmeye karşı değişmeyen türü kullanılmıştır (uniform rotation invariant) [14]. YİÖ farklı çap ve nokta sayısı kombinasyonları ile elde edilmiş doğruluk oranları Tablo 3’te gösterilmiştir.

**Tablo 3.** Doğrulama verisi için YİÖ’nün farklı çap ve nokta sayısı kullanılarak öznelik çıkarımı için gerekli süreler ve YSA sınıflandırıcı ile elde edilen doğruluk oranları.

Çap ve Nokta sayısı	Öznelik çıkarım süresi (s)	Kitle doğruluk oranı (%)	Kalsifikasyon doğruluk oranı (%)	Doğruluk oranı (%)
(3,16)	0.12	84.42	67.27	76.23
(1,8), (2,8), (2,16), (4,16)	0.28	84.8	69.55	77.62
(1,8)...(4,8), (2,16)...(6,16)	0.6	83.95	71.37	77.95
(1,8), (2,16)...(8,16)	0.65	84.62	71.91	78.63
(1,8), (2,16)...(15,16)	1.18	86.99	75.22	81.32
(1,8), (2,16)...(15;16), (16,20)...(19,20), (20,24)...(23,24), (24,28)...(27,28)	2.55	<b>88.58</b>	<b>82.41</b>	<b>85.54</b>

Öznelik çıkarımından sonra YSA oluşturulmuştur. YSA’da iki gizli katman ve her bir gizli katmanda sırası ile 40 ve 10 nöron kullanılmıştır. Levenberg-Marquardt (LM) eğitim algoritması ve performans fonksiyonu olarak hatanın karesi ortalaması (MSE) kullanılmıştır [31-32]. Eğitim algoritması olarak LM kullanılması durumunda RPROP ve ölçekli eşlenik gradyan geri yayılımı algoritmasına göre daha yüksek sınıflandırma doğruluğu elde edilmiştir. Gizli katmanlarda aktivasyon fonksiyonu olarak tanjant sigmoid ve çıktı katmanında softmax fonksiyonu kullanılmıştır. Verilerin %15’i test amaçlı kullanılmıştır, test verileri bir defa rastgele belirlenmiştir ve sonrasında hep aynı test verileri kullanılmıştır böylelikle test verisine rassal olarak erişim engellenmiştir. Farklı çap ve nokta sayısı ile çıkarılan öznelik vektörleri eklendikçe doğrulama verisi doğruluk oranının arttığı görülmüştür. Sonuç olarak YİÖ çapının 1’den 27’ye kadar birer piksel artırılması ve nokta sayısının ise 8, 16, 20, 24, 28 olarak belirlenmesi ile elde edilen öznelik vektörlerinin birleştirilmesi sonucunda doğrulama verisi için %85.54 ve test verisi için %85.74 doğruluk oranı elde edilmiştir. Nihai öznelik vektörünün uzunluğu 574 olarak elde edilmiştir. Öznelik çıkarım süresi mamografi başına 2.55 saniyedir. Test verisi için kitle sınıflandırma oranı %88.58 ve kalsifikasyon sınıflandırma oranı %82.41’dir. Şekil 2’de YSA performans eğrisi ile eğitim ve test verisi karışıklık matrisleri verilmiştir. Şekilde görüldüğü gibi 70 iterasyon sonucunda eğitim durdurulmuştur.



Şekil 2. YSA sınıflandırıcı sonuçları. a) performans eğrisi, b) doğrulama verisi için karışıklık matrisi, c) test verisi için karışıklık matrisi

#### 4. Sonuç ve Öneriler

Mamografi görüntülerinde düşük kontrast ve gürültü bulunmasından dolayı mamografilerin sınıflandırılmaları zor bir problemdir. Mamografilerin bit derinliğinin nispeten yüksek (16 bit) olması sınıflandırmayı daha da zorlaştırmaktadır. Bu çalışmada, CBIS-DDSM veri tabanındaki mamografiler YİÖ ve varyantları ile öznelik çıkarımından sonra kitle ve kalsifikasyon olarak sınıflandırılmıştır. Bu sınıflandırma problemi için en yüksek test verisi doğruluk oranı YİÖ ile farklı çaplarda öznelik vektörleri füzyonu takiben verilerin YSA ile sınıflandırılmasıyla elde edilmiştir. Kitle ve kalsifikasyon sınıflandırılmasında rastgele belirlenmiş test verisi için elde edilen en yüksek doğruluk oranı %85.74'tür. Literatürde bugüne kadar CBIS-DDSM veri setinde bu çalışma ile benzer sınıflandırma çalışmalardan birisi [2] nolu çalışmadır ve elde edilen en yüksek doğruluk oranı %92.53'tür. Ancak makalede CBIS-DDSM veri setinin hepsi kullanılmamıştır ve veri setinin nasıl oluşturulduğu hakkında detaylı bilgi verilmemiştir. Bu çalışmada bahsi geçen çalışmadan daha fazla görüntü kullanılmıştır. Ayrıca makalede veri setindeki görüntülerin boyut indirgeme yöntemi açıklanmamıştır. Bu yüzden elde edilen doğruluk oranının bu çalışmanın doğruluk oranı ile doğrudan kıyaslanması mümkün değildir. Başka bir çalışmada ise CBIS-DDSM veri seti ve derin öğrenme kullanılmış ve en fazla %75.5 doğruluk oranı elde edilmiştir [33]. Veri seti olarak CBIS-DDSM ve öznelik olarak YİÖ ile HOG (histogram of oriented gradients) kullanılan 2019 yılında yapılan bir çalışmada ise en yüksek %64.35 doğruluk oranı rapor edilmiştir [34]. Bu değerler göz önüne alındığında bu çalışmada literatürde bugüne kadar elde edilen doğruluk oranlarının ikisinden daha yüksek ve biri ile kıyaslanabilir sonuçlar elde edilmiştir.

Bu çalışmada CBIS-DDSM meme kanseri veri setinde YİÖ ve varyantları ile öznelik çıkarımının sınıflandırma performansı ortaya konmuştur. Ayrıntılı histogram YİÖ ve varyantlarının "GHIM 10000" ve "COREL 1000" görüntü veri tabanlarında yönteminin daha üstün olduğu görülmüştür [23]. Ayrıntılı histogram ile elde edilen sonuçlar YİÖ ve varyantları ile elde edilen doğruluk oranlarından daha yüksektir. Ancak en yüksek doğruluk oranı farklı çap ve nokta sayısı için hesaplanmış YİÖ özneliği vektörlerinin füzyonu ve sınıflandırıcı olarak YSA kullanıldığında elde edilmiştir. YİÖ ve varyantlarının öznelik olarak seçilmesinin nedeni mamografilerdeki anormalliklerin yerel komşuluk bazında olmasından ve YİÖ özneliklerinin ise yerel komşuluğa bağlı örüntüleri genel olarak iyi sınıflandırmasından dolayıdır. Farklı çap ve nokta sayısının kullanılmasının nedeni ise anormalliklerin kapladığı alanın değişkenlik göstermesinden dolayıdır. DVM ve YSA sınıflandırıcıların kullanılma sebebi bu yöntemlerin görüntü sınıflandırma alanında en iyi yöntemlerden birkaçı olmasıdır. YSA eğitim algoritması olarak LM algoritması kullanılması durumunda RPROP ve ölçekli eşlenik gradyan geri yayılımına göre daha yüksek doğruluk oranı elde edilmiştir. Bununla birlikte, LM eğitim algoritması bu iki yöntemle göre daha fazla bilgisayar hafızası gerektirmektedir.

Makine öğrenmesi araştırmalarında dikkat edilmesi gereken hususlardan bir tanesi genelleştirmedir. Eğitim aşamasında test verisinden elde edilen sonuçlara göre sistem hiperparametreleri

değiştirilmemelidir. Bu amaç için doğrulama verisi kullanılmalıdır. Bu çalışmada genelleştirme için test verisi yalnızca test aşamasında kullanılmış ve doğrulama verisi doğruluk oranı ile test verisi doğruluk oranının birbirine yakın olduğu görülmüştür. Bu oranların yakın olması genelleştirme konusunda başarılı olduğunu göstermektedir.

Günümüzde makine öğrenme çalışmaları biyomedikal görüntü işleme alanında başarı ile kullanılmaktadır. Makine öğrenme araştırmaları içinde bulunan görüntü sınıflandırma problemleri çok boyutluluk içermektedir. Çok boyutlu görüntülerden sınıflandırma performansını artıracak öznelik seçimi ve öznelik çıkarımı araştırmaları büyük önem arz etmektedir. Mamografi sınıflandırma önemli biyomedikal görüntü sınıflandırma problemlerinden bir tanesidir. Bu çalışmada önerilen mamografi sınıflandırma yaklaşımı meme kanseri anormalliklerinin sınıflandırılmasında sağlık personelinin karar vermesine destek olabilecek, karar aşamasındaki özneliği azaltabilecek ve anormallik bulunması ile ilgili hataların azaltılmasına yardımcı olabilecektir.

## **Teşekkür**

CBIS-DDSM veri setini hazırladıkları ve araştırma amaçlı olarak herkesin erişimine açtıkları için Lee ve diğerlerine [6] teşekkür ederim.

## **Yazarların Katkısı**

Bu makalenin hazırlanması için gerekli çalışmaların tümü yazar tarafından yapılmıştır.

## **Çıkar Çatışması Beyanı**

Yazarlar arasında herhangi bir çıkar çatışması bulunmamaktadır.

## **Etik Kurallar**

Yapılan çalışmada, araştırma ve yayın etiğine uyulmuştur.

## **Kaynaklar**

- [1] Özmen V. 2014. Breast Cancer in Turkey: Clinical and Histopathological Characteristics (Analysis of 13.240 Patients). *Journal of Breast Health*, 10 (2): 98-105.
- [2] Xi P., Shu C., Goubran R. 2018. Abnormality Detection in Mammography using Deep Convolutional Neural Networks. 2018 IEEE International Symposium on Medical Measurements and Applications, 11-13 Haziran, Roma, İtalya, 1-6.
- [3] Tang J., Rangayyan R.M., Xu J., Naqa I.E., Yang Y. 2009. Computer aided detection and diagnosis of breast cancer with mammography: Recent advances. *IEEE Transactions on Information Technology in Biomedicine*, 13 (2): 236-251.
- [4] Miller R.G. 2001. Breast cancer screening. *Journal of General Internal Medicine*, 16 (3): 206-207.
- [5] Lee R.S., Gimenez F., Hoogi A., Rubin D. 2016. The Cancer Imaging Archive. "Curated Breast Imaging Subset of DDSM". <http://dx.doi.org/10.7937/K9/TCIA.2016.7O02S9CY> (Erişim Tarihi: 01.02.2019).
- [5] Lee R.S., Gimenez F., Hoogi A., Miyake K.K., Gorovoy M., Rubin D.L. 2017. A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data*, 4 (170177): 1-9.
- [6] Clark K., Vendt B., Smith K., Freymann J., Kirby J., Koppel P., Moore S., Phillips S., Maffitt D., Pringle M., Tarbox L., Prior F. 2013. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *Journal of Digital Imaging*, 26 (6): 1045-1057.
- [7] Suckling J., Parker J., Dance D., Astley S., Hutt I., Boggis C., Savage J. 1994. The Mammographic Image Analysis Society digital mammogram database. 2nd International Workshop on Digital Mammography, Amsterdam, 375-378.
- [8] Ackerman L.V., Gose E.E. 1972. Breast lesion classification by computer and xeroradiograph. *Cancer*, 30 (4): 1025-1035.

- [9] Kutluk S., Günsel B. 2013. Tissue density classification in mammographic images using local features. 21st Signal Processing and Communications Applications Conference, 24-26 Nisan, Haspolat, Türkiye, 1-4.
- [10] Talha M. 2016. Classification of mammograms for breast cancer detection using fusion of discrete cosine transform and discrete wavelet transform features. *Biomedical Research*, 27 (2): 322-327.
- [11] Spanhol F.A., Oliveira L.S., Cavalin P.R., Petitjean C., Heutte L. 2017. Deep features for breast cancer histopathological image classification. *International Conference on Systems, Man, and Cybernetics*, 5-8 Ekim, Alberta, Kanada, 1868-1873.
- [12] Qayyum A., Basit A. 2016. Automatic breast segmentation and cancer detection via SVM in mammograms. *International Conference on Emerging Technologies*. 18-19 Ekim, Islamabad, Pakistan, 1-6.
- [13] Ojala T., Pietikäinen M., Mäenpää T. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (7): 971-987.
- [14] Burçin K., Vasif N.V. 2011. Down syndrome recognition using local binary patterns and statistical evaluation of the system. *Expert Systems with Applications*, 38 (7): 8690-8695.
- [15] Kaya Y., Uyar M., Tekin R., Yıldırım S. 2014. 1D-local binary pattern based feature extraction for classification of epileptic EEG signals. *Applied Mathematics and Computation*, 243 (2014): 209-219.
- [16] Oppedal K., Eftestøl T., Engan K., Beyer M.K., Aarsland D. 2015. Classifying dementia using local binary patterns from different regions in magnetic resonance images. *International Journal of Biomedical Imaging*, 572567: 1-14.
- [17] Lenc L., Kral P. 2016. LBP Features for Breast Cancer Detection. *IEEE International Conference on Image Processing*, 25-28 Eylül, Phoenix, Arizona, ABD, 2643-2647.
- [18] Nahid A.A., Kong Y. 2018. Histopathological Breast-Image Classification Using Local and Frequency Domains by Convolutional Neural Network. *Information*, 9 (1): 19-45.
- [19] Esener İ.I., Ergin S., Yüksel T. 2015. Göğüs Kanseri Teşhisinde Bir Öznitelik Seçim Analizi. *Tıp Teknolojileri Ulusal Kongresi*, 15-18 Ekim, Muğla, 416-419.
- [20] Kadiroğlu Z., Şengür A., Deniz E. 2018. Classification of Histopathological Breast Cancer Images With Low Level Texture Features. *International Engineering and Natural Sciences Conference*, Kasım 2018, Diyarbakır, 1-8.
- [21] Reston V.A. 2003. *ACR BI-RADS-mammography, ultrasound and magnetic resonance imaging*. 4th ed., American College of Radiology.
- [22] Ojala T., Pietikäinen M., Mäenpää T. 2000. Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *European Conference on Computer Vision*, Berlin, 404-420.
- [23] Zhang B., Gao Y., Zhao S., Liu J. 2010. Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor. *IEEE Transactions on Image Processing*, 19 (2): 533-544.
- [24] Murala S., Maheshwari R.P., Balasubramanian R. 2012. Local tetra patterns: a new feature descriptor for content-based image retrieval. *IEEE Transactions on Image Processing*, 21 (5): 2874-2886.
- [25] Ren J., Jiang X., Yuan J. 2013. Noise-resistant local binary pattern with an embedded error-correction mechanism. *IEEE Transactions on Image Processing*, 22 (10): 4049-4060.
- [26] Tiwari A.T., Kanhangad V., Pachori R.B. 2017. Histogram refinement for texture descriptor based image retrieval. *Signal Processing: Image Communication*, 53: 73-85.
- [27] Cristianini N., Shawe-Taylor J. 2000. *An Introduction to Support Vector Machines and other kernel based learning methods*. AI Magazine, Cambridge University Press, Cambridge.
- [28] Schölkopf B., Smola, A.J. 2001. *Learning with kernels: Support Vector Machines, Regularization, Optimization and Beyond*. The MIT Press, 644s.
- [29] Suykens J.A.K., Van Gestel T., De Brabanter J., De Moor B., Vandewalle J. 2002. *Least Squares Support Vector Machines*. World Scientific, Singapur, 291s.
- [30] Marquardt D. 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *SIAM Journal on Applied Mathematics*, 11 (2): 431-441.



- [31] Hagan M.T., Menhaj M. 1994. Training feed-forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, 5 (6): 989–993.
- [32] Tsochatzidis L., Costaridou L., Pratikakis I. 2019. Deep Learning for Breast Cancer Diagnosis from Mammograms-A Comparative Study. *Journal of Imaging*, 5 (37): 1-11.
- [33] Alhakeem Z., Jang S.A. 2019. Convolution-Free Lbp-Hog Descriptor for Mammogram Classification. *arXiv preprint arXiv: 1904.00187*: 1-5.